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. Here, we present *Cat_Wiz*, a *Coot*-integrated, stereochemistry guided toolkit that facilitates the localization, diagnosis, correction and annotation of Mg^{2+} binding sites. *Cat_Wiz* comprises three modules: *MG_diagnosis which* validates and regularizes existing assignments; *MG_detect which* identifies unmodelled ion binding sites; and *MG_clamp which* classifies recurrent Mg^{2+} clamp motifs. *Cat_Wiz* also includes a complete binding site annotation system. The stereochemical principles implemented in *Cat_Wiz* were derived from an earlier analysis of the 1.55 Å resolution *Escherichia coli* ribosome and from surveys of the *Cambridge Structural Database.* These principles provide a robust experimental foundation for characterizing Mg^{2+} binding sites. Applications to ribosomes, hammerhead ribozymes, group I introns, and quaternary RNA assemblies demonstrate that *Cat_Wiz* rapidly locates overlooked 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 identification of unique Mg^{2+}-dependent backbone folds. *Cat_Wiz* is also applicable to Mg^{2+} binding sites in proteins and all related biomolecular systems since it is based on universal stereochemical principles.

INTRODUCTION

Hexahydrated and partially dehydrated magnesium (Mg^{2+}) ions are central to RNA structure, folding, and catalysis (1-7). Accurate placement and subsequent validation of Mg^{2+} ions in experimental models are therefore critical for mechanistic interpretation, comparative mining and other computational use. While several studies have investigated Mg^{2+} coordination to RNA (5,8-14), misassignments remain common in deposited models where ions are often placed in environments incompatible with their characteristic octahedral geometry, 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). Consequently, the **8b0x** model was revised to correct Mg^{2+} assignments and missing water molecules mandatory to complete octahedral coordination shells. This process significantly increased the number of authenticated Mg^{2+} binding sites and revealed 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 ranging from 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 can further support 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 unmodelled 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 *Cat_Wiz* on four representative cases spanning different experimental modalities and resolutions: *(i)* the 1.55 Å resolution cryo-EM **8b0x** ribosome model (13), *(ii)* historically debated Mg^{2+} binding in hammerhead ribozymes (15,32), *(iii)* 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 rectified incorrect assignments, recovered missing hydrated ions and differentiated Mg^{2+} from K^{+} by enforcing representative coordination distances, thereby improving the fidelity of deposited structures. Additionally, *Cat_Wiz* can assist in the refinement and interpretation of nucleic acid and protein systems.

The stereochemical Mg^{2+} priors we defined in **Figure 2** were derived from small-molecule statistics through the *Cambridge Structural Database* (CSD) and 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 this context, CMM and *Cat_Wiz* are complementary tools.

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 offers a set of restraints to stabilize Mg^{2+} octahedra in lower-resolution regions. By embedding prior knowledge rather than learning from noisy structures, *Cat_Wiz* provides a fast, transparent refinement solution for solvent particles and Mg^{2+} ions that can enhance machine learning (ML) and artificial intelligence (AI) workflows. In the following sections, we *(i)* outline the current *Cat_Wiz* stereochemical nomenclature in detail, *(ii)* describe the different *Cat_Wiz* modules, *(iii)* apply the toolkit to the case studies introduced above and *(iv)* 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, a coordinate file (PDB or mmcif) and a corresponding map file derived from X-ray or cryo-EM experiments must first be loaded into Coot. The user can then navigate to the *"Calculate"* menu, select *"Run scripts",* and *choose* the desired script from the local directory. Upon first use, particularly with large structures such as 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, structure factor files (.sf) provided by the PDB must be converted into an *.mtz* file using the SF-tool web server (https://sf-tool.wwpdb.org) (47). Default cutoff parameters were empirically derived and can be adjusted in the setup header of the scripts that were tested with *Coot* version 0.9.8.98. A version adapted for *Coot* 1.xxx is currently under development. For additional guidance, users are referred to the updated help pages provided with each module. It is advised to exercise caution when working with *"over"*-sharpened maps that might alter results of the *MG_detect* module and occasionally lead to zero or negative *B-factors* at the refinement stage in Phenix. Next, we first describe additions to a previously reported Mg^{2+} binding site nomenclature (13) before detailing the 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 summarized in (**Figure 3**). This nomenclature reflects the strict octahedral coordination geometry of Mg^{2+} characterized by $d(Mg^{2+}...O/N)$ coordination distances of 2.07/2.19 Å, with the slightly longer distance for corresponding nitrogen coordination. The system uses *cis-/trans-/fac-/mer-/CIS-/TRANS-*prefixes together with the following atom type identifiers: *"O_{ph}"* for OP1/OP2 phosphate anionic oxygens; *"O_r"* for O2'/O4' ribose oxygens and O3'/O5' phosphate oxygens; *"O_b"* for O2/O4/O6 nucleobase oxygens; and *"N_b"* for non-protonated N1/N3/N7 nucleobase nitrogens.

• *General nomenclature:* As an example, $2O_{ph}.4O_w$ denotes a hexacoordinated ion bound to two O_{ph} atoms and four water molecules (**Figure 2C**). The prefixes *cis-* and *trans-* describe the relative orientation of the two O_{ph} atoms. When three non-water atoms adopt a *cis-* arrangement, the corresponding isoforms is noted *fac*-$3O_{ph}.3O_w$ (*facial*). When one O_{ph} pair is in *trans-* (10,14) configuration, the site is named *mer*-$3O_{ph}.3O_w$ (*meridional)*. When four non-water atoms coordinate the ion, as in $4O_{ph}.2O_w$, the orientation of the two water molecules defines the *cis-/trans-* terminology. To distinguish these binding types from the more common *cis-/trans*-$2O_{ph}.4O_w$ types, the capital prefixes *CIS-/TRANS-* 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 (replacing 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 direct Mg^{2+} contact and a water-mediated interaction (**Figure 3**). Hence, the corresponding $O_{ph}.5O_w$ or $O_b.5O_w$ magnesium binding type become *"water-mediated"*

clamps. We previously hypothesized that some of these patterns may represent precursors to more standard clamp geometries (14). Due to structural compaction, these water-mediated contacts also appear in more complex motifs but are not specifically annotated by the *Cat_Wiz* modules for the sake of nomenclature simplification.

• *Ribose O3'/O5' atoms and Mg^{2+} coordination:* Recent investigations on group I and II introns have described catalytic sites with suspected Mg^{2+}…O3' contacts (48,49). This observation is puzzling as direct coordination of divalent ions to O3' atoms have not been observed in the ensemble of structures we analyzed. Therefore, such contacts are not included in **Figure 3**. The only exception identified to date involves terminal nucleotides where the O3'-H group can coordinate to a Mg^{2+} ion, either 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). In this case, 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 distinct binding rules related to two terminal-residue hydroxyl groups and may contribute to stabilization of leaving groups during catalysis.

• *Non-hydrated Mg^{2+} ions:* Coordination of six non-water atoms has not yet been observed in nucleic acid structures, likely due to steric constraints imposed by phosphate and nucleobase groups (**Figure 3**). This coordination mode has, however, been observed in a protein structure in which the six oxygen atoms belong to protein carbonyl and carboxylate groups as well as to three phosphate groups from a 2'-deoxyguanosine-5'-triphosphate (DGT) ligand (PDBid: **8ef9**) (51). Both monohydrated and fully dehydrated patterns are shown in grey in **Figure 3** as they have not been observed in representative nucleic acid systems to date.

• 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 components 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, form polyatomic clusters with Mg^{2+}, as observed 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, the *Cat_Wiz* 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-type identifiers. For example, the tetracycline-bound ion in PDB entry **8cgj** would be described as a *TRANS*-$2O_b.2O_L.2O_w$ or, more specifically, *TRANS*-$2O_b.2O_{TAC}.2O_w$ when using the PDB ligand code for tetracycline (TAC). This example illustrates the versatility and extensibility of the Mg^{2+} annotation code.

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

All *Cat_Wiz* scripts, developed for *Coot* version 0.9.8.98, generate *GtkTables.* The table produced by the *MG_diagnosis* script provides precise stereochemical assessments of structural Mg^{2+} ions to enable the discrimination between improperly from properly assigned ions (**Figure 4**). The table is

organized into multiple tabs that guide users through inspection, validation and refinement. The first two tabs can be exported as a comma-separated value (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 flag potential issues, three coordination distance ranges are defined: 1.9–2.3 Å, 2.3–2.6 Å and 2.6–3.2 Å. Each range associates with a coordination number (CN) indicator. An ideal Mg^{2+} ion (11,12,14) is characterized by CN = 6 indicating that all six coordination distances fall within the 1.9–2.3 Å range (**Figure 2**). Such ions are listed in the *"2. Mg^{2+} (validated)"* tab (**Figure 4**). A CN = 5/1/0 reflects the presence of a slightly out-of-range coordination distance, while CN = 0/0/5 suggests a non-Mg^{2+} solvent particle, often consistent with K^+ when present in the experimental buffer (11,14,57). Re-inspection and local refinements are often necessary to validate the assigned ion identity. Color-coding facilitates interpretation: red highlights ions with coordination distances in the 2.6–3.2 Å range typical of K^+ (tab1), while green marks validated octahedral Mg^{2+} ions with CN = 6 (tab2).

