

## Metal Complex Geometries in Small-Molecule Crystals

A. GUY ORPEN

School of Chemistry, University of Bristol, Bristol BS8 1TS, England. E-mail: guy.orpen@bristol.ac.uk

(Received 20 May 1998; accepted 1 June 1998)

### Abstract

The origins, scope and utility of compilations of metal–ligand and intraligand bond lengths based on the Cambridge Structural Database are discussed. The limitations on the apparent uncertainty of metal–ligand bond lengths derived from crystallographic data and recent evidence of metal-assisted hydrogen bonding involving ligands are reviewed in the light of the transferability of bond-length values from one crystal structure determination.

### 1. Introduction

In modern chemistry crystallography has played a crucial role in providing increasingly rapid means of identifying molecules and their geometries, usually as a result of single-crystal X-ray diffraction studies. These studies are typically carried out as part of a broader programme of research in which the synthesis of new molecules or assemblies is attempted. The result of the widespread use of single-crystal studies as the final arbiter in the determination of molecular structure is a plethora of geometric data on molecules, their stereochemistries and both intramolecular and intermolecular geometries. Hundreds of thousands of crystal structure determinations in the fields of organic, inorganic and biological chemistry have now been carried out. The (published) outcomes of these studies have been collected in computer databases (Allen & Kennard, 1993; Allen, Bergerhoff *et al.*, 1987) divided by subject area [small-molecule, inorganic (extended-array), macromolecular, and metals and alloys]. The availability of these databases has led to the development of new fields of science in which the purpose is to extract new knowledge from collections of structures rather than explore the import of an individual structure. In this paper a specific sort of study of this type is considered, in which we seek to establish transferable typical dimensions for metal-containing moieties.

The principal collection of information on small-molecule crystal structure analyses is the Cambridge Structural Database (CSD) (Allen & Kennard, 1993). At present (February 1998) this contains data on over 175 000 crystal structures. The content of this database is highly chemically diverse, spanning the fields of organic

chemistry (including amino acids, oligopeptides, steroids, sugars and oligosaccharides); coordination and organometallic complex chemistry (of the *d*- and *f*-block metals – the transition elements), and the chemistry of the *s*- and *p*-block elements (main-group chemistry). The majority of the structures contain fully refined three-dimensional atomic coordinates based on relatively high resolution diffractometer data (typically to better than 0.9 Å). In this respect the structural data differ from those in, for example, the Protein Data Bank (PDB). In the PDB the structures are (or the diffraction data from which they are derived are) primarily of lower resolution and cover a smaller range of chemical species. In many of these structures and in related macromolecular species, metals of the *d* and *f* blocks are present, for example in metalloproteins or in heavy-atom derivatives of proteins.

The construction of refinement models for such macromolecular structures, the use of restraints on their refinement to ensure that they converge to ‘normal’ geometries, or the validation of refined structures against known geometries, will inevitably rely on a store of reliable transferable data on *inter alia* metal–ligand and intraligand atomic distances. The CSD provides such data albeit in a form that requires processing before it can be used. In this paper the transferability of such data and the origin and utility of one major source of processed data (Orpen *et al.*, 1989), or perhaps better ‘knowledge’, about the geometry of metal–ligand interactions is reviewed.

### 2. Results

Fig. 1 shows the growth of entries in the CSD over the period 1950–1995. As may be seen growth is approximately exponential with the number of structures in the database doubling every 7 or 8 years. Of the structures in the CSD, over half are formally inorganic by virtue of containing metallic elements. In the vast majority of these cases the metal is from the *d* or *f* block. The abundance of data on both transition-metal complexes and organic compounds allowed compilations of bond-length tables for organic compounds (Allen, Kennard *et al.*, 1987) and *d*- and *f*-block coordination and organometallic complexes (Orpen *et al.*, 1989). These tables

were subsequently reproduced in Volume *C* of *International Tables for Crystallography* (Wilson, 1992) and (in part) in Appendix A of *Structure Correlation* (Bürgi & Dunitz, 1994). They have attracted wide attention and use in fields ranging from computational chemistry (*e.g.* Pidun & Frenking, 1995) and surface science (*e.g.*, Gierer *et al.*, 1997) to structural biology whether using crystallographic (*e.g.* Djinovic *et al.*, 1992) or X-ray absorption spectroscopic methods (Blackburn *et al.*, 1991).

