Published online 6 December 2016

Nucleic Acids Research, 2017, Vol. 45, No. 2 ***987–1004***
doi: 10.1093/nar/gkw1175

Mg^{2+} ions: do they bind to nucleobase nitrogens?

Filip Leonarski[1,2], Luigi D'Ascenzo[1] and Pascal Auffinger[1,*]

[1]Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR9002, F-67000 Strasbourg, France and [2]Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

Received August 21, 2016; Revised November 10, 2016; Editorial Decision November 11, 2016; Accepted November 14, 2016

ABSTRACT

Given the many roles proposed for Mg^{2+} in nucleic acids, it is essential to accurately determine their binding modes. Here, we surveyed the PDB to classify Mg^{2+} inner-sphere binding patterns to nucleobase imine N1/N3/N7 atoms. Among those, purine N7 atoms are considered to be the best nucleobase binding sites for divalent metals. Further, Mg^{2+} coordination to N7 has been implied in several ribozyme catalytic mechanisms. We report that Mg^{2+} assigned near imine nitrogens derive mostly from poor interpretations of electron density patterns and are most often misidentified Na^+, K^+, NH_4^+ ions, water molecules or spurious density peaks. Consequently, apart from few documented exceptions, Mg^{2+} ions do not bind to N7 atoms. Without much of a surprise, Mn^{2+}, Zn^{2+} and Cd^{2+}, which have a higher affinity for nitrogens, may contact N7 atoms when present in crystallization buffers. In this respect, we describe for the first time a potential Zn^{2+} ribosomal binding site involving two purine N7 atoms. Further, we provide a set of guidelines to help in the assignment of Mg^{2+} in crystallographic, cryo-EM, NMR and model building practices and discuss implications of our findings related to ion substitution experiments.

INTRODUCTION

Magnesium has unique physicochemical properties (1,2) and is recognized as the most important divalent ion for RNA folding, structure and function (3–8). Next to monovalent cations and polyamines (9–11), the main Mg^{2+} function is to counterbalance the high concentration of charged phosphate groups present in nucleic acids, but also to assist folding and function through specific binding modes. As such, it is critical to precisely characterize these binding modes.

A recent PDB survey, available through the MgRNA web site (12), which followed earlier efforts to understand Mg^{2+} binding to RNA (13–16), established a classification of these binding sites. Based on these data, for inner-sphere binding, it was found that Mg^{2+} coordination to phosphate and carbonyl groups dominate followed by a still significant number of coordination patterns to imine sites comprising principally purine N7 and less often N1/N3 atoms. Likewise, other studies relay the opinion that N7 positions make for significant nucleobase metal binding sites (17–19). These views contrast with the understanding, based on the pioneering work of RPJ Williams, that alkali earth metals—including Mg^{2+}—poorly bind to imine atoms, unlike transition metals such as Mn^{2+}, Zn^{2+} or Cd^{2+} (8,20–25). These facts cast doubt on the involvement of nitrogen-bound Mg^{2+} ions in catalytic mechanisms, as previously proposed for hammerhead and pistol ribozymes (18,26–33).

In general, the assignment of ions and other solvent molecules in crystallographic structures is a complex undertaking which seems to lead to harmless attribution errors. After all, ion-binding sites are often believed to play a mere structural role. However, Mg^{2+} are sometimes also identified at key locations where misidentifications can dramatically alter our perception of how biomolecular systems perform their tasks. Many of these errors have been described and are related to the fact that Mg^{2+} is isoelectronic with water and Na^+ (34–37). Yet, despite this awareness and other studies reporting recurrent crystallographic misinterpretations (23,35,38–45), errors are still present in many recently deposited PDB structures while older ones are rarely amended (40,46–51).

Identifying errors in structures is a difficult, tedious, but essential undertaking since, if not corrected, these errors will persist in databases and silently affect the outcome of later studies. Further, they can contaminate the results of database surveys (43,52). For example, an RNA polymerase structure with 485 Mg^{2+} and 5 476 waters (PDB code: 1IW7; resolution: 2.6 Å) was released by the PDB in 2002 (cited by (38)) and a *T. thermophilus* 70S ribosome structure, with ≈1 330 Mg^{2+} per assembly, was released in 2014 (PDB code: 4V83; resolution: 3.5 Å). Given their medium to low-resolution range, these structures contain an excessive number of Mg^{2+} and water molecules, since it was suggested that ions assigned to solvent electron densities at resolutions lower than 2.5 Å are not particularly reliable (23,53).

Here, we critically re-examine inner-sphere binding of Mg^{2+} to imine N1/N3/N7 atoms in RNA, DNA and

*To whom correspondence should be addressed. Tel: +33 388 41 70 49; Fax: +33 388 60 22 18; Email: p.auffinger@ibmc-cnrs.unistra.fr

purine containing metabolites (PDB; May 2016; resolution ≤ 3.0 Å). We limited our investigations to a small subset of all potential Mg^{2+} binding sites in nucleic acids, since analyzing with the same level of details all potential Mg^{2+} binding modes would have been too lengthy. For the same reasons, we did not analyze outer-sphere binding involving water mediated Mg^{2+} to N7 contacts, which is often considered as significant (12,14,54).

Based on our survey, we established that reliable inner-sphere binding occurrences to imine nitrogen atoms are much less frequent than assumed up to now. We notably reduced the number of Mg^{2+} to nitrogen binding types described in earlier classifications (12). Concomitantly, we characterized a large array of misattribution errors and identified some of the underlying reasons that led to them. Not the least of those is a tendency of experimentalists to want to see ions in their density patterns. This causes an overall bias in the database, because experimentalists have systematically interpreted ambiguous information in a given direction, that is toward unjustified or weakly justified identification of solvent peaks as ions and especially Mg^{2+}. These findings call for a more thorough examination of all ion binding sites found in newly deposited crystal structures, and for a re-examination of the Mg^{2+} assignment process for nucleic acid structures. They also advocate for the more systematic use of anomalous diffraction data to identify heavy ions. Such comprehensive and detailed studies are required to move forward on a subject that received already so much attention. As recently claimed (8), we are still at the beginning of understanding the complex interrelationships that link metals to nucleic acid systems.

We conclude this study by providing a set of rules to facilitate ion binding pattern identification. For example, particular care should be taken to respect the octahedral coordination of Mg^{2+} ions and to avoid the overuse of crystallographic restraints that may lead to a confusion between Mg^{2+} and Na^{+} since the main criterion allowing to distinguish them, namely their respectively 2.07 and 2.40 Å coordination distances, is altered. Therefore, we suggest that a significant number of the electron density patterns attributed to Mg^{2+} are generated by other solvent species such as Na^{+}, K^{+}, NH_4^{+}, polyamines or water. In support to these considerations, we present examples where it is indeed the case and stress the necessity to critically examine solvent density maps with a thorough knowledge of all the types of solvent particles present in purification and crystallization buffers (55,56).

MATERIALS AND METHODS

PDB survey

All ≈5 500 nucleic acid crystal structures deposited to the Protein Data Bank (PDB; May 2016; resolution ≤ 3.0 Å) were searched for Mg^{2+} binding to purine and pyrimidine imine N1/N3/N7 atoms (or N_b atoms as defined in reference (12)). To determine cut-off distances for the identification of Mg^{2+} bound to imine nitrogens, we relied on a histogram derived from a CSD search (CSD: Cambridge Structural Database, Version 5.37, February 2016) that identified precise Mg^{2+} to water coordination distances as well as ion exclusion zones (Figure 1). Note that the CSD (57) is a repository for small molecule crystallographic structure that were solved with much better accuracy and, in general, at much higher resolution than those from the PDB. These data parallel those derived from quantum mechanical calculations (58), other PDB surveys (23,59) and first principles molecular dynamics simulations of Mg^{2+} in aqueous solution (60–62) that all suggest that: (i) the $Mg^{2+}\ldots OH_2$ coordination distance is slightly below 2.1 Å; (ii) no water oxygens are found within a d($Mg^{2+}\ldots$Ow) ≈2.2–3.8 Å 'exclusion zone'; (iii) the second coordination shell starts at a 3.8 Å distance from Mg^{2+} and peaks at 4.2 Å. However, since we mostly deal with medium to low-resolution crystallographic structures (3.0 Å ≥ resolution ≥ 2.0 Å), we used more relaxed criteria to identify solvent species around imine nitrogens. Further, we need to consider that, although the most appropriate $Mg^{2+}\ldots$O coordination distance is in the 2.06–2.08 Å range, the default value in the libraries used by the PHENIX (63) and REFMAC (64) refinement programs for d($Mg^{2+}\ldots$Ow) is 2.18 Å. In some instances, this overestimated coordination distance induces serious stereochemical approximations (see below). Bearing in mind that we focus on Mg^{2+} to nitrogen distances, we have also to consider that some authors estimate that the $Mg^{2+}\ldots$N distance is slightly longer (≈2.2 Å) than the $Mg^{2+}\ldots$O distance in agreement with quantum mechanical calculations and PDB/CSD surveys (12,2125). Thus, to a first approximation, our procedures place Mg^{2+} with d($Mg^{2+}\ldots$N) ≤ 2.4 Å in the pool of possible direct binders, while those with distances in the 2.4–3.8 Å exclusion zone were inspected for misidentification.

Since CSD surveys established that divalent ions directly interacting with a purine or imidazole nitrogen lone pair are located in the C–N=C plane (25,65), we applied a 1.0 Å cut-off on the distance between the ion and the nucleobase plane. This criterion applies to divalent ions and not to the less strongly bound alkali (Na^{+}, K^{+}) and the larger alkali earth ions (Ca^{2+}, Sr^{2+}) that display a greater propensity to lie out-of-plane. The searches included also contacts generated by applying crystallographic symmetry operations.

In the ≤ 3.0 Å resolution range, ions with *B*-factors ≥ 79 $Å^2$ were excluded from our statistics since such high *B*-factors do not warrant unequivocal binding site characterizations. Further, we excluded ions with *B*-factors ≤ 1.0 $Å^2$ that are definitely not reliable for Mg^{2+} and hint to the presence of a more electron rich atom (see below). Only Mg^{2+} with occupancy of 1.0 were considered unless otherwise specified. Finally, for all Mg^{2+} ions close to imine nitrogens that we identified as suspect, the F_o-F_c and $2F_o-F_c$ electron density maps deposited to the Uppsala Electron Density Server (EDS) were visualized (66). When these maps were not available—typically for large ribosomal structures—we calculated them with phenix.maps by using the structure factors retrieved from the PDB (63).

Non-redundant Mg^{2+} binding sites were identified as follows. If two nucleotides from different structures involved in a similar Mg^{2+} binding event shared the same residue numbers, chain codes, trinucleotide sequences, ribose puckers, backbone dihedral angle sequences (we used the g+, g−, t categorization) and *syn*/*anti* conformations, they were considered as similar and the one with the best resolution was marked as non-redundant. In case of matching resolu-

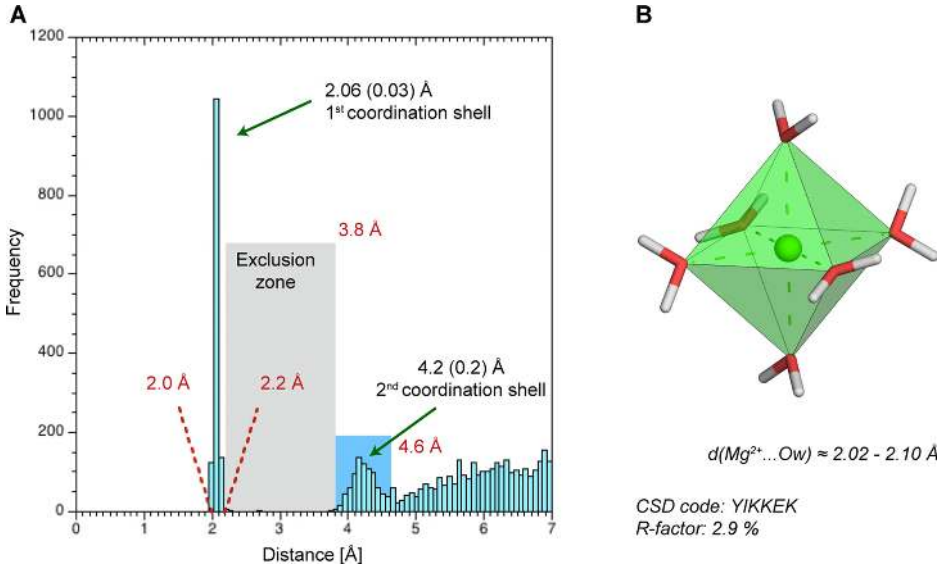

Figure 1. The Mg^{2+} first hydration shell is strictly defined as deduced from high-resolution crystal structures. (**A**) $d(Mg^{2+}\ldots Ow)$ histogram derived from the CSD (version 5.37, update February 2016; *R*-factors ≤ 5%). No disordered, error containing, polymeric or powder structures were included. Standard deviations for the average $Mg^{2+}\ldots Ow$ coordination distances are given in parenthesis. The water exclusion zone and the second Mg^{2+} coordination shell are marked by a grey and a light blue rectangle, respectively. (**B**) An ultra high-resolution $Mg[H_2O]_6^{2+}$ CSD x-ray structure examplifyes the strict octahedral water arrangement around Mg^{2+}(125).

tions, the nucleotide with the lowest *B*-factor was selected. Likewise, if in the same structure two nucleotides involved in a similar Mg^{2+} binding event shared the same residue numbers and trinucleotide sequences (with different chain codes) as well as ribose puckers, backbone dihedral angle sequences and *syn*/*anti* conformations, they were considered as similar and the one corresponding to the first biological unit was marked as non-redundant. To further limit redundancy in the largest ribosomal structures, we restricted our analysis to a single biological assembly when more than one was present (see Supplementary Material for selection criteria).

Two non-redundant sets were calculated with a 2.4 and a 3.5 Å $d(Mg^{2+}\ldots N1/N3/N7)$ distance cutoff, respectively (Table 1). Note that it is impossible to completely eliminate redundancy from such a complex structural ensemble without eliminating at the same time relevant data. Here, we provide an upper limit for a truly 'non-redundant' set. Redundancy issues are further complicated by some systematic assignment errors such as the nucleotide misidentification identified in the first *H. marismortui* 50S structures that leads to the characterization of two distinct structural ensembles (Supplementary Table S1 and Figure S1).

RESULTS AND DISCUSSION

Mg^{2+} to imine N1/N3/N7 contacts are rare

As of May 2016, ≈56 000 Mg^{2+} ions are assigned in ≈1 000 nucleic acid crystallographic structures from the PDB (resolution ≤ 3.0 Å). This corresponds to a ratio of roughly one Mg^{2+} per eight nucleotides. Comparatively, under the same resolution threshold, ≈25 000 Mg^{2+} are found in ≈8 500 proteins. With no resolution limit, the number of Mg^{2+} rises to ≈100 000 in nucleic acids and only ≈30 000 in proteins. In nucleic acids, the largest number of ions comes from low-resolution ribosomal structures that do not allow reliable identification of light solvent particles.

