Open access • Journal Article • DOI:10.1107/S002188980802476X

MESPEUS: a database of the geometry of metal sites in proteins

Kun-Yi Hsin, Y G Sheng, Marjorie M. Harding, P. Taylor ...+1 more authors

Institutions: University of Edinburgh

Published on: 01 Oct 2008 - Journal of Applied Crystallography (International Union of Crystallography)

Topics: Protein Data Bank and Protein Data Bank (RCSB PDB)

MESPEUS

Journal of
Applied Crystallography

ISSN 0021-8898

Received 17 May 2008
Accepted 1 August 2008

MESPEUS: a database of the geometry of metal sites in proteins

K. Hsin, Y. Sheng, M. M. Harding,* P. Taylor and M. D. Walkinshaw

Centre for Translational and Chemical Biology, University of Edinburgh, Michael Swann Building, Mayfield Road, Edinburgh EH9 3JR, UK. Correspondence e-mail: marjorie.harding@ed.ac.uk

A database with details of the geometry of metal sites in proteins has been set up. The data are derived from metalloprotein structures that are in the Protein Data Bank [PDB; Berman, Henrick, Nakamura & Markley (2006). *Nucleic Acids Res.* **35**, D301–D303] and have been determined at 2.5 Å resolution or better. The database contains all contacts within the crystal asymmetric unit considered to be chemical bonds to any of the metals Na, Mg, K, Ca, Mn, Fe, Co, Ni, Cu or Zn. The stored information includes PDB code, crystal data, resolution of structure determination, refinement program and *R* factor, protein class (from PDB header), contact distances, atom names of metal and interacting atoms as they appear in the PDB, site occupancies, *B* values, coordination numbers, information on coordination shapes, and metal–metal distances. This may be accessed by SQL queries, or by a user-friendly web interface which searches for contacts between specified types of atoms [for example Ca and carboxylate O of aspartate, Co and imidazole Nδ of histidine] or which delivers details of all the metal sites in a specified protein. The web interface allows graphical display of the metal site, on its own or within the whole protein molecule, and may be accessed at http://eduliss.bch.ed.ac.uk/MESPEUS/. Some applications are briefly described, including a study of the characteristics of Mg sites that bind adenosine triphosphate, the derivation of an average Mg—$O_{phosphate}$ distance and some problems that arise when average bond distances with high precision are required.

1. Introduction

Metal atoms occur in many proteins and may be essential for catalytic function, or for the maintenance of structure, or for as yet unidentified reasons. Metal sites can be well characterized by X-ray crystal structure determination. In the relational database MESPEUS (metal sites in proteins at Edinburgh University), we have assembled geometric information for ten biologically common metals from files in the Protein Data Bank (PDB; Berman *et al.*, 2006). The database now includes sites in crystal structures determined at resolution 2.5 Å or better, by diffraction methods, for the metals Na, Mg, K, Ca, Mn, Fe, Co, Ni, Cu and Zn, and in the PDB at 1 January 2007. There are 34 896 metal sites in 10 919 structures; nucleic acid structures are included, as well as proteins. Geometric information can be useful (i) in the interpretation and fitting of models to electron-density maps, (ii) in the validation of structures or in restrained structure refinement when only low-resolution diffraction data are available, and (iii) for defining a 'docking pocket' for virtual screening in structure-based drug design (see Taylor *et al.*, 2008). Furthermore, when a metal site has been identified in a new structure, related sites in known structures can be compared quickly. Na and K sites are included because, although their interactions with protein are better described as electrostatic interactions than as bonds, knowledge of the geometry of the sites can be useful. When they occur in metalloproteins, the metals V, Mo and W are normally present in anionic forms, such as MoO_4^{3-}, and often in association with cofactors; their interaction with the protein structure and its amino-acid side chains is quite different from that of the metals that are present as cations, and so they are not included in the database.

Previous work has extracted characteristic bond distances and metal-site geometry from the PDB (Harding, 2000, 2001, 2006) and made comparisons with small molecule structures in the Cambridge Structural Database (CSD; Allen, 2002). These publications did not give access to the individual observations. MESPEUS now contains these observations from structures in the PDB at 1 January 2007. Castagnetto *et al.* (2002) described a different database, Metallo-Scripps MDB, also available on the Web and concerned with metals in protein structures, but new material does not appear to have been added to this since 2004.

For each metal atom, interactions with surrounding atoms which are within chemically bonding distances were extracted using the program *MP* described by Harding (2000); this distance is the 'target' distance, already characterized (Harding, 2006), plus a tolerance, 0.75 Å, to allow for coordinate errors in structure determination. These 'donor' atoms in the metal coordination group are often O, N or S from functional groups in the amino-acid side chains of the protein, or of the main chain carbonyl group, but they may also be from water molecules, substrate analogues or other non-protein molecules present in the crystal. Thus, for each metal site, a coordination number and a description of the geometry could be derived and stored (see the next section for details of the information stored).

