Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Applications of the Cambridge Structural Database to molecular inorganic chemistry

A. Guy Orpen

School of Chemistry, University of Bristol, Bristol BS8 1TS, England

Correspondence e-mail: guy.orpen@bristol.ac.uk

Received 20 December 2001
Accepted 21 February 2002

Applications of the data in the Cambridge Structural Database (CSD) to knowledge acquisition and fundamental research in molecular inorganic chemistry are reviewed. Various classes of application are identified, including the derivation of typical molecular dimensions and their variability and transferability, the derivation and testing of theories of molecular structure and bonding, the identification of reaction paths and related conformational analyses based on the structure correlation hypothesis, and the identification of common and presumably energetically favourable intermolecular interactions. In many of these areas, the availability of plentiful structural data from the CSD is set against the emergence of high-quality computational data on the geometry and energy of inorganic complexes.

1. Introduction

This paper reviews the scope of applications of the Cambridge Structural Database (CSD; Allen & Kennard, 1993; Allen, 2002; Bruno *et al.*, 2002) to molecular inorganic chemistry. Excellent reviews of this field have appeared in 1994 (Auf der Heyde, 1994*b*) and 1998 (Bürgi, 1998) and the main focus of this paper will be on work reported since 1993. The range of uses to which the CSD has been put in this field reflects the growing maturity of the crystallographic literature. Single-crystal X-ray diffraction analysis has, in many areas of molecular inorganic chemistry, become the method of choice for the identification of new products from synthetic chemistry programs. Indeed, such is the speed and power of the technique that it is often quicker and more reliable to determine the crystal structure (data collection in the morning with an area detector instrument, solution and refinement by the end of the afternoon!) than it is to measure and interpret multinuclear or multidimensional NMR spectra or other 'sporting' characterization data. This tendency is particularly pronounced in those areas of inorganic chemistry where spectroscopic data are often of limited value in structure assignment (*p*-block chemistry of elements other than P or F; paramagnetic species; cluster chemistry *etc.*).

In reality, of course, these techniques are in many ways complementary, but the upshot of modern practice is a flood of three-dimensional crystal structure data into the CSD, which continues to double in size every 7–8 years and now holds over 250000 structures of carbon-containing molecular species (see Fig. 1). Over 50% of these are metal-containing and fall within

the remit of this review. These, nearly 140000, structures represent an enormous store of information on coordination, organometallic and main-group element chemistry, *i.e.* the chemistry of the *s*-, *p*-, *d*- and *f*-block elements. The focus of the work discussed in this paper is therefore the study of collections of structures, and the relationship of these structures to one another and to other aspects of the chemistry of these compounds, rather than a discussion of individual structures. The growing and parallel maturity of various aspects of theoretical and computational chemistry offers new opportunities to use the CSD as a complement to these techniques, and to reinforce arguments based on those techniques by comparison with the experimental evidence available in the CSD.

The report below is divided into various themes, which are, in part, overlapping and interconnected. They cover applications ranging from the prosaic (extracting a typical bond length value *etc.*) to the more ambitious, such as establishing a reaction mechanism on the basis of a series of related structures. These share the common theme of being based on taking a view of a series of structures with some common component rather than focusing on individual structures.

2. Molecular dimensions from single-crystal X-ray data and their reliability

The CSD provides a database of molecular dimensions which may be of value in situations where direct determination of the geometry of a molecular species is either unreliable or impossible. When processed into knowledge (*i.e.* appropriate data extracted, errors corrected, data validated *etc.*, statistics determined) then a range of applications in molecular model building may be envisaged. Two major studies (Allen *et al.*, 1987; Orpen *et al.*, 1989; Orpen, 1998) presented typical intramolecular dimensions for, firstly, organic, and secondly, *d*- and *f*-block-element compounds. These analyses were carried out on a case-by-case basis for earlier and much smaller versions of the CSD that contained about 20% of the data now available. More recent attempts to distil the contents of the CSD into useful knowledge bases have focused on the automated generation of intermolecular geometry and intramolecular geometry libraries: ISOSTAR (Bruno *et al.*, 1997) and MOGUL (Taylor *et al.*, 2001), respectively. Here the objective is to build into the knowledge base suitable subsets of the data present in the CSD arranged by chemical context.

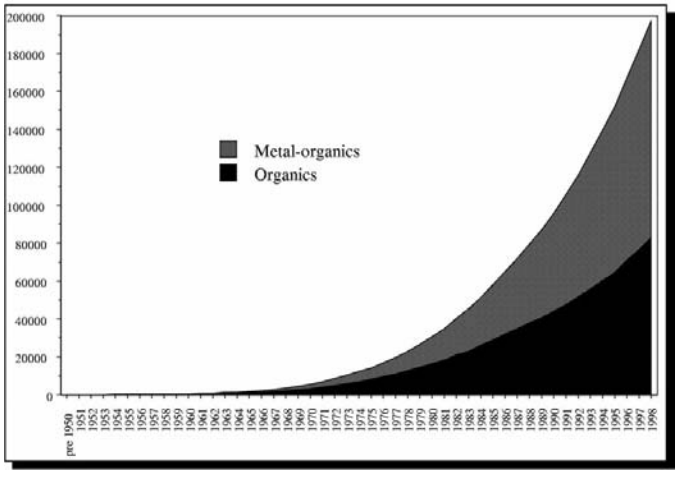

Figure 1
Growth in the number of structures recorded in the CSD, 1950–1998.

These knowledge bases serve vital, if prosaic, functions, as follows.

(i) Data validation through comparison of molecular geometry in a partly or recently determined crystal structure, or X-ray absorption fine structure (XAFS), NMR spectroscopy or computationally derived structure determination. Many such studies have been reported, notably when developing new force fields for molecular-mechanics procedures or when evaluating the accuracy of quantum-mechanical methods of geometry optimization (see Cundari *et al.*, 2000; Burton & Deeth, 1995; Burton *et al.*, 1995). On occasion, the converse occurs and misleading crystallographically determined data are corrected using computational information (Kahr *et al.*, 1995).

(ii) Model building: the provision of geometry for incorporation into a structural model, *e.g.* for use in constrained protein structure refinement based on an 'ideal' fragment geometry, or in the solution and refinement of structures from powder diffraction data ('*ab initio*' but based on known and essentially rigid molecular or sub-molecular components). The development of reliable models for protein structure refinement on the basis of the relatively high resolution single-crystal diffraction data on small molecules in the CSD has been the main focus of such studies.