Out of the ≈56 000 Mg^{2+} found in nucleic acid structures, ≈1 000 (≈1.8%) display partial occupancies, 59 are associated with occupancies above 1.0, ≈4 100 (≈7%) display *B*-factors ≥ 79 $Å^2$ and ≈480 (< 1%) display *B*-factors ≤ 1.0 $Å^2$. We excluded these ions from statistics in Table 1. In the remaining pool, around 3 900 (≈7%) ions display $d(Mg^{2+}\ldots N)$ ≤ 3.5 Å. If we consider the more stringent $d(Mg^{2+}\ldots N)$ ≤ 2.4 Å criterion that is more in line with the coordination distance derived from the CSD (Figure 1), only 293 (≈0.5%) Mg^{2+} are contacting imine nitrogens. Most of these Mg^{2+} are close to 108 adenine and 69 guanine N7 atoms, with only 20 close to N1/N3 positions. This number drops to 126 (≈0.3%) if we consider only non-redundant Mg^{2+} binding sites. These values are to be compared with the ≈8 500 (≈15%) Mg^{2+} in direct contact (≤ 2.4 Å) with phosphate oxygens that are considered to be the primary nucleic acid binding sites for Mg^{2+}.

$d(Mg^{2+}\ldots N)$ histograms reveal unrealistic coordination distances

As stated in the Methods section, the coordination geometry of Mg^{2+} to water and other ligands is strictly defined. Ideally, the $d(Mg^{2+}\ldots N)$ and the $d(Mg^{2+}\ldots Ow)$ PDB and CSD histograms should display a similar profile. However, in the former (Figure 2), we could not identify a clear peak around 2.1 Å. Furthermore, the exclusion zone identified in Figure 1 is significantly populated in the PDB data, suggesting ion misidentifications. Mg^{2+} assignments with coordination distances in the 2.4–2.6 Å range may correspond to Na^+ that are frequently present in crystallization buffers—for example as sodium cacodylate—and have coordination distances to water around 2.4 Å, as shown in a

Table 1. Number of non-redundant Mg^{2+}…N1/N3/N7 contacts in structures from the PDB (resolution ≤ 3.0 Å)

	$d(Mg^{2+}…N1)$		$d(Mg^{2+}…N3)$		$d(Mg^{2+}…N7)$	
	≤ 3.5 Å	≤ 2.4 Å	≤ 3.5 Å	≤ 2.4 Å	≤ 3.5 Å	≤ 2.4 Å
DNA						
DA	—	—	—	—	3 (3)	—
DG	*NR*	*NR*	1 (1)	—	23 (23)	8 (8)
DC	*NR*	*NR*	1 (1)	—	*NR*	*NR*
RNA						
A	116 (243)	5 (5)	131 (198)	2 (6)	245 (783)	24 (108)
G	*NR*	*NR*	84 (121)	5 (6)	1324 (2386)	84 (191)
C	*NR*	*NR*	122 (192)	3 (3)	*NR*	*NR*
Total:	116 (243)	5 (5)	339 (513)	10 (15)	1595 (3195)	116 (307)

The total number of occurrences is given in parenthesis ('*NR*' stands for 'Non-Relevant'). Ions with *B*-factors ≤ 1.0 $Å^2$ and ≥ 79 $Å^2$ were not counted.

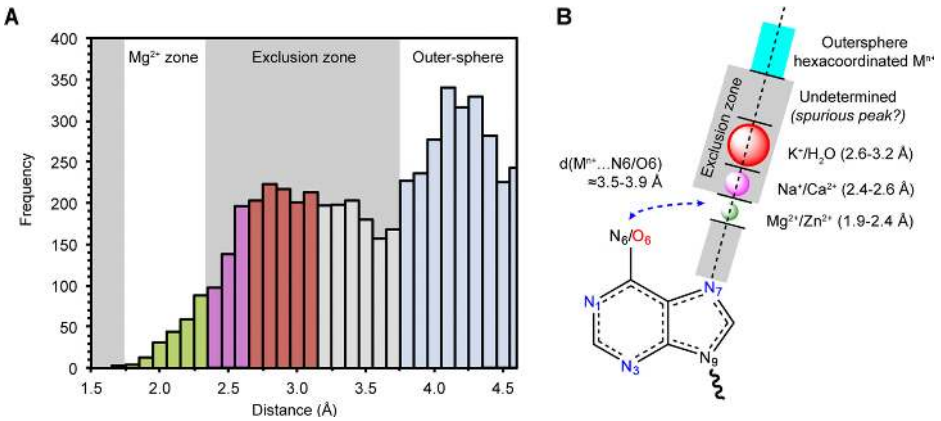

Figure 2. Mg^{2+} coordination to purine N7 atoms derived from PDB structures. (**A**) $d(Mg^{2+}…N7)$ histogram (derived from the PDB; May 2016; resolution ≤ 3.0 Å). Ions with *B*-factors ≤ 1.0 $Å^2$ and ≥ 79 $Å^2$ were excluded. The different ion binding zones are colored according to Figures 1 and 2B. (**B**) Scheme showing the different ion binding zones in front of the purine N7 atom (the oxygen and nitrogen atoms able to associate with a cation are shown in red and blue, respectively). The $d(M^{n+}…N6/O6)$ expected distance range is also indicated. Note that these cutoff distances are indicative and simply suggest that the potential ion assignments close to the mentioned limits should be considered with greater care.

hammerhead ribozyme structure (PDB code: 3ZP8; resolution: 1.55 Å) (67). Note that Ca^{2+} with similar coordination distances are mentioned in the crystallization conditions of some *T. thermophilus* 30S structures (see Supplementary Material). Mg^{2+} assignments with coordination distances in the 2.6–3.2 Å range may correspond to K^+, NH_4^+ or water, all with coordination distances around 2.8 Å. Mg^{2+} assignments in the 3.2–3.8 Å range may be related to the presence of anions (68,69), solvent molecules pertaining to the purification and crystallization buffers, contaminants or may be related to crystallographic artifacts (40,70). In accordance with CSD data, the broad peak around 4.2 Å is attributable to Mg^{2+} interacting with imine sites through their first hydration shell. Interestingly, the abnormalities in the $d(Mg^{2+}…N7)$ histogram become more apparent when compared to the $d(Mn^{2+}…N7)$ histograms that show a clear first shell peak in the 2.1–2.6 Å range (Figure 2 and Supplementary Figure S2).

A second binding criterion, derived from CSD searches (25), specifies that when a transition metal or Mg^{2+} binds to a purine N7 atom, the allowed $d(M^{2+}…N6/O6)$ should be in the 3.5–3.9 Å range (Figure 2B). Out of 111 non-redundant ions with $d(Mg^{2+}…N7) \leq 2.4$ Å, only 62 ions satisfy this criterion. Thus, we infer that the majority of Mg^{2+} close to imine nitrogens are misidentified (Supplementary Figure S3).

With a few exceptions, all direct binding occurrences with $d(Mg^{2+}…N7) \leq 2.4$ Å are suspect

Mg^{2+} singly bound to adenine N7 atoms is not observed. Direct ion binding to the N7 position of adenine is complicated by the presence of the nucleobase amino group that imposes steric and electrostatic constraints (25). Although pentahydrated ion-to-N7 binding has been observed at high-resolution in an adenine/nickel complex (CSD code: ZZZAAF01; Supplementary Figure S4), only few partially hydrated and no $Mg(H_2O)_5^{2+}$ were located close to (A)N7 in the PDB. In a hammerhead ribozyme (29), the conformation of a nucleobase that involves a $d(Mg^{2+}…N6/N7) \leq$ 2.0 Å coordination was probably incorrectly modeled (71) (Figure 3A). Elsewhere in the same structure, completing the hydration sphere of an adenine-bound Mg^{2+} resulted in severe clashes with adjacent nucleotides (Supplementary Figure S5). In a glmS ribozyme (PDB code: 2GCV; resolution: 2.1 Å), the *B*-factor is higher for the metal ion than for the attached nucleobase and waters, which hints at refinement issues combined to ion misidentification (Figure 3B). Moreover, a Mg^{2+} has been placed at 3.1 Å from a cytosine N4 atom (most probably a Cl^- ion (68,69); Supplementary

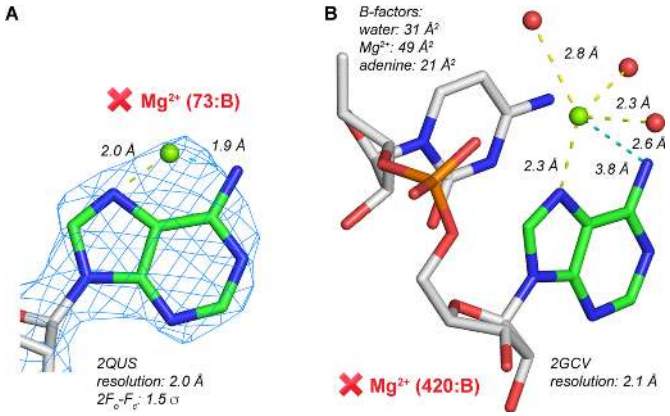

Figure 3. Unrealistic binding of Mg^{2+} to adenine N7 atoms. Mg^{2+} to N6 coordination distances are shown in cyan. The red cross is used to mark misidentified Mg^{2+} ions. (**A**) Ill-placed Mg^{2+} according to electron density patterns and coordination distances. (**B**) The Mg^{2+} hydration sphere is incomplete with erratic coordination distances. Fixing the hydration sphere of this ion with a proper hexacoordinated geometry would result in clashes similar to those shown in Supplementary Figure S5.

Figure S6A), suggesting to critically reexamine all other solvent assignments proposed for this glmS ribozyme. Finally, a suspicious Mg^{2+} bound to a N7 atom—more probably Na^+—is found in a ribosome where $d(Mg^{2+}\ldots N7)$ and $d(Mg^{2+}\ldots Ow)$ are ≈2.4 and ≈2.6 Å, respectively (Supplementary Figure S6B).

This quasi-absence of reliable $Mg(H_2O)_5{}^{2+}$ to (A)N7 contacts suggests a low Mg^{2+} affinity for this site (25). However, in rare instances, secondary Mg^{2+} contacts to N7 atoms complemented by primary contacts to anionic phosphate or amino acid oxygens may be associated with the formation of appropriate but rare Mg^{2+} binding pockets (see below).

Mg^{2+} singly bound to guanine N7 atoms: is it more probable? Binding of divalent metals to (G)N7 atoms has been reported more frequently in both the CSD and the PDB as a probable result of the larger electronegativity of guanine versus adenine Hoogsteen edges (25). In the PDB, 41 non-redundant Mg^{2+} bind solely to (G)N7 atoms (no other direct contact to DNA/RNA atoms). Out of those, only 20 Mg^{2+} comprising three $Mg(H_2O)_5{}^{2+}$ satisfy the $3.5 \leq d(M^{2+}\ldots O6) \leq 3.9$ Å criterion (Figure 2). Two of these $Mg(H_2O)_5{}^{2+}$ are present in the same synthase/tRNA structure (PDB code: 4YCO; resolution: 2.1 Å) and have been modeled based on octahedral densities displaying merged water peaks. $d(Mg^{2+}\ldots Ow) = 2.18$ Å restraints were used during refinement (Figure 4A). As a result, Na^+ can be fitted with a similar level of confidence into these density patterns (see below).

The remaining 21 ions with outlier $d(Mg^{2+}\ldots O6)$ distances were assigned without proper care for stereochemistry and hydration patterns. In a few instances, a complete hydration shell was modeled. However, without well-defined solvent density patterns, these hydration shells display poor geometry. In a twister ribozyme structure (PDB code: 5DUN; resolution: 2.6 Å), five waters were fitted in a density pattern lacking octahedral symmetry and the Mg^{2+} *B*-factor is larger than that of the bound nucleobase (Figure 4B). In a group II intron (PDB code: 4E8N; resolution: 3.0 Å), a Mg^{2+} is placed in front of an N7 atom and displays a non-octahedral coordination (Figure 4C).

Mg^{2+} binding to (G)N7 in DNA is rare. Ion coordination to the N7 of a terminal guanine was identified in five Z-DNA hexamers with resolution ≈1.0 Å (Figure 4D) while in a few structures (PDB codes: 4HIG, 4HIF, 1D39), Mn^{2+}, Zn^{2+} and Cu^{2+} replace Mg^{2+} (72,73). All these structures belong to the $P2_12_12_1$, space group. However, these hexamers crystallize in the $P2_12_12_1$ space group in both a Mg^{2+} and a spermine form (74–76) suggesting that Mg^{2+} is not specifically required to stabilize the crystal. Further, under high $MgCl_2$ and $CaCl_2$ concentrations (500 mM), this hexamer crystallizes in a *P*32 space group and, surprisingly, no direct ion binding to N7 was reported (77) as in another Z-DNA structure with much lower divalent ion concentrations (30 mM), stressing the difficulty to predict such ion binding patterns (78).

Pentahydrated Mg^{2+} coordination to (G)N7 was reported in only one B-DNA structure (PDB code: 1DCR; resolution: 1.6 Å). It shows clearly identifiable solvent density peaks as well as coordination distances and geometry consistent with Mg^{2+} binding. Similarly, Co^{2+} and Zn^{2+} bind to a terminal guanine of two DNA hexamers (PDB codes: 1FD5, 1P26; resolutions: 1.10, 2.92 Å) (79). Again, the terminal position is favored since steric hindrance prevents pentahydrated metals to bind to N7 within a B-DNA helical context (79,80) (Supplementary Figure S4). It is surprising that $Mg(H_2O)_5{}^{2+}$ binds to terminal Z- and B-DNA but not to RNA nucleobases.

M^{n+} bound to two N7 atoms of stacked purines: Mg^{2+}, Zn^{2+} or a monovalent cation? Next to pentahydrated ion binding sites, we searched for an atypical pattern involving the coordination of a divalent cation to two N7 atoms belonging to purines arranged in a stacked head-to-tail manner (Figure 5A), a pattern that we identified earlier in the CSD (25) and was also described by others (8,12,14,81). We identified 45 cases with $d(Mg^{2+}\ldots N7) \leq 2.4$ Å, all associated with four prokaryotic ribosome sites (Table 2 and Figure 5B). Sites I, III and IV were observed in important structural elements—three-way junction for site I and bulges for sites III and IV— while site II is constitutive of ribosomal helix 52. Hence, we analyzed more systematically the 289 potential binding sites in the 126 prokaryotic ribosomes we surveyed (Supplementary Table S1). We excluded 18 instances where both purine *B-factors* are larger than 79 Å^2 leading thus to a total of 271 binding sites. Among them, Mg^{2+} was assigned in 243 instances and Sr^{2+} in 28 instances since this ion was used in crystallization buffers of *H. marismortui* large ribosomal subunits (see below).