• *Binding site quality:* Each Mg^{2+} binding site is evaluated using three quality metrics *"rmsd ion"*, *"distorted octahedron (Dist. oct.)"* and *"Clash"* which together define the *"high-", "fair-" and 'low-quality"* tags. The cells in 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 primary quality marker. Sites with rmsd values exceeding 10% of the highest peak are considered high-quality (*green*). Values below 5% are tagged as *"low-quality"* (*grey*) while intermediate values (5–10%) are deemed *"fair"* (*pink*). Map-dependent rmsd thresholds are noted in tab5 (*"Map info")*. Importantly, rmsd values are computed from the original experimental map rather than from the smoothed map.

• *Distorted coordination octahedron (Dist. oct.) indicator:* This indicator evaluates distortion of the Mg^{2+} coordination geometry by measuring distances between first-shell atoms which average 2.9 Å in an ideal octahedron (**Figure 2A**). As a proxy for overall coordination quality, only the shortest of these inter-atom distances is reported. Sites are classified as *"high-quality"* when this distance exceeds 2.6 Å (green), *"fair-quality"* between 2.4 and 2.6 Å (pink), and *"low-quality"* below 2.4 Å (grey; **Figure 5**).

• *Clash indicator:* Strict Mg^{2+} coordination distances (< 2.3 Å) generate a 2.3–3.4 Å exclusion zone around the ion (**Figure 2B**). The presence of atoms within the 3.2–3.4 Å range induces a severe clash (grey). In some convoluted binding patterns, distances of 3.4–3.6 Å are categorized as mild clashes (pink). Such clashes generally require deletion or reassigment of the Mg^{2+} ion, remodeling of the binding site, or re-evaluation of ion identity. Representative examples are discussed in the **Results** section.

• *Global quality indicator:* A global quality score is derived by aggregating the individual indicators described above. If any single metric indicates *"low-quality"*, the entire site is classified accordingly. Otherwise, sites are labeled as *"fair-"* or *"high-quality"*. This global indicator applies only to *"validated"* Mg^{2+} ions in tab2 and is reported 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, provides visualization and refinement options (**Figure 4D**).

Selecting *"Zoom to this ion"* centers the view on an ion and fascilitates inspection of *d*(Mg^{2+}…O/N) distances, missing water molecules, or incomplete octahedral shells. Slightly elongated coordination distances can often be corrected through local real-space refinement in *Coot*, while missing water molecules can be added using density-guided placement tools of *Coot*. When solvent densities are located at more than 4.3 Å from solute atoms and lack clear octahedral density patterns, distinguishing hydrated $Mg[H_2O]_6^{2+}$ ion from alternative solvent particles becomes challenging. In such cases, careful examination of second-shell contacts, steric clashes, and experimental buffer composition is required to validate ion identities (14).

• *Spherical refinement of the ion binding site: Coot* real-space refinement tools can be used at any stage. The *"sphere refine"* command, accessible by a right-mouse click on a tab1 or tab2 entry, performs unrestrainted real-space refinement of atoms within specified radius (default = 3.5 Å). When density patterns are well-defined, this typically yields near-ideal octahedral geometries. However, in regions with blurred densities, unrestrained refinement may introduce distortions (**Figure 6**). In such cases, the *"sphere refine with restraint"* option is recommended, as it enforces 2.07/2.19 Å restraints for the six *d*(Mg^{2+}…O/N) coordination distances.

• *Additional distance restraints:* In some instances, and for less-well resolved sites, additional restraints may be required to stabilize the binding site and avoid too-close Mg^{2+}…O/N contacts. This can be done by navigating to the *"Calculate → Scripting → Python"* menu using the *Coot* command *"add_extra_bond_restraints"*, which requires thirteen arguments as listed in the **SI**.

• *Mg^{2+}…Mg^{2+}/K^+ ion pairs:* When two residues coordinate the same Mg^{2+} ion, the "*sphere refine with restraint"* option may induce slight deviation from ideal 2.07/2.19 Å coordination distances, typically remaining within *d*(Mg^{2+}…O/N) < 2.3 Å. Such deviations are commonly observed in Mg^{2+}…Mg^{2+} and other ion pairs, which require a dedicated inspection. *MG_diagnosis* identifies all Mg^{2+}…Mg^{2+}/K^+ pairs with inter-ion distance ≤ 6.0 Å. Pairs with *d*(Mg^{2+}… Mg^{2+}/K^+) < 4.2 Å (orange) are currently poorly documented and require careful analysis, while those in the 4.2–6.0 Å range (cyan) should be visually inspected to assess potential Mg^{2+} μ-clusters involving shared phosphate atoms (14).

• *Doted spheres to visualize clashes:* For ambiguous cases, *MG_diagnosis* allows visualization of doted spheres with radii of 2.3, 3.2, and 3.6 Å. A site is considered stereochemically valid if all six coordinating atoms fall within the 2.3 Å sphere. Atoms within the 2.3–3.2 Å range indicate the need for correction, while solute atoms penetrating the 3.2 Å sphere are incompatible with hydrated Mg^{2+} ion requires that no solute atoms enter the 3.2 Å sphere. Ideally, no atoms should enter the 3.6 Å sphere, as second shell coordination distances peak around this value (**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. Together, the first three tabs provide an appropriate overview of the quality of the Mg^{2+} assignments for users who do not intend to further refine the ionic coordination shells.

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+} ions, 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 thereby isolate unassigned density spots (**Figure 7**).

MG_detect_RNA and *MG_detect_Prot* inspect density spots in proximity to nucleotide and protein atoms. For nucleotides, these include phosphate and nucleobase atoms (OP1/OP2/O2/O4/O6/O2'/N7/N3/N1), while for proteins they include side-chain atoms (OD1/OD2/OE1/OE2/ND1/NE2/N/NZ/NE/NH1/NH2) as well as 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 occasionally coordinate transition metals as detailed below. For modified nucleotides or amino acids that involve different atom names, nearby density spots must be inspected manually. Since *MG_detect_RNA* and *MG_detect_Prot* operate analogously, the description below focuses on *MG_detect_RNA.*

The generated mask is searched for the highest-density peaks using *Coot "map_peaks_around_molecule"* command. In all tabs associated with OP1/OP2/O2/O4/O6/O2'/N7/N3/N1 atoms, peaks are ranked by decreasing rmsd values in the *"Mg site"* column. To avoid overlaps with existing atoms, only peaks located within 1.8–2.4 Å of any assigned atom are retained as candidates for inner-sphere contacts. The *"Residues < 2.4 Å"* column lists all residues located near each 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 excluded while peaks 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 cannot be associated with Mg^{2+}, K^{+} or water and instead indicates a metal-mediated base pair interaction (**Figures 7E–G**). Similarly, 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 do not indicate Mg^{2+} binding sites but may reveal the presence of transition metals or problematic local modeling.

Interestingly, some peaks also highlight poorly modelled backbones or reveal the presence of alternate conformations. We therefore intentionally avoid excessive filtering this list to allow such issues to be detected and corrected. In addition, *MG_detect* can also locate unassigned densities

consistent with octahedral or tetrahedral transition metals (Mn^{2+}, Zn^{2+}, etc.) that share Mg^{2+} coordination distances. As a note of caution, these modules also report peaks that associated with diffuse or ambiguous densities, which may arise from disorder, unidentified solvent particles or insufficient knowledge.

Besides inner-shell contacts, *MG_detect_RNA* can identify hexahydrated Mg^{2+} ions of the $6O_w$ type, provided that water molecules have been assigned. For that, automated procedures implemented in *Coot* or *Phenix* are recommended. 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 density peaks that may correspond to Mg^{2+} ions or transition metals while excluding peaks associated with water and monovalent ions (e.g., K^+ or NH_4^+) that exhibit coordination distances in the 2.6–3.2 Å range. Following *MG_detect,* all newly assigned Mg^{2+} positions must be evaluated using *MG_diagnosis* to validate ion identity, correct 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 approach is highly effective for most high-resolution structures (< 2.0 Å). However, in lower-resolution structures or in poorly resolved regions of large complexes, the experimental density patterns may be insufficiently defined for *MG_detect* to be fully reliable. In such cases, *MG_clamp provides* a complementary identification strategy by searching 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 separated by *d*(O/N…O/N) ≤ 3.4 Å can only be stabilized by coordination to Mg^{2+} or to other transition metals. Such stabilization typically occurs via 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 using appropriate protocols in *Phenix* or *Molprobity* (60,61), Asn or Gln side chains may revert to an inappropriate orientation (**Figures 8A/B/E**). *MG_clamp* therefore identifies all *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 *Coot "flip 180°"* 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 reduce 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** belong to 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 frequently indicate incompletely modelled sites. These pairs are retained for visual inspection.

• *Tab3 ("Mg^{2+} clamps"):* The third tab identifies and annotates all d(O/N…O/N) ≤ 3.4 Å contacts that are associated with an assigned Mg^{2+} ion. Clamps involving a hexacoordinated Mg^{2+} ion (CN = 6) are highlighted in green, whereas sites requiring adjustments or corrections are shown in cyan. The latter can typically be remodelled straightforwardly in *Coot.* The *"n…n+1"* column identifies clamps in which 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 and annotated in the *"Comment"* column (**Figure 3**). Finally, as with all newly assigned ions, visual inspection of the binding site and the corresponding experimental density using *MG_diagnosis* is required prior to validation.

