Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The geometry of metal–ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal–donor interactions

Marjorie M. Harding

Institute of Cell and Molecular Biology, University of Edinburgh, Michael Swann Building, Edinburgh EH9 3JR, Scotland

Correspondence e-mail: marjorie.harding@ed.ac.uk

Received 6 January 2000
Accepted 17 April 2000

Geometrical data which could be of relevance in the structure determination, structure refinement, assessment or understanding of metalloproteins have been extracted from the Cambridge Structural Database (CSD). The CSD contains crystallographic data from 'small-molecule' structures determined by X-ray or neutron diffraction to an accuracy and precision much better than that of most current protein structure determinations. Structures of Mg, Mn, Fe, Cu and Zn complexes with ligands whose donor atoms may be only N, O, S or Cl have been selected and analysed in terms of the geometry of the metal coordination group – octahedral, tetrahedral, tetragonal pyramidal *etc*. The r.m.s. deviation of all the interbond angles around the metal atom provides a measure, δ, of the deviation from ideal geometry. Average values of δ are tabulated for the different metals in each type of complex. For simple non-chelated complexes of Mn, Fe and Zn, distortions of up to 5° in octahedral complexes and 10° in tetrahedral complexes are found to be normal and seem likely to be a consequence of packing effects, ligand bulk or intramolecular effects. Substantially larger distortions are found for some other metals and geometries and are common for chelated complexes. Brief comments on six-, seven- and eight-coordinate Ca complexes are included. Tables are also presented showing that for four- and five-coordinate complexes of Zn and Cu it is quite common to find additional weakly coordinated ligands, usually with N or O donor atoms and with $M\cdots$N,O distances longer than a normal bond length but shorter than a van der Waals contact, *e.g.* in the range 2.4–3.0 Å for Zn and 2.6–3.0 Å for Cu. Although the contributions to bond valency or bonding energy of such interactions may not be great, their effect on geometry can be considerable; they can, for example, cause much larger distortions of tetrahedral Zn complexes than indicated above.

1. Introduction

Metalloproteins occur widely and have many important functions. In some, the metal atom or ion is a part of the active site for a catalytic process; in others, the metal appears to have a role in maintaining structure. The present studies are concerned with the geometry of the interaction of a metal atom with the ligand groups around it, using the accurate information which can be obtained about this geometry from the Cambridge Structural Database (CSD; Allen & Kennard, 1993*a*,*b*). In the previous paper (Harding, 1999), tables giving interatomic distances and some details of ligand geometry were assembled for the metals Ca, Mg, Mn, Fe, Cu and Zn in their complexes with water molecules, carboxylate groups, imidazole groups *etc*., ligands which are analogues of the

amino-acid side chains available in proteins. The objective of that study was to make conveniently available to protein crystallographers accurate information on preferred bond distances for a selection of metals and ligands which commonly occur in proteins. This information could be of use in the interpretation and fitting of models to electron-density maps calculated with limited resolution data, for target distances in restrained refinement or in the validation of protein structural data (but not of course all three at once!). It could also be hoped that the analysis would contribute to the basic understanding of the function of different metals in metalloproteins.

It seemed desirable to follow this with a similar analysis of the angles around the metal atom in its complexes to indicate, for example, the extent to which ML_6 deviates from regular octahedral geometry, where L is a ligand donor atom and the ligand may consist of one atom or a group of atoms. The present analysis deals mainly with Mg, Mn, Fe, Cu and Zn as ML_6, ML_5 and ML_4, but some briefer comments on Ca are included in six-, seven- and eight-coordinate complexes. The ligand donor atoms are restricted to N, O, S or Cl; Cl is included, although it is probably not biologically very important, because of its resemblance to S in donor properties. The other aspect of metal–ligand geometry addressed here is that of weak interactions between metal atoms and N or O donor atoms in ligands, interactions that appear to be substantially longer than the usual M—O or M—N single bonds but significantly shorter than a van der Waals contact. A few examples of the occurrence of such interactions were given in the previous paper, but it proved essential to examine them more systematically because of their effects on the stereochemistry at the metal. Others have also commented on these interactions (Auf der Heyde & Burgi, 1989) and Auf der Heyde & Nassimbeni (1984) point out that zinc shows a smooth progression from four- to five-coordinate states.

A variety of different indicators of distortion from regular polyhedral geometry have been used by others. The indicator of Zabrodsky *et al.* (1993) is appropriate to any polyhedron and is the mean-square displacement of all the vertices from the nearest regular polyhedron. Howard *et al.* (1998) have used a measure Rc(χ) which measures angular distortions only, as a percentage of their mean values. The measure δ used here is very similar to Rc(χ) – it is the r.m.s. deviation of the angles from those in the regular polyhedron and is directly proportional to Rc(χ); the constant of proportionality is the r.m.s. angle at the metal atom in the ideal geometry (tetrahedral, 109.5°; square planar, 127.3°; trigonal bipyramidal, 111.4°; octahedral, 113.8°) and so is slightly different for different ideal geometries. Thus δ and Rc(χ) deal only with angular distortions, whereas the measure of Zabrodsky *et al.* (1993) also includes variations in the distances of ligand donor atoms from the central metal atom. Further, in evaluating δ or Rc(χ) no preliminary fitting (*i.e.* reorientation) of the observed polyhedron to an ideal one is performed; if this were performed, as it is by Zabrodsky *et al.* (1993), and then r.m.s. deviations of bonds from the ideal directions used for δ or Rc(χ), their values would be numerically a little smaller. Some more extensive studies on selected metals have analysed distortion in relation to the symmetry elements of ML_4 or ML_5