However, for the largest number of sites, Mg^{2+} attribution is inappropriate since only less than one out of every six ion satisfies the $d(Mg^{2+}\ldots N7) \leq 2.4$ Å criterion (Table 2 and Figure 5B). Therefore, these sites appear to be occupied by waters or non-Mg^{2+} ions for which the coordination distance to N7 is > 2.4 Å. These data illustrate the difficulty of thoroughly analyzing the current pool of ribosomal structures for which we do not only have to deal with various resolution levels but also with a large gamut of crystallization protocols, refinement procedures and interpretation habits

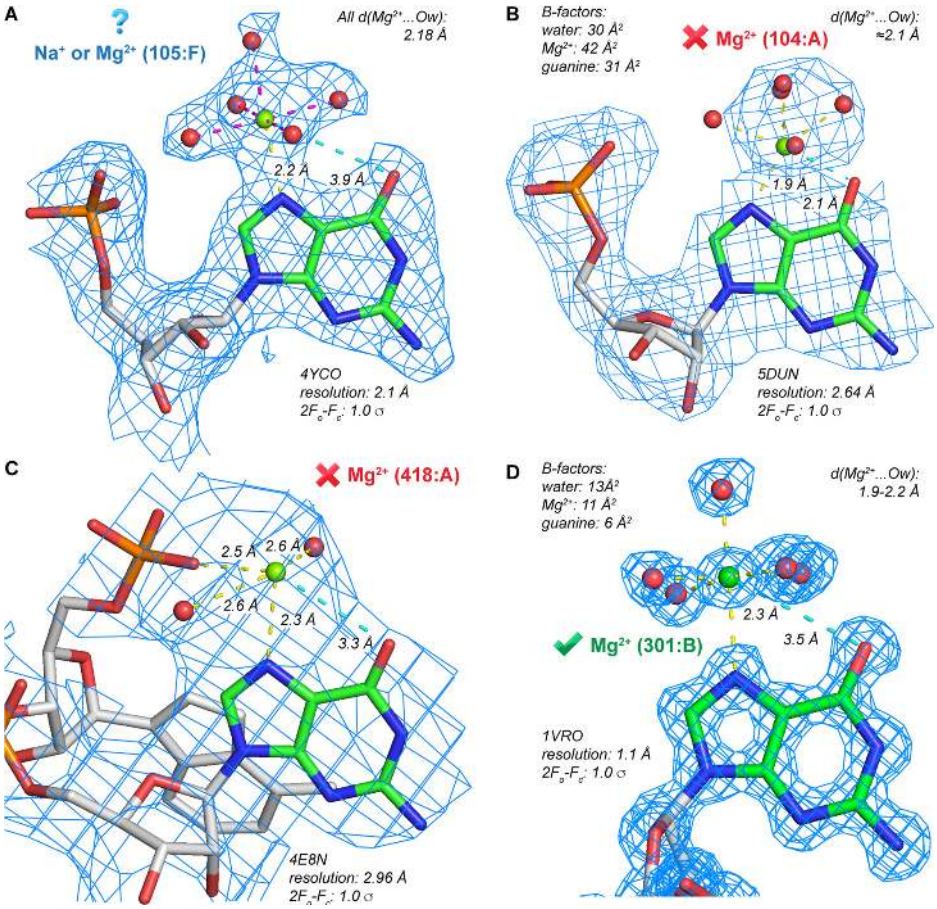

Figure 4. Mg^{2+} close to guanine N7 atoms in PDB structures. The cyan question mark, red cross and green mark are used to identify: (i) sites where either Na^{+} or Mg^{2+} match the electron density, (ii) a misidentified and (iii) a correctly placed ion. (**A**) The $d(Mg^{2+}\ldots Ow)$ coordination distances, shown in magenta, were irrealistically modeled to 2.18 Å. A larger Na^{+} could equally fit into this density pattern. (**B**) Mg^{2+} is distant from the electron density center, leading to underestimated $d(Mg^{2+}\ldots N7)$ and $d(Mg^{2+}\ldots O6)$ distances. (**C**) Incomplete and poorly defined Mg^{2+} coordination shell. The coordination distances suggest the presence of Na^{+} or water but not Mg^{2+}. (**D**) A reliable but rare pentahydrated coordination pattern with separate densities for water and Mg^{2+}. Similar patterns are reported in a set of high-resolution Z-DNA crystal structure (PDB codes: 1VRO, 292D, 2DCG, 2ELG, 336D).

(37,82–83). Going through such a demanding process is at least necessary for some sites. Such coordination distance spreads could certainly not have been foreseen otherwise.

In the following, we focus on the 43 ions with $d(Mg^{2+}\ldots N7) \leq 2.4$ Å found at site I and IV. Although their coordination distances seem appropriate, we noticed several inconsistencies that led us to question their identity. The first is associated with ion *B-factors* that show a propensity to be lower than those of the attached purines (Table 2 and Figure 5C): in one instance, for site I, the Mg^{2+} *B*-factor was set to zero (PDB code: 4U20); in 10 instances, the ion *B*-factors were set to ≤ 1.0 Å^2. Such low *B*-factors usually appear when the atom at the origin of the observed density has more electrons than the one used in the model. We identified additional strategies used by experimentalists to absorb excess density in the F_o-F_c maps at site IV: first, an occupancy of 1.3 was assigned to Mg^{2+} (PDB code: 1N32; resolution: 3.0 Å); second, two Mg^{2+} separated by 1.7 Å and each with 1.0 occupancy were modeled (PDB code: 4B3M; resolution: 2.9 Å). Both of these scenarios are physically impossible. Consequently, we pondered about which other ion could explain such density patterns and realized that, by analogy with zinc-fingers where Zn^{2+} binds to two histidine residues, this double N7 motif represents also an appropriate binding site for Zn^{2+}, a metal that should not be perceived as a trace element given high Zn^{2+} intracellular concentrations (84). We realized also that Zn^{2+} has been identified in the ribosomal proteins (85–87) of almost all *T. thermophilus*, *E. coli* and *S. cerevisae* structures (Supplementary Table S1). Further, although Zn^{2+} displays a tetrahedral coordination in zinc-fingers, this ion binds sometimes to (G)N7 with an octahedral coordination (see PDB code: 4HIF; resolution: 0.85 Å) (73). In order to validate our hypothesis of Zn^{2+} replacing Mg^{2+}, we examined the peak height of the electron densities associated with some of these ions. In 11 instances out of 43, the density of the ion remains visible at sigma levels above those corresponding to neighboring phosphorus atoms, therefore strongly suggesting the presence of a transition metal (Figure 5D). When we attempted to re-refine some of these structures by replacing Mg^{2+} by Zn^{2+}, reasonable ion *B*-factors and no abnormal positive or negative peaks in the F_o-F_c density maps were obtained.

As mentioned above, 28 structures of the large *H. marismortui* ribosomal subunit where crystallized in the pres-

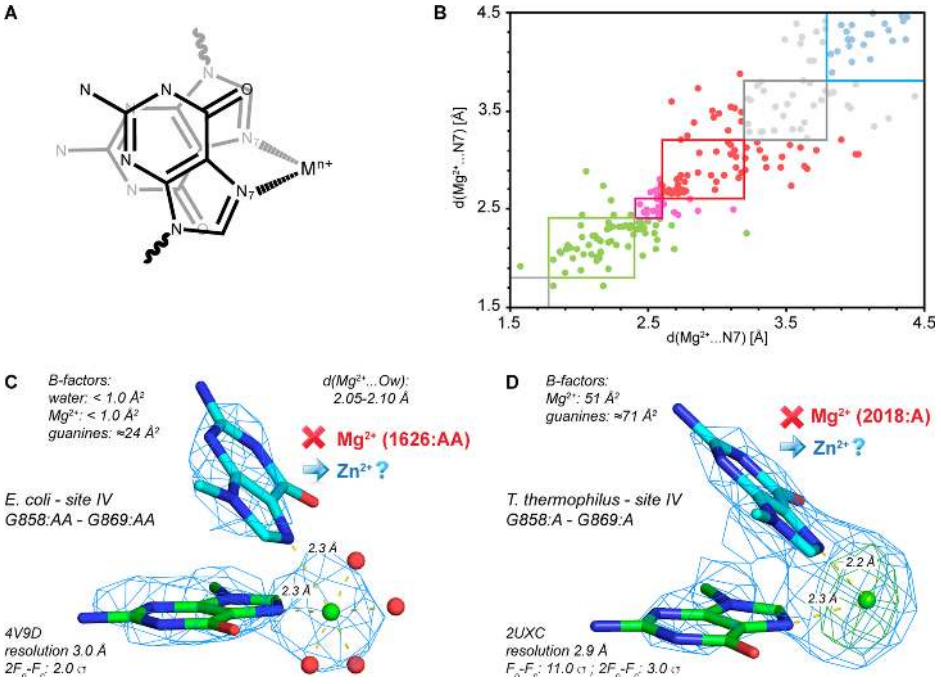

Figure 5. M^{n+} bound to a double N7 motif. (**A**) Schematical representation of this ion binding pattern with two guanines. Guanine-adenine combinations were also identified (Table 2). (**B**) 2D diagram showing Mg^{2+} distances with respect to each of the bound N7 atoms. The Mg^{2+} dots are colored according to their distance to the closest N7 atom (see Figure 2). The colored rectangular boxes frame the ions with respect to both coordination distances. (**C**) A Mg^{2+} placed close to site IV in an *E. coli* structure (Table 2). Though the coordination distance is correct, this ion assignment is ambiguous since ion and water *B*-factors < 1.0 Å^2 are not consistent with those of the guanines. These facts hint to the presence of a more electron dense ion such as Zn^{2+}. (**D**) Mg^{2+} placed close to site IV in a *T. thermophilus* structure. This ion assignement is ambiguous since, although the coordination distances are in agreement with those of Mg^{2+}, the F_o-F_c density (in orange) points to the presence of a more electron dense ion, possibly Zn^{2+}.

Table 2. Occurrence of Mg^{2+} and Sr^{2+} ions bound to two N7 atoms of stacked purines in ribosomal structures (resolution ≤ 3.0 Å; see Supplementary Table S1 and Figure 5)

Site	Org.[a]	Res.[b]	$d(Mg^{2+}\ldots N7) \leq 2.4$ Å[c]	$d(Mg^{2+}\ldots N7) > 2.4$ Å[d]	Second-shell Mg^{2+} [e]	Sr^{2+} [f]	Empty site [g]	Total [h]
Large ribosomal subunit (LSU)								
I.	*HM*	G:824-G:854	27 (18)	1 (1)	—	28 (0)	1	57
	TT	G:733-A:761	4 (3)	37 (27)	—	*NR*	1	42
	EC	G:733-A:761	5 (4)	7 (5)	—	*NR*	—	12
	DR	G:733-A:761	—	—	—	*NR*	—	—
II.[i]	*TT*	G:1358-G:1371	—	13 (6)	26 (4)	*NR*	3	42
	EC	G:1358-G:1371	—	12 (8)	—	*NR*	—	12
	DR	G:1358-G:1371	—	—	—[j]	*NR*	1	1
Small ribosomal subunit (SSU)								
III.[i]	*TT*	G:581-G:758	2 (2)	44 (17)	2 (1)	*NR*	4	52
IV.	*TT*	G:858-G:869	6 (6)	35 (26)	—	*NR*	—	41
	EC	G:858-G:869	1 (1)	11 (10)	—	*NR*	—	12
Total**:**			45 (33)	160 (99)	28 (5)	28 (0)	10	271

The number of ions with an inappropriate *B*-factor—lower than that of at least one of its bound N7 atoms—is given in parenthesis. A site is counted only when the *B*-factors of any of the two purines are below 79 Å^2; no *B*-factor criterion was applied to the ion.
[a]Organisms in which these motifs occur; HH, TT, EC and DR stand for *H. marismortui*, *T. thermophilus*, *E. coli* and *D. radiodurans*, respectively.
[b]The residue numberings are those found in the most representative PDB structure for each organism as noted in Supplementary Table S1, namely 4V9F, 4Y4O, 4YBB, and 5DM6 for HM, TT, EC and DR, respectively.
[c]Both $d(Mg^{2+}\ldots N7)$ distances have to be below 2.4 Å.
[d]One of both $d(Mg^{2+}\ldots N7)$ distances has to be in the 2.4–3.8 Å range.
[e]Both $d(Mg^{2+}\ldots N7)$ distances have to be in the 3.8–4.6 Å range.
[f]Sr^{2+} ions are only present in some HH–LSU structures (see Supplementary Table S1). Sr^{2+} ions are considered if $d(Mg^{2+}\ldots N7) \leq 3.0$ Å.
[g]No ions with the criteria defined above are found at these sites.
[h]Total number of identified sites.
[i]Site II. in HM and site III. in EC have no double N7 ion binding motif.
[j]Without the *B*-factor < 79 Å^2 criterion, two second shell Mg^{2+} ions are reported in DR structures (PDB codes: 5DM6, 5DM7).

ence of Sr^{2+}. In site I, Mg^{2+} is systematically replaced by Sr^{2+} that accounts better for the observed electron density and results in ion *B-factors* larger than those of the nucleobases (Table 2). However, a close inspection of the crystallographic data suggests that Sr^{2+} is incompatible with the observed electron density: *B-factors* are twice those of the nucleobase with some $d(Sr^{2+}\ldots N7)$ distances as short as 2.2 Å, while the CSD estimated Sr^{2+} coordination distance is ≈2.6 Å and the preferred ligands are oxygens (23,88). Moreover, all alkali earth metals, including Sr^{2+}, are poor N7 binders (25). Rightfully, the authors of these structures did not envisage the binding of Cd^{2+}, an ion that is present in all *H. marismortui* structures (Supplementary Table S1) and has an excellent affinity for nitrogens but has also 10 and 18 more electrons than Sr^{2+} and Zn^{2+}, respectively.

We are aware that the data we gathered are not sufficient to unambiguously identify the ions present at these locations. However, the possibility that site I and IV bind Zn^{2+} is strongly supported by our analysis and should be further investigated. Based on EXAFS experiments, it has been proposed that the *E. coli* 70S ribosomes tightly bind to 8 equivalents of Zn^{2+} (87,89). The authors of this study suggested that, next to zing-finger motifs, another strong Zn^{2+} binding site was associated with ribosomal RNA but were unable to characterize it. Therefore, we hypothesize that these double N7 binding sites, that are poor binding sites for alkali earth ions, are the best ribosomal locations for Zn^{2+} and other transition metals and bind eventually monovalent cations when transition metals are not present. Definite answers regarding the identity of these ions will have to wait for multi-wavelength anomalous diffraction measurements (36,56,90).