(iii) Information to allow restrained geometry refinement for macromolecular or quasi-macromolecular structure analysis by X-ray methods. Here the object is to add information content to the refinement process while allowing greater flexibility of the refinement models than is possible with a constrained model (see Harding, 1999, 2000, 2001).

(iv) Generating models for optimization using say *ab initio* molecular orbital (MO) or other theories starting from a reasonable molecular geometry based on data observed in the CSD.

(v) Improved parameterization of empirical or semi-empirical modelling procedures by incorporation of structural data in the optimization process (see *e.g.* Halgren, 1996).

All of these functions are based on the fundamental assumptions that the (partial) molecular geometry excised from one structure is somehow transferable to a different crystal structure or molecule. The validity of this assumption may be questioned and it must be tested in detail, although over 30 years ago Kitaigorodskii (1970) had no doubts about its value in the context of organic chemistry, stating without apparent fear that 'the crystalline field does not change the bond lengths [or angles] of organic molecules'. It has been tested whether the dimensions of organometallics, coordination complexes and related species are indeed independent of the crystal environment in which they are determined (Martin

& Orpen, 1996; Cotton & Yokochi, 1997; Orpen & Quayle, 2001). The outcome is that there appears to be a minimum baseline uncertainty in metal–ligand bond lengths of *ca* 0.01 Å and in valence angles of *ca* 1–2°. These values apply to the most rigid and kinetically inert of metal–ligand bonds ($[PtCl_4]^{2-}$ for example) and much higher values of uncertainty are observed for more easily deformed entities (such as $[CuCl_4]^{2-}$, low-order metal–metal bonds, 14e *p*-block species such as $[TeCl_6]^{2-}$ *etc.*; see Fig. 2 and Orpen & Quayle, 2001). It appears that in addition to the usual experimental uncertainties, which derive from uncertainties in the observations of the diffraction experiment, there are significant contributions from the effects of the crystalline field – the notorious 'packing effects' – on the molecular structure. There is some (very weak) correlation between the softness (*i.e.* inverse of force constant) of a molecular parameter and the magnitude of the uncertainty that is associated with it, when evaluated by comparison with examples of the same parameter in different crystal environments, and this effect is most notable for torsion angles. This assumption of transferability is that a bond parameter has an ideal 'natural' value, presumably close to that which the structure would exhibit in the gas phase. This assumption has been tested with regard to torsion angles by Allen *et al.* (1996).

3. Models of structure and bonding

The CSD offers a wealth of structural evidence against which to test a given structural hypothesis or to provide inspiration for the development of new structural models. These models may take a qualitative form (the hypothesis being of the type: bond length *x* will increase as bond angle *y* decreases) or a quantitative form, in which a more explicit correlation of molecular dimensions is posited or a computational model based on, for example, quantum-mechanical methods is tested against the experimental data available. A variety of areas have been addressed in this spirit.

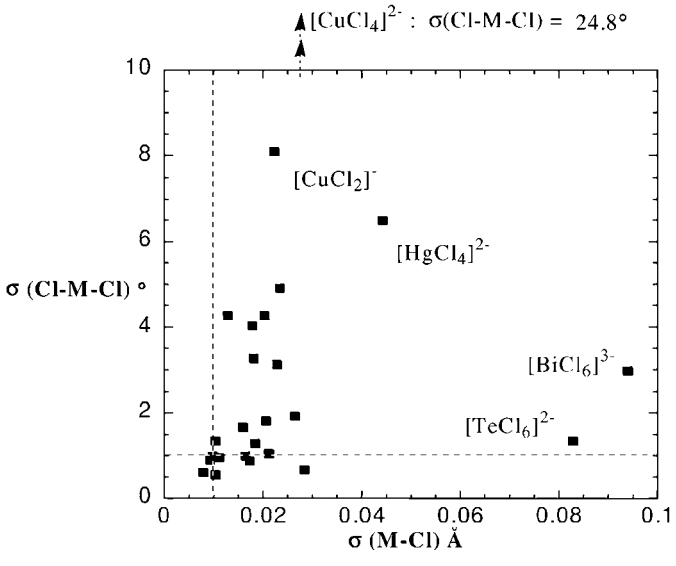

Figure 2
Uncertainties of $M-$Cl bond lengths and Cl$-M-$Cl bond angles in salts of perhalometallate anions (Orpen & Quayle, 2001).

3.1. Metal–ligand bonding

One important family of studies in this area concerned the effects of a metal (and its co-ligands) on the geometry of a ligand or family of ligands. This has led to the development of models of the nature of the metal–ligand bond that can be tested against the observed patterns of behaviour in the data extracted from the CSD. The thesis is that we can probe the nature of the L'_nM-L interaction by varying M (and L' too, in principle). The effects of M (and its oxidation state, electron count, co-ligands *etc.*) on the geometry of the ligand L and the $M-L$ bond length can yield useful information on the nature of the $M-L$ interaction. In some instances, it may be helpful to include a wider range of moieties at the M site and not just M = transition metal, so that one can inspect cases in which M is one of a broader range of Lewis acids, such as H^+, O, CR_3^+. The qualitative models used in this class of study are typically based on two aspects of the $M-L$ interaction, as follows.

(i) L'_nM-L σ donation. This is typically the primary bonding interaction and is likely to be the main influence on the geometry of L. However, it is noteworthy that the effects on ligand geometry may not be directly related to the effects on $M-L$ bond energy.

(ii) L'_nM-L π bonding (often back donation). This is typically a secondary effect but may be important both in terms of the $M-L$ interaction and in its effects on the geometry of L.

Interaction (i) gives information on the ligand highest occupied molecular orbital (HOMO), while (ii) gives information on the lowest occupied molecular orbital(s) [LUMO(s)] of the ligand, when it acts as a π acceptor. Thus, variation of M (and/or L') will lead to modification of the extent of interactions (i) and (ii), and hence provide an opportunity for monitoring the effect of depopulating the ligand HOMO (as M becomes a better electron pair σ acceptor) or the population of the ligand LUMO(s) when the ML'_n fragment acts as a better π donor.

