The assignment and validation of metal oxidation states in the Cambridge Structural Database

Gregory P. Shields,[a,b] Paul R. Raithby,[b] Frank H. Allen[a] and W. D. Samuel Motherwell[a]*

[a]Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England, and [b]Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, England

Correspondence e-mail:
motherwell@ccdc.cam.ac.uk

A methodology has been developed for the semi-automatic assignment and checking of formal oxidation states for metal atoms in the majority of metallo-organic complexes stored in the Cambridge Structural Database (CSD). The method uses both chemical connectivity and bond-length data, *via* ligand donor group templates and bond-valence sums, respectively. In order to use bond-length data, the CSD program *QUEST* has been modified to allow the coordination sphere of metal atoms to be recalculated using user-defined criteria at search time. The new methodology has been used successfully to validate the +1, +2 and +3 oxidation states in 743 four-coordinate copper complexes in the CSD for which atomic coordinates are available in *ca* 99% of structures using one or other method, and both succeed for >86% of structures.

1. Introduction

The formal concept of metal oxidation state is used routinely by coordination chemists, and is invaluable in classifying and understanding the geometries of metal coordination spheres, *e.g.* in structure correlation analyses that make use of the Cambridge Structural Database (CSD; Allen *et al.*, 1991; Allen & Kennard, 1993). While oxidation-state data are available in the CSD, they are entered as components of the chemical compound name, a text field, and only when the authors' assignments are provided unambiguously in published papers. Thus, in structure correlation analyses, it is necessary to use manual editing to add oxidation-state data to the matrix of geometrical parameters that normally underpins such analyses. Over time, we have noted that not only is the oxidation-state data incomplete, it is also prone to error in the published literature, presumably due to typographical or other mistakes.

It would be infinitely preferable if a common definition for oxidation state was applied, and values expressed numerically and associated with the appropriate metal in the chemical connectivity stored within the CSD. This would allow oxidation state to be used as a criterion in chemical substructure searches of coordination compounds, and to be tabulated as part of the data matrix used in structure correlation analyses. Data entry could be carried out by CCDC editorial staff, on the basis of published values and using their own chemical expertise. However, as will be shown below, oxidation-state assignments are not always straightforward, particularly in multi-nuclear complexes, *catena* structures *etc.* Further, while the existing backlog of text data could be translated into numerical form, it is both incomplete and liable to error.

To address these fundamental problems of chemical data representation within the CSD, we have developed computer

procedures that use (*a*) chemical connectivities, and (*b*) bond-valence methods, to determine metal oxidation states. It is hoped that, by combining both approaches, greater accuracy and certainty in assignment can be achieved through the cross-validation of the alternative methods. In the vast majority of cases (*ca* 95% in the case of a large sample of Cu complexes), the procedure is able to make an unambiguous assignment. In other cases it is able to guide the CSD editors in making an assignment. In testing this procedure we have already noted cases where published assignments were in error, and other cases in which the CSD chemical connectivity representation itself was incorrect, often in minor but important respects. Thus, the procedures described here not only yield oxidation-state data that are assigned using a consistent set of rules, but also provide a check on the correctness of published data and a further valuable cross check on the correctness of CSD connectivity representations in an area of chemistry where uncertainty sometimes prevails.

2. Chemical and structural background

A number of definitions have been given for the formal concept of oxidation state (or number), such as those frequently provided in introductory texts, *e.g.* 'formal charge remaining on an element when all other atoms (or molecules) have been removed as their normal ions' (Greenwood & Earnshaw, 1994) or 'number of charges an atom would have in a molecule if electrons were completely transferred in the direction indicated by the differences in electronegativity' (Chang, 1981). The first definition differs in that it allows the possibility that a ligand may not have complete electron transfer in its usual ionic state, *i.e.* ligand atoms may not have a complete valence shell and exist in stable oxidized or reduced forms, which is not allowed by total electron transfer in the direction indicated by differences in electronegativity. The second definition is closely related to that proposed for the formal ionic charge, *viz.* 'the charge on the anion is equal to the charge needed to complete the electron octet of the valence shell. The charge on the cations is then determined by the requirement of electroneutrality. In many cases this will also leave the cation with a filled shell configuration' (Brown, 1999*a*).

The definition of Greenwood & Earnshaw (1984) suggests that oxidation states may be derived directly from chemical connectivity representations, provided that there is no ambiguity in the oxidation states of the ligands (*i.e.* in the normal ions which they normally adopt). For the majority of ligand types the premise that particular ligand donor groups always give rise to a particular oxidation state and valence electron count applies. However, this assumption does not hold for a small proportion of ligands which can exist in reduced or oxidized radical forms (*e.g.* 2,2′-bipyridine, *ortho*-quinone, dinitrogen, peroxo, superoxo and porphyrin ligands), and this may not be discernible from the connectivity representation. However, this situation is comparatively rare and restricted to a few particular classes of ligands which can thus be readily identified. (It is necessary to consult the original publications in cases where the oxidation state is not given in the name field, since spectroscopic techniques may have allowed the oxidation state to be assigned unambiguously.) As a result, a consideration of the nature of the ligands around the metal atoms provides a convenient basis for attempting to assign oxidation state automatically from the chemical connectivity.

Geometric data may also be used in oxidation-state assignment. In order to use such data, it is necessary to have a consistent description of metal coordination. Unfortunately, there is not always agreement as to which interactions should be considered as bonds, or what distance criteria can be used. The connectivity of structures in the CSD is usually stored as defined by the author(s) of the paper in which the structure was reported. Different authors have used different criteria to decide whether a particular interaction should be considered as a bond. Thus, an interatomic distance of 2.8 Å between a Cu(+2) ion and an O atom may be described as a bond in one paper, or as a semi-coordinated or non-bonded interaction in another. The metals of groups 1 and 2 are particularly difficult to treat consistently in CSD analyses. In some cases no formal bonds are assigned at all in the CSD, with the metal being encoded as a free cation. Here, and in other cases, analysis of CSD data is problematical since, by following the published criteria of individual authors, metal–ligand bonds are not encoded in a consistent manner. Such situations can be alleviated by using the non-bonded search facilities of the program *QUEST*3D (CSD, 1994) to locate metal–ligand contacts within specific distance ranges, but this feature is inadequate when more than one contact is being sought around a central metal atom.

A further complication is that two connectivity descriptions are stored for each structure in the CSD, a chemical and crystallographic connectivity. The chemical connectivity and diagram provide a two-dimensional description of the molecules in the structure, including element types, bond types, number of connected H atoms and formal atomic charges (these are constrained by the overall charge on the molecule or ion but do not reliably correspond with the formal ionic charge as described above), but no stereochemical or geometrical information. Conversely, the crystal connectivity consists of atom positions and connections, but no additional chemical data. In general, the two descriptions are completely matched, such that each atom in the chemical connectivity is mapped to one or more atoms in the crystal connectivity (since chemically identical but crystallograpically inequivalent molecules in the structure are represented only once in the chemical connectivity). However, this may not be possible if atoms were not located in the crystal structure, or where molecules are disordered. For such structures, some or all molecules or ions may be unmatched, so there is no means of relating the two connectivity descriptions for these molecules or ions. Since connectivity is searched *via* the chemical connectivity in the CSD, geometrical data cannot be retrieved for molecules which are unmatched.