The web interface for this MESPEUS database provides the possibility of searching for different combinations of metal and donor atom and retrieving the distances. The mean distance (and standard deviation) can be evaluated; the maximum crystal structure resolution used can be restricted; individual metal sites can be examined by graphical display, with or without the surrounding protein structure.

2. Methodology

2.1. Details of information stored

Some information has been extracted directly from PDB files and some with the program *MP* (Harding, 2000, where a fuller description of some of the terms can be found). For each protein or nucleic acid structure in the PDB, determined by diffraction methods at a resolution of 2.5 Å or better, and containing one or more metal atoms, the PDB code is stored, together with the name of the protein, the class of protein (HEADER in PDB file), the data resolution (Å), the space group and cell dimensions, the refinement program used, the *R* factor, and $R_{\rm free}$. Within each structure, the names of all metal and donor atoms are stored as they appear in the PDB file; for metal atoms the coordination number, *B* value and information on the shape of the site (*e.g.* the average deviation from tetrahedral or square planar for sites with coordination number 4) are stored, and for donor atoms their distances from the metal atom, their occupancies and *B* values. The r.m.s. difference between the actual metal donor atom distances and the target values is stored, as an indicator of the quality of the geometry at each metal site. Metal coordination numbers are stored, as evaluated by the program *MP* from the PDB file. Symmetry-related atoms are not stored in the PDB and no attempt has been made in this version of the MESPEUS database to generate them; occasionally the metal coordination group should include atoms in neighbouring symmetry-related units, *e.g.* when the metal atom lies on a two-, three- or fourfold symmetry axis in the crystal, and in these cases the value found for the coordination number will be too low. Also stored are the E.C. numbers for enzymes (where given in the PDB), metal–metal distances, indicators for bidentate carboxylate groups, angles between bonds at the metal sites, and error information, *e.g.* the presence of disorder at the site. All the information can be accessed by SQL queries [Structured Query Language, see http://dev.mysql.com/doc/refman/5.0/en/ or Chamberlin *et al.* (1976)]. Further details of the tables are available at http://eduliss.bch.ed.ac.uk/MESPEUS/query_SQL.jsp.

2.2. The construction of the database

The Fortran program *MP* (Harding, 2001) examined each PDB file in turn and gave a log file summarizing the geometric data for each metal site found. Using Perl scripts these data were written to the database tables. Information not given in the log files (*e.g.* structure refinement program used, *R* factors) was extracted directly from the PDB files. In cases where two alternative positions for a side chain have been detected in the refinement and indicated by *A* and *B* as the last character of the atom names, the *B* atoms (with lower occupancy) have been removed and an error flag set for the associated metal atom. About one-tenth of the metal sites are affected in this way. (At other metal sites where a number of low-occupancy donor atoms are reported there may also be disorder.)

2.3. Construction of the web interface

The web interface of MESPEUS was established utilizing a Java web-based approach. In order to enable changes in functionality and for ease of maintenance, the web site was constructed using Model-View-Controller (MVC) software architecture. MVC separates the data model, user phase and control logic into three individual components so that modifications to one component can be made with minimal impact to the others. The web interface allows the client to set series of query criteria to access the MESPEUS database without requiring any knowledge of SQL. It displays the accessible information of metal coordination groups and displays the individual metal site with distances, angles and coordination geometry.

