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

Filip Leonarski\textsuperscript{1,2}, Luigi D’Ascenzo\textsuperscript{1} and Pascal Auffinger\textsuperscript{1,*}

\textsuperscript{1}Université de Strasbourg, CNRS, Architecture et Réactivité de l’ARN, UPR9002, F-67000 Strasbourg, France and \textsuperscript{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\textsuperscript{2+} in nucleic acids, it is essential to accurately determine their binding modes. Here, we surveyed the PDB to classify Mg\textsuperscript{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\textsuperscript{2+} coordination to N7 has been implied in several ribozyme catalytic mechanisms. We report that Mg\textsuperscript{2+} assigned near imine nitrogens derive mostly from poor interpretations of electron density patterns and are most often misidentified Na\textsuperscript{+}, K\textsuperscript{+}, NH\textsubscript{4}\textsuperscript{+} ions, water molecules or spurious density peaks. Consequently, apart from few documented exceptions, Mg\textsuperscript{2+} ions do not bind to N7 atoms. Without much of a surprise, Mn\textsuperscript{2+}, Zn\textsuperscript{2+} and Cd\textsuperscript{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\textsuperscript{2+} ribosomal binding site involving two purine N7 atoms. Further, we provide a set of guidelines to help in the assignment of Mg\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{2+}—poorly bind to imine atoms, unlike transition metals such as Mn\textsuperscript{2+}, Zn\textsuperscript{2+} or Cd\textsuperscript{2+} (8,20–25). These facts cast doubt on the involvement of nitrogen-bound Mg\textsuperscript{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\textsuperscript{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\textsuperscript{2+} is isoelectronic with water and Na\textsuperscript{+} (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\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{2+} to imine N1/N3/N7 atoms in RNA, DNA and...
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 these 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 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 respective 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\(^{2+}\)...OH\(_2\) coordination distance is slightly below 2.1 Å; (ii) no water oxygens are found within a d(Mg\(^{2+}\)...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\(^{2+}\)...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\(^{2+}\)...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\(^{2+}\)...N distance is slightly longer (≈2.2 Å) than the Mg\(^{2+}\)...O distance in agreement with quantum mechanical calculations and PDB/CSD surveys (12,21,25).

Thus, to a first approximation, our procedures place Mg\(^{2+}\) with d(Mg\(^{2+}\)...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 characterization. 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$^{2+}$...Ow) histogram derived from the CSD (version 5.37, update February 2016; $R$-factors $\leq 5\%$). No disordered, error containing, polymeric or powder structures were included. Standard deviations for the average Mg$^{2+}$...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$^{2+}$ CSD x-ray structure exemplifies the strict octahedral water arrangement around Mg$^{2+}$.

RESULTS AND DISCUSSION

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

As of May 2016, $\approx$56 000 Mg$^{2+}$ ions are assigned in $\approx$1 000 nucleic acid crystallographic structures from the PDB (resolution $\leq$ 3.0 Å). This corresponds to a ratio of roughly one Mg$^{2+}$ per eight nucleotides. Comparatively, under the same resolution threshold, $\approx$25 000 Mg$^{2+}$ are found in $\approx$8 500 proteins. With no resolution limit, the number of Mg$^{2+}$ rises to $\approx$100 000 in nucleic acids and only $\approx$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 $\approx$56 000 Mg$^{2+}$ found in nucleic acid structures, $\approx$1 000 ($\approx$1.8%) display partial occupancies, 59 are associated with occupancies above 1.0, $\approx$4 100 ($\approx$7%) display $B$-factors $\geq 79 \AA^2$ and $\approx$480 ($<1\%$) display $B$-factors $\leq 1.0 \AA^2$. We excluded these ions from statistics in Table 1. In the remaining pool, around 3 900 ($\approx$7%) ions display $d$(Mg$^{2+}$...N) $\leq 3.5 \AA$. If we consider the more stringent $d$(Mg$^{2+}$...N) $\leq 2.4 \AA$ criterion that is more in line with the coordination distance derived from the CSD (Figure 1), only 293 ($\approx$0.5%) Mg$^{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 ($\approx$0.3%) if we consider only non-redundant Mg$^{2+}$ binding sites. These values are to be compared with the $\approx$8 500 ($\approx$15%) Mg$^{2+}$ in direct contact ($\leq 2.4 \AA$) with phosphate oxygens that are considered to be the primary nucleic acid binding sites for Mg$^{2+}$.