• *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 Mg^{2+} assignments are introduced or existing ion positions are modified, *B-factors* must be updated prior to PDB deposition. This can be achieved by refining individual *B-factors* using *Phenix* or an equivalent refinement program. Care is required when working with over-sharpened maps, which may otherwise lead to zero or negative *B-factors* during refinement.

RESULTS

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

Analysis of the *E. coli* ribosome structure determined at 1.55 Å resolution (PDBid: **8b0x**) serves as the primary case study for evaluating the performance *Cat_Wiz*. This structure was previously investigated in detail by us, revealing the presence of 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 formed the basis for establishing a systematic classification of Mg^{2+} binding sites in RNA and RNA/protein systems (14). Besides identifying new binding sites, this earlier study corrected misassignments by reclassifying several Mg^{2+} ions as water molecules or K^+ ions and by deleting others. Using *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. No novel Mg^{2+} binding motifs beyond those summarized in **Figure 3** were detected during this process. The revised structure, deposited in the PDB under the identifier **9q87**, reveals significant differences in the ionic shell compared with the original **8b0x** and the new structure (**Table 1**). In **8b0x,** only 100 Mg^{2+} ions exhibit a hexacoordinated geometry that met our stereochemical criteria. In contrast, the refined **9q87** structure contains 406 validated hexacoordinated Mg^{2+} binding sites.

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

MG_detect further improved completeness of the **amended 8b0x** (14) model by identifying 12 additional Mg^{2+} ions while nine previously assigned Mg^{2+} ions were flagged for anomalous features (**Table 1**). One puzzling ion exhibited a coordination shell of seven water molecules with coordination distances close to 2.07 Å, a geometry more consistent with a K^+ assignment than Mg^{2+}. Eight additional ions were associated with Mg^{2+}…Mg^{2+} ion pairs and displayed atypical coordination patterns. Approximately 50 *"low-quality"* instances were identified (**Table 1**). Although these sites passed inspection and did not exhibit steric clashes, they are not suitable for inclusion in statistical surveys or AI training sets. Nevertheless, documenting these sites remains important for achieving a complete description of the ribosome solvent shell. Finally, *MG_detect identified* one ion with a coordination number of five that appears 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, as annotated in tab1 (**Figures 8A/B**). This script flagged 33 contacts 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+} binding motif involving the small ribosomal subunit, ribosomal protein (r-protein) S7, and a Mg^{2+} ion at position a:3219. MG_clamp required approximately 10 minutes to run on a standard modern laptop.

A few side notes are necessary for a better understanding of the **8b0x/9q87** structures. First, the *E. coli* ribosome sequence used in this structure originates from strain B rather than the more frequently studied MRE 600 (62,63) or K12 strains (64), resulting in systematic 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 currently be annotated using standard 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) which likely represent metal-mediated base pairs, as well as between a:G_{1366}:N7 and the N-terminal w:Ser_2 of r-protein bl28. An additional CU site was assigned to a density located between v:His_{57}:ND1…v:Cys_{53}:SG of r-protein L27 (**Fig. S1**). The observed coordination geometries are consistent with Cu^{+} ions(67-70). If present, these Cu^{+}/Cu^{2+} ions represent buffer contaminants potentially originating from grid-related contamination. Importantly, *MG_detect* successfully localized these non-Mg^{2+} densities.

Finally, although *Cat_Wiz* identified the majority of prominent Mg^{2+} clamps in this ribosome (14), it remains difficult to recover the location of all $6O_w$ and $5O_w$ sites. Diffuse ion-binding events, remain currently beyond the reach of *Cat_Wiz* and of available experimental approaches.

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

As a second example, *Cat_Wiz* was applied 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 earliest 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, arguing that Mg^{2+} ions are not directly involved in the catalytic mechanism (73-75). To date, structural studies including X-ray crystallography have not provided sufficient evidence for conserved or catalytically essential Mg^{2+} binding sitesin the hammerhead ribozyme. Current hypothesis instead propose an acid-base catalytic mechanism, in which a hydroxyl group within the active site functions 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 is now more commonly classified as an acid-base catalysed ribozyme rather than a true metalloenzyme.

A challenge in analyzing these X-ray structures arises from the presence of submolar concentrations of cacodylate ($AsO_2(CH_3)_2^{-}$) and, in some instances, high concentrations of sulfate (SO_4^{2-}) in crystallization buffers (**Tables 2/S1 and Figure 9**). Indeed, several originally assigned Mg^{2+} ions are stereochemically inconsistent with Mg^{2+} coordination and are more plausibly explained as

sulfate or cacodylate anions (82,83) which could, in principe, have been detected through their strong anomalous X-ray diffraction signals (84,85).

To reassess these assignments, we analyzed all 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 explicitly assigned Mg^{2+} ions, reporting five such ions. Four of these ions are located more than 3.2 Å away from any nucleotide, consistent with fully solvated ions like $Mg[H_2O]_6^{2+}$, although their corresponding densities could equally be attributed to sulfate or cacodylate ions (**Figure 9**). A fifth ion, positioned at a distance of 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 hammerhead dimer structures **2quw** and **2qus**, solved at resolutions between 2.2 and 2.4 Å, contain 17 and 16 assigned Mg^{2+} ions, respectively. In **2quw** *Cat Wiz* identified only a single Mg^{2+} with a 2.18 Å distance and sufficient space to accommodate a hexahydrated shell (**Figure 11B**). However, unlike the 2.0 Å resolution Mn^{2+}-containing structure **2oeu** (**Table S1**), no clear octahedral water densities were observed. Five of the remaining metal densities in **2quw** could be assigned to hydrated Mg^{2+} ions or to anions. On the other hand, the remaining ten assigned Mg^{2+} ions are poorly positioned and sterically incompatible with Mg^{2+} binding according to *Cat Wiz.* Placement of a hydrated Mg^{2+} ion at these sites would induce severe clashes with neighboring residues (**Figure 11C**).

The absence of well-defined Mg^{2+} binding sites, particularly when contrasted with structures containing transition metals or Na^+, complicates mechanistic interpretations on currently assigned ion positions. Future investigations of these hammerhead ribozymes may therefore be required to clarify the catalytic mechanisms of hammerhead ribozymes(74,87,88). At present, available hammerhead structures underscore the challenges associated with interpreting solvation shells in low-resolution systems that contain stereochemically unsupported Mg^{2+} assignments which at best can only support speculative interpretations.

Another set of hammerhead structures containing questionable Mg^{2+} assignments include **5dh6**, **5dqk**, and **5eao**. In **5dh6,** the proposed OP2:103 Mg^{2+} binding provides weak support for specific Mg^{2+} coordination as $d(Mg^{2+}\ldots O) \approx 2.48$ Å exceeds typical Mg^{2+}...O distance (**Figure 11D**). Similarly, **5dqk** and **5eao** show densities close to 2.45 Å with no octahedral coordination geometry (**Figure 11E**), suggesting that these features are more consistent with other solvent particles than Mg^{2+} ions. The large number of poorly assigned Mg^{2+} ions in these structures significantly further diminishes confidence in these interpretations. By contrast, the 1.55 Å resolution structure **3zp8**, crystallized in a 1.7 M sodium malonate buffer, displays several clearly defined octahedrally coordinated Na^+ ions (**Figure 11F**).

In retrospect, early stereochemical validation of these metal-binding sites could have substantially clarified mechanistic interpretations of hammerhead ribozyme. These observations highlight the importance of comprehensive analysis across all available structures of a given system, as well as the

need to verify that metal ion assignments conform to fundamental stereochemical principles. A detailed understanding of crystallization buffer composition is equally critical to avoiding misassignments. Overall, *Cat_Wiz* proved highly effective for reevaluating 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 previously 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 under PDBids **9cbx** and **9cby** at 2.20 Å and 2.30 Å resolutions, respectively, was analyzed. In these structures, Mg^{2+} ions and water molecules were automatically assigned using a segmentation-guided water and ion modelling protocol called SWIM (22,59,91), and they display slightly different P9.2 stem configurations. Companion structures (**9cbu** and **9cbw**) which contain only the 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 that meet our stereochemical criteria and 12 that do not (**Table 3**). Of the validated ions, only four possess 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 modelled 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 Mg^{2+} binding sites, 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:* The **9cbu** and **9cbx** structures contain 22 (consensus) and 54 Mg^{2+} ions, respectively. Surprisingly, hydration shells were not modelled for these ions, 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 coordination analysis of **9cbu**, we examined each of the 22 assigned Mg^{2+} ions individually. 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 the **1hr2** structure were absent in **9cbu**, despite clear densities supporting their presence (**Figures 12C/D/E** and **Table 3**). More concerningly, 16 Mg^{2+} ions were placed at locations lacking visible octahedral density. Modeling a full hydration shell at these locations result in steric clashes with neighboring residues (**Figure 12F**) suggesting that these sites are unlikely to represent *bona fide* Mg^{2+} binding sites.

Regarding non-consensus ions, we found that the ion modelled in **Figures 12C/D/E** is absent 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 complete hydration sphere induces clashes. The *MG_detect* tool failed to identify any other credible positions for Mg^{2+} ions, despite the presence of numerous unassigned densities.

Taken together, our findings suggest that the *Tetrahymena* group I ribozyme requires only a few precisely positioned Mg^{2+} ions to adopt its active conformation. This analysis demonstrates how *Cat_Wiz* a critical inspection of solvation shells in moderate resolution structures and effectively discriminates between improbable and valid ion assignments. A more comprehensive view of the solvation shell of these group I introns will require higher-resolution structures, preferably obtained using K^+ rather than Na^+ containing buffers, as used in most deposited structures.