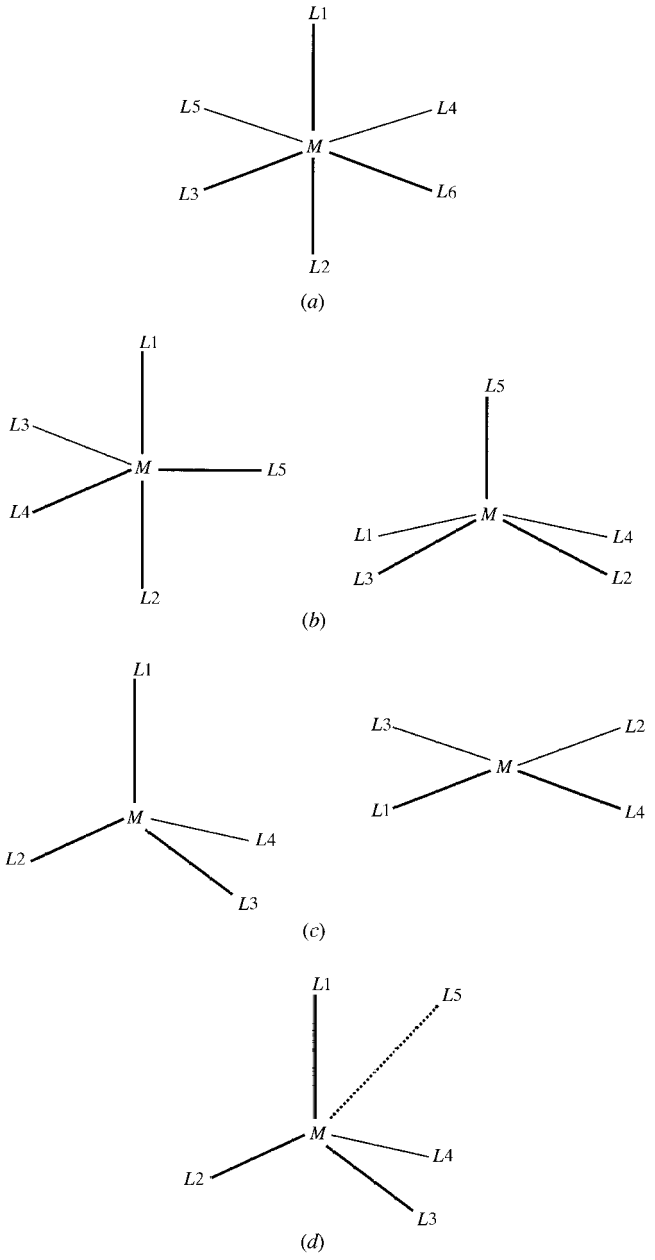

Figure 1
Search fragments. L_1, L_2 *etc.* may be N, O, S or Cl. Bonds are of type 'any'. Other atoms or groups may be connected to L but not to M. The complex is described as unchelated if M is an acyclic atom and chelated if it is a cyclic atom. a_{ij} is the angle L_i–M–L_j. (*a*) M has total coordination number 6 and the labels are assigned so that a_{12} is the largest of the 15 angles and a_{34} is the next largest. (*b*) M has a total coordination number 5 and the labels are assigned so that a_{12} is the largest of the ten angles and a_{34} is the next largest. The resultant labelling of a trigonal bipyramid (left) or square pyramid (right) is shown. (*c*) M has a total coordination number 4. The labels are assigned so that a_{12} is the largest of the six angles. A tetrahedral (left) and a square-planar arrangement (right) are shown. (*d*) ML_4 fragment with an additional partial bond. The ML_4 fragment is defined as in Fig. 1(*c*). $M\cdots L_5$ is a 'non-bonded contact', less than 3.0 Å in length, with L_5 = N or O. (A fragment found may have more than one such non-bonded contact.)

Table 1
ML_6 complexes, numbers (N_{obs}) and mean values of δ_{oct}, in degrees, from ideal geometry, together with their *sample* standard deviations.

The ligand donor atom may be N, O, S or Cl. The October 1999 release of the CSD was used and the connectivity recalculated (in *QUEST* 5.17) assuming the maximum M—N,O distances for a bond shown; M—S and M—Cl may be longer by 0.34 and 0.31 Å, respectively. The standard deviations of the means may be obtained by dividing the sample standard deviation by $N_{obs}^{1/2}$.

	Maximum distance assumed for M—N,O bond (Å)	Non-chelated		Chelated	
		N_{obs}	$\langle\delta_{oct}\rangle$ (σ)	N_{obs}	$\langle\delta_{oct}\rangle$ (σ)
Mg	2.45	120	1.7 (1.1)	123	8 (6)
Mn	2.45	86	2.8 (1.8)	759	9 (6)
Fe	2.35	69	2.2 (1.5)	1094	8 (4)
Cu	2.45	60	1.9 (1.4)	258	8 (5)
Zn	2.35	79	2.4 (1.4)	242	9 (6)
Ca	2.80	41	5.4 (5.7)	33	16 (11)

and/or the transformation paths between tetrahedral and square planar (Klebe & Weber, 1994; Raithby *et al.*, 2000) or trigonal bipyramidal and tetragonal pyramidal (Auf der Heyde & Burgi, 1989; Auf der Heyde, 1994).

References to related studies of metal–ligand geometry are given in the previous paper (Harding, 1999); to these should be added a recent valuable comparison of many properties of Mn with those of Mg and Zn, which are relevant to the possibility of their interchange in metalloenzyme systems (Bock *et al.*, 1999).

2. Methodology

The programs *QUEST* and *VISTA* (Allen & Kennard, 1993*a*,*b*) were used for search and analysis of the October 1998 release of the CSD for four- and five-coordinated complexes and of the October 1999 release for six-, seven- and eight-coordinated complexes. All searches required $R \leq 0.10$ and accepted only non-polymeric structures with no disorder and no errors unresolved at the time of their inclusion in the database. The search fragments used are shown in Fig. 1. N, O, S or Cl were allowed as ligand donor atoms. Recalculation of the metal-atom connectivity at the time of search (*QUEST* version 5.17 and later versions) was performed throughout, using maximum allowable metal–donor atom distances chosen in the light of the distributions of distances actually observed in a variety of relevant complexes (Harding, 1999).

For Cu^{II} compounds, entries were only accepted if they contained the text (ii) in the compound name and no other oxidation-state indicators, (i), (iii) *etc.*; oxidation states were similarly specified for Cu^{I}. An attempt was made to deal with Mn^{II}, Mn^{III} and Fe^{II}, Fe^{III} similarly, but unfortunately this very substantially decreases the number of entries available – only 28% of Mn entries and 20% of Fe entries have oxidation state specified, whereas 84% of Cu entries do. Therefore, all Mn and Fe entries were used and some attempt was made at the end of the analysis to see what significance oxidation state might have. For Mg and Zn, all database entries were used and assumed to be oxidation state 2.

Table 2
ML_4 complexes.

(*a*) Numbers of ML_4 complexes (N_{obs}) with and without additional weak interactions $M\cdots$N,O up to 3.0 Å.

	Mg	Mn†	Fe‡	Cu^{I}	Cu^{II}	Zn
Without additional interactions	40	76	369	237	858	569
With one additional interaction	0	1	4	4	707	49
With two additional interactions	0	2	2	0	1850	86
Max. length assumed for M—N,O bond (Å)	2.45	2.45	2.35	2.45	2.25	2.35

(*b*) ML_4 complexes without additional weak interactions, numbers and mean values of δ_{tet} and δ_{sqp}, in °, from ideal geometry, together with their *sample* standard deviations. The standard deviations of the means may be obtained by dividing by $N_{obs}^{1/2}$.

		Non-chelated		Chelated	
		Tetrahedral	Square planar	Tetrahedral	Square planar
Mg	N_{obs}	8	0	32	0
	$\langle\delta\rangle$ (σ)	11 (3)		16.7 (5.4)	
Mn	N_{obs}	31	0	34	11
	$\langle\delta\rangle$ (σ)	4.0 (3.2)		12 (5)	1.1 (1.0)
Fe	N_{obs}	150	1	207	12
	$\langle\delta\rangle$ (σ)	2.6 (2.6)	3.1	6.1 (2.9)	1.4 (2.1)
Cu^{I}	N_{obs}	39	0	188	10
	$\langle\delta\rangle$ (σ)	6.2 (5.0)		15 (6)	17 (10)
Cu^{II}	N_{obs}	79	87	54	649
	$\langle\delta\rangle$ (σ)	17 (3)	5.5 (7.1)	18.4 (4)	7.0 (5.4)
Zn	N_{obs}	237	0	260	71
	$\langle\delta\rangle$ (σ)	4.3 (2.2)		11 (5)	2.1 (5.6)

(*c*) ML_4 complexes with additional weak interactions, numbers and mean values of δ_{tet} and δ_{sqp}, in °, from ideal geometry, together with their *sample* standard deviations. The standard deviations of the means may be obtained by dividing by $N_{obs}^{1/2}$.

		Non-chelated		Chelated	
		Tetrahedral	Square planar	Tetrahedral	Square planar
Cu^{II}	N_{obs}	5	266	6	1632
	$\langle\delta\rangle$ (σ)	17 (8)	2.5 (3.3)	25 (4)	8.0 (4.1)
Zn	N_{obs}	43	1	91	44
	$\langle\delta\rangle$ (σ)	10 (4)	3.0	13 (5)	0.8 (2.0)

† Mn^{II}, 38 without, seven with additional contacts; Mn^{III}, no ML_4 complexes. ‡ Fe^{II}, 49 without, 13 with additional contacts. Fe^{III}, 77 without, none with additional contacts.

ML_6 fragments were found using the maximum allowable metal–donor atom distances shown in Table 1 and were separated into two groups: those with no chelation (metal atom acyclic) and those with some chelation (metal atom cyclic). For each ML_6 fragment, *QUEST* extracted the 15 interbond angles a_{ij} (see Fig. 1*a*). With a separate local program, δ_{oct}, the r.m.s deviation of these 15 angles from their ideal values was evaluated:

$$\delta_{oct} = \left[\sum_{i=1}^{15}(a_i - a_{ideal})^2/15\right]^{1/2}.$$

If the ligands are labelled so that a_{12} is the largest of the 15 angles and a_{34} is the next largest, the ideal value for a_{12}, a_{34} and a_{56} is 180° and the ideal value for all the other angles is 90°.