$d$(Mg$^{2+}$...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$^{2+}$...N) and the $d$(Mg$^{2+}$...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$^{2+}\ldots$N1/N3/N7 contacts in structures from the PDB (resolution $\leq$ 3.0 Å)

| DNA  | $d$(Mg$^{2+}\ldots$N1) $\leq$ 3.5 Å | $d$(Mg$^{2+}\ldots$N3) $\leq$ 2.4 Å | $d$(Mg$^{2+}\ldots$N7) $\leq$ 2.4 Å |
|------|----------------------------------|----------------------------------|----------------------------------|
| DA   | 116 (243)                        | 5 (5)                            | 131 (198)                        |
| DG   | NR                               | NR                               | 2 (6)                            |
| DC   | NR                               | 1 (1)                            | 3 (3)                            |
| RNA  | 116 (243)                        | 5 (5)                            | 139 (3195)                       |

A 116 (243) 5 (5) 131 (198) 2 (6) 245 (783) 24 (108)
G 84 (121) 5 (6) 1324 (2386) 84 (191)
C 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 $\leq$ 1.0 Å$^2$ and $\geq$ 79 Å$^2$ were not counted.

Figure 2. Mg$^{2+}$ coordination to purine N7 atoms derived from PDB structures. (A) $d$(Mg$^{2+}\ldots$N7) histogram (derived from the PDB; May 2016; resolution $\leq$ 3.0 Å). Ions with $B$-factors $\leq$ 1.0 Å$^2$ and $\geq$ 79 Å$^2$ were excluded. The different ion binding zones are colored according to Figures 1A and 1B. (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$^{n+}\ldots$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. thermophila 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$^{2+}\ldots$N7) histogram become more apparent when compared to the $d$(Mn$^{2+}\ldots$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$^{n+}\ldots$N6/O6) should be in the 3.5–3.9 Å range (Figure 2B). Out of 111 non-redundant ions with $d$(Mg$^{2+}\ldots$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$^{2+}\ldots$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: ZZAAAF01; Supplementary Figure S4), only few partially hydrated and no Mg(H$_2$O)$_5^{2+}$ were located close to (A)N7 in the PDB. In a hammerhead ribozyme (29), the conformation of a nucleobase that involves a $d$(M$^{n+}\ldots$N6/O6) $\leq$ 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\textsuperscript{2+} to adenine N7 atoms. Mg\textsuperscript{2+} to N6 coordination distances are shown in cyan. The red cross is used to mark misidentified Mg\textsuperscript{2+} ions. (A) Ill-placed Mg\textsuperscript{2+} according to electron density patterns and coordination distances. (B) The Mg\textsuperscript{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.

This quasi-absence of reliable Mg(H\textsubscript{2}O)\textsubscript{5}\textsuperscript{2+} to (A)N7 contacts suggests a low Mg\textsuperscript{2+} affinity for this site (25). However, in rare instances, secondary Mg\textsuperscript{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\textsuperscript{2+} binding pockets (see below).

Mg\textsuperscript{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\textsuperscript{2+} bind solely to (G)N7 atoms (no other direct contact to DNA/RNA atoms). Out of those, only 20 Mg\textsuperscript{2+} comprising three Mg(H\textsubscript{2}O)\textsubscript{6}\textsuperscript{2+} satisfy the 3.5 ≤ d(Mg\textsuperscript{2+}...O6) ≤ 3.9 Å criterion (Figure 2). Two of these Mg(H\textsubscript{2}O)\textsubscript{6}\textsuperscript{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\textsuperscript{2+}...Ow) = 2.18 Å restraints were used during refinement (Figure 4A). As a result, Na\textsuperscript{+} can be fitted with a similar level of confidence into these density patterns (see below).

The remaining 21 ions with outlier d(Mg\textsuperscript{2+}...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 twisted ribozyme structure (PDB code: 5DUN; resolution: 2.6 Å), five waters were fitted in a density pattern lacking octahedral symmetry and the Mg\textsuperscript{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\textsuperscript{2+} is placed in front of an N7 atom and displays a non-octahedral coordination (Figure 4C).