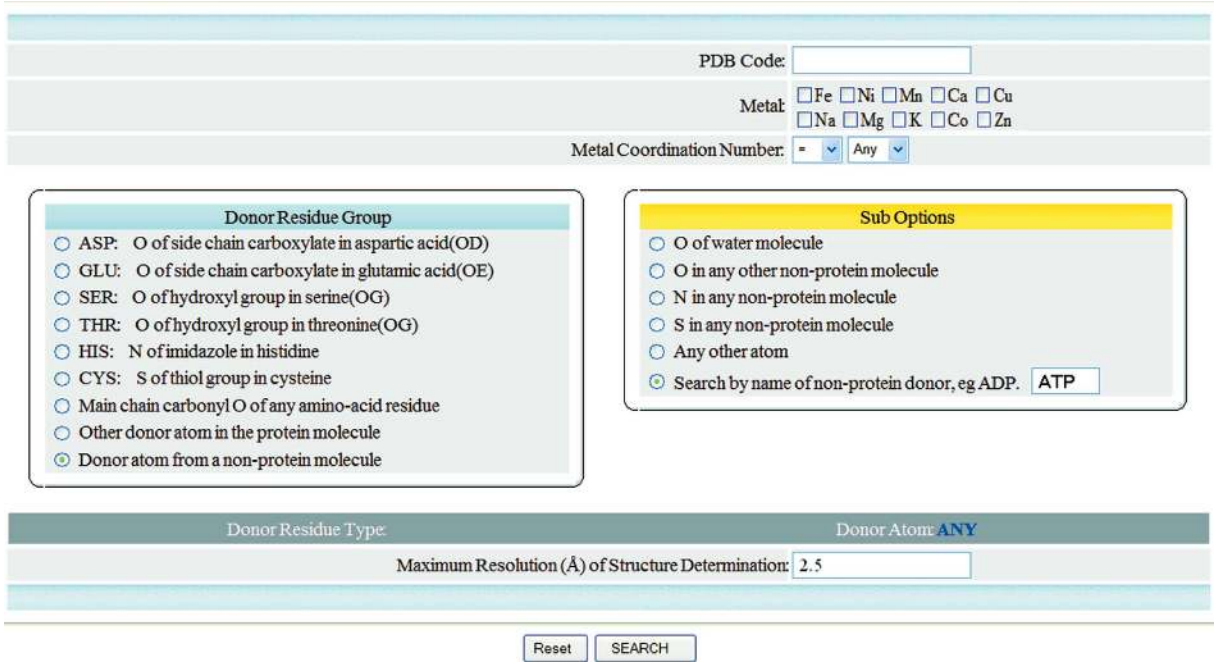


Figure 1
The search screen of the web interface, filled in for query about Mg links to adenosine triphosphate (ATP).

Table 1
Numbers of metal sites in the MESPEUS database.

Note that a significant number of metal atoms are listed in PDB files which do not appear to be within chemical bonding distance (target distance plus 0.75 Å) of any appropriate atoms; in many cases there are not even any appropriate atoms within 3.6 Å. This does not make chemical sense and can only be regarded as incomplete structure determination; these metal atoms are not included in the database.

Metal	(1) All sites in database	(2) Sites with metal and donor occupancy = 1.0	(3) Sites from column (2) with metal–protein interactions	(4) Sites from column (2) without metal–protein interactions
Na	2372	2069	1768	301
Mg	5561	4953	3676	1277
K	1424	1259	1108	151
Ca	7120	6425	6018	407
Mn	2118	1810	1688	122
Fe	7763	7226	7003	223
Co	637	506	356	150
Ni	534	398	366	32
Cu	1327	1105	1088	17
Zn	6031	5073	4979	94
Total	34 896	30 824	28 050	2774

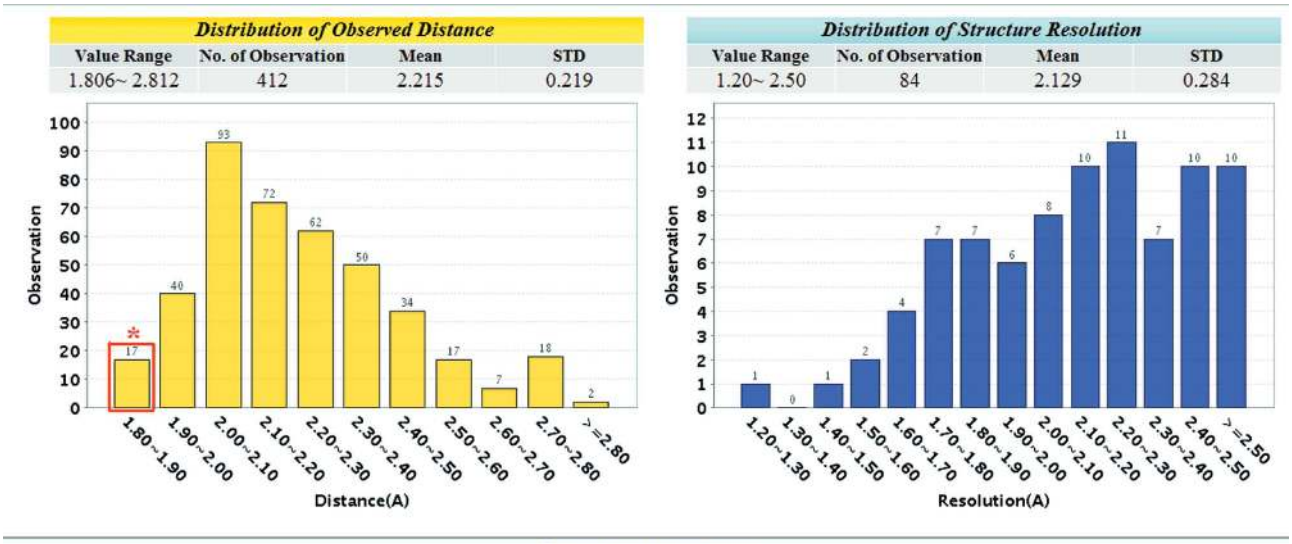
3. Contents of database

At present, the database contains material from 10 919 structures in the PDB, with numbers of metal sites as shown in Table 1. Some metal or donor atoms are listed with low occupancy, and in some cases there were pairs of disordered sites only one of which has been retained. Coordinate errors can be large when atom positions have low occupancy, and the presence of disordered sites gives meaningless results for coordination numbers. These sites can be identified in the database and excluded from searches, leaving the numbers shown in column (2) of Table 1 (and except where stated otherwise all the applications below use fully occupied atom positions). Not all the metal sites interact directly with protein. The number of these is particularly large for Mg [see Table 1, column (4)]. Some are simply hydrated ions such as $Mg(OH_2)_6^{2+}$; others interact only with ligand molecules, or with RNA or DNA (about 300 of the PDB files used for our database describe RNA or DNA structures without protein).