M^{n+} bound to N7/O6 atoms of stacked guanines: Mg^{2+} or a monovalent cation? Contrary to the double N7 binding site described above (Figure 5), a pattern where both N7 and O6 atoms belonging to stacked purines coordinate Mg^{2+} has not been identified in the CSD. In the PDB, such a motif with $d(Mg^{2+}\ldots N7/O6) \leq 2.4$ Å is found 8 times and is associated with ApG and GpG steps (Figure 6A). This motif has been first described in a P4–P6 group I intron structure (Figure 6B) and is since cited as a good example of a well-defined Mg^{2+} binding pocket (91–93). However, even at 2.25 Å resolution (PDB code: 1HR2), the ion density is merged to that of the attached waters prohibiting unequivocal Mg^{2+}/Na^{+} identification. In support to this assumption, a water has replaced Mg^{2+} in a related P4-P6 structure (PDB code: 2R8S; resolution: 1.95 Å) marking a poor divalent binding site. Moreover, this binding pattern is reminiscent of that of monovalent ions to carbonyl groups in DNA/RNA quadruplexes where an ion bridges two O6 atoms of 'stacked' guanines. Indeed, a 'semi-quadruplex' binding pattern with Mg^{2+} bound to the carbonyl groups of a GpU step has been identified in the same group I intron fragment (Figure 6C). This Mg^{2+} is more probably Na^{+} given $d(M^{n+}\ldots Ow)$ in the 2.3–2.5 Å range.

Thus, we hypothesize that this site is not occupied by Mg^{2+} but rather by monovalent cations or transition metals as suggested by the binding of a hexacoordinated cobalt ion to a B-DNA structure (PDB code: 4R4A, resolution: 1.45 Å). Such sites could also be occupied by Mn^{2+}, questioning results from ion substitution experiments. To summarize, the significance of this site is limited since the binding of Mg^{2+} to consecutive purines has only been reported in 8 instances, although every ribosome contains on average > 200 similar purine-purine steps.

N7: a secondary Mg^{2+} binding site next to primary anionic oxygens. As discussed above, Mg^{2+} to N7 binding is rare and existing assignments are often questionable. However, in some instances, N7 may correspond to a secondary Mg^{2+} coordination site when the ion is primarily bound to anionic phosphate or carboxylate oxygens. We gathered evidence from CSD structures that when such multiple binding occurs, the distance to the anionic oxygen over the nitrogen atom is systematically shorter by 0.1-0.2 Å (25).

In the PDB, we identified 80 sites with $d(Mg^{2+}\ldots N7/OP) \leq 2.4$ Å for which 71 satisfy the $d(Mg^{2+}\ldots OP) < d(Mg^{2+}\ldots N7)$ criterion. Among those, 51 and 20 sites involve an N7 atom belonging to an adenine and guanine nucleobase, respectively. Yet, a large number of them are redundant. For example, 48 out of 51 adenines are located in a loop capping helix 11 of the large ribosomal subunit and involve an N7 atom and two phosphate groups (Figure 7A). This site is present in all four ribosome families (Supplementary Table S1). Elsewhere in ribosomes, we identified only 8 weak non-redundant sites. These sites are at best occupied by Mg^{2+} satisfying our stereochemical criteria in five instances in the 134 surveyed ribosome structures. Hence, Mg^{2+} directly bound to phosphate groups are rarely establishing direct contacts to N7 atoms given the paucity of appropriate structural contexts in RNA and DNA.

Furthermore, it is important to consider that solvation conditions in the surveyed ensemble of ribosomal structures are very heterogeneous. While the highest populated site (Supplementary Figure S7A) points to Mg^{2+} in high-resolution *H. marismortui* structures (see PDB code: 1VQ8; resolution: 2.2 Å), in two other high-resolution structures from *E. coli* (PDB code: 4YBB; resolution: 2.1 Å) and *T. thermophilus* (PDB code: 4Y4O; resolution: 2.3 Å), $d(Mg^{2+}\ldots N7)$ distances are more consistent with the presence of Na^{+} (Supplementary Figure S7B and C). Further studies are necessary to isolate the factors that favor the binding of one or the other ion to this location.

Besides ribosomal structures, a binding pattern involving a single phosphate group with $d(Mg^{2+}\ldots N7/OP) \leq 2.4$ Å was reported in only two instances. In a c-di-GMP riboswitch (PDB code: 3Q3Z; resolution: 2.51 Å), the bound OP atom is in an equatorial (*cis*) position with respect to the N7 atom (Supplementary Figure S8A) while it is opposite (*trans*) to the N7 atom in a DNA quadruplex (Figure 7B and Supplementary Figure S8B). Interestingly, these patterns involve crystal contacts that are part of the ion coordination shell. Lastly, in a CRISPR-Cas RNA complex (Figure 7C), a single hexacoordinated Mg^{2+} bound to a Glu carboxylate ligand in *trans* has been identified (94). This ion is found in a tight binding pocket at the RNA/protein interface. As shown elsewhere, one of the Mg^{2+} first shell water forms a hydrogen bond with an oxygen of the carboxylate group (95).

The ion placement should be checked carefully when $d(Mg^{2+}\ldots OP) > d(Mg^{2+}\ldots N7)$. For example, see Sup-

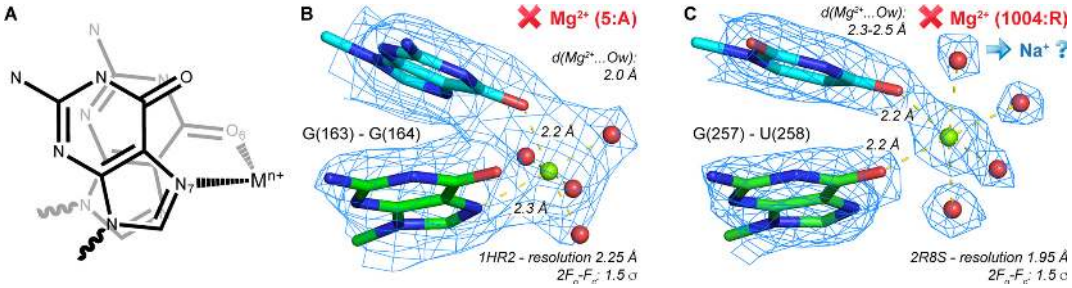

Figure 6. M^{n+} bound to N7/O6 atoms of RpG steps. (**A**) Schematical representation of this ion binding pattern. (**B**) Mg^{2+} binding as reported in a group I intron structure. (**C**) Probable Na^+ binding observed in a group I intron structure of slightly better resolution.

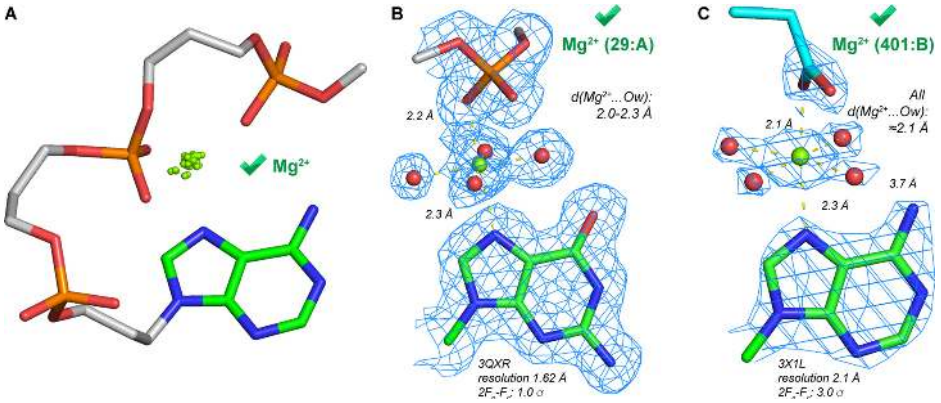

Figure 7. Mg^{2+} bound to N7 and anionic oxygens. (**A**) Overlap of 51 Mg^{2+} found in helix 11 of large ribosomal subunits with $d(Mg^{2+}\ldots N7/OP) \leq 2.4$ Å. Loop configuration is taken from a *H. marismortui* structure (PDB code: 4V9F; resolution 2.4 Å). All structures were superimposed on the adenine base. (**B**) Mg^{2+} bound to a (G)N7 and a phosphate group in B-DNA (crystal contact); separate density peaks for ion and water allow for more reliable ion identification. (**C**) Mg^{2+} bound to a (G)N7 and a glutamate carboxyl group in a RNA/protein complex.

plementary Figure S8C where $d(Mg^{2+}\ldots N7) \approx 2.3$ Å and $d(Mg^{2+}\ldots OP) \approx 2.4$ Å (PDB code: 462D; resolution: 2.3 Å). Here, the coordination distances strongly suggest the presence of Na^+. Note that the distances to phosphate oxygens (when not restrained) are much more reliable and accurate than those to nitrogens. Yet, these examples of simultaneous binding to N7 and anionic oxygens remain exceptional.

Suspicious Mg^{2+} binding occurrences in the 2.6-3.2 Å range: Na^+, K^+, NH_4^+ or water?

As stressed by Table 1 and Figure 2, Mg^{2+} are often placed in the 2.4–3.2 Å exclusion zone that corresponds to the coordination distance range for $Na^+/K^+/NH_4^+$ and water. Na^+ ions were clearly identified in several structures with resolution $\leq$ 2.0 Å. For example, in a hammerhead ribozyme (PDB code: 3ZP8; resolution: 1.55 Å) (67), two out of sixteen Na^+ are bound to (G)N7 and one is bound to (A)N7. The associated octahedral coordination patterns are similar to those for Mg^{2+} in Z-DNA (Figure 4D) with, however, $d(Na^+\ldots N7) \approx 2.4$–2.6 Å. These Na^+ are associated with density patterns showing clearly identifiable metal-bound water molecules. One of these residues (G10; Figure 8A) is often linked to the hammerhead ribozyme catalytic mechanism (32,96). This residue is also associated with direct Mn^{2+} binding but was never unambiguously shown to be in direct contact with Mg^{2+}. Further, 12 examples of Na^+ to N7 contacts, where Na^+ displays an octahedral coordination, are found in 8 structures with resolutions $\leq$ 2.0 Å (PDB codes: 2R1S, 2R20, 2R21, 2R22, 3ND4, 3ZP8, 3DIL, 3PNC). Interestingly, besides Z-DNA structures, no Mg^{2+} to N7 contacts with separate water densities were identified. It is possible that the binding of these hydrated Na^+ is induced by the crystallization buffers since a 1.7 M sodium malonate or NaCl buffer were used to crystallize a hammerhead ribozyme and a *H. marismortui* large ribosomal subunit, respectively (PDB codes: 3ZP8, 1S72; resolutions: 1.55, 2.4 Å). On the contrary, the authors of a lysine riboswitch structure containing 29 well-resolved Na^+ ions (PDB code: 3DIL; resolution: 1.9 Å) mentioned the use of a ≈0.1 M sodium citrate buffer (97). These data shake the common idea that Na^+ octahedral coordination is difficult to observe due to a weaker stability of its hydration shell compared to hydrated Mg^{2+}. Such a belief might have led to misidentifications in *H. marismortui* where octahedral densities with Na^+ coordination distances in the 2.4-2.6 Å range were attributed to Mg^{2+} and where Na^+ labels were used for species with coordination distances in the 2.8–3.2 Å range that are more typical for K^+ ions (14). Indeed, the octahedral Mg^{2+}/Na^+ coordination geometries are difficult to distinguish when the refinement protocols involve distance restraints (Figure 8B).

Next to Na^+, binding of K^+ to nucleic acid N7 atoms is rarely observed. We identified only 79 instances with $d(K^+\ldots N7)$ in the 2.6–3.2 Å range; 10 of those are found in structures with resolution $\leq$ 2.0 Å (PDB codes: 5EW4,

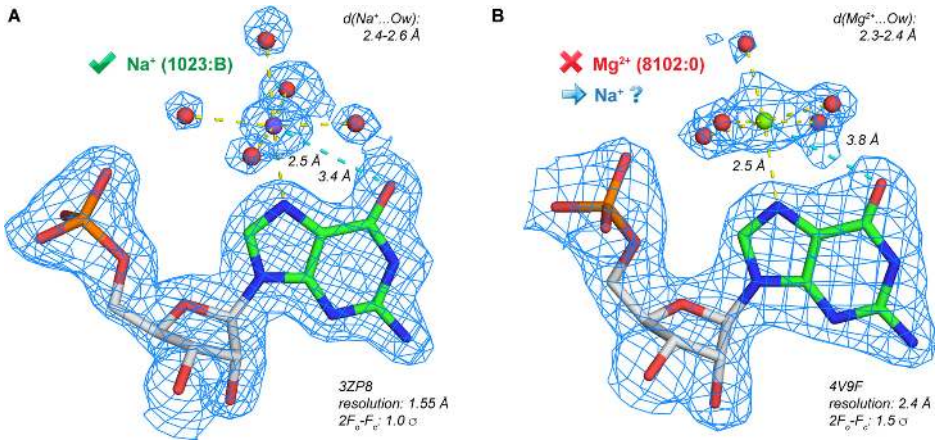

Figure 8. Na^+ coordination. (**A**) Pentahydrated Na^+ bound to (G10)N7 in a hammerhead ribozyme structure. (**B**) Based on coordination distances in the 2.3–2.5 Å range, this pentahydrated Mg^{2+} is probably a Na^+. Note that some water molecules display isolated density blobs even at a 2.4 Å resolution.

1HQ1, 5EW7, 1DUL, 4WO2, 4CN5, 4YAZ). The detection of such ions is complicated by their weak binding affinity and a less nicely defined non-octahedral coordination shell involving preferentially eight ligands with coordination distances similar to those of water molecules (≈2.8 Å). Therefore, K^+ is difficult to distinguish from water especially in case of mixed water/K^+ occupancy. When K^+ is mentioned in the crystallization conditions, anomalous diffraction experiments should systematically be conducted to detect its presence (36,56,98,99).

NH_4^+ ions are common in crystallization buffers due to the recurrent use of $(NH_4)_2SO_4$. Hence, many water molecules close to nucleotides could correspond to hidden NH_4^+ ions and this hypothesis should be seriously considered (14,50). When it occurs, binding of NH_4^+ resembles that of water although, instead of being surrounded by two donor and two acceptor atoms, these ions should be surrounded by four acceptor atoms. Such differences are very subtle and NH_4^+ was assigned in only four structures at resolutions $\leq$ 2.5 Å (100).

Mg^{2+} replacing co-solvent molecules like polyamines: can this happen? When resolution is insufficient and/or data treatment inappropriate, co-solvent molecules like polyamines might remain hidden. In a refinement using low temperature data, isolated 'water' peaks converted into a 'tube' of electron density and resulted in the correct placement of a spermine molecule. It was inferred that at room temperature, the methylene groups were thermally disordered while the more ordered amino groups, which are stabilized through direct hydrogen bonds, appeared as spheres of electron density (35).