Mg\textsuperscript{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\textsuperscript{2+}, Zn\textsuperscript{2+} and Cu\textsuperscript{2+} replace Mg\textsuperscript{2+} (72,73). All these structures belong to the P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1} space group. However, these hexamers crystallize in the P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1} space group in both a Mg\textsuperscript{2+} and a spermine form (74–76) suggesting that Mg\textsuperscript{2+} is not specifically required to stabilize the crystal. Further, under high MgCl\textsubscript{2} and CaCl\textsubscript{2} concentrations (500 mM), this hexamer crystallizes in a P32 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\textsuperscript{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\textsuperscript{2+} binding. Similarly, Co\textsuperscript{2+} and Zn\textsuperscript{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\textsubscript{2}O)\textsubscript{5}\textsuperscript{2+} binds to terminal Z- and B-DNA but not to RNA nucleobases.

Mn\textsuperscript{2+} bound to two N7 atoms of stacked purines: Mg\textsuperscript{2+}, Zn\textsuperscript{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\textsuperscript{2+}...N7) ≤ 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 leading thus to a total of 271 binding sites. Among them, Mg\textsuperscript{2+} was assigned in 243 instances and Sr\textsuperscript{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\textsuperscript{2+} attribution is inappropriate since only less than one out of every six ion satisfies the d(Mg\textsuperscript{2+}...N7) ≤ 2.4 Å criterion (Table 2 and Figure 5B). Therefore, these sites appear to be occupied by waters or non-Mg\textsuperscript{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$^{2+}$...Ow) coordination distances, shown in magenta, were unrealistically 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$^{2+}$...N7) and $d$(Mg$^{2+}$...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$^{2+}$...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 $\leq 1.0$ 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 electron 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$^{2+}$ bound to a double N7 motif. (A) Schematic 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 <$1.0$ $\AA^2$ 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 assignment 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 $\leq$ 3.0 $\AA$; see Supplementary Table S1 and Figure 5)

| Site     | Org.a | Res.b | $d$(Mg$^{2+}$...N7) $\leq$ 2.4 $\AA$ | $d$(Mg$^{2+}$...N7) > 2.4 $\AA$ | Second-shell Mg$^{2+}$ | Sr$^{2+}$ | 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 | —                                | —                               | —                      | —       |
| II.      | 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 | —                                | —                               | —                      | NR      | 1          | 1       |
| **Small ribosomal subunit (SSU)** |
| III.     | 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 $\AA^2$; no $B$-factor criterion was applied to the ion.

Organisms in which these motifs occur; HH, TT, EC and DR stand for H. marismortui, T. thermophilus, E. coli and D. radiodurans, respectively.

The residue numberings are those found in the most representative PDB structure for each organism as noted in Supplementary Table S1, namely 4V9F, 4Yo4, 4YBB, and 5DM6 for HM, TT, EC and DR, respectively.

Both $d$(Mg$^{2+}$...N7) distances have to be below 2.4 $\AA$.

One of both $d$(Mg$^{2+}$...N7) distances has to be in the 2.4–3.8 $\AA$ range.

Both $d$(Mg$^{2+}$...N7) distances have to be in the 3.8–4.6 $\AA$ range.

Sr$^{2+}$ ions are only present in some HH–LSU structures (see Supplementary Table S1). Sr$^{2+}$ ions are considered if $d$(Mg$^{2+}$...N7) $\leq$ 3.0 $\AA$.

No ions with the criteria defined above are found at these sites.

Total number of identified sites.

Site II. in HM and site III. in EC have no double N7 ion binding motif.