For the purposes of the present work, and to improve the general search capabilities of *QUEST3D*, we have modified the program to permit metal coordination spheres to be

recalculated at search time, ensuring that the chemical and crystallographic connectivities are matched where possible. These facilities are summarized in §3 below.

The bond-valence model (BVM; Brown, 1992) is the most widely used means of utilizing bond-length data in oxidation-state assignment. This method is based on the concept of bond-valence sums (BVS), *i.e.* that the sum of valences of individual bonds is equal to the atomic valence (*i.e.* the formal ionic charge), and has its origins in Pauling's second rule (Pauling, 1929). The bond valence, s_{ij} (Donnay & Allmann, 1970; Donnay & Donnay, 1973), may be related to the length R_{ij} of a given bond (over the limited distance ranges observed for most bonds) by a number of functional forms, of which an inverse exponential or negative power have been most frequently employed. This leads to the following expressions for the BVS, respectively,

$$V_i = \sum \exp[(R_0 - R_{ij})/B], \qquad (1)$$

$$V_i = \sum (R_{ij}/R_0)^{-N'}. \qquad (2)$$

The parameters R_0, B and N' are constants fitted empirically. B has frequently been taken as a fixed constant 0.37 (Brown & Altermatt, 1985). A semi-empirical justification for the BVM and these functional forms has been provided by Urusov (1995), and more recently the BVM has been placed on a secure theoretical foundation, based on an ionic bonding model, by Preiser *et al.* (1999). Strictly, the BVM is applicable to compounds with a bipartite bond graph (*i.e.* if atoms are considered to be anions and cations there are only bonds between ions of opposite sign; Brown, 1992). In this case the BVS around the anions and cations in the structure should be equal (the electroneutrality principle), which provides a means of verifying R_0 values (since the BVS for anions, *e.g.* O^{2-}, are known). In metal complexes it is not generally possible to determine the BVS around the ligands (the ligand bond graphs are not necessarily bipartite), and R_0 values are usually fitted to predicted formal ionic charges for the metal atoms.

Brown & Altermatt (1985) provide R_0 values for commonly occurring atom pairs, in specific metal and anion oxidation states, derived from the Inorganic Crystal Structure Database (ICSD; Bergerhoff *et al.*, 1983) and these tables were extended by Brese & O'Keeffe (1991). O'Keeffe & Brese (1991) also showed that approximate R_0 values could be derived from size r_i and electronegativity parameters c_i for atom-type pairs for use when more accurate values were not available, and found that the functional form

$$R_{ij} = r_i + r_j - \left[r_i r_j \left(c_i^{1/2} - c_j^{1/2} \right)^2 \right] / c_i r_i + c_j r_j \qquad (3)$$

provided the most satisfactory fit. However, the R_0 values derived by (3) may differ significantly from those fitted empirically for a particular atom-type pair, and should be used with caution.

The program *VALENCE* (Brown, 1996) allows the calculation and manipulation of bond-valence parameters and BVS values interactively, using the R_0 values of Brown & Altermatt

(1985) as defaults, and derives them from the atomic parameters of O'Keeffe & Brese (1991) where bond parameters are unavailable. As pointed out by Hati & Datta (1995), expressions (1) and (2) are unable to accommodate zero or negative oxidation states. As a result they suggested a modified form of (2) which could accommodate such structures,

$$V_i + K = (1 + K'/n) \sum R_{KK'}/r_{ij}^{N'}, \qquad (4)$$

where n is the coordination number and K, K' and N' are adjustable constants (Naskar *et al.*, 1997). However, this formula has no physical justification, and it has been argued (Urusov, 1995) that it is not meaningful to calculate a zero or negative BVS. As a result we have not attempted to use such formulae in this work.

While the BVM has been applied extensively to inorganic structures, as reviewed by Brown (1992), it has been less widely used in molecular coordination chemistry. Liu & Thorp (1993) and Hati & Datta (1995) found that R_0 values were significantly dependent on the oxidation state when expression (1) was used, and R_0 typically increased with increasing oxidation state, *e.g.* for Ni—S bonds Ni(0)—S 1.098, Ni(1)—S 1.649, Ni(2)—S 1.693, Ni(3)—S 2.040 Å. It is notable that the differences are rather smaller for higher oxidation states. This is not surprising in view of the fact that (1) requires mean bond lengths to differ by $0.37\ln(1/2) = 0.256$ Å between Cu(1) and Cu(2) and $0.37\ln(2/3) = 0.150$ Å for Cu(2) and Cu(3) if a common oxidation-state-independent R_0 value is used. These differences are significantly larger than the differences in mean Cu—X bond lengths observed experimentally for most non-metal elements X, these being more typically of the order of 0.1 Å or less.

A further complication is provided by differing spin states of transition-metal ions, since these will affect the differences in R_0 values for different oxidation states. Furthermore, compounds in the same oxidation state may have different spin states [*e.g.* low-spin *versus* high-spin Fe(2)] and in some circumstances it may be necessary to take this into account. Naskar *et al.* (1997) applied (4) to Cu and Ni complexes, with $K = 4$, $K' = 20$, $N = 1$, using oxidation-specific R_0 values, and encountered problems in some structures where several oxidation states fitted within ±0.25. Furthermore, the correct oxidation state was that which agreed most closely with the oxidation-state-specific BVS value in only *ca* 90% of cases, *i.e.* it was not always possible to assign the oxidation state unambiguously. The R_0 values used in this expression vary approximately linearly with oxidation state for Ni—N and Ni—C bonds.

A number of recent studies by Palenik (Palenik, 1997a,b,c; Kanowitz & Palenik, 1998; Wood & Palenik, 1998) have explored the use of oxidation-state-independent R_0 values as an aid to oxidation-state assignment and validation in coordination compounds for metals in oxidation states (2)–(7) with N, O and Cl ligands, with reasonable success. For example, in the case of Fe—O bonds, satisfactory agreement (*e.g.* BVS within 0.25 of the predicted integral value) was obtained for both the (2) and (3) oxidation states using an R_0 value of 1.745,

although a more precise fit was obtained using the oxidation-state-specific values of 1.713 and 1.751 for Fe(2) and Fe(3) bonds, respectively (Kanowitz & Palenik, 1998). Interestingly, this difference is comparable with that found by Hati & Datta (1995) for Ni(2)—S and Ni(3)—S bonds. Steric effects, particularly in low-coordinate metal structures, tend to lead to a lower BVS than expected (Hati & Datta, 1995; Naskar *et al.*, 1997; Palenik, 1997*b*), sometimes deviating by as much as 0.5. The work performed so far in this area has demonstrated the potential for routine use of BVS in validating metallo-organic structures in the CSD, but suggests that oxidation-state-dependent values may be required, at least for low (+1 or less) oxidation states where the R_0 values differ markedly.