In a fragment of a human signal recognition particle (101), we identified several Mg^{2+} at 2.8 Å from N7 atoms (Figure 9A) and none in the appropriate 2.1–2.4 Å coordination range. A closer examination of the $2F_o-F_c$ maps revealed a tube of density that could be interpreted as resulting from the presence of a polyamine. Tentatively, we placed a spermine molecule into this density and suggest that this model, supported by the presence of spermine in the crystallization buffer, constitutes a reasonable working hypothesis (Figure 9B). In a combined x-ray/neutron Z-DNA diffraction structure in complex with a spermine molecule (PDB codes: 1WOE, 1V9G; resolutions: 1.5, 1.8 Å), ammonium

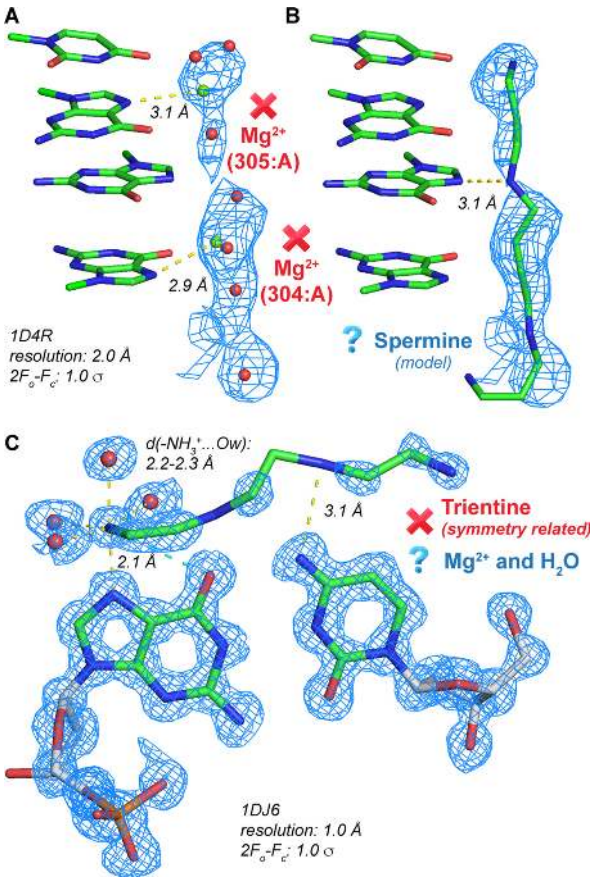

Figure 9. Polyamine misattributions. (**A**) Mg^{2+} ions with inapropriate coordination distances are close to N7 atoms in a human SRP helix 6 structure. (**B**) A spermine molecule—spermine is mentioned in the crystallization conditions—has tentativelly been fitted into the electron density in place of the original Mg^{2+} and water molecules. (**C**) A misplaced symmetry related polyamine lined up on the major groove of a Z-DNA G=C pair. Note the coordination pattern of the hydrated $-NH_3^+$ head that fits a pentahydrated Mg^{2+} (see Figure 4D).

groups are at hydrogen bond distance to both guanine N7 and phosphate groups (78) (see also PDB code: 4HIG, 2F8W; resolutions: 0.8, 1.2 Å) indicating that N7 sites are good docking spots for ammonium groups.

We traced also the opposite type of misidentification, namely a polyamine positioned at Mg^{2+} binding sites. In a Z-DNA hexamer (102), a symmetry related polyamine is unusually lined up on the major groove of a terminal G=C pair. Here, a $-NH_3^+$ group is at 2.1 Å from a N7 atom (Figure 9C) and the coordination is similar to that shown Figure 4D. Further, the Mg^{2+} ion placed in this structure is at 2.6 Å from the closest oxygen and its coordination shell is not octahedral. This observation stresses that odd solvent density interpretations occur even in high-resolution structures.

Mixed Mg^{2+} and monovalent cation/water occupancies: are they meaningful? It has been reported that some DNA major groove hexahydrated Mg^{2+} binding sites are not fully occupied but that a monovalent cation can partially occupy such a site (35). The latter event was identified through anomalous diffraction experiments involving Tl^+ ions (103). Furthermore, K^+ or NH_4^+ could overlap with inner-sphere water molecules of a hexacoordinated Mg^{2+} (104,105). However, it is less likely that water overlays with Mg^{2+} in direct contact to a N7 atom. Yet, this was reported in a Z-DNA structure (PDB code: 1ICK; resolution: 0.95 Å) where a 0.24 occupancy water and a 0.76 occupancy Mg^{2+} share the same position (106). This site is similar to the Z-DNA Mg^{2+} binding site described for Z-DNA (Figure 4D) and illustrates the outcomes of unusual protocols employed to satisfy crystallographic constraints. Although such quirks are rare, they are present in high-resolution crystal structures as mentioned above, a fact that should not be ignored when surveying structural databases.

Direct Mg^{2+} binding to imine N1/N3 nitrogens is unlikely

Mg^{2+} binding to N1 and N3 sites is observed in only nine non-redundant instances and appears at best marginal (Table 1). None of these Mg^{2+} binding occurrences passes visual scrutiny. All Mg^{2+} have one or more contacts in the 2.5–3.0 Å exclusion zone. Titration experiments suggest that Mg^{2+} and Ca^{2+} close to these imine atoms form outer sphere complexes (107). Hence, these inner-sphere contacts represent obvious misattributions.

For instance, an ill-placed Mg^{2+} ion in front of an (A)N1 atom is found in a *T. thermophilus* ribosome structure (108) (see the density extension on the adenine Watson–Crick edge; Figure 10A). Here, the 2.9 Å distance to an amino group of an arginine amino acid suggests the binding of a water molecule. Similarly, in a thiamine riboswitch, a direct binding of Mg^{2+} to N3 with a tetrahedral coordination and long Mg^{2+} to ligand distances suggests the presence of a water molecule rather than an ion (Figure 10B).

Mg^{2+} do not bind to the N7 atom of purine containing metabolites

We checked if Mg^{2+} to N7 binding could be associated with purine containing metabolites, such as ATP, by using the Relibase+ program to search the PDB (109). In the ≤ 3.0 Å resolution range, close to 25 000 such metabolites were identified but only four binding sites with $d(Mg^{2+}\ldots N7)$ in the 3.0–3.5 Å range, and none with $d(Mg^{2+}\ldots N7) \leq 3.0$ Å. In the best resolution structure (PDB code: 4H2U; resolution: 2.1 Å), Mg^{2+} is at a 2.8–3.0 Å distance from all its ligands including a phosphate oxygen and a positively charged Arg side chain. This unambiguous result strongly illustrates the poor Mg^{2+} binding potential of N7 atoms.

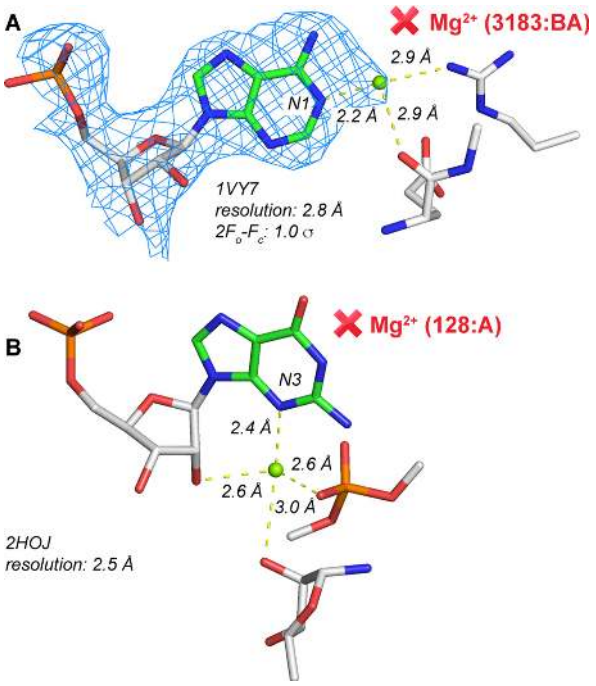

Figure 10. Missassigned Mg^{2+} ions close to imine N1/N3 atoms. (**A**) This figure illustrates the pitfalls of placing ions into poorly defined density patterns. See, for example, the unrealistic Mg^{2+} to arginine contact. (**B**) The tetrahedral coordination inferred from the solvent density at the N3 site and $d(Mg^{2+}\ldots N/O)$ in the 2.4–3.0 Å range suggest the presence of a water molecule and excludes that of Mg^{2+}.

Resolutions > 3.0 Å

If serious identification issues arise at resolutions better that 3.0 Å (≤ 3.0 Å), such issues are certainly much more severe at lower resolutions. In that respect, it is important to note that the PDB contains a significant number of structures with resolutions lower than 3.0 Å. Many of these structures comprise Mg^{2+}, Na^+ and even NH_4^+ ions (110). For instance, we counted in this resolution range 84 nucleic acid structures containing Na^+ (including 12 ribosomes) while 43 Mg^{2+} and 2 K^+ containing structures at resolutions lower than 4.0 Å as well as 63 cryo-EM structures containing Mg^{2+} with resolutions > 3.0 Å were deposited to the PDB. Although crystallography is making significant progress, we believe that assigning light mono- and divalent ions at such resolutions can be detrimental to the crystallographic process and problematic in the development of data mining tools since misinterpretation odds are too high (43,52). In this resolution range (> 3.0 Å), Na^+, Mg^{2+} and their hydration shell are essentially modeled (see below) and, consequently, should be excluded from database surveys.

About the use of coordination distance restraints and modeled hydration spheres

As noted above, the default refinement procedures often involve the use of restraints to place water molecules coordinated to Mg^{2+}. Therefore, their positions are approximated or sometimes entirely modeled. The most obvious example comes from water with $d(Mg^{2+}\ldots Ow) = 2.18$ Å. Such water molecules represent a large part of those that are bound to Mg^{2+} in the PDB (Figure 11). The use of restraints might help to position properly the octahedral coordination shell. However, it has several important drawbacks. The first is that this coordination distance is not appropriate for Mg^{2+} since it is intermediate between the 2.07 Å coordination distance expected for $d(Mg^{2+}\ldots Ow)$ and the ≈2.40 Å coordination distance expected for $d(Na^{+}\ldots Ow)$. Thus, the use of restraints to model the Mg^{2+} hydration shell might make impossible the unambiguous assignment of the electron density peak to Mg^{2+} or Na^{+}. This is especially true when water and ion densities are merged. In those cases, since we identified well defined coordination shells for Mg^{2+} and Na^{+} in structures with resolution ≤ 2.0 Å, the possibility of Na^{+} coordination should at least be considered during the refinement process.

The use and implications of crystallographic restraints have already been noted elsewhere as well as the less frequent but more appropriate use of ≈2.1 Å restraints (12). Finally, we note that restraints are mainly used for Mg^{2+} and rarely for other ions such as Na^{+} and Mn^{2+}, as deduced from the $d(Na^{+}/Mn^{2+}\ldots Ow)$ histograms (Figure 11 and Supplementary Figure S9). For these ions and at least in nucleic acid structures, restraints do not seem necessary.

Because of the issues mentioned here, it appears worthwhile to tag modeled water molecules associated with the systematic use of restraints especially at resolutions > 3.0 Å. Occupancies could be set to zero as is already done by some authors for polyatomic ligands. Specific identifiers could be added to the more accommodating mmCIF files. But, as noted elsewhere, non-crystallographers visualizing a biomolecular system might not be aware of the presence of modeled waters (40). Therefore, we suggest that visualization programs should include an option to turn 'on' the modeled part of the structure that should remain hidden when the structure is first opened. Turning on the visualization of the modeled part of the structure should require a voluntary action. Lastly, the use of restraints should be systematically mentioned in PDB headers and validation reports.

Ion substitution experiments

To identify Na^{+} or Mg^{2+} when the resolution is insufficient, replacement strategies are used (14,111,112). However, they do not supersede direct evidence obtained from high-resolution structures. This is especially true when Mn^{2+} ions are used as substitutes since the affinity of Mn^{2+} over Mg^{2+} for N7 is higher (20,25). Further, although rarely described, Mn^{2+}/Mg^{2+} substitutions can induce significant structural changes; see for instance a crystallographic study of a signal recognition particle (104). There, Mn^{2+} changed the conformation of a nucleotide by linking the N7 to a phosphate oxygen from a neighboring residue. For another RNA structure for which soaking with 13 different metals was performed, a similar conformational change induced by the binding of Mn^{2+}, Zn^{2+} and Co^{2+} with respect to the native Mg^{2+} structure was reported, resulting in a direct N7 to M^{n+} contact (113). Indeed, Mn^{2+} are not perfect substitutes for Mg^{2+} and replacement of $MgCl_2$ by $MnCl_2$ in *H. marismortui* crystallization buffers resulted systematically into twinned crystals (14).

Such ion-induced conformational changes might be more frequent than expected in crystallographic structures with insufficient resolution or in spectroscopic experiments such as electron paramagnetic resonance (EPR) and NMR (8). Drawbacks of substitution experiments might even become worse when larger transition metals like Zn^{2+} or Cd^{2+} are used. Indeed, strong binding of transition metals to N7 sites significantly affects the nucleobase chemical properties (21). It has been shown that Cd^{2+} binding to (G)N7 leads to an acidification of the N1 imino group that can consequently deprotonate at physiological pH and affect the interpretation of biochemical experiments (24).

Thus, in the best resolution structure of $tRNA^{Phe}$ (PDB code: 1EHZ; resolution: 1.93 Å), anomalous data derived from crystals soaked with $MnCl_2$ and $CoCl_2$ were used to identify ion binding sites. In this structure, all four sites close to N7 atoms were associated with Mn^{2+} or Co^{2+} and unrealistically short 2.0 Å metal-water restraints were used. In the parent $tRNA^{Phe}$ structure that was obtained without the use of soaking procedures (PDB code: 1EVV; resolution: 2.0 Å), no N7-bound ions were reported (114).

Further, numerous soaking experiments were performed on the hammerhead ribozyme showing consistently the presence of a transition metal bound to (G10)N7 and, in a high-resolution structure, the presence of Na^{+} (32,96). Despite significant efforts, no direct evidence of Mg^{2+} binding to this site has been reported (Figure 8A).

Hence, binding sites presenting new coordination topologies should not be proposed based uniquely on substitution experiments combined or not with anomalous diffraction data unless a similar binding site backing up the proposed topology has unambiguously been identified in unrelated high-resolution structures (115–117). Given the affinity of Mn^{2+} for N7, this transition metal could eventually replace Mg^{2+} but more probably also Na^{+} commending great care in the interpretation of ion substitution experiments. This is not purely speculative since we are aware of at least one example of substitution of Na^{+} by Mn^{2+} in a protein crystal structure (118).