Case IV: RNA quaternary structures

• *Exploring metal binding sites in RNA scaffold structures:* Recent advances in cryo-EM have opened new avenues 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. At these resolutions, metal hydration shells were not modelled. Noteworthy, in the publication describing these quaternary structures (34), the assigned ions were labelled as “Metals” or (M), except in the supplementary material. In the corresponding PDB file, some of these ions are incorrectly annotated as “Mg” highlighting how automatic database surveys can be compromised by annotations that do not adhere to basic stereochemical criteria. We therefore used these structures to evaluate *Cat_Wiz* performance on low-resolution models and to assess its ability to identify plausible Mg^{2+} binding sites and clamps. In addition, we analyzed a higher-resolution structure of a *Ligi lactobacillus salivarius* ROOL monomer (PDBid: **9m79**), solved at 1.96 Å, for which no ions or water molecules were assigned despite the improved 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 the 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 at this site, which are more consistent with K^+ binding rather than with Mg^{2+} (**Figure 13A**). The remaining 38 ions, located more than 3.2 Å from RNA atoms, are difficult to assign without additional experimental evidence. While they could relate to hexahydrated Mg^{2+} ions, they may also represent other buffer components. For example, 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 absence of clear octahedral density raises doubts about this assignment (**Figure 13B**).

• *The **9lcr** Clostridium botulinum OLE dimer (2.6 Å):* In this dimer, hydration shells were not modelled for the 28 assigned metals (14 per monomer). Among the 14 ions in monomer “A”, one valid *mer-*

$3O_{ph}.3O_w$ coordination site (M2) was identified. Here, metal coordination distances are below 2.3 Å and the *cis-d*(O_{ph}…O_{ph}) distances are 2.7/3.1 Å. Four additional 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, only remodel a subset of the assigned metals could be remodelled. At the *mer*-$3O_{ph}.3O_w$ binding site (M2), densities corresponding to three missing water molecules were visible (**Figure 13C**). Using *Coot*, we placed water molecules at their presumed location and reran *MG_diagnosis* followed by positional optimization using the *"Sphere refine with restraints"* option. In a similar manner, we added water to a presumed hexahydrated Mg^{2+} ion at the M1 site and to an M5 ion at a *trans*-$2O_{ph}.4O_w$ site. This procedure yielded satisfactory hydration sphere model, validating the presence of octahedral coordinated Mg^{2+} ions. A more in-depth remodeling of these structures, including the evaluation of additional 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 modelled in this structure (92). Yet *MG_clamp* identified 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 assemblies indicate that Mg^{2+} and K^+ ions in low-resolution structures can in some cases, be differentiated or invalidated based on density patterns and stereochemical considerations. Such prior knowledge-based approaches are essential to limit the propagation of misassignments in structural databases. In these examples, *MG_detect* and *MG_clamp* did not identify additional Mg^{2+} binding sites beyond those discussed. Besides providing a clear survey of the assigned ions, *MG_diagnosis* proved particularly useful for modeling Mg^{2+} hydration shells in these challenging low-resolution structures.

DISCUSSION

Accurate assignment of Mg^{2+} binding sites remains one of the most significant challenges in RNA structural biology. Misplaced ions are common in the PDB even in high-resolution models where stereochemical constraints are violated or hydration shells are incomplete (11,12,14,94). Such errors and incompleteness lead to inaccurate mechanistic interpretations can mislead 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 challenges by integrating stereochemical binding principles into a unified analysis 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 experimental density or backbone modeling may be inadequate. Herein, we demonstrated applications to ribosomes, hammerhead ribozymes, group I introns, and RNA quaternary assemblies, illustrating how *Cat_Wiz* can recover incomplete ionic hydration shells, correct misassigned cations, and annotate ion motifs such as Mg^{2+} clamps. Importantly, these corrections can be achieved within hours rather than days of manual examination. *Cat_Wiz* is applicable to both X-ray and cryo-EM structures across a broad resolution range and can be used for nucleic acids as well as proteins.

While *Cat_Wiz* provides valuable insights into RNA structural biology it shares limitations common to all structure-based approaches. At resolutions poorer than ≈2.5 Å, ion coordination becomes increasingly difficult to distinguish and errors in RNA backbone modeling may propagate into ion placement. To mitigate these issues, *Cat_Wiz* provides restrained refinement options that enforce realistic Mg^{2+} coordination distances while preserving structural flexibility. These restraints are conceptually analogous to those routinely applied to nucleotides and amino acids and 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 including cacodylate and sulfate (82,83) are frequently present in experimental buffers. Their coordination distances can overlap with those of Mg^{2+}, complicating density interpretation and increasing the likelihood of misassignment. *Cat_Wiz* explicitly highlights cases where such ions have been misinterpreted as Mg^{2+}, as illustrated here for 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 ensure physiologically meaningful assignments. Notably, the use of *"simplified"* buffers does not preclude the determination of high-resolution structures and is compatible with *in vitro* autonomous ribosome biogenesis (19,102). In the future, more systematic inclusion of polyamines in experimental buffers may further advance our understanding of biological solvation shells (62,103-106). Although chloride (Cl^-) is the most abundant intracellular anion (3,82,83), it is often absent from modern experimental buffers and replaced by other anionic components including acetate, malonate, HEPES or TRIS molecules. Such substitutions may perturb RNA solvation shells in ways that remain poorly understood.

A third concern relates to the deposition and annotation of structural data (15,18,107,108). Misassigned ions lacking appropriate annotations persist in the PDB, despite repeated identifications of such errors (15,32). Greater and more consistent user of the CAVEAT remark would help flag known issues in deposited PDB structures (18). Such inaccuracies not only mislead newcomers to the field but also comprise automated surveys and machine-learning training datasets. Herein, we documented how such issues affect hammerhead and recent group I intron ribozyme structures potentially jeopardizing collective efforts to understand and model their ionic structure and catalytic properties (22). Broader adaption of annotation tools such as *Cat_Wiz*, together with improved PDB-level flagging, would significantly enhance the overall quality of publicly available structural models. As a side note, the PDB currently lacks specific residue codes for anionic U(-) and G(-) nucleotides recently identified in bacterial ribosomes (e.g., **8b0x**) despite the existence of dedicated codes for protonated A(+) and C(+) residues ("AP7" and "CH", respectively) (14,65). Addressing this limitation would further improve ion and charge-state annotation.

Cat_Wiz also highlights the inherent limits of automation. Alternate nucleotide conformations, ambiguous solvent densities, and unusual motifs such as metal-mediated base pairs or diffuse ion populations, require expert visual inspection prior to publication. *Cat_Wiz* is designed to flag these instances, yet it does not aim at automatically resolving them, as careful human interpretation remains essential for understanding and documenting these features. Even with *Cat_Wiz*-assisted inspection, analyses of poorly resolved regions may still raise more questions than definitive answers.

Finally, while machine-learning ion identification approaches are gaining popularity (20,109,110), current methods struggle to correct stereochemical errors and misassignments already present in existing models. Similarly, MD simulations that lack polarization and charge transfer effects must be interpreted cautiously when modelling Mg^{2+} ions (22,111,112). By contrast, *Cat_Wiz* provides a crucial bridge between experimental data and modelling efforts. By incorporating rigorous stereochemical principles and generating high-quality ion assignments, it supplies curated datasets that can improve the accuracy and robustness of next generation AI tools and molecular dynamics force fields (113-117).

To summarize, *Cat_Wiz* offers a rapid, and stereochemistry-guided framework for diagnosing, annotating and correcting Mg^{2+} assignments in biomolecular structures. By improving the fidelity of solvation-shell models and producing carefully curated datasets, *Cat_Wiz* benefits experimentalists seeking accurate mechanistic interpretations and computational scientists developing predictive models of ion-RNA interactions. Its integration into *Coot* makes it broadly accessible, positioning *Cat_Wiz* as a practical, fast, and transparent tool supplementing both manual refinement and emerging autonomous structural biology approaches.

Disclaimer

We made every effort to design *Cat_Wiz* rigorously and to assign Mg^{2+} ions in the **8b0x/9q87** structure as accurately as possible. The quality of these assignments necessarily depends on structural resolution, preceding refinement steps, and the state of prior knowledge available at the time of structure deposition, particularly when disentangling complex or unforeseen solvation patterns.

This remains an ongoing process. Accordingly, we do not claim: *i)* that all Mg^{2+} ions were identified, *ii)* that particularly complex or poorly resolved solvation patterns did not occasionally lead to ambiguous interpretations, or *iii)* that all reported coordination patterns are free of modeling imperfections. However, our analyses were conducted using the best available structural data and current stereochemical knowledge. We are confident that all major Mg^{2+} binding sites in the **8b0x** ribosomal structure were identified and that the most relevant Mg^{2+} binding types across the RNA systems examined were captured, enabling a near-complete and internally consistent annotation framework.

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. We also thank Prof. Giovanni Bussi that organized a CECAM meeting in Paris where P.A and S.K. first met and started this collaborative research project.