(Strictly, only nine of the 15 angles are independent – the other six are related by spherical trigonometry; if ML_6 is on a crystallographic inversion centre, only three of the angles are independent.)

Five- and four-coordinate complexes ML_5 and ML_4 were similarly found using the cutoff distances shown in Table 2 and were each separated into two groups: those which have and those which do not have additional weak $M\cdots$O or $M\cdots$N interactions – identified in *QUEST* searches as 'non-bonded contacts' up to 3.0 Å in length. They were then further subdivided into non-chelated and chelated complexes.

For ML_5 complexes, the ideal geometry may be a trigonal bipyramid or a tetragonal pyramid and δ_{tbp} and δ_{tetp} were similarly evaluated as the r.m.s. deviations of the ten interbond angles. As shown in Fig. 1(*b*), the ligands were labelled so that a_{12} and a_{34} are the largest and second largest angles, respectively; the second largest is irrelevant for the trigonal bipyramid, but for the tetragonal pyramid this labelling makes L_5 the apex. In the trigonal bipyramid, the ideal values are 180° for a_{12}, 120° for a_{34}, a_{45} and a_{35}, and 90° for the others. In the tetragonal pyramid it is useful to define b_m, the mean of the four angles between the apical bond and the basal bonds, a_{15}, a_{25}, a_{35} and a_{45}. The ideal value for these four angles is then b_m; the ideal value for a_{12} and a_{34} is $(360^\circ - 2b_m)$ and the ideal value for a_{13}, a_{23}, a_{14} and a_{24} is $2\sin^{-1}\{2^{-1/2}[\sin(180^\circ - b_m)]\}$. Then,

$$\delta_{tbp} \text{ or } \delta_{tetp} = \left[\sum_{i=1}^{10}(a_i - a_{ideal})^2/10\right]^{1/2}.$$

Strictly, and in the absence of symmetry, only seven of the angles are independent. Again, *QUEST* extracted values for the ten angles and a separate local program evaluated δ_{tbp} and δ_{tetp}; it also assigned the fragment as trigonal bipyramidal if $\delta_{tbp} < \delta_{tetp}$, otherwise it was assigned as tetragonal pyramidal.

For an ML_4 complex (Fig. 1*c*), two ideal geometries are considered, tetrahedral and square planar, and for each ML_4 complex δ_{tet} and δ_{sqp} were evaluated, the r.m.s. deviations of the six interbond angles from their ideal geometries (strictly, only five angles are independent). For tetrahedral geometry, the ideal value of each angle is 109.5°. The ligand labels are chosen so that a_{12} is the largest angle, so for square-planar geometry the ideal value of a_{12} and a_{34} is 180° and the ideal value for the other angles is 90°. The complex is tetrahedral if $\delta_{tet} < \delta_{sqp}$; otherwise it is square planar. For ML_4 complexes it was possible to extract the values of angles and carry out the calculation of δ_{tet} and δ_{sqp} and the assignment of geometry within *QUEST*. However, an additional local program was written to evaluate the deviation of each angle set from various symmetry elements of the tetrahedral or square planar arrangement, the fourfold inversion axis (of both), the threefold rotation axis (of tetrahedral) *etc*., $\delta_{\bar{4}}$, δ_3 *etc*.

After δ had been evaluated for each ML_n fragment (n = 4, 5 or 6), *VISTA* was used to evaluate the mean and the sample standard deviation of δ; these are given in the tables, together with the number, N_{obs}, of crystallographically independent observations. [Since δ for each fragment is already a root-mean-square value for six or more angular distortions, the sample standard deviations overestimate the spread of the individual angular distortions. If we wished to obtain a standard deviation representing the spread of all the angular distortions in the sample, *e.g.* if we want the standard deviation of the 24 angular distortions in four ML_4 groups, not just the standard deviation of four δ values, we would need to multiply by $m^{-1/2}$, where m is the number of independent angle values in one fragment. When no crystallographic symmetry is present, m is five for ML_4, seven for ML_5 and nine for ML_6, but when an inversion centre is present within a fragment m is one for ML_4 and three for ML_6. (Other symmetry elements give intermediate reductions in m.) Two-thirds of the non-chelated ML_6 complexes do lie on crystallographic inversion centres.]

CaL_6 complexes have been examined along with other ML_6. CaL_7 and CaL_8 complexes have been extracted from the CSD using search queries analogous to that in Fig. 1(*a*). Detailed analysis of distortions of the geometry around Ca has not been performed, but the distributions of O—M—O angles and the distributions of O$\cdots$O distances within the first coordination sphere around Ca have been derived.

Ranges of δ values are not quoted here, but were recorded and are usually a little less than $\langle\delta\rangle \pm 3\sigma$. In all the searches R factors and sigf (a *QUEST* indicator of bond-length e.s.d.s) were also extracted and it was checked that there were no trends in δ values with R or sigf. Many individual structures were examined as chemical and stereochemical diagrams. Outliers occasionally indicated errors; for example, a Cu^{II} structure wrongly named as Cu^{I}; in such cases the structure was removed. More often, the inspection of outliers helped in understanding the factors that affect the distributions.

3. Results and discussion

Tables 1, 2 and 3 show the mean values of δ resulting from the various searches and the sample standard deviations representing the spread of values found for ML_6, ML_4 and ML_5 complexes, respectively. The numbers of observations in each category give a rough indication of the relative importance of different geometries for different metals. The δ values represent r.m.s. distortions of interbond angles from the ideal values in the different shapes of complexes. These distortions might be a consequence of one or more of the following: (i) experimental uncertainties in determination of crystal structure, (ii) intramolecular effects, electronic or steric, (iii) intermolecular effects, sometimes called 'packing forces', (iv) the existence of additional partial $M\cdots L$ bonds [this could be considered as a part of (ii) or (iii), but it is found here to be an important factor and is therefore listed separately].

It will be shown that for the metals studied here the distortions in ML_6 complexes are comparatively small and consistent with (i) to (iii) as causes, whereas in ML_4 and ML_5 complexes larger distortions occur. In quite a number of ML_4 and ML_5 complexes substantial distortions can be associated with the presence of a fifth or sixth more weakly interacting ligand. With Cu^{II} and Mn^{III} some longer $M\cdots$O,N distances

are expected as a result of the Jahn–Teller effect. They also occur for Zn, for which there can be no Jahn–Teller effect. All complexes in which such weak interactions occur are treated as a separate group and the main analysis of distortion in ML_4 and ML_5 complexes excludes complexes in which there are additional weak interactions.

3.1. Experimental uncertainties and packing effects

(i) and (iii) are fairly small. The absence of a correlation between δ values and crystallographic R factor or sigf suggests that experimental uncertainty in crystal structure determination, (i), is not a substantial contributor to the effects seen. An estimate has been made by Martin & Orpen (1996) for complexes of transition metals with chloride, acetylacetonate and pyridine ligands by examining duplicate structure determinations and determinations where identical chemical units occur more than once in the crystallographic asymmetric unit, *i.e.* cases where (ii) and (iv) are absent. They selected fairly high accuracy structure determinations ($R < 0.07$ *etc.*) and found bond-angle variations, expressed as standard deviations, of $\sim$1.5°; this represents mainly or entirely 'packing forces',