Comparison with the MgRNA database

The abundance of poorly modeled or incorrectly identified Mg^{2+} ions in nucleic acids has already been noted and was taken into account in an attempt to build an exhaustive and comprehensive classification of Mg^{2+} binding sites, including 41 inner-sphere coordination patterns among which eight are associated with nitrogen sites (12). The MgRNA database defines a set of rules to separate good from bad ion assignments (52). Based on a complex combination of geometrical and crystallographic criteria derived from those used by the CheckMyMetal web server (37), benchmark sets for each binding pattern were defined, embracing 15% of the

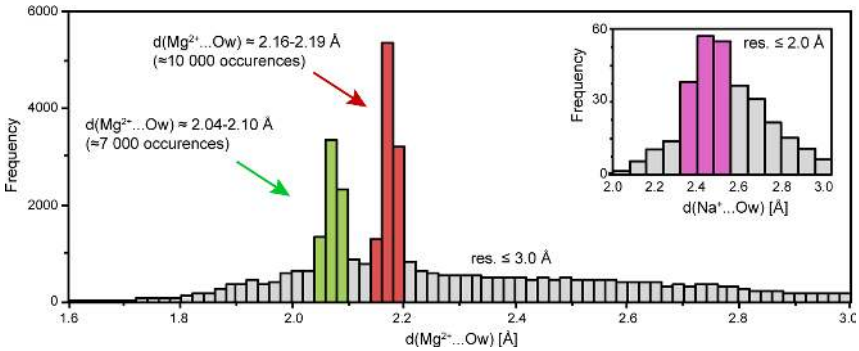

Figure 11. d(Mg^{2+}…Ow) histogram for nucleic acid crystal structures (PDB; May 2016; resolution ≤ 3.0 Å) that emphasize the systematic use of crystallographic restraints around 2.07 and 2.18 Å. The *d*(Na^{+}…Ow) histogram peaks around 2.4 Å and no peaks associated with crystallographic restraints are apparent (see insert; resolution ≤ 2.0 Å). Supplementary Figure S9 displays *d*(Mn^{2+}/Na^{+}…Ow) histograms where peaks due to the use of restraints are also absent (resolution ≤ 3.0 Å).

full dataset—for details, see (12). However, we believe that these criteria were not restrictive enough to exclude all dubious coordination patterns, resulting in a still considerable overestimation of the Mg^{2+} to N7 binding.

In this section, we identify shortcomings in the criteria defined by MgRNA that make further investigations necessary and describe methods to improve them. We have to stress that the numbers provided by MgRNA are not directly comparable to ours since we use resolutions ≤ 3.0 Å on a May 2016 dataset while no resolution limits are applied in MgRNA on the September 2014 dataset. Here, we adopt the MgRNA nomenclature where O_{ph} corresponds to phosphate oxygens (OP1/OP2); O_r to O2'/O4'/O3'/O5' ribose oxygens; O_b to nucleobase oxygens and N_b to nucleobase nitrogens.

In MgRNA, 284 Mg^{2+} to N1/N3/N7 (N_b) contacts were identified and placed in the benchmark dataset (Figure 4). The authors chose a representative of this category in the 2QOU ribosome (resolution: 3.93 Å) that shows a perfectly modeled hydration shell with *d*(Mg^{2+}…N7) = 2.18 Å and *d*(Mg^{2+}…Ow) = 2.08 Å, raising once more the issue of modeled water molecules (see above). Cleary, such a structure including a modeled hydration shell is not representative. Overall, the N_b benchmark dataset contains 77 structures with resolution ≤ 3.0 Å and 92 structures with resolutions > 3.0 Å. The lowest resolution structure is 4V5Y (resolution: 4.45 Å). Among the 77 structures with resolution ≤ 3.0 Å, 127 Mg^{2+} were identified but only 12 of them satisfy the d(Mg^{2+}…N7) ≤ 2.4 Å criterion. The largest d(Mg^{2+}…N7) is 3.17 Å (PDB code: 4PEA; resolution: 2.95 Å) and the average d(Mg^{2+}…N7) of 2.64±0.40 Å is too long for Mg^{2+} to N7 contacts. Based on these data, only 12 out of the 284 sites satisfy our criteria although their coordination shell is far from being strictly octahedral.

For the double N7 site discussed above ($2N_b$), called 'purine N7-seat' in MgRNA (Figure 5), 158 occurrences constitute the benchmark set. The representative site is extracted from an *E. coli* ribosome structure (PDB code: 2I2V; resolution: 3.22 Å), that used d(Mg^{2+}…N7) ≈2.08 Å restraints. Monovalent cation or transition metal binding was not considered. This double N7 binding site was tagged as a novel Mg^{2+}-binding motif, although it has been mentioned elsewhere (8,14).

We already addressed the direct binding of an ion to N7 and a nucleobase oxygen atom (O_b.N_b) (Figure 6). The representative site was taken from 1HR2 (resolution: 2.25 Å). To us, it is unlikely that Mg^{2+} binds to this site, which is most probably involved in the binding of a monovalent cation, possibly Na^{+}. Further, only 26 instances are found in the MgRNA benchmark dataset stressing its limited relevance.

Occurrence of simultaneous binding to N7 and two base or sugar atoms were categorized into three binding types ($2O_b$.N_b, O_b.$2N_b$, $2O_r$.N_b) derived from only five crystal structures with resolution in the 3.3-3.9 Å range. No structure with sufficient resolution to interpret solvent binding details is available to support the genuineness of these modeled sites. Therefore, they should not be labeled as Mg^{2+} binding sites.

The representative O_{ph}.N_b site (PDB code: 3R8S; resolution: 3.0 Å) displays good coordination distances to water, phosphate oxygen and N7 atoms. However, a closer examination revealed that all the $Mg(H_2O)_4^{2+}$ but also all the neighboring nucleotide *B*-factors display an unrealistic 0.01 $Å^2$ value. In 1VQ8 (resolution: 2.2 Å), *d*(Mg^{2+}…N7) is stretched to 3.12 Å. Indeed, all these occurrences in the benchmark dataset have not been identified by us mainly because of inappropriate resolution and/or *d*(Mg^{2+}…N7) > 2.4 Å. Thus, this Mg^{2+}-binding motif should be excluded from the MgRNA classification.

The benchmark set for the last MgRNA site, labelled *cis*-$2O_{ph}$.N_b, comprises 118 occurrences. The representative site is extracted from the 1VS6 ribosome structure with a 3.46 Å resolution and *d*(Mg^{2+}…Ow) = 2.08 Å. Overall, for this dataset, the level of redundancy is high. The 118 sites, exclusively identified in ribosomes, are found at two locations involving (A)N7 or (G)N7. We identified 49 out of 118 sites in structures with resolution ≤ 3.0 Å. Among those with (A)N7, 18 have *d*(Mg^{2+}…N7) ≤ 2.4 Å and 15 have *d*(Mg^{2+}…N7) > 2.4 Å. The remaining 16 occurrences involve (G)N7 with *d*(Mg^{2+}…N7) > 2.4 Å. One of these MgRNA sites, called '10-member ring with purine N7', has been identified in the present study (Figure 7). Again, the significance of this site is low and there is not enough evidence to suggest its presence other than in rare and highly specific ribosomal pockets. This site can probably also accommodate Na^{+} ions (Supplementary Figure S7).

Consequently, MgRNA still contains a large number of misidentified Mg^{2+} and was not successful in creating reliable benchmark datasets. We identified several factors that led to such issues. First, the process could be improved if structures with resolutions > 3.0 Å were excluded (38,40). Second, strict enforcement of $d(Mg^{2+}\ldots N/O)$ cutoffs would lead to a significant reduction of false positive. Although uncertainties in the coordination distances are difficult to estimate (23,37), it seems problematic to accept distances >2.4 Å to validate new Mg^{2+} binding sites. Third, coordination distance issues involving restrained water molecules have to be identified more systematically (12,23,37). In that respect, indicators based on the bond valence theory should be considered with caution for resolutions > 2.0 Å (119). Such indicators should not be used when restraints on coordination bonds are present. As an outcome, when stricter criteria are used, the eight binding sites described by MgRNA, reduce to two for which we found a limited number of convincing occurrences, namely N_b and *cis*-$2O_{ph}.N_b$.

Mg^{2+} ion assignment and validation checklist

In order to facilitate the ion assignment process, we defined a set of rules regarding the placement of ions in solvent electron densities next to N7 atoms that can easily be extended to the binding of ions to other sites (Table 3). In that perspective, we would like to stress a few points that we consider of importance. First, numerous competing ionic species might be present in crystallization buffers, sometimes as contaminants and should be taken into account (70). For instance, in our survey, it was not immediately apparent that Zn^{2+} could bind to a specific ribosome site (Figure 5) especially since Zn^{2+} is not mentioned in the crystallographic conditions. Therefore, it is important to integrate excess electron density that can reveal the presence of transition metals or electron rich K^+. Second, an ion *B*-factor lower than those of the bound nucleobase or water molecules or an ion occupancy significantly higher than 1.0 should hint to the presence of an electron rich atom. In such instances, anomalous diffraction data should be collected at the appropriate wavelengths. It has to be noted that a large excess of unassigned electron density might affect not only the position of the ion and its hydration shell, but can also wrongfully force nucleobases to come closer to the excess electron density center leading to unreliable coordination distances (43,117). On the other hand, weak electron density patterns manifested by high *B*-factors or negative F_o-F_c peaks suggest 'wishful' ion attributions. Anions such as Cl^-, SO_4^{2-} or even cacodylate are also often disregarded (68,69). If the identity of an ion is inferred from binding sites observed in a different structure, the original data should be carefully checked including the electron density peak height, *B-factor* value, coordination number, bond distances and angles as well as the $2F_o-F_c$ and F_o-F_c maps in order to avoid replicating errors. Finally, when no reasonable solution emerges, protonated and tautomeric forms of the coordinated nucleobase or the surrounding residues should be considered (120).

In case of doubt, and especially in the $d(Mg^{2+}\ldots N7)$ ≈3.2–3.8 Å exclusion range (Figures 1 and 2), density patterns should not be assigned to Mg^{2+}. Such patterns are probably related to the presence of other ionic or molecular species present in the crystallization buffer or as contaminants. X-ray data are also prone to experimental errors that might result in weak/spurious electron density peaks (121,122). In those instances, density assignment is counterproductive even if it reduces R_{work} and R_{free} values. It can here be reminded that the PDB allows to use the UNK code for placing an atom at positions where atom identity is uncertain.

We have to stress that the chosen cutoff distances are merely indicative. They are less stringent for Mg^2 binding to nitrogen than to oxygen. These distances will be refined in further studies when more high-quality data become available. A rule of thumb is that, at least in nucleic acids, $d(Mg^{2+}\ldots N7) \leq 2.4$ Å and $d(Mg^{2+}\ldots N7) > d(Mg^{2+}\ldots O)$. Caution should be exerted when restraints are used, especially with the 2.18 Å default value, suggesting that the data do not allow to differentiate Mg^{2+} from Na^+ and that further refinement without restraints should be conducted.

Further reasons can lead to bad ion assignments among which we list: (i) the possibility that ions were placed automatically or without great care into density blobs in order to lower the R_{free} value; (ii) existing stereochemical knowledge was ignored; (iii) wishful thinking; (iv) the replication of errors already present in PDB structures and (v) overestimation of the amount of information that can be extracted from low-resolution structures. Hence, it is suggested to exclude structures displaying obvious ion identification errors from database surveys, at least as far as ion placement is concerned (40).

CONCLUSION AND PERSPECTIVES

Based on the data we gathered, we conclude that nearly all the Mg^{2+} to N7 contacts reported in PDB structures need to be reexamined and propose a Mg^{2+} assignment checklist to facilitate this endeavor. Indeed, non-ambiguous examples of Mg^{2+} binding to N7 are excessively rare and are limited to a few occurrences where Mg^{2+} binding seems to result from unique crystallization conditions and/or is associated with primary contacts with anionic oxygens. Additionally, we noted that none of the 25 000 purine metabolites from the PDB establish Mg^{2+} to N7 contacts, an additional strong evidence that Mg^{2+} does rarely bind to purine N7 sites. Consequently, we conclude that almost all Mg^{2+} assignments to solvent density in front of N7 atoms, as found in PDB structures, are incorrect. This outcome significantly diverges from that presented by the MgRNA survey that identified 8 binding modes involving imine nitrogens in opposition to barely two by us (12).

Interestingly, we characterized a potential Zn^{2+} binding site in prokaryotic ribosomal structures that involves two head-to-tail stacked purines in the core of a three-way junction, a finding that opens a new window on the complexity of the metal/nucleic acid ecosystem (123,124). However, we were unable to establish if these transition metal binding sites are populated *in vivo* or if they may only be found under specific *in crystallo* conditions.