Author contributions: Nawavi Naleem (Conceptualization [equal], Methodology [equal], Software [lead], Validation [equal], Visualization [equal], [Writing - review & editing [supporting]), Anja Henning-Knechtel (Conceptualization [supporting], Methodology [supporting], Writing - review & editing [supporting]), Serdal Kirmizialtin (Conceptualization [supporting], Funding acquisition [lead], Methodology [supporting], Project administration [lead], Resources [lead], Software [supporting], Supervision [supporting], Writing - review & editing [equal]), Pascal Auffinger (Conceptualization [lead], Methodology [equal], Project administration [supporting], Software [supporting], Supervision [lead], Validation [equal], Visualization [equal], Writing - original draft [lead], Writing - review & editing [equal]). d

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.*

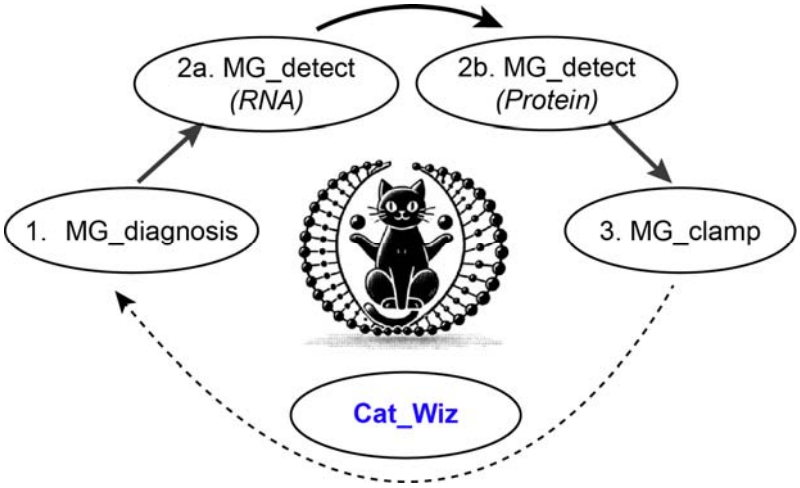


Figure 1. *Cat_Wiz* workflow. The diagram illustrates the recommended sequence for using the four independent Mg^{2+} related *Cat_Wiz* scripts. It is advised to re-run the *MG_diagnosis* script whenever structural changes were made to ensure the accuracy of the final model.

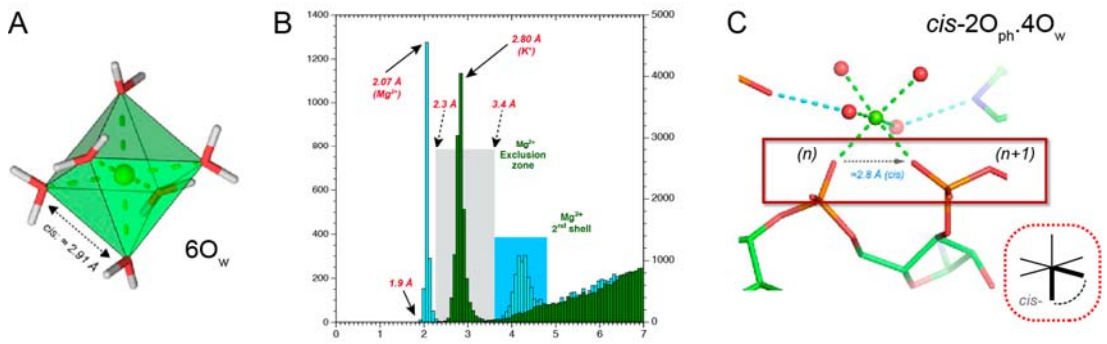


Figure 2. Stereochemical features of Mg^{2+} ions and Mg^{2+} clamps. (**A**) A hexahydrated Mg^{2+} ion ($6O_w$) from a high-resolution *Cambridge Structural Database* (CSD) entry, showing characteristic $d(O_w \ldots O_w)$ distances around 2.9 Å (14). (**B**) Histograms of Mg^{2+} (cyan) and K^+ (green) to oxygen coordination distance from the CSD. The distinct ~0.7 Å distance between the two peaks is 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.

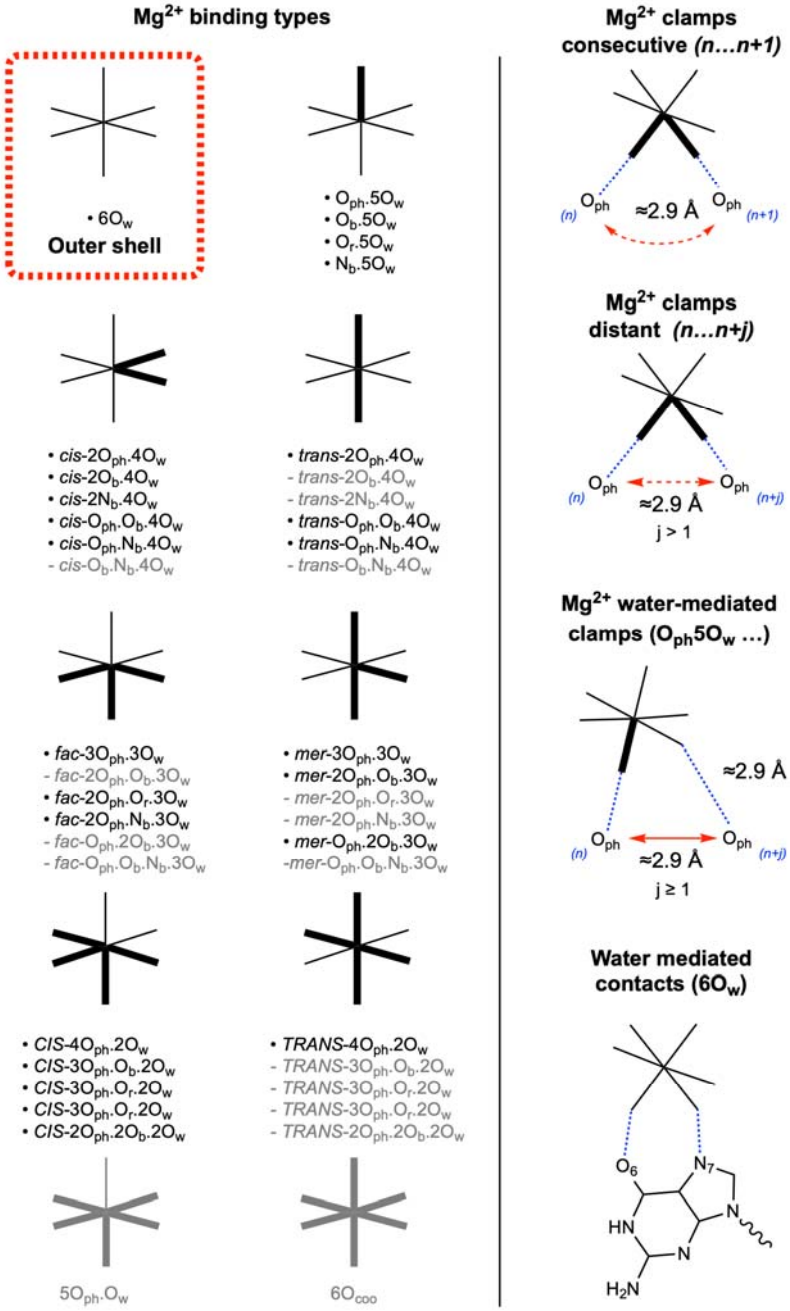


Figure 3. Updated view of Mg^{2+} binding motifs. (Left) This panel presents all inner- and outer-shell Mg^{2+} binding motifs identified in **9q87** and is an update of the list derived from **8b0x** (14). For clarity, motifs involving ribosomal proteins have been excluded. Greyed out binding types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