In the tables of data for metal complexes (Orpen *et al.*, 1989) the metal–ligand contact distances for coordination complexes ( $M-N$ ,  $M-O$ ,  $M-S$  *etc.*) and organometallics ( $M-C$ ) are arranged first by ligand contact atom and subsequently by the type and bonding mode of ligand present. The ligands were assigned numerical classes. For example all oxygen ligands belong to class 5; carboxylate ligands to class 5.5, alcohols to class 5.23, ethers to class 5.24, and so on. Particular ligands or groups of ligands are further specified. So, for example the formate ligand ( $O_2CH$ ) is ligand 5.5.1 and acetates ( $O_2CMe$ ) form ligand class 5.5.2. Alternative bonding modes of a ligand are denoted by a further index. Therefore, terminal  $\sigma$ -bonded acetates form class 5.5.2.1, while chelating acetates are in class 5.5.2.2 and bridging acetates ( $\mu-O_2CMe$ ) in class 5.5.2.3. The ligand coverage spans organometallic chemistry (ligands such as CO,  $\sigma$ -alkyls,  $\eta-C_5H_5$  *etc.*) as well as coordination chemistry [nitrogen, oxygen and sulfur ligands such as amines, pyridines, and imines, hydroxide, nitrate, nitric oxide (NO), carboxylates and thiolates *etc.*] and bioinorganic chemistry (including ligands such as amino-acid derivatives, corrins and porphyrins as well as simpler ligands such as water and sulfate).

Bond-length data for the table were extracted from the September 1985 version of the CSD, which contained 49 854 structures (*cf.* > 175 000 12 years later). Of these, 9802 structures were used, those that satisfied the stringent quality criteria (and contained *d*- or *f*-block metals) noted below.

- (i) Atomic coordinates were available.
- (ii) The structure was determined from diffractometer data.
- (iii) The CSD entry contained no unresolved numerical errors.
- (iv) Either the crystallographic *R* factor was <0.05 or it was <0.07 and the mean e.s.d. of the C–C bond lengths was <0.03 Å.
- (v) Only the most precise determination of a given structure was used.

Once retrieved, bond-length distributions and ligands in a given ligand class for bonds of a given metal were inspected and outliers lying more than four standard deviations from the mean of a unimodal distribution were discarded. Subdivisions of distributions according to chemical criteria were attempted when there were sufficient data, so as to yield unimodal distributions with

small sample standard deviations (typically <0.04 Å). The statistics compiled for those cases having four or more data were as below.

- (i) The unweighted sample mean, *d*.
- (ii) The sample median, *m*.
- (iii) The sample standard deviation, *s*.
- (iv) The lower quartile,  $q_l$ .
- (v) The upper quartile,  $q_u$ .
- (vi) The number of observations in the sample, *n*.

Statistics *d* and *m* provide measures of ‘typical’ or central values in the distributions, while *s* (and  $q_l$  and  $q_u$ ) are measures of the dispersion or spread of values. Finally, *n* is an important indication of the reliability of these statistics. Notably, in cases where *n* is small, the statistics may not be robust or reliable. If insufficient data were available individual bond lengths and the CSD refcodes (and their literature citations) were given.

Where possible the distribution for a given metal–ligand bond length was subdivided according to metal oxidation state and/or coordination number. In a few cases ( $O_2$ , *o*-quinones *etc.*) the oxidation state of the ligand was used to subdivide bond-length distributions. In other cases statistics are given for the distribution excluding a special group of structures (*e.g.* those in which the bond is *trans* to a metal–metal bond).