When more than one copy of the protein molecule with metal site(s) is present in the crystal asymmetric unit, all are listed in the database – for metal to donor atom distances they should represent separate observations. For this reason the number of distinct metal sites is only *circa* 60% of the total number of sites listed. Furthermore, the PDB contains many groups of very similar proteins within which the metal sites are similar or identical.

3.1. The web interface

The MESPEUS web interface has been designed to allow straightforward searching for particular kinds of interactions, with the possibility of selecting only higher-resolution structures, as shown in Fig. 1. Alternatively, a PDB code can be given to yield all the metal sites in that protein. The first result is a list of all the metal–donor



Download Query Result
Number of Hits: 412

No.	Distance	Coordination Number	Shape	Metal Name	Donor Residue Name	PDB	Reslu.	r.m.s.d.(Å)	Difference	Delete
1	2.370	6	δ oct(16.1°) **	MG 534 A	O2G ATP A 535	1A49	2.10	0.156	0.300	☐
2	2.011	5	δ tbp(21.7°) or δ tetp(12.1°)	MG 536 A	O1G ATP A 535	***		0.151	-0.059	☐
3	2.330	5	δ tbp(21.7°) or δ tetp(12.1°)	_	O2B ATP A 535	_			0.260	☐
4	2.013	5	δ tbp(21.7°) or δ tetp(12.1°)	_	O1A ATP A 535	_			-0.057	☐
5	2.420	5	δ tbp(27.3°) or δ tetp(32.1°)	MG 1734 C	O2G ATP C 1735	_		0.172	0.350	☐
6	2.081	6	δ oct(11.4°)	MG 1736 C	O1G ATP C 1735	_		0.088	0.011	☐
7	2.202	6	δ oct(11.4°)	_	O2B ATP C 1735	_			0.132	☐
8	1.920	6	δ oct(11.4°)	_	O1A ATP C 1735	_			-0.150	☐
9	2.272	6	δ oct(10.8°)	MG 2334 D	O2G ATP D 2335	_		0.110	0.202	☐

Figure 2
Results of the search for Mg linked to ATP; this yielded 412 examples in 84 proteins or DNA/RNA structures. Average Mg—O distances and their distributions can be displayed. *: Clicking here selects this distance range only. **: Links to a page like Fig. 3 showing details of the specified metal site. ***: Links to a page like Fig. 4 showing the whole metal-containing protein.

atom distances satisfying the search query, together with information on coordination number of the metal, shape, atom names *etc.* as shown in Fig. 2. Mean distances and distributions can be obtained, or the list may be downloaded for other use. Any metal site selected can be displayed graphically and its position in the whole protein structure shown (Figs. 3 and 4).

4. Applications of database and web interface

4.1. The importance of near atomic resolution in deriving mean bond distances

We have been concerned to derive the best average values for distances from metal atoms to different kinds of donor atoms in structures deposited in the PDB, for use in validation, electron-density map interpretation, model fitting or restrained refinement. The average bond distance can be evaluated from all the data in the database, or from the higher-resolution structure determinations only. We and others (*e.g.* Meyer Klaucke, 2007) have noticed that often the average found for a particular distance, for example Zn—N of histidine with coordination number 4, shown in Table 2, is slightly smaller when only high-resolution structure determination results are used than when all results are used. The standard deviation (s.d.), representing the scatter of results, is, as expected, smaller too.

The most precise averages, *i.e.* those with smallest standard deviation, are those for the highest-resolution results, better than 1.3 Å at least, *i.e.* near atomic resolution; these agree best with the CSD average for equivalent interactions, 2.00 (2) Å for 25 observations (Harding, 2006). Fig. 5 shows the wide scatter of reported distances from poor-resolution structures; it is only at fairly high resolution that there is good consistency of observations. On chemical grounds very little variation in this particular distance is expected (say < 0.02 Å); nearly all the variation above must be due to coordinate errors remaining after structure refinement. (Other kinds of distance, *e.g.* M—$O_{carboxylate}$, can show more chemical variation.) It is strongly recommended that for such averages only the highest-resolution results practicable be used, consistent with a reasonable number of observations.

Table 2
Effect of resolution on mean distance found for Zn—N of histidine, with Zn coordination number 4 (Nδ and Nε of histidine are both included; if they are kept separate, no significant difference in their means is found).

