Cat_Wiz: A stereochemistry-guided toolkit for locating, diagnosing and annotating Mg^{2+} ions in RNA structures.

Nawavi Naleem[1], Anja Henning-Knechtel[1], Serdal Kirmizialtin[1,2*], Pascal Auffinger[3]*

[1]Chemistry Program, Science Division, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

[2]Department of Chemistry, New York University, USA

[3]Université de Strasbourg, Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, 2 Allée Konrad Roentgen, 67084 Strasbourg, France

* To whom correspondence should be addressed. Tel: +33 3 88 41 70 00; Fax: +33 3 88 60 22 18;
Email: serdal@nyu.edu; p.auffinger@ibmc-cnrs.unistra.fr

ABSTRACT

Misassigned Mg^{2+} ions are pervasive in RNA structural databases, obscuring mechanistic interpretation, undermining comparative analyses and compromising machine-learning training sets. We present *Cat_Wiz*, a *Coot*-integrated and stereochemistry-guided toolkit that automates the diagnosis, correction and annotation of Mg^{2+} binding sites. *Cat_Wiz* combines three modules: *MG_diagnosis* validates and regularizes existing assignments, *MG_detect* identifies unmodeled ions, *MG_clamp* classifies recurrent clamp motifs. *Cat_Wiz* includes also a precise binding site annotation system. The implemented stereochemical principles that provide a robust experimental foundation for characterizing Mg^{2+} binding sites were derived from an analysis of the 1.55 Å resolution *Escherichia coli* ribosome and from surveys of the Cambridge Structural Database. Applications to ribosomes, hammerhead ribozymes, group I introns, and quaternary RNA assemblies demonstrate that *Cat_Wiz* rapidly recovers missing ions, corrects misassignments, and improves stereochemical fidelity in minutes rather than days. Beyond refinement, *Cat_Wiz* generates curated data that can seed diverse machine-learning and AI models. This transparent, cost-effective framework establishes reproducible standards for RNA ion assignments and will drive progress in the automatic design of RNA 3D architectures through the description of unique Mg^{2+}-dependent backbone folds.

INTRODUCTION

Magnesium ions (Mg^{2+}), interacting as hexahydrated as well as partially dehydrated forms, are central to RNA structure, folding, and catalysis (1-7). Accurate placement and validation of Mg^{2+} in experimental models is therefore critical for mechanistic interpretation, comparative mining and other computational use. While several studies have explored Mg^{2+} coordination to RNA (5,8-14), misassignments remain common in deposited models where ions are often placed in local environments that are incompatible with the strict octahedral geometry of Mg^{2+} ions thereby compromising database surveys and inference (15-18).

In 2022, a cryo-EM structure of a 70S *Escherichia coli* ribosome was solved at 1.55 Å (PDBid: **8b0x**). This high-resolution structure provided an excellent opportunity to codify Mg^{2+} binding principles and establish a robust experimentally based coordination pattern index (14,19). This survey led to a revised **8b0x** model based on the amendment of inaccurate Mg^{2+} assignments and the addition of missing waters to complete octahedral coordination shells. This process significantly increased the number of authenticated Mg^{2+} binding sites while revealing recurring assignment pathologies.

Manual curation at such scales is particularly time-consuming, especially for high-resolution structures like ribosomes that can comprise over 20,000 solvent particles (14). Automated methods based on deep learning classifiers to segmentation-guided solvent placement and molecular dynamics (MD) simulations remain vulnerable to local map heterogeneity and violation of well-established stereochemical boundaries (14,20-22). These drawbacks motivated the development of a transparent, user-friendly and stereochemistry-guided approach for routine Mg^{2+} assignment and validation of X-ray and cryo-EM models. The data gathered on Mg^{2+} binding motifs will certainly advance the development of accurate RNA three-dimensional structure prediction tools (23-27).

Herein, we present *Cat_Wiz* ("Cat" for cation and "Wiz" for wizard), an open suite of Python scripts developed for the *Coot* molecular modeling package (28). *Cat_Wiz* comprises three modules (**Figure 1**): *MG_diagnosis* identifies assignment issues and provides an annotated summary of stereochemically validated Mg^{2+} binding sites while providing tools to regularize Mg^{2+} coordination shells; *MG_detect* scans masked maps around coordinating O/N atoms in RNA (*MG_detect_RNA*) and proteins (*MG_detect_Prot*) to find unmodeled Mg^{2+} ions; *MG_clamp* identifies, classifies, and annotates Mg^{2+} clamps (**Figure 2C**), a recurrent motif in RNA systems (3,5,8,13,14,29-31).

We demonstrate the use of *Cat_Wiz* on four representative cases spanning differences in experimental modalities and resolutions: *(i)* the 1.55 Å resolution cryo-EM **8b0x** ribosome model (13), *(ii)* the historically debated Mg^{2+} binding sites in hammerhead ribozymes that were considered important for catalysis (15,32), *(iii)* the X-ray and near-atomic cryo-EM *Tetrahymena* group I intron structures, the latter having been used as a target in the 2024 CASP competition (26,33) and *(iv)* two large cryo-EM quaternary RNA assemblies with resolutions above 2.6 Å (26). Across all these cases, *Cat_Wiz* successfully resolved incorrect assignments, recovered missing hydrated ions and differentiated Mg^{2+} from K^{+} by enforcing realistic coordination distances, thereby significantly improving interpretability and fidelity of deposited structures. *Cat_Wiz* also assists in the refinement and interpretation of the ionic shell of systems during refinement.

The stereochemical Mg^{2+} priors we defined (**Figure 2**) were derived from small-molecule statistics through an exploration of the Cambridge Structural Database (CSD) and were described elsewhere (11,12,14,34). Several community resources already assist with metal-site identification or diagnosis. Among them, CheckMyMetal (CMM) performs geometry and valence-based validation of metal-binding sites while providing an interactive platform applicable to X-ray and cryo-EM models (35-43). Hence, CMM and *Cat_Wiz,* that provide extensive refinement tools, complement each other.

Cat_Wiz also complements popular refinement programs such as REFMAC5 and Phenix (44,45) by providing stereochemically-aware Mg^{2+} binding site evaluation. The suite of scripts offers optional *d*(Mg^{2+}···O/N) restraints to stabilize Mg^{2+} octahedra in lower-resolution regions. By embedding prior knowledge rather than learning from noisy structures, *Cat_Wiz* provides a fast, transparent refinement solution for solvent and Mg^{2+} ions that can also improve machine learning (ML) and artificial intelligence (AI) workflows. Next, we *(i)* outline the stereochemical nomenclature derived from the **8b0x** analysis, *(ii)* describe separately each module, *(iii)* apply the toolkit to the four systems introduced above and *(iv)* discuss a few *Cat_Wiz* limitations. The current scripts are optimized for modelling Mg^{2+} ions in RNA systems obtained with buffers containing exclusively K^+ as a monovalent ion. Additional strategies to precisely assign K^+ ions based on stereochemical criteria are under development.

METHODS

Cat_Wiz installation in *Coot*

Cat_Wiz is a suite of Python scripts developed for the molecular modeling program *Coot* (28). This suite consists of several independent modules (**Figure 1**). To use the scripts, load in *Coot* a coordinate and a map file derived from X-ray or cryo-EM experiments, then navigate to the *"Calculate"* menu, select *"Run scripts",* and pick the desired script file from your local directories. Upon first use, especially with large structures like ribosomes and to improve visualization of density patterns, *Cat_Wiz* requires a few moments to generate a map smoothed by a factor of 1.25. During this process, the *Coot* console will display "waiting" messages. The scripts were tested with the *Coot* version 0.9.8.96. A version adapted for *Coot* 1.xxx is currently under development. These scripts can analyse X-ray and cryo-EM structures. For X-ray structures, the PDB often provides structure factor files (.sf) instead of *.mtz* files. To use the *Cat_Wiz* scripts, these *.sf* files must be converted into a *.mtz* file using the SF-Tool web server (https://sf-tool.wwpdb.org) adapted from the CCP4 (46). The default cutoffs were determined empirically and therefore comprise a margin of arbitrariness. If interested, the user can modify these defaults in the setup header of the scripts. For further details, please consult the help pages generated by each module. Thereafter, we first describe the nomenclature used for Mg^{2+} binding sites and the added improvements to its preceding version (13), before describing specific features of each module.

Description and additions to the Mg^{2+} binding site nomenclature

The *Cat_Wiz* scripts are based on a fine-tuned nomenclature for Mg^{2+} binding sites (**Figure 3**), derived from the MgRNA study (10-12,14). This nomenclature is based on the strict octahedral coordination of Mg^{2+} with characteristic *d*(Mg^{2+}...O/N) coordination distances of 2.07/2.19 Å. The system uses *cis-/trans-/fac-/mer-/CIS-/TRANS-* prefixes and the following atom type names: "O_{ph}" for OP1/OP2 phosphate anionic oxygens, "O_r" for O2'/O4' ribose oxygens and O3'/O5' for phosphate oxygens, "O_b" for O2/O4/O6 nucleobase oxygens, and "N_b" for non-protonated N1/N3/N7 nucleobase nitrogens.

• *General nomenclature:* As an example, $2O_{ph}.4O_w$ symbolizes a hexacoordinated ion bound to two O_{ph} atoms and four water molecules (**Figure 2C**). The prefixes *cis-* and *trans-* are used to describe the orientation of the two O_{ph} atoms relative to each other (e.g., *cis*-$2O_{ph}.4O_w$ and *trans*-$2O_{ph}.4O_w$). When three non-water atoms are in *cis-* to each other, the corresponding isoforms are named *fac*-$3O_{ph}.3O_w$ (*facial*), and *mer*-$3O_{ph}.3O_w$ (*meridional*) when one O_{ph} pair is in *trans-* (10,14). When four non-water atoms are bound to the ion, as in $4O_{ph}.2O_w$, the orientation of the two bound water molecules is used to define the *cis-/trans-* terminology. To distinguish these from the *cis-/trans-* $2O_{ph}.4O_w$ types, the *CIS-* and *TRANS-* prefixes (in capitals) are used. This nomenclature can be extended to any combination of $O_{ph}/O_b/O_r/N_b$ atoms. For proteins, additional atom types are defined, i.e., O_{bb} for amino acid backbone oxygens in place of previous O_{back} (10,14), O_{coo} for carboxyl oxygens (Asp/Glu), O_{cno} for carbonyl oxygens (Asn/Gln), O_{coh} for hydroxyl oxygens (Thr/Ser/Tyr), and N_{his} for non-protonated histidine nitrogens (ND1/NE2).

• *Water mediated Mg^{2+} clamps:* Rarely, interactions between O/N atoms separated by < 3.4 Å can be stabilized by both, a water-mediated and a direct contact to the Mg^{2+} ion (**Figure 3**). Hence, the

corresponding $O_{ph}.5O_w$ or $O_b.5O_w$ magnesium binding type become "water-mediated" clamps. Previously; we hypothesized that some of these patterns are forerunners to the formation of more standard clamps (14). Because of structural compaction, these water-mediated contacts also appear in complex motifs but, for the sake of nomenclature simplification, are not specifically marked by the *Cat_Wiz* scripts.

• *Ribose O3'/O5' atoms and Mg^{2+} coordination:* Recent group I and II investigations described catalytic sites with alleged Mg^{2+}…O3' contacts (47,48). This is a puzzling observation given that direct coordination of divalent ions to O3' atoms has not been observed in the ensemble of structures we analysed. Therefore, we did not include them in the **Figure 3** summary. To the best of our knowledge, the only exception we came across (see below) was related to terminal residues where the terminal O3' hydroxyl group can coordinate to a Mg^{2+} ion, alone or in combination with the O2' hydroxyl group of the same ribose (49).

• *Monohydrated Mg^{2+} ions:* Monohydrated Mg^{2+} binding patterns involving five non-water molecules are uncommon (**Figure 3**). One such site was modelled in the X-ray structure of *Oceanobacillus iheyensis* group II intron (PDBid: **5j01)** at 3.39 Å resolution (49). The ion coordinates to the 2'-OH and 3'-OH hydroxyl groups of a 3'-terminal uridine near the catalytic site. The organization of this motif may follow different binding rules related to the coordination of two hydroxyl groups from a terminal residue and may play a role in stabilizing leaving groups during the catalytic process.

• *Non-hydrated Mg^{2+} ions:* Coordination of six non-water molecules has not yet been observed in nucleic acids, likely due to the bulkiness of phosphate and nucleobase groups (**Figure 3**). This coordination mode has been observed in a protein structure where the six oxygen atoms belong to protein carbonyl and carboxylate groups, as well as to three phosphate groups from a 2'-deoxyguanosine-5'-triphosphate or DGT ligand (PDBid: **8ef9**) (50). Both these monohydrated and completely dehydrated patterns are included in **Figure 3** but are greyed out since to the best of our knowledge, they have not been observed in representative RNA systems.

• *Coordination to biological and buffer molecules:* Mg^{2+} can also coordinate with biological or buffer ligands. Examples include coordination to the tetracycline antibiotic in a ribosome structure (PDBid: **8cgj**) (51), to ZMP in a riboswitch (PDBid: **4xwf**) (52), or more commonly to di- and triphosphate metabolites (5). Mg^{2+} can also bind to oxygen atoms from buffer molecules like (4S)-2-methyl-2,4-pentanediol (MPD; PDBid: **7mky**) (53) or acetates (PDBid: **1yvp**) (54). Occasionally, non-oxygen atoms, such as fluoride, bind to Mg^{2+} ions forming polyatomic clusters as seen in fluoride riboswitches (PDBid: **4enc**) (55). Direct coordination of Mg^{2+} to a cacodylate anion (CAC: $AsO_2(CH_3)_2^-$) has also been reported (PDBid: **4xvf**).

• *Ligand nomenclature:* Currently, our scripts do not consider Mg^{2+} to ligand binding but do account for most modified nucleotides. For describing sites involving ligands, we suggest using the O_L and N_L atom types. Hence, the tetracycline-bound ion in **8cgj** would be described as a *TRANS*-$2O_b.2O_L.2O_w$ or, more specifically, as a *TRANS*-$2O_b.2O_{TAC}.2O_w$ type, if the PDB code for tetracycline (TAC) is used. This illustrates the versatility of this Mg^{2+} annotation code.

***MG_diagnosis* (script 1): identifying/correcting/validating/annotating Mg^{2+} ions**

The *MG_diagnosis* script generates a *GtkTable* that provides precise stereochemical assessment of ions in a PDB structure helping to distinguish improperly from properly assigned Mg^{2+} ions (**Figure 4**). The table is organized into multiple tabs to guide users through the evaluation and refinement process. The first three tabs can be exported as a CSV file.

• *Tab1 (Mg^{2+} assignment issues) and tab2 (Validated Mg^{2+} ions):* The *"1. Mg^{2+} (check)"* tab identifies ions displaying one or more stereochemical discrepancies. To isolate these issues, we defined three coordination distance ranges: 1.9–2.3 Å, 2.3–2.6 Å and 2.6–3.2 Å, each associated with a coordination number (CN) indicator. An ideal Mg^{2+} ion (11,12,14) is identified by CN = 6 where all six coordination distances fall within the 1.9–2.3 Å range (**Figure 2**). These ions are listed in the *"2. Mg^{2+} (validated)"* tab (**Figure 4**). A CN = 5/1/0 implies a slightly out-of-range coordination distance, while CN = 0/0/5 suggests a non-Mg^{2+} solvent particle which could correspond to K^+ if this ion is present in the buffer (11,14,56). Re-inspection and local refinements are often necessary to validate the assigned ion identity. Colour-coding is used to facilitate data analysis. For instance, red highlights misassigned ions with coordination distances in the 2.6–3.2 Å range typical for K^+ (tab1) while green marks validated octahedral Mg^{2+} ions with CN = 6 (tab2).

• *Binding site quality:* Each Mg^{2+} binding site is evaluated using the tags: *"low", "fair" and 'high-quality"*, which correspond to the three *"rmsd ion"*, *"distorted octahedron (Dist. oct.)"* and *"Clash"* columns. The cells in each of these columns are color-coded in grey (low), pink (fair) or green (high) to facilitate a rapid assessment.

• *"rmsd" indicator:* The "rmsd" value calculated by *Coot* is used as a quality marker. Sites with rmsd values above 10% of the highest peak are considered high-quality (*green*). A value below 5% of the highest peak is arbitrarily tagged as *"low"* (*grey*) while values in the 5–10% range are considered *"fair"* (*pink*). The map-dependent threshold rmsd values are provided in the *"Map info"* tab5. The rmsd values are derived from the original PDB map and not from the visualized smoothed map.