3. Assignment of metal coordination sphere

The difficulties in defining metal coordination sphere connectivities were described in §2. To enable more consistent searches of the CSD to be performed for this work, new facilities have been added to the CSD search program *QUEST3D* for modifying the connectivity of the coordination sphere around metal atoms. All existing bonds between given pairs of atom types, *e.g.* Cu—O bonds, may be removed and new bonds added to the crystal connectivity description according to a user-defined distance limit, existing bonds not involving these atom-type pairs in the crystal connectivity being retained. This takes the form of a tolerance value, the adjustment to the covalent radii (r_1, r_2) of the atoms which determines whether a Cu—O contact d_{12} is a bond, *i.e.* when $d_{12} \leq r_1 + r_2 +$ tol. The tolerance may be negative, *e.g.* unwanted Cu(2)—Cu(2) contacts could be removed with a tolerance of -2.0 Å. Normally the CSD elemental radii are used for recalculating the crystal connectivity (CSD, 1994). Alternatively, the radius value for a particular element may be given explicitly, this value only applying to recalculated connections. The crystal connectivity is recalculated and converted to a chemical connectivity description and a revised chemical diagram for each structure as a search proceeds, ensuring that the chemical and connectivity descriptions are completely matched. These new facilities have already proved useful in the analysis of some metal–ligand interactions relevant to proteins (Harding, 1999).

If the original chemical and crystallographic connectivities were completely matched, the existing chemical bond types, number of terminal H atoms and formal charges are retained in the new chemical connectivity. New bonds added in the recalculation are assumed to be single or π-bonds, and this assignment is made automatically. In partially matched structures, *i.e.* those in which the chemical and crystallographic connectivities of only some of the molecules or ions matched exactly, bond types from the original chemical connectivity are retained for the matched molecules or ions. For molecules or ions which were unmatched, bond types and formal charges are assigned algorithmically, using methods developed for deriving chemical diagrams from geometrical data.

In some cases, particularly some disordered structures, the CSD contains atomic coordinates but no crystal connectivity. Here the connectivity is first evaluated for all atoms using a default covalent radius tolerance of 0.4 Å, as in the unique molecule procedure (Allen *et al.*, 1974) before recalculating bonds between specific atom-type pairs, *e.g.* Cu—O. The results obtained may not always be chemically sensible as there may be misplaced atoms, or unresolved disorder. However, the first coordination shell of the metal atoms may be correct in many cases.

Polymeric (*catena*) structures present a particular problem for automatic connectivity generation. By default, polymers will not be expanded and the base unit (asymmetric unit plus adjacent atoms) will be generated. This unit may include incomplete ligands, although the first connections to the unique metal atoms will be complete. Alternatively, a polymer may be expanded to include the full ligand coordination sphere around each symmetry-independent metal atom in a structure. This is achieved by building the connectivity of the ligands as separate molecules first, and then connecting the metal atoms and ligands such that each symmetry-independent metal atom exists in the expanded connectivity with its complete ligand coordination shell. While this may create a large polymer unit in some cases, it does ensure that all ligands are complete, which facilitates the matching of polymeric structures with ligand templates.

4. Oxidation-state assignment methodology

The calculation of oxidation states and bond-valence sums was performed using a modified version of *QUEST*3D, implemented according to the following scheme:

(*a*) read a set of ligand donor group templates, and the oxidation-state contributions associated with them;

(*b*) locate a metal-containing structure in the CSD using a suitable search criteria;

(*c*) recalculate the metal coordination environment where necessary, using the methodology described in §3;

(*d*) generate a metal/ligand connectivity table, and a list of which donor atoms belong to each ligand, from the chemical connectivity;

(*e*) derive ligand formulae and formal ligand charges (sum of formal CSD charges on the ligand atoms);

(*f*) attempt to match ligand donor group templates with the chemical connectivity, until all the donor atoms in the structure are matched or the ligand templates are exhausted;

(*g*) provided all donor atoms are identified by the ligand templates, assign metal oxidation states on the basis of the ligand template contributions and CSD formal charges on the metals and ligands;

(*h*) calculate bond-valence sums where the crystallographic connectivity is available and matched with chemical connectivity;

(*i*) balance oxidation-state contributions from bridging ligands using BVS values where possible;

(*j*) attempt to assess whether the oxidation state is reasonable, and whether there are likely sources of error.

4.1. Ligand templates

The donor group templates comprise both ligand fragments and the metal(s) to which they are attached, and are encoded as *QUEST3D* connectivity search fragments, with extra fields defining the contribution of each template donor atom to the oxidation state of each metal included in the template. For some ligands which have been taken as having different oxidation states (*e.g.* η^5-C_5H_5 which may be described as a neutral $5\pi e^-$ radical or a $6\pi e^-$ anion) it was necessary to make a decision of which to adopt.

Donor group fragments rather than complete ligands were defined, and some of these include generic donor atoms or substituents to be defined. This allows more ligands to be described with the same template. For example, an EX_2 template (E = S, Se, Te; Fig. 1a) would match the four S donor atoms in the ligands shown in Fig. 1(*b*). It is not necessary to have a separate template for ligands with different numbers of the same SR_2 donor group, since the way in which the donor groups are connected in an unsaturated carbon ligand backbone does not affect the oxidation contributions each make to a metal (zero in this case). Furthermore, complicated or unusual ligands would be difficult to include given the limit on the total number of templates which can be defined in the program (currently 250). In conjugated/delocalized systems this is not generally the case and the effect of peripheral non-coordinated groups must be taken into account in designing the ligand templates, particularly those involving N atoms.

We have designed 250 donor group templates that are sufficiently general to include a large proportion of ligand types commonly encountered in organometallic and coordination chemistry, while being specific enough to exclude ligands which, although similar, make different oxidation-state
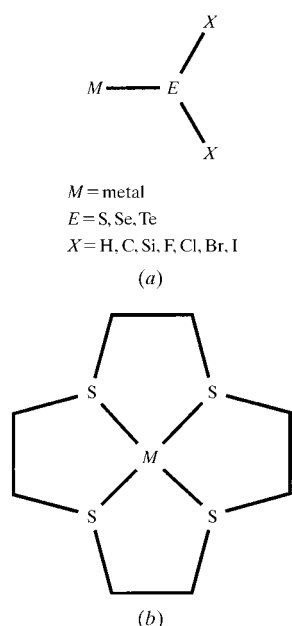
contributions. These have been deposited as supplementary material.[1] Lack of standardization of bond types in some common functional groups in the CSD necessitated the use of variable bond types in some templates. With the exception of simple ligands (*e.g.* terminal halide or carbonyl) the templates were verified individually by searching the CSD, ensuring that only the desired fragments were matched and amending the templates where necessary. Some potential templates succeeded for fewer than ten structures and these were eliminated from the 250 finally used (in some cases these were included in more general templates, for others no template is currently provided). The complete set of templates were trialled on a sample of the first 200 transition metal/lanthanide/actinide complexes in the CSD. Further templates were introduced where these trials indicated that particular classes of ligands were not accounted for by the existing set. Ligand donor groups which commonly occur in more than one oxidation state (*e.g.* peroxo) were flagged in the template, causing the program to print an appropriate warning.

Once all donor atoms have been identified, the remaining ligand templates are ignored. If the same ligand donor atom has been matched to more than one template (this may occur occasionally if the templates are not mutually exclusive), the automatic assignment process is abandoned for the structure, an appropriate warning is given and the oxidation states must then be assigned manually. Similarly, if the ligand templates are exhausted and some donor atoms are left unmatched, either the entry may be abandoned, oxidation states assigned manually or the contributions from each donor atom which had not been matched may be entered separately, the appropriate metal–ligand bond being highlighted in the chemical diagram presented to the operator.