As a result of these procedures these tables (Orpen *et al.*, 1989) contain bond lengths for 45 metals (Sc–Zn, Y–Cd, La–Lu, Hf–Hg, Th and U), 15 contact atoms (H, B–F, Si–Cl, As–Br, Te and I) and a total of 325 ligand types. Where appropriate intraligand bond lengths are also tabulated. In addition the table provides literature references to over 1650 individual structures, and 164 surveys and reviews of relevant areas of structural chemistry up to 1987. In this respect the table is

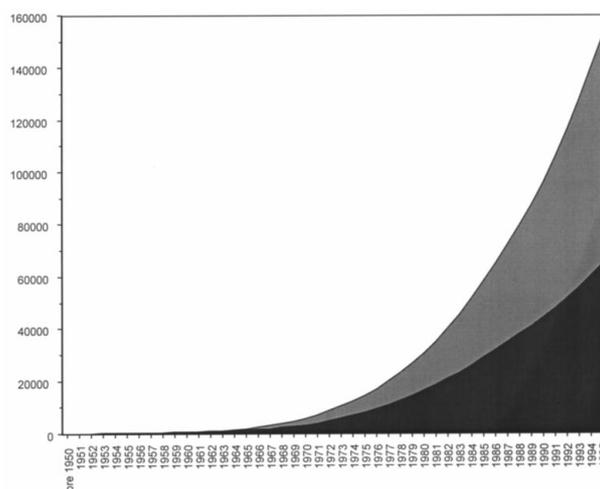


Fig. 1. The number of organic (black) and inorganic (shaded) structures in the CSD over the period 1950–1995.

both a source of bond-length data and entry into the literature.

### 3. Discussion

The tabulation for the *d*- and *f*-block metals provides a large compendium of highly processed data (or knowledge) for use in model building, restraint or validation. In many instances further work is required to explore the utility of the data in the CSD, *e.g.*, as a means of testing theories of structure and bonding whether qualitative or quantitative. We and others have taken this approach. For example, we have sought to test the hypothesis that metal–phosphine bonding has  $\sigma$  and  $p$  components, the latter based in part at least on the involvement of P–C (or P–O in phosphites)  $\sigma^*$  orbitals (Orpen & Connelly, 1990; Dunne *et al.*, 1991; Garner & Orpen, 1993; Crispini *et al.*, 1996).

However, more fundamental questions arise from the attempt to tabulate ‘standard’ or ‘typical’ bond lengths. In particular the transferability of the mean (or median) values to other chemical circumstances is critical to their usefulness. The bond-length distributions themselves show indications of variable reliability in this respect. Thus, for example, the  $\sigma$  value for Fe–PPh<sub>3</sub> bonds (0.038 Å) is substantially larger than that for the P–C bonds in PPh<sub>3</sub> ligands in general (0.011 Å), despite the wide variety of metals bonded to P in this latter distribution. A number of causes may be posited for this phenomenon. Thus, in many cases, substituent effects at the metal remain unresolved (*e.g.* the *trans* or *cis* influence of other ligands, the stereochemistry of the metal). In contrast the substituent pattern at the ligand is often fully defined, as in the case of PPh<sub>3</sub>. Also, the random errors (or precision) in the X-ray diffraction determined locations of lighter (ligand) atoms in the presence of many elements (such as the *4d*, *5d* and *5f* elements) is poorer than in the structures of the lighter metals. Thus, while the overall  $\sigma$  value for C–O bond lengths in terminal carbon monoxide ligands is 0.020 Å (10 022 cases), the corresponding values for individual *3d* metals are in the range 0.011–0.024 Å, while for the *5d* metals they are in the range 0.023–0.035 Å. Finally, one might also expect that metal containing bond lengths would have lower force constants and hence be more susceptible to the ‘packing effects’ often cited by chemical crystallographers confronted by otherwise inexplicable geometrical variations. The discussion below addresses this point.

More generally one might ask, can we expect that a given bond in a given molecule has the same length in a different crystal environment? In 1970 Kitaigorodskii had no doubts on this point, at least for organic molecules. He wrote ‘the crystalline field does not change the bond lengths of organic molecules’ (Kitaigorodskii, 1970). On the basis of our table of bond lengths of organic lengths (Allen, Kennard *et al.*, 1987) there is

little evidence that this assertion does not hold since many distributions show  $\sigma$  values of 0.01 Å or less, even before requiring that the molecular environment of a bond be identical.