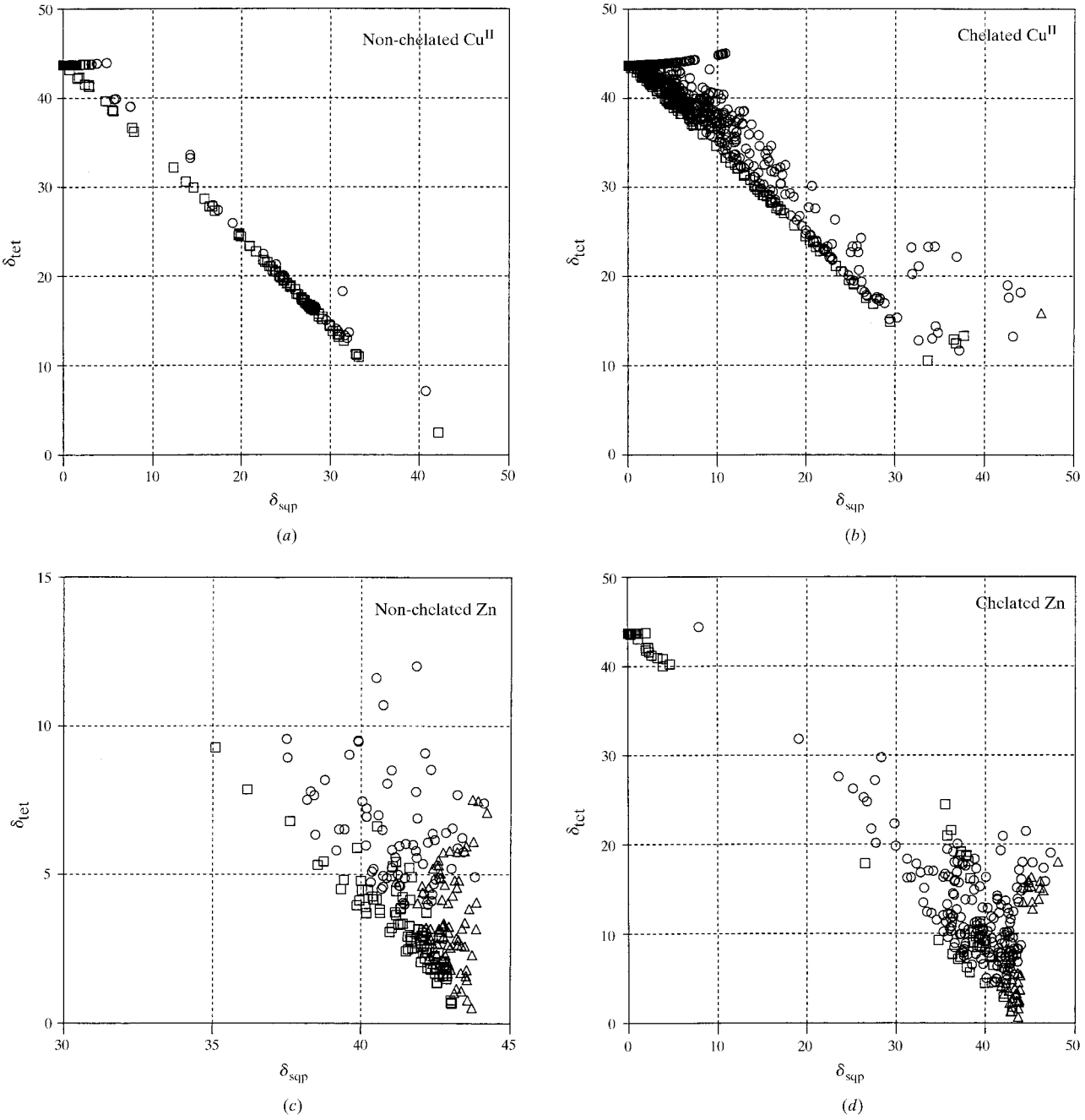

Figure 2
Plots of δ_{tet} *versus* δ_{sqp} for ML_4 complexes of (*a*) non-chelated Cu^{II}, (*b*) chelated Cu^{II}, (*c*) non-chelated Zn, (*d*) chelated Zn. An ideal tetrahedral complex has $\delta_{sqp} = 44°$, $\delta_{tet} = 0$; a perfect square-planar complex has $\delta_{sqp} = 0$, $\delta_{tet} = 44°$. Complexes with $\delta_3 < 3°$ are represented by triangles, complexes with $\delta_{\bar{4}} < 3°$ are represented by squares and all other complexes are represented by circles. Note that the ranges in (*c*) are different from those in the other plots.

since the experimental uncertainty in crystal structure determination is much smaller than this. A similar estimate has been made here by examining examples of $[Mg(OH_2)_6]^{2+}$ in the CSD; their mean δ_{oct} is found to be 1.7 (1.0)°.

3.2. Chelation

Chelated and non-chelated complexes have been separated in the tabulations below and it will be seen that distortions found in the chelated complexes are often much larger (two or three times as great) than those in unchelated complexes. In most of these, the chelate ring includes the metal atom and three, four or five other atoms (C, N, O, S *etc.*); in this case, closure of the ring requires substantial distortion of bond angles and the bond angles at the metal atom are likely to be the most easily distorted. Metalloproteins are usually also formally chelated, but if the donor atoms belong to two or more side chains of a polypeptide chain the ring involves 10–15 atoms or many more; in this case, ring closure should be achieved by distortion of torsion angles from their ideal values rather than by distortion of bond angles. Thus, the acceptable distortions tabulated for non-chelated complexes should be those normally relevant to metalloproteins, apart from the cases where the metal coordination group includes bidentate carboxylate. A distortion of magnitude comparable to that found in chelated complexes may suggest activation of the metal towards a catalytically active state.

3.3. Six-coordinate complexes

For ML_6 with the metals Mg–Zn, δ_{oct} should give a reasonable indication of the variations (ii), intramolecular effects, added to those of experimental uncertainty and crystal packing effects, (i) and (iii). Values are given in Table 1. For all the non-chelated complexes the mean δ_{oct} is 2.2°, which does not greatly exceed the Martin & Orpen estimate of (i) and (iii), the experimental uncertainty and crystal packing effects, and suggests that intramolecular effects are small. However, for chelated complexes the mean δ_{oct} increases to 8.5°, a clear indication that the favourable energy change of chelate formation can easily outweigh the unfavourable energy of distorting an $L-M-L$ bond from its ideal value. The variations from metal to metal are small, but Mn appears to be a little more distortable than any of the other metal ions. In CaL_6 the distortions are twice as great and this is discussed subsequently.

3.4. Four-coordinate complexes, ML_4, without additional weak interactions

It is not difficult to classify these as tetrahedral or square planar according to whether δ_{tet} or δ_{sqp} is smaller. Table 2(*b*) shows, as is already well known, that for all the metals except Cu^{II} a tetrahedral configuration is clearly preferred and is found for virtually all the non-chelated complexes; chelating ligands can force the metal into the less favourable square-planar geometry. For Cu^{II} square-planar geometry is clearly preferred, but bulky or chelating ligands can force it to be near-tetrahedral. For the 465 non-chelated complexes of metals other than Cu^{II} the mean δ_{tet} is 4.0 (2.8)°. Distortions of these ML_4 from tetrahedral are thus twice as great as the average distortion of ML_6 from octahedral. Examination of individual structures shows that large δ_{tet} values generally correspond to complexes with bulky ligands. For reasons similar to those for ML_6 complexes, the distortions in chelated ML_4 are substantially larger than in non-chelated ML_4; the mean δ_{tet} is 11 (5)°. Chelation can often force the geometry to be square planar or nearly so; for many of the examples in which this occurs the ligand is a porphyrin. For non-chelated complexes, variations of $\langle\delta_{tet}\rangle$ from metal to metal are again small, but follow the same pattern as those of $\langle\delta_{oct}\rangle$; the differences are statistically significant when the standard deviation of the mean is considered rather than the sample standard deviation given in the table. A very detailed analysis of distortions in Cu^{II} L_4 complexes taking fuller account of the symmetry of the distortions has recently been published (Raithby *et al.*, 2000).

The existence of complexes which lie on 'the transformation pathway' of conversion from tetrahedral to square planar has been extensively discussed (Klebe & Weber, 1994). Here, they are illustrated in Figs. 2(*a*) and 2(*b*) for Cu^{II} complexes using δ_{tet} and δ_{sqp}. The simplest transformation path involves compression of the tetrahedral complex along its fourfold inversion axis with maintenance of the $\overline{4}$ symmetry. The figures show that the complexes lying close to the line joining the ideal tetrahedral to the ideal square-planar configuration do have small $\delta_{\overline{4}}$, confirming that the line does indeed represent this transformation path. For non-chelated Cu^{II} most complexes lie close to this path and show that a wide range of positions along it is acceptable; many of these are $CuCl_4^{2-}$ in combination with different cations. A small branch near the square-planar position represents another kind of distortion which maintains the inversion centre of the square-planar configuration. The distribution for chelated Cu^{II} is similar, but with larger distortions and many complexes which fall some way off the ideal transformation path, no doubt as a result of the chelating ligands. Zinc complexes are similarly displayed in Figs. 2(*c*) and 2(*d*). The non-chelated complexes do not extend very far up the transformation path. Only a small proportion of the distortions conform to $\overline{4}$ symmetry; some appear to maintain approximate threefold symmetry and the rest of the distortions should probably be regarded as random. Generally, large δ_{tet} values (>8°) can be attributed to bulky ligands, often branched at the donor atom. With chelated ligands, the characteristics of the distribution are similar, though with larger δ_{tet} values and a little more extension up the $\overline{4}$ transformation path; a group of examples near the ideal square-planar position are also seen where porphyrins or similar ligands have forced this geometry. The distribution for Zn is thus strikingly different from that for Cu^{II}. Distributions for Cu^{I}, Mn, Fe and Mg are very similar to those for Zn.

3.5. Five-coordinate complexes

Five-coordinate complexes ML_5 can be classified as trigonal bipyramidal (tbp) or tetragonal pyramidal (tetp) according to

whether δ_{tbp} or δ_{tetp} is smaller. The number of non-chelated complexes is small. Non-chelated Cu^{II} shows the full range from tbp to tetp, but with tetp much commoner (Fig. 3*a*); there is little distortion away from the tbp to tetp path, assuming this is represented by the straight line joining the tbp and tetp positions. (For a detailed discussion of the tbp to tetp transformation, see Auf der Heyde & Burgi, 1989; Auf der Heyde, 1994; Ferretti *et al*., 1996.). Chelated ML_5 are much commoner and tetp is clearly favoured. The distributions for chelated Cu^{II} and Zn are illustrated in Figs. 3(*b*) and 3(*c*); the other distributions are similar. A large proportion of chelated ML_5 have $\delta_{tetp} < 8°$; it is noticeable that none have $\delta_{tbp} < 3.0°$. There are also substantial distortions away from the tbp to tetp path, the greatest of which occur when the metal is part of a four-membered ring. Larger numbers of S or Cl donors tend to favour tbp.

3.6. ML_4 and ML_5 complexes with additional weak interactions

The numbers of these complexes are given in Tables 2(*a*) and 3(*a*), showing how significant these can be, especially with Zn and Cu. The interactions can be intermolecular or intramolecular. The commonest donors are carboxylate O, nitrate or nitrite O, but perchlorate O, water O, heterocyclic O or N, thiocyanate N, secondary or tertiary N of alicyclic ligands and a variety of others are found. Fig. 4(*a*) shows the distribution of observed distances in ZnL_4 complexes; this is continuous up to 3.0 Å. (3.0 Å is an arbitrary cutoff; interactions beyond this may be possible, but they are weak and difficult to distinguish from van der Waals contacts.) In tetrahedral ZnL_4 complexes, the mean distortion of the non-chelated complexes is about twice that of complexes without additional weak interactions (Table 2*c*). If they are treated as ML_5 or ML_6 they are still quite distorted. If there are two additional weak interactions, resulting in a pseudo-ML_6 complex, the angle between the two M···N,O directions ranges from 28 to 180°, with no clear division into *cis* and *trans*. In square-planar CuL_4 and ZnL_4 complexes the distortion is not great because the new donors occupy the fifth or fifth and sixth positions of a tetragonal pyramid or octahedron. Likewise, since most ML_5 complexes are tetragonal pyramidal, an additional weakly interacting donor need not cause much distortion of this symmetry, although it may change the base angle b_m of the complex.

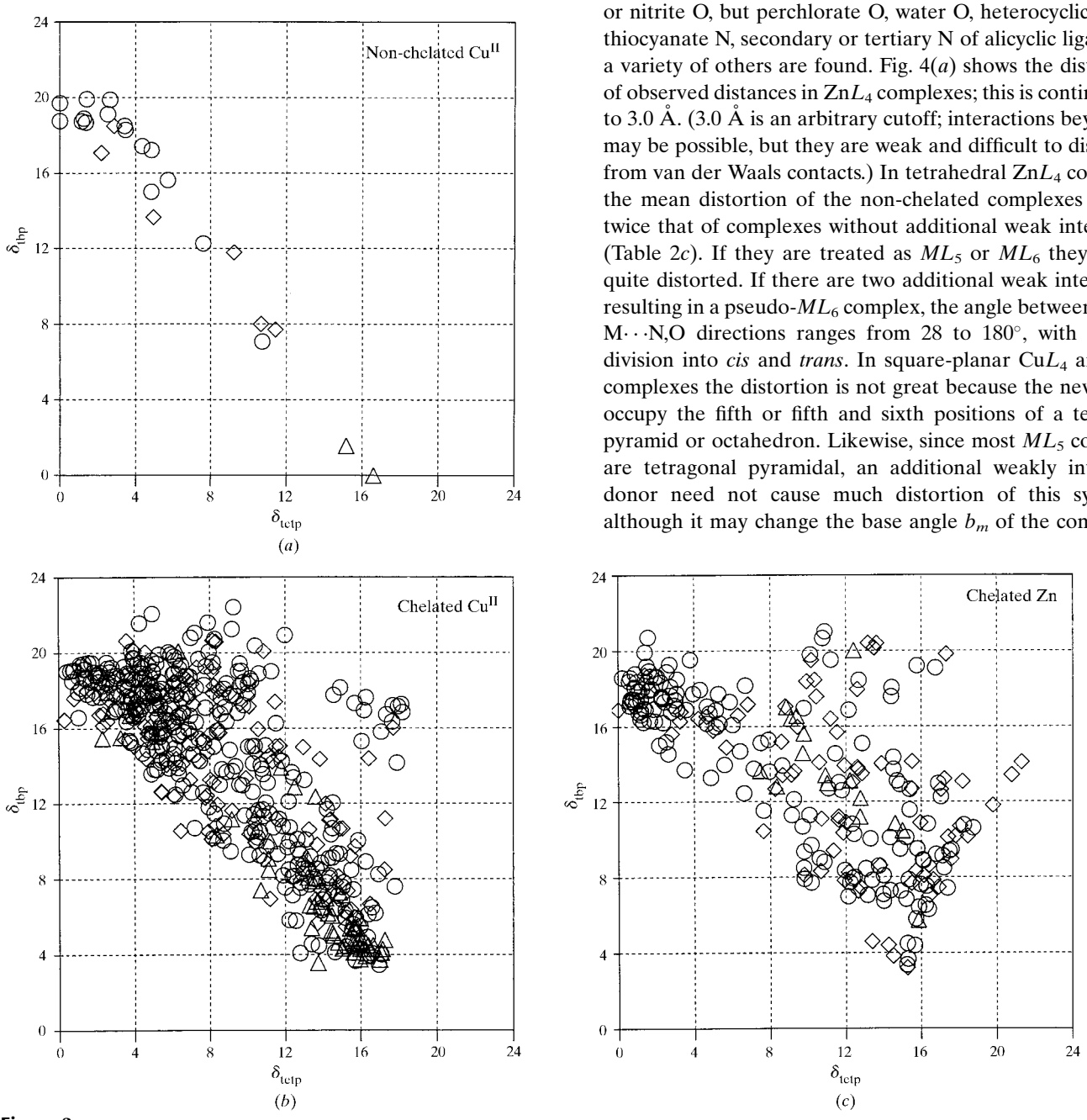

Figure 3
Plots of δ_{tbp} *versus* δ_{tetp} for ML_5 complexes of (*a*) non-chelated Cu^{II}, (*b*) chelated Cu^{II}, (*c*) chelated Zn. An ideal trigonal bipyramidal complex has $\delta_{tbp} = 0$, $\delta_{tetp} = 17°$; an ideal tetragonal pyramidal complex with $b_m = 105°$ has $\delta_{tbp} = 17°$, $\delta_{tetp} = 0$. A circle represents a ligand with N, O donors only, a rhombus one with one or two S or Cl donors and a triangle one with three or more S or Cl donors. The distributions for Mn and Fe are very similar.

3.7. Weak interactions – bond-length and bond-valence considerations

The continuous range of observed Zn—O distances is shown in Fig. 5 for carboxylate coordination in the complexes $(RCO_2)_2Zn(OH_2)_2$ and $(RCO_2)_2Zn(imidazole)_2$. The structures have been fairly accurately determined ($R \leq 0.070$) and the distances Zn—O1, Zn—O2 to the two O atoms in one carboxylate group are clearly inversely correlated.

It is useful to consider weak interactions as weak bonds and compare their bond valence with that of 'normal' M—N or M—O bonds. Brown uses the concept of bond-valence sum at each cation or anion (Brown, 1992; see also See *et al.*, 1998, who estimate bond orders in simple metal coordination complexes). For each bond,

$$\text{bond valence} = \exp[(R_{ij} - \text{bond distance})/0.37],$$

where R_{ij} is a sum of anion and cation radii. The bond-valence sum at a divalent metal ion should be ~2, whether it is made up of four bonds of valence ~0.5 or six of ~0.33 or some other combination. For ionic compounds, bond-valence sums are in good agreement with expectations. For partially covalent complexes, such as the metal complexes here, there is rough agreement, although to obtain more precise agreement the Lewis base strength of the ligand would need to be taken into account (and possibly other factors). According to the above equation and with R_{ij} = 1.66 Å, a normal Zn—O bond of length 2.00–2.10 Å would have a bond valence 0.40–0.30, while a bond of length 2.75 Å would have a bond valence 0.05. Thus, the bond-valence contribution of this long bond is quite small, but the presence of the weakly bonded atom can have a substantial effect on the geometry, particularly in distorting Zn complexes from regular tetrahedral (see Table 2*c*). In Cu^{II} complexes the Jahn–Teller effect can provide an explanation for some of the longer distances; their distribution (Fig. 4*b*) is very different from those of Zn (Fig. 4*a*). It could well be appropriate to regard Cu···N,O distances up to 2.5 or 2.6 Å as 'normal' for axial bonds in ML_5 or ML_6, but there is still a very

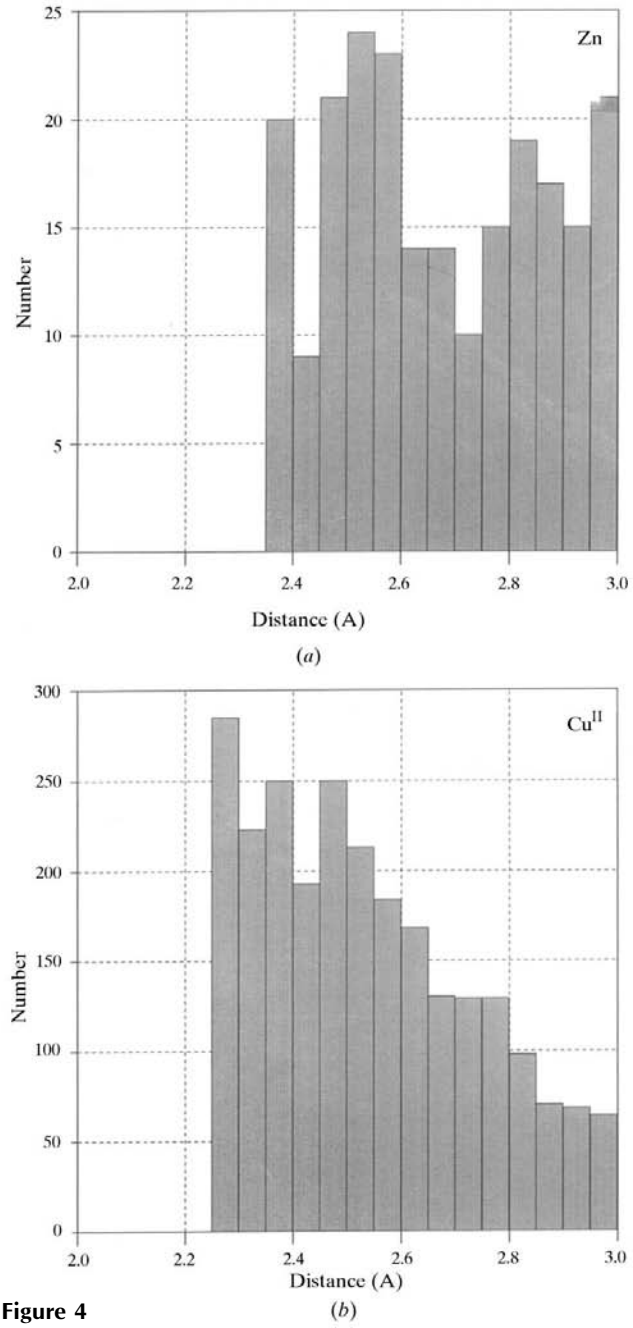

Figure 4
For ML_4 complexes, distribution of $M\cdots$N,O distances greater than the cutoff distances given in Table 2(*a*) (*a*) for Zn and (*b*) for Cu(II).

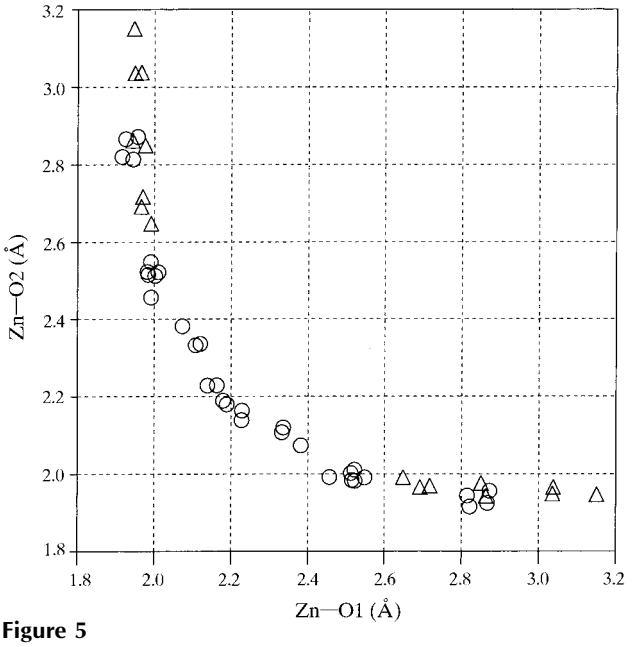

Figure 5
Distances Zn—O1 and Zn—O2 to the two O atoms in one carboxylate group in the complexes $(RCO_2)_2Zn(OH_2)_2$ (circles) and $(RCO_2)_2Zn(imidazole)_2$ (triangles). All structures were determined with crystallographic $R < 0.070$. The refcodes, chemical composition, references and distances for the 17 compounds are deposited as supplementary material.[1]

[1] Supplementary materials are available from the IUCr electronic archive (Reference: ad0108). Services for accessing these data are described at the back of the journal.

Table 3
ML_5 complexes.

(*a*) Numbers of ML_5 complexes (N_{obs}) with and without additional weak interactions $M\cdots$N,O up to 3.0 Å. The maximum distances assumed for $M-$N,O bonds are the same as those for ML_4 complexes.

	Mg	Mn	Fe	Cu^{I}	Cu^{II}	Zn
Without additional interactions	28	87	224	13	575	259
With one additional interaction	3	46	20	0	91	39
With two or more additional interactions	0	6	5	1	3	7

(*b*) ML_5 complexes without additional weak interactions, numbers and mean values of δ_{tbp} and δ_{tetp}, in °, from ideal geometry, together with their *sample* standard deviations. The standard deviations of the means may be obtained by dividing by $N_{obs}^{1/2}$.

		Non-chelated		Chelated	
		Trigonal bipyramidal	Tetragonal pyramidal	Trigonal bipyramidal	Tetragonal pyramidal
Mg	N_{obs}	0	1	8	19
	$\langle\delta\rangle$ (σ)	—	4	11 (5)	5 (4)
Mn	N_{obs}	1	3	21	62
	$\langle\delta\rangle$ (σ)	7	6 (3)	9 (4)	5 (4)
Fe	N_{obs}	4	0	39	185
	$\langle\delta\rangle$ (σ)	1 (1)	—	7 (3)	4 (3)
Cu^{I}	N_{obs}	0	0	4	9
	$\langle\delta\rangle$ (σ)	—	—	11 (8)	8 (2)
Cu^{II}	N_{obs}	5	19	177	374
	$\langle\delta\rangle$ (σ)	5 (4)	3 (3)	8 (3)	6 (3)
Zn	N_{obs}	0	0	103	156
	$\langle\delta\rangle$ (σ)	—	—	9 (3)	6 (5)

(*c*) ML_5 complexes with additional weak interactions, numbers and mean values of δ_{tbp} and δ_{tetp}, in °, from ideal geometry, together with their *sample* standard deviations. The standard deviations of the means may be obtained by dividing by $N_{obs}^{1/2}$.

		Trigonal bipyramidal	Tetragonal pyramidal	Trigonal bipyramidal	Tetragonal pyramidal
Mn	N_{obs}	—	—	11	41
	$\langle\delta\rangle$ (σ)			11 (3)	7 (5)
Fe	N_{obs}	—	—	3	23
	$\langle\delta\rangle$ (σ)			14 (4)	9 (5)
Cu^{II}	N_{obs}	0	6	8	85
	$\langle\delta\rangle$ (σ)		4 (1)	11 (2)	8 (3)
Zn	N_{obs}	3	2	7	36
	$\langle\delta\rangle$ (σ)	7 (6)	10 (?)	11 (4)	9 (5)

significant number of observed Cu···N,O distances in the range 2.6–3.0 Å. Mn^{III} should also show the Jahn–Teller effect. None of the MnL_4 complexes with additional weak interactions (Table 2*a*) was identifiable as Mn^{III}; the MnL_5 complexes (Table 3*a*) include both Mn^{II} and Mn^{III}, which seem to have similar distributions of weak interaction distances.

3.8. Weak $M\cdots$N,O interactions – some general observations

Fig. 6 shows the results of searching in the whole database (April 1999 release) for $M\cdots$O contacts longer than an acceptable $M-$O bond distance (plus a small tolerance for experimental uncertainty in crystal structure determination and restricted to structures with $R < 0.065$). Rowland & Taylor (1996) found histograms like this useful for evaluating the van der Waals radii of a variety of light atoms and halogens. In the absence of hydrogen bonding, Cl···Cl contacts give a distribution similar in characteristics to that shown for Fe···O and with no contacts less than 3.1 Å. Inspection of the histograms of Fig. 6 shows striking differences between different metals. While there are very very few Fe···O contacts shorter than 3.0 Å, Mg and Mn show a few and Cu and Zn show very large numbers. $M\cdots$O contacts just larger than the lower cutoff distance assumed for a bond may simply represent a slightly poorly chosen limit and nearly 'normal' bonding, but distances in the range 2.5–3.0 Å must represent some weak bonding. Even if Cu···O distances up to 2.6 Å are regarded as normal bonds in a complex with Jahn–Teller distortion, the histogram shows that distances of 2.6–3.0 Å also occur quite frequently. It is clear that weak bonds are quite common with Cu and Zn and are much less common with Mn and Fe (Tables 2*a*, 3*a* and Fig. 6), but it is not clear why this should be so; presumably, it would be necessary to look for an explanation in terms of electronic structure.

Table 4
Calcium complexes, some comparative data.

(*a*) Numbers of complexes found (in October 1999 release of CSD), L = N, O, S, Cl (but mostly O).

	Non-chelated	Chelated
CaL_6	41	33 (see Table 1)
CaL_7	16	71
CaL_8	1	97

(*b*) Nearest neighbour O···O distances (in Å) in first coordination sphere around M.

	Minimum	Mean
CaL_6, non-chelated	3.00	3.28 (12)
CaL_7, non-chelated	2.72	3.06 (17)
CaL_8, non-chelated	2.86	3.05 (14) [one only, $Ca(OH_2)_8^{2+}$]
MgL_6, non-chelated	2.78	2.93 (6)
MnL_6, non-chelated	2.83	3.08 (10)
ZnL_6, non-chelated	2.66	2.96 (8)
CaL_7, CaL_8 chelated	2.05 (nitrate)	
	2.17 (carboxylate)	
	2.4 (phosphate)	
	2.6 (other bi or multi-dentate ligands)	

3.9. Simple geometrical considerations

The maximum coordination number (CN) for Mg–Zn in unchelated complexes is normally six, a simple geometrical consequence of the relative size of the O or N donor atoms/ions and the metal. (Coordination number seven does occur in chelated complexes with multidentate ligands, *e.g.* EDTA complexes.) Thus, when CN = 6 the donor atoms/ions are nearly close packed around M and there is little distortability of angles from ideal octahedral. When CN = 4 or 5, the donor atoms/ions are not close packed and there is the possibility of additional weak coordination; both these factors may greatly

increase the distortability of the angles from ideal tetrahedral, square planar *etc*. as shown by the $\langle\delta\rangle$ values given here.

3.10. Calcium complexes

In calcium complexes the commonest coordination numbers are seven and eight, but CaL_6 also occurs (Table 4*a*); O donors are far commoner than any others. The geometry is determined by the possibilities of packing donor atoms (negative ions or atoms with available lone pairs of electrons) around the Ca^{2+} cation; *i.e.* the bonding is predominantly ionic. This is in contrast to ML_4 and ML_5 with M a later transition metal, where a substantial directional contribution from covalency is expected. CaL_6 are octahedral, but with much larger distortions, δ_{oct}, than are found for any of the other metals Mg–Zn (see Table 1); this is easily understandable as a consequence of the greater size of the Ca ion. For CaL_7, an analysis of the angular distortions, δ, from pentagonal bipyramid (PBP) and capped trigonal prism (CTP) has not been performed here, but Howard *et al.* (1998) have described how this can be performed and have given distributions equivalent to δ_{PBP} and δ_{CTP} for all seven-coordinate