• *Distorted coordination octahedron (Dist. oct.) indicator:* This indicator assesses the distortion of Mg^{2+} coordination octahedron by measuring the distances between first-shell atoms, which average to 2.9 Å for an ideal octahedron (**Figure 2A**). As a proxy for the overall quality of the modelled coordination sphere, we report only the shortest of these distances. When the minimum distance is above 2.6 Å, the site is arbitrarily considered "*high-quality"* (green), in the 2.4–2.6 Å range it is tagged as "*fair-quality"* (pink), and "*low-quality"* when below 2.4 Å (grey; **Figure 5**).

• *Clash indicator:* The Mg^{2+} hexacoordination shell, with its strict coordination distances (< 2.3 Å), generates a 2.3–3.4 Å exclusion zone (**Figure 2B**). If an atom is detected in the 3.2–3.4 Å range, it induces a severe clash (grey). In some intricate patterns, distances in the 3.4–3.6 Å range are categorized as a mild clash (pink). Such clashes require either deleting the assigned Mg^{2+} ion, remodelling the binding site, or changing the ion's identity. Examples of clashes are discussed in the **Results** section.

• *Global quality indicator:* To create a global quality score, the individual indicators mentioned above are aggregated. If any single indicator points to *"low-quality"*, the entire site is labelled accordingly. Otherwise, the site is marked *"fair"* or *"high-quality"*. This global indicator is only used for

the "validated" Mg^{2+} ions in tab2 and is to be found in the "*Quality*" column. It is not pertinent for the non-octahedral coordination patterns found in tab1.

• *Visualization and remodelling of Mg^{2+} binding sites:* A pop-up window, accessible by a right-mouse click on a tab1 or tab2 line, provides visualization and refinement options (**Figure 4D**). *"Zoom to this ion"* allows to quickly center on an atom and visualize any coordination shell issues, such as missing water molecules or *d*(Mg^{2+}…O/N) distances above 2.3 Å. When coordination distances are slightly too large, a simple local refinement in *Coot* can resolve the issue. For water missing molecules, *Coot* offers options to add them into appropriate density patterns. When a Mg^{2+} ion is located at more than 4.3 Å from solute atoms, it is often difficult to assign the density spot to a hydrated $Mg[H_2O]_6^{2+}$ ion or to another solvent particle type in the absence of octahedral density patterns. A subsequent examination of second-shell contacts and eventual clashes can help validate the presence of $Mg[H_2O]_6^{2+}$ ions. A precise evaluation of the experimental buffer composition is also required to evaluate if other buffer component can account for the observed density patterns (14).

• *Spherical refinement of the ion binding site:* The *Coot* real-space refinement tools can be used at any stage of the assignment process. The *"sphere refine"* command, accessible by a right-mouse click on any tab1/tab2 line, performs real-space refinement without restraints on residues comprising atoms within a specified radius (default = 3.5 Å). When density patterns are well-defined, this option typically results in an optimal octahedral geometry with slight coordination distance variations. However, with locally blurred densities, this tool can sometimes lead to more distorted coordination octahedra (**Figure 6**). In such cases, the *"sphere refine with restraint"* option, accessible by a right-mouse click, is recommended as it uses 2.07/2.19 Å restraints for each of the six *d*(Mg^{2+}…O/N) distances.

• *Extra distance restraints:* In some instances, and for less-well resolved sites it might be necessary to add extra distance restraints to stabilize the binding site and avoid too-close Mg^{2+}…O/N contacts. This can be done by navigating to the *"Calculate"* menu and selecting *"Scripting…"* followed by *"Python…"* and use the command *"add_extra_bond_restraints"* that needs thirteen arguments as listed in the **SI**.

• *Mg^{2+}…Mg^{2+} ion pairs:* When two residues are connected to the same Mg^{2+} ion, using the "*sphere refine with restraint"* option may cause a slight deviation from ideal 2.07/2.19 Å coordination distances for at least one of the ions, though they generally remain in the *d*(Mg^{2+}…O/N) < 2.3 Å range. This is often observed for Mg^{2+}…Mg^{2+} and other ion pairs, which require a specific inspection. *MG_diagnosis* identifies and lists all Mg^{2+}…Mg^{2+} ion pairs with a maximum distance of 6.0 Å between the ions. The pairs with *d*(Mg^{2+}… Mg^{2+}) < 4.2 Å (orange) are currently not well documented and require special inspection. The pairs with *d*(Mg^{2+}… Mg^{2+}) in the 4.2–6.0 Å range (cyan) correspond to Mg^{2+} μ-clusters when the same phosphate group binds two Mg^{2+} ions as described in (14).

• *Doted spheres to visualize clashes:* For ambiguous cases, *MG_diagnosis* offers the possibility to visualize spheres of dots with 2.3, 3.2, and 3.6 Å radii. If the six coordinating atoms are within the 2.3 Å sphere, the site is validated while correction is needed for atoms in the 2.3–3.2 Å range. If the application of restraints does not lead to a stereochemically acceptable configuration, it might be necessary to delete, change the ion identity, or remodel the entire binding site. Stereochemical

compatibility with a hydrated Mg^{2+} ion requires that no solute atoms entering the 3.2 Å sphere, and preferably the 3.6 Å sphere since the coordination distance peak identifying second shell atoms starts around 3.6 Å with a maximum at 4.3 Å (**Figure 2B**).

• *Tab3 summary:* The third tab *"3. Mg^{2+} (summary)",* provides an overview of the principal characteristics of validated ions including a list of their binding site types. The first three tabs provide a sufficient overview of the quality of the Mg^{2+} assignments for users who do not wish to further refine the ionic shell of their system.

Tab4 (Help) and tab5 (Map info): The fourth and fifth tabs, *"4. Mg^{2+} (help)"* and *"5. Map_info",* provide general information on the *MG_diagnosis* scripts and the experimental maps, respectively.

MG_detect (script 2): localizing unassigned Mg^{2+} ions

While *MG_diagnosis* annotates all assigned Mg^{2+}, it is not designed to locate unassigned ions. This task is devoted to the second and third *Cat_Wiz* modules which are ideally run in the order shown in **Figure 1**. *MG_detect* is the second of these modules and comprises the *MG_detect_RNA* and *MG_detect_Prot* scripts. These scripts use the *Coot "mask_map_by_molecule"* command with an atom radius of 1.6 Å to create a masked map around the selected molecule by suppressing densities already assigned and isolating unassigned density spots (**Figure 7**).

MG_detect_RNA and *MG_detect_Prot* inspect density spots close to nucleotide atoms (OP1/OP2/O2/O4/O6/O2'/N7) and protein side-chain atoms (OD1/OD2/OE1/OE2/ND1/NE2) in addition to protein backbone atoms (O_{bb}). For modified nucleotides and amino acids often involving different atom names users should manually inspect nearby density. Since *MG_detect_RNA* and *MG_detect_Prot* work similarly, the description below focuses on *MG_detect_RNA.*

The generated mask is searched for the highest peaks from a list obtained using the *Coot* "*map_peaks_around_molecule*" command. In all tabs associated with the OP1/OP2/O2/O4/O6/O2'/N7 atoms, these peaks are sorted by their rmsd values from high to low in the "*Mg site*" column. To avoid overlaps with existing atoms, only peaks within the 1.8–2.4 Å range of any assigned biomolecular or solvent atom are retained as candidates for inner-sphere contacts. The "*Residues < 2.4 Å*" column lists all residues that are near the potential "*Mg site*" peaks. Thus, peaks within 1.8–2.4 Å of an O/N atom are retained for further inspection as potential inner-sphere bound Mg^{2+} ion locations.

The lower-ranked peaks inevitably include spurious positions caused by map irregularities or noise rather than genuine solvent particles. The number of these peaks depends on the treatment (smoothing or sharpening) applied to the map (57,58). To manage this, a pop-up menu has been added that is accessible by a right-mouse click on any line, allowing to select peak ranges and control list length.

The highest-ranked peaks are typically associated with unassigned Mg^{2+} ions that directly contact RNA atoms (**Figures 7A–D**). However, some peaks may correspond to *"unknown"* atoms, such as those previously identified in **8b0x** that appear to be metal-mediated base pair components (14). For instance, in **8b0x**, the top-ranked peak associated with N7 atoms points to a metal-mediated base pair (**Figures 7E–G**). The identity of the metal remains unknown but is clearly not Mg^{2+} or K^{+}. In **9q87**,

we tentatively used the CU1 residue that corresponds to Cu^+ based on the rational that copper grids were employed and that these ions typically exhibit in-plane coordination with CN = 4.

Interestingly, some of these peaks may also highlight poorly modeled backbones or reveal the presence of alternate conformations. We have therefore chosen not to over-filter this list, so that such issues can be identified and corrected. In addition, *MG_detect* can also locate densities corresponding to tetrahedrally or octahedrally coordinated transition metals (Mn^{2+}, Zn^{2+}, etc.) with Mg^{2+}-like coordination distances. As a note of caution, these modules identify also peaks that are associated with fuzzy densities that are particularly difficult to interpret due to possible disorder, presence of unidentified solvent particles or to insufficient knowledge.

Besides inner-shell contacts, *MG_detect_RNA* can also identify hexahydrated Mg^{2+} ions of the $6O_w$ type, provided water molecules were assigned. For that, the automated procedures implemented in *Coot* or *Phenix* are useful and several cutoffs might be used for an optimal screening. In these instances, the *MG_detect* scripts locate densities close to already assigned water molecules with distances below 2.3 Å. The list of peaks is provided in the "*Check O_w*" column.

In summary, *MG_detect_RNA* and *MG_detect_Prot* identify unassigned peaks that may correspond to Mg^{2+} ions or other transition metals, while filtering out peaks associated with water and monovalent K^+ or NH_4^+ ions that exhibit coordination distances in the 2.6–3.2 Å range. It is essential to run *MG_diagnosis* to validate any newly assigned Mg^{2+} positions, correct their stereochemistry and complete their coordination shell.

MG_clamp (script 3): detecting and annotating Mg^{2+} clamps

Several techniques are available to detect Mg^{2+} binding clamps (**Figure 2**). The first involves the *MG_detect_RNA* script, which captures densities near $O_{ph}/O_b/O_r/N_b$ atoms that can subsequently be assigned to Mg^{2+} ions. This technique is highly effective for most high-resolution structures (< 2.0 Å). However, in lower-resolution structures or in less-well resolved regions of large complexes, the experimental density patterns may not be precise enough for *MG_detect_RNA* to be 100% successful. In such cases, a complementary identification technique, implemented in *MG_clamp,* searches for pairs of atoms with *d*(O/N…O/N) ≤ 3.4 Å (**Figure 8**). This helps locating Mg^{2+} binding motifs with weak ion densities but sometimes also identifies poorly modelled RNA backbone. Woth noting, two non-protonated O/N atoms with *d*(O/N…O/N) ≤ 3.4 Å can only be stabilized by coordination to Mg^{2+} or to other transition metals. To date, no examples of clamp motifs stabilized by monovalent ions have been reported (8,30).

• *Tab1 (Fix Asn/Gln sidechains):* Misoriented protein side chains can occasionally generate artefactual contacts that mimic genuine Mg^{2+} clamps. Even after refinement with appropriate options in *Phenix* or *Molprobity* (59,60), some Asn or Gln residues may revert to an inappropriate orientation (**Figures 8A/B/E**). To address this, *MG_clamp* identifies *d*(O/N…O/N) ≤ 3.4 Å distances involving Asn/Gln side chains. When a misorientation is detected, it is recommended to invert the Asn/Gln(N/O) positions with the "*flip sidechain 180°" Coot* tool to improve the model.

• *Tab2 (Potential Mg^{2+} clamps):* The second tab lists all non-water contacts with d(O/N…O/N) ≤ 3.4 Å that are not yet associated with an Mg^{2+} ion. because most hydrogen bond contacts fall into this category, several criteria are used to limit false positive. For example, the identification process

prioritizes atom pairs that include at least one O_{ph} atom since over 99% of the clamps detected in **8b0x** are of this type (**Table 1** exceptions include $2N_b.4O_w$ and $2O_b.4O_w$ binding types). Atom pairs with low *rmsd* values are also flagged, as they often point to incompletely modelled sites. These pairs were kept to allow visual inspection.

• *Tab3 (Mg^{2+} clamps):* The third tab identifies all contacts with d(O/N…O/N) ≤ 3.4 Å that are associated with an assigned Mg^{2+} ion. Clamps involving a hexacoordinated Mg^{2+} ion (CN = 6) are shown in green, while others requiring correction are shown in cyan. The latter can be easily remodelled in *Coot.* The *"n…n+1"* column identifies clamps where Mg^{2+} binds to consecutive residues. For some nucleic acids, water-mediated $O_{ph}.5O_w$ and $O_b.5O_w$ clamps are also included (**Figure 3**). These rare motifs are annoted in the *"Comment"* column. A final visual inspection of all identified binding sites and associated experimental density patterns should always be performed before definitively assigning an ion. Again, *MG_diagnosis* should be used to validate the final modifications made to the model.

• *Tab4 (Help) and tab5 (Map info):* The fourth and fifth tabs, *"4. Mg^{2+} (help)"* and *"5. Map_info"*, provide general information on the various processes associated with the *MG_clamp* scripts and the experimental maps, respectively.

Generating B-factors to end the assignment process

If new assignments or corrections to ion positions are made, *Cat_Wiz* affects the position of atoms and therefore B-factors are no longer accurate. One mean to get updated B-factors is, at the end of the reassignment process, to refine individual the B-factors in *Phenix* or any other program of choice.

RESULTS

Case I: The 1.55 Å *Escherichia coli* ribosome.

Our analysis of the *E. coli* ribosome at 1.55 Å resolution (PDBid: **8b0x**) serves as the main case study to evaluate *Cat_Wiz*. In a previous study, we systematically re-evaluated this structure, which originally contained 361 Mg^{2+} ions (14,19). A visual re-inspection led to the reassignment and validation of 403 Mg^{2+} ions, a process that was critical for generating a comprehensive classification of Mg^{2+} binding sites in RNA and RNA/protein systems (14). Besides identifying new binding sites, this work also corrected misassignments by reclassifying some Mg^{2+} as water or K^{+} and deleting others. Using the *Cat_Wiz* tools, we re-examined the previously **revised-8b0x** structure. This second revision identified an additional seven Mg^{2+} binding sites, while validating the positions of Mg^{2+} ions in a few seconds – a task that otherwise requires days of manual inspection. We stress that we did not identify new binding motifs besides those shown in **Figure 3**. The revised structure, that now contains 415 Mg^{2+} ions, has been deposited to the PDB with **9q87** as identifier.

As detailed in **Table 1**, our analysis revealed significant differences between the ionic shell of the original **8b0x** and the new **9q87** structure. In the original **8b0x,** only 100 Mg^{2+} ions exhibited a hexacoordinated pattern that met our stereochemical criteria. In contrast, the refined **9q87** structure contains 415 ions meeting our criteria.

MG_diagnosis proved invaluable for this revision process. The script flagged 261 ions in the **8b0x** structure that required re-evaluation of either their identity or their coordination shell. Out of these, 42 were identified as potentially misassigned K^{+} ions or other solvent particles, while 220 required coordination-shell adjustments. Among the latter, 85 needed only minor corrections (e.g., adjusting coordination distances), whereas the remaining 135 required more in-depth remodelling. The *Cat_Wiz* ability to rapidly diagnose these issues and facilitate their correction in Coot significantly streamlined the refinement process, even for users with minimal familiarity with *Coot* tools.

An inspection of the **9q87** structure with *MG_diagnosis* (**Table 1**) stressed the quality of our re-assignments: 406 of the 415 assigned Mg^{2+} ions displayed suitable stereochemical features. The *MG_detect* scripts further contributed to the completeness of the model by identifying 12 additional Mg^{2+} ions to complement the 403 Mg^{2+} ions included in the final **amended 8b0x** (14).

As shown in **Table 1**, nine specific Mg^{2+} ions were noted for their anomalous features. One ion, with a coordination shell of seven water molecules, was more consistent with a K^{+} assignment also coordination distances are close to 2.07 Å. Additionally, eight other ions were associated with Mg^{2+}…Mg^{2+} ion pairs and displayed atypical coordination patterns. Approximately 50 *"low-quality"* instances were identified (**Table 1**). While these sites passed inspection and did not exhibit steric clashes, they are nevertheless not suitable for being integrated in statistical surveys or in AI training sets. Documenting them is, however, important for a complete description of the ribosome solvent shell.