Maximum resolution (Å)	No. of observations	Mean (s.d.) (Å)
2.5	3583	2.10 (13)
1.8	1007	2.07 (9)
1.5	261	2.06 (6)
1.3	88	2.03 (4)
1.1	21	2.02 (4)

4.2. Application: a survey of interactions between Mg and adenosine triphosphate (ATP)

Phosphate–magnesium interactions are crucial in many enzyme mechanisms. There is a particularly rich source of structural infor-

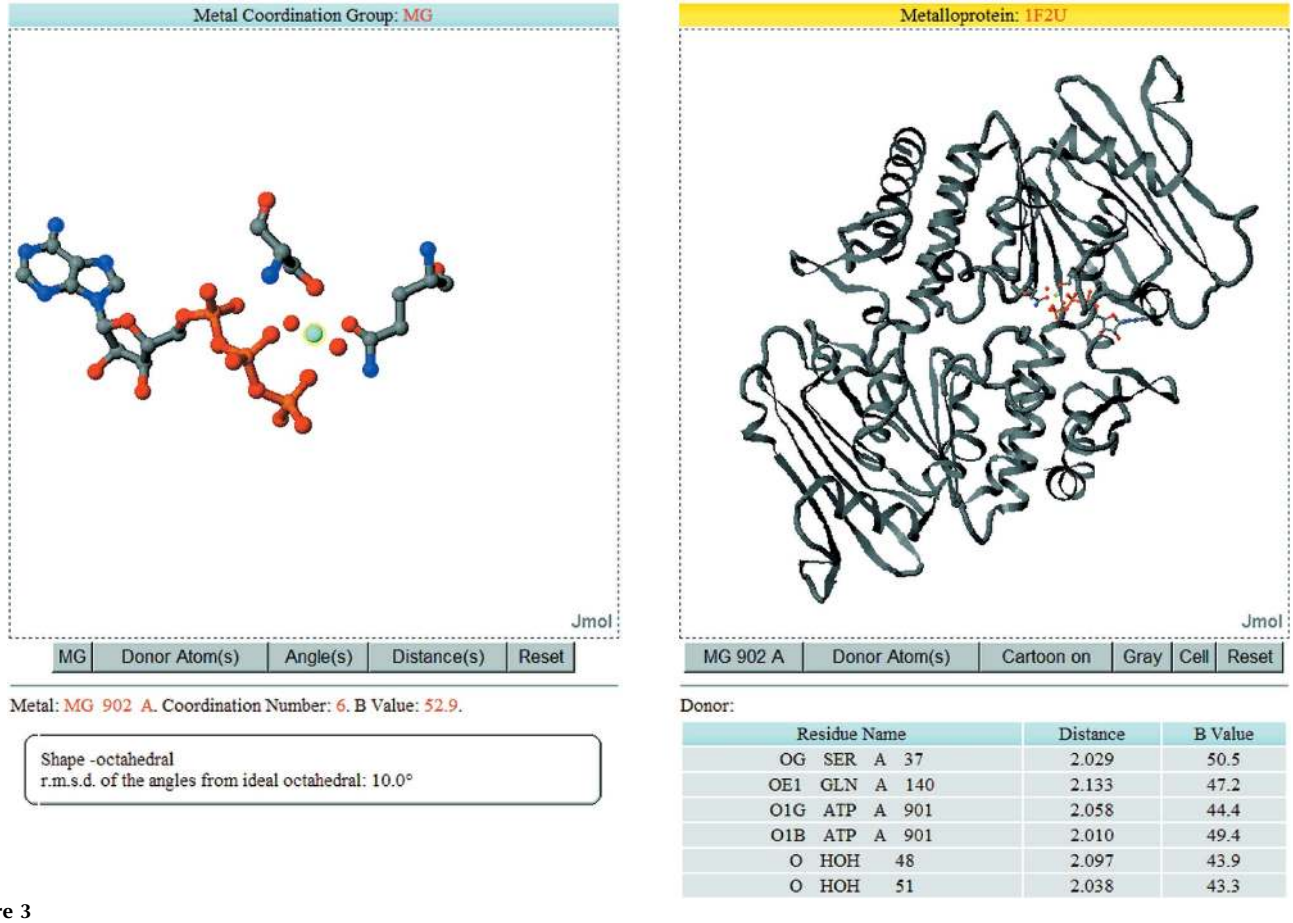
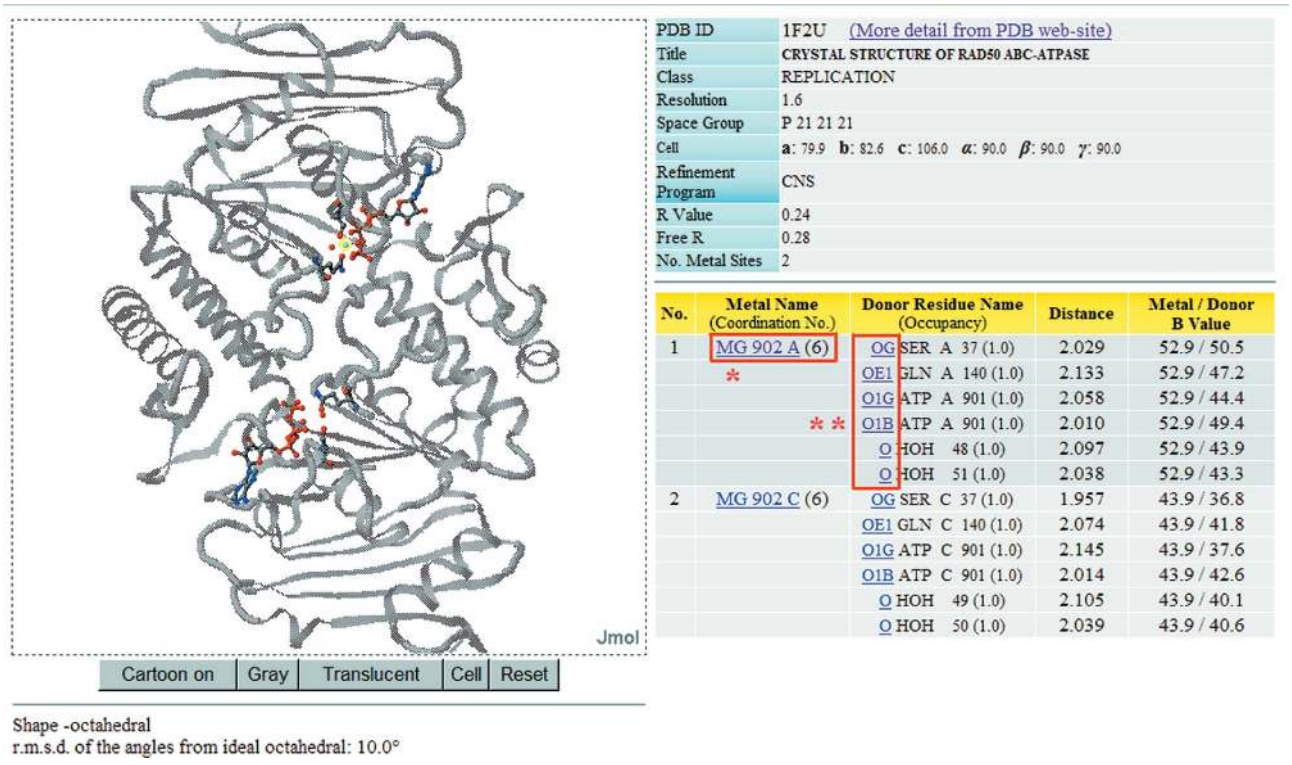