The situation as regards metal–ligand bonds is less clear, although in general given their high atomic numbers the metals’ locations (and hence their bond lengths) are rather precisely determined by X-ray diffraction measurements (Martín & Orpen, 1996).

In general the variance ( $\sigma^2$ ) of a bond length between two elements (*e.g.* Pt–Cl) may be described as in equation (1)

$$\sigma^2 = \sigma_c^2 + \sigma_p^2 + \sigma_e^2. \quad (1)$$

Here  $\sigma_c^2$  is the variance due to the variation in molecular environment (*trans* influence; metal oxidation state; coordination number *etc.*),  $\sigma_p^2$  is the variance due to variations in crystal environment (packing effects!) and  $\sigma_e^2$  is the variance due to experimental uncertainties. In general  $\sigma_c^2$  is much the largest contributor to  $\sigma^2$ . By limiting the structures studied to those of relatively high precision  $\sigma_e^2$  may be reduced to a small proportion of  $\sigma^2$ . Furthermore by studying the same bond in the same molecule in different crystal environments  $\sigma_c^2$  can be made exactly zero. Observations of the effect of crystal packing (and other possible sources of inaccuracy in structure determination) may be made by study of the same bond in cases (i)–(vi) below (Martín & Orpen, 1996).

- (i) Polymorphs of the given compound.
- (ii) Structures where the molecular symmetry is higher than the crystallographic site symmetry.
- (iii) Structures in which there is more than one molecule in the asymmetric unit.
- (iv) So-called pseudo-polymorphs in which a molecular species forms multiple solvates and/or co-crystals.
- (v) Salts of complex ions.
- (vi) Combinations of (i)–(v).

Furthermore, in principle, comparison between crystallographically determined geometries and those in the gas phase whether experimentally or computationally determined would provide estimates of  $\sigma_p$ . As noted by Bernstein (Bernstein, 1992), Kitaigorodskii suggested a number of these approaches.

In published work using methods (i)–(iv) we (Martín & Orpen, 1996) and others (Cotton & Yokochi, 1997) have obtained  $\sigma$  values (and  $\sigma_p$  estimates) for metal complexes. Our work (based on data collected from the CSD) (Martín & Orpen, 1996) and that based on in-house data (Cotton & Yokochi, 1997) showed that metal–ligand bond lengths show  $\sigma$  (and  $\sigma_p$ ) values of the order of 0.01–0.02 Å across a broad range of ligand types (*M–X* bonds, where *X* = Cl, O, N, S, P, C *etc.*). Furthermore, *L–M–L* angles show  $\sigma$  of the order 1–2° for these bonds, much larger than the  $\sigma_e$  values reported

on the basis of conventional crystallographic least squares.

More recently by method (vi), we have shown that perchlorometallate salts show similar  $\sigma$  values for a wide range of *p*- and *d*-block metals for which  $\sigma_e$  values are very much lower (Orpen & Quayle, 1998). In some cases (*e.g.*  $[\text{BiCl}_6]^{3-}$ ,  $[\text{TeCl}_6]^{2-}$ ,  $[\text{HgCl}_4]^{2-}$ ) the  $\sigma$  values are very much larger (0.04–0.10 Å) in accordance with the well known softness of the potential-energy surface of these ions (Carmalt *et al.*, 1995). This is in qualitative agreement with the expectation that bonds with lower

force constants would show larger  $\sigma$  values. It has been noted that this expectation is not borne out quantitatively (Martín & Orpen, 1996; Cotton & Yokochi, 1997).