This class of study has implications for ligand design in that as the nature of the $M-L$ bond becomes better understood, it may be possible to design the ligand L so as to increase, for example, the strength of the $M-L$ bond at the expense of one or more of the $M-L'$ interactions and hence increase reactivity at the $M-L'$ bond. In our hands, these studies have focused on metal–phosphine and related bonds (see *e.g.* Garner & Orpen, 1993; Crispini *et al.*, 1996). Others have studied the behaviour of ligands that have variable oxidation state, which can be determined from the structural data on the basis of the ligand geometry (Carugo, 1994). Cundari has recently reported novel approaches to the study of metal–ligand multiple bonds (Cundari & Russo, 2001).

3.2. Ligand effects on Lewis acid geometries

The converse approach to that described above may be based on the same bonding model [interactions (i) and (ii) as above], but now the information (i) is on the L'_nM fragment geometry and its σ acceptor orbital, while (ii) leads to information on the L'_nM fragment orbitals of $M-L$ π symmetry.

The geometry of the Ph_3B-NC interaction in $[Mn(CNBPh_3)(PR_3)(NO)(\eta\text{-}C_5H_4Me)]$ and $L\cdot BPh_3$ fragments in the CSD (where L is a neutral or anionic ligand bound to B) was studied recently (Bellamy, Brown *et al.*, 1999). The donor strength of L was characterized according to its ability to distort the BPh_3 unit from planarity. Stronger donor ligands lead to occupancy of the BPh_3 LUMO and cause distortion towards a pyramidal geometry with smaller Ph—B—Ph angles and longer B—C bonds. A negative correlation between B—C bond length and C—B—C bond angles is observed. In a similar study of the structure of $[Mn(CNSbCl_5)(CO)\{P(OEt)_3\}(dppm)]$ and other $L\cdot SbCl_5$ complexes in the CSD, we sought to use the geometry of the $SbCl_5$ fragment as a measure of the ligand donor strength (Bellamy, Connelly *et al.*, 1999). On the basis of these studies, the cyanomanganese ligand was attributed as a weak donor as it only caused a slight distortion of the BPh_3 and $SbCl_5$ moieties. In both cases (and the phosphine studies noted above), bonding models that used Walsh diagrams were employed to give an understanding of the link between orbital occupancy and geometry.

The study of $L\cdot SbCl_5$ complexes also led to the observation that *cis* and *trans* Sb—Cl bond lengths vary almost equally with variation in L. This is in marked contrast to the received wisdom about the dominant role of *trans* influence over *cis* influence in metal complex chemistry. Recently, a simple qualitative model was presented to interpret the variation in the relative magnitudes of *trans* and *cis* influence observed as a function of electron configuration in a range of *d*- and *p*-block metal complexes (Anderson & Orpen, 2001). This moves on from the traditional view of *trans* influence, which is based mainly on experience derived from d^8-ML_4 square-planar (Bugarcic *et al.*, 1993) and to a lesser extent d^6-ML_6 octahedral species (Coe & Glenwright, 2000).

In a major series of studies, Alvarez and co-workers have focused on a variety of aspects of metal–metal and metal–ligand interactions to develop and test models based on both qualitative and quantitative MO theory (Aullon & Alvarez, 1993, 1996, 1997; Alvarez & Aullon, 1999; Aullon *et al.*, 2000; Liu & Alvarez, 1997, Mota *et al.*, 1993; Palacios *et al.*, 2000). In general, the focus has been on offering qualitative understanding as well as testing more quantitative levels of theory against the crystallographic data presented in the CSD.

4. Quantitative models

A number of groups have exploited Pauling's semi-empirical Bond Valence Model to analyse bonding and electronic structure. This allows (i) rationalization of the variation of bond lengths in transition and other metal complexes (See *et al.*, 1998; Palenik, 1997; Wood *et al.*, 2000; Wood & Palenik, 1998, 1999; Jensen *et al.*, 2001) and (ii) assignment of oxidation states in metal complexes (notably in cases where the ligand identity is ambiguous or unhelpful in assigning the oxidation state of the metal in a structure; Shields *et al.*, 2000).

As noted above, the CSD is an invaluable data store for the parameterization, calibration or validation of quantitative computational models of structure, whether they are based on molecular mechanics or on one or more flavours of quantum mechanics, or even combinations of these methods (Burton & Deeth, 1995; Burton *et al.*, 1995; Rappe *et al.*, 1992).

In many areas, a key goal is the development of a link between the structure of a complex and its properties. Perhaps the two most striking areas of application for our purposes are in the context of the magnetic behaviour of transition metal complexes and the catalytic behaviour of phosphine complexes of the late transition elements.

4.1. Magnetism

The correlation of magnetic behaviour and magnetic coupling has long been established in general, and a number of studies have sought to place this on a sounder footing by inspecting the correlation between aspects of the molecular geometry of copper(II) (and other transition metal) complexes and their observed magnetic properties (Ruiz *et al.*, 1997; Cano *et al.*, 2000).

4.2. Phosphine catalysts

Phosphine ligands are of enormous significance in homogeneous catalysis, notably in commercialized hydrogenation, carbonylation and hydrocyanation processes. Their success is based on the fact that the electronic and steric properties of the phosphine can be readily manipulated by adjustment of the substituents at the P atom. The range of substituent types is enormous and the consequent breadth of phosphine (and related phosphite, phosphinite *etc.*) chemistry that has been developed is considerable. While it is clear that relatively few 'active' catalysts have themselves been the subject of crystal structure analysis, in many instances the ligand that is responsible for controlling the catalytic behaviour of the metal species has been studied quite frequently. This paves the way for studies with the long-term goal of developing models that enable the design of improved ligand systems to support catalysis.

Some key parameters that have been cited as important in determining the nature of the phosphine–metal bond and the catalytic properties of the complexes, characterized by crystallography, can be established from data in the CSD. These include the following.

(i) Ligand cone angles (developed by Tolman, 1977, as a measure of ligand steric influence) may be readily determined from CSD data (Mueller & Mingos, 1995; Smith *et al.*, 1997).

(ii) The 'natural' bite angles of chelating ligands. This conceptual device has been studied by van Leeuwen and others (Dierkes & van Leeuwen, 1999) as a determining factor in selectivity in hydroformylation reactions. The real bite angles and their ranges and variability can be derived directly

from the CSD and compared with computational models of the same ligand systems to provide corroborating evidence for the model(s).