Figure 3
By clicking on the metal name in the results shown in Fig. 2, the metal site can be displayed by *Jmol* (http://jmol.sourceforge.net/), as well as its relation to the whole protein molecule. Here, the metal, Mg (cyan coloured), is coordinated by the O atoms from Ser and Gln, two phosphate O atoms (one β, one γ) of ATP, and two water molecules. The buttons below the pictures allow users to centre and rotate the view and add other information.



PDB ID	1F2U (More detail from PDB web-site)
Title	CRYSTAL STRUCTURE OF RAD50 ABC-ATPASE
Class	REPLICATION
Resolution	1.6
Space Group	P 21 21 21
Cell	a: 79.9 b: 82.6 c: 106.0 α: 90.0 β: 90.0 γ: 90.0
Refinement Program	CNS
R Value	0.24
Free R	0.28
No. Metal Sites	2

No.	Metal Name (Coordination No.)	Donor Residue Name (Occupancy)	Distance	Metal / Donor B Value
1	MG 902 A (6)	OG SER A 37 (1.0)	2.029	52.9 / 50.5
	*	OE1 GLN A 140 (1.0)	2.133	52.9 / 47.2
		O1G ATP A 901 (1.0)	2.058	52.9 / 44.4
	**	O1B ATP A 901 (1.0)	2.010	52.9 / 49.4
		O HOH 48 (1.0)	2.097	52.9 / 43.9
		O HOH 51 (1.0)	2.038	52.9 / 43.3
2	MG 902 C (6)	OG SER C 37 (1.0)	1.957	43.9 / 36.8
		OE1 GLN C 140 (1.0)	2.074	43.9 / 41.8
		O1G ATP C 901 (1.0)	2.145	43.9 / 37.6
		O1B ATP C 901 (1.0)	2.014	43.9 / 42.6
		O HOH 49 (1.0)	2.105	43.9 / 40.1
		O HOH 50 (1.0)	2.039	43.9 / 40.6

Figure 4
By clicking on the protein name in the results shown in Fig. 2, information about the protein is tabulated, as well as details of each metal site; the whole protein molecule is displayed by *Jmol*. Here the protein has two chains, *A* and *C*, with one Mg–ATP site on each. *: A dynamic link to centre this metal site in the view on the left. **: Links to a page like Fig. 3 showing the specified metal site.