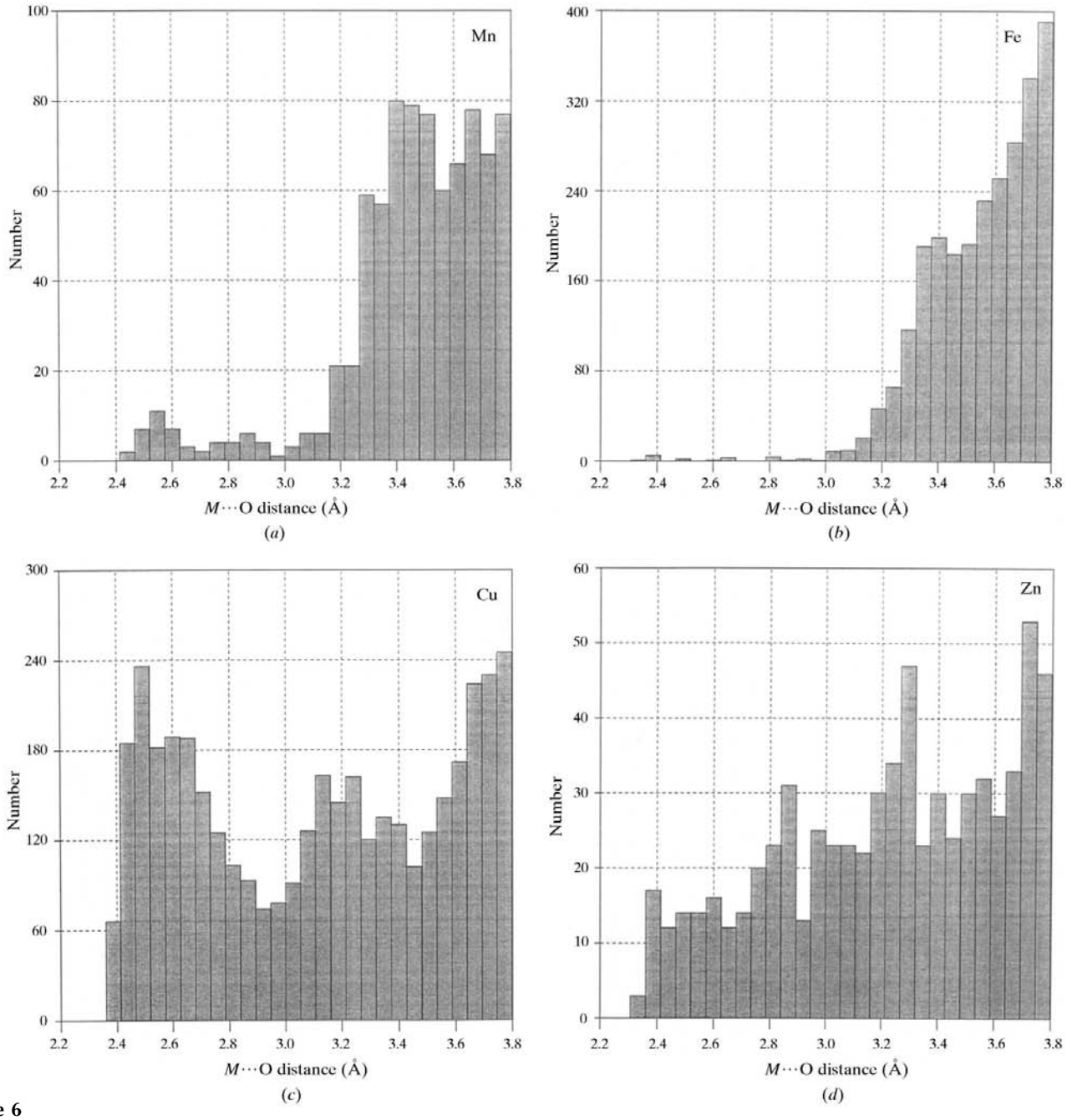

Figure 6
Distributions of $M\cdots$O distances in the CSD (April 1999 release, $R < 0.065$) greater than the cutoff distances given in Table 2(*a*). Effectively, for Mn^{II}, Fe and Zn the distances shown here are longer than those usually regarded as bonding; for Cu^{II} and Mn^{III} distances in the range up to 2.6 A might well be regarded as normal bonding distances in ML_5 or ML_6 when Jahn–Teller distortion is taken into account. (For Mg the distribution is not greatly different form that for Mn, but the total number of observations is much smaller.)

complexes in the CSD, treated together. The distribution of O—*M*—O angles found in CaL_7 complexes is consistent with a mixture of PBP and CTP complexes, with appreciable distortions from the ideal angles of at least 5–10° in non-chelated complexes and more in chelated ones. Rather than considering these angles in detail, it is probably more useful here to consider the donor O atoms distributed on the surface of a sphere around Ca with radius equal to the Ca—O distance. This distribution is governed by the allowed O···O approach distance. Table 4(*b*) makes some comparisons. In non-chelated CaL_7, MgL_6, MnL_6 and ZnL_6 the normal maximum coordination number has been reached and the distances between 'touching' O atoms in the first coordination sphere around *M* are very similar, with a minimum of 2.7 Å and a mean of 3.0 Å. With chelating ligands, a much closer O···O approach may occur; this can result in larger distortions from ideal geometry or an increase to eight of the number of donor atoms around Ca – this is particularly common when carboxylate groups are present. Note that there is only one example of non-chelated CaL_8 in the CSD, $Ca(OH_2)_8^{2+}$. In CaL_6 the O···O distances are larger than those which correspond to 'touching'.

4. Conclusions

For Mg, Mn, Fe, Cu and Zn tabulations of $\langle\delta\rangle$ and the sample standard deviations give a guide to the distortions from ideal values found for interbond angles in four-, five- and six-coordinate complexes with N, O, S and Cl donor ligands. In non-chelated complexes of Mn, Fe and Zn, a cautious view would accept that distortions giving δ up to 5° in six-coordinate complexes and 10° in four-coordinate complexes are normal and are a consequence of packing effects, ligand bulk or other intramolecular effects (experimental uncertainty in crystal structure determination is also included within this, but is comparatively small). Six-coordinate complexes of Cu^{II} are similar. In Mg complexes and in four-coordinate complexes of Cu^{I} somewhat greater distortions are not unusual, say up to 15°. In four-coordinate Cu^{II} complexes, especially those with Cl or S ligands, nearly the whole path from tetrahedral to square-planar geometry appears to be acceptable, and likewise nearly the whole path from trigonal bipyramidal to tetragonal pyramidal geometry in five-coordinate Cu^{II}. Non-chelated ML_5 complexes of the other metals are too small in number for useful conclusions. The geometry of chelated complexes (other than those with bidentate carboxylate) is probably less relevant to metalloproteins, but in them the distortions can be substantially larger.

It is intended that the present analysis of metal–ligand geometry in small molecules will be followed by an analysis of the geometry found in metalloproteins using structural data in the Protein Data Bank (PDB; Bernstein *et al.*, 1977). There are some difficulties in this analysis, because of the much larger uncertainties in most protein structures resulting from lower resolution diffraction data and from the use of restraints in refinement, but it is very desirable to see how far the 'predictions' based on small-molecule structures hold in metalloproteins.

Weak bonding to additional ligands is observed in a substantial number of structures of Zn and Cu complexes, ML_4 and ML_5; a continuous range of $M\cdots$N,O distance between a bonding distance and a van der Waals contact seems to be possible. Similar geometry could well occur in metalloproteins; although the contribution to bond valence or bond energy is not great, the effect on geometry can be considerable (*e.g.* 10–20° distortions in ZnL_4) and this could well be relevant to the understanding of catalytic activity. It would also be important to recognize the possibility of such interactions when interpreting electron-density maps of a metalloprotein.

I am grateful to Dr Jenny Glusker for encouraging my interest in this area, to Professor Malcolm Walkinshaw for laboratory space and computing facilities and Dr Paul Taylor for computational support and to the Cambridge Crystallographic Data Centre for their database and search systems.

References

Allen, F. H. & Kennard, O. (1993*a*). *Chem. Des. Autom. News*, **8**, 1.
Allen, F. H. & Kennard, O. (1993*b*). *Chem. Des. Autom. News*, **8**, 31–37.
Auf der Heyde, T. (1994). *Angew. Chem. Int. Ed. Engl.* **33**, 823–839.
Auf der Heyde, T. P. E. & Burgi, H. B. (1989). *Inorg. Chem.* **28**, 3960–3969.
Auf der Heyde, T. & Nassimbeni, L. R. (1984). *Acta Cryst.* B**40**, 582–590.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Bock, C. W., Katz, A. K., Markham, G. D. & Glusker, J. P. (1999). *J. Am. Chem. Soc.* **121**, 7360–7372.
Brown, I. D. (1992). *Acta Cryst.* B**48**, 553–572.
Ferretti, V., Gilli, P., Bertolasi, V. & Gilli, G. (1996). *Cryst. Rev.* **5**, 3–104.
Harding, M. M. (1999). *Acta Cryst.* D**55**, 1432–1443.
Howard, J. A. K., Copley, R. C. B., Yao, J. W. & Allen, F. H. (1998). *J. Chem. Soc. Chem. Commun.* pp. 2175–2176.
Klebe, G. & Weber, F. (1994). *Acta Cryst.* B**50**, 50–59.
Martin, A. & Orpen, A. G. O. (1996). *J. Am. Chem. Soc.* **118**, 1464–1470.
Raithby, P. R., Shields, G., Allen, F. H. & Motherwell, W. D. S. (2000). *Acta Cryst.* B**56**, 444–454.
Rowland, R. S. & Taylor, R. (1996). *J. Phys. Chem.* **100**, 7384–7391.
See, R. F., Kruse, R. A. & Strub, W. M. (1998). *Inorg. Chem.* **37**, 5369–5375.
Zabrodsky, H., Peleg, S. & Avnir, D. (1993). *J. Am. Chem. Soc.* **115**, 8278–8289.