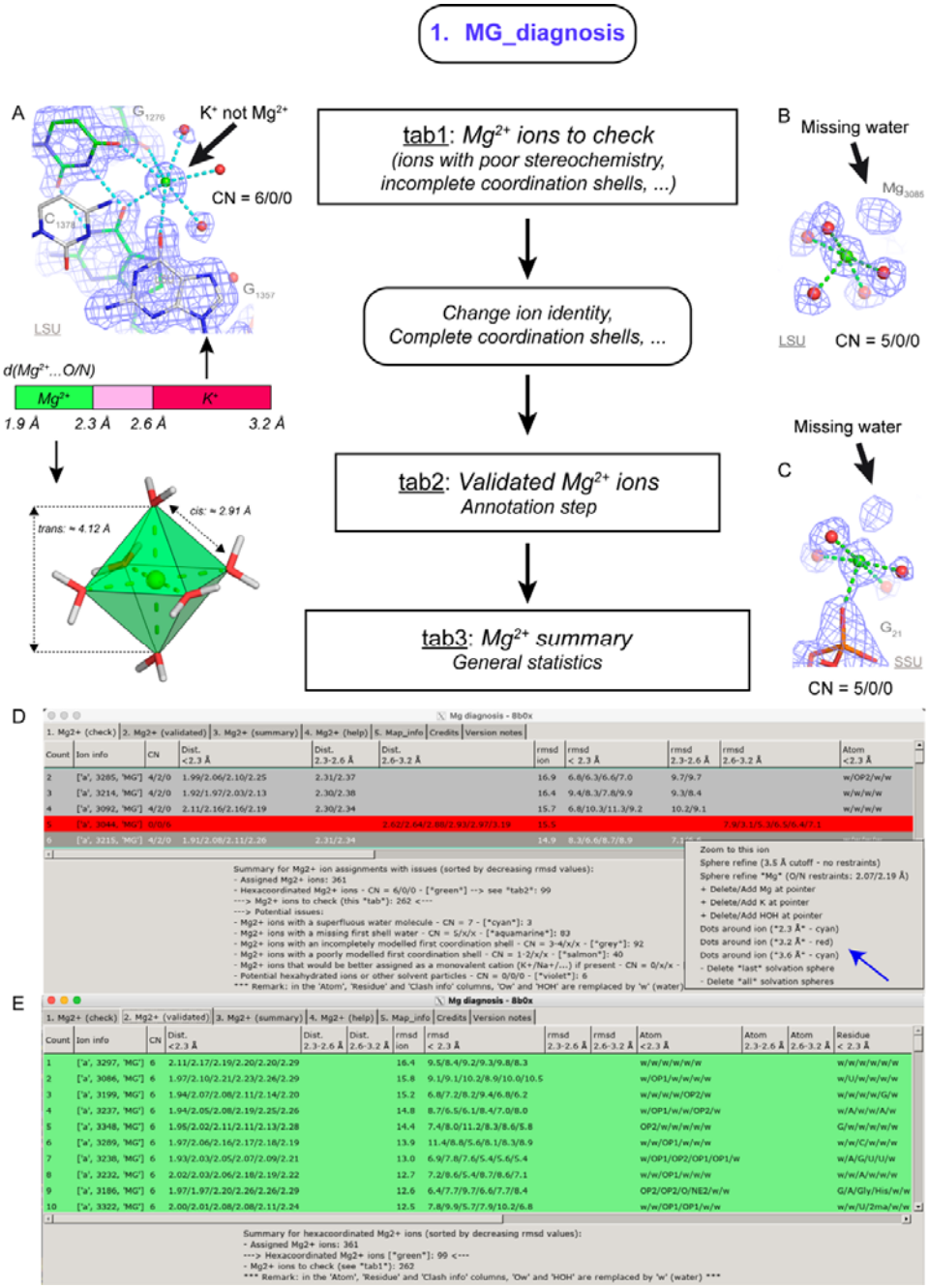


Figure 4. Workflow of the *MG_diagnosis* script. The central column illustrates the *MG_diagnosis* primary functions, i.e., *i)* assessing the stereochemical quality of assigned ions and *ii)* recommending changes based on specific coordination number (CN) and distance criteria. These examples are taken 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).

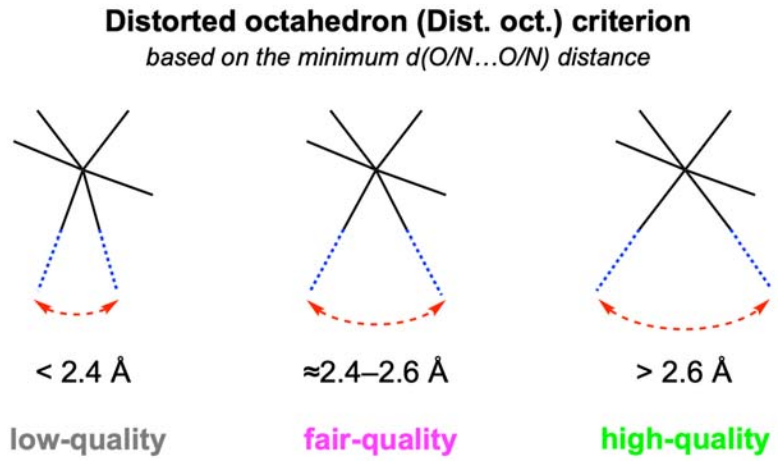


Figure 5. The distorted octahedron (Dist. oct.) criterion. The quality of a binding site is *"low"*, *"fair",* or *"high"* based on the smallest distance between any two first shell coordinating atoms. This criterion is exclusively applied in tab2 of *MG_diagnosis*. These distance cutoffs were empirically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

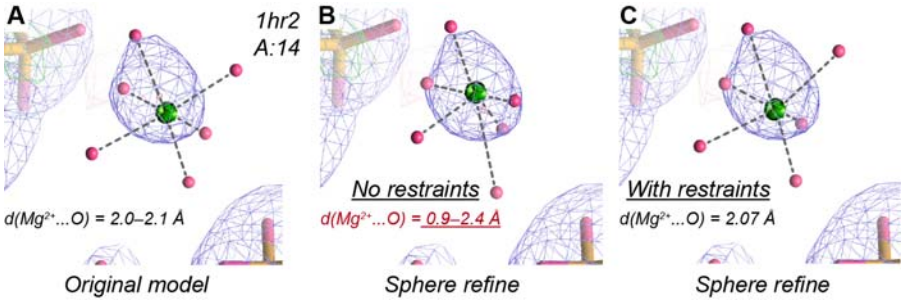


Figure 6. Effect of restraints on the *Coot "Sphere refine"* option. Different outcomes are observed when using the *Coot "Sphere refine"* command with and without distance restraints. (**A**) A hexahydrated Mg^{2+} ion modelled into a non-octahedral density pattern (89). (**B**) Refining the same position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

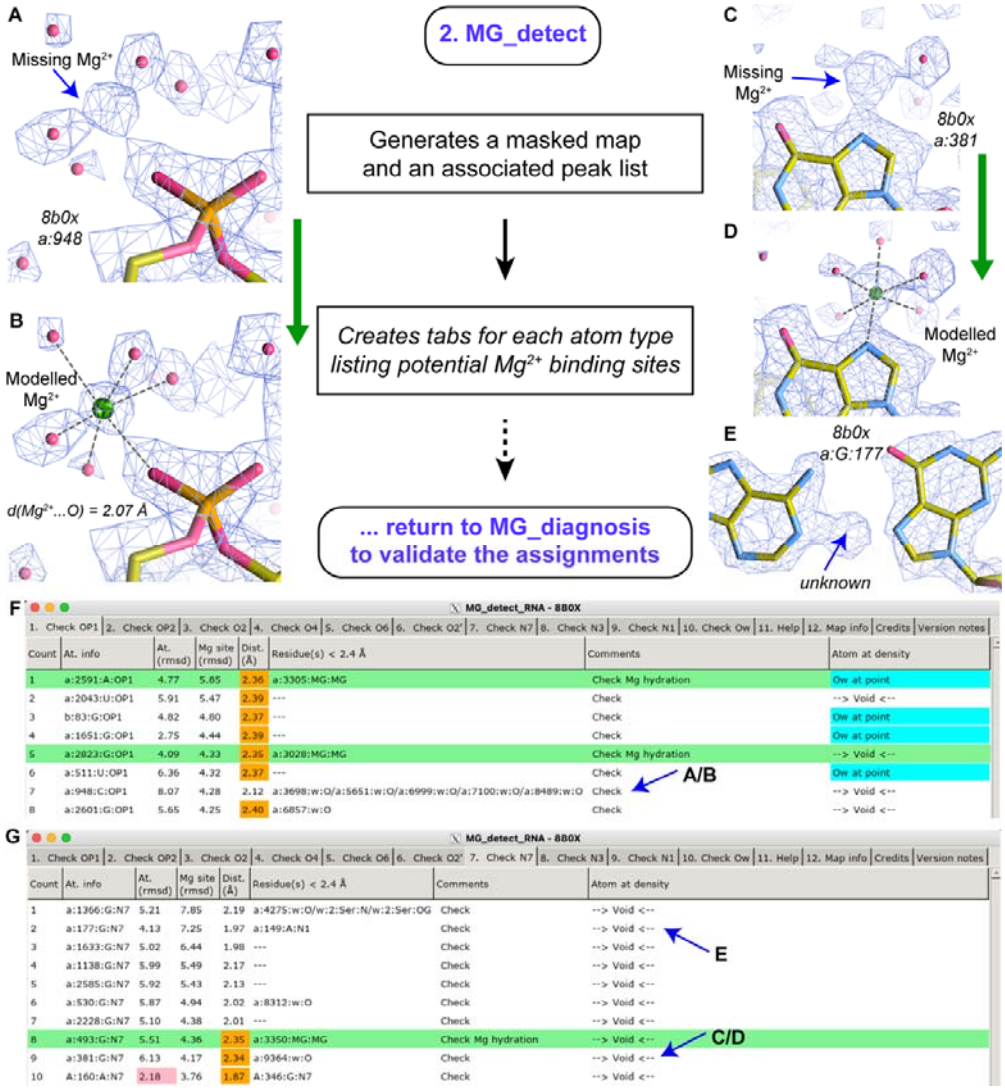
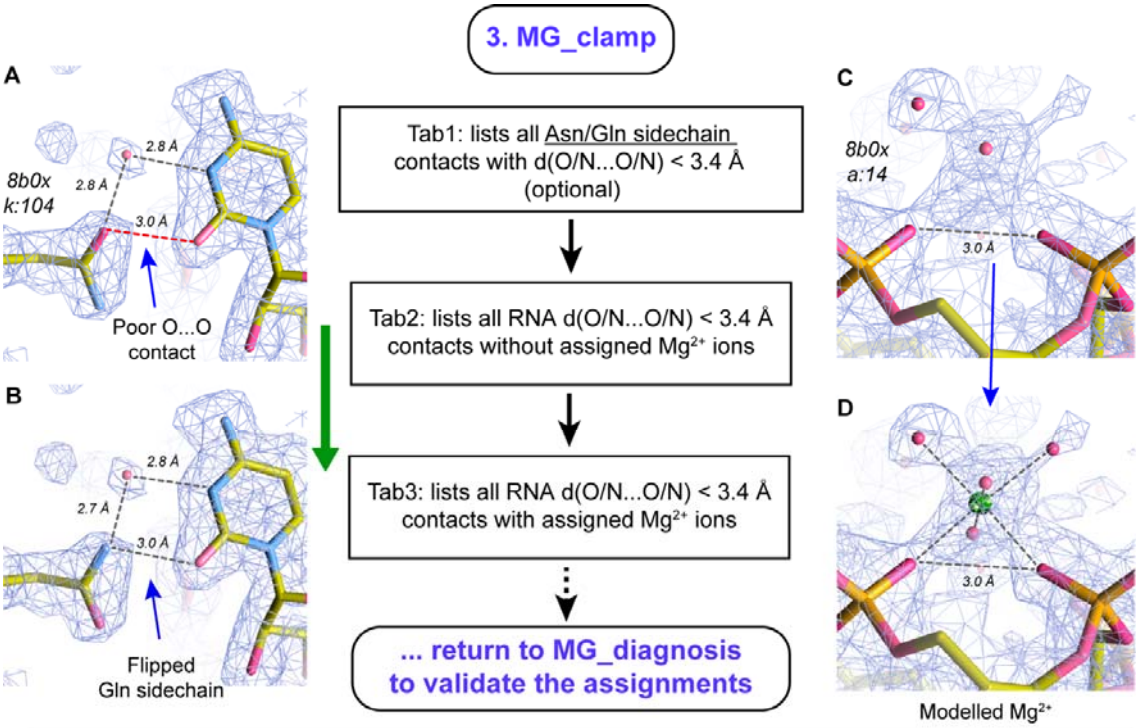