mation available in the PDB for metal–ATP interactions and this survey provides an overview of the interaction geometries found from the database for protein–Mg–ATP complexes. The web interface can find all links from Mg to O of ATP. For the detailed analysis, it was more convenient to use one SQL query to find all Mg with ATP links, and a second to pick up all the other atoms, protein or otherwise, linked to each Mg; Perl scripts were then used to present the results more conveniently. This yielded 136 Mg sites from 46 proteins; transferases were the commonest protein types. After removal of identical sites within crystal asymmetric units, 66 sites remained and are listed, together with some statistics, in Table D (deposited[1]). In these, Mg is normally linked to protein through one or two amino-acid side chains, most commonly Asp, Glu, Asn, Gln, Ser and Thr, but there are two examples with three links to protein, and three with links to main chain carbonyl oxygen. ATP may be linked to Mg through one, two or all three phosphate groups; the commonest pattern (23 examples) is linkage through the β- and γ-phosphate groups, but all other possibilities occur, although there are only two cases where linkage is through two O atoms of the same phosphate group. The Mg coordination group may also include one to four water molecules, and occasionally another small molecule like oxalate (as a bidentate ion). More than half have total coordination number 6, the expected coordination number for Mg. A surprisingly large number, 20, appear to have coordination number ≤ 4; this suggests that the reliability of some of the data is questionable – some of the analyses may be incomplete, or some of the atoms identified as Mg may in fact be water molecules. Figs. 3 and 4 show an example of an Mg site with ATP. Seven observations in structures at resolution ≤ 1.5 Å give a mean Mg$-$O$_{\rm phosphate}$ distance of 2.05 (7) Å, but for a reliable average more observations are desirable.

4.3. Application: Mg interactions with other phosphates

To establish an Mg$-$O$_{\rm phosphate}$ distance from a larger number of higher-resolution observations, it was necessary to include other phosphate ligands. Identifying phosphates from the atom names in the PDB is not straightforward, but by selecting with SQL queries O atoms whose residue names (het group names, see 'het groups' section of PDBSUM, http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) include the letter P and whose atom names include one of the letters A, B or G, or O atoms whose atom names include the letter P, over 2000 Mg$-$O$_{\rm phosphate}$ links were found, 75 of which are from

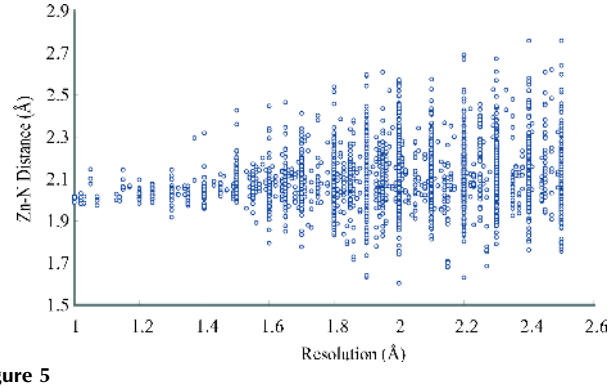


Figure 5
Reported values of Zn $-$N$_{\rm his}$ distances for Zn coordination number 4, as a function of resolution.

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: KK5027). Services for accessing these data are described at the back of the journal.

structures with resolution 1.3 Å or better. These have an average Mg—O distance of 2.06 (8) Å, in good agreement with the value above [and also close to the target distance given for $Mg—O_{carboxylate}$ (Harding, 2006)]. The commonest links found are to adenosine diphosphate (ADP; 453 links to Mg), ATP (410), atrial natriuretic peptide (ANP; 299), and guanosine disphosphate (GDP; 219), imidotriphosphate (GNP; 195) and triphosphate (GTP; 113). It is again clear that linkage through two O atoms of one $–PO_4–$ group is very rare, and that for the triphosphates ATP, GTP and CTP (cytidine) the linkage is often bidentate, through the β- and γ-phosphate groups, not the α group.