Without the $B$-factor $< 79 \AA^2$ criterion, two second shell Mg$^{2+}$ ions are reported in DR structures (PDB codes: 5DM6, 5DM7).
ence of Sr2+. In site I, Mg2+ is systematically replaced by Sr2+ 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 Sr2+ is incompatible with the observed electron density: B-factors are twice those of the nucleobase with some d(Sr2+...N7) distances as short as 2.2 Å, while the CSD estimated Sr2+ coordination distance is ≈2.6 Å and the preferred ligands are oxygens (23,88). Moreover, all alkali earth metals, including Sr2+, are poor N7 binders (25). Rightfully, the authors of these structures did not envisage the binding of Cd2+, 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 Sr2+ and Zn2+, 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 Zn2+ 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 Zn2+ (87,89). The authors of this study suggested that, next to zing-finger motifs, another strong Zn2+ bindervisualizes the B-factors of the attached waters prohibiting univalent Mg2+ in the 2.3–2.5 Å range. For example, see Supplementary Figure S7B and C). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures (Supplementary Figure S8B). Interestingly, these structures.
Figure 6. Mn\(^{2+}\) 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 \(d(Mg^{2+}...N7/OP) \leq 2.4\) 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.

Supplementary Figure S8C where \(d(Mg^{2+}...N7) \approx 2.3\) Å and \(d(Mg^{2+}...OP) \approx 2.4\) 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, \(d(Na^{+}...N7) \approx 2.4–2.6\) 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 \(\approx 0.1\) 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 \(d(K^{+}...N7) \leq 2.0\) Å (PDB codes: 5EW4,
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\(_4\))\(_2\)SO\(_4\). 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 ≤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 2\(F_o\)–\(F_c\) 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 inappropriate 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 tentatively 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 Mg2+ 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 –NH3+ group is at 2.1 Å from a N7 atom (Figure 9C) and the coordination is similar to that shown Figure 4D. Further, the Mg2+ 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 Mg2+ and monovalent cation/water occupancies: are they meaningful?** It has been reported that some DNA major groove hexahydrated Mg2+ 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 NH4+ could overlap with inner-sphere water molecules of a hexacoordinated Mg2+ (104,105). However, it is less likely that water overlays with Mg2+ 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 Mg2+ share the same position (106). This site is similar to the Z-DNA Mg2+ 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 Mg2+ binding to imine N1/N3 nitrogens is unlikely**

Mg2+ binding to N1 and N3 sites is observed in only nine non-redundant instances and appears at best marginal (Table 1). None of these Mg2+ binding occurrences passes visual scrutiny. All Mg2+ have one or more contacts in the 2.5–3.0 Å exclusion zone. Titration experiments suggest that Mg2+ and Ca2+ close to these imine atoms form outer sphere complexes (107). Hence, these inner-sphere contacts represent obvious misattributions.

For instance, an ill-placed Mg2+ ion in front of an (A)N1 atom is found in a T. thermophilus ribosome structure (108) (see the density extension on the adenosine 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 Mg2+ to N3 with a tetrahedral coordination and long Mg2+ to ligand distances suggests the presence of a water molecule rather than an ion (Figure 10B).

**Mg2+ do not bind to the N7 atom of purine containing metabolites**

We checked if Mg2+ 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(Mg2+...N7) in the 3.0–3.5 Å range, and none with d(Mg2+...N7) ≤ 3.0 Å. In the best resolution structure (PDB code: 4H2U; resolution: 2.1 Å), Mg2+ 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 Mg2+ binding potential of N7 atoms.

**Resolutions > 3.0 Å**

If serious identification issues arise at resolutions better than 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 Mg2+, Na+ and even NH4+ ions (110). For instance, we counted in this resolution range 84 nucleic acid structures containing Na+ (including 12 ribosomes) while 43 Mg2+ and 2 K+ containing structures at resolutions lower than 4.0 Å as well as 63 cryo-EM structures containing Mg2+ 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+, Mg2+ 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 \(d(\text{Mg}^{2+} \ldots \text{Ow}) = 2.18\) 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 \(d(\text{Mg}^{2+} \ldots \text{Ow})\) and the \(\approx 2.40\) Å coordination distance expected for \(d(\text{Na}^{+} \ldots \text{Ow})\). 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 \(\leq 2.0\) 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 \(\approx 2.1\) 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 \(d(\text{Na}^{+}/\text{Mn}^{2+} \ldots \text{Ow})\) 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\(^{2+}\) for Mg\(^{2+}\) for N7 is higher (20,25). Further, although rarely described, Mn\(^{2+}/\text{Mg}^{2+}\) 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\(^{2+}\) to Mg\(^{2+}\) 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 Mn\(^{2+}\) 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

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
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 Oph corresponds to phosphate oxygens (OP1/OP2); O1' to O2'/O4'/O3'/O5' ribose oxygens; Ob to nucleobase oxygens and Nb to nucleobase nitrogens.