4.2. Metal and ligand charges

The ligand template oxidation-state contribution is the opposite of the total charge on the ligand template donor atoms when the ligand is in its usual ionic state (*i.e.* the charges which allow the template atoms to fulfil their usual valencies). The contribution of residual metal charge, *i.e.* the hypothetical formal charge remaining on the metal atom in order to satisfy the overall charge on the molecule or ion if the atoms in the ligand templates were formally uncharged, must also be added to the calculated oxidation state. As a result, the CSD formal charges on the metal and on the ligand atoms must also be considered, since the residual charge is also the difference between the formal charges which are actually stored in the CSD for the metals and ligand atoms matched by the donor group templates, provided that the formal charges are correct for atoms on the periphery of ligands which are not involved in the template matching. CSD charges on the donor atoms and their neighbours and next-nearest neighbours are taken into account in the calculation of the formal CSD ligand charge, which is derived for each ligand as a whole rather than for

Figure 1
(*a*) Thio/seleno/telluro ether ligand template and (*b*) macrocyclic ether ligand matched four times by this template.

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: BR0092). Services for accessing these data are described at the back of the journal.

research papers

each donor atom, since the neighbours and nearest-neighbours may be common to more than one donor atom. The contribution of the ligand CSD charges is divided equally between the number of metal atoms bridged by the ligand in these calculations. This procedure is illustrated in Fig. 2 for a possible CSD representation of a hypothetical Cu(2) complex.

Unfortunately, the assignment of ligand charges is not always consistent in the CSD. Problems are presented by charged groups (*e.g.* $-NH_3^+$, COO^-, $^-OSO_3$) on the periphery of ligands which are not involved in the donor template matching, as conventional charges may not always have been included in the CSD chemical connectivity description. For example, *para*-tosyl ligands may not possess a conventional negative charge on an O atom, or quaternary nitrogen may not bear a positive charge. In such cases we have applied automatic corrections and adjusted the CSD formal metal charges accordingly.

4.3. Metal–metal bonds and bridging ligands

Metal–metal bonds are considered to be electron-pair bonds, *i.e.* each metal contributes one electron to a single metal–metal bond and two to a double bond. As a result, metal–metal bonds are assumed to have no mutual influence on oxidation state (see below). Cu(2) presents difficulties in this respect, since the inter-metal distances in Cu(2) dimers may be less than the sum of the covalent radii although the Cu atoms may not be bonded directly, merely interacting *via* a superexchange mechanism. The assignment of a bond in the CSD generally reflects the comments of the author(s) or the manner in which the molecule was represented in the original publication. This difficulty can be overcome by including metal coordination recalculation methodology (§3) in the search, to ensure that such interactions are treated consistently.

An ambiguity in the oxidation state exists when a metal is associated with metal–metal bonds or bridging ligands, since it can neither be assumed that the two metals formally contribute the same number of electrons to a metal–metal bond, nor that a bridging ligand which is chemically symmetric will donate the same number of electrons to each metal in a formal sense (*e.g.* a carboxylate ligand bridging two different metals Cu and Ni). Furthermore, the apportioning of CSD formal charges to individual metal atoms within a molecule or molecular ion containing several metal atoms may be incorrect in some structures. In these cases it is not possible to derive the oxidation state automatically from the connectivity description alone and manual intervention is required in order to apportion correctly the oxidation state and electron count contributions between the metal atoms. Use of BVS provides a means of resolving this ambiguity (§4.4).

Even for structures in which each molecule or ion contains only a single metal atom, if the total molecular charges are not assigned correctly the oxidation states and electron counts will be in error. Similar problems exist for polymeric (*catena*) structures, since it is not possible to determine to which ligand donor atoms from incomplete ligands belong, and no attempt

was made to assign an oxidation state automatically in these cases. However, this difficulty is resolved if the recalculated expanded connectivity is used (§3), since only complete ligands are included. The oxidation states of metals having incomplete coordination shells are not calculated, although their ligand donor template contributions are calculated.

Non-integral oxidation states may be derived when the total charge on a molecule is not divisible by the number of metal atoms and in mixed oxidation-state species. Although metal charges may be formally non-integral in practice, the CSD only stores formal integral charges. A simple algorithm attempts to balance the oxidation states of two or more atoms of the same element which have non-integral oxidation states in the molecule, by adding/subtracting oxidation-state remainders. The process is based on the assumption that a particular element occurs in only one oxidation state in a molecule. No attempt has been made to correct non-integral oxidation states for different elements or situations with unbalanced oxidation states involving integral oxidation states at this stage [*e.g.* structures with two Cu(2.5) and one Cu(1) centre which are all actually Cu(2)], although this can be achieved using BVS data.
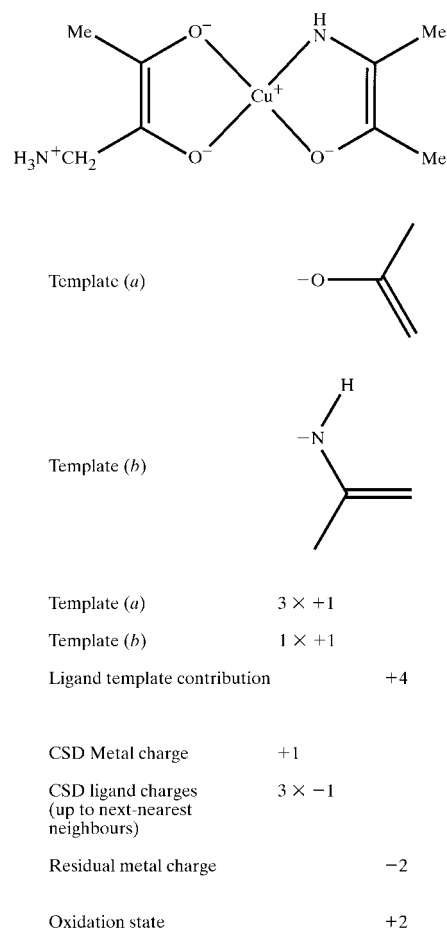

Figure 2
Hypothetical Cu(2) complex matching two donor group templates (*a*) and (*b*), demonstrating the ligand template method of oxidation-state calculation.

Table 1
CSD elemental radii.

Element	r (Å)
Cu	1.52
C	0.68
N	0.68
O	0.68
F	0.64
P	1.05
S	1.02
Cl	0.99
As	1.21
Se	1.22
Br	1.21
Te	1.47
I	1.40

4.4. Bond-valence sums

The calculation of BVS is performed after the ligand template matching, charge corrections and non-integral oxidation-state balancing steps have been performed. Expression (1) with a fixed B value of 0.37 is used. BVS were calculated for each metal; where oxidation-specific R_0 values were available, the R_0 value appropriate to the oxidation state found from ligand template matching was tried first. If the BVS agreed within a tolerance of ±0.15, the value was retained, otherwise the R_0 values for all oxidation states were tested and the R_0 which gave a BVS closest to that oxidation state was retained, following the approach of Naskar *et al.* (1997). Where atomic coordinates are not available for a metal-containing structure, and where the coordinates are not matched with the chemical connectivity, the BVS cannot be calculated and the oxidation states must be calculated on the basis of the ligand donor templates alone.

Where the sum of the nearest integers to each BVS value for a structure is equal to the total of the metal oxidation states predicted using the ligand templates, the oxidation-state contributions are apportioned between the metal centres according to the BVS values, even where all the oxidation states derived using the ligand templates were integral. This allowed the ambiguity in structures with bridging ligands to be overcome, and corrected errors due to missing charges on molecules or ions which occur only occasionally in some structures.

4.5. Program operation

The interactive mode of operation permits four courses of action for any CSD entry: (*a*) the oxidation states derived by the procedure outlined above may be accepted, (*b*) the oxidation state and valence electron count contributions may be re-apportioned manually (with reference to information provided in the original publication where necessary), (*c*) the assignments may be overridden, or (*d*) the entry may be abandoned. The re-apportioning method has the advantage that a set of oxidation states is obtained which is consistent with the total contribution from all the ligands, and may be

useful when re-apportioning oxidation-state contributions using BVS values (§4.4) has failed. Similarly, where some ligands were not matched, their oxidation-state contributions can be entered by hand. In batch mode, the assignments are always accepted and BVS values alone are used when some ligands do not match the ligand templates. This ensures that all structures have oxidation states assigned in the output file, which lists the CSD reference code, metal type, derived oxidation state and BVS for each metal centre.

5. Experimental

5.1. Data retrieval

The methodology was tested on a CSD subset of 743 structures containing exactly four-coordinate Cu complexes for which atomic coordinates were stored, excluding determinations having bond-length errors of >0.02 Å, or an R-factor of ≥10%. These were retrieved using *QUEST3D* (April 1998 release). The Cu coordination spheres were recalculated using the methodology described above (§3). Cu-transition metal (including Cu) bonds were deleted from the connectivity description and Cu−L bonds were included which were less than the sum of the CSD covalent radii (Table 1) within a tolerance of 0.4 Å for C, 0.5 Å (N, P, As, Sb, Bi), 0.6 Å (S, Se, Te, Po and F, Cl, Br, I, At) or 0.8 Å (O). Bonds between other atom-type pairs in the existing chemical connectivity were unchanged. This eliminated Cu(2) complexes which were nominally four-coordinate in the CSD, but which have additional semi-coordinated (Brown *et al.*, 1967) ligands within these distance limits.

5.2. Derivation of bond-valence parameters

Oxidation-state-dependent R_0 values were used only for Cu. Oxidation-state-independent R_0 values derived by Palenik (1997*b,c*), Wood & Palenik (1998) and Kanowitz & Palenik (1998) were used for other metals when available, otherwise they were derived from the atomic parameters of O'Keeffe & Brese (1991).

The R_0 values for Cu were calculated from CSD data using Fortran programs supplied by Professor G. Palenik (Palenik, 1999). Geometric *QUEST3D* searches were performed using the same methodology and search criteria as for the four-coordinate subset (§5.1), but for complexes having a given coordination number (between 2 and 6) and donor atom type. The resulting output was processed to produce a list of the Cu−L bond lengths for each donor type and coordination number and grouped according to oxidation state. The oxidation state was assigned by inspection with reference to the compound name in the CSD and a rough BVS calculated with tentative R_0 values based on the atomic parameters of O'Keeffe & Brese (1991). For a given oxidation state the R_0 value which minimized the differences between predicted and calculated BVS values was calculated. Where there were no examples of complexes with only one donor element available, bond lengths in mixed complexes were fitted using R_0 values

Table 2
R_0 values for Cu derived from CSD data.

[] indicates coordination number of Cu.

	Cu(1)		Cu(2)		Cu(3)	
	R_0	N_{obs}	R_0	N_{obs}	R_0	N_{obs}
Cu—C						
[2]	1.716	83				
[3] (μ_1-C)	1.521†	7				
[3] (μ_2-C)	1.734†	6				
[4]	1.446	8	1.716	17	1.844	3
Cu—N						
[2]	1.623	36				
[3]	1.582	60				
[4]	1.525	60	1.716	63	1.768	8
[5]			1.705	64		
[6]			1.715	153		
Overall	1.571	195	1.713	280	1.768	8
Cu—O						
[2]	1.578	51				
[3]		3	1.692	2		
[4]	1.534‡	34	1.657	147		
[5]			1.652	138		
[6]			1.656	246		
[7]			1.658	25		
[8]			1.657	12		
Overall	1.567	88	1.655	570		
Cu—P						
[2]	1.979	19				
[3]	1.897	45				
[4]	1.774	31				
Overall	1.844	99	2.053‡	5		
Cu—S						
[2]	1.907	36				
[3]	1.843	132				
[4]	1.811	123	2.022	37	2.078	27
[5]			2.031	8		
[6]			2.026	5		
Overall	1.834	291	2.024	50	2.078	27
Cu—Cl						
[2]	1.829	36				
[3]	1.849	13				
[4]	1.858	16	1.992	156		
[5]			2.004	58		
[6]			2.008	79		
Overall			1.999	293		
Cu—As						
[4]	1.856‡	111				
Cu—Se						
[4]	1.900	3				
Overall	1.900	3	2.124	6		
Cu—Br						
[2]	1.963	13				
[3]	1.971	12				
[4]	1.964	7	2.124	27		
[5]			2.136	4		
[6]			2.148	20		
Overall	1.967	32	2.134	51		
Cu—I						
[2]	2.132	2				
[3]	2.146	14				
[4]	2.155	32				
Overall	2.153	48	2.36	6		

† Bridging and terminal C donors fitted separately due to large difference in R_0 value. ‡ Values derived by fitting complexes with other donor atom types also present.

already derived from the other atom pairs. The results are summarized in Table 2.

From this table it will be seen that the R_0 values are strongly oxidation-state dependent for Cu. It is clear from Table 3 that the ideal bond lengths, R_{ideal}, *i.e.* those which would give a BVS equal to the oxidation state for a homoleptic complex with equal bond lengths, do not differ as much as would be predicted from equation (1). For some elements (especially C), R_{ideal} values are essentially independent of oxidation state, hence BVS based on these elements alone will not be able to differentiate between possible oxidation states. R_{ideal} values are most discriminating for O and the halogens.

For Cu(1) the mean bond lengths in homoleptic three-coordinate complexes are expected to be $0.37\ln(4/3) = 0.106$ Å shorter than in four-coordinate complexes, and $0.37\ln(3/2) = 0.150$ Å shorter still in two-coordinate complexes according to equation (1) if a coordination-number-independent R_0 value is used. Whereas the mean bond lengths for N, S and P do not differ as much as this, those for Cl and I differ by more, and as a consequence the coordination-number-specific R_0 values increase with coordination number for N, S and P donors, whereas the opposite trend is shown by the halogens. This suggests that the effect may, in part, be steric, since the N, S and P ligands usually have bulky substituents (*e.g.* ^{t}Bu) whereas the halogens cannot, as found previously for four-coordinate metal centres (*e.g.* Palenik, 1997*b*). Alternatively, it may indicate that a different value of B should be used, as has recently been suggested by Adams (1999) in cases where R_0 values are strongly dependent on coordination number when $B = 0.37$ is used. Adams (1999) found that B values can be broadly correlated with the difference in softness of the anion and cation. In our work a common R_0 value can be obtained for Cu(1) bonds if B values of 0.64 (P), 0.52 (S), 0.50 (N), 0.44 (O), 0.37 (Br), 0.34 (I) and 0.33 Å (Cl) are used (Brown, 1999*b*). However, since the B values are difficult to determine accurately and further investigations are needed, we have continued to use 0.37 Å for this work. Using $B = 0.37$ Å for N, P and S donors, satisfactory agreement for four-coordinate Cu(1) complexes could only be obtained if a coordination-number-specific R_0 value (Table 3) was used. For Cu(2) the R_0 values did not vary significantly with coordination number (and the ideal bond lengths are not predicted to vary as much) and the overall values (Table 3) were used in the oxidation-state assignment program. The overall values were calculated by fitting the bond lengths for all coordination numbers at the same time, and do not correspond to the arithmetic mean of the coordination-number-specific values.

In addition, it was found that there was a large spread of BVS for CuN_x complexes; in particular, Cu(2) complexes with amine ligands gave BVS less than 1.80. Considerably better agreement was obtained using the values in Table 3 for N donor atoms subdivided according to the coordination number of N atoms within the ligand, *i.e.* excluding metal—N bonds. The problems of assigning a single R_0 value for M—N bonds have been identified previously by See *et al.* (1998), the donor ability of N ligands differing sufficiently that a common parameter is not appropriate, and separate parameters for

Table 3
R_0 values used in oxidation-state assignment for four-coordinate Cu complexes with corresponding R_{ideal} bond length which would give predicted BVS with this R_0 in a homoleptic complex.

[] indicates coordination number of donor atom in ligand, *i.e.* ignoring bonds to metal atoms.

	Cu(1)		Cu(2)		Cu(3)	
	R_0	R_{ideal}	R_0	R_{ideal}	R_0	R_{ideal}
C	1.446	1.96	1.716†	1.97	1.844†	1.95
N	1.525	2.04	1.713	1.97	1.768	1.87
N[1]	1.480	1.99	1.709†	1.97	–	–
N[2]	1.520	2.03	1.704	1.96	–	–
N[3]	1.630	2.14	1.763	2.02	–	–
O	1.504†	2.02	1.655	1.91	–	–
P	1.774	2.29	2.053†	2.31	–	–
S	1.811	2.32	2.024	2.28	2.078	2.19
Cl	1.858	2.37	1.999	2.25	–	–
As	1.856†	2.37	–	–	–	–
Se	1.900	2.41	2.124†	2.38	–	–
Br	1.964	2.48	2.134	2.39	–	–
I	2.155	2.67	2.36†	2.62	–	–

† Values derived by fitting complexes with other donor atom types also present.

one-, two- and three-coordinate N were used for oxidation-state checking. For Cu(2)N$_4$ the R_0 values for two- and three-coordinate N differ by *ca* 0.06 Å, which corresponds to BVS differing by a factor of *ca* 0.85.

6. Results and discussion

6.1. Oxidation-state assignment

The validation results are summarized in Table 4. While the oxidation states of other metals were evaluated, the results for Cu alone are summarized, as this is the only metal for which sufficient accurate R_0 values have been derived. These results are encouraging in that, for 88% of the metal structures assigned, the oxidation states determined using ligand templates agreed with the BVS values. Of these, five Cu(1) and ten Cu(2) structures required oxidation-state contributions to be balanced over several metal centres (in some cases metals other than Cu) using the best BVS fit atomic valences. In 63 complexes the ligands were not matched by the standard templates and oxidation-state assignment had to be performed by BVS alone. These oxidation states were verified by manual inspection, with reference to the CSD compound name field. The oxidation state was calculated incorrectly for the bis(1,2,6,7-tetracyano-3,5-diimino-3,5-dihydropyrrolizinido-*N*,*N'*) Cu(2) complexes VAVKOS [Bonamico *et al.*, 1989; BVS 1.13 using the R_0 value for Cu(1), 1.84 R_0 of Cu(2)], WEGDUH [Bonamico *et al.*, 1993; 1.12 R_0 of Cu(1), 1.82 R_0 of Cu(2)] and the 5,10,15,20-tetramethylporphyrinato Cu(2) complex JILGAM [McGhee *et al.*, 1991; BVS for R_0 of Cu(1) 1.10, R_0 of Cu(2) 1.78], although the BVS with Cu(2) R_0 values are in acceptable agreement. Similar problems were encountered for Ni complexes using equation (4) (Naskar *et al.*, 1997).

The BVS and ligand template methods both fail to give the correct result in only three examples, (py)$_2$Cu(μ_2-S)$_2$-Mo(μ_2S)$_2$Cu(py)$_2$ DIBTEN11 (Lin & Lin, 1991), [η^2-(EtO)$_2$PS$_2$]$_3$W$_3$(NCMe)(μ_2-PhCOO)(μ_3-S)$_4$CuI JAXLUP (Zheng, Zhan *et al.*, 1989) and [η^2-(EtO)$_2$PS$_2$]$_3$Mo$_3$(py)(μ_2-HCOO)(μ_3-S)$_4$CuI KEFNIS (Zheng, Zhu & Wu, 1989). The oxidation-state contributions in these clusters are not apportioned correctly by the ligand templates. The BVS values suggest Mo/W(5)Cu(2) for JAXLUP and KEFNIS, inconsistent with the total ligand oxidation-state contribution, whereas the correct description is M(4)Cu(1). Cu(1) R_0 parameters give BVS values of 1.23 and 1.27, and Cu(2) R_0 values give 2.10 and 2.17, respectively. It is probable that some atom positions may not be determined to sufficient precision in the presence of the I atom in each structure. These examples show that balancing ligand contributions using BVS values will fail if the correct oxidation state does not give the best BVS fit. The structures of the mixed Fe/Cu complexes FIHFEH (Brewer & Sinn, 1987) and YAKMIG (Serr *et al.*, 1992) provide further examples of this, the Cu oxidation state being incorrect due to the default Fe R_0 values providing a poor fit.

The most interesting cases are those where the ligand templates and BVS values do not agree, and in such instances it is necessary to analyse the individual structures in detail. Occasionally this is due to the presence of a ligand in a reduced oxidation state, and provides a convenient means of identifying this situation, *e.g.* *N,S,S'*-C$_5$H$_4$(CH)$_2$S(CH$_2$)-S(CH$_2$)C$_5$H$_4$N Cu(1) (*N*-tetracyanoquinidimethane) GIVBES (Humphrey *et al.*, 1988) for which the ligand template suggestion is Cu(0). Occasionally it indicates an error in the CSD chemical connectivity. JEXXAL [Wang *et al.*, 1989; Cu(2), BVS 2.00] has two extra H atoms, such that the ligand appears as a diketone rather than a diketonate in the CSD. Conversely, PIBVOL [Carrier *et al.*, 1993; Cu(1), BVS 0.92 *versus* 1.49 for Cu(2)] has a missing H atom in the CSD (both in the formula and chemical connectivity) such that the ligand templates suggest Cu(2). These discrepancies will be corrected in future releases of the CSD. In [(*N,N'*-2,2'-bipyridyl)(η^2-ethene)Cu] FOMHEU and [(*N,N'*-1,10-phenanthroline)(η^2-ethene)Cu] FOMHIY (Masuda *et al.*, 1987) the alkene ligands are represented as dialkyls, implying a Cu^{III} oxidation state, whereas the BVS are 1.12 and 1.02, respectively, for the correct Cu(1) oxidation state. Some caution should be exercised, however, since the BVS parameters for [Cu(3)(CF$_3$)$_4$] fit for both Cu(3) and Cu(1). The ligand templates may fail to give the correct result if the original chemical and crystal connectivity were not matched, so that the charges are not carried over into the new connectivity: BGCUCB (Coghi *et al.*, 1978), GEHDIG (Lei *et al.*, 1988) and GEHDIG10 (Huang *et al.*, 1990) suffer from this sort of error which is an unfortunate artefact of the recalculation procedure (§3).

In some cases discrepancies may suggest a reinterpretation of the original data. LAPTOL (Nyburg *et al.*, 1993) is described as a Cu(2) amido complex, *viz.* [(Ph$_3$P)(NH-CH^{i}Pr$_2$)$_2$(μ_2-Cl)$_2$Cu$_2$]. However, BVS suggests a Cu(1) complex (BVS 1.00). Examination of the crystal data shows

Table 4
Oxidation-state validation statistics for 743 four-coordinate Cu complexes.

| | Assigned using ligand templates (LT) and BVS | | | | Assigned using BVS alone | | |
| | BVS correct | | BVS incorrect | | | | |
	LTs correct	LTs incorrect	LTs correct	LTs incorrect	Correct	Incorrect	Total
Cu(1)	362	6	15	0	40	3	426
Cu(2)	269	7	6	3	18	0	303
Cu(3)	12	0	0	0	2	0	14
Total	443	13	21	3	60	3	743
%	86.5	1.8	2.8	0.4	8.1	0.4	100

that the geometry of the α-carbon is trigonal planar (the valence angles about the α-carbon are 119.2, 119.8 and 121.0°) and it would appear that the additional H atom on this carbon should not have been included in the model. This, and the 1.30 Å N=C bond length, suggest a Cu(1) imine, which would provide a much better fit to the Cu—L bond lengths and bond angles. In particular, there is little evidence for the D_{2d} flattening observed in other Cu(2) bridged dimers, *e.g.* [(2,2,6,6-piperidinoxo-O)$_2$Cl$_2$(μ_2-Cl)$_2$Cu$_2$] JOGKEV (Laugier *et al.*, 1991). However, the bulk properties (colour and paramagnetism) suggest Cu(2); it is possible that the crystal does not represent the majority composition of the bulk sample. The situation in [(Ph$_2$C$_5$H$_5$NP)$_2$Cl$_2$(μ_2-Cl)$_2$Cu] FAHHIF (Wang *et al.*, 1986) is similar; whilst apparently Cu(2), the BVS is closer for Cu(1) (BVS 1.06, 1.04) and the Cl$_2$ dimer geometry does not suggest Cu(2). It is possible that the pyridyl N atoms are in fact protonated. The disagreement in CATVUO (Corain *et al.*, 1983) is due to the O and N atoms in the ligand being misassigned (*cf.* CATVUO10 for correction: Basato *et al.*, 1992).

There are several Cu(1) and Cu(2) structures in which the BVS deviates by more than ±0.15 from the predicted value, and fits the incorrect oxidation state more closely. In some cases this is due to the constraints imposed by rigid or semi-rigid macrocyclic ligands, *e.g.* porphyrinato complexes C$_{44}$Cl$_8$F$_{20}$N$_4$Cu LALROF (Schaefer *et al.*, 1993) and VUHROF (Polyakova *et al.*, 1991). In other structures the reason for the discrepancy is less clear, and may indicate a problem with the structure determination. Whilst the structures have not been examined in detail, low BVS values for Cu(2) may indicate unusually long bonds (outside the distance limits described in §5). In a few instances the best-fit BVS deviates substantially from an integer value. In Cu(1)(PPh$_3$)$_4$ KIDFUY (Bowmaker *et al.*, 1990) (BVS 0.51) this can probably be attributed to steric effects [this complex was omitted from the derivation of R_0 values for Cu(1)—P bonds]. However, in other structures, *e.g.* BOJYAA [(2.39), macrocyclic N_4 ring: Matsumoto *et al.*, 1982], BUACUM ({[S,S'-NCC(S)C(S)CN]$_2$Cu(2)$^{2-}$}, 3.134: Forrester *et al.*, 1964), CUVYUN (*O,N,N',O* chelating ligand, 2.59: Atkins *et al.*, 1985), JUKXES (CuBr$_4$$^{2-}$ counteranion, 2.47: Kahwa *et al.*, 1993) and GATDUA {[Br$_2$Cu(*O,N*-2,2,6,6-tetramethylpiperidnyl-1-oxy)], 2.493: Caneschi *et al.*, 1988}, it may indicate disorder, high librational motion or problems with the structure determination.

7. Conclusions

250 ligand fragment templates have been derived which accurately identify substructural metal–ligand fragments in metal-containing structures. Oxidation states and valence electron counts are calculated for structures having all donor atoms identified, on the basis of the donor properties of the matched ligand fragments, the metal and ligand charges and the metal–metal connectivity. These values are cross-validated with bond-valence sums, based on metal coordination environments recalculated in a consistent manner. The program has been tested successfully on a subset of four-coordinate Cu compounds, providing correct assignments for *ca* 99% of entries, including those with complex multidentate ligands. This figure is better than that which would have been achieved with ligand templates or BVS alone; in addition, structures in which there were errors in the CSD chemical connectivity description were readily identified, and a reinterpretation of the original structure determination was suggested in a few cases. Whilst the approach is not entirely automatic, it does distinguish between structures which can have oxidation states readily assigned (*i.e.* the 86.5% of structures for which ligand templates and BVS agree) and those which should be investigated in more detail.

The main problems arose with charge assignments in structures in which the original chemical and crystallographic connectivities were not matched, structures containing other metals for which good R_0 values were not available, and complexes having bond lengths severely strained by interligand steric effects. In order to apply the BVS aspect of the methodology to metals other than Cu, it will be necessary to derive accurate R_0 values for other elements. Research in this area by a number of workers is in progress, including studies on *p*-block metals (Palenik, 1999), and these parameters could be incorporated as they become available. In the meantime, better results could probably be obtained by using the R_0 values of Brown & Altermatt (1985) and Brese & O'Keeffe (1991) rather than atom-based parameters (O'Keeffe & Brese, 1991).

The changes to *QUEST3D* to permit user-definition of metal coordination sphere connectivities, described in §3, now form part of standard releases of the CSD system. The code for oxidation-state assignment, which requires some further development as indicated above, will be introduced into CSD validation and database creation operations as soon as practicable.

References

Adams, S. (1999). Private communication.

Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.

Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 1, 31–37.

Allen, F. H., Kennard, O., Motherwell, W. D. S., Town, W. G., Watson, D. G., Scott, T. J. & Larson, A. C. (1974). *J. Appl. Cryst.* **7**, 73–78.

Atkins, R., Brewer, G., Kokot, E., Mockler, G. M. & Sinn, E. (1985). *Inorg. Chem.* **24**, 127–134.

Basato, M., Corain, B. & Zanotti, G. (1992). *J. Chem. Soc. Dalton Trans.* pp. 3003–3007.

Bergerhoff, G., Hundt, R., Sievers, R. & Brown, I. D. (1983). *J. Chem. Inf. Comput. Sci.* **23**, 66–69.

Bonamico, M., Fares, V., Flamini, A., Imperatori, P. & Poli, N. (1989). *Angew. Chem. Int. Ed. Engl.* **28**, 1049–1050.

Bonamico, M., Fares, V., Flamini, A., Poli, N. & Mattei, G. (1993). *Polyhedron*, **12**, 1209–1214.

Bowmaker, G. A., Healy, P. C., Engelhardt, L. M., Kildea, J. D., Skelton, B. W. & White, A. H. (1990). *Aust. J. Chem.* **43**, 1697–1705.

Brese, N. E. & O'Keeffe, M. (1991). *Acta Cryst.* B**47**, 192–197.

Brewer, G. A. & Sinn, E. (1987). *Inorg. Chem.* **26**, 1529–1535.

Brown, D. S., Lee, J. D., Melsom, B. G. A., Hathaway, B. J., Procter, I. M. & Tomlinson, A. A. G. (1967). *J. Chem. Soc. Chem. Commun.* pp. 369–371.

Brown, I. D. (1992). *Acta Cryst.* B**48**, 553–572.

Brown, I. D. (1996). *J. Appl. Cryst.* **29**, 479–480.

Brown, I. D. (1999). Private communication.

Brown, I. D. (2000). *J. Chem. Educ.* Submitted.

Brown, I. D. & Altermatt, D. (1985). *Acta Cryst.* B**41**, 244–247.

Cambridge Structural Database (1994). *CSD User's Manual*. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England.

Caneschi, A., Grand, A., Laugier, J., Rey, P. & Subra, R. (1988). *J. Am. Chem. Soc.* **110**, 2307–2309.

Carrier, S. M., Ruggiero, C. E., Houser, R. P. & Tolman, W. B. (1993). *Inorg. Chem.* **32**, 4889–4899.

Chang, R. (1981). *Chemistry*, 2nd ed., p. 335. New York: Random House.

Coghi, L., Lanfranchi, M., Pelizzi, G. & Tarasconi, P. (1978). *Transit. Met. Chem.* **3**, 69–76.

Corain, B., Basato, M., Del Zotto, A. & Zanotti, G. (1983). *Inorg. Chem.* **22**, 2744–2749.

Donnay, G. & Allmann, R. (1970). *Am. Mineral.* **55**, 1003–1015.

Donnay, G. & Donnay, J. H. D. (1973). *Acta Cryst.* B**29**, 1417–1425.

Forrester, J. D., Zalkin, A. & Templeton, D. H. (1964). *Inorg. Chem.* **3**, 1507–1515.

Greenwood, N. N. & Earnshaw, A. (1984). *Chemistry of the Elements*, p. 31. Oxford: Pergamon Press.

Harding, M. H. (1999). *Acta Cryst.* D**55**, 1432–1443.

Hati, S. & Datta, D. (1995). *J. Chem. Soc. Dalton Trans.* pp. 1177–1182.

Huang, Z., Lei, X., Kang, B., Liu, J., Liu, Q., Hong, M. & Liu, H. (1990). *Inorg. Chim. Acta*, **169**, 25–29.

Humphrey, D. G., Fallon, G. D. & Murray, K. S. (1988). *Chem. Commun.* pp. 1356–1358.

Kahwa, I. A., Miller, D., Mitchel, M. & Fronczek, F. R. (1993). *Acta Cryst.* C**49**, 320–321.

Kanowitz, S. M. & Palenik, G. J. (1998). *Inorg. Chem.* **37**, 2086–2089.

Laugier, J., Latour, J.-M., Caneschi, A. & Rey, P. (1991). *Inorg. Chem.* **30**, 4474–4477.

Lei, X., Liu, Q. & Liu, H. (1988). *Jiegou Huaxue (J. Struct. Chem.)*, **7**, 99–102.

Lin, F. & Lin, C. (1991). *Wuji Huaxue Xuebao (J. Inorg. Chem.)*, **7**, 389–393.

Liu, W. & Thorp, H. H. (1993). *Inorg. Chem.* **32**, 4102–4105.

McGhee, E. M., Godfrey, M. R., Hoffman, B. M. & Ibers, J. A. (1991). *Inorg. Chem.* **30**, 803–808.

Masuda, H., Yamamoto, N., Taga, T., Machida, K., Kitagawa, S. & Munakata, M. (1987). *J. Organomet. Chem.* **322**, 121–129.

Matsumoto, N., Imaizumi, M. & Ohyoshi, A. (1982). *Bull. Chem. Soc. Jpn*, **55**, 3170–3174.

Naskar, J. P., Hati, S. & Datta, D. (1997). *Acta Cryst.* B**53**, 885–894.

Nyburg, S. C., Parkins, A. W. & Sidi-Boumedine, M. (1993). *Polyhedron*, **12**, 1119–1122.

O'Keeffe, M. & Brese, N. E. (1991). *J. Am. Chem. Soc.* **113**, 3226–3229.

Palenik, G. J. (1997a). *Inorg. Chem.* **36**, 122.

Palenik, G. J. (1997b). *Inorg. Chem.* **36**, 3394–3397.

Palenik, G. J. (1997c). *Inorg. Chem.* **36**, 4888–4890.

Palenik, G. J. (1999). Personal communication.

Pauling, L. (1929). *J. Am. Chem. Soc.* **51**, 1010–1026.

Polyakova, I. N., Starikova, Z. A. & Fakeeva, O. A. (1991). *Kristallografiya*, **36**, 1167–1171.

Presier, C., Lösel, J., Brown, I. D., Kunz, M. & Skowron, A. (1999). *Acta Cryst.* B**55**, 698–711.

Schaefer, W. P., Hodge, J. A., Hughes, M. E., Gray, H. B., Lyons, J. E., Ellis, P. E. Jr & Wagner, R. W. (1993). *Acta Cryst.* C**49**, 1342–1345.

See, R. F., Kruse, R. A. & Strub, W. M. (1998). *Inorg. Chem.* **37**, 5369–5375.

Serr, B. R., Headford, C. E. L., Anderson, O. P., Elliott, C. M., Spartalian, K., Fainzilberg, V. E., Hatfield, W. E., Rohrs, B. R., Eaton, S. S. & Eaton, G. R. (1992). *Inorg. Chem.* **31**, 5450–5465.

Urusov, V. S. (1995). *Acta Cryst.* B**51**, 641–649.

Wang, B., Wu, J., Zheng, P., Zhang, H. & Zhang, H. (1989). *Fud. Xue. Ziran Kex. (Fudan U.)*, **28**, 24–28.

Wang, H., Wang, R., Zhang, Z. & Wang, X., (1986). *Jiegou Huaxue (J. Struct. Chem.)*, **5**, 94–96.

Wood, R. M. & Palenik, G. J. (1998). *Inorg. Chem.* **37**, 4149–4151.

Zheng, Y., Zhan, H., Wu, X. & Lu, J. (1989). *Transit. Met. Chem.* **14**, 161–164.

Zheng, Y., Zhu, N. & Wu, X. (1989). *Acta Cryst.* C**45**, 1988–1990.