One ion located by *MG_detect* with a coordination number of five appeared to be directly linked to an alternate conformation of a guanine nucleobase. As illustrated in **Figure 10**, the high-resolution structure reveals two alternate conformations. Yet, these alternate conformations are particularly challenging to interpret and remain outside the scope of this study.

We also used *MG_clamp* to identify short *d*(O...O/N) distances (≤ 3.4 Å) involving Asn/Gln residues annotated in tab1 (**Figures 8A/B**). *MG_clamp* flagged 33 problematic distances in **8b0x** and that incited the flip of at least 16 Asn/Gln . In one instance, we identified a rare $2O_{ph}.O_{coo}.3O_w$ RNA/protein-Mg^{2+} contact involving the small ribosomal subunit, ribosomal protein S7, and a Mg^{2+} ion at position a:3219. *MG_clamp* takes roughly 10 minutes to run on a recent laptop computer.

A few side notes are necessary for a better understanding of the **8b0x/9q87** structures. First, this *E.* coli ribosome sequence stems from strain B and not from the more conventional MRE 600 (61,62) or K12 strains (63), inducing numbering shifts (14). Second, four residues (A:U:677; A:G:741; A:U:1086; a:U:2316) are anionic (14,64). Unfortunately, the PDB does not provide specific codes for those. Third, a recurrent a:2050:TKW residue (5-hydroxycytidine) has been annotated for the first time in bacterial ribosome structures (65). Fourth, we tentatively included five CU1 residues that account for Cu^+ at positions (A:A:160…A:G:346; a:A:149…a:G177; a:1309...a:1608; a:U1854…a:A:1893) that correspond to presumed metal-mediated base pairs and at position a:G:1366:N7 where Cu^+ links to the terminal w:Ser:2 of r-protein bl28. The observed coordination geometry is consistent with that of Cu^+ (66-69). If present, Cu^+ must be considered as a buffer contaminant. *MG_detect* localized these probable non-Mg^{2+} metals.

Finally, also probably most major Mg^{2+} binding locations involving Mg^{2+} binding clamps were uncovered in this ribosome (14) through the use of *Cat_Wiz*, we can difficultly claim that all of the more difficult to locate $6O_w$ and $5O_w$ sites were recovered. This is especially true in the low density regions. Moreover, it is important to state that diffuse ion binding events are out of *Cat*_Wiz reach. Also, besides well-defined binding sites, numerous other patterns were impossible to interpret in a rational way for now. Lastly, we state that at the present refinement stage, 235 K^+ ions were assigned.

Case II: The hammerhead ribozyme as an example of poorly assigned Mg^{2+} ions

As a second example, *Cat_Wiz* was used to analyse Mg^{2+} binding sites in hammerhead ribozyme structures (**Table 2**). This RNA has long been regarded as the first RNA metalloenzyme. While the early published hammerhead structure (PDBid: **1hmh**) did not mention any Mg^{2+} ions (70), a subsequent X-ray structure (PDBid: **1mme**) reported five Mg^{2+} ions aligning with biochemical findings (71). However, later research contradicted these early conclusions, indicating that Mg^{2+} ions are not directly involved in the catalytic mechanism (72-74) which contrasts with conclusions from several molecular dynamics (MD) studies (75-79). Despite significant effort, X-ray crystallography has not provided sufficient evidence for the direct involvement of Mg^{2+} in the hammerhead ribozyme's catalytic mechanism or in the presence of conserved Mg^{2+} binding sites. Current data suggest that a hydroxyl group from the catalytic site acts as a general acid, likely activated by a nearby Mg^{2+} ion, while a guanine N1 acts as a base (74,80). Consequently, the hammerhead ribozyme's classification has shifted from a metalloenzyme to an acid-base catalysed ribozyme.

A challenge in analyzing these structures is related to the presence of submolar cacodylate ($AsO_2(CH_3)_2^-$) and, in some instances, high concentrations of SO_4^{2-} anions in crystallization buffers (**Tables 2/S2 and Figure 9**). Indeed, some of the originally assigned Mg^{2+} ions are stereochemically inconsistent with Mg^{2+} and correspond probably to SO_4^{2-} or cacodylate anions (81,82) which could have been identified through their strong anomalous X-ray diffraction signals (83,84).

To evaluate structural evidence for Mg^{2+} binding, we analyzed the structures mentioned in **Table 2**. The **1hmh, 1mme,** and **299d** structures lack assigned Mg^{2+} ions although the related papers discussed ion dependent catalytic mechanisms (70,71,85) and *Cat_Wiz* did not locate Mg^{2+} binding positions. **301d** is the first deposited structure with five assigned Mg^{2+} ions. Four of these ions are more than 3.2 Å away from any nucleotide, which is consistent with the presence of solvated ions like $Mg[H_2O]_6^{2+}$, but their densities could also represent SO_4^{2-} or cacodylate ions (**Figure 9**). A fifth ion, located 2.4 Å from A:11:OP2 (**Figure 11A**), is a weak assignment and is barely supported by experimental density. Interestingly, the assigned occupancy for all five ions is 0.5 while in the parent **359d** structure, Tb^{3+} occupancies are below 0.5 indicating unresolved refinement issues.

The **2quw** and **2qus** hammerhead dimer structures, with resolutions between 2.2 and 2.4 Å, contain 17 and 16 Mg^{2+} ions, respectively. In **2quw**, *Cat_Wiz* finds only one Mg^{2+} with a reasonable distance of 2.18 Å and sufficient space to accommodate a hexahydrated shell (**Figure 11B**). However, no clear octahedral densities for the water molecules were observed (in the 2.0 Å resolution **2oeu** Mn^{2+}-containing structure, water densities are present; **Tables S2**). Five of the remaining metal densities could be assigned to hydrated Mg^{2+} ions or anions, but the other ten are poorly placed and sterically incompatible with Mg^{2+} binding. Placing a hydrated ion at these locations would cause severe clashes with neighbouring residues (**Figure 11C**).

The absence of assigned Mg^{2+} hydration shells, in stark contrast to structures containing transition metals or Na^+, makes it challenging to base mechanistic hypotheses on the currently assigned ion positions. Future investigations of these hammerhead ribozymes may reveal more precise and informative details regarding their catalytic mechanisms (73,86,87). For now, existing hammerhead structures essentially highlight the difficulties of understanding the solvation shell of low-resolution systems containing stereochemically unsupported Mg^{2+} assignments which can at best lead to confusing interpretations.

Other structures also contain questionable Mg^{2+} assignments. In **5dh6,** a density with $d(Mg^{2+}\ldots O)$ ≈2.48 Å from OP2:103 is difficult to assign as Mg^{2+} (**Figure 11D**). Similarly, **5dqk** shows a density at a distance of 2.45 Å,(**Figure 11E**) and **5eao** has similar densities, all of which could correspond to other solvent particles since none of them exhibit the expected octahedral shape. Moreover, the high number of poorly assigned Mg^{2+} ions in these structures significantly reduces the confidence one can have in these assignments. In contrast, the 1.55 Å resolution **3zp8** structure, crystallized in a 1.7M sodium-malonate buffer, shows several clearly identifiable octahedrally coordinated Na^+ ions (**Figure 11F**).

The absence of assigned Mg^{2+} hydration shells, in stark contrast to structures with transition metals or Na^+, makes it challenging to formulate mechanistic hypotheses based on the currently assigned ion positions. Future investigations on these hammerhead ribozymes may reveal more precise and informative details about their catalytic mechanisms (73,86,87). For now, existing hammerhead structures primarily highlight the difficulties in interpreting the solvation shell of low-resolution systems containing stereochemically unsupported Mg^{2+} assignments which can lead to confusing conclusions.

Our findings suggest that an early stereochemical assessment of these Mg^{2+} binding sites would have provided clearer guidance for initial investigations into the ribozyme catalytic mechanism. The *Cat_Wiz* scripts proved highly useful for interpreting the solvent structure of these systems. As shown in **Figure 11G**, the **2quw** output does not identify any hexacoordinated ion. Furthermore, *MG_detect_RNA* and *MG_clamp* did not locate unassigned Mg^{2+} binding sites in these hammerhead structures. This underlines also the necessity to explore all structures of a given system at hand to cross-validate assignments that must obey strict stereochemical rules and to have a clear understanding of the buffer conditions that were used to prevent misassignments.

Case III: group I *Tetrahymena* ribozyme structures

Using *Cat_Wiz*, we analysed different *Tetrahymena* group I ribozyme structures to re-inspect their solvation shell. The first is an X-ray structure of the 160-nucleotide P4-P6 ribozyme core (PDBid: **1hr2**) solved at 2.25 Å resolution. This fragment is part of the full-length 387-nucleotide ribozyme (88). Analysis of this and related P4-P6 structures has led to the description of a well-defined Mg^{2+} core with 42 assigned divalent ions (89). A second set of cryo-EM structures of the full-length *Tetrahymena* ribozyme deposited with PDBids **9cbx** and **9cby** at 2.20 Å and 2.30 Å resolutions were also analysed. These structures have slightly different P9.2 stem configurations. Mg^{2+} and water molecules were automatically assigned using a segmentation-guided water and ion modelling protocol called SWIM (22,58,90). The companion structures (**9cbu** and **9cbw)** which contain only consensus solvent molecules common to **9cbx** and **9cby** were also deposited.

• *The P4-P6 mutant solvation shell:* The **1hr2** X-ray structure of this 160-nucleotide group I intron fragment comprises two biological assemblies (88). Chain "A" contains 12 hexacoordinated Mg^{2+} ions that meet our stereochemical criteria and 12 that do not (**Table 3**). Of these, only four have a complete hydration shell and contact directly the RNA. One of these binding site is of the *mer*-$3O_{ph}.3O_w$ type (**Figure 12A**). The coordination spheres of three other ions could be adjusted to satisfy the $d(Mg^{2+}\ldots O/N) < 2.3$ Å criterion. Additionally, eight other hexahydrated Mg^{2+} were modeled to form hydrogen bonds with guanine O6/N7 atoms. The hydration shells of the remaining nine ions were not modelled, making it difficult to determine whether they are hexahydrated Mg^{2+} ions, cacodylate anions, or fragments of HEPES molecules based solely on the available experimental density maps. The *Cat_Wiz* analysis confirmed the Mg^{2+} placements for the hydrated ions, while the *MG_detect* tool did not identify additional inner-sphere binding Mg^{2+} ions, suggesting that all directly bound Mg^{2+} ions in the P4-P6 structure have been accounted for.

• *The full-length **9cbu** and **9cbx** ribozyme solvation shells:* **9cbu** and **9cbx** contain 22 (consensus) and 54 Mg^{2+} ions, respectively. Surprisingly, their hydration shells were not modeled, even though coordinated water molecule densities are apparent for six of these ions in **9cbu**. Because of this, the positions of these ions are not validated by *MG_diagnosis*. To complete the **9cbu** coordination shell, we examined each of the 22 assigned Mg^{2+} ions. The densities for the three water molecules of the *fac/mer*-$3O_{ph}.3O_w$ metals and the four water molecules of the *cis*-$2O_{ph}.4O_w$ metals were visible, supporting these Mg^{2+} assignments. Yet at least four Mg^{2+} ions identified in **1hr2** were missing in **9cbu**, despite clear densities supporting their placement (**Figures 12C/D/E** and **Table 3**). More

concerningly, 16 Mg^{2+} ions were placed at locations where no octahedral densities are visible. Modeling a hydration shell at these sites results in steric clashes with neighbouring residues (**Figure 12F**). These problems highlight the limitations of methods that do not account for stereochemical criteria.

Regarding the non-consensus ions, we found that the ion shown in **Figures 12C/D/E** was missing in **9cbx** like in **9cbu**. Furthermore, none of the 32 additional identified Mg^{2+} ions fit our stereochemical criteria (14). *MG_diagnosis* helped visualize clashes through a sphere of dots, as shown in **Figure 12F**. Our analysis shows that these ions do not display identifiable octahedral densities and are primarily located at sites where a full hydration sphere generates clashes. The *MG_detect* tool was unable to locate any other credible positions for Mg^{2+} ions, despite many additional unassigned densities. Our findings suggest that the *Tetrahymena* group I ribozyme requires only a few precisely positioned inner-sphere binding Mg^{2+} ions to fold into its active conformation. Overall, this analysis demonstrates how the *Cat_Wiz* scripts provide support to inspect the solvation shells of moderate resolution structures and can sort out improbable versus valid assignments. To get a more comprehensive view of the solvation shell of these group I introns, higher-resolution structures are needed, preferably with K^+ instead of Na^+ as monovalent ions.

Case IV: RNA quaternary structures

• *Exploring metal binding sites in RNA scaffold structures:* Cryo-EM techniques have recently oppened a new avenue for investigating RNA quaternary assemblies (26). Metal ions were assigned in two such structures: a *Lactobacillus salivarius* ROOL hexamer (PDBid: **9j6y**), resolved at 3.25 Å and a *Clostridium botulinum* OLE dimer (PDBid: **9lcr**), resolved at 2.6 Å. At these resolutions, no metal hydration shells were modeled. Nevertheless, we used these structures to evaluate how the *Cat_Wiz* tools perform with low-resolution metal-containing models.

• *The **9j6y** Lactobacillus salivarius ROOL hexamer:* In this structure, 56 non-hydrated Mg^{2+} ions were assigned. Of these, 18 form contacts with the RNA in the 2.6–3.2 Å range. These coordination distances are idiosyncratic to K^+ ions. The authors suggested in their supplementary material that a few of the assigned metals form a "*magnesium-clamp*" involving the P13 and P15 secondary structure elements (M6 site). However, *Cat_Wiz* identifies $d(O_{ph}\ldots O_{ph})$ distances of approximately 3.8 Å that are consistent with K^+ binding (**Figure 13A**).

Without additional evidence, the remaining 38 ions at more than 3.2 Å from the RNA are difficult to assign. Despite the absence of octahedral density patterns, they could correspond to hexahydrated Mg^{2+} ions but also to other buffer molecules. For example, the ions placed on the three-fold symmetry axis (M10 site) might correspond to $Mg[H_2O]_6^{2+}$ ions given the cavity size. However, the lack of an octahedral density casts doubts on this assignment (**Figure 12B**). Higher-resolution structures would facilitate such assignments.

• *The **9lcr** Clostridium botulinum OLE dimer:* For this 2.6 Å resolution dimer, no hydration shells were modelled around the 28 assigned metals (14 per monomer). Among the 14 monomer "A" ions, a valid *mer*-$3O_{ph}.3O_w$ coordination (M2 site) was identified. The metal coordination distances are below 2.3 Å and the *cis-d*$(O_{ph}\ldots O_{ph})$ distances are 2.7/3.1 Å. Four other faintly octahedral metal sites may correspond to hexahydrated Mg^{2+} ions, while the remaining nine occupy K^+ sites. These sites,

especially M6, M7, M9, M11, M12 and M14, are located near the major groove of G•U pairs which are known to be preferential K^+ binding sites (3,11,91).

Given the limited resolution of this structure, we were only able to remodel a few of the assigned metals to illustrate how missing waters can be added to ions with octahedral densities. The ion at the *mer*-$3O_{ph}.3O_w$ binding site (M2) shows densities for three missing water molecules (**Figure 13C**). Using *Coot*, we placed waters near their presumed location and ran again *MG_diagnosis*. We then used the *"Sphere refine with restraints"* option to optimise their positions. Similarly, we added water to a presumed hexahydrated Mg^{2+} ion at the M1 site and a M5 ion at a *trans*-$2O_{ph}.4O_w$ site. This procedure resulted in a local geometry optimisation and a satisfactory model of the ion hydration sphere, validating the presence of octahedrally coordinated Mg^{2+} ions. Yet, a more in-depth remodeling of these structures, including an evaluation of several orphan densities in the experimental maps is beyond the scope of this paper. However, in the publication describing these quaternary structures (26), the assigned ions were named "Metals" or (M), except in a supplementary material table. Yet, the PDB identifies these ions as "Mg". This discrepancy highlights how automatic database surveys can be compromised by annotations made without proper consideration of stereochemical criteria.

These findings suggest that Mg^{2+} and K^+ ions in low-resolution structures can be differentiated or invalidated based on density patterns and prior stereochemical knowledge. Such prior-knowledge based approaches help prevent the propagation of incorrect interpretations. In these examples, *MG_detect* and *MG_clamp* did not identify any additional Mg^{2+} binding sites. However, besides providing a clear survey of the assigned ions, *MG_diagnosis* was of help in modeling Mg^{2+} hydration shells.

DISCUSSION

Accurate assignments of Mg^{2+} binding sites remain one of the most significant challenges in RNA structural biology. Misplaced ions are common in the PDB even in high-resolution models where stereochemical constraints are violated or hydration shells are incomplete (11,12,14,92). Such errors lead to inaccuracies in the mechanistic interpretation of structural models, deceiving functional hypotheses, and introduce noise into databases used for comparative analyses or machine-learning workflows (10,20,21,40-42,93).

Cat_Wiz addresses these problems by combining stereochemical binding principles into a pipeline. By enforcing canonical distances and octahedral coordination geometries derived from the CSD and from a 1.55 Å resolution structure of the *E. coli* ribosome (14), *Cat_Wiz* identifies inconsistent assignments, suggests alternative interpretations and highlights regions where density or backbone modeling may be inadequate. Herein, we showed applications from ribosomes, hammerhead ribozymes, and group I introns, as well as RNA quaternary structures to illustrate how the toolkit works and recovers incomplete hydration shells of ions, corrects misassigned cations, and annotates recurrent ion motifs such as Mg^{2+} clamps. Notably, these corrections can be completed in minutes rather than taking days of manual examination. *Cat_Wiz* can be used with low- to high-resolution X-ray and cryo-EM structures.

While *Cat_Wiz* aims at providing insights into RNA structural biology it also shares limitations with other approaches. We note that at resolutions poorer than ≈2.5 Å, ion coordination becomes difficult to discern, and errors in RNA backbone modeling propagate leading to ambiguous solvation shells. To mitigate this, *Cat_Wiz* provides restrained refinement options that enforce realistic Mg^{2+} coordination distances while maintaining flexibility. These options impose stereochemical distance restraints that are commonly applied to nucleotides and amino acids, with the aim to ensure chemically plausible structural models (94-96).

Buffer composition presents a second challenge. Non-physiological intracellular cations (97-99) such as Ca^{2+}, NH_4^+, Na^+, and anions such as cacodylates and sulphates (81,82) are often present in experimental buffers, complicating experimental map interpretations. *Cat_Wiz* analyses and highlights cases where these ions have been misinterpreted as in Mg^{2+} in hammerhead ribozymes. These findings reinforce the importance of using simplified and physiologically relevant buffers containing primarily Mg^{2+} and K^+, to ensure better density map interpretability and physiologically relevant assignments. We note that the use of simplified buffers does not hinder the obtention of high-resolution structures and allows *in vitro* autonomous ribosome biogenesis (19,100). In the future, systematic inclusion of physiologically relevant polyamines will represent an important addition to experimental buffers and advance our understanding of biological solvation shells (101-104). Although chloride is the most common intracellular anion (3,81,82), it is often absent from experimental buffers and replaced by other anionic components (e.g. acetate; malonate, HEPES or TRIS molecules). It is important to keep in mind that these substitutes may perturb to an unknown extent the RNA solvation shell.

A third concern lies in the deposition and annotation of structural data (15,18,105,106). Misassigned ions persist in the Protein Data Bank without warning annotations, despite repeated identification of errors (15,32). In this regard, the CAVEAT remark should be used more consistently to flag known issues in deposited PDB structures (18). Such inaccuracies not only mislead newcomers but have also the potential to twist automated surveys and machine-learning training sets. For instance, this issue affects hammerhead ribozymes and recent group I intron ribozyme structures potentially jeopardizing collective efforts to understand and model the ionic structure and catalytic features of these systems (22). Wider use of annotation tools such as *Cat_Wiz*, together with PDB-level flagging, would improve the quality of the models accessible to the community. As a side note, the PDB has no specific code for identifying anionic U and G nucleotides recently annotated in bacterial ribosomes and observed in **8b0x**, while a specific AP7 code is used for protonated adenines (14,64). A fix by the PDB would drive progress in understanding these systems.

Cat_Wiz also underscores the limits of automation. Alternate nucleotide conformations, ambiguous solvent densities, and unusual motifs, such as metal-mediated base pairs or diffuse ion populations, require expert inspection before being published. *Cat_Wiz* is designed to highlight these instances, yet it does not aim to automatically resolve them since visual inspection is key to understand undocumented features although even visual inspection assisted by *Cat_Wiz* of less-well resolved regions, leaves us sometimes with more interrogations than answers.

Finally, while machine-learning approaches are becoming popular as ion identification techniques (20,107,108), current methods struggle to correct poor assignments and stereochemical errors present in existing models. Similarly, MD simulations that lack polarization and charge transfer effects should be interpreted with caution when modelling Mg^{2+} ions (22,109,110). Awaiting to extract more information from high-resolution structures (111-115), *Cat_Wiz* provides an essential bridge between experimental data and modelling. By incorporating stereochemical principles and producing high-quality ion assignments, it can supply more accurate training sets to improve the accuracy and reliability of future AI tools and forcefields.

In summary, *Cat_Wiz* offers a rapid, and stereochemistry-guided framework for diagnosing and correcting Mg^{2+} assignments. By improving the structural fidelity of the solvation shell of biomolecules and generating curated datasets, the *Cat_Wiz* toolkit benefits both experimentalists seeking for accurate solvent models and computational scientists developing predictive algorithms to study ion-RNA interactions. Its integration into *Coot* makes it accessible to a wide community and positions it as a practical, fast, and transparent tool that complements both manual refinement and emerging autonomous structural biology approaches.

Disclaimer

We did our best in designing *Cat_Wiz* and in assigning Mg^{2+} ions especially in **8b0x**. The quality of these assignments is strongly dependent on the resolution and the refinement that has been done as well as on the accumulated prior-knowledge at disposal that helps disentangle complex and unforeseen solvation patters. This is an ongoing process and we do not claim that we assigned all Mg^{2+} ions, that some solvation patterns did not misled us, and that all are accurate. However, we did

our best to assign these ions by using the best available knowledge and structures, one of the most helpful being actually **8b0x** and we are confident that we located probably all important Mg^{2+} binding sites in ribosomal structures and all important binding site types in RNA systems.

AVAILABILITY

During the submission process, scripts are deposited to GitHub as a password protected '.zip' folder at the https://github.com/pauffinger/Cat_Wiz address. Password is provided upon request. After manuscript acceptance, these scripts will become publicly available.

ACCESSION NUMBERS

Atomic coordinates and structure factors for the reported crystal structures have been deposited to the PDB under accession number **9q87**.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENT

The authors thank Prof. Eric Westhof and Dr. Eric Ennifar for helpful discussions, Dr. Ahmad Jomaa and Dr. Emir Maldosevic for providing info on cryo-EM structures, Dr. Paul Elmsley and Dr. Lucrezia Catapano for help and advice on *Coot* functionalities, as well as Dr Aliza Fedorenko, Dr Anat Bashan and Dr Jiro Kondo for discussions on the possible presence of Cu^+ as contaminant in the experimental process.

FUNDING

Centre National de la Recherche Scientifique; NYUAD [AD181 to S.K. and N.N]; REF [RE317 to S.K. and N.N].

CONFLICT OF INTEREST

None

References:

1. Draper, D.E. (2004) A guide to ions and RNA structure. *RNA*, **10**, 335-43.
2. Draper, D.E., Grilley, D. and Soto, A.M. (2005) Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.*, **34**, 221-43.
3. 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**, 1366-79.
4. Nierhaus, K.H. (2014) Mg^{2+}, K^{+}, and the ribosome. *J. Bacteriol.*, **196**, 3817-9.
5. Bowman, J.C., A.S., P. and Williams, L.D. (2018) Role of cations in RNA folding and function. In Kolb, V. M. (ed.), *Handbook of Astrobiology*. CRC Press, Boca Raton, pp. 421-35.
6. Akanuma, G. (2021) Diverse relationships between metal ions and the ribosome. *Biosci. Biotechnol. Biochem.*, **85**, 1582-93.
7. Helmann, J.D. (2025) Metals in motion: understanding labile metal pools in bacteria. *Biochemistry*, **64**, 329-45.
8. Bowman, J.C., Lenz, T.K., Hud, N.V. and Williams, L.D. (2012) Cations in charge: magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.*, **22**, 262-72.
9. Marcia, M. and Pyle, A.M. (2014) Principles of ion recognition in RNA: insights from the group II intron structures. *RNA*, **20**, 516-27.
10. 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**, 3789-801.
11. Leonarski, F., D'Ascenzo, L. and Auffinger, P. (2019) Nucleobase carbonyl groups are poor Mg^{2+} inner-sphere binders but excellent monovalent ion binders - A critical PDB survey. *RNA*, **25**, 173-92.
12. Leonarski, F., D'Ascenzo, L. and Auffinger, P. (2017) Mg^{2+} ions: do they bind to nucleobase nitrogens? *Nucleic Acids Res.*, **45**, 987-1004.
13. Cowan, J.A. (2024) Understanding the thermodynamics of magnesium binding to RNA structural motifs. *Life (Basel)*, **14**.
14. Leonarski, F., Henning-Knechtel, A., Kirmizialtin, S., Ennifar, E. and Auffinger, P. (2025) Principles of ion binding to RNA inferred from the analysis of a 1.55 A resolution bacterial ribosome structure - Part I: Mg2. *Nucleic Acids Res*, **53**.
15. Minor, W., Dauter, Z., Helliwell, J.R., Jaskolski, M. and Wlodawer, A. (2016) Safeguarding structural data repositories against bad apples. *Structure*, **24**, 216-20.
16. Wlodawer, A., Dauter, Z., Porebski, P.J., Minor, W., Stanfield, R., Jaskolski, M., Pozharski, E., Weichenberger, C.X. and Rupp, B. (2018) Detect, correct, retract: How to manage incorrect structural models. *FEBS J.*, **285**, 444-66.
17. Wlodawer, A., Dauter, Z., Lubkowski, J., Loch, J.I., Brzezinski, D., Gilski, M. and Jaskolski, M. (2024) Towards a dependable data set of structures for L-asparaginase research. *Acta Crystallogr. D Struct. Biol.*, **80**, 506-27.
18. Wlodawer, A., Dauter, Z., Rubach, P., Minor, W., Loch, J.I., Brzezinski, D., Gilski, M. and Jaskolski, M. (2024) Waterless structures in the Protein Data Bank. *IUCrJ*, **11**, 966-76.
19. Fromm, S.A., O'Connor, K.M., Purdy, M., Bhatt, P.R., Loughran, G., Atkins, J.F., Jomaa, A. and Mattei, S. (2023) The translating bacterial ribosome at 1.55 A resolution generated by cryo-EM imaging services. *Nat. Commun.*, **14**, 1095.
20. Zhou, Y. and Chen, S.J. (2022) Graph deep learning locates magnesium ions in RNA. *QRB Discov.*, **3**.
21. Wang, K., Yin, Z., Sang, C., Xia, W., Wang, Y., Sun, T. and Xu, X. (2024) Geometric deep learning for the prediction of magnesium-binding sites in RNA structures. *Int. J. Biol. Macromol.*, **262**, 130150.
22. Kretsch, R.C., Li, S., Pintilie, G., Palo, M.Z., Case, D.A., Das, R., Zhang, K. and Chiu, W. (2025) Complex water networks visualized by cryogenic electron microscopy of RNA. *Nature*, **642**, 250-9.
23. Mukherjee, S., Moafinejad, S.N., Badepally, N.G., Merdas, K. and Bujnicki, J.M. (2024) Advances in the field of RNA 3D structure prediction and modeling, with purely theoretical approaches, and with the use of experimental data. *Structure*, **32**, 1860-76.
24. Shen, T., Hu, Z., Sun, S., Liu, D., Wong, F., Wang, J., Chen, J., Wang, Y., Hong, L., Xiao, J. *et al.* (2024) Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nat. Methods*, **21**, 2287-98.

25. Bernard, C., Postic, G., Ghannay, S. and Tahi, F. (2025) Has AlphaFold3 achieved success for RNA? *Acta Crystallogr. D Struct. Biol.*, **81**, 49-62.
26. Bu, F., Adam, Y., Adamiak, R.W., Antczak, M., de Aquino, B.R.H., Badepally, N.G., Batey, R.T., Baulin, E.F., Boinski, P., Boniecki, M.J. *et al.* (2025) RNA-Puzzles Round V: blind predictions of 23 RNA structures. *Nat. Methods*, **22**, 399-411.
27. Xia, Y., Liang, Z., Du, X., Cao, D., Li, J., Sun, L., Huo, Y.X. and Guo, S. (2025) Design of function-regulating RNA via deep learning and AlphaFold 3. *Brief. Bioinform.*, **26**.
28. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Struct. Biol.*, **66**, 486-501.
29. Hsiao, C. and Williams, L.D. (2009) A recurrent magnesium-binding motif provides a framework for the ribosomal peptidyl transferase center. *Nucleic Acids Res.*, **37**, 3134-42.
30. Petrov, A.S., Bowman, J.C., Harvey, S.C. and Williams, L.D. (2011) Bidentate RNA-magnesium clamps: on the origin of the special role of magnesium in RNA folding. *RNA*, **17**, 291-7.
31. Hsiao, C., Tannenbaum, E., VanDeusen, H., Hershkovitz, E., Perng, G., Tannenbaum, A.R. and Williams, L.D. (2009) Complexes of nucleic acids with group I and II cations. In Hud, N. V. (ed.), *Nucleic Acid-Metal Ion Interactions*. Royal Society of Chemistry, Cambridge, pp. 1-38.
32. Rupp, B., Wlodawer, A., Minor, W., Helliwell, J.R. and Jaskolski, M. (2016) Correcting the record of structural publications requires joint effort of the community and journal editors. *FEBS J.*, **283**, 4452-7.
33. Su, Z., Zhang, K., Kappel, K., Li, S., Palo, M.Z., Pintilie, G.D., Rangan, R., Luo, B., Wei, Y., Das, R. *et al.* (2021) Cryo-EM structures of full-length *Tetrahymena* ribozyme at 3.1 Å resolution. *Nature*, **596**, 603-7.
34. Taylor, R. and Wood, P.A. (2019) A million crystal structures: The whole is greater than the sum of its parts. *Chem. Rev.*, **119**, 9427-77.
35. Zheng, H., Cooper, D.R., Porebski, P.J., Shabalin, I.G., Handing, K.B. and Minor, W. (2017) CheckMyMetal: a macromolecular metal-binding validation tool. *Acta Crystallogr. D Struct. Biol.*, **73**, 223-33.
36. Handing, K.B., Niedzialkowska, E., Shabalin, I.G., Kuhn, M.L., Zheng, H. and Minor, W. (2018) Characterizing metal-binding sites in proteins with X-ray crystallography. *Nat. Protoc.*, **13**, 1062-90.
37. Porebski, P.J., Sroka, P., Zheng, H.P., Cooper, D.R. and Minor, W. (2018) Molstack-Interactive visualization tool for presentation, interpretation, and validation of macromolecules and electron density maps. *Protein Science*, **27**, 86-94.
38. Porebski, P.J., Bokota, G., Venkataramany, B.S. and Minor, W. (2020) Molstack: A platform for interactive presentations of electron density and cryo-EM maps and their interpretations. *Protein Sci*, **29**, 120-7.
39. Brzezinski, D., Porebski, P.J., Kowiel, M., Macnar, J.M. and Minor, W. (2021) Recognizing and validating ligands with CheckMyBlob. *Nucleic Acids Res*, **49**, W86-W92.
40. Gucwa, M., Lenkiewicz, J., Zheng, H., Cymborowski, M., Cooper, D.R., Murzyn, K. and Minor, W. (2023) CMM-An enhanced platform for interactive validation of metal binding sites. *Protein Sci.*, **32**, e4525.
41. Majorek, K.A., Gucwa, M., Murzyn, K. and Minor, W. (2024) Metal ions in biomedically relevant macromolecular structures. *Front. Chem.*, **12**, 1426211.
42. Gucwa, M., Bijak, V., Zheng, H., Murzyn, K. and Minor, W. (2024) CheckMyMetal (CMM): validating metal-binding sites in X-ray and cryo-EM data. *IUCrJ*, **11**, 871-7.
43. Zhang, H., Zhong, J., Gucwa, M., Zhang, Y., Ma, H., Deng, L., Mao, L., Minor, W., Wang, N. and Zheng, H. (2025) PinMyMetal: a hybrid learning system to accurately model transition metal binding sites in macromolecules. *Nat. Commun.*, **16**, 3043.
44. 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 Crystallogr. D Struct. Biol.*, **67**, 355-67.
45. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J. *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.*, **D75**, 861-77.
46. Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C.B., Basle, A., Beilsten-Edmands, J., Borges, R.J., Brown, D.G., Burgos-Marmol, J.J., Berrisford, J.M. *et al.* (2023) The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.*, **79**, 449-61.

47. Luo, B., Zhang, C., Ling, X., Mukherjee, S., Jia, G., Xie, J., Jia, X., Liu, L., Baulin, E.F., Luo, Y. *et al.* (2023) Cryo-EM reveals dynamics of *Tetrahymena* group I intron self-splicing. *Nat. Catal.*, **6**, 298-309.
48. Haack, D.B., Rudolfs, B., Jin, S., Khitun, A., Weeks, K.M. and Toor, N. (2025) Scaffold-enabled high-resolution cryo-EM structure determination of RNA. *Nat. Commun.*, **16**, 880.
49. Costa, M., Walbott, H., Monachello, D., Westhof, E. and Michel, F. (2016) Crystal structures of a group II intron lariat primed for reverse splicing. *Science*, **354**.
50. Li, C., Zhu, H., Jin, S., Maksoud, L.M., Jain, N., Sun, J. and Gao, Y. (2023) Structural basis of DNA polymerase theta mediated DNA end joining. *Nucleic Acids Res.*, **51**, 463-74.
51. Paternoga, H., Crowe-McAuliffe, C., Bock, L.V., Koller, T.O., Morici, M., Beckert, B., Myasnikov, A.G., Grubmuller, H., Novacek, J. and Wilson, D.N. (2023) Structural conservation of antibiotic interaction with ribosomes. *Nat. Struct. Mol. Biol.*, **30**, 1380-92.
52. Trausch, J.J., Marcano-Velazquez, J.G., Matyjasik, M.M. and Batey, R.T. (2015) Metal ion-mediated nucleobase recognition by the ZTP riboswitch. *Chem. Biol.*, **22**, 829-37.
53. Grotz, K.K. and Schwierz, N. (2022) Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties in SPC/E, TIP3P-fb, TIP4P/2005, TIP4P-Ew, and TIP4P-D. *J Chem Theory Comput*, **18**, 526-37.
54. Stein, A.J., Fuchs, G., Fu, C., Wolin, S.L. and Reinisch, K.M. (2005) Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity. *Cell*, **121**, 529-39.
55. Ren, A., Rajashankar, K.R. and Patel, D.J. (2012) Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch. *Nature*, **486**, 85-9.
56. Rozov, A., Khusainov, I., El Omari, K., Duman, R., Mykhaylyk, V., Yusupov, M., Westhof, E., Wagner, A. and Yusupova, G. (2019) Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction. *Nat. Commun.*, **10**, 2519.
57. Wlodawer, A., Li, M. and Dauter, Z. (2017) High-resolution cryo-EM maps and models: A crystallographer's perspective. *Structure*, **25**, 1589-97 e1.
58. Pintilie, G. and Chiu, W. (2021) Validation, analysis and annotation of cryo-EM structures. *Acta Crystallogr. D Struct. Biol.*, **77**, 1142-52.
59. Word, J.M., Lovell, S.C., Richardson, J.S. and Richardson, D.C. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.*, **285**, 1735-47.
60. Richardson, J.S., Williams, C.J., Hintze, B.J., Chen, V.B., Prisant, M.G., Videau, L.L. and Richardson, D.C. (2018) Model validation: local diagnosis, correction and when to quit. *Acta Crystallogr. D Struct. Biol.*, **74**, 132-42.
61. Nissley, A.J., Kamal, T.S. and Cate, J.H.D. (2023) Interactions between terminal ribosomal RNA helices stabilize the *E. coli* large ribosomal subunit. *RNA*, **29**, 1500-8.
62. Parajuli, N.P., Emmerich, A., Mandava, C.S., Pavlov, M.Y. and Sanyal, S. (2023) Antibiotic thermorubin tethers ribosomal subunits and impedes A-site interactions to perturb protein synthesis in bacteria. *Nat. Commun.*, **14**, 918.
63. Noeske, J., Wasserman, M.R., Terry, D.S., Altman, R.B., Blanchard, S.C. and Cate, J.H. (2015) High-resolution structure of the *Escherichia coli* ribosome. *Nat. Struct. Mol. Biol.*, **22**, 336-41.
64. Westhof, E., Watson, Z.L., Zirbel, C.L. and Cate, J.H.D. (2023) Anionic G*U pairs in bacterial ribosomal rRNAs. *RNA*, **29**, 1069-76.
65. Fasnacht, M., Gallo, S., Sharma, P., Himmelstoss, M., Limbach, P.A., Willi, J. and Polacek, N. (2022) Dynamic 23S rRNA modification ho5C2501 benefits *Escherichia coli* under oxidative stress. *Nucleic Acids Res.*, **50**, 473-89.
66. 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**, 20175-84.
67. Solomon, E.I., Heppner, D.E., Johnston, E.M., Ginsbach, J.W., Cirera, J., Qayyum, M., Kieber-Emmons, M.T., Kjaergaard, C.H., Hadt, R.G. and Tian, L. (2014) Copper active sites in biology. *Chem. Rev.*, **114**, 3659-853.
68. Iwamura, M., Takeuchi, S. and Tahara, T. (2015) Ultrafast excited-state dynamics of copper(I) complexes. *Acc. Chem. Res.*, **48**, 782-91.
69. Beaudelot, J., Oger, S., Perusko, S., Phan, T.A., Teunens, T., Moucheron, C. and Evano, G. (2022) Photoactive copper complexes: properties and applications. *Chem. Rev.*, **122**, 16365-609.

70. Pley, H.M., Flaherty, K.M. and McKay, D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68-74.
71. 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.
72. Leclerc, F. (2010) Hammerhead ribozymes: true metal or nucleobase catalysis? Where is the catalytic power from? *Molecules*, **15**, 5389-407.
73. Herschlag, D. (2015) Learning from ribozymes. *RNA*, **21**, 527-8.
74. Deng, J., Shi, Y., Peng, X., He, Y., Chen, X., Li, M., Lin, X., Liao, W., Huang, Y., Jiang, T. *et al.* (2023) Ribocentre: a database of ribozymes. *Nucleic Acids Res.*, **51**, D262-D8.
75. Hermann, T., Auffinger, P. and Westhof, E. (1998) Molecular dynamics investigations of the hammerhead ribozyme RNA. *Eur. J. Biophys.*, **27**, 153-65.
76. Hermann, T., Auffinger, P., Scott, W.G. and Westhof, E. (1997) Evidence for a hydroxide ion bridging two magnesium ions at the active site of the hammerhead ribozyme. *Nucleic Acids Res.*, **25**, 3421-7.
77. Lee, T.S., Silva Lopez, C., Giambasu, G.M., Martick, M., Scott, W.G. and York, D.M. (2008) Role of Mg^{2+} in hammerhead ribozyme catalysis from molecular simulation. *J. Am. Chem. Soc.*, **130**, 3053-64.
78. Chen, H., Giese, T.J., Golden, B.L. and York, D.M. (2017) Divalent metal ion activation of a guanine general base in the Hammerhead ribozyme: Insights from molecular simulations. *Biochemistry*, **56**, 2985-94.
79. Ekesan, S., McCarthy, E., Case, D.A. and York, D.M. (2022) RNA electrostatics: How ribozymes engineer active sites to enable catalysis. *J. Phys. Chem. B*, **126**, 5982-90.
80. Zhan, X., Wilson, T.J., Li, Z., Zhang, J., Yang, Y., Lilley, D.M.J. and Liu, Y. (2024) The structure and catalytic mechanism of a pseudoknot-containing hammerhead ribozyme. *Nat. Commun.*, **15**, 6628.
81. Auffinger, P., Bielecki, L. and Westhof, E. (2004) Anion binding to nucleic acids. *Structure*, **12**, 379-88.
82. D'Ascenzo, L. and Auffinger, P. (2016) Anions in nucleic acid crystallography. *Methods Mol. Biol.*, **1320**, 337-51.
83. Doukov, T., Herschlag, D. and Yabukarski, F. (2023) Obtaining anomalous and ensemble information from protein crystals from 220 K up to physiological temperatures. *Acta Crystallogr. D Struct. Biol.*, **79**, 212-23.
84. El Omari, K., Forsyth, I., Duman, R., Orr, C.M., Mykhaylyk, V., Mancini, E.J. and Wagner, A. (2024) Utilizing anomalous signals for element identification in macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.*, **80**, 713-21.
85. Scott, W.G., Murray, J.B., Arnold, J.R.P., Stoddard, B.L. and Klug, A. (1996) Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science*, **274**, 2065-9.
86. Hammann, C. and Westhof, E. (2009) The unforeseeable hammerhead ribozyme. *F1000 Biol. Rep.*, **1**, 6.
87. Hammann, C., Luptak, A., Perreault, J. and de la Pena, M. (2012) The ubiquitous hammerhead ribozyme. *RNA*, **18**, 871-85.
88. 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**, 221-31.
89. Cate, J.H., Hanna, R.L. and Doudna, J.A. (1997) A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.*, **4**, 553-8.
90. Zhang, K., Pintilie, G.D., Li, S., Schmid, M.F. and Chiu, W. (2020) Resolving individual atoms of protein complex by cryo-electron microscopy. *Cell Res.*, **30**, 1136-9.
91. Westhof, E., Yusupov, M. and Yusupova, G. (2019) The multiple flavors of GoU pairs in RNA. *J. Mol. Recognit.*, **32**, e2782.
92. Auffinger, P., Ennifar, E. and D'Ascenzo, L. (2021) Deflating the RNA Mg^{2+} bubble. Stereochemistry to the rescue! *RNA*, **27**, 242-52.
93. Yu, T. and Chen, S.J. (2018) Hexahydrated Mg(2+) binding and outer-shell dehydration on RNA surface. *Biophys. J.*, **114**, 1274-84.
94. Wlodawer, A. (2017) Stereochemistry and validation of macromolecular structures. *Methods Mol. Biol.*, **1607**, 595-610.
95. Gilski, M., Zhao, J., Kowiel, M., Brzezinski, D., Turner, D.H. and Jaskolski, M. (2019) Accurate geometrical restraints for Watson-Crick base pairs. *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater*, **75**, 235-45.

96. Babai, K.H., Long, F., Maly, M., Yamashita, K. and Murshudov, G.N. (2024) Improving macromolecular structure refinement with metal-coordination restraints. *Acta Crystallogr. D Struct. Biol.*, **80**, 821-33.
97. Auffinger, P., Grover, N. and Westhof, E. (2011) Metal ion binding to RNA. *Met. Ions Life Sci.*, **9**, 1-35.
98. Auffinger, P., D'Ascenzo, L. and Ennifar, E. (2016) Sodium and potassium interactions with nucleic acids. *Met. Ions Life Sci.*, **16**, 167-201.
99. Remick, K.A. and Helmann, J.D. (2023) The elements of life: A biocentric tour of the periodic table. *Adv. Microb. Physiol.*, **82**, 1-127.
100. Kosaka, Y., Miyawaki, Y., Mori, M., Aburaya, S., Nishizawa, C., Chujo, T., Niwa, T., Miyazaki, T., Sugita, T., Fukuyama, M. *et al.* (2025) Autonomous ribosome biogenesis in vitro. *Nat. Commun.*, **16**, 514.
101. Michael, A.J. (2016) Polyamines in Eukaryotes, Bacteria, and Archaea. *J. Biol. Chem.*, **291**, 14896-903.
102. Gevrekci, A.O. (2017) The roles of polyamines in microorganisms. *World J. Microbiol. Biotechnol.*, **33**, 204.
103. Nishio, T., Sugino, K., Yoshikawa, Y., Matsumoto, M., Oe, Y., Sadakane, K. and Yoshikawa, K. (2020) K^+ promotes the favorable effect of polyamine on gene expression better than Na^+. *PLoS One*, **15**, e0238447.
104. Nissley, A.J., Shulgina, Y., Kivimae, R.W., Downing, B.E., Penev, P.I., Banfield, J.F., Nayak, D.D. and Cate, J.H.D. (2025) Structure of an archaeal ribosome reveals a divergent active site and hibernation factor. *Nat. Microbiol.*, **10**, 1940-53.
105. 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**, 179-93.
106. Brzezinski, D., Dauter, Z., Minor, W. and Jaskolski, M. (2020) On the evolution of the quality of macromolecular models in the PDB. *FEBS J.*, **287**, 2685-98.
107. Ma, H., Pham, P., Luo, B., Rangan, R., Kappel, K., Su, Z. and Das, R. (2023) Auto-DRRAFTER: Automated RNA modeling based on cryo-EM density. *Methods Mol. Biol.*, **2568**, 193-211.
108. Li, T., He, J., Cao, H., Zhang, Y., Chen, J., Xiao, Y. and Huang, S.Y. (2025) All-atom RNA structure determination from cryo-EM maps. *Nat. Biotechnol.*, **43**, 97-105.
109. Puyo-Fourtine, J., Zgarbova, M., Kantin, L., Jurecka, P. and Duboue-Dijon, E. (2025) Implicit electronic polarization via charge scaling in Amber OLx nucleic acid force fields improves ion pairing without backbone reparametrization. *J. Chem. Theory Comput.*, **21**, 8583-600.
110. Fan, S., Mason, P.E., Chamorro, V.C., Shanks, B.L., Martinez-Seara, H. and Jungwirth, P. (2025) Charge scaling force field for biologically relevant ions utilizing a global optimization method. *J. Chem. Theory Comput.*, **21**, 9023-34.
111. Maki-Yonekura, S., Kawakami, K., Takaba, K., Hamaguchi, T. and Yonekura, K. (2023) Measurement of charges and chemical bonding in a cryo-EM structure. *Commun. Chem.*, **6**, 98.
112. Yonekura, K., Matsuoka, R., Yamashita, Y., Yamane, T., Ikeguchi, M., Kidera, A. and Maki-Yonekura, S. (2018) Ionic scattering factors of atoms that compose biological molecules. *IUCrJ*, **5**, 348-53.
113. Yonekura, K. and Maki-Yonekura, S. (2016) Refinement of cryo-EM structures using scattering factors of charged atoms. *J. Appl. Cryst.*, **49**, 1517-23.
114. Marques, M.A., Purdy, M.D. and Yeager, M. (2019) CryoEM maps are full of potential. *Curr. Opin. Struct. Biol.*, **58**, 214-23.
115. Bick, T., Dominiak, P.M. and Wendler, P. (2024) Exploiting the full potential of cryo-EM maps. *BBA Adv.*, **5**, 100113.
116. 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.
117. Anderson, M., Schultz, E.P., Martick, M. and Scott, W.G. (2013) Active-site monovalent cations revealed in a 1.55-Å-resolution hammerhead ribozyme structure. *J. Mol. Biol.*, **425**, 3790-8.
118. 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**, 6369-81.

Table 1. Occurrence of well-defined RNA octahedral Mg^{2+} binding sites in the 8b0x and 9q87 structures. For each Mg^{2+} binding site, *"high", "fair"* and "*low*" quality binding site observed occurrences are provided in parenthesis.

	8b0x		**9q87**	
Outer-shell				
$6O_w$	**21** (17/1/4)		**118** (48/53/17)	
Pentahydrated				
$O_{ph}.5O_w$	**37** (28/7/2)		**149** (71/57/21)	
$O_b.5O_w$	**3** (3/–/–)		**16** (8/7/1)	
$O_r.5O_w$	**1** (1/–/–)		**1** (–/1/–)	
$N_b.5O_w$	–		**9** (1/5/3)	
$O_{coo}.5O_w$ (with r-protein)	–		**1** (1/–/–)	
Tetrahydrated	*cis-*	*trans-*	*cis-*	*trans-*
$2O_{ph}.4O_w$	**13** (9/3/1)	**3** (2/-/1)	**46** (25/12/9)	**15** (4/6/5)
$2O_b.4O_w$	–	–	**1** (–/1/–)	–
$2N_b.4O_w$	–	–	**3** (1/2/–)	–
$O_{ph}.O_b.4O_w$	**3** (2/–/1)	**1** (–/–/1)	**10** (7/1/2)	**1** (1/–/–)
$O_{ph}.N_b.4O_w$	–	–	**2** (2/–/–)	**2** (1/–/1)
$O_{ph}.N_{his}.4O_w$ (with r-protein)	**1** (1/–/–)	–	**2** (2/–/–)	–
Trihydrated	*fac-*	*mer-*	*fac-*	*mer-*
$3O_{ph}.3O_w$	**3** (2/1/–)	**6** (6/–/–)	**7** (5/2/–)	**9** (9/–/–)
$2O_{ph}.O_b.3O_w$	–	**1** (1/–/–)	–	**2** (2/–/–)
$2O_{ph}.O_r.3O_w$	–	–	**1** (1/–/–)	–
$2O_{ph}.N_b.3O_w$	**1** (1/–/–)	–	**2** (2/–/–)	–
$O_{ph}.2O_b.3O_w$	–	–	–	**1** (–/1/–)
$2O_{ph}.O_{cno}.3O_w$ (with r-protein)	–	**1** (1/–/–)	–	**1** (1–/–)
Dihydrated	*CIS-*	*TRANS-*	*CIS-*	*TRANS-*
$4O_{ph}.2O_w$	**1** (1/–/–)	**1** (1/–/–)	**1** (1/–/–)	**1** (1/–/–)
$3O_{ph}.O_b.2O_w$	–	–	**1** (1/–/–)	–
$3O_{ph}.O_r.2O_w$	–	–	**1** (1/–/–)	–
$2O_{ph}.2O_b.2O_w$	**1** (–/–/1)	–	**1** (–/1/–)	–
$2O_{ph}.O_{bb}.N_{his}.2O_w$ (with r-protein)	**1** (1/–/–)	–	**1** (1/–/–)	–
$O_{ph}.3O_{bb}.2O_w$ (with r-protein)	–	–	**1** (–/–/1)	–
$3O_{coo}.N_{his}.2O_w$ (r-protein)	–	–	**1** (–/1/–)	–
Sub-total:	**100** (77/12/11)		**406** (197/150/60)	
With issues (not in table)**:**	**261**		**9**	
Total:	**361**		**415**	

Table 2. List of X-ray structures of the hammerhead ribozyme with assigned Mg^{2+} ions. The table lists the six PDB- structures with assigned Mg^{2+} ions. The six associated buffers contain cacodylate anions (CAC, $AsO_2(CH_3)_2^-$). Structures with no assigned divalent ions, despite the presence of Mg^{2+} in the buffer, are highlighted in grey. A complete list of all 26 hammerhead structures is provided in **Table S2**.

	PDB	Date	Res.(Å)	Ions	Nucl.	Buffer components	Ref
1.	1hmh	1995	2.6	–	34+13	2.4M Li_2SO_4; 10mM $MgCl_2$; CAC;	(70)
2.	1mme	1996	3.1	–	25+16	100mM NH_4OAc; 10mM $Mg(OAc)_2$; CAC;	(71)
3.	299d	1997	3.0	–	25+16	1.8M Li_2SO_4; 10mM $MgSO_4$; CAC;	(85)
4.	301d	1997	3.0	5 Mg^{2+}	25+16	1.8M Li_2SO_4; 10mM $MgSO_4$; CAC;	(85)
5.	2quw	2008	2.2	17 Mg^{2+}	57+12	28mM $MgCl_2$, CAC;	(116)
6.	2qus	2008	2.4	16 Mg^{2+}	69	28mM $MgCl_2$, CAC;	(116)
7.	3zp8	2013	1.55	16 Na^+	43+20	1.7M Na-malonate; 10mM $MgCl_2$;	(117)
8.	5dh6	2015	2.78	5 Mg^{2+}	48+20	0.9 M KCl; 10mM $MgCl_2$; CAC;	(118)
9.	5dqk	2016	2.71	1 Mg^{2+} + 4 K^+	48+20	50mM $Mg(OAc)_2$; CAC;	(118)
10.	5eao	2016	2.99	8 Mg^{2+}	48+20	0.4M KCl; 50mM $Mg(OAc)_2$; CAC;	(118)

Table 3. Comparison of Mg^{2+} binding sites in the 9cbu cryo-EM and 1hr2 X-ray *Tetrahymena* group I intron structures. An asterisk indicates the presence of *d*(Mg…O/N) distances in the second 2.3–2.6 Å range. In the **1hr2** structure (88), the 10 chain "A" ions (Mg:51 to Mg:64) are distant from RNA atoms and do not show octahedral densities, marking their hexahydrated Mg^{2+} assignment as speculative. These densities should more likely be attributed to other buffer components. The eight assigned $Mg[H_2O]_6^{2+}$ chain "A" (Mg:14 to Mg:27) ions are valid, as some of their water molecules form hydrogen bonds with guanine O6/N7 atoms or OP atoms. In **9cbu** (22), 16 out of 22 Mg^{2+} ions display incorrect stereochemistry and are associated with clashes when hydrated. The ions present in consensus **9cbx** structure are shown in the last column.

9cbu/1hr2		9cbu	1hr2[a]	Contact atom
	Cryo-EM assignment	Model	X-ray assignment	
Mg:501/–	*mer*-$3O_{ph}$*→	mer-$3O_{ph}.3O_w$	*Not in P4-P6*	OP1:257 (–/9cbx)
Mg:502/Mg:1	*fac*-$3O_{ph}$ →	*fac*-$3O_{ph}.3O_w$	*fac*-$3O_{ph}.3O_w$*	OP1:183 (1hr2/9cbx)
Mg:503/Mg:3	*mer*-$3O_{ph}$ →	*mer*-$3O_{ph}.3O_w$ *or Trans*-$4O_{ph}.2O_w$	*mer*-$3O_{ph}.3O_w$	OP1:184 (1hr2/9cbx)
Mg:504/–	*cis*-$2O_{ph}$ →	*cis*-$2O_{ph}.4O_w$	*Not in P4-P6*	OP1:301 (1hr2/9cbx)
Mg:505/–	N_7	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:506/–	**N_7***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:507/–	*cis*-$2O_{ph}$* →	*cis*-$2O_{ph}.4O_w$	*Not in P4-P6*	OP1:256 (–/9cbx)
Mg:508/–	*cis*-$2O_{ph}$ →	*cis*-$2O_{ph}.4O_w$	*Not in P4-P6*	OP1:304 (–/9cbx)
Mg:509/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:510/–	**O_{ph}***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:511/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:512/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:513/–	**O_6***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:514/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:515/–	**O_{ph}***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:516/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:517/–	**N_7***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:518/–	**O_{ph}***	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:519/–	O_{ph}	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:520/–	O2'	**Clash/No octahedral density**	*Not in P4-P6*	–
Mg:521/–	**O2'***	**Clash**	*Not in P4-P6*	–
Mg:522/–	**O_{ph}***	**No octahedral density**	*Not in P4-P6*	–
–/Mg:5	Can be modelled →	*cis*-$O_b.N_b.4O_w$	*cis*-$O_b.N_b.4O_w$*	N7:163 (1hr2/9cbx)
–/Mg:7	Can be modelled →	$O_{ph}.5O_w$	$O_{ph}.5O_w$	OP2:171 (1hr2/–)
–/Mg:9	Can be modelled →	*cis*-$2O_{ph}.4O_w$	$O_{ph}.5O_w$	OP2:173 (1hr2/9cbx)
–/Mg:11	Can be modelled →	$O_b.5O_w$	$O_b.5O_w$	O6:188 (1hr2/–)
–/Mg:57	–	–	*fac*-$2O_{ph}.O_r.3O_w$*	OP2:249 (1hr2/–)

[a] *1hr2 is a 160 nucleotide-long P4-P6 fragment taken from the 387 nucleotide-long group I intron (residue G:103 to C:260). 1hr2 is a C:109 deletion mutant.*

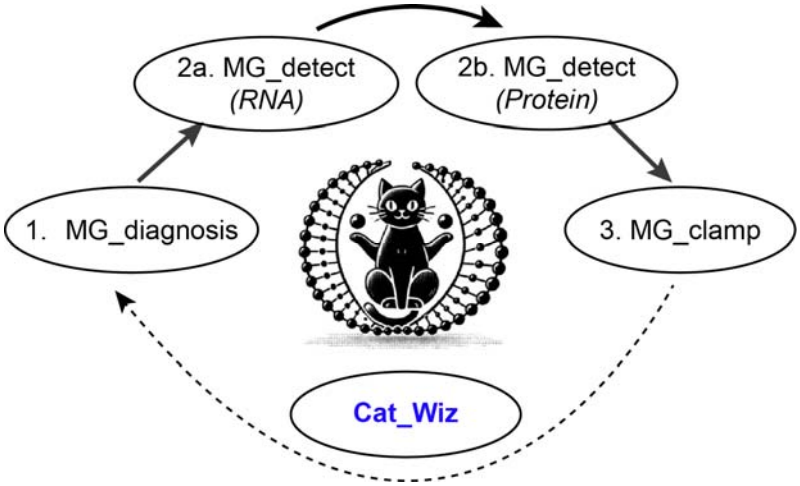

Figure 1. Workflow of the *Cat_Wiz* Scripts. This diagram illustrates the recommended sequence for using the four independent *Cat_Wiz* scripts. To ensure the accuracy of the model, it is advised to re-run the *MG_diagnosis* script whenever structural changes were made.

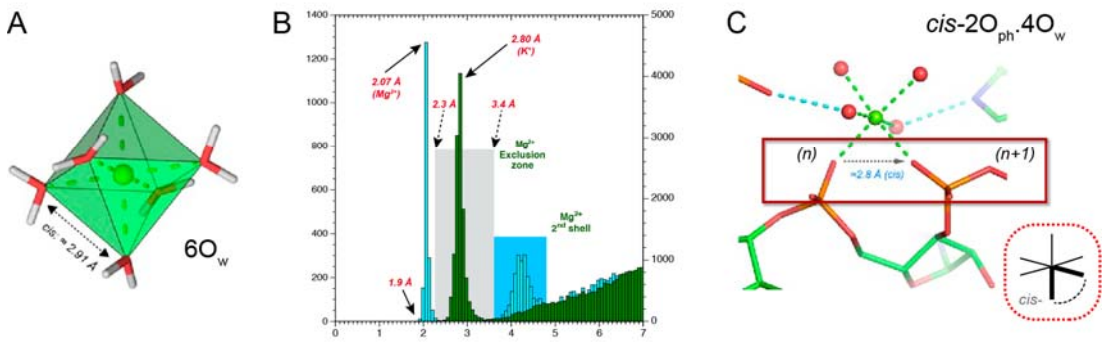

Figure 2. Stereochemical features of Mg^{2+} ions and Mg^{2+} clamps. (**A**) A hexahydrated Mg^{2+} ion ($6O_w$) from a high-resolution Cambridge Structural Database (CSD) entry, showing characteristic $d(O_w \ldots O_w)$ distances around 2.9 Å (14). (**B**) Histograms of Mg^{2+} (cyan) and K^+ (green) to oxygen coordination distance from the CSD. The distinct ~0.7 Å distance between the two peaks is an efficient criterion for distinguishing these ions. The grey rectangle indicates the Mg^{2+} exclusion zone, while the blue rectangle defines the second coordination shell. (**C**) The geometry of a *cis*-$2O_{ph}.4O_w$ magnesium clamp, which involves two consecutive phosphates.

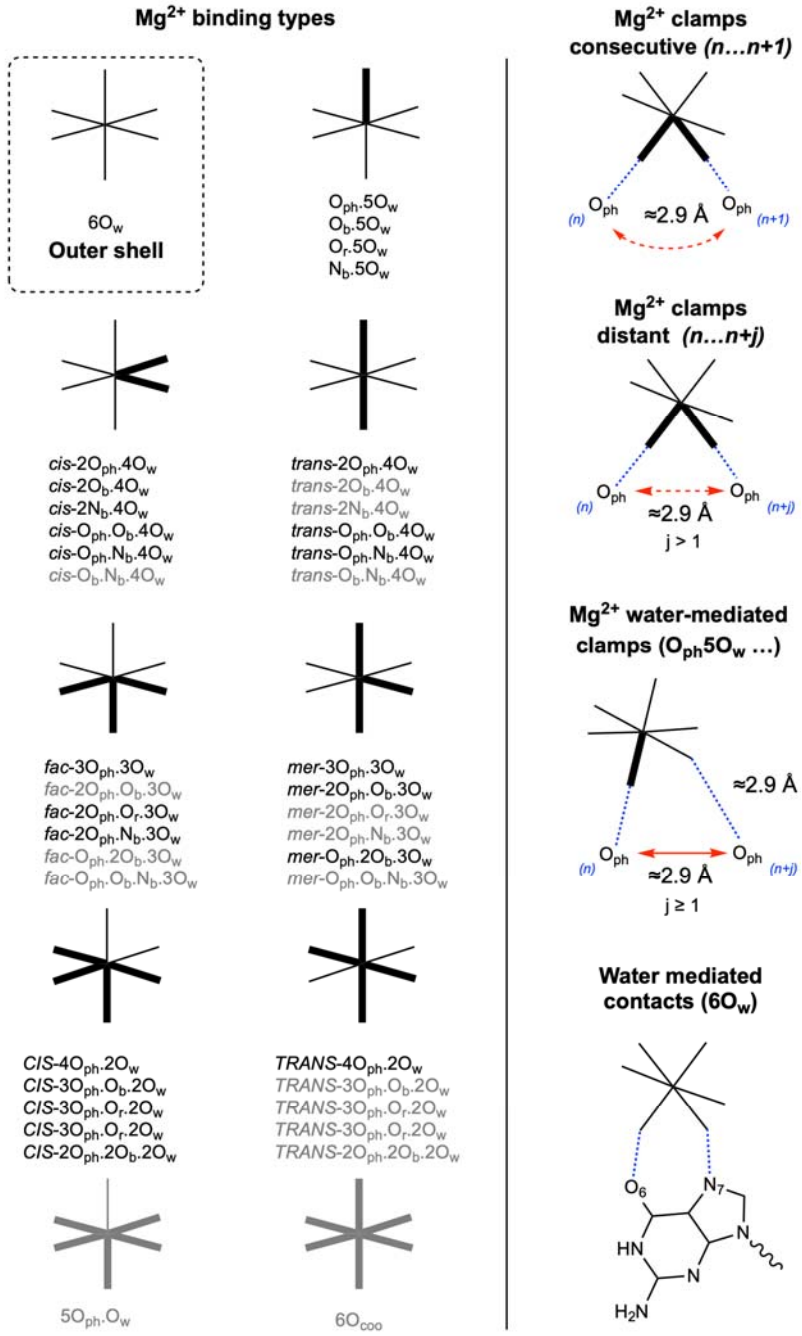

Figure 3. Updated view of Mg^{2+} binding motifs. (Left) This panel presents all inner- and outer-shell Mg^{2+} binding motifs identified in **9q87** and is an update of the list derived from **8b0x** (14). For clarity, motifs involving ribosomal proteins have been excluded. Greyed out binding types have until now weak, insufficient, or no structural evidence. Binding to O3' occurs only in exceptional configurations involving terminal residues and is not mentioned in the figure. Binding to O5' has not been observed (Right) The top two figures illustrate a classic Mg^{2+} clamp, in which the ion coordinates in *cis*- to two phosphate atoms, either from consecutive (j=1) or distant nucleotides (j>1). The third figure depicts a water-mediated $O_{ph}.5O_w$ motif. The bottom figure shows an outer-shell $6O_w$ ion forming a guanine Hoogsteen-edge contact.

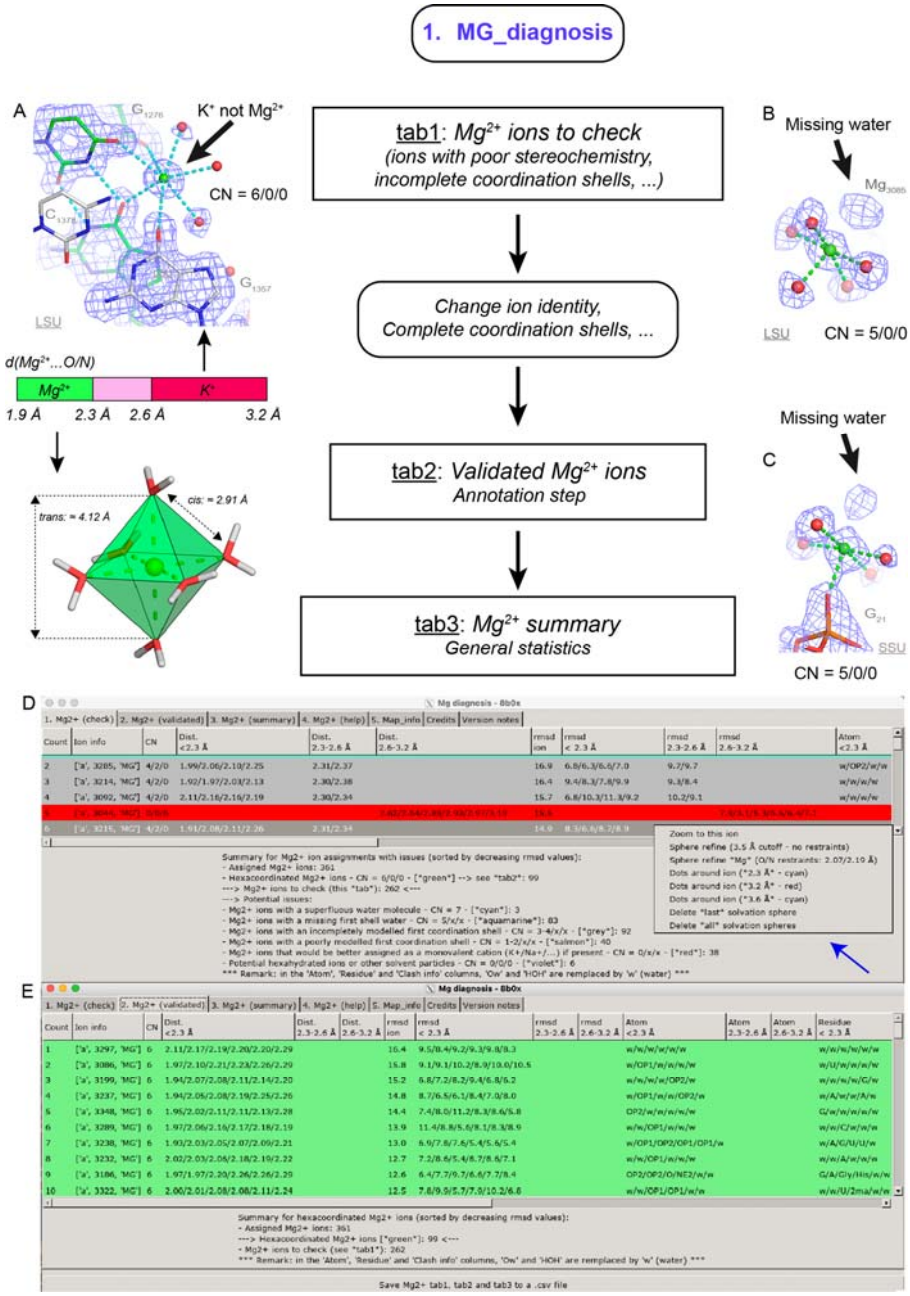

Figure 4. Workflow and features of the *MG_diagnosis* script. The central column illustrates the *MG_diagnosis* primary functions, i.e., assessing the stereochemical quality of pre-assigned ions and recommending changes based on specific coordination number (CN) and distance criteria. (**A**) A color-coded scale is used to evaluate coordination distances. Green corresponds to an ideal 1.9–2.3 Å range, pink to an improbable 2.3–2.6 Å range and red to a 2.6–3.2 Å range characteristic of misassigned K^+ ions. The top-right panel shows a misidentified K^+ ion with CN = 6/0/0 (14). (**B, C**) Two Mg^{2+} ions with CN = 5/0/0, pointing to a missing first shell water as revealed by density maps. (**D**) Partial view of the *MG_diagnosis* tab1, which lists all identified issues. A pop-up menu (see blue arrow) accessible through a right-mouse click, provides context-specific options. (**E**) Partial view of the *MG_diagnosis* tab2 (the tab1 and tab2 extensions to the right are not shown).

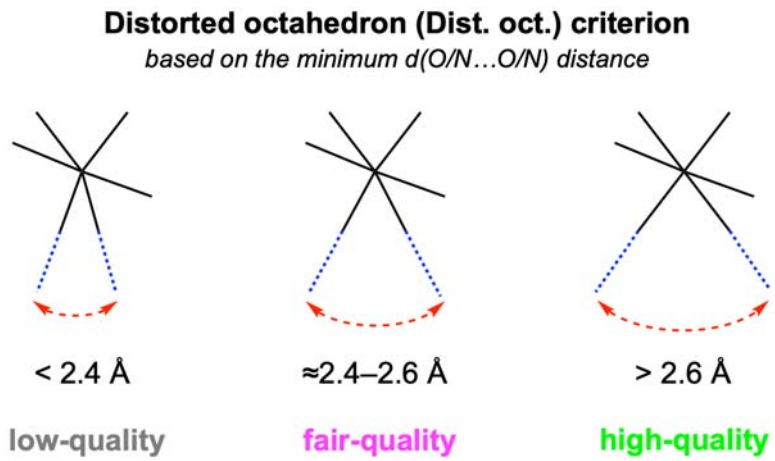

Figure 5. The distorted octahedron (Dist. Oct.) criterion. The quality of a binding site is *"low"*, *"fair",* or *"high"* based on the smallest distance between any two first-shell coordinating atoms. It is important to note that this criterion is exclusively applied in tab2 of *MG_diagnosis*. These arbitrary distance cutoffs were empirically inferred.

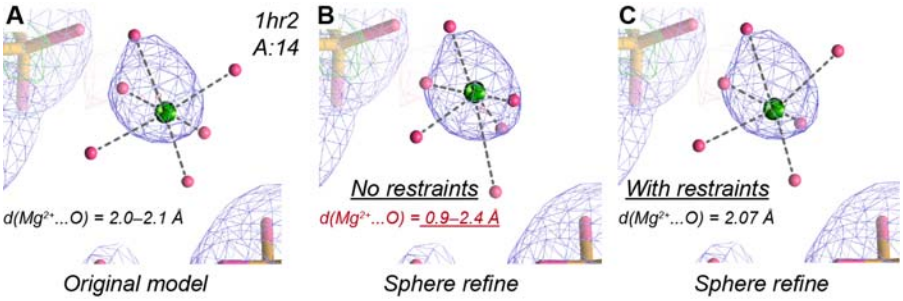

Figure 6. The effect of restraints on the *Coot "Sphere refine"* option. Different outcomes are observed when using the *Coot "Sphere refine"* command with and without distance restraints. (**A**) A hexahydrated Mg^{2+} ion modelled into a non-octahedral density pattern (88). (**B**) Refining the same position using the *"sphere_refine"* command without added restraints clearly shows a deformed coordination octahedron with out-of-range coordination distances. (**C**) When the refinement is performed with added restraints, the octahedral geometry of the hydrated Mg^{2+} ion is preserved, despite minor deformation.

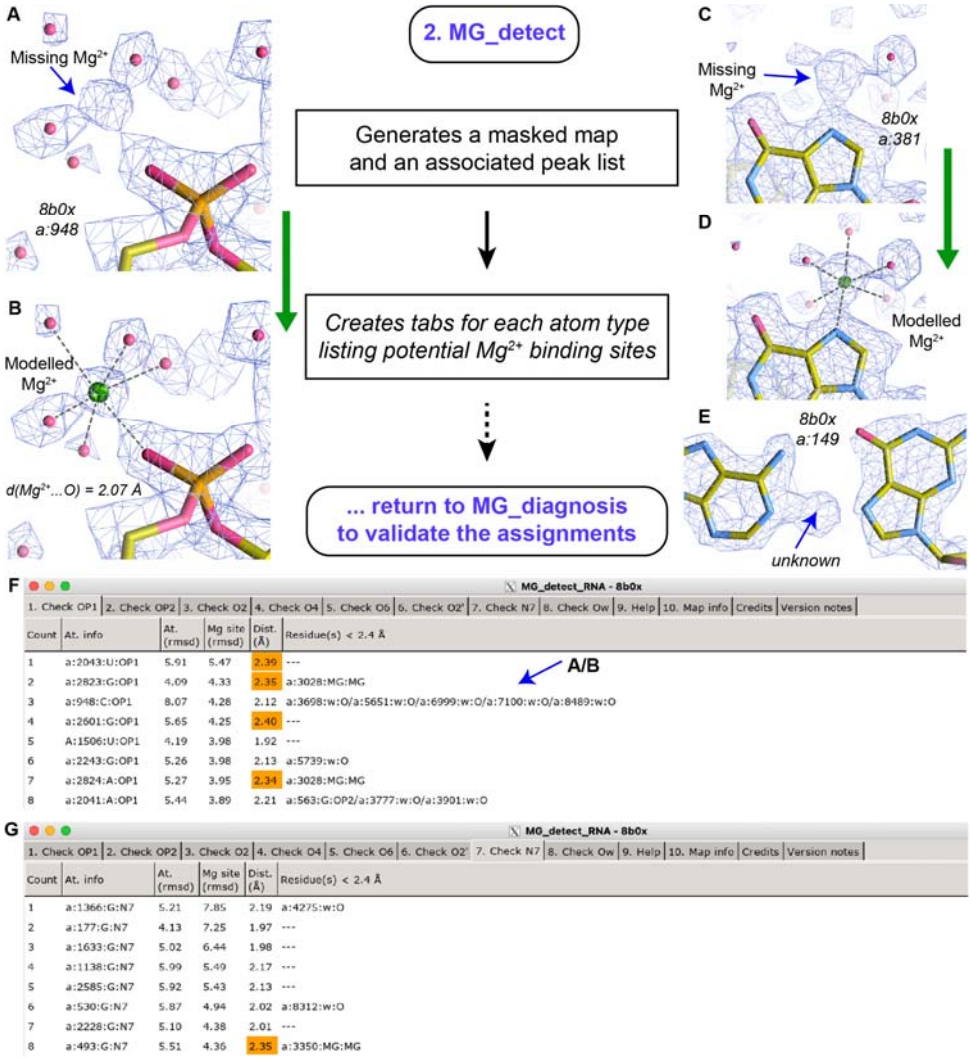

Figure 7. Organisation of the *MG_detect_RNA* script. *MG_detect_RNA* identifies potential Mg^{2+} binding sites by collecting the highest peaks from a masked map. (**A**) Shows a missing Mg^{2+} ion that is part of an $O_{ph}.5O_w$ binding site successfully identified by *MG_detect_RNA*. (**B**) Shows subsequent modelling of this site using *MG_diagnosis* options. (**C/D**) Show a similar process for a missing Mg^{2+} that is part of an $N_7.5O_w$ site. (**E**) Shows density associated with an unidentified ion that is part of a metal-mediated base pair as described in (14). (**E/G**) Show partial views of the generated tabs which identify potential Mg^{2+} binding sites near OP1 and N7 atoms, respectively (marked by blue arrows). Cells in orange mark distances from the potential Mg^{2+} binding site to the selected atom that are above 2.3 Å or below 1.9 Å.

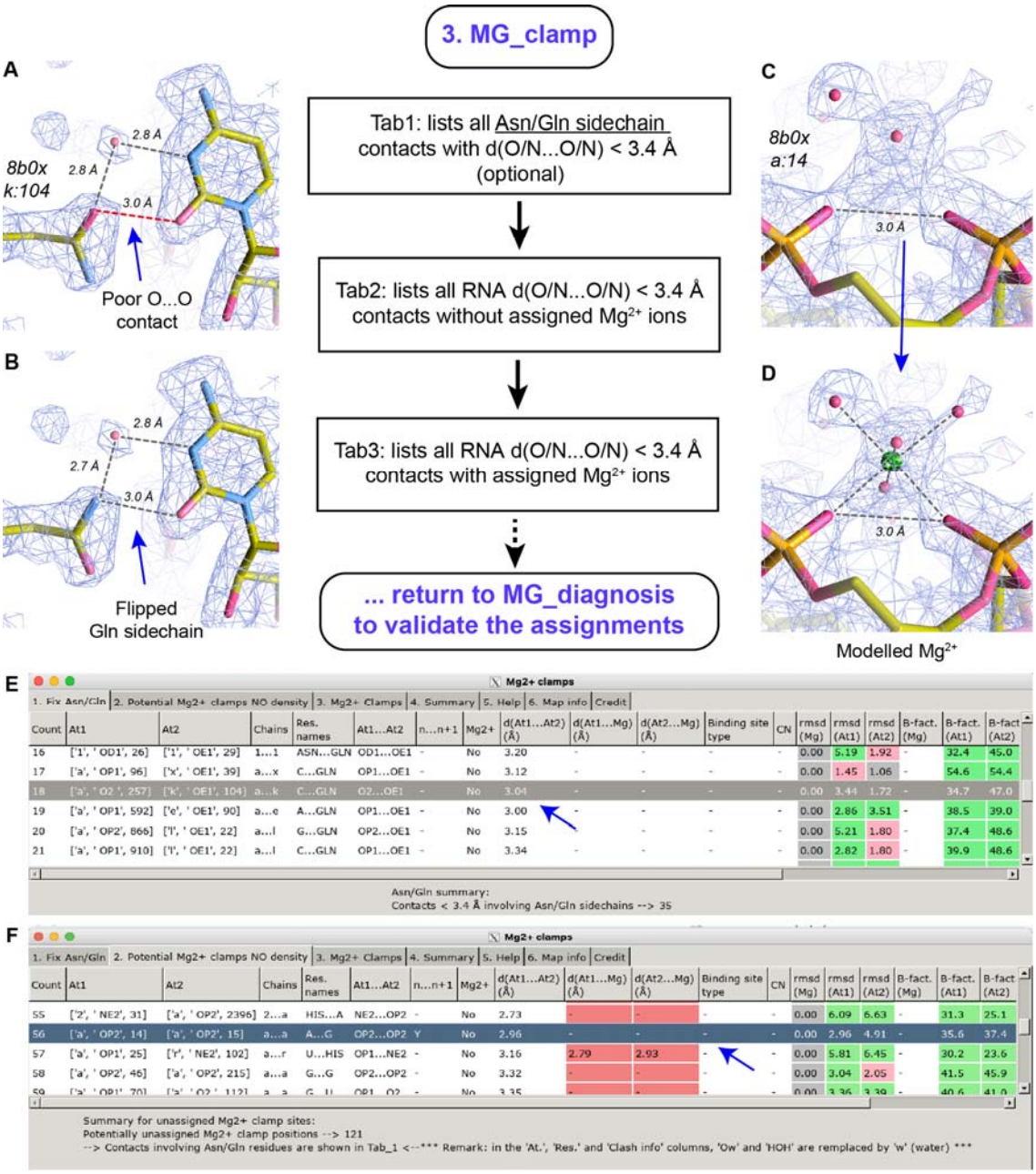

Figure 8. Organisation of the *MG_clamp* script. *MG_clamp* locates non-hydrogen-bonded pairs of atoms with *d*(O/N…O/N) below 3.4 Å. These pairs are common to all higher-binding motifs (except *trans*-$2O_{ph}.4O_w$; see **Figure 3**). When such pairs are identified, a visual inspection of the associated density patterns is required to determine whether the site can accommodate a hydrated Mg^{2+} ion. (**A**) Shows a misoriented Gln side chain that creates an unrealistic *d*(O…O) contact near 3.0 Å. (**B**) Shows the correct side chain orientation obtained by using the *Coot "flip_sidechain_180°"* tool. (**C**) Illustrates how a $2O_{ph}.4O_w$ binding pattern with *d*(O…O) ≈ 3.0 Å can be identified by *MG_clamp*. (**D**) Shows the modelled binding pattern. (**E/F**) Show a partial view of tab1 and tab2 generated by *MG_clamp*. In panel **E**, a blue arrow marks the short *d*(O…O) distance caused by an incorrectly oriented Gln side chain (see **A/B**). In panel **F**, a blue arrow marks a short *d*(O…O) distance detected by *MG_clamp* that is associated with an unassigned *cis*-$2O_{ph}.4O_w$ Mg^{2+} ion (see **C/D**).

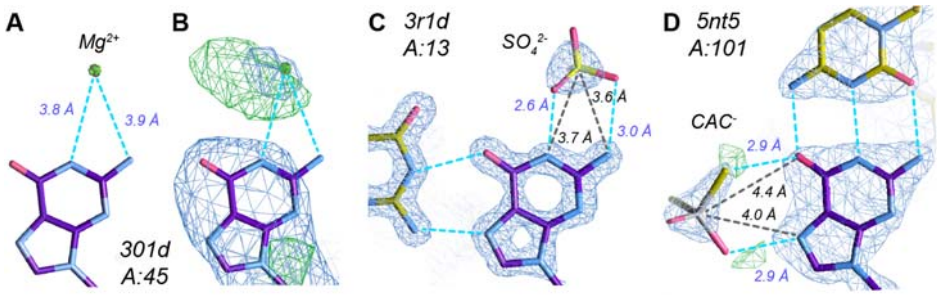

Figure 9. Anion alternatives to questionable Mg^{2+} assignments. This figure presents examples of how questionable Mg^{2+} assignments may have alternative explanations. (**A/B**) In the **301d** hammerhead structure, the Mg^{2+} ion (Mg:45) is unconvincingly positioned near an electropositive Watson-Crick edge (81,82). A more plausible assignment for this density would be a sulfate anion (SO_4^{2-}), as observed in the 1.45 Å resolution structure **3r1d** (**C**), or a cacodylate anion ($AsO_2(CH_3)_2^-$), as seen in the 2.30 Å resolution structure **5nt5** (**D**). Panel **B** shows that *fo-fc* density patterns (3.0 σ; green) suggest the presence of a more electron-dense particle than Mg^{2+}. The *2fo-fc* density patterns (1.5 σ; blue) are shown in all panels.

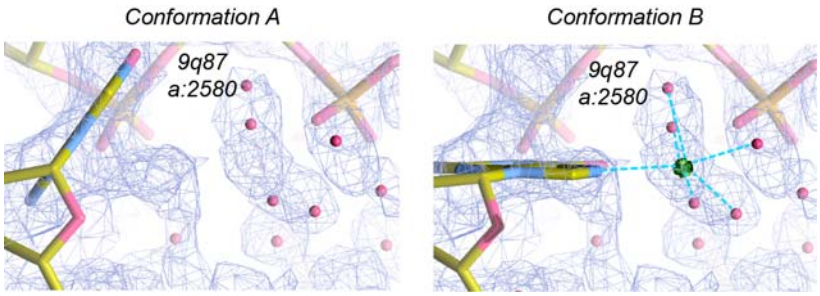

Figure 10. Alternative conformation of the a:2580 residue in 9q87. Conformation A shows the major orientation of a:G:2580. Conformation B shows an alternate conformation involving a well-defined N7.5O_w contact.

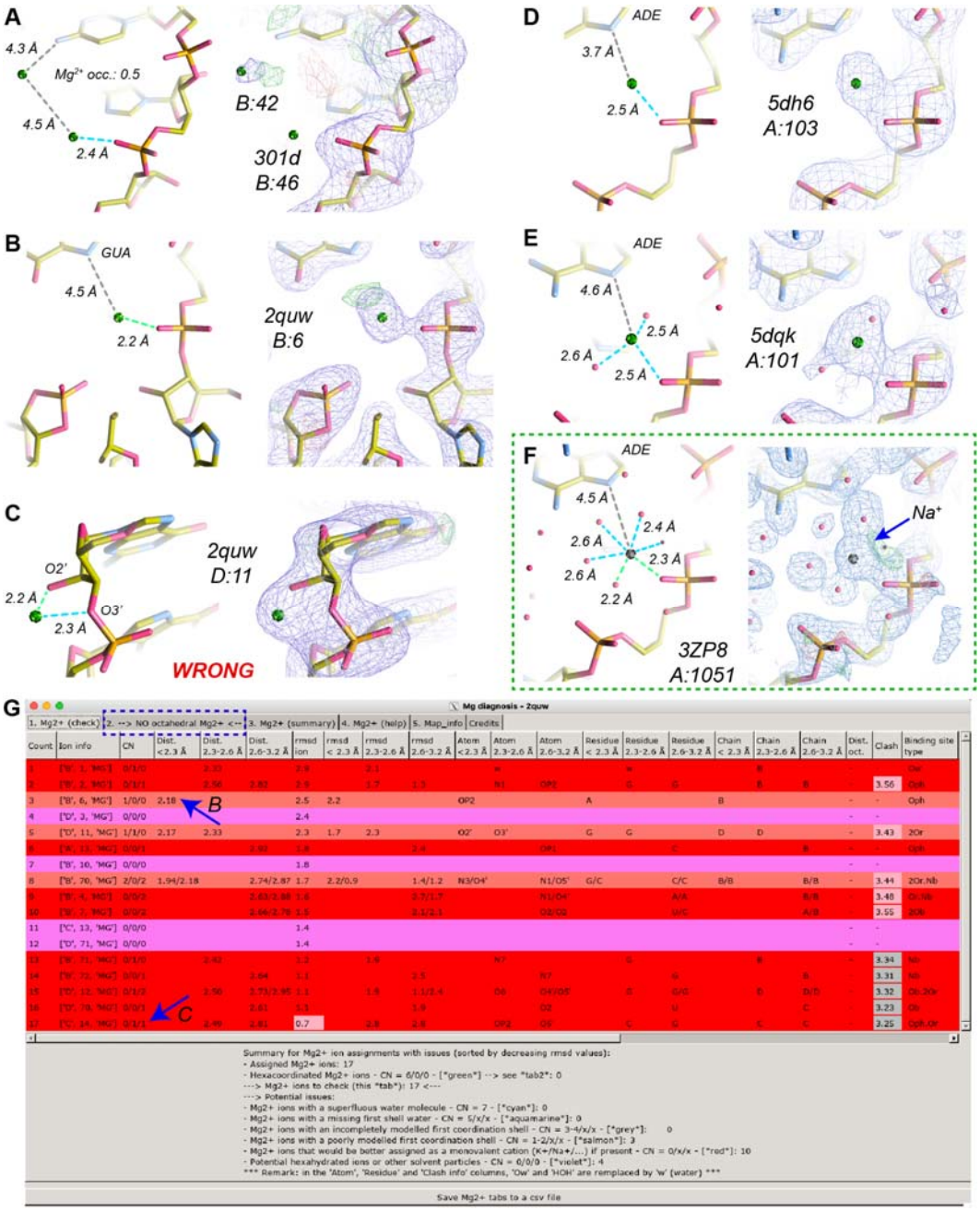

Figure 11. Ion binding assignments in hammerhead ribozyme X-ray structures. (**A**) Mg^{2+} ion tentatively assigned to a weak density spot in the 3.0 Å resolution **301d** structure (occupancy = 0.5). (**B**) A more realistic Mg^{2+} ion from the 2.2 Å resolution **2quw** structure, although lacking octahedral density. (**C**) A stereochemically improbable Mg^{2+} ion also found in **2quw**. (**D,E**) Two similar assignments from the 2.78 and 2.71 Å resolution **5hd6** and **5dqk** structures, respectively, neither of which display octahedral densities. In **5dqk**, the $d(O_{ph}\ldots O_w)$ > 2.6 Å distances further cast doubt on the assignments. (**F**) An equivalent position in the 1.55 Å resolution **3zp8** structure is assigned as an Na^+ ion occupying a $O_{ph}.5O_w$ site. (G) A partial view of the *MG_diagnosis* "tab1" output, showing the structural environment of the 17 Mg^{2+} ions assigned in **2quw**, with blue arrows indicating the ions highlighted in panels (**B**) and (**C**).

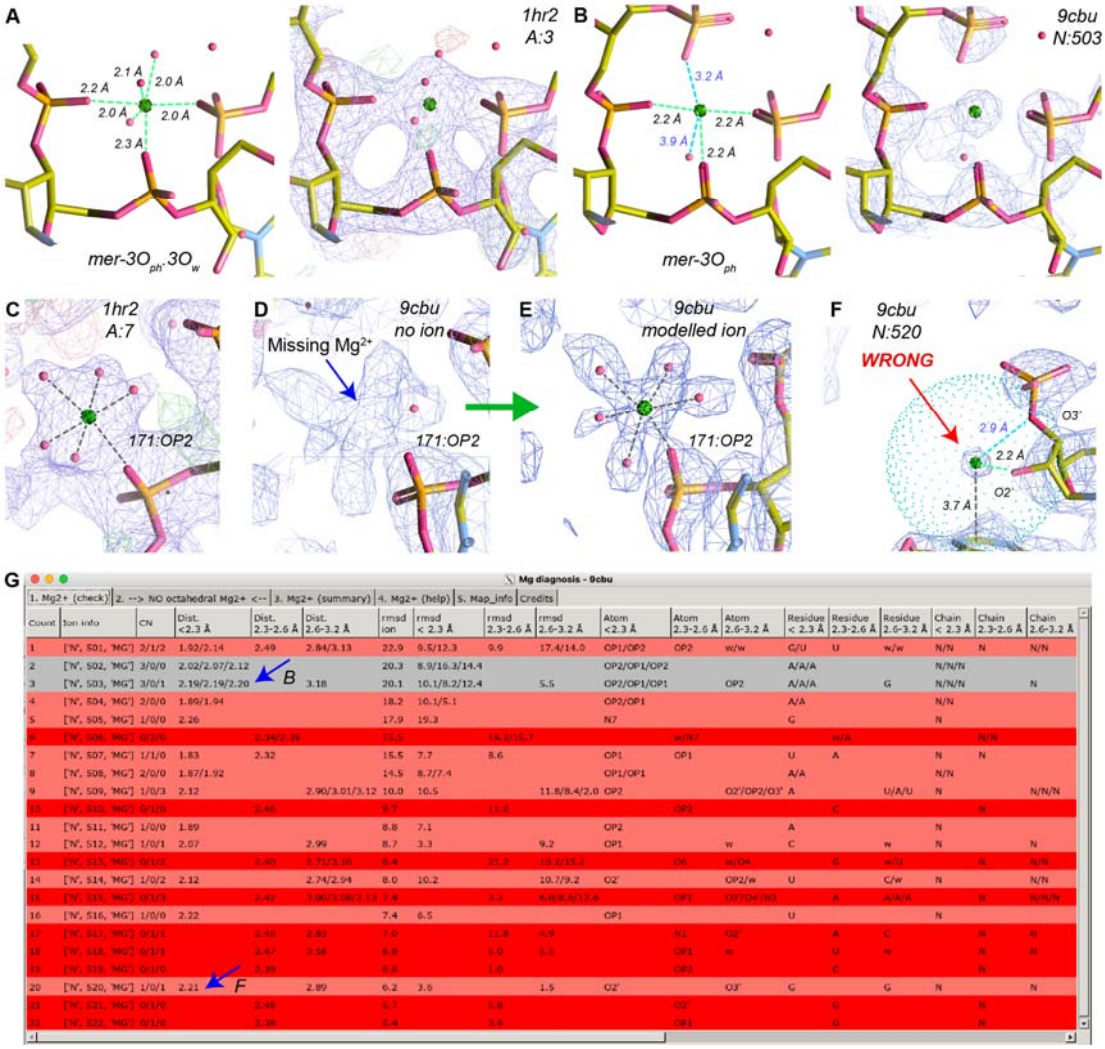

Figure 12. Ion binding assignments in two *Tetrahymena* group I intron structures: (**A**) A *mer*-3O$_{ph}$.3O$_w$ binding site in the 2.25 Å resolution **1hr2** X-ray structure (88), and (**B**) the equivalent site modelled in the 2.20 Å resolution **9cbu** cryo-EM structure (22). (**C**) An O$_{ph}$.5O$_w$ binding site within an octahedral density pattern in **1hr2**, and (**D-E**) the corresponding site in **9cbu**, which is unoccupied despite clear octahedral density (**D**) but can be correctly modelled with an O$_{ph}$.5O$_w$ magnesium ion using *MG_diagnosis* (**E**). (**F**) A 3.6 Å radius sphere, generated by *MG_diagnosis*, emphasizes improbable contacts established by an incorrectly modelled Mg^{2+} ion in **9cbu** positioned too close to an O3' atom and unrealistically stacking over an adenine nucleobase. (**G**) A partial view of the *MG_diagnosis* "tab1" summarizing the structural environment of the 22 Mg^{2+} ions assigned in **9cbu**, with blue arrows marking the ions shown in panels (**B)** and **(F**).

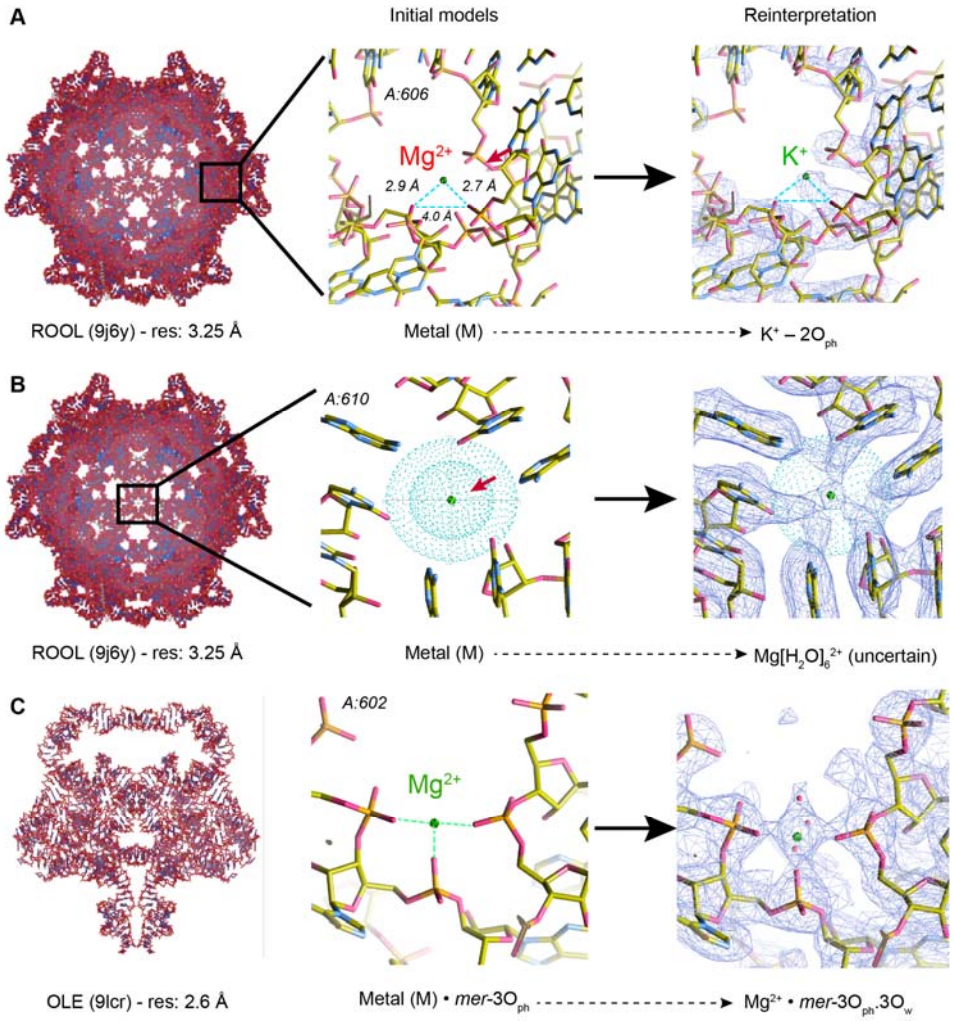

Figure 13. Ion placements in two low-resolution quaternary structures. (A) In the **9j6y** PDB structure (26), Mg^{2+} ions with coordination distances in the 2.6–2.9 Å range should be reassigned as K^+ ions. **(B)** This ion is located on a three-fold symmetry axis. Although the resolution of the structure results densities with no octahedral geometry that are difficult to interpret, the distances to the neighbouring residues allow for the placement of a $Mg[H_2O]_6^{2+}$ ion. This is visualized by spheres of dots with 2.3 and 3.6 Å radii which can be viewed using the *MG_diagnosis* script. Nevertheless, alternative assignments should be considered. **(C)** In this 2.6 Å resolution OLE structure a density pattern near three O_{ph} atoms results in a *mer*-$3O_{ph}.3O_w$ assignment. The three missing water molecules can be modelled without introducing clashes by using the *Coot* tools and the *Cat_Wiz* *"Sphere refine with restraints"* option.