From a purely methodological point of view, the most interesting outcome of this study resides in the recognition of

Table 3. Ion to N7 assignment and validation checklist

Ion to N7 assignment and validation checklist			
$d(M^{n+}…N7) \leq 2.4$ Å	$2.4 \leq d(M^{n+}…N7) \leq 2.6$ Å	$2.6 \leq d(M^{n+}…N7) \leq 3.2$ Å	
→ Mg^{2+} • Octahedral coordination • In plane • $d(Mg^{2+}…N6/O6) \approx 3.8$ Å • $d(Mg^{2+}…Ow) \approx 2.07$ Å • $d(Mg^{2+}…N7) > d(Mg^{2+}…O)$ **→ Transition metals** • Check for excess electron density • Use anomalous data when possible	**→ Na^{+}** • Octahedral coordination • In/out of plane • $d(Na^{+}…O6/N6) \approx 3.8$ Å • $d(Na^{+}…Ow) \approx 2.4$ Å • $d(Na^{+}…N7) \approx d(Na^{+}…O)$	**→ K^{+}** • Coordination > 6 • $d(K^{+}…Ow) \approx 2.8$ Å • $d(K^{+}…N7) \approx d(K^{+}…O)$ • Possibility of partial occupancy (higher than expected *B-factor*) • Check for excess electron density • Use anomalous data when possible	**→ NH_4^{+}** • Coordination 4 • Tetrahedral (4 acceptors) • $d(NH_4^{+}…Ow) \approx 2.8$ Å • $d(NH_4^{+}…N7) \approx 2.8$-3.2 Å **→ H_2O** • Coordination 4 • Tetrahedral (2 acceptors – 2 donors) • In plane • $d(Ow…Ow) \approx 2.8$ Å • $d(Ow…N7) \approx 2.8$-3.2 Å
General rules about resolution: • Avoid placing light ions (Na^{+}, Mg^{2+}) in structures with resolutions > 3.0 Å; be very careful in the 2.5-3.0 Å range; in this resolution range, it is almost impossible to distinguish Mg^{2+} from water and Na^{+}. Eventually, consider placing ions at locations for which prior-knowledge has been gathered from several independent high-resolution structures. Always keep in mind that both, Mg^{2+} and Na^{+}, can fit equally well the electron density; **General rules about ion substitutions:** • Consider that transition metals (Mn^{2+}, Zn^{2+}, …) might locally induce conformational changes; • Consider that Na^{+} can be replaced by transition metals; **General rules about crystallization conditions:** • Check for all ions and solvent molecules that might be present in the crystallization buffers; • Do not exclude possible contaminants; • A badly interpreted polyatomic solvent density might correspond to ions and/or water; **General rules about crystallographic parameters:** • In all instances, *B-factor* (nucleobase) < *B-factor* (ion) < *B-factor* (water); • Check for unusual occupancies; occupancies significantly larger than one can hide excess densities; • In case of doubt; always check $2F_o$-F_c and F_o-F_c maps; • Questionable electron density peaks might result from experimental noise; some peaks are better to be left unassigned; UNK keyword is a viable option (see PDB format recommendations); **Specific rules for Mg^{2+} ions:** • When the coordination shell is not complete, check if completing it generates clashes; • Ion binding to two N7 atoms hints to the presence of a transition metal or a monovalent cation; • If $d(Mg^{2+}…Ow) = 2.18$ Å restraints are used, consider that the densities could also fit a Na^{+} ion; • Try $d(Mg^{2+}…Ow) = 2.07$ Å and $d(Na^{+}…Ow) = 2.40$ Å instead; • Similar Mg^{2+} binding sites should recurrently be observed in high-resolution structures; • In the $d(Mg^{2+}…N7) \approx 3.2$-3.8 Å, the solvent electron densities should not be assigned to Mg^{2+} but rather to buffer molecules if possible or left unassigned;		

specific and frequent octahedral Na^{+} coordination modes identified in structures with resolution $\leq$ 2.0 Å. Those are very difficult to distinguish from octahedral Mg^{2+} coordination modes, especially when 2.18 Å coordination distance restraints are used during the refinement process. Undeniably, such restraints often combined with poor resolutions do not allow to distinguish Na^{+} from Mg^{2+} based on their respective 2.40 and 2.07 Å coordination distances. We suggest that the use of restraints on Mg^{2+} coordination distances is probably at the origin of the large number of Mg^{2+} misidentifications in nucleic acids and that Na^{+} binding should always be considered as an alternative.

To conclude, we strongly believe that careful visual examination of crystallographic data is needed in order to create a reliable 'prior knowledge' dataset before developing or using automatic assignment protocols, pattern detection algorithms or machine learning tools (12,47). Further, 'prior-knowledge' should only be based on Mg^{2+} binding motifs unambiguously characterized in multiple unrelated high-resolution structures and not on circumstantial evidences as it is too often the case. Currently, automatic re-refinement workflows such as PDB-REDO cannot resolve solvent attribution issues that remain one of the last major bottlenecks in the interpretation of crystallographic data (37,44,46,48,49). This work should provide a more solid experimental ground for the development of molecular dynamics force-fields that sometimes rely on the erroneous assumption that N7 is an appropriate Mg^{2+} binding site (12,31).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

P.A. wishes to thank Prof. Eric Westhof for ongoing support and helpful discussions, as well as Prof. Neocles Leontis, Prof. Richard Giegé, Dr Eric Ennifar and Dr Quentin Vicens for careful reading of the manuscript and for beneficial discussions. The authors wish also to thank Prof. Jamie Cate and Dr Albert Weixlbaumer for constructive discussions on the difficult 'Zn^{2+}' topic as well as Prof. Alexander Serganov on more general crystallographic issues. We thank the referees for their generous comments.

FUNDING

French 'Ministère de la recherche et de l'enseignement' (to L.D.); Polish Ministry of Higher Education and Science (Mobility Plus programme) [1103/MOB/2013/0 to F.L.]. Funding for open access charge: French National Research Agency [ANR-15-CE11-0021-02].
Conflict of interest statement. None declared.

REFERENCES

1. Erat,M.C. and Sigel,R.K. (2011) Methods to detect and characterize metal ion binding sites in RNA. *Met. Ions Life Sci.*, **9**, 37–100.
2. Pechlaner,M. and Sigel,R.K. (2012) Characterization of metal ion-nucleic acid interactions in solution. *Met. Ions Life Sci.*, **10**, 1–42.
3. Pyle,A.M. (2002) Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.*, **7**, 679–690.
4. Woodson,S.A. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.*, **9**, 104–109.
5. Draper,D.E. (2008) RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.*, **95**, 5489–5495.
6. Auffinger,P., Grover,N. and Westhof,E. (2011) Metal ion binding to RNA. *Met. Ions Life Sci.*, **9**, 1–35.
7. Erat,M.C., Coles,J., Finazzo,C., Knobloch,B. and Sigel,R.K. (2012) Accurate analysis of Mg^{2+} binding to RNA: From classical methods to a novel iterative calculation procedure. *Coord. Chem. Rev.*, **256**, 279–288.
8. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reedjik,J and Poeppelmeier,K (eds). *Comprehensive Inorganic Chemistry II*. Elsevier, Oxford, Vol. **3**, pp. 623–660.
9. Auffinger,P., D'Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. *Met. Ions Life Sci.*, **16**, 167–201.
10. Miller-Fleming,L., Olin-Sandoval,V., Campbell,K. and Ralser,M. (2015) Remaining mysteries of molecular biology: The role of polyamines in the cell. *J. Mol. Biol.*, **427**, 3389–3406.
11. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of diamine and Mg^{2+} with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. *Biochemistry*, **52**, 5911–5919.
12. Zheng,H., Shabalin,I.G., Handing,K.B., Bujnicki,J.M. and Minor,W. (2015) Magnesium-binding architectures in RNA crystal structures: validation, binding preferences, classification and motif detection. *Nucleic Acids Res.*, **43**, 3789–3801.
13. Banatao,D.R., Altman,R.B. and Klein,T.E. (2003) Microenvironment analysis and identification of magnesium binding sites in RNA. *Nucleic Acids Res.*, **31**, 4450–4460.
14. Klein,D.J., Moore,P.B. and Steitz,T.A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, **10**, 1366–1379.
15. Stefan,L.R., Zhang,R., Levitan,A.G., Hendrix,D.K., Brenner,S.E. and Holbrook,S.R. (2006) MeRNA: a database of metal ion binding sites in RNA structures. *Nucleic Acids Res.*, **34**, D131–D134.
16. Schnabl,J., Suter,P. and Sigel,R.K. (2012) MINAS—a database of metal ions in nucleic acidS. *Nucleic Acids Res.*, **40**, D434–D438.
17. Lippert,B. (2000) Multiplicity of metal ion binding patterns to nucleobases. *Coord. Chem. Rev.*, **200-202**, 487–516.
18. Erat,M.C., Kovacs,H. and Sigel,R.K. (2010) Metal ion-N7 coordination in a ribozyme branch domain by NMR. *J. Inorg. Biochem.*, **104**, 611–613.
19. Bartova,S., Pechlaner,M., Donghi,D. and Sigel,R.K. (2016) Studying metal ion binding properties of a three-way junction RNA by heteronuclear NMR. *J. Biol. Inorg. Chem.*, **21**, 319–328.
20. Bock,C.W., Katz,A.K., Markham,G.D. and Glusker,J.P. (1999) Manganese as a replacement for magnesium and zinc: functional comparison of the divalent ions. *J. Am. Chem. Soc.*, **121**, 7360–7372.
21. Sponer,J., Sabat,M., Gorb,L., Leszczynski,J., Lippert,B. and Hobza,P. (2000) The effect of metal binding to the N7 site of purine nucleotides on their structure, energy, and involvement in base pairing. *J. Chem. Phys. B*, **104**, 7535–7544.
22. Wang,G., Gaffney,B.L. and Jones,R.A. (2004) Differential binding of Mg^{2+}, Zn^{2+}, and Cd^{2+} at two sites in a hammerhead ribozyme motif, determined by ^{15}N NMR. *J. Am. Chem. Soc.*, **126**, 8908–8909.
23. Harding,M.J., Nowicki,M.W. and Walkinshaw,M.D. (2010) Metals in protein structures: a review of their principal features. *Cryst. Rev.*, **16**, 247–302.
24. Sigel,R.K., Skilandat,M., Sigel,A., Operschall,B.P. and Sigel,H. (2013) Complex formation of cadmium with sugar residues, nucleobases, phosphates, nucleotides, and nucleic acids. *Met. Ions Life Sci.*, **11**, 191–274.
25. Leonarski,F., D'Ascenzo,L. and Auffinger,P. (2016) Binding of metals to purine N7 nitrogen atoms and implications for nucleic acids: a CSD survey. *Inorg. Chim. Acta.*, **452**, 82-89.
26. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell*, **81**, 991–1002.
27. Sigel,R.K. and Pyle,A.M. (2007) Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. *Chem. Rev.*, **107**, 97–113.
28. Ward,W.L., Plakos,K. and DeRose,V.J. (2014) Nucleic acid catalysis: metals, nucleobases, and other cofactors. *Chem. Rev.*, **114**, 4318–4342.
29. Chi,Y.I., Martick,M., Lares,M., Kim,R., Scott,W.G. and Kim,S.H. (2008) Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol.*, **6**, e234.
30. Mir,A., Chen,J., Robinson,K., Lendy,E., Goodman,J., Neau,D. and Golden,B.L. (2015) Two divalent metal ions and conformational changes play roles in the hammerhead ribozyme cleavage reaction. *Biochemistry*, **54**, 6369–6381.
31. Panteva,M.T., Giambasu,G.M. and York,D.M. (2015) Force field for Mg^{2+}, Mn^{2+}, Zn^{2+}, and Cd^{2+} ions that have balanced interactions with nucleic acids. *J. Phys. Chem. B*, **119**, 15460–15470.
32. Mir,A. and Golden,B.L. (2016) Two active site divalent ions in the crystal structure of the hammerhead ribozyme bound to a transition state analogue. *Biochemistry*, **55**, 633–636.
33. Ren,A., Vusurovic,N., Gebetsberger,J., Gao,P., Juen,M., Kreutz,C., Micura,R. and Patel,D.J. (2016) Pistol ribozyme adopts a pseudoknot fold facilitating site-specific in-line cleavage. *Nat. Chem. Biol.*, **12**, 702–708.
34. Nayal,M. and Di Cera,E. (1996) Valence Screening of water in protein crystals reveals potential Na^{+} binding sites. *J. Mol. Biol.*, **256**, 228–234.
35. Williams,L.D. (2005) Between objectivity and whim: nucleic acid structural biology. *Top. Curr. Chem.*, **253**, 77–88.
36. Echols,N., Morshed,N., Afonine,P.V., McCoy,A.J., Miller,M.D., Read,R.J., Richardson,J.S., Terwilliger,T.C. and Adams,P.D. (2014) Automated identification of elemental ions in macromolecular crystal structures. *Acta Cryst.*, **D70**, 1104–1114.
37. Zheng,H., Chordia,M.D., Cooper,D.R., Chruszcz,M., Muller,P., Sheldrick,G.M. and Minor,W. (2014) Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. *Nat. Protoc.*, **9**, 156–170.
38. Wlodawer,A., Minor,W., Dauter,Z. and Jaskolski,M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.*, **275**, 1–21.
39. Zheng,H., Chruszcz,M., Lasota,P., Lebioda,L. and Minor,W. (2008) Data mining of metal ion environments present in protein structures. *J. Inorg. Biochem.*, **102**, 1765–1776.
40. Cooper,D.R., Porebski,P.J., Chruszcz,M. and Minor,W. (2011) X-ray crystallography: assessment and validation of protein-small molecule complexes for drug discovery. *Expert Opin. Drug Dis.*, **6**, 771–782.
41. Pozharski,E., Weichenberger,C.X. and Rupp,B. (2013) Techniques, tools and best practices for ligand electron-density analysis and results from their application to deposited crystal structures. *Acta Cryst.*, **D69**, 150–167.
42. Wlodawer,A., Minor,W., Dauter,Z. and Jaskolski,M. (2013) Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.*, **280**, 5705–5736.
43. Dauter,Z., Wlodawer,A., Minor,W., Jaskolski,M. and Rupp,B. (2014) Avoidable errors in deposited macromolecular structures: an impediment to efficient data mining. *IUCrJ*, **1**, 179–193.