The fact that  $\sigma_p$  values are in general much larger than  $\sigma_e$  has implications for how much reliance can be placed on an individual bond length from a single structure determination. As we have noted (Martín & Orpen, 1996) it seems likely that packing effects (*i.e.* intermolecular interactions) are largely responsible for the  $\sigma_p$  values observed, although systematic errors in structure analysis may also be responsible

One significant intermolecular interaction that may lead to the variations in bond lengths seen is, of course, the hydrogen bond. In recent work (Aullón *et al.*, 1998) we have shown that metal chloride complexes can act as effective hydrogen-bond acceptors through  $M-\text{Cl}\cdots\text{HO}$  or  $M-\text{Cl}\cdots\text{HN}$  interactions. The bonds formed are of similar length (modal value *ca* 2.3–2.5 Å) to those formed by free chloride ions, and considerably shorter than the shortest  $\text{C}-\text{Cl}\cdots\text{HO}$  or  $\text{C}-\text{Cl}\cdots\text{HN}$  interactions (see Fig. 2). The stereochemistry of both  $M-\text{Cl}\cdots\text{HO/N}$  and  $\text{C}-\text{Cl}\cdots\text{HO/N}$  bonds is pronounced. In both cases, while near linear geometries are evidently favourable at hydrogen, the most likely angle at chlorine is markedly bent (typically 90–120°, see Fig. 3). The implications of the strength and pronounced angle preference in  $M\text{Cl}\cdots\text{HX}$  interactions implies that

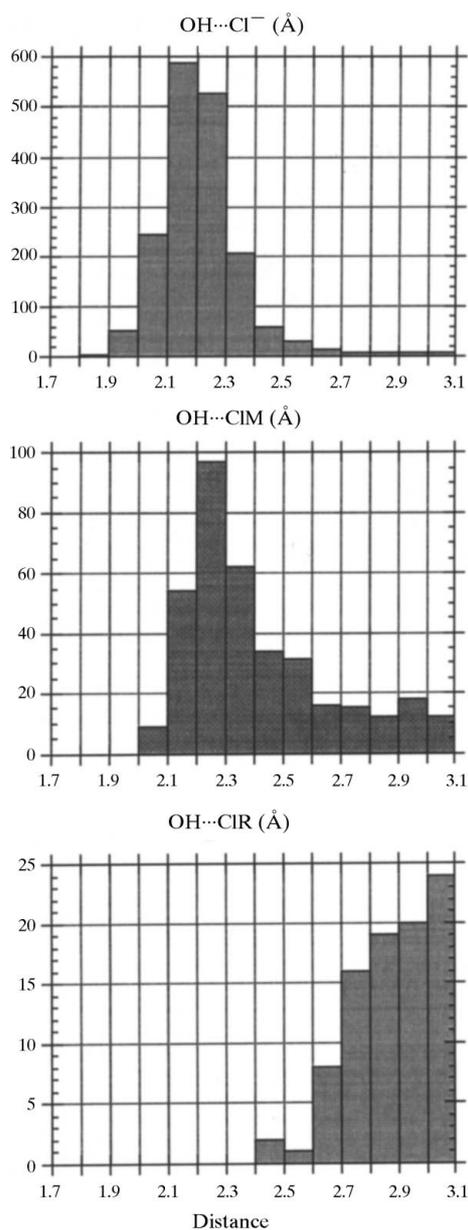


Fig. 2. Histograms of intermolecular  $\text{Cl}\cdots\text{HO}$ ,  $M\text{Cl}\cdots\text{HO}$  and  $\text{CCl}\cdots\text{HO}$  distances (Å) less than 3.1 Å in the CSD.

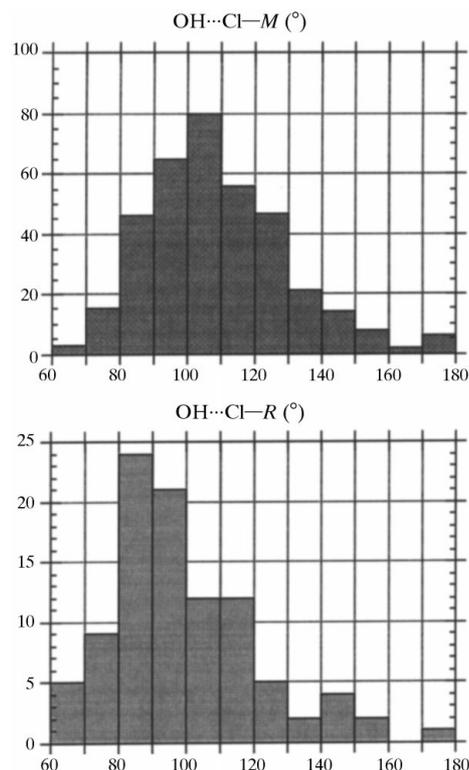


Fig. 3. Histograms of intermolecular  $M-\text{Cl}\cdots\text{HO}$  and  $\text{C}-\text{Cl}\cdots\text{HO}$  angles (°) in the CSD.