4.4. Application: Mg compared with Mn and Co

DNA polymerases, RNAses, transposases and integrases all require one or two Mg ions as an integral part of their enzymatic mechanism (Steitz & Steitz, 1993). The effect of substituting Mn or Co for Mg frequently results in modified activities, either partial or complete inhibition of activity or in some cases enhancement of activity and broadening of substrate specificity (Frank & Woodgate, 2007). It may be possible to relate at least some of these differences in enzyme activity to differences in coordination geometry of the bound metal in the active site. Mg and Co divalent ions are very similar in size, while Mn^{2+} is significantly larger, *e.g.* Mg—O and Co—O are ∼2.07 Å while Mn—O is ∼2.17 Å. They have somewhat different preferences for amino-acid donors; Co and Mn are much less likely than Mg to interact with serine, threonine or main chain carbonyl oxygen, and more likely to interact with histidine [see http://tanna.bch.ed.ac.uk/, item 4, '..frequency of occurrence..', or Dokmanić *et al.* (2008)]. Here, coordination numbers (Table 3) and shapes have been compared in structures with resolution 1.8 Å or better. For each metal, more than half of the metal sites have coordination number 6, and the next most important coordination number is 5. Coordination numbers higher than 6 are almost always associated with multidentate ligands, and coordination numbers less than 4 or 5 may represent incomplete structure determinations, or cases where additional donor atoms are present but are in neighbouring asymmetric units not listed in the PDB files, or, particularly in the case of Mg, unreliable data in the PDB, because of the difficulty of locating Mg or distinguishing it from O of water. When the coordination number is 6 the shape is octahedral, but the average distortion from ideal octahedral is considerably less for Mg, 7 (5)°, and Co, 8 (8)°, than for Mn, 11 (6)° [see Harding (2000) for a discussion of distortions]. When the coordination number is 5 the ideal shapes may be described as trigonal bipyramidal or tetragonal pyramidal; real metal sites are often between these two, and the description is based on which ideal version it is least distorted from. For Mg the tetragonal pyramid is strongly favoured (more than 95% of observations) with average distortion 8 (6)°; for Mn and Co there are near equal numbers of each shape with greater distortions. (All this information can be gathered from repeated queries of the web interface and some manual averaging, or it can be found very efficiently with SQL queries of the database.)

4.5. Application: listing of metal coordination groups

These lists were described by Harding (2004). Metal coordination groups (metal sites) were categorized by the sequence of amino-acid donors and the relative positions of these amino acids in the polypeptide chain, using one-letter codes for the amino acids, and the differences between the successive residue numbers. For example, CHCC Zn 2 18 3 describes a coordination group in which Zn is coordinated to the thiolate S of Cys(n), where n is the amino-acid residue number, an imidazole N of His(n + 2), and the thiolate S atoms of Cys(n + 2 + 18) and Cys(n + 2 + 18 + 3). Coordination number and information on other non-protein donors present are also given. The new database has allowed the preparation of new lists including all metal sites in the PDB at January 2007. For each metal atom in the database, the full names of all its donor atoms were extracted with an SQL query; Perl scripts were used to convert these to one-letter amino-acid codes, check that all donor atoms have site occupancy > 0.7, work out sequence separations *etc.*, as well as remove duplicate coordination groups within the asymmetric unit of any one crystal. The lists are available at http://tanna.bch.ed.ac.uk (item 3, 'New lists..'). They allow quick identification of other proteins with coordination groups identical or similar to a specified one, and could provide a resource for the study of patterns of metal coordination and protein evolution. A direct link to this database is planned.

Table 3
Coordination numbers found for Mg, Mn and Co in structures determined at resolution 1.8 Å or better.

	1	2	3	4	5	6	7	8	9
Mg	42	84	91	101	195	956	25	2	1
Mn	3	6	8	26	78	246	19	3	
Co	2	5	3	11	52	131			

5. Conclusions, summary

The database, with its user-friendly web interface, allows immediate identification and display of the metal sites in any specified protein whose structure is in the PDB, determined at resolution 2.5 Å or better. Alternatively, the web interface can identify all metal sites with a particular kind of contact, giving distance, coordination number and atom names for each, as well as average distances and distribution of distances. Furthermore, as shown in several applications, SQL queries to the database can extract additional information about the interactions of different metals with proteins.

References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Berman, H., Henrick, K., Nakamura, H. & Markley, J. L. (2006). *Nucleic Acids Res.* **35**, D301–D303.
Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A. & Pique, M. E. (2002). *Nucleic Acids Res.* **30**, 379–382.
Chamberlin, D. D., Astrahan, M. M., Eswaran, K. P., Griffiths, P. P., Lorie, R. A., Mehl, J. W., Reisner, P. & Wade, B. W. (1976). *IBM J. Res. Dev.* **20**, 560–575.
Dokmanić, I., Šikić, M. & Tomić, S. (2008). *Acta Cryst.* D**64**, 257–263.
Frank, E. G. & Woodgate, R. (2007). *J. Biol. Chem.* **282**, 24689–24696.
Harding, M. M. (2000). *Acta Cryst.* D**56**, 857–867.
Harding, M. M. (2001). *Acta Cryst.* D**57**, 401–411.
Harding, M. M. (2004). *Acta Cryst.* D**60**, 849–859.
Harding, M. M. (2006). *Acta Cryst.* D**62**, 678–682.
Meyer Klaucke, W. (2007). Personal communication.
Steitz, T. A. & Steitz, J. A. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 6498–6502.
Taylor, P., Blackburn, E., Sheng, Y. G., Harding, S., Hsin, K. Y., Kan, D., Shave, S. & Walkinshaw, M. D. (2008). *Br. J. Pharmacol.* **153**, S55–S67.