Validation of nature and coordination of metals in protein structures by matching against the data in the CSD

Aim

To compare the structure of a protein-bound metal and its coordination shell, with similar coordination structures in the CSD. This analysis may be useful in identifying errors in metal assignment or in coordination geometry in X-Ray derived protein models

Introduction

The X-ray crystallographer faces many problems in assigning metal ions and the correct coordination of ligands around these ions. Consequently it is perhaps not surprising that a recent survey of metal ion environments in protein structures suggested that a substantial number of structures contained either an incorrect metal ion assignment or an unusual coordination pattern.[1] This study used metal-chelating atom bond-length information from the Cambridge Structural Database to compare with similar information derived from the PDB.

It would be useful for both crystallographers and molecular modelers to easily and quickly identify a problematic metal ion within a crystal structure and, where the element itself may be in error, suggest an alternative choice that fits the data better. We will use the Packing Feature Search tool in Materials Mercury to see if this tool can be used in this way. The example used has the PDB code *2as8.* This structure contains a magnesium ion which is reported by Zheng *et al*[1] to be very likely miss-assigned according to analysis of metal coordination bond lengths (Figure 1). They claim it should in fact be a calcium ion.

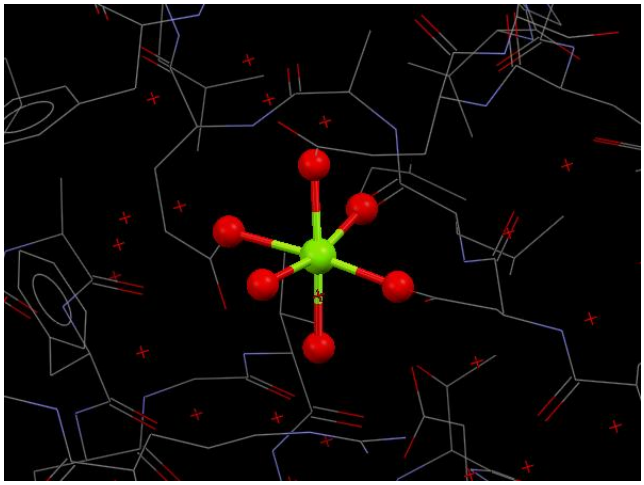

Figure 1 - Coordination sphere for purported Mg ion in *2as8.*

Method

The PDB file *2as8* is opened in Mercury. The magnesium ion (Mg2001), sits in an octahedral coordination sphere of oxygen atoms which, at first sight does not look atypical for cationic magnesium. Before carrying out the searches, the structure was edited so that the six coordinating oxygen atoms all were linked to the magnesium by a single bond. This is to ensure the bond typing is rendered in an equivalent manner to most of the coordination structures in the CSD. The seven atoms of the coordination sphere including the ion, are selected in the Packing Feature Search wizard of Materials Mercury. The search is carried out with no specification of number of hydrogens, or of the total number of atoms, bonded to each of the coordinating oxygen atoms. The quality of fit of the hit structures is set as 'low'. A strict set of quality filters was applied: 3D coordinates determined, factor <5%, not disordered, no errors, no catena bonds, no ions, no powder structures. CSD refcodes of those hits found, along with the Root Mean Square Deviations (RMSD) of hit atoms to the query atoms, were saved (166 hits in all).

A second search on the same set of atoms was carried out, except that the atom type of the magnesium atom was changed to calcium (37 hits retrieved in all).

A packing feature search was also carried out on the octahedral coordination sphere around the magnesium ion in CSD entry CIRVAA01 (166 hits retrieved) a randomly chosen high quality complex with six oxygens coordinating magnesium. Similarly a search was carried out on the octahedral coordination sphere around the calcium ion in CSD entry EFOJIS (37 hits retrieved).

Frequency histograms were created in Excel to compare the RMSD distributions of the two Mg and two Ca coordination sphere searches (Figures 2 and 3).