(iii) The range of conformation of flexible ligands. The great majority of phosphine ligands that are of interest in homogeneous catalysis are flexible by virtue of the range of conformations that the substituents at P can adopt or the flexibility of saturated chelate rings. Here we investigate the likely conformations and their range of availability and hence obtain some idea of the 'pocket' into which substrates bind in the catalytic reaction cycles of metal–phosphine complexes (see *e.g.* Kadyrov *et al.*, 1999; Brunner *et al.*, 1998; Harris *et al.*, 2001; Beyreuther *et al.*, 1996; Morton & Orpen, 1992).

The outcome of these studies is progress towards qualitative, and eventually quantitative models that relate activity, and other parameters indicating usefulness of a catalyst system, to the structure. Such models, for example the quantitative analysis of ligand effects (QALE) by Prock, Giering and co-workers (see *e.g.* Fernandez *et al.*, 2000), are challenging, not least because of the fleeting nature of some catalytic cycles and their transition states. In contrast to biological catalysts, where the complexity and bulk of the enzyme preclude dramatic changes in the rate-determining step or other aspects of the mechanism, in homogeneous catalysis a reaction step may well cease to be rate-determining as the catalyst (or the ligand on which it is based) is elaborated away from the known system.

5. Reaction pathway analysis

The core of the structure correlation hypothesis, advanced by Bürgi & Dunitz in some classic papers from 1973 onwards (Bürgi & Dunitz, 1983), is that for a sub-molecular fragment, the local molecular or crystal environment acts as a perturbation on the fragment or molecular geometry and therefore samples the range of energetically accessible geometries that the fragment might adopt. This hypothesis has been developed to afford insight into the nature of the structure–energy correlation and applied to a range of systems, both organic and organometallic in nature.

One major strand of these studies has been the variation of geometry within the coordination sphere of metals. This has taken a variety of forms, derived from studies of large collections of four-, five- and higher coordinate metals. The central aims have been: (i) to investigate S_N1, S_N2 and other reaction mechanisms of ligand substitution and exchange (see *e.g.* Bürgi & Dunitz, 1983; Yao *et al.*, 2001) and (ii) to follow unimolecular reactions involving *e.g.* the interconversion of alternative coordination geometries (see *e.g.* Alvarez & Llunell, 2000; Alvarez, Pinsky, Llunell & Avnir, 2001; Alvarez, Pinsky & Avnir, 2001; Raithby *et al.*, 2000).

In earlier studies, Bürgi, Dunitz and others exploited symmetry-adapted distortion measures (see Dunitz, 1979) and principal component analysis and other statistical techniques to extract information from the mass of data (see Auf der Heyde, 1994*a*).

In more recent examples, Avnir & Alvarez and their co-workers have used continuous symmetry techniques to quantify the range of distortions from ideal geometries (Alvarez & Llunell, 2000; Alvarez, Pinsky, Llunell & Avnir, 2001; Alvarez, Pinsky & Avnir, 2001; Zabrodsky & Avnir, 1995; Keinan & Avnir, 2001*a*,*b*). Allen, Howard and co-workers have used alternative methods to quantify the distortions of coordination complexes from archetype geometries (Yao *et al.*, 2001).

In work that did not explicitly use the CSD, Crabtree & Lavin (1986) studied the trajectory of CO exchange in dinuclear iron complexes using the structure correlation method, looking at the geometry of a single carbonyl ligand in a variety of di-iron species. In contrast, Johnson, Mann and others have used structure correlation methods to investigate the mechanism of CO exchange processes in metal cluster carbonyl species in which multiple CO ligands undergo site-exchange simultaneously (see Mann, 1997; Johnson, 1997, and references cited therein)

Intermolecular reaction studies on inorganic species have been less common and include the work of Crabtree on alkane C—H activation through direct interaction with transition metals (Crabtree *et al.*, 1985) and that of Brammer on proton transfer in NH$\cdots$Co and N$\cdots$H–Co species (Brammer *et al.*, 2000).

6. Conformational analysis

The crystal structure analysis of an individual complex only determines the equilibrium geometry of that species in a particular crystal field. There is no information about possible conformations far from this equilibrium position. As noted above, studies of collections of crystal structures containing the same or similar (sub-) molecular fragments offers the opportunity to examine how the crystal field perturbs the geometry of the fragment. In the case of conformationally flexible molecules (which are ubiquitous in inorganic chemistry) the variation in crystal environment will allow a more complete sampling of the conformational potential-energy hypersurface than in many of the 'reaction path' analyses noted above. Bürgi & Dunitz (1983) noted this early and studied some key conformation interconversions. The objectives in this class of study include: (i) identification of common (preferred) conformers; (ii) interpretation of the pattern of common conformers and trails of observed structures between them in terms of the characteristics of the potential-energy hypersurface; (iii) comparison of the distribution of structures with the computed potential-energy hypersurface; (iv) links between the observed conformations and other aspects of the molecular geometry, such as the coordination stereochemistry of the metal to which a flexible ligand is coordinated.

Most studies of this sort have focused on the variation in conformation in metal complexes of flexible ligands. For the most part, the ligand systems studied have been either (saturated) chelate species, including macrocycles (see Harris *et al.*, 2001; Beyreuther *et al.*, 1996; Morton & Orpen, 1992; Donnelly & Zimmer, 1999; Zimmer, 2001; Raithby *et al.*, 1997*a*,*b*;

Leuwerink *et al.*, 1993) or the coupled rotations of aryl substituents attached to the ligand contact atoms (Barker & Orpen, 1999; Costello & Davies, 1998; Costello *et al.*, 1999; Hunger *et al.*, 1998). Relatively few studies have been conducted on ligand systems with alkyl substituents (Smith & Coville, 2001; Smith *et al.*, 1998) or rather rigid systems such as porphyrins (but see Cullen *et al.*, 2001).

The range of statistical methods employed is notable: principal component analysis, neural networks and cluster analysis of various sorts have been used in addition to more traditional methods, such as graphical (scatterplots) and descriptive statistics based on torsion angles. Other specially derived approaches have been used for five-membered rings (Altona–Sundaralingam; Altona & Sundaralingam, 1972) and other cyclic systems (Cremer–Pople puckering parameters; Cremer & Pople, 1975). In all cases, the symmetry of the conformation space, as described by Longuet-Higgins, Mislow and others (Longuet-Higgins, 1963; Mislow, 1966), and applied to this field of study by Bürgi, Dunitz and others (Dunitz, 1979; Bürgi & Dunitz, 1983) must be considered. The objectives of using principal component analysis is to identify the 'best' way to view the data set by projecting the usually high-dimensional data set onto a more efficient set of mutually orthogonal axes which are linear combinations of the original (torsion angle) axes. In cluster analysis, objective groupings of structures are sought, usually in order to identify sets of similar conformations corresponding to distinct conformers. Neural-network methods have been used to map the connections between such groups of structures (Beyreuther *et al.*, 1996).