44. Weichenberger,C.X., Afonine,P.V., Kantardjieff,K. and Rupp,B. (2015) The solvent component of macromolecular crystals. *Acta Cryst.*, **D71**, 1023–1038.
45. Raczynska,J.E., Wlodawer,A. and Jaskolski,M. (2016) Prior knowledge or freedom of interpretation? A critical look at a recently published classification of 'novel' Zn binding sites. *Proteins*, **84**, 770–776.
46. Joosten,R.P., Womack,T., Vriend,G. and Bricogne,G. (2009) Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Cryst.*, **D65**, 176–185.
47. Read,R.J., Adams,P.D., Arendall,W.B. 3rd, Brunger,A.T., Emsley,P., Joosten,R.P., Kleywegt,G.J., Krissinel,E.B., Lutteke,T., Otwinowski,Z. *et al.* (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure*, **19**, 1395–1412.
48. Joosten,R.P., Joosten,K., Murshudov,G.N. and Perrakis,A. (2012) PDB_REDO: constructive validation, more than just looking for errors. *Acta Cryst.*, **D68**, 484–496.
49. Joosten,R.P., Long,F., Murshudov,G.N. and Perrakis,A. (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ*, **1**, 213–220.
50. Touw,W.G., Joosten,R.P. and Vriend,G. (2016) New biological insights from better structure models. *J. Mol. Biol.*, **428**, 1375–1393.
51. van Beusekom,B., Perrakis,A. and Joosten,R.P. (2016) Data mining of macromolecular structures. *Methods Mol. Biol.*, **1415**, 107–138.
52. Minor,W., Dauter,Z., Helliwell,J.R., Jaskolski,M. and Wlodawer,A. (2016) Safeguarding structural data repositories against bad apples. *Structure*, **24**, 216–220.
53. Harding,M.M. and Hsin,K.Y. (2014) Mespeus–a database of metal interactions with proteins. *Methods Mol. Biol.*, **1091**, 333–342.
54. Sigel,R.K. and Sigel,H. (2010) A stability concept for metal ion coordination to single-stranded nucleic acids and affinities of individual sites. *Acc. Chem. Res.*, **43**, 974–984.
55. Lamb,A.L., Kappock,T.J. and Silvaggi,N.R. (2015) You are lost without a map: Navigating the sea of protein structures. *Biochim. Biophys. Acta*, **1854**, 258–268.
56. Mueller-Dieckmann,C., Panjikar,S., Schmidt,A., Mueller,S., Kuper,J., Geerlof,A., Wilmanns,M., Singh,R.K., Tucker,P.A. and Weiss,M.S. (2007) On the routine use of soft X-rays in macromolecular crystallography. Part IV. Efficient determination of anomalous substructures in biomacromolecules using longer X-ray wavelengths. *Acta Cryst.*, **D63**, 366–380.
57. Groom,C.R. and Allen,F.H. (2014) The Cambridge Structural Database in retrospect and prospect. *Angew. Chem. Int. Ed. Engl.*, **53**, 662–671.
58. Markham,G.D., Glusker,J.P. and Bock,C.W. (2002) The arrangement of first and second-sphere water molecules in divalent magnesium complexes: results from molecular orbital and density functional theory and from structural crystallography. *J. Phys. Chem. B*, **106**, 5118–5134.
59. Kuppuraj,G., Dudev,M. and Lim,C. (2009) Factors governing metal-ligand distances and coordination geometries of metal complexes. *J. Phys. Chem. B*, **113**, 2952–2960.
60. Martinez,J.M., Pappalardo,R.R. and Marcos,E.S. (1999) First-principles ion-water interaction potentials for highly charge monoatomic cations. Computer simulations of Al^{3+}, Mg^{2+}, and Be^{2+} in water. *J. Am. Chem. Soc.*, **121**, 3175–3184.
61. Lightstone,F.C., Schwegler,E., Hood,R.Q., Gygi,F. and Galli,G. (2001) A first principle molecular dynamics simulation of the hydrated magnesium ion. *Chem. Phys. Lett.*, 549–555.
62. Bhattacharjee,A., Pribil,A.B., Randolf,B.R., Rode,B.M. and Hofer,T.S. (2012) Hydration of Mg^{2+} and its influence on the water hydrogen bonding network via ab initio QMCF MD. *Chem. Phys. Lett.*, **536**, 39–44.
63. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.*, **D66**, 213–221.
64. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst.*, **D67**, 355–367.
65. Harding,M.H. (1999) The geometry of metal-ligand interactions relevant to proteins. *Acta Cryst.*, **D55**, 1432–1443.
66. Kleywegt,G.J., Harris,M.R., Zou,J.Y., Taylor,T.C., Wahlby,A. and Jones,T.A. (2004) The Uppsala electron-density server. *Acta Cryst.*, **D60**, 2240–2249.
67. Anderson,M., Schultz,E.P., Martick,M. and Scott,W.G. (2013) Active-site monovalent cations revealed in a 1.55-A-resolution hammerhead ribozyme structure. *J. Mol. Biol.*, **425**, 3790–3798.
68. Auffinger,P., Bielecki,L. and Westhof,E. (2004) Anion binding to nucleic acids. *Structure*, **12**, 379–388.
69. D'Ascenzo,L. and Auffinger,P. (2016) Anions in nucleic acid crystallography. *Methods Mol. Biol.*, **1320**, 337–351.
70. Giegé,R. (2013) A historical perspective on protein crystallization from 1840 to the present day. *FEBS J.*, **280**, 6456–6497.
71. DePristo,M.A., De Bakker,P.I. and Blundell,T.L. (2004) Heterogeneity and inaccuracy in protein structures solved by x-ray crystallography. *Structure*, **12**, 831–838.
72. Kagawa,T.F., Geierstanger,B.H., Wang,A.H.J. and Ho,P.S. (1991) Covalent modification of guanine bases in double-stranded DNA. The 1.2 Å Z-DNA structure of d(CGCGCG) in the presence of $CuCl_2$. *J. Biomol. Chem.*, **266**, 20175–20184.
73. Drozdzal,P., Gilski,M., Kierzek,R., Lomozik,L. and Jaskolski,M. (2013) Ultrahigh-resolution crystal structures of Z-DNA in complex with Mn^{2+} and Zn^{2+} ions. *Acta Cryst.*, **D69**, 1180–1190.
74. Gessner,R.V., Frederick,C.A., Quigley,G.J., Rich,A. and Wang,A.H.J. (1989) The molecular structure of the left-handed Z-DNA double helix at 1.0 Å atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem.*, **264**, 7912–7935.
75. Ho,P.S. and Mooers,B.H.M. (1997) Z-DNA crystallography. *Biopolymers*, **44**, 65–90.
76. Brzezinski,K., Brzuszkiewicz,A., Dauter,M., Kubicki,M., Jaskolski,M. and Dauter,Z. (2011) High regularity of Z-DNA revealed by ultra high-resolution crystal structure at 0.55 Å. *Nucleic Acids Res.*, **39**, 6238–6248.
77. Chatake,T. and Sunami,T. (2013) Direct interactions between Z-DNA and alkaline earth cations, discovered in the presence of high concentrations of $MgCl_2$ and $CaCl_2$. *J. Inorg. Biochem.*, **124**, 15–25.
78. Chatake,T., Tanaka,I., Umino,H., Arai,S. and Niimura,N. (2005) The hydration structure of a Z-DNA hexameric duplex determined by a neutron diffraction technique. *Acta Cryst.*, **D61**, 1088–1098.
79. Yang,X.L., Robinson,H., Gao,Y.G. and Wang,A.H. (2000) Binding of a macrocyclic bisacridine and ametantrone to CGTACG involves similar unusual intercalation platforms. *Biochemistry*, **39**, 10950–10957.
80. Labiuk,S.L., Delbaere,L.T. and Lee,J.S. (2003) Cobalt(II), nickel(II) and zinc(II) do not bind to intra-helical N(7) guanine positions in the B-form crystal structure of d(GGCGCC). *J. Biol. Inorg. Chem.*, **8**, 715–720.
81. Egli,M., Gessner,R.V., Williams,L.D., Quigley,G.J., Vandermarel,G.A., Vanboom,J.H., Rich,A. and Frederick,C.A. (1990) Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 3235–3239.
82. McLellan,T.J., Marr,E.S., Wondrack,L.M., Subashi,T.A., Aeed,P.A., Han,S., Xu,Z., Wang,I.K. and Maguire,B.A. (2009) A systematic study of 50S ribosomal subunit purification enabling robust crystallization. *Acta Cryst.*, **D65**, 1270–1282.
83. Khatter,H., Myasnikov,A.G., Mastio,L., Billas,I.M., Birck,C., Stella,S. and Klaholz,B.P. (2014) Purification, characterization and crystallization of the human 80S ribosome. *Nucleic Acids Res.*, **42**, e49.
84. Maret,W. (2013) Zinc biochemistry: from a single zinc enzyme to a key element of life. *Adv. Nutr.*, **4**, 82–91.
85. Laity,J.H., Lee,B.M. and Wright,P.E. (2001) Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.*, **11**, 39–46.
86. Dresios,J., Chan,Y.L. and Wool,I.G. (2005) Ribosomal zinc finger proteins: the structure and the function of yeast YL37a. In: UIuchi,S and Kuldell,N (eds). *Zinc Finger Proteins: From Atomic Contact to Cellular Function*. Springer, Boston, pp. 91–98.
87. Hensley,M.P., Tierney,D.L. and Crowder,M.W. (2011) Zn(II) binding to Escherichia coli 70S ribosomes. *Biochemistry*, **50**, 9937–9939.

88. Zhang,J. and Ferre-D'Amare,A.R. (2014) Dramatic improvement of crystals of large RNAs by cation replacement and dehydration. *Structure*, **22**, 1363–1371.
89. Gunasekera,T., Easton,J.A., Sugerbaker,S.A., Klingbeil,L. and Crowder,M.W. (2009) Zn(II) homeostasis in *E. coli*. In: Long,EC and Baldwin,MJ (eds). *Bioinorganic Chemistry*. American Chemical Society, pp. 81–95.
90. Thorn,A. and Sheldrick,G.M. (2011) ANODE: anomalous and heavy-atom density calculation. *J. Appl. Cryst.*, **44**, 1285–1287.
91. Juneau,K., Podell,E., Harrington,D.J. and Cech,T.R. (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA-solvent interactions. *Structure*, **9**, 221–231.
92. Freisinger,E. and Sigel,R.K. (2007) From nucleotides to ribozymes - A comparison of their metal ion binding properties. *Coord. Chem. Rev.*, **251**, 1834–1851.
93. Frederiksen,J.K., Li,N.S., Das,R., Herschlag,D. and Piccirilli,J.A. (2012) Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA*, **18**, 1123–1141.
94. Osawa,T., Inanaga,H., Sato,C. and Numata,T. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol. Cell.*, **58**, 418–430.
95. Dudev,T. and Lim,C. (2004) Monodentate versus bidentate carboxylate binding in magnesium and calcium proteins: what are the basic principles? *J. Phys. Chem. B*, **108**, 4546–4557.
96. Martick,M., Lee,T.S., York,D.M. and Scott,W.G. (2008) Solvent structure and hammerhead ribozyme catalysis. *Chem. Biol.*, **15**, 332–342.
97. Serganov,A., Huang,L. and Patel,D.J. (2008) Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature*, **455**, 1263–1268.
98. Tereshko,V., Wilds,C.J., Minasov,G., Prakash,T.P., Maier,M.A., Howard,A., Wawrzak,Z., Manoharan,M. and Egli,M. (2001) Detection of alkali metal ions in DNA crystals using state-of-the-art X-ray diffraction experiments. *Nucleic Acids Res.*, **29**, 1208–1215.
99. Ennifar,E., Walter,P. and Dumas,P. (2003) A crystallographic study of the binding of 13 metal ions to two related RNA duplexes. *Nucleic Acids Res.*, **31**, 2671–2682.
100. Safaee,N., Noronha,A.M., Rodionov,D., Kozlov,G., Wilds,C.J., Sheldrick,G.M. and Gehring,K. (2013) Structure of the parallel duplex of poly(A) RNA: evaluation of a 50 year-old prediction. *Angew. Chem. Int. Ed. Engl.*, **52**, 10370–10373.
101. Wild,K., Weichenrieder,O., Leonard,G.A. and Cusack,S. (1999) The 2 Å structure of helix 6 of the human signal recognition particle RNA. *Structure Fold. Des.*, **7**, 1345–1352.
102. Ohishi,H., Suzuki,K., Ohtsuchi,M., Hakoshima,T. and Rich,A. (2002) The crystal structure of N(1)-[2-(2-amino-ethylamino)-ethyl]-ethane-1,2-diamine (polyamines) binding to the minor groove of d(CGCGCG)(2), hexamer at room temperature. *FEBS Lett.*, **523**, 29–34.
103. Howerton,S.B., Sines,C.C., VanDerveer,D. and Williams,L.D. (2001) Locating monovalent cations in the grooves of B-DNA. *Biochemistry*, **40**, 10023–10031.
104. Batey,R.T. and Doudna,J.A. (2002) Structural and energetics of metal ions essential to SRP signal recognition domain assembly. *Biochemistry*, **41**, 11703–11710.
105. Auffinger,P., Bielecki,L. and Westhof,E. (2004) Symmetric K^+ and Mg^{2+} ion binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. *J. Mol. Biol.*, **335**, 555–571.
106. Dauter,Z. and Adamiak,D.A. (2001) Anomalous signal of phosphorus used for phasing DNA oligomer: importance of data redundancy. *Acta Cryst*., **D57**, 990–995.
107. Knobloch,B., Linert,W. and Sigel,H. (2005) Metal ion-binding properties of (N3)-deprotonated uridine, thymidine, and related pyrimidine nucleosides in aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 7459–7464.
108. Polikanov,Y.S., Steitz,T.A. and Innis,C.A. (2014) A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat. Struct. Mol. Biol.*, **21**, 787–793.
109. Hendlich,M., Bergner,A., Gunther,J. and Klebe,G. (2003) Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.*, **326**, 607–620.
110. Marcia,M. and Pyle,A.M. (2012) Visualizing group II intron catalysis through the stages of splicing. *Cell*, **151**, 497–507.
111. Huang,L., Serganov,A. and Patel,D.J. (2010) Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Mol. Cell.*, **40**, 774–786.
112. Bowman,S.E., Bridwell-Rabb,J. and Drennan,C.L. (2016) Metalloprotein crystallography: more than a structure. *Acc. Chem. Res.*, **49**, 695–702.
113. Ennifar,E., Walter,P. and Dumas,P. (2010) Cation-dependent cleavage of the duplex form of the subtype-B HIV-1 RNA dimerization initiation site. *Nucleic Acids Res.*, **38**, 5807–5816.
114. Jovine,L., Djordjevic,S. and Rhodes,D. (2000) The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg^{2+} in 15-year old crystals. *J. Mol. Biol.*, **301**, 401–414.
115. Robart,A.R., Chan,R.T., Peters,J.K., Rajashankar,K.R. and Toor,N. (2014) Crystal structure of a eukaryotic group II intron lariat. *Nature*, **514**, 193–197.
116. Marcia,M. and Pyle,A.M. (2014) Principles of ion recognition in RNA: insights from the group II intron structures. *RNA*, **20**, 516–527.
117. Wang,J. (2010) Inclusion of weak high-resolution X-ray data for improvement of a group II intron structure. *Acta Cryst.*, **D66**, 988–1000.
118. Kawamura,T., Kobayashi,T. and Watanabe,N. (2015) Analysis of the HindIII-catalyzed reaction by time-resolved crystallography. *Acta Cryst.*, **D71**, 256–265.
119. Muller,P., Kopke,S. and Sheldrick,G.M. (2003) Is the bond-valence method able to identify metal atoms in protein structures? *Acta Cryst.*, **D59**, 32–37.
120. Borbulevych,O., Martin,R.I., Tickle,I.J. and Westerhoff,L.M. (2016) XModeScore: a novel method for accurate protonation/tautomer-state determination using quantum-mechanically driven macromolecular X-ray crystallographic refinement. *Acta Cryst.*, **D72**, 586–598.
121. Borek,D., Minor,W. and Otwinowski,Z. (2003) Measurement errors and their consequences in protein crystallography. *Acta Cryst.*, **D59**, 2031–2038.
122. Diederichs,K. (2016) Crystallographic data and model quality. In: Ennifar,E (ed). *Nucleic Acid Crystallography: Methods and Protocols*. Springer, NY, pp. 147-173.
123. Williams,R.J.P. (2001) Chemical selection of elements by cells. *Coord. Chem. Rev.*, **216**, 583–595.
124. Hill,H.A. and Sadler,P.J. (2016) Bringing inorganic chemistry to life with inspiration from R. J. P. Williams. *J. Biol. Inorg. Chem.*, **21**, 5–12.
125. Hennings,E., Schmidt,H. and Voigt,W. (2013) Crystal structures of hydrates of simple inorganic salts. I. Water-rich magnesium halide hydrates $MgCl_2.8H_2O$, $MgCl_2.12H_2O$, $MgBr_2.6H_2O$, $MgBr_2.9H_2O$, $MgI_2.8H_2O$ and $MgI_2.9H_2O$. *Acta Cryst.*, **C69**, 1292–1300.