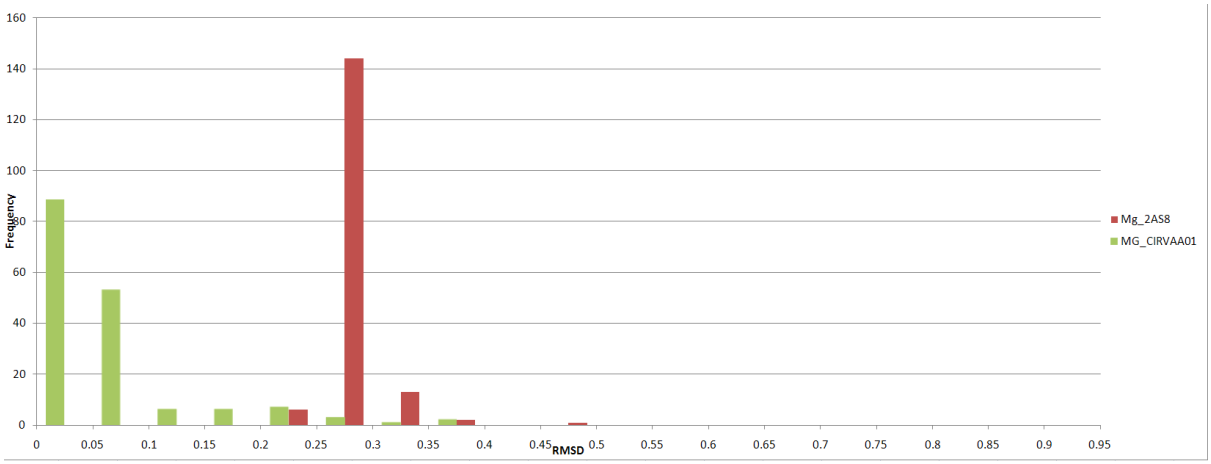

Figure 2 - RMSD distribution of matches in the CSD to Mg Coordination spheres in *2as8* (red) and CIRVAA01 (green).

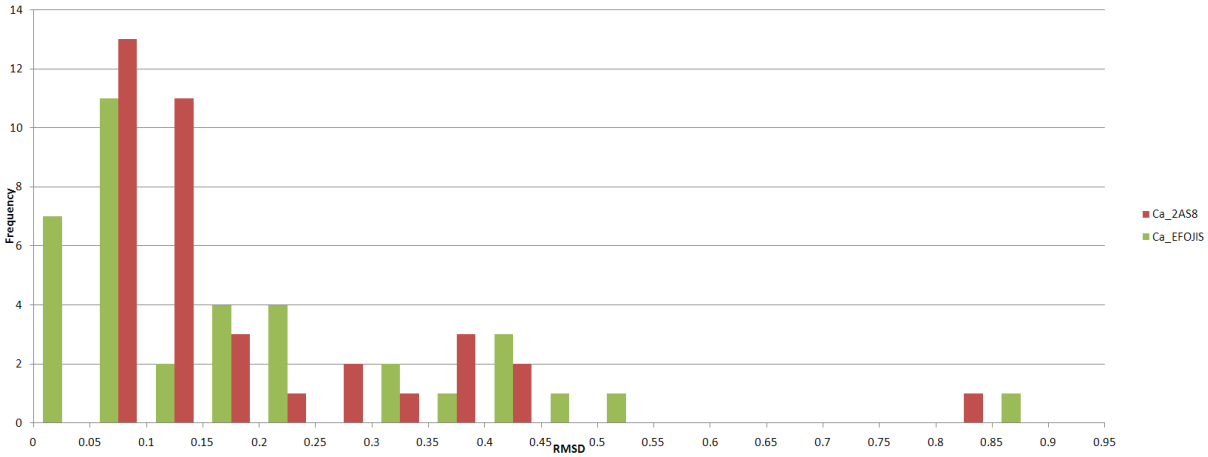

Figure 3 - RMSD distribution of matches in the CSD to Ca coordination spheres in *2as8* (red) and EFOJIS (green).

Results

Matches less than 0.05Å RMSD in Figures 2 and 3, represent very close matches to the query in all 7 matching atoms of the hit. Very close matches are found for the magnesium coordination sphere search from CIRVAA01 and the calcium coordination sphere search from EFOJIS, suggesting that these have geometries characteristic for the corresponding octahedral coordination spheres. The matches found in the two searches based on the PDB structure, *2as8*, do not match the hit structure as closely. The search where the metal ion is assigned as calcium gives a preponderance of matches at 0.10 to 0.15Å RMSD. The corresponding magnesium based search, on the other hand, has a sharp peak much further out at 0.25-0.30Å. This observation is consistent with the view that the assignment of this atom as a magnesium is incorrect, in agreement with the conclusions of Zheng *et al.*[1] Although the match of the two calcium based searches is not perfect, it is not bad. It should be considered that the resolution of the protein structure (1.95Å) may not be sufficient to expect a better match than this to be observed. In addition the configuration of the metal ligands may be slightly distorted by the constraints imposed by the protein tertiary structure.

Conclusions

Use of the Packing Feature tool in Materials Mercury appears to have value in the validation of metal assignment coordination sphere geometry in protein structures. Atom-atom mappings of coordination spheres onto similar spheres found in the CSD, allows a quick and easy assessment of validity and can also be used to investigate alternative metal assignments in cases of doubt.

References

1. H. Zheng, M. Chruszcz, P. Lasota, L. Lebida, W Minor, *J. Inorg. Biochem.*, **102**, 2008, 1765

Products

CSD – the world's only comprehensive, fully curated database of crystal structures, containing over 500,000 entries

Materials module of Mercury – a powerful exploration and comparison tool for solid state structures

Mercury – a versatile and feature-rich visualisation tool for molecular structures

For further information please contact Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK. Tel: +44 1223 336408, Fax: +44 1223 336033, Email: admin@ccdc.cam.ac.uk