7. Intermolecular interactions in crystals

Crystallography provides the best quality information available on the geometry and popularity of interactions between functional groups. This has meant that many studies have been made of the geometry and prevalence of various intermolecular interactions in crystals, perhaps most notably the hydrogen bond in its various forms. The rapid rise of supramolecular chemistry and related fields, such as pharmacophore design, crystal engineering (or more specifically crystal synthesis based on molecular species) and the like, has reinforced interest in this well established type of database study. In many instances, these studies have attempted to quantify the importance and characteristics (primarily geometric) of these interactions. By far the most studied interaction type is hydrogen bonding in its many manifestations, but a range of more exotic interaction types peculiar to inorganic chemistry have also attracted attention.

7.1. Hydrogen bonds

While there is an immense literature on hydrogen bonding based on solid-state organic chemistry, there is a substantial body which has its origins in molecular inorganic chemistry. The inorganic context has proved, as one might instinctively expect, to be a rich environment in which to explore exotic hydrogen-bond types. Among those that have been studied are: (i) metals as hydrogen-bond acceptors ($M\cdots HD$; see Brammer *et al.*, 1995; Braga, Grepioni, Tedesco, Biradha & Desiraju, 1997); (ii) metal and other hydrides as hydrogen-bond acceptors (so-called dihydrogen bonds; $MH\cdots HD$, $BH\cdots HN$ *etc.*; see Braga *et al.*, 1998; Klooster *et al.*, 1999); (iii) metal-bound ligands as hydrogen-bond acceptors ($M-X\cdots HD$, X = halide, CO, NCS *etc.*; Aullon *et al.*, 1998; Brammer *et al.*, 1999, 2001; Tchertanov & Pascard, 1997; Dadon & Bernstein, 1997); (iv) Organometallic species as hydrogen-bond donors ($M-CH\cdots A$ *etc.*; Braga *et al.*, 1995); (v) 'organic' hydrogen bonds, in which the functional groups ($-CO_2H$ *etc.*) are familiar from organic chemistry and are present as substituents on ligands attached to organometallic or coordination complexes of inorganic elements (Biradha *et al.*, 1996; Braga, Grepioni, Walther *et al.* 1997; Jones & Ahrens, 1998).

7.2. Secondary bonding

The seminal work of Alcock (1972) indicated that what he termed secondary bonding, namely the hypervalent intermolecular interactions of the heavier *p*-block elements, was the functional equivalent of hydrogen bonding in the chemistry of those elements. This idea has proved remarkably prescient and has been explored in both database and experimental studies seeking to establish its validity. These interactions have much in common with hydrogen bonds, both conceptually, *i.e.* as three-centre four-electron systems (Landrum & Hoffmann, 1998; Starbuck *et al.*, 1999), and practically, in their ability to form networks with a high probability of occurrence.

7.3. Aurophilic interactions

The work of Schmidbaur (2001) and others has emphasized the attractive interactions that exist between gold(I) centres in molecular chemistry, and Pyykko (1997) and others have provided an understanding of the origins of this effect. Desiraju has explored the influence of such interactions in crystal structures of gold(I) species (Pathaneni & Desiraju, 1993), while analogous $Ag^I\cdots Ag^I$ and $Cu^I\cdots Cu^I$ (Liu *et al.*, 1998) interactions have been noted.

7.4. Interligand interactions

The importance of dipole–dipole interactions in both organic and inorganic carbonyl species has been noted (Allen *et al.*, 1998). Dance and co-workers have explored the intermolecular chemistry of a range of polyphenylated systems, usually involving metal complexes (salts of PPh_4^+ and PPN^+, complexes of 2,2′-bipy, *o*-phen and PPh_3 *etc.*; Dance & Scudder, 1996, 2000*a*,*b*, 2001; Lewis & Dance, 2000*a*,*b*; Russell *et al.*, 2001; Scudder & Dance, 1998*a*,*b*, 2000; Steiner, 2000). Crystal structures of these species show remarkably persistent motifs in which attractive (by computation) edge–face interactions between sets of phenyls are present. These interactions might be termed hydrogen bonds by some workers, but irrespective of terminology, they seem influential in determining

the molecular packing arrangement adopted in the crystal structure.

7.5. Ligand–*s*-block metals

There has been considerable interest in the 'intermolecular' interactions of biologically important metals, largely from the *s* block (Ca, Mg, but also Zn and the Group 1 metals; see Bock *et al.*, 1995; Katz *et al.*, 1996; Pidcock & Moore, 2001). These interactions are arguably within the *s*-block metal's coordination sphere and are therefore properly called intramolecular in character. However, for both historical and practical reasons (the lability of these bonds in solution for example) these interactions have been termed intermolecular and are classified as such.

7.6. Agostic interactions

In contrast to the prevalence of the electron-rich (three-centre four-electron) intermolecular hydrogen bond, there is relatively little information available on intermolecular electron-deficient (three-centre two-electron) agostic bonding ($M\cdots$HC *etc.*). Nevertheless, studies have been conducted to establish patterns of geometry in the data that are available (Braga *et al.*, 1996).

7.7. New approaches

While many studies have focused on individual types of interaction, the Cambridge Crystallographic Data Centre produces an automatically compiled knowledge base, ISOSTAR (Bruno *et al.*, 1997), in which interactions between (organic) functional groups drawn from both the CSD and the Protein Databank (PDB; Berman, 2002) are recorded, as well as some computational data. The information derived is presented in a number of ways, including 'propensity' plots in which the number of intermolecular contacts per voxel is compared with that to be expected in a random interaction geometry. These plots together with the geometric information may be generated for more orthodox collections of structural data harvested from the CSD.

8. Space groups and molecular packing

Prediction of the way in which molecules pack in crystal structures is a key unsolved problem in modern chemical crystallography. Understanding the packing of organometallic and inorganic molecular complexes is no less troublesome than the same problem in organic solid-state chemistry, and indeed may be more complex because of the greater softness of organometallics and the problems of adequately modelling both the intramolecular and intermolecular aspects of geometry. While computational approaches to the prediction of crystal structures (see *e.g.* Schmidt & Englert, 1996) and the study of alternative packings (Braga *et al.*, 1994) have met with some success, more energy has been focused on the rationalization and analysis of observed packings. Mingos (see Mingos & Rohl, 1991; Rohl & Mingos, 1993) and Braga and co-workers (see Braga *et al.*, 1993) have studied the way in which the overall shape and size of ions in complex salts is related to their packing motifs.

9. Conclusions

The diversity of applications noted in this review indicates the growing maturity of the CSD in respect of molecular inorganic chemistry. However, it is clear that much more will be possible in due course as the volume of data on the inorganic elements grows. In crude terms, there are complexes and compounds of approaching 100 'inorganic' elements to be studied on the basis of less than 140000 structures in the present CSD; this contrasts with >110000 structures in the CSD on which to formulate studies of the structural chemistry of some 15 'organic' elements (*i.e.* those non-metals above the Zintl line in the *p*-block which form substantial numbers of compounds). The great variety of molecular chemistry, both intramolecular and intermolecular, offers endless challenges to the structural chemist in mining the CSD for knowledge that might afford deeper insights into the geometric behaviour of these species, and in testing the best computational methodologies that might be employed to help convert this knowledge into understanding. Thus, our objective must be to develop tools for the analysis of rapidly increasing quantities of structural information, and to move towards a synthesis of the new understanding in a way that facilitates a new and knowledge-based approach to chemistry.

I thank the many chemists and crystallographers who synthesized the compounds and determined the structures discussed in this paper, and Helena Dumycz for assistance in its preparation.

References

Alcock, N. W. (1972). *Adv. Inorg. Chem. Radiochem.* **15**, 1–58.
Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 1, 31–37.
Allen, F. H., Baalham, C. A., Lommerse, J. P. M. & Raithby, P. R. (1998). *Acta Cryst.* B**54**, 320–329.
Allen, F. H., Harris, S. E. & Taylor, R. (1996). *J. Comput.-Aided Mol. Des.* **10**, 247–254.
Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1987). *J. Chem. Soc. Perkin Trans. 2*, pp. S1–S19.
Altona, C. & Sundaralingam, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.
Alvarez, S. & Aullon, G. (1999). *Met. Clusters Chem.* **1**, 308–322.
Alvarez, S. & Llunell, M. (2000). *J. Chem. Soc. Dalton Trans.* pp. 3288–3303.
Alvarez, S., Pinsky, M. & Avnir, D. (2001). *Eur. J. Inorg. Chem.* pp. 1499–1503.
Alvarez, S., Pinsky, M., Llunell, M. & Avnir, D. (2001). *Cryst. Eng.* **4**, 179–200.
Anderson, K. M. & Orpen, A. G. (2001). *Chem. Commun.* pp. 2682–2683.
Auf der Heyde, T. P. E. (1994*a*). *Angew. Chem.* **106**, 871–888.
Auf der Heyde, T. P. E. (1994*b*). *Structure Correlation*, edited by H.-B. Bürgi & J. D. Dunitz, pp. 337–368. Weinheim: Verlag Chemie.
Aullon, G. & Alvarez, S. (1993). *Inorg. Chem.* **32**, 3712–3719.
Aullon, G. & Alvarez, S. (1996). *Inorg. Chem.* **35**, 3137–3144.

Aullon, G. & Alvarez, S. (1997). *J. Chem. Soc. Dalton Trans.* pp. 2681–2687.

Aullon, G., Bellamy, D., Orpen, G. A., Brammer, L. & Bruton, E. A. (1998). *Chem. Commun.* pp. 653–654.

Aullon, G., Lledos, A. & Alvarez, S. (2000). *Inorg. Chem.* **39**, 906–916.

Barker, J. J. & Orpen, A. G. (1999). *Acta Cryst.* B**55**, 203–208.

Bellamy, D., Brown, N. C., Connelly, N. G. & Orpen, A. G. (1999). *J. Chem. Soc. Dalton Trans.* pp. 3191–3195.

Bellamy, D., Connelly, N. G., Hicks, O. M. & Orpen, A. G. (1999). *J. Chem. Soc. Dalton Trans.* pp. 3185–3190.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Ravichanran, V., Schneider, B., Thanki, N., Padilla, D., Weissig, H., Westbrook, J. D. & Zardecki, C. (2002). *Acta Cryst.* B**58**, 899–907.

Beyreuther, S., Hunger, J., Huttner, G., Mann, S. & Zsolnai, L. (1996). *Chem. Ber.* **129**, 745.

Biradha, K., Desiraju, G. R., Braga, D. & Grepioni, F. (1996). *Organometallics*, **15**, 1284–1295.

Bock, C. W., Katz, A. K. & Glusker, J. P. (1995). *J. Am. Chem. Soc.* **117**, 3754–3765.

Braga, D., Grepioni, F., Biradha, K. & Desiraju, G. R. (1996). *J. Chem. Soc. Dalton Trans.* pp. 3925–3930.

Braga, D., Grepioni, F., Biradha, K., Pedireddi, V. R. & Desiraju, G. R. (1995). *J. Am. Chem. Soc.* **117**, 3156–3166.

Braga, D., Grepioni, F., Milne, P. & Parisini, E. (1993). *J. Am. Chem. Soc.* **115**, 5115–5122.

Braga, D., Grepioni, F. & Orpen, A. G. (1994). *Organometallics*, **13**, 3544–3556.

Braga, D., Grepioni, F., Tedesco, E., Biradha, K. & Desiraju, G. R. (1997). *Organometallics*, **16**, 1846–1856.

Braga, D., Grepioni, F., Walther, D., Heubach, K., Schmidt, A., Imhof, W., Goerls, H. & Klettke, T. (1997). *Organometallics*, **16**, 4910–4919.

Braga, D., Leonardis, P. D., Grepioni, F., Tedesco, E. & Calhorda, M. J. (1998). *Inorg. Chem.* **37**, 3337–3348.

Brammer, L., Bruton, E. A. & Sherwood, P. (1999). *New J. Chem.* **23**, 965–968.

Brammer, L., Bruton, E. A. & Sherwood, P. (2001). *Cryst. Growth Des.* **1**, 277–290.

Brammer, L., Rivas, J. C. M. & Spilling, C. D. (2000). *J. Organomet. Chem.* **609**, 36–43.

Brammer, L., Zhao, D., Ladipo, F. T. & Braddock-Wilking, J. (1995). *Acta Cryst.* B**51**, 632–640.

Brunner, H., Winter, A. & Breu, J. (1998). *J. Organomet. Chem.* **553**, 285–306.

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., Pearson, J. & Taylor, R. (2002). *Acta Cryst.* B**58**, 389–397.

Bruno, I. J., Cole, J. C., Lommerse, J. P. M., Rowland, R. S., Taylor, R. & Verdonk, M. L. (1997). *J. Comput.-Aided Mol. Des.* **11**, 525–537.

Bugarcic, Z., Lovqvist, K. & Oskarsson, A. (1993). *Acta Chem. Scand.* **47**, 554–559.

Bürgi, H.-B. (1998). *Acta Cryst.* A**54**, 873–885.

Bürgi, H.-B. & Dunitz, J. D. (1983). *Acc. Chem. Res.* **16**, 153–161.

Burton, V. J. & Deeth, R. J. (1995). *J. Chem. Soc. Chem. Commun.* pp. 573–574.

Burton, V. J., Deeth, R. J., Kemp, C. M. & Gilbert, P. J. (1995). *J. Am. Chem. Soc.* **117**, 8407–8415.

Cano, J., Rodriguez-Fortea, A., Alemany, P., Alvarez, S. & Ruiz, E. (2000). *Chem. Eur. J.* **6**, 327–333.

Carugo, O. (1994). *Inorg. Chim. Acta*, **215**, 219–223.

Coe, B. J. & Glenwright, S. J. (2000). *Coord. Chem. Rev.* **203**, 5–80.

Costello, J. F. & Davies, S. G. (1998). *J. Chem. Soc. Perkin Trans. 2*, pp. 1683–1689.

Costello, J. F., Davies, S. G. & McNally, D. (1999). *J. Chem. Soc. Perkin Trans. 2*, pp. 465–474.

Cotton, F. A. & Yokochi, A. (1997). *Inorg. Chem.* **36**, 2461–2462.

Crabtree, R. H., Holt, E. M., Lavin, M. & Morehouse, S. M. (1985). *Inorg. Chem.* **24**, 1986–1992.

Crabtree, R. H. & Lavin, M. (1986). *Inorg. Chem.* **25**, 805–812.

Cremer, D. & Pople, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.

Crispini, A., Harrison, K. N., Orpen, A. G., Pringle, P. G. & Wheatcroft, J. (1996). *J. Chem. Soc. Dalton Trans.* pp. 1069–1076.

Cullen, D. L., Desai, L. V., Shelnutt, J. A. & Zimmer, M. (2001). *Struct. Chem.* **12**, 127–136.

Cundari, T. R., Deng, J. & Fu, W. (2000). *Int. J. Quantum Chem.* **77**, 421–432.

Cundari, T. R. & Russo, M. (2001). *J. Chem. Inf. Comput. Sci.* **41**, 281–287.

Dadon, H. & Bernstein, J. (1997). *Inorg. Chem.* **36**, 2898–2900.

Dance, I. & Scudder, M. (1996). *Chem. Eur. J.* **2**, 481–486.

Dance, I. & Scudder, M. (2000*a*). *J. Chem. Soc. Dalton Trans.* pp. 1587–1594.

Dance, I. & Scudder, M. (2000*b*). *J. Chem. Soc. Dalton Trans.* pp. 1579–1585.

Dance, I. & Scudder, M. (2001). *New J. Chem.* **25**, 1500–1509.

Dierkes, P. & van Leeuwen, P. W. N. M. (1999). *J. Chem. Soc. Dalton Trans.* pp. 1519–1530.

Donnelly, M. A. & Zimmer, M. (1999). *Inorg. Chem.* **38**, 1650–1658.

Dunitz, J. D. (1979). *X-ray Analysis and the Structure of Organic Molecules.* Ithaca, New York: Cornell University Press.

Fernandez, A. L., Reyes, C., Prock, A. & Giering, W. P. (2000). *J. Chem. Soc. Perkin Trans. 2*, pp. 1033–1041.

Garner, S. E. & Orpen, A. G. (1993). *J. Chem. Soc. Dalton Trans.* pp. 533–541.

Halgren, T. A. (1996). *J. Comput. Chem.* **17**, 616–641.

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

Harding, M. M. (2000). *Acta Cryst.* D**56**, 857–867.

Harding, M. M. (2001). *Acta Cryst.* D**57**, 401–411.

Harris, S. E., Pascual, I. & Orpen, A. G. (2001). *J. Chem. Soc. Dalton Trans.* pp. 2996–3009.

Hunger, J., Beyreuther, S., Huttner, G., Allinger, K., Radelof, U. & Zsolnai, L. (1998). *Eur. J. Inorg. Chem.* pp. 693–702.

Jensen, W. P., Palenik, G. J. & Tiekink, E. R. T. (2001). *Polyhedron*, **20**, 2137–2143.

Johnson, B. F. G. (1997). *J. Chem. Soc. Dalton Trans.* pp. 1473–1479.

Jones, P. G. & Ahrens, B. (1998). *Chem. Commun.* pp. 2307–2308.

Kadyrov, R., Boerner, A. & Selke, R. (1999). *Eur. J. Inorg. Chem.* pp. 705–711.

Kahr, B., Mitchell, C. A., Chance, J. M., Clark, R. V., Gantzel, P., Baldridge, K. K. & Siegel, J. S. (1995). *J. Am. Chem. Soc.* **117**, 4479–4482.

Katz, A. K., Glusker, J. P., Beebe, S. A. & Bock, C. W. (1996). *J. Am. Chem. Soc.* **118**, 5752–5763.

Keinan, S. & Avnir, D. (2001*a*). *Inorg. Chem.* **40**, 318–323.

Keinan, S. & Avnir, D. (2001*b*). *J. Chem. Soc. Dalton Trans.* pp. 941–947.

Kitaigorodskii, A. I. (1970). *Adv. Struct. Res. Diffr. Methods*, **3**, 173.

Klooster, W. T., Koetzle, T. F., Siegbahn, P. E. M., Richardson, T. B. & Crabtree, R. H. (1999). *J. Am. Chem. Soc.* **121**, 6337–6343.

Landrum, G. A. & Hoffmann, R. (1998). *Angew. Chem. Int. Ed. Engl.* **37**, 1887–1890.

Leuwerink, F. T. H., Harkema, S., Briels, W. J. & Feil, D. (1993). *J. Comput. Chem.* **14**, 899–906.

Lewis, G. R. & Dance, I. (2000*a*). *Inorg. Chim. Acta*, **306**, 160–167.

Lewis, G. R. & Dance, I. (2000*b*). *J. Chem. Soc. Dalton Trans.* pp. 299–306.

Liu, X.-Y. & Alvarez, S. (1997). *Inorg. Chem.* **36**, 1055–1060.

Liu, X.-Y., Mota, F., Alemany, P., Novoa, J. J. & Alvarez, S. (1998). *Chem. Commun.* pp. 1149–1150.

Longuet-Higgins, H. C. (1963). *Mol. Phys.* **6**, 445–460.

Mann, B. E. (1997). *J. Chem. Soc. Dalton Trans.* pp. 1457–1471.

Martin, A. & Orpen, A. G. (1996). *J. Am. Chem. Soc.* **118**, 1464–1470.

Mingos, D. M. P. & Rohl, A. L. (1991). *J. Chem. Soc. Dalton Trans.* pp. 3419–3425.

Mislow, K. (1966). *Introduction to Stereochemistry*. New York: W. A. Benjamin, Inc.
Morton, D. A. V. & Orpen, A. G. (1992). *J. Chem. Soc. Dalton Trans.* pp. 641–653.
Mota, F., Novoa, J. J., Losada, J., Alvarez, S., Hoffmann, R. & Silvestre, J. (1993). *J. Am. Chem. Soc.* **115**, 6216–6229.
Mueller, T. E. & Mingos, D. M. P. (1995). *Transit. Met. Chem. (London)*, **20**, 533–539.
Orpen, A. G. (1998). *Acta Cryst.* D**54**, 1194–1198.
Orpen, A. G., Brammer, L., Allen, F. H., Kennard, O., Watson, D. G. & Taylor, R. (1989). *J. Chem. Soc. Dalton Trans.* pp. S1–S83.
Orpen, A. G. & Quayle, M. J. (2001). *J. Chem. Soc. Dalton Trans.* pp. 1601–1610.
Palacios, A. A., Aullon, G., Alemany, P. & Alvarez, S. (2000). *Inorg. Chem.* **39**, 3166–3175.
Palenik, G. J. (1997). *Inorg. Chem.* **36**, 4888–4890.
Pathaneni, S. S. & Desiraju, G. R. (1993). *J. Chem. Soc. Dalton Trans.* pp. 319–322.
Pidcock, E. & Moore, G. R. (2001). *J. Biol. Inorg. Chem.* **6**, 479–489.
Pyykko, P. (1997). *Chem Rev.* **97**, 597.
Raithby, P. R., Shields, G. P. & Allen, F. H. (1997*a*). *Acta Cryst.* B**53**, 476–489.
Raithby, P. R., Shields, G. P. & Allen, F. H. (1997*b*). *Acta Cryst.* B**53**, 241–251.
Raithby, P. R., Shields, G. P., Allen, F. H. & Motherwell, W. D. S. (2000). *Acta Cryst.* B**56**, 444–454.
Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. III & Skiff, W. M. (1992). *J. Am. Chem. Soc.* **114**, 10024–10034.
Rohl, A. L. & Mingos, D. M. P. (1993). *Inorg. Chim. Acta*, **212**, 5–13.
Ruiz, E., Alemany, P., Alvarez, S. & Cano, J. (1997). *Inorg. Chem.* **36**, 3683–3688.
Russell, V., Scudder, M. & Dance, I. (2001). *J. Chem. Soc. Dalton Trans.* pp. 789–799.
Schmidbaur, H. (2001). *Nature (London)*, **413**, 31–33.
Schmidt, M. U. & Englert, U. (1996). *J. Chem. Soc. Dalton Trans.* pp. 2077–2082.
Scudder, M. & Dance, I. (1998*a*). *J. Chem. Soc. Dalton Trans.* pp. 3167–3176.
Scudder, M. & Dance, I. (1998*b*). *J. Chem. Soc. Dalton Trans.* pp. 329–344.
Scudder, M. & Dance, I. (2000). *J. Chem. Soc. Dalton Trans.* pp. 2909–2915.
See, R. F., Kruse, R. A. & Strub, W. M. (1998). *Inorg. Chem.* **37**, 5369–5375.
Shields, G. P., Raithby, P. R., Allen, F. H. & Motherwell, W. D. S. (2000). *Acta Cryst.* B**56**, 455–465.
Smith, J. M. & Coville, N. J. (2001). *Organometallics*, **20**, 1210–1215.
Smith, J. M., Taverner, B. C. & Coville, N. J. (1997). *J. Organomet. Chem.* **530**, 131–140.
Smith, M. B., Orpen, A. G., Pringle, P. G. & Worboys, K. (1998). *J. Organomet. Chem.* **550**, 255–266.
Starbuck, J., Norman, N. C. & Orpen, A. G. (1999). *New J. Chem.* **23**, 969–972.
Steiner, T. (2000). *New J. Chem.* **24**, 137–142.
Taylor, R., Cole, J. C., Kessler, M., Luo, J., Smith, B. R., Harris, S. E. & Orpen, A. G. (2001). *Abstracts of Papers, 222nd American Chemical Society National Meeting*, Chicago, IL, USA, August 26–30, 2001. CINF-011.
Tchertanov, L. & Pascard, C. (1997). *Acta Cryst.* B**53**, 904–915.
Tolman, C. A. (1977). *Chem. Rev.* **77**, 313–348.
Wood, R. M., Abboud, K. A., Palenik, R. C. & Palenik, G. J. (2000). *Inorg. Chem.* **39**, 2065–2068.
Wood, R. M. & Palenik, G. J. (1998). *Inorg. Chem.* **37**, 4149–4151.
Wood, R. M. & Palenik, G. J. (1999). *Inorg. Chem.* **38**, 3926–3930.
Yao, J. W., Copley, R. C. B., Howard, J. A. K., Allen, F. H. & Motherwell, W. D. S. (2001). *Acta Cryst.* B**57**, 251–260.
Zabrodsky, H. & Avnir, D. (1995). *Adv. Mol. Struct. Res.* **1**, 1–31.
Zimmer, M. (2001). *Coord. Chem. Rev.* **212**, 133–163.