metal-assisted hydrogen bonding of this sort may have useful applications and consequences in the supramolecular chemistry of metal complexes of chloride and related ligands. Complementary work based on the CSD on the hydrogen bonding of organic carboxylate groups (not bound to metals) has been reported (Görbitz & Etter, 1992), in which the biological chemistry of such interactions was emphasised.

#### 4. Conclusions

The availability of substantial and highly validated tabulations of intramolecular geometries (bond lengths) is of great value to structural scientists in many fields, including biology. The compilations discussed here provide good coverage of large areas of the periodic table but need to be used appropriately in the light of caveats expressed here about the robustness of crystallographically determined bond lengths and angles, at least for metal complexes.

There is a clear and present need for computerized (searchable and accessible) compilations of this sort. More complete coverage of the periodic table, to cover *s* and heavy *p*-block elements for example, is desirable. In addition, the scope of these compilations needs broadening to incorporate non-bonded (intramolecular) distances, bond angles and conformational descriptors. The recent publication of the *ISOSTAR* knowledge base as part of the CSD system (Bruno *et al.*, 1997) has led the way in providing a validated knowledge base for intermolecular interactions, of the sort now needed for intramolecular data.

The contributions of Professor L. Brammer, Dr A. Martín and Drs F. H. Allen and R. Taylor to this work are gratefully acknowledged. The many synthetic chemists and crystallographers who prepared and characterized the structures in the CSD are due many thanks.

#### References

- Allen, F. H., Bergerhoff, G. & Sievers, R. (1987). Editors. *Crystallographic Databases*. Bonn: IUCr.
- Allen, F. H. & Kennard, O. (1993). *Chem. Design Autom. News*, **8**, 31–37.
- Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1987). *J. Chem. Soc. Perkin Trans. II*, S1–S19.
- Aullón, G., Bellamy, D., Brammer, L., Bruton, E. A. & Orpen, A. G. (1998). *Chem. Commun.* pp. 653–654.
- Bernstein, J. (1992). In *Accurate Molecular Structures: Their Determination and Importance*, edited by A. Domenicano & I. Hargittai, pp. 469–497. Oxford University Press.
- Blackburn, N. J., Hasnain, S. S., Pettingill, T. M. & Strange, R. W. (1991). *J. Biol. Chem.* **266**, 23120–23127.
- Bruno, I. J., Cole, J. C., Lommerse, J. P. M., Rowland, R. S., Taylor R. & Verdonk, M. L. (1997). *J. Comput. Mol. Design*, **11**, 525–537.
- Bürgi, H.-B. & Dunitz, J. D. (1994). Editors. *Structure Correlation*, Appendix A, Vol. 2. Weinheim: Verlag Chemie.
- Carmalt, C. J., Norman, N. C. & Farrugia, L. J. (1995). *Polyhedron*, **14**, 1405–1413.
- Cotton, F. A. & Yokochi, A. (1997). *Inorg. Chem.* **36**, 2461–2462.
- Crispini, A., Harrison, K. N., Orpen, A. G., Pringle, P. G. & Wheatcroft, J. (1996). *J. Chem. Soc. Dalton Trans.* pp. 1069–1076.
- Djinovic, K., Gatti, G., Coda, A., Antolini, L., Pelosi, G., Desideri, A., Falconi, M., Marmocchi, F., Rotilio, G. & Bolognesi, M. (1992). *J. Mol. Biol.* **225**, 791–809.
- Dunne, B. J., Morris, R. B. & Orpen, A. G. (1991). *J. Chem. Soc. Dalton Trans.* pp. 653–662.
- Garner, S. E. & Orpen, A. G. (1993). *J. Chem. Soc. Dalton Trans.* pp. 533–541.
- Gierer, M., Barbieri, A., