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 facilitates the localization, 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 ion binding sites, *MG_clamp* classifies recurrent clamp motifs. *Cat_Wiz* includes also a complete binding site annotation system. The implemented stereochemical principles that provide a robust experimental foundation for characterizing Mg^{2+} binding sites were derived from an earlier 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 locates unnoticed ions, corrects misassignments, and improves stereochemical fidelity in hours 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 design of RNA 3D architectures through the description of unique Mg^{2+}-dependent backbone folds and works also be for locating/validating Mg^{2+} binding sites in proteins.

INTRODUCTION

Hexahydrated and partially dehydrated magnesium (Mg^{2+}) ions are central to RNA structure, folding, and catalysis (1-7). Therefore, accurate placement and subsequent validation of Mg^{2+} in experimental models is critical for mechanistic interpretation, comparative mining and other computational use. While several studies investigated Mg^{2+} coordination to RNA (5,8-14), misassignments remain common in deposited models where ions are often placed in environments incompatible with the characteristic octahedral geometry of these ions thereby compromising database surveys and associated inferences (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 that prompted the creation of a robust experimentally based coordination pattern index (14,19). In turn, this 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 recurrent 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 help in promoting the development of accurate RNA three-dimensional structure prediction tools (23-27).

Herein, we present *Cat_Wiz* (*"Cat"* stands 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 complete and 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 locate unmodeled Mg^{2+} ions; finally, *MG_clamp* identifies, classifies, and annotates Mg^{2+} clamps (see **Figure 2**), 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 in hammerhead ribozymes (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)* cryo-EM quaternary RNA assemblies with resolutions above 2.6 Å (34). Across all these cases, *Cat_Wiz* successfully resolved incorrect assignments, recovered missing hydrated ions and differentiated Mg^{2+} from K^{+} by enforcing representative coordination distances, thereby improving the fidelity of deposited structures. Moreover, *Cat_Wiz* can also assist in the refinement and interpretation of nucleic acid and protein systems at the refinement stage.

The stereochemical Mg^{2+} priors we defined in **Figure 2** were derived from small-molecule statistics through an exploration of the *Cambridge Structural Database* (CSD) and were described elsewhere (11,12,14,35). 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 (36-44). In that regard, CMM and *Cat_Wiz,* complement each other.

Cat_Wiz also complements popular refinement programs such as REFMAC5 and Phenix (45,46) by providing stereochemically-aware Mg^{2+} binding site evaluation. The suite of scripts offers a set of restraints to help 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 particles and Mg^{2+} ions that can improve machine learning (ML) and artificial intelligence (AI) workflows. Next, we *(i)* outline the current *Cat_Wiz* stereochemical nomenclature, *(ii)* describe the different *Cat_Wiz* modules, *(iii)* apply the toolkit to different case studies introduced above and finally *(iv)* we address *Cat_Wiz* limitations and future extensions.

METHODS

Cat_Wiz installation in *Coot*

Cat_Wiz is a suite of independent Python scripts (**Figure 1**) developed for the molecular modeling program *Coot* (28). To use the scripts, load in *Coot* a coordinate (PDB or mmcif) 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 *Cat_Wiz* requires a few minutes to generate a map smoothed by a factor of 1.25 (default value). During this process, the *Coot* console displays *"waiting"* messages. For X-ray structures, the PDB often provides structure factor files (.sf) that need to be converted into a *.mtz* file using the SF-tool web server (https://sf-tool.wwpdb.org) (47). Default cutoffs were empirically derived and can be adjusted in the setup header of the scripts. The scripts were tested with the *Coot* version 0.9.8.98. A version adapted for *Coot* 1.xxx is currently under development. For further details, please consult the updated help pages provided by each module. It is advised to work with caution with *"over"* sharpened maps that might alter results of the *MG_detect* module and sometimes lead to zero or negative *B-factors* at the refinement stage in Phenix. Next, we first describe additions to a preceding version of our Mg^{2+} binding site nomenclature (13), before addressing specific features of each module.

Mg^{2+} binding site nomenclature description and additions

Cat_Wiz is based on a Mg^{2+} binding site nomenclature derived from the MgRNA study (10-12,14) and shown in (**Figure 3**). 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; note the slightly longer *d*(Mg^{2+}…N) coordination distance. 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' and O3'/O5' ribose and phosphate oxygens, respectively; *"O_b"* for O2/O4/O6 nucleobase oxygens; and *"N_b"* for non-protonated N1/N3/N7 nucleobase nitrogens.

• *General nomenclature:* As an example, $2O_{ph}.4O_w$ refers to 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. When three non-water atoms are in *cis-* to each other, the corresponding isoforms are named *fac*-$3O_{ph}.3O_w$ (*facial*). When one O_{ph} pair is in *trans-* (10,14) it is named *mer*-$3O_{ph}.3O_w$ (*meridional)*. 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 binding types from the more common *cis-/trans*-$2O_{ph}.4O_w$ types, the *CIS-/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 the previously used O_{back} denomination (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:* In rare instances, interactions between O/N atoms separated by < 3.4 Å are 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). Due to structural compaction, these water-mediated contacts also appear in complex motifs but, for the sake of nomenclature simplification, they are not specifically annotated by the *Cat_Wiz* modules.

• *Ribose O3'/O5' atoms and Mg^{2+} coordination:* Recent investigations on group I and II introns described catalytic sites with suspected Mg^{2+}…O3' contacts (48,49). This is a puzzling observation since direct coordination of divalent ions to O3' atoms have not been observed in the ensemble of structures we analyzed. Therefore, such contacts were not included in **Figure 3**. The only exception we came across (see below) is related to terminal nucleotides where the O3'-H group can coordinate to a Mg^{2+} ion, alone or in combination with the O2'-H of the same ribose (50).

• *Monohydrated Mg^{2+} ions:* 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 (50). 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 atoms have 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**) (51). Both monohydrated and completely dehydrated patterns are greyed out in **Figure 3** since to the best of our knowledge, they have not been observed in representative nucleic acid systems.

• Mg^{2+} *coordination to metabolites, drugs and buffer molecules:* Examples include coordination to the tetracycline antibiotic in a ribosome structure (TAC; PDBid: **8cgj**) (52), to ZMP in a riboswitch (PDBid: **4xwf**) (53), 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**) (54) or acetates (ACT/ACY; PDBid: **1yvp**) (55). Occasionally, non-oxygen atoms such as fluoride, bind to Mg^{2+} ions forming polyatomic clusters as seen in fluoride riboswitches (PDBid: **4enc**) (56). Direct coordination of Mg^{2+} to a cacodylate anion (CAC: $AsO_2(CH_3)_2^-$; PDBid: **4xvf**) has also been reported.

• *Ligand nomenclature:* Currently, our scripts account for most modified nucleotides but do not consider Mg^{2+} binding to ligands. For describing such sites, 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

All *Cat_Wiz* scripts based on the *Coot* 0.9.8.98 version, generate a set of *GtkTables.* The table generated by the *MG_diagnosis* script provides precise stereochemical assessments of structural Mg^{2+}

ions to help distinguishing improperly from properly assigned ions (**Figure 4**). The table involves several tabs to guide users through the refinement process. The first two 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 identify potential 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 meaning 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 (11,14,57). Re-inspection and local refinements are often necessary to validate the assigned ion identity. Color-coding is used to facilitate data analysis: red color highlights the misassigned ions with coordination distances in the 2.6–3.2 Å range which is typical for K^+ (tab1) while green color marks validated octahedral Mg^{2+} ions with CN = 6 (tab2).

• *Binding site quality:* Each Mg^{2+} binding site is evaluated using the *"low-", "fair-" and 'high-quality"* tags, 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 tagged as *"low"* (*grey*) while values in the 5–10% range are considered *"fair"* (*pink*). The map-dependent threshold rmsd values are noted in the *"Map info"* tab5. The rmsd values noted in the tabs are derived from the original PDB map and not from the visualized smoothed map.

• *Distorted coordination octahedron (Dist. oct.) indicator:* This indicator evaluates the distortion of Mg^{2+} coordination octahedra by measuring distances between first-shell atoms which is averaged to 2.9 Å for an ideal Mg^{2+} 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 considered of *"high-quality"* (green), in the 2.4–2.6 Å range it is tagged as of *"fair-quality"* (pink), and of *"low-quality"* when below 2.4 Å (grey; **Figure 5**).

• *Clash indicator:* The Mg^{2+} hexacoordination shell with 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 convoluted binding patterns, distances in the 3.4–3.6 Å range are categorized as mild clashes (pink). Such clashes require either deleting the assigned Mg^{2+} ion, remodeling of the binding site, or changing the ion's identity. Examples of clashes are discussed in the **Results** section.

• *Global quality indicator:* To generate 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 of *"fair-"* or *"high-quality"*. This global indicator is only used for the *"validated"* Mg^{2+} ions in tab2 and is found in the *"Quality"* column. This indicator is not relevant for the non-octahedral coordination patterns noted in tab1.

• *Visualization and remodeling of Mg^{2+} binding sites:* A pop-up window, accessible by a right-mouse click on a tab1 or tab2 line, offers visualization and refinement options (**Figure 4D**). *"Zoom to this ion"* allows to quickly center on an atom and visualize 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* is used to resolve the issue. For missing water molecules, *Coot* offers options to add them to appropriate density patterns. When solvent molecule densities are located at more than 4.3 Å from solute atoms and in the absence of clear octahedral density patterns, it is often difficult to assign these spots to a hydrated $Mg[H_2O]_6^{2+}$ ion or to other solvent particles. A thorough 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 assess if other buffer components 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 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 use of 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 is 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 using the *Coot* command *"add_extra_bond_restraints"* that needs thirteen arguments as listed in the **SI**.

• *Mg^{2+}…Mg^{2+}/K^+ ion pairs:* When two residues are connected to the same Mg^{2+} ion, the "*sphere refine with restraint"* option may induce 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+}/K^+ ion pairs with a maximum inter-ion distance of 6.0 Å. The pairs with *d*(Mg^{2+}… Mg^{2+}/K^+) < 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) need visual inspection to verify if they belong to Mg^{2+} μ-clusters involving at least two Mg^{2+} ions binding to a same phosphate group (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 corrections are 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 enter the 3.2 Å sphere, and

preferably the 3.6 Å sphere since the coordination distance peak identifying second shell atoms starts at around 3.6 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 an appropriate 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 *MG_detect* and *MG_clamp* modules which are ideally run in the order shown in **Figure 1**. *MG_detect* comprises *MG_detect_RNA* and *MG_detect_Prot*. 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 and, therefore, isolate unassigned density spots (**Figure 7**).

MG_detect_RNA and *MG_detect_Prot* inspect density spots close to nucleotide (OP1/OP2/O2/O4/O6/O2'/N7/N3/N1) and protein side chain atoms (OD1/OD2/OE1/OE2/ND1/NE2/N/NZ/NE/NH1/NH2) in addition to protein backbone atoms (O_{bb}). Note that the nucleobase N1/N3 and protein N/NZ/NE/NH1/NH2 atoms are not involved in Mg^{2+} binding but can sometimes associate with transition metals as detailed below. For modified nucleotides and amino acids that involve different atom names, users should manually inspect nearby density spots. 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 in a list generated by the *Coot "map_peaks_around_molecule"* command. In all tabs associated with the OP1/OP2/O2/O4/O6/O2'/N7/N3/N1 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 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 listed as potential inner-sphere Mg^{2+} ion locations.

The lower-ranked peaks (rmsd below 5% of the highest peak value) 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 (58,59). Thus, peaks with rmsd values below 5% of that of the highest peak are not shown and peaks with rmsd in the 5–10% range are shown in pink.

The highest-ranked peaks (rmsd above 10% of the highest peak value) 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**) since it cannot be associated with Mg^{2+}, K^{+} or water. In **9q87**, we tentatively used the "CU" PDB residue identifier that corresponds to copper,

based on the rationale that copper grids were employed and that these ions typically exhibit mostly in-plane coordination with CN = 4. Peaks near nucleobase N3/N1 atoms are not pointing to Mg^{2+} binding sites but could reveal the presence of transition metals or to problematic local modeling.

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 unassigned densities corresponding octahedral or tetrahedral transition metals (Mn^{2+}, Zn^{2+}, etc.) that share Mg^{2+} 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, the presence of unidentified solvent particles or 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. Different cutoffs might be used for an optimal screening. In these instances, *MG_detect* locates densities close to already assigned water molecules with distances below 2.4 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 transition metals while filtering out peaks associated with water and monovalent ions such as K^+ or NH_4^+ that exhibit coordination distances in the 2.6–3.2 Å range. After running *MG_detect,* it is essential to use *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**). For instance, *MG_detect_RNA* 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* 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 backbones. Worth 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 most commonly by bridging two O/N atoms through two direct O/N…Mg^{2+} contacts or more rarely through one direct and one water-mediated contact (**Figure 3**). To date, no examples of clamp motifs stabilized by monovalent ions have been reported (8,30).

• *Tab1 ("Fix Asn/Gln" side chains):* Misoriented protein side chains can occasionally generate artefactual contacts that mimic genuine Mg^{2+} clamps with *d*(O…O) ≤ 3.4 Å. Even after refinement with appropriate options in *Phenix* or *Molprobity* (60,61), some Asn or Gln side chains 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 180°" Coot* tool to improve the model.

• *Tab2 ("Potential clamps with no Mg^{2+} density"):* 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 retained for visual inspection.

• *Tab3 ("Mg^{2+} clamps"):* The third tab identifies and annotates 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 adjustments/corrections are shown in cyan. The latter can be easily remodeled 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 annotated in the *"Comment"* column. Finally, a visual inspection of binding sites and associated experimental density patterns should always be conducted with *MG_diagnosis* before validating an ion position.

• *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.

Updating B-factors at the end of the assignment process

If new assignments or corrections to ion positions are made, *B-factors* need to be updated before PDB submission. This can be done by refining individual *B-factors* in *Phenix* or any other program of choice. Care has to be taken when over sharpened maps are used that can lead to zero or negative *B-factors* at the refinement stage.

RESULTS

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

Analysis of the *E. coli* ribosome at 1.55 Å resolution (PDBid: **8b0x**) serves as the main case study for evaluating *Cat_Wiz*. Previously, we investigated this structure, which contains 361 Mg^{2+} ions (14,19). A comprehensive re-inspection of the structure led to the reassignment and validation of 403 Mg^{2+} ions, a process that was critical for generating the classification of Mg^{2+} binding sites in RNA and RNA/protein systems (14). Besides identifying new binding sites, this study also aimed to correct misassignments by reclassifying some Mg^{2+} as water or K^+ and deleting others. Through *Cat_Wiz*, we re-examined the previously **revised_8b0x** structure. This second revision identified seven additional Mg^{2+} binding sites, while validating the positions of existing Mg^{2+} ions in the amended structure in just a few minutes – a task that otherwise requires days of manual inspection. During this process, we did not identify new binding motifs besides those shown in **Figure 3**. The revised structure that has been deposited in the PDB under the **9q87** identifier revealed significant differences between the ionic shell of the original **8b0x** and the new structure (**Table 1**). In **8b0x,** only 100 Mg^{2+} ions exhibit a hexacoordinated pattern that met our stereochemical criteria. In contrast, the refined **9q87** structure contains 406 validated hexacoordinated patterns.

MG_diagnosis proved invaluable for this revision process. The script flagged 261 ions that required re-evaluation of their identity or their coordination shell. Of those, 42 were identified as potentially misassigned K^+ ions or other solvent particles and 220 ions required coordination-shell adjustments. Among the latter, 85 needed only minor corrections (e.g., coordination distance adjustments), whereas the remaining 135 required a complete remodeling. The ability of *Cat_Wiz* to rapidly diagnose these issues and facilitate their correction significantly streamlined the refinement process, even for users with minimal familiarity with *Coot*.

MG_detect contributed to the completeness of the **amended 8b0x** (14) model by identifying 12 additional Mg^{2+} ions and nine specific Mg^{2+} ions were noted for a few anomalous features (**Table 1**). One puzzling ion, with a coordination shell of seven water molecules and coordination distances close to 2.07 Å was more consistent with a K^+ assignment. Additionally, eight other ions were associated with Mg^{2+}…Mg^{2+} ion pairs and displayed atypical coordination patterns. Approximately 50 *"low-quality"* instances were also identified (**Table 1**). While these sites passed inspection and did not exhibit steric clashes, they are not suitable for inclusion in statistical surveys or AI training sets. Documenting them is, however, important for a complete description of the ribosome solvent shell. Finally, 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. These conformations are 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**). This script flagged 33 problematic distances in **8b0x** that resulted in the flip of at least 16 Asn/Gln side chains. In one instance, we identified a rare $2O_{ph}.O_{coo}.3O_w$ RNA/protein-Mg^{2+} contact involving the small ribosomal subunit, r-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 rather than the more frequently used MRE 600 (62,63) or K12 strains (64), resulting in residue numbering shifts (14). Second, four nucleobases (A:U:677; A:G:741; A:U:1086; a:U:2316) are anionic (14,65) and cannot be annotated with the current PDB residue nomenclature that does not include such "details". Third, a recurrent a:2050:TKW residue (5-hydroxycytidine) has been annotated for the first time in bacterial ribosome structures (66). Fourth, we tentatively included six CU residues that account for Cu^{+} (PDBid: CU1) or Cu^{2+} (PDBid: CU) at positions (A:A_{160}:N7…A:G_{346}:N7; a:A_{149}:N1…a:G_{177}:N7; a:U_{1309}:N7...a:C_{1608}:N3; a:U_{1854}:N3…a:A_{1893}:N7) that correspond to presumed metal-mediated base pairs and between a:G_{1366}:N7 and the N-terminal w:Ser_2 of r-protein bl28. The last CU residue was assigned to a density spot located between v:His_{57}:ND1…v:Cys_{53}:SG of r-protein L27 (**Fig. S1**). The observed coordination geometry of these ions is consistent with that of Cu^{+} (67-70). If present, Cu^{+}/Cu^{2+} ions must be considered as a buffer contaminant probably originating from buffer poisoning by the grid. *MG_detect* does localize these non-Mg^{2+} densities.

Finally, although *Cat_Wiz* identified most major binding sites involving Mg^{2+} clamps in this ribosome (14), it is hard to claim that all more difficult to locate $6O_w$ and $5O_w$ sites were recovered. Noteworthy, diffuse ion binding events are currently beyond the reach of *Cat_Wiz* and most other experimental techniques.

Case II: *Cat Wiz* as an independent tool to investigate metal ion chemistry of hammerhead ribozyme

As a second example, *Cat_Wiz* was used to analyze Mg^{2+} binding sites in hammerhead ribozyme structures deposited to the PDB and listed in **Table 2/S1**. 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 (71), a subsequent X-ray structure (PDBid: **1mme**) reported the position of five Mg^{2+} ions supported by biochemical inferences (72) and subsequent molecular dynamics (MD) studies that proposed catalytic roles for these Mg^{2+} ions (76-80). However, later research challenged these early findings, suggesting that Mg^{2+} ions are not directly involved in the catalytic mechanism (73-75). Up to now, structural studies including X-ray crystallography have 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 hypothesis 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 (75,81). Consequently, the hammerhead ribozyme's classification has shifted from a metalloenzyme to a more consensual acid-base catalysed ribozyme.

A challenge in analyzing these X-ray 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/S1 and Figure 9**). Indeed, some of the originally assigned Mg^{2+} ions are stereochemically inconsistent with Mg^{2+} binding and correspond more likely to SO_4^{2-} or cacodylate anions (82,83) which could have been detected through their strong anomalous X-ray diffraction signals (84,85).

To test this, we analyzed the structures mentioned in **Table 2** using *Cat_Wiz*. The **1hmh, 1mme,** and **299d** structures lack assigned Mg^{2+} ions although related papers discussed ion dependent catalytic

mechanisms (71,72,86). **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 be assigned to SO_4^{2-} or cacodylate ions (**Figure 9**). A fifth ion, located 2.4 Å from A:11:OP2 (**Figure 11A**), 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. For all those structures, *Cat Wiz* was unable to propose any strong Mg^{2+} binding site (75,81).

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* identified only one Mg^{2+} with a 2.18 Å distance and sufficient space to accommodate a hexahydrated shell (**Figure 11B**). However, no clear octahedral densities for the water molecules were observed as in the 2.0 Å resolution **2oeu** Mn^{2+}-containing structure (**Table S1**). Five of the remaining metal densities in **2quw** could be assigned to hydrated Mg^{2+} ions or to anions. On the other hand, the other ten assigned ions are poorly placed and sterically incompatible with Mg^{2+} binding according to *Cat Wiz.* Placing a hydrated Mg^{2+} ion at these locations would induce severe clashes with neighboring residues (**Figure 11C**).

The absence of clear Mg^{2+} binding sites, 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 be more informative regarding their catalytic mechanisms (74,87,88). For now, existing hammerhead structures highlight the difficulties of understanding the solvation shell of low-resolution systems containing stereochemically unsupported Mg^{2+} assignments which can at best foster further discussions and interpretations.

Another set of hammerhead structures containing questionable Mg^{2+} assignments include **5dh6**, **5dqk**, and **5eao**. In **5dh6,** the proposed OP2:103 Mg binding difficultly supports specific Mg^{2+} binding as $d(Mg^{2+}\ldots O)$ ≈2.48 Å is longer than a typical Mg^{2+}...O distance (**Figure 11D**). Similarly, **5dqk** and **5eao** show densities close to 2.45 Å with no octahedral geometry (**Figure 11E**), suggesting that these densities could be better assigned to other solvent particles. Moreover, the high number of poorly assigned Mg^{2+} ions in these structures significantly reduces confidence 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**).

We envision that early stereochemical assessments of these Mg^{2+} binding sites would have provided a clearer picture for exploring the ribozyme catalytic mechanism. This also underscores the need for an in-depth examination of all available structures of a given system and for checking whether ion assignments obey all basic stereochemical rules. In addition, a clear understanding of the buffer conditions used in experiments may prevent some misassignments. *Cat_Wiz* proved highly useful for re-interpreting these hammerhead systems.

Case III: group I *Tetrahymena* ribozyme structures

Cat_Wiz was also used to reinspect the solvation shell of different *Tetrahymena* group I ribozyme structures. The first example 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 (89). Its analysis has led to the description of a well-defined Mg^{2+} core with 42 assigned divalent ions (90).

Further, 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, respectively, was analyzed. These structures where Mg^{2+} and water molecules were automatically assigned using a segmentation-guided water and ion modelling protocol called SWIM (22,59,91) display slightly different P9.2 stem configurations. Companion structures (**9cbu** and **9cbw**) which only contain consensus solvent molecules common to **9cbx** and **9cby** were also deposited to the PDB.

• *The P4-P6 mutant solvation shell:* The **1hr2** X-ray structure of this 160-nucleotide group I intron fragment comprises two biological assemblies (89). Chain "A" contains 12 hexacoordinated Mg^{2+} ions meeting our stereochemical criteria and 12 that do not (**Table 3**). Of the former, only four have a complete hydration shell and establish direct contacts to the RNA. One of these binding sites 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. In addition, eight other hexahydrated Mg^{2+} ions 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 $Mg[H_2O]_6^{2+}$ ions, cacodylate anions, or HEPES fragments based solely on the available density maps. The *Cat_Wiz* analysis confirmed the Mg^{2+} placements for the hydrated ions, while the *MG_detect* tool did not identify any 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 modelled, even though coordinated water molecule densities are apparent for six of these ions. Thus, the positions of these ions could not be 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 with no visible octahedral densities. Modeling a hydration shell at these locations result in steric clashes with neighboring residues (**Figure 12F**) suggesting they these sites are not *bona fide* Mg^{2+} binding sites.

Regarding non-consensus ions, we found that the ion modeled in **Figures 12C/D/E** is missing in all four parent structures. Furthermore, none of the 32 additionally 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 induces 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 Mg^{2+} ions to fold into its active conformation. Overall, this analysis demonstrates how *Cat_Wiz* provides support to inspect the solvation shells of moderate resolution structures and can sort 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 obtained with K^+ instead of Na^+ containing buffers as in most of the deposited structures.

Case IV: RNA quaternary structures

• *Exploring metal binding sites in RNA scaffold structures:* Cryo-EM techniques have recently opened a new avenue for investigating RNA quaternary assemblies (34). Metal ions were assigned in two such structures: a *Lactobacillus salivarius* ROOL hexamer (PDBid: **9j6y**) and a *Clostridium botulinum* OLE dimer (PDBid: **9lcr**), resolved at 3.25 and 2.6 Å, respectively. Given these low resolutions, no metal hydration shells were modeled. Noteworthy, in the publication describing these quaternary structures (34), the assigned ions were named "Metals" or (M), except in the attached supplementary material. In the corresponding PDB file, some of these ions are wrongly identified as "Mg" highlighting how automatic database surveys can be compromised by annotations made without proper consideration of stereochemical criteria. We used these structures to evaluate how *Cat_Wiz* performs with low-resolution models and locates unassigned Mg^{2+} binding sites and clamps. We further investigated the 1.96 Å high-resolution structure of a *Ligi lactobacillus salivarius* ROOL monomer (PDBid: **9m79**) for which no ions or waters were assigned despite the better resolution(92).

• *The **9j6y** Lactobacillus salivarius ROOL hexamer (3.25 Å):* In this structure, 56 non-hydrated Mg^{2+} ions were assigned. Of those, 18 form contacts with the RNA in the 2.6–3.2 Å range with coordination distances specific 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* identified $d(O_{ph}\ldots O_{ph})$ distances of approximately 3.8 Å suggestive of K^+ rather than Mg^{2+} binding (**Figure 13A**). The remaining 38 ions located at more than 3.2 Å from RNA atoms are more difficult to assign without additional experimental evidence. They could relate 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 13B**).

• *The **9lcr** Clostridium botulinum OLE dimer (2.6 Å):* For this dimer, no hydration shells were modelled around the 28 assigned metals (14 per monomer). Among the 14 ions in monomer "A", a valid *mer*-$3O_{ph}.3O_w$ coordination (M2 site) was identified. Here, 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 unclear octahedral metal sites could 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 (3,11,93).

Given the limited resolution of this structure, we could only remodel a few of the assigned metals. 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 optimize their positions. In the same maner, 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 satisfactory hydration sphere model, validating the presence of octahedral Mg^{2+} ions. 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.

• ***9m79*** *Ligi lactobacillus salivarius ROOL monomer (1.96 Å):* Despite its higher resolution, no ions or water molecules were modeled (92). Yet, *MG_clamp* identifies at least three *cis*-$2O_{ph}.4O_w$ clamps, one *cis*-$O_{ph}.N_b.4O_w$ and one relatively well-resolved *fac*-$3O_{ph}.3O_w$ binding site, demonstrating the usefulness of these ion detection techniques.

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

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 incomplete hydration shells are seen (11,12,14,94). Such errors and incompleteness lead to inaccuracies in mechanistic interpretations of structural models, deceiving functional hypotheses, and introduce noise into databases used for comparative analyses or machine-learning workflows (10,20,21,41-43,95).

Cat_Wiz addresses these problems by combining stereochemical binding principles into a pipeline. By enforcing established coordination distances and geometries derived from the CSD and from a 1.55 Å resolution structure of an *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, group I introns, and RNA quaternary structures to illustrate how *Cat_Wiz* works to recover incomplete ionic hydration shells, correct misassigned cations, and annotate ion motifs such as Mg^{2+} clamps. These corrections can be completed in hours rather than days of manual examination. *Cat_Wiz* can be used with low- to high-resolution X-ray and cryo-EM structures and can be used for nucleic acids as well as proteins.

While *Cat_Wiz* aims at providing insights into RNA structural biology it also shares limitations with other approaches. At resolutions poorer than ≈2.5 Å, ion coordination becomes difficult to distinguish and errors in RNA backbone modeling might propagate. 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 generate chemically acceptable structural models (96-98).

Buffer composition presents a second challenge. Non-physiological intracellular cations (99-101) such as NH_4^+, Na^+, and anions such as cacodylate and sulphate (82,83) are often present in experimental buffers, complicating experimental map interpretations due to coordination distance overlaps. *Cat_Wiz* analyses and highlights cases where these ions have been misinterpreted as for instance Mg^{2+} in hammerhead ribozymes. These findings strengthen the importance of using simplified and physiologically relevant buffers containing primarily Mg^{2+} and K^+, to warrant 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 (19,102). In the future, more systematic inclusion of polyamines will represent an important addition to experimental buffers and advance our understanding of the functions of biological solvation shells (62,103-106). Although chloride (Cl^-) is the most common intracellular anion (3,82,83), it is often absent from recent experimental buffers and replaced by other anionic components including acetate, malonate, HEPES or TRIS molecules. These latter substitutes may perturb to a yet unknown extent the RNA solvation shell.

A third concern lies in the deposition and annotation of structural data (15,18,107,108). Misassigned ions without proper annotations persist in the PDB, despite repeated error identifications (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 also have the potential to misguide automated surveys and affect machine-learning training sets. Herein, we documented how these issues affect hammerhead and recent group I intron ribozyme structures potentially jeopardizing collective efforts to understand and model the ionic structure and catalytic properties of these systems (22). Wider use of annotation tools such as *Cat_Wiz*, together with PDB-level flagging, would significantly 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 specific "AP7" and "CH" codes are used for protonated A(+) and C(+), respectively (14,65). A fix by the PDB would drive progress in understanding and annotating 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 at automatically resolving them since visual inspection is key to understand and describe such undocumented features. Yet, even visual inspection assisted by *Cat_Wiz* in less-well resolved regions leaves us sometimes with more interrogations than answers.

Finally, while machine-learning ion identification approaches are becoming popular (20,109,110), current methods struggle to correct poor assignments and stereochemical errors present in existing models. MD simulations that lack polarization and charge transfer effects should be interpreted with caution when modelling Mg^{2+} ions (22,111,112). While anticipating to extract more information from high-resolution structures (113-117), *Cat_Wiz* provides a crucial bridge between experimental data and modelling. By incorporating stereochemical principles and producing high-quality ion assignments, it supplies more accurate training sets to improve the accuracy and reliability of next generation AI tools and molecular dynamics force fields.

To summarize, *Cat_Wiz* offers a rapid, and stereochemistry-guided framework for diagnosing, annotating and correcting Mg^{2+} assignments. By improving the structural fidelity of biomolecular solvation shells and generating curated datasets, *Cat_Wiz* benefits both experimentalists seeking for accurate solvation models, interpretation of catalytic mechanisms and computational scientists developing predictive algorithms to study ion-RNA interactions. Its integration into *Coot* makes it accessible to a wide community by positioning *Cat_Wiz* as a practical, fast, and transparent tool supplementing both manual refinement and emerging autonomous structural biology approaches.

Disclaimer

We did our best in designing *Cat_Wiz* and in assigning Mg^{2+} ions in **8b0x/9q87**. The quality of these assignments is strongly dependent on the resolution and preceding refinement steps as well as on the prior knowledge at disposal at the time of the structure deposition used to disentangle complex and unforeseen solvation patterns. This is an ongoing process and we do not claim: *i)* that we assigned all Mg^{2+} ions, *ii)* that some of the more complex or less-well resolved solvation patterns did not mislead us, and *iii)* that all reported patterns were accurately modelled. However, we used the best available knowledge and structures. We are confident that we located all major Mg^{2+} binding sites in the **8b0x**

ribosomal structure and all-important binding types in most relevant RNA systems we investigated allowing for a close to complete annotation system.

AVAILABILITY

The scripts are deposited to Zenodo as password protected files at the following address: https://zenodo.org/records/18144017.

ACCESSION NUMBERS

Atomic coordinates and structure factors for the reported crystal structures have been deposited to the PDB under accession number **9q87** (version 3).

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 Emsley 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 copper 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. Wang, L., Xie, J., Gong, T., Wu, H., Tu, Y., Peng, X., Shang, S., Jia, X., Ma, H., Zou, J. *et al.* (2025) Cryo-EM reveals mechanisms of natural RNA multivalency. *Science*, **388**, 545-50.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. 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.
41. 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.
42. Majorek, K.A., Gucwa, M., Murzyn, K. and Minor, W. (2024) Metal ions in biomedically relevant macromolecular structures. *Front. Chem.*, **12**, 1426211.
43. 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.
44. 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.
45. Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst.*, **D67**, 355-67.
46. 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 Cryst.*, **D75**, 861-77.
47. 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.

48. 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.
49. 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.
50. 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**.
51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.
57. 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.
58. Wlodawer, A., Li, M. and Dauter, Z. (2017) High-resolution cryo-EM maps and models: A crystallographer's perspective. *Structure*, **25**, 1589-97 e1.
59. Pintilie, G. and Chiu, W. (2021) Validation, analysis and annotation of cryo-EM structures. *Acta Crystallogr. D Struct. Biol.*, **77**, 1142-52.
60. 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.
61. 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 Cryst.*, **D74**, 132-42.
62. 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.
63. 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.
64. 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.
65. 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.
66. 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.
67. 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.
68. 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.
69. Iwamura, M., Takeuchi, S. and Tahara, T. (2015) Ultrafast excited-state dynamics of copper(I) complexes. *Acc. Chem. Res.*, **48**, 782-91.
70. 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.

71. Pley, H.M., Flaherty, K.M. and McKay, D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68-74.
72. 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.
73. Leclerc, F. (2010) Hammerhead ribozymes: true metal or nucleobase catalysis? Where is the catalytic power from? *Molecules*, **15**, 5389-407.
74. Herschlag, D. (2015) Learning from ribozymes. *RNA*, **21**, 527-8.
75. 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.
76. Hermann, T., Auffinger, P. and Westhof, E. (1998) Molecular dynamics investigations of the hammerhead ribozyme RNA. *Eur. J. Biophys.*, **27**, 153-65.
77. 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.
78. 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.
79. 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.
80. 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.
81. 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.
82. Auffinger, P., Bielecki, L. and Westhof, E. (2004) Anion binding to nucleic acids. *Structure*, **12**, 379-88.
83. D'Ascenzo, L. and Auffinger, P. (2016) Anions in nucleic acid crystallography. *Methods Mol. Biol.*, **1320**, 337-51.
84. 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.
85. 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.
86. 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.
87. Hammann, C. and Westhof, E. (2009) The unforeseeable hammerhead ribozyme. *F1000 Biol. Rep.*, **1**, 6.
88. Hammann, C., Luptak, A., Perreault, J. and de la Pena, M. (2012) The ubiquitous hammerhead ribozyme. *RNA*, **18**, 871-85.
89. 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.
90. 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.
91. 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.
92. Zhang, S., Yi, R., An, L., Liu, J., Yao, X., Li, S. and Zhang, K. (2025) Structural insights into higher-order natural RNA-only multimers. *Nat. Struct. Mol. Biol.*, **32**, 2012-21.
93. Westhof, E., Yusupov, M. and Yusupova, G. (2019) The multiple flavors of GoU pairs in RNA. *J. Mol. Recognit.*, **32**, e2782.
94. Auffinger, P., Ennifar, E. and D'Ascenzo, L. (2021) Deflating the RNA Mg^{2+} bubble. Stereochemistry to the rescue! *RNA*, **27**, 242-52.
95. Yu, T. and Chen, S.J. (2018) Hexahydrated Mg(2+) Binding and Outer-Shell Dehydration on RNA Surface. *Biophys J*, **114**, 1274-84.
96. Wlodawer, A. (2017) Stereochemistry and validation of macromolecular structures. *Methods Mol. Biol.*, **1607**, 595-610.

97. 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.
98. 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.
99. Auffinger, P., Grover, N. and Westhof, E. (2011) Metal ion binding to RNA. *Met. Ions Life Sci.*, **9**, 1-35.
100. Auffinger, P., D'Ascenzo, L. and Ennifar, E. (2016) Sodium and potassium interactions with nucleic acids. *Met. Ions Life Sci.*, **16**, 167-201.
101. 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.
102. 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.
103. Miller-Fleming, L., Olin-Sandoval, V., Campbell, K. and Ralser, M. (2015) Remaining mysteries of molecular biology: The role of polyamines in the cell. *J. Mol. Biol.*, **427**, 3389-406.
104. Michael, A.J. (2016) Polyamines in Eukaryotes, Bacteria, and Archaea. *J Biol Chem*, **291**, 14896-903.
105. Gevrekci, A.O. (2017) The roles of polyamines in microorganisms. *World J. Microbiol. Biotechnol.*, **33**, 204.
106. 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.
107. 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.
108. 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.
109. 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.
110. 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.
111. 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.
112. 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.
113. 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.
114. 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.
115. Yonekura, K. and Maki-Yonekura, S. (2016) Refinement of cryo-EM structures using scattering factors of charged atoms. *J. Appl. Cryst.*, **49**, 1517-23.
116. Marques, M.A., Purdy, M.D. and Yeager, M. (2019) CryoEM maps are full of potential. *Curr. Opin. Struct. Biol.*, **58**, 214-23.
117. Bick, T., Dominiak, P.M. and Wendler, P. (2024) Exploiting the full potential of cryo-EM maps. *BBA Adv.*, **5**, 100113.
118. 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.
119. 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.
120. 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. *"High", "fair"* and "*low"* quality binding site occurrences are provided in parenthesis.

	8b0x		**9q87**	
Outer-shell				
$6O_w$	**21** (17/1/4)		**118** (48/53/17)	
Pentahydrated				
$O_{ph}.5O_w$	**38** (28/7/3)		**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-*
$2O_{ph}.4O_w$	**12** (9/2/1)	**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:	**99** (76/12/11)		**406** (197/150/60)	
With issues (not in table)**:**	**262**		**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 associated buffers contain cacodylate or CAC, $AsO_2(CH_3)_2^-$ anions. Structures with no assigned divalent ions, despite the presence of Mg^{2+} in the buffer, are shown in grey. A complete list of all 26 hammerhead structures is provided in **Table S1**.

	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;	(71)
2.	1mme	1996	3.1	–	25+16	100mM NH_4OAc; 10mM $Mg(OAc)_2$; CAC;	(72)
3.	299d	1997	3.0	–	25+16	1.8M Li_2SO_4; 10mM $MgSO_4$; CAC;	(86)
4.	301d	1997	3.0	5 Mg^{2+}	25+16	1.8M Li_2SO_4; 10mM $MgSO_4$; CAC;	(86)
5.	2quw	2008	2.2	17 Mg^{2+}	57+12	28mM $MgCl_2$, CAC;	(118)
6.	2qus	2008	2.4	16 Mg^{2+}	69	28mM $MgCl_2$, CAC;	(118)
7.	3zp8	2013	1.55	16 Na^+	43+20	1.7M Na-malonate; 10mM $MgCl_2$;	(119)
8.	5dh6	2015	2.78	5 Mg^{2+}	48+20	0.9 M KCl; 10mM $MgCl_2$; CAC;	(120)
9.	5dqk	2016	2.71	1 Mg^{2+} + 4 K^+	48+20	50mM $Mg(OAc)_2$; CAC;	(120)
10.	5eao	2016	2.99	8 Mg^{2+}	48+20	0.4M KCl; 50mM $Mg(OAc)_2$; CAC;	(120)

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 2.3–2.6 Å range. In the **1hr2** structure (89), the 10 chain "A" ions (Mg:51 to Mg:64) with no octahedral densities are distant from RNA and at best 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 in the 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.*

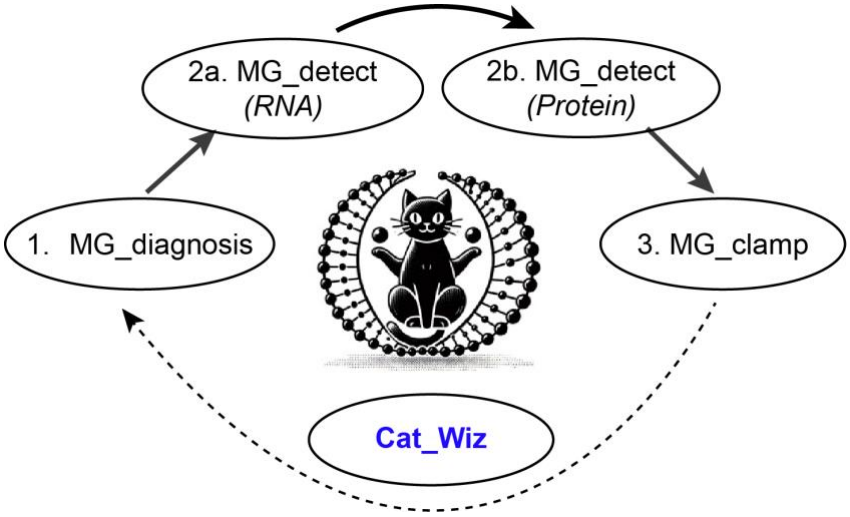


Figure 1. *Cat_Wiz* workflow. The diagram illustrates the recommended sequence for using the four independent Mg^{2+} related *Cat_Wiz* scripts. It is advised to re-run the *MG_diagnosis* script whenever structural changes were made to ensure the accuracy of the final model.

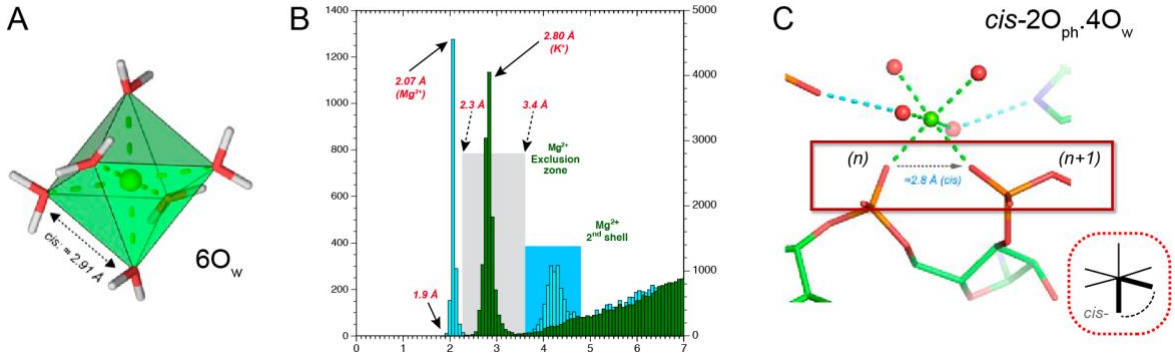


Figure 2. Stereochemical features of Mg^{2+} ions and Mg^{2+} clamps. (**A**) A hexahydrated Mg^{2+} ion ($6O_w$) from a high-resolution *Cambridge Structural Database* (CSD) entry, showing characteristic *d*(O_w…O_w) distances around 2.9 Å (14). (**B**) Histograms of Mg^{2+} (cyan) and K^+ (green) to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

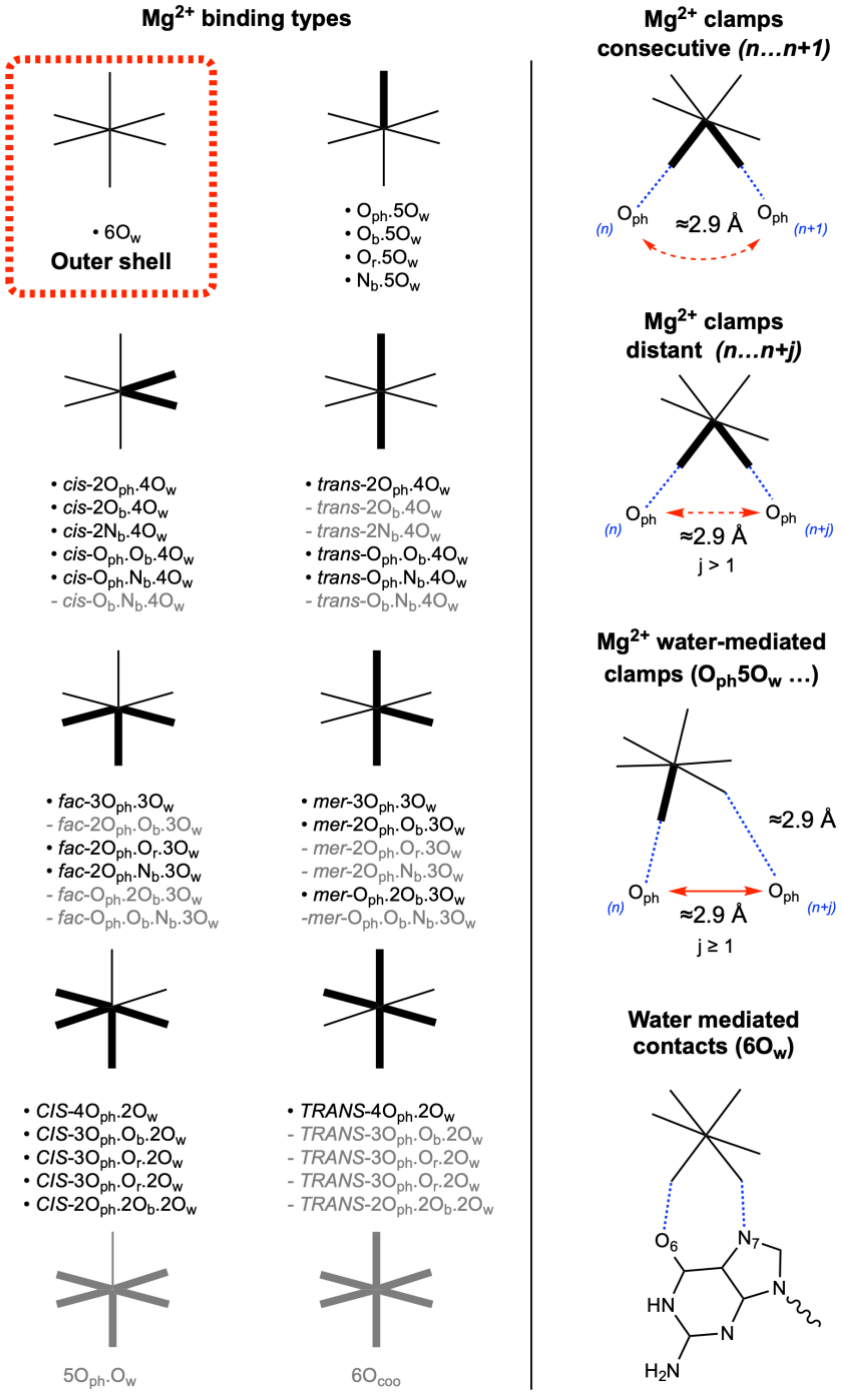


Figure 3. Updated view of Mg^{2+} binding motifs. (Left) This panel presents all inner- and outer-shell Mg^{2+} binding motifs identified in **9q87** and is an update of the list derived from **8b0x** (14). For clarity, motifs involving ribosomal proteins have been excluded. Greyed out binding types are until now associated with 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 O4'/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 through 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.

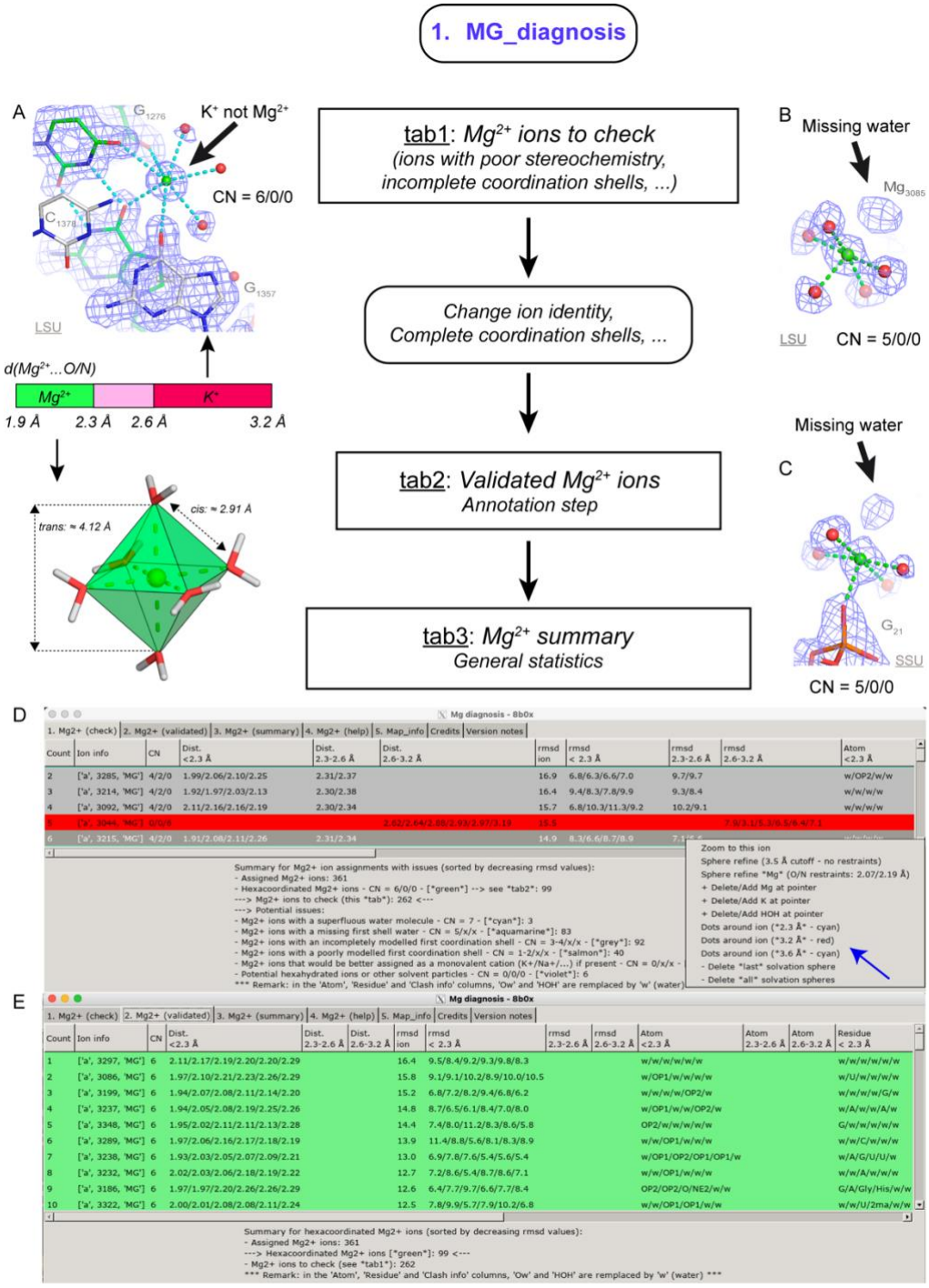


Figure 4. Workflow of the *MG_diagnosis* script. The central column illustrates the *MG_diagnosis* primary functions, i.e., *i)* assessing the stereochemical quality of assigned ions and *ii)* recommending changes based on specific coordination number (CN) and distance criteria. These examples are taken from the **8b0x** structure. (**A**) A color-coded scale is used to evaluate coordination distances. Green corresponds to an ideal 1.9–2.3 Å range for Mg^{2+}, 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-left 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 exposed by the density map. (**D**) Partial view of the *MG_diagnosis* tab1, which lists all identified issues for a given structure. A pop-up menu (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).

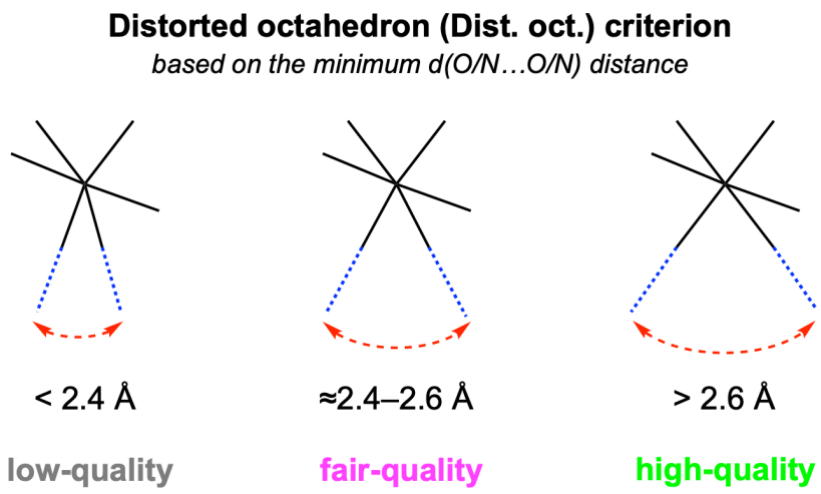


Figure 5. The distorted octahedron (Dist. oct.) criterion. The quality of a binding site is *"low"*, *"fair",* or *"high"* based on the smallest distance between any two first shell coordinating atoms. This criterion is exclusively applied in tab2 of *MG_diagnosis*. These distance cutoffs were empirically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

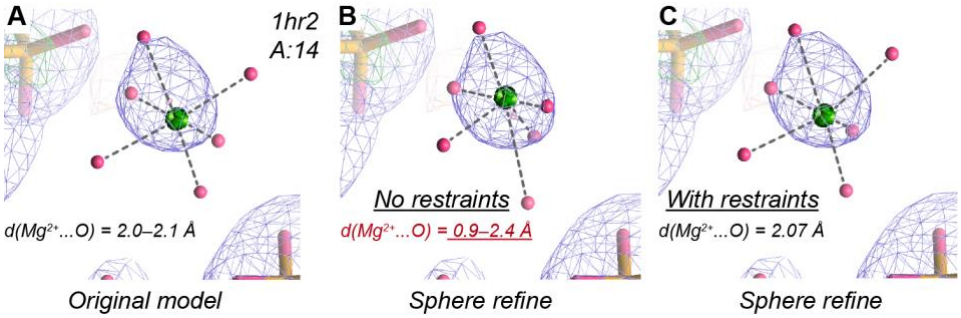


Figure 6. Effect of restraints on the *Coot "Sphere refine"* option. Different outcomes are observed when using the *Coot "Sphere refine"* command with and without distance restraints. (**A**) A hexahydrated Mg^{2+} ion modelled into a non-octahedral density pattern (89). (**B**) Refining the same position using the *"sphere_refine"* command without added restraints 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.

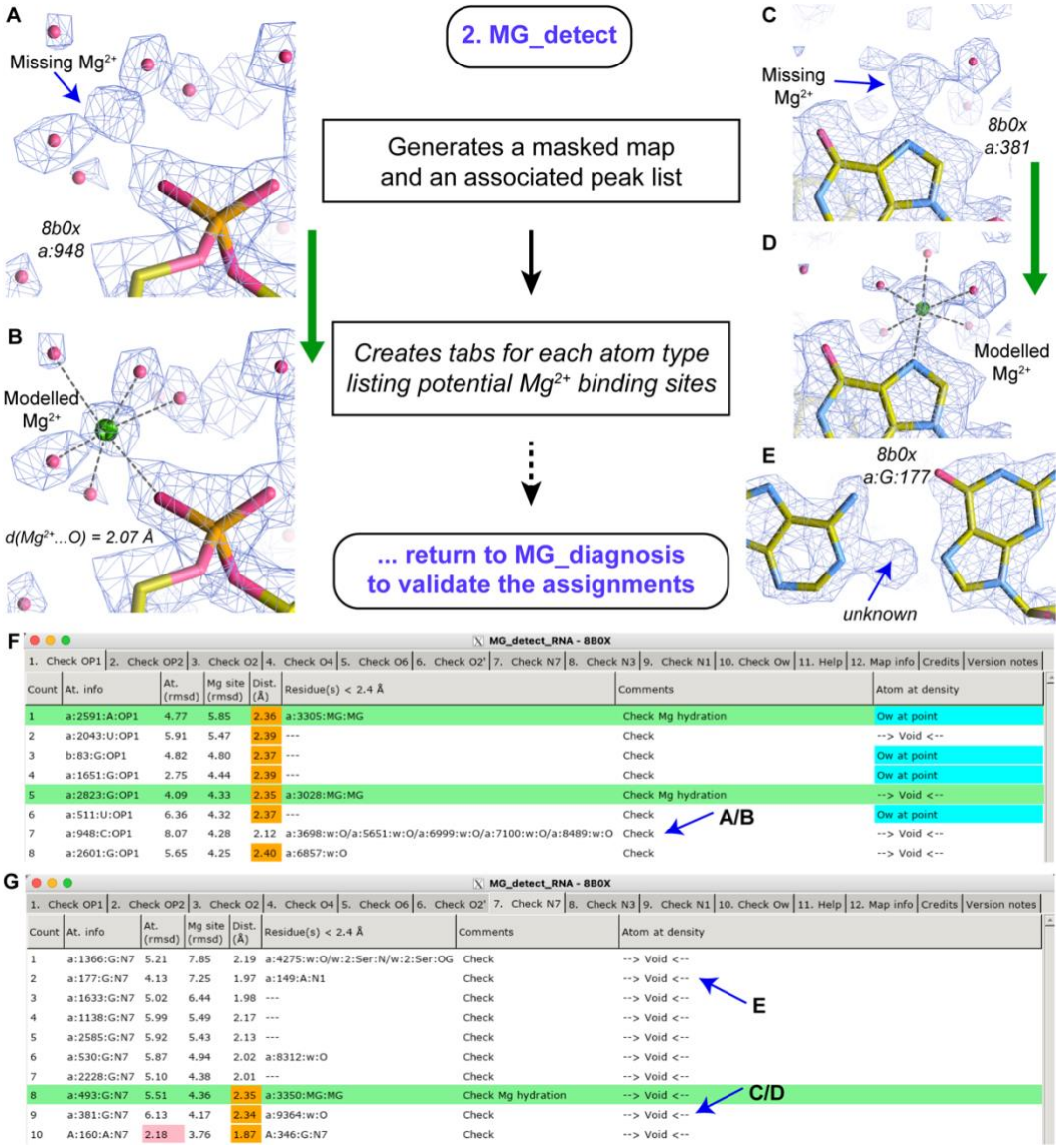


Figure 7. Organization of the *MG_detect_RNA* script. *MG_detect_RNA* identifies potential Mg^{2+} binding sites by collecting the highest non-assigned peaks from a masked map. (**A**) Shows a missing Mg^{2+} ion that is part of an O_{ph}.5O_w binding site successfully identified by *MG_detect_RNA*. (**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 A…G pair as described in (14). (**F/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). Orange cells mark distances from the identified density peak to the selected atom that are in the 2.3–2.4 Å range or below 1.9 Å; pink cells mark low rmsd values.

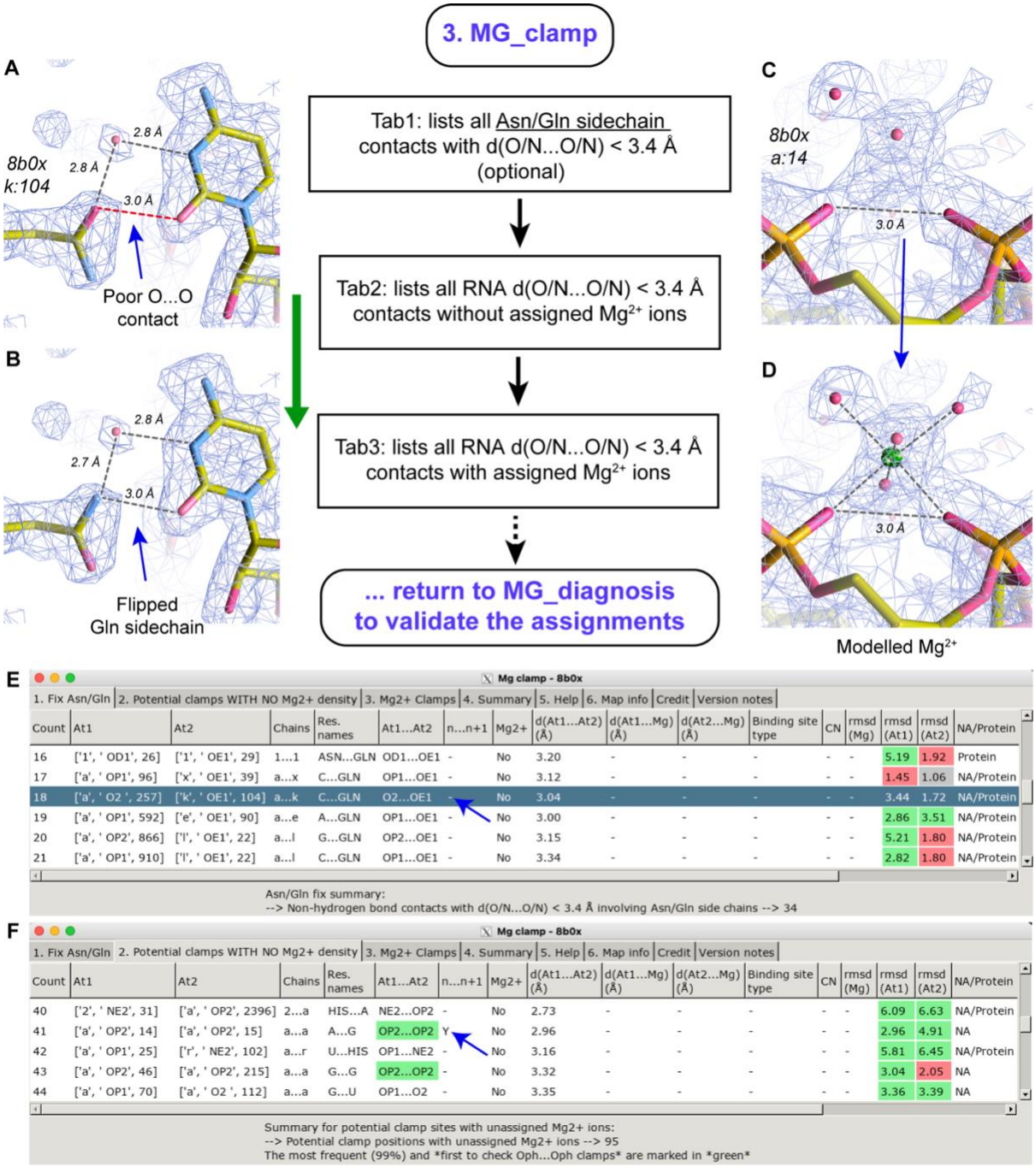


Figure 8. Organization of the *MG_clamp* script. *MG_clamp* locates non-hydrogen-bonded pairs of atoms with *d*(O/N…O/N) below 3.4 Å. These pairs are common to all higher-binding motifs (except *trans*-$2O_{ph}.4O_w$ and related; see **Figure 3**). When such pairs are identified, a visual inspection 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**).

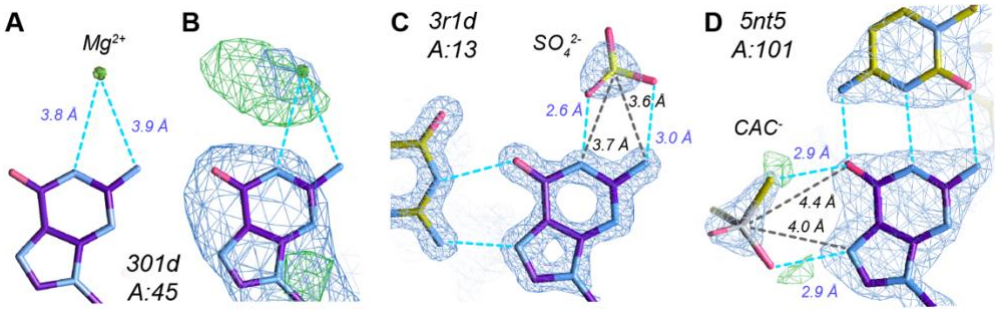


Figure 9. Anionic alternatives to questionable Mg^{2+} assignments. This figure presents examples of alternative explanations to questionable Mg^{2+} assignments. (**A/B**) In the **301d** hammerhead structure, the Mg^{2+} ion (Mg:45) is unconvincingly positioned near an electropositive Watson-Crick edge (82,83). 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 modelled in the 2.30 Å resolution structure **5nt5** (**D**). Panel **B** shows *fo-fc* density patterns (3.0 σ; green) suggesting 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.

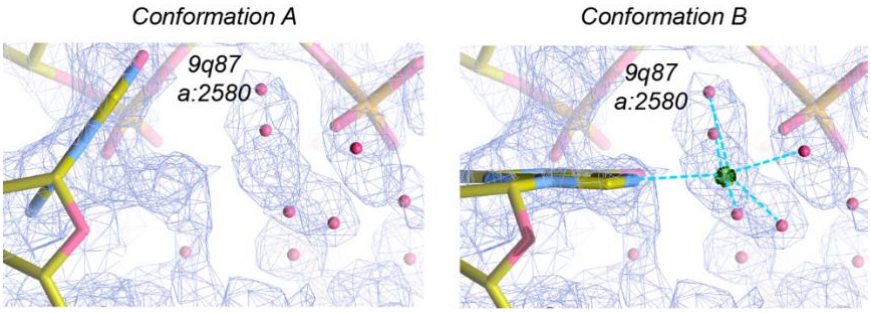


Figure 10. Alternative conformation of the a:2580 residue in 9q87. Conformation A shows the favoured orientation of a:G:2580. Conformation B shows an alternate conformation involving a well-defined N7.5O_w contact with Mg^{2+}.

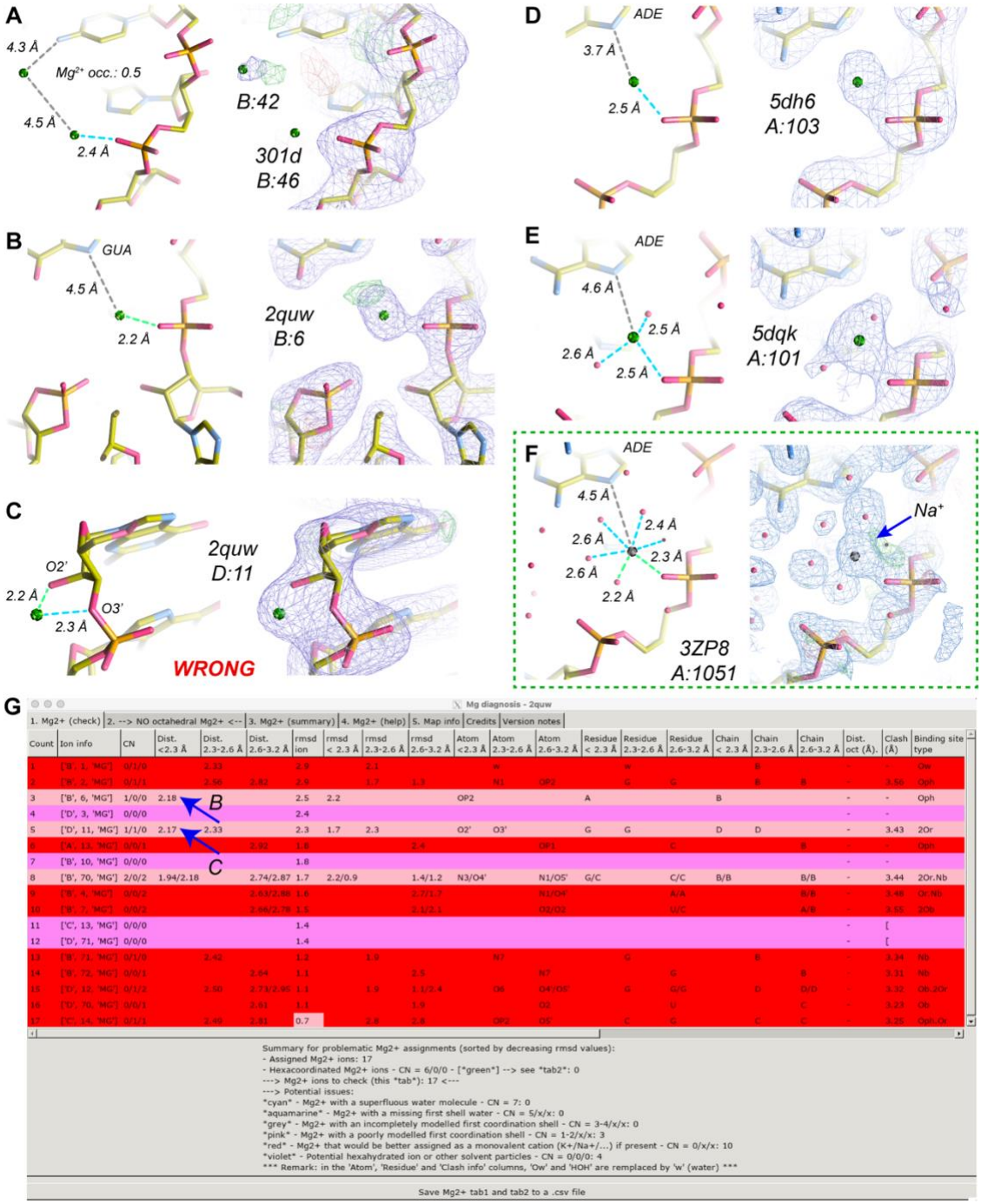
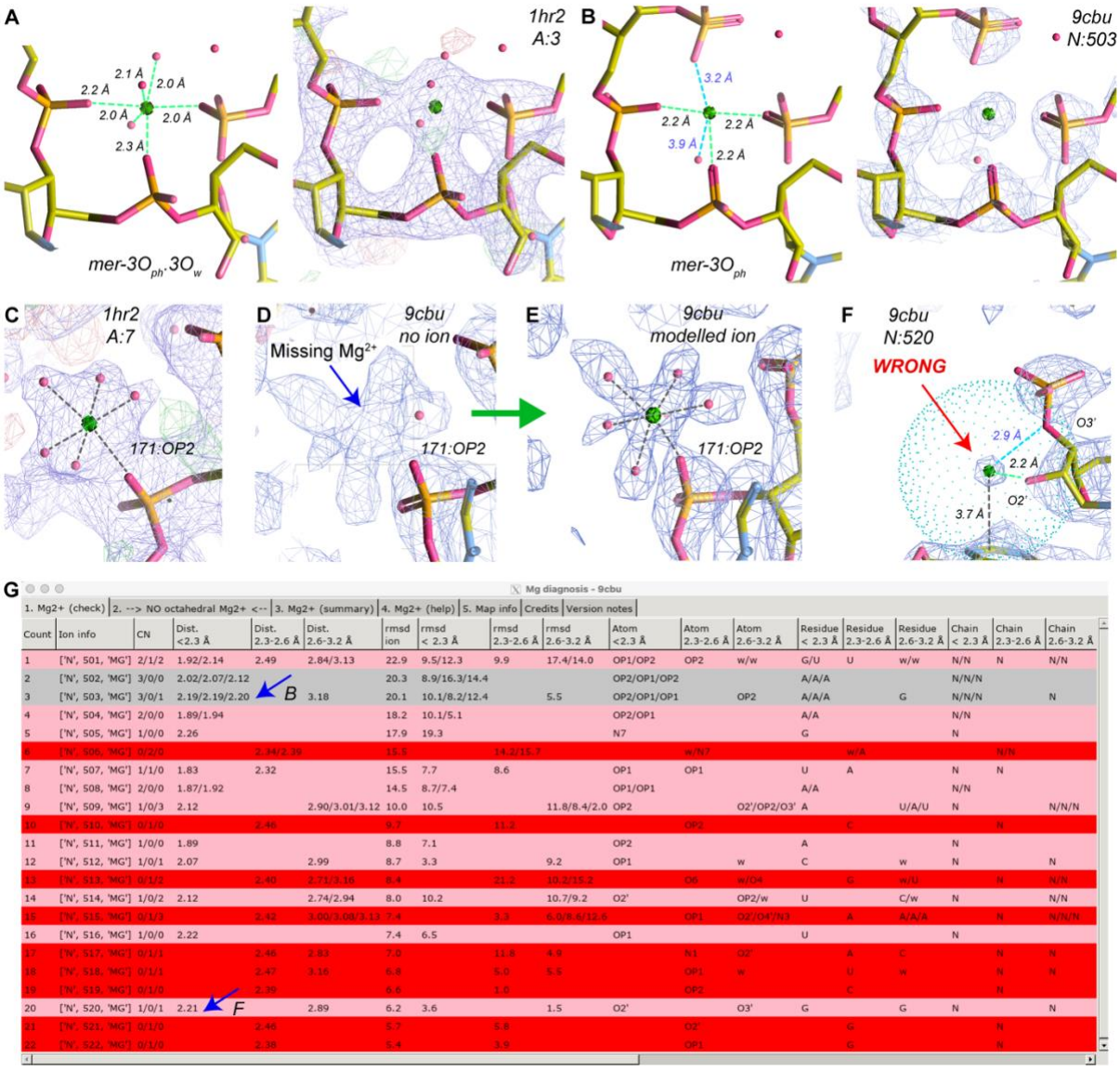


Figure 11. Ion binding assignments in hammerhead ribozyme X-ray structures. (**A**) Mg^{2+} ion tentatively assigned to a weak density spot in the 3.0 Å resolution **301d** structure (occupancy = 0.5). (**B**) A more realistic Mg^{2+} ion from the 2.2 Å resolution **2quw** structure, although lacking octahedral density. (**C**) A stereochemically improbable Mg^{2+} ion. (**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**).



Count	Ion info	CN	Dist. <2.3 Å	Dist. 2.3-2.6 Å	Dist. 2.6-3.2 Å	rmsd ion	rmsd < 2.3 Å	rmsd 2.3-2.6 Å	rmsd 2.6-3.2 Å	Atom <2.3 Å	Atom 2.3-2.6 Å	Atom 2.6-3.2 Å	Residue < 2.3 Å	Residue 2.3-2.6 Å	Residue 2.6-3.2 Å	Chain < 2.3 Å	Chain 2.3-2.6 Å	Chain 2.6-3.2 Å
1	['N', 501, 'MG']	2/1/2	1.92/2.14	2.49	2.84/3.13	22.9	9.5/12.3	9.9	17.4/14.0	OP1/OP2	OP2	w/w	G/U	U	w/w	N/N	N	N/N
2	['N', 502, 'MG']	3/0/0	2.02/2.07/2.12			20.3	8.9/16.3/14.4			OP2/OP1/OP2			A/A/A			N/N/N		
3	['N', 503, 'MG']	3/0/1	2.19/2.19/2.20		3.18	20.1	10.1/8.2/12.4		5.5	OP2/OP1/OP1		OP2	A/A/A		G	N/N/N		N
4	['N', 504, 'MG']	2/0/0	1.89/1.94			18.2	10.1/5.1			OP2/OP1			A/A			N/N		
5	['N', 505, 'MG']	1/0/0	2.26			17.9	19.3			N7			G			N		
6	['N', 506, 'MG']	0/2/0		2.34/2.39		15.5		14.2/15.7			w/N7			w/A			N/N	
7	['N', 507, 'MG']	1/1/0	1.83	2.32		15.5	7.7	8.6		OP1	OP1		U	A		N	N	
8	['N', 508, 'MG']	2/0/0	1.87/1.92			14.5	8.7/7.4			OP1/OP1			A/A			N/N		
9	['N', 509, 'MG']	1/0/3	2.12		2.90/3.01/3.12	10.0	10.5		11.8/8.4/2.0	OP2		O2'/OP2/O3'	A		U/A/U	N		N/N/N
10	['N', 510, 'MG']	0/1/0		2.46		9.7		11.2			OP2			C			N	
11	['N', 511, 'MG']	1/0/0	1.89			8.8	7.1			OP2			A			N		
12	['N', 512, 'MG']	1/0/1	2.07		2.99	8.7	3.3		9.2	OP1		w	C		w	N		N
13	['N', 513, 'MG']	0/1/2		2.40	2.71/3.16	8.4		21.2	10.2/15.2		O6	w/O4		G	w/U		N	N/N
14	['N', 514, 'MG']	1/0/2	2.12		2.74/2.94	8.0	10.2		10.7/9.2	O2'		OP2/w	U		C/w	N		N/N
15	['N', 515, 'MG']	0/1/3		2.42	3.00/3.08/3.13	7.4		3.3	6.0/8.6/12.6		OP1	O2'/O4'/N3		A	A/A/A		N	N/N/N
16	['N', 516, 'MG']	1/0/0	2.22			7.4	6.5			OP1			U			N		
17	['N', 517, 'MG']	0/1/1		2.46	2.83	7.0		11.8	4.9		N1	O2'		A	C		N	N
18	['N', 518, 'MG']	0/1/1		2.47	3.16	6.8		5.0	5.5		OP1	w		U	w		N	N
19	['N', 519, 'MG']	0/1/0		2.39		6.6		1.0			OP2			C			N	
20	['N', 520, 'MG']	1/0/1	2.21		2.89	6.2	3.6		1.5	O2'		O3'	G		G	N		N
21	['N', 521, 'MG']	0/1/0		2.46		5.7		5.8			O2'			G			N	
22	['N', 522, 'MG']	0/1/0		2.38		5.4		3.9			OP1			G			N	

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 (89), 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 remains unoccupied despite clear octahedral density (**D**) but can be modelled as O_{ph}.5O_w using *MG_diagnosis* (**E**). (**F**) A 3.6 Å radius sphere, generated by *MG_diagnosis*, emphasizes 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**).

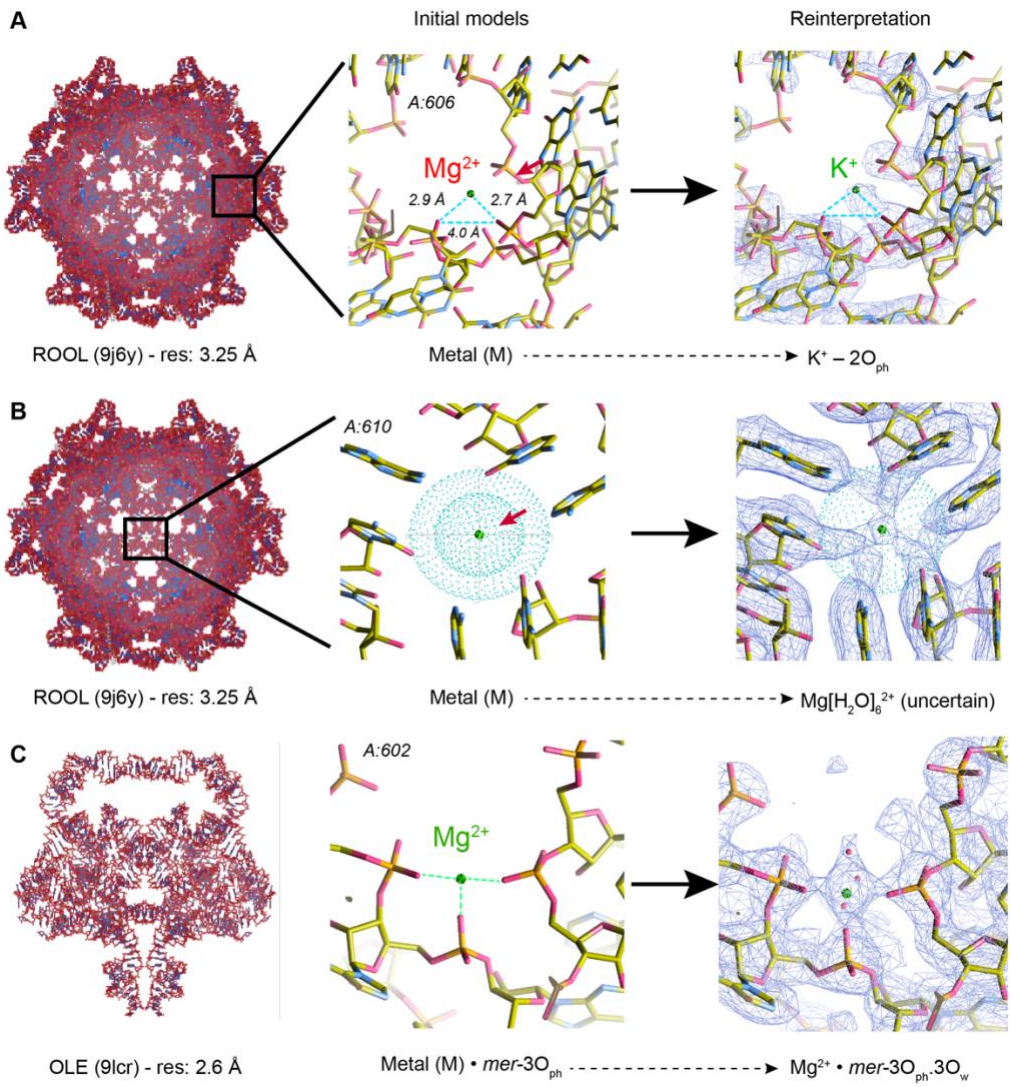


Figure 13. Ion placements in two low and one high-resolution quaternary structures. (A) In the **9j6y** structure (34), Mg^{2+} ions with coordination distances in the 2.6–2.9 Å range should be assigned as K^+ ions. **(B)** This ion is located on a three-fold symmetry axis. The associated non-octahedral densities are difficult to interpret, As visualized by spheres of dots with 2.3 and 3.6 Å radii generated by *MG_diagnosis*, the distances to the neighboring residues allow for the placement of a $Mg[H_2O]_6^{2+}$ ion. However, 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 valid *mer*-$3O_{ph}.3O_w$ assignment. The three missing water molecules can be modelled without introducing clashes by using *Coot* tools and *Cat_Wiz "Sphere refine with restraints"* options.