Figure 7. Organization of the *MG_detect_RNA* script. *MG_detect_RNA* identifies potential Mg^{2+} binding sites by collecting the highest non-assigned peaks from a masked map. (**A**) Shows a missing Mg^{2+} ion that is part of an $O_{ph}.5O_w$ binding site successfully identified by *MG_detect_RNA*. (**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.



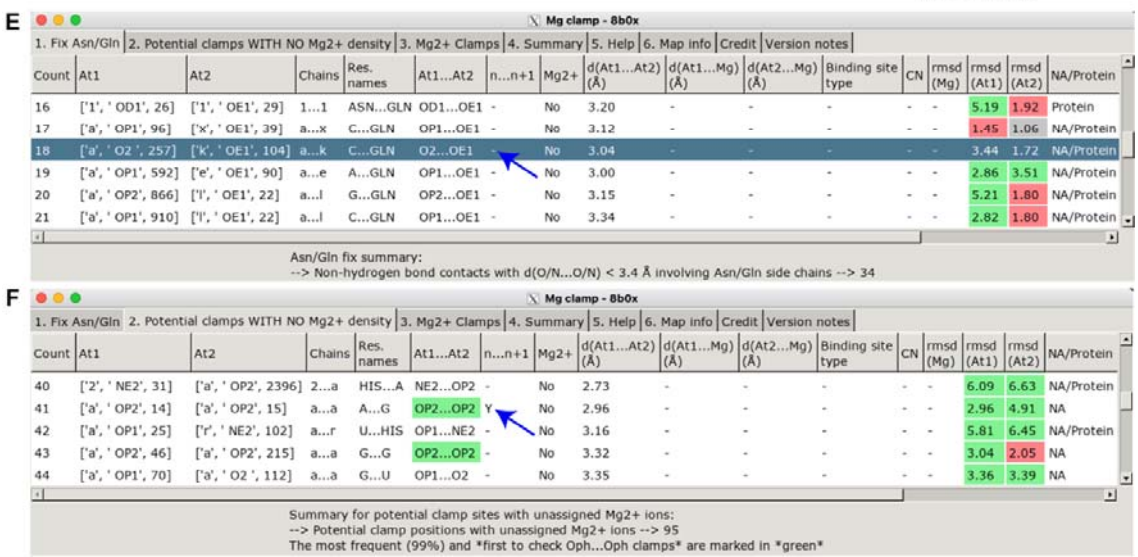


Figure 8. Organization of the *MG_clamp* script. *MG_clamp* locates non-hydrogen-bonded pairs of atoms with *d*(O/N…O/N) below 3.4 Å. These pairs are common to all higher-binding motifs (except *trans*-$2O_{ph}.4O_w$ and related; see **Figure 3**). When such pairs are identified, a visual inspection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

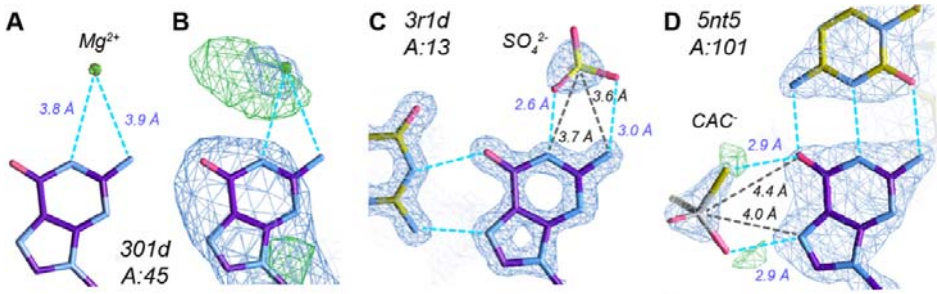


Figure 9. Anionic alternatives to questionable Mg^{2+} assignments. This figure presents examples of alternative explanations to questionable Mg^{2+} assignments. (**A/B**) In the **301d** hammerhead structure, the Mg^{2+} ion (Mg:45) is unconvincingly positioned near an electropositive Watson-Crick 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.

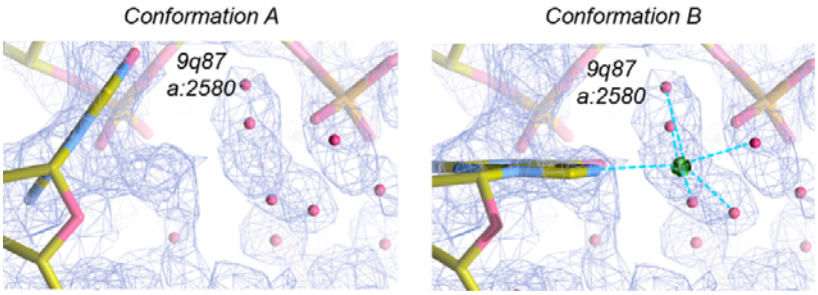


Figure 10. Alternative conformation of the a:2580 residue in 9q87. Conformation A shows the favoured orientation of a:G:2580. Conformation B shows an alternate conformation involving a well-defined N7.5O_w contact with Mg^{2+}.