In MgRNA, 284 Mg$^{2+}$ to N1/N3/N7 (Nb) 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 non-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).

We already addressed the direct binding of an ion to N7 and a nucleobase oxygen atom (Ob,Nb) (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 or three atoms in MgRNA were categorized into three binding types (2Ob,Nb, Oc,2Nb, 2Or,Nb) 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 Oph,Nb 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$_2$O)$_4^{2+}$ 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-2Oph,Nb, 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$^{2+}$...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$_{ph, N}$.

**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 data set (Tables 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$^{2+}$...N7) ≤ 2.4 Å and $d$(Mg$^{2+}$...N7) > $d$(Mg$^{2+}$...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
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 distances 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 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 Leonitis, Prof. Richard Gieger, 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.
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. (2013) 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: Reeedijk,J. and Poepelmeier,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 dinium 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., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

13. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

14. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

15. 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.

16. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of dinium and Mg$^{2+}$ with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. Biochemistry, 52, 5911–5919.

17. Zheng,H., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

18. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

19. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

20. 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.

21. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of dinium and Mg$^{2+}$ with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. Biochemistry, 52, 5911–5919.

22. Zheng,H., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

23. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

24. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

25. 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.

26. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of dinium and Mg$^{2+}$ with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. Biochemistry, 52, 5911–5919.

27. Zheng,H., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

28. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

29. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

30. 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.

31. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of dinium and Mg$^{2+}$ with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. Biochemistry, 52, 5911–5919.

32. Zheng,H., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

33. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

34. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

35. 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.

36. Trachman,R.J. 3rd and Draper,D.E. (2013) Comparison of interactions of dinium and Mg$^{2+}$ with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. Biochemistry, 52, 5911–5919.

37. Zheng,H., Shabalim,I.G., Handing,K.B., Bujnicki,J.M. and Proaschell,B. and Sigel,R.K. (2012) Analysis of Mg$^{2+}$ binding to RNA: From classical methods to a novel iterative calculation procedure. Coord. Chem. Rev., 256, 279–288.

38. Sigel,R.K. and Sigel,H. (2013) Metal-ion interactions with nucleic acids and their constituents. In: Reeedijk,J. and Poepelmeier,K (eds). Comprehensive Inorganic Chemistry II. Elsevier, Oxford, Vol. 3, pp. 623–660.

39. Auffinger,P., D’Ascenzo,L. and Ennifar,E. (2016) Sodium and potassium interactions with nucleic acids. Met. Ions Life Sci., 16, 167–201.

40. 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.
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., Wlodaver, 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., Otwonowicz, 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. Toulw, 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 relevant to proteins. Trends Biochem. Sci., 41, 182–196.

52. Minor, W., Dauter, Z., Hellwell, 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 Schmidt, A. (2015) Covalent modification of guanine bases in double-stranded DNA. Proc. Natl. Acad. Sci. U.S.A., 112, 10950–10957.

57. Chatake, T., Tanaka, I., Umino, H., Arasi, 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.

58. 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, 6238–6248.

59. 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, I. 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, E.C and Baldwin, M.J (eds). *Bioinorganic Chemistry*. American Chemical Society, pp. 81–95.

90. Thom, 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 and Baldwin, M.J (eds). *Metalloprotein crystallography: more than a structure*. *Acc. Chem. Res.*, 49, 695–702.

97. 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.

98. Juneau, K., Podell, E., Harrington, D.J. and Cech, T.R. (2001) Locating monovalent cations in the grooves of B-DNA. *J. Mol. Biol.*, 313, 10023–10031.

99. Ennifar, E., Walter, P. and Dumas, P. (2012) Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA*, 18, 1123–1141.

100. Safaee, N., Noronha, A.M., Rodionov, D., Kozlov, G., Wilds, C.J., 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. *Nat. Struct. Mol. Biol.*, 17, 787–793.