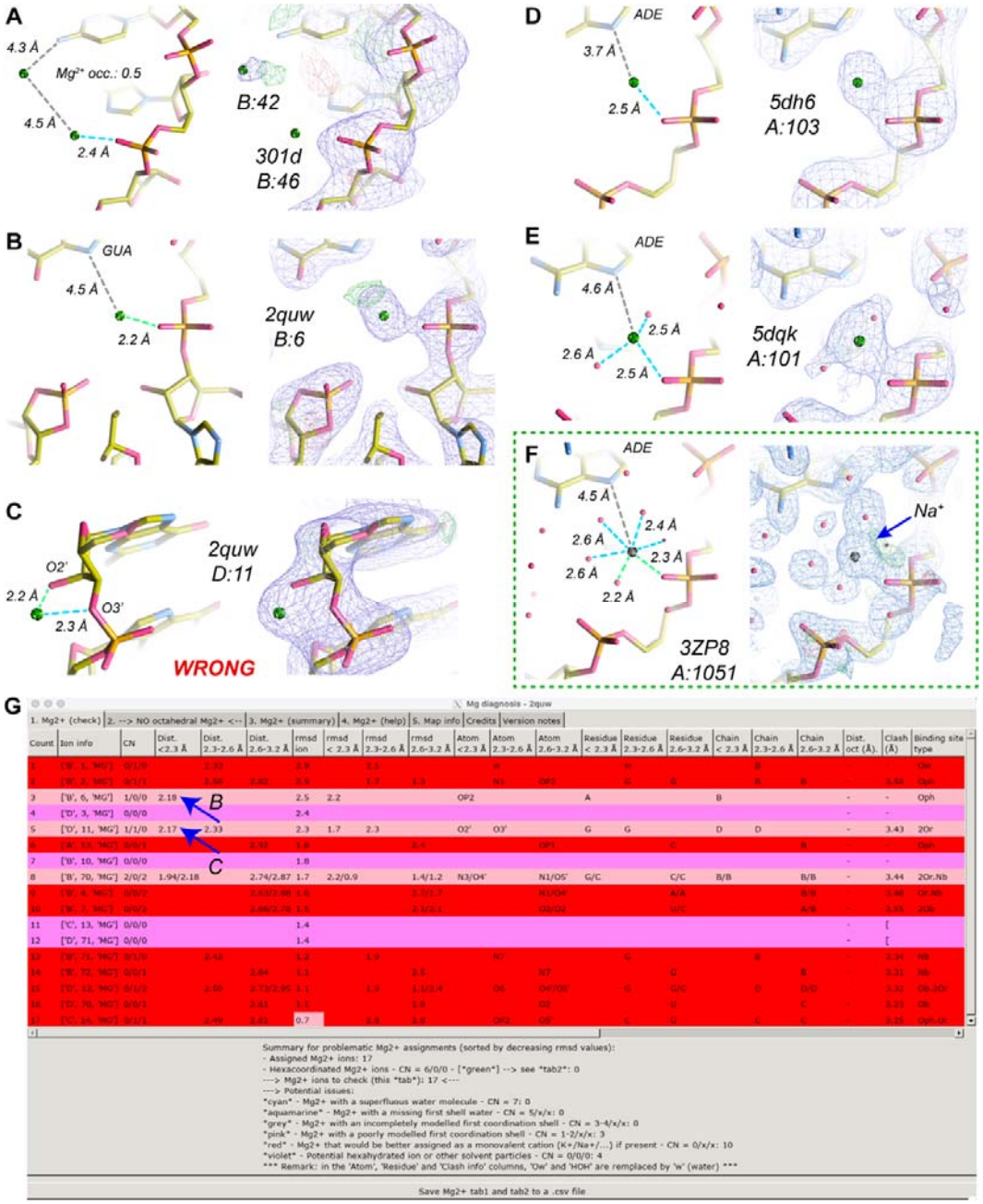


Figure 11. Ion binding assignments in hammerhead ribozyme X-ray structures. (**A**) Mg^{2+} ion tentatively assigned to a weak density spot in the 3.0 Å resolution **301d** structure (occupancy = 0.5). (**B**) A more realistic Mg^{2+} ion from the 2.2 Å resolution **2quw** structure, although lacking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

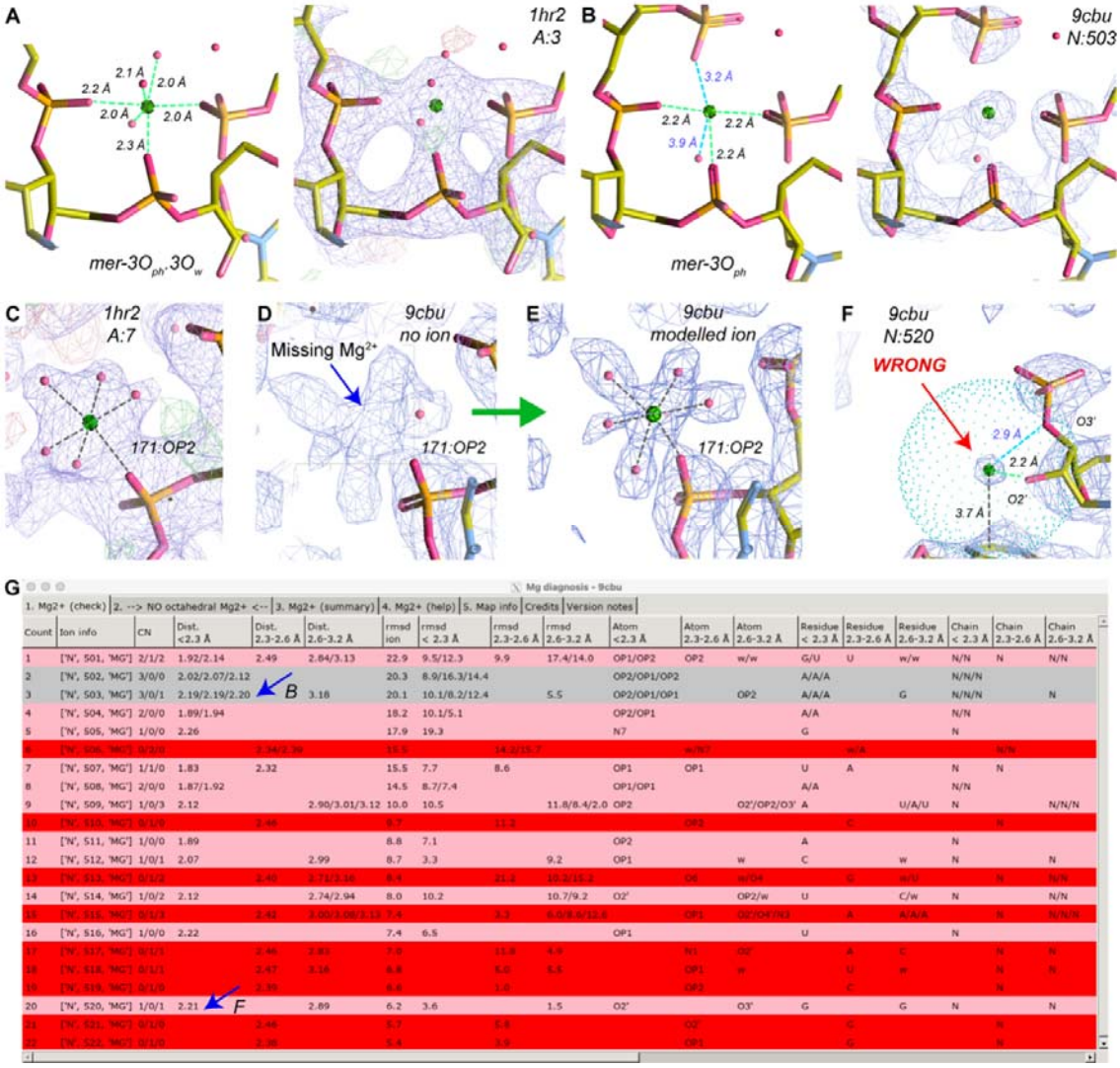


Figure 12. Ion binding assignments in two *Tetrahymena* group I intron structures: (**A**) A *mer*-3O$_{ph}$.3O$_w$ binding site in the 2.25 Å resolution **1hr2** X-ray structure (89), and (**B**) the equivalent site modelled in the 2.20 Å resolution **9cbu** cryo-EM structure (22). (**C**) An O$_{ph}$.5O$_w$ 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**).

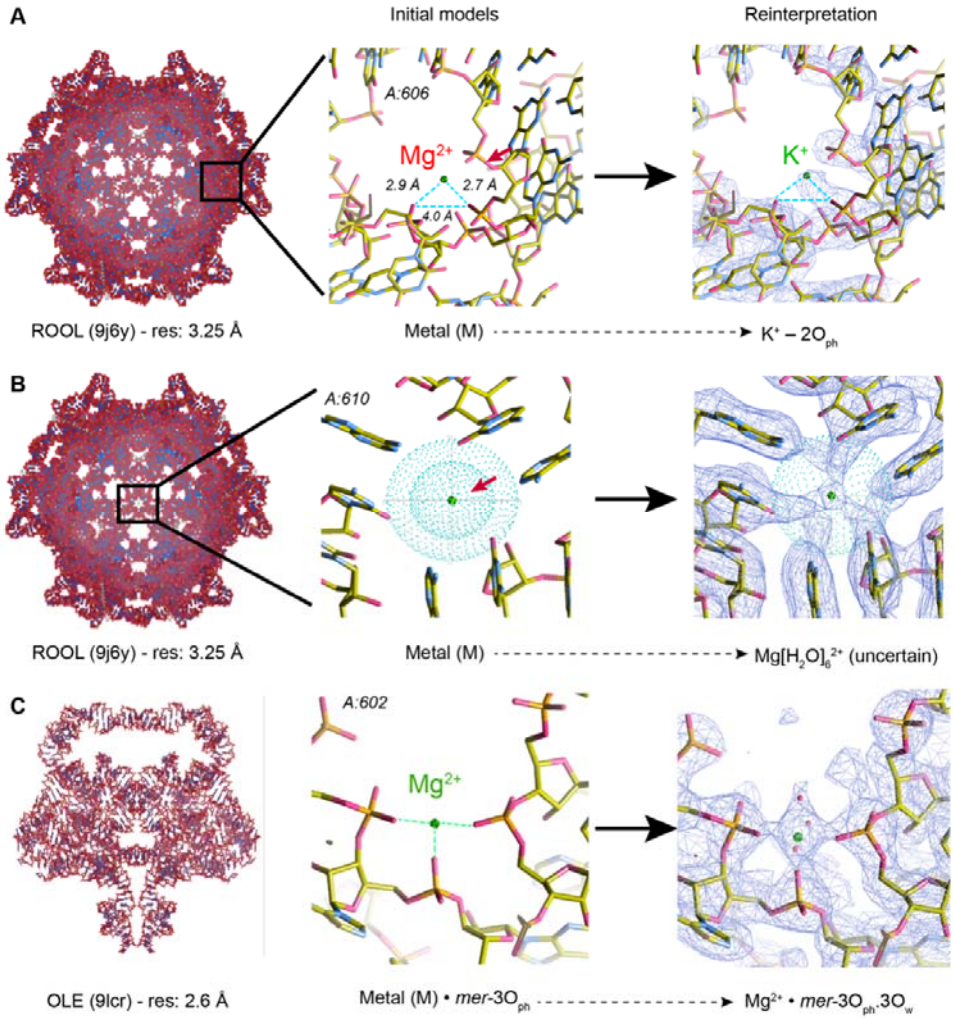


Figure 13. Ion placements in two low and one high-resolution quaternary structures. (A) In the **9j6y** structure (34), Mg^{2+} ions with coordination distances in the 2.6–2.9 Å range should be assigned as K^+ ions. **(B)** This ion is located on a three-fold symmetry axis. The associated non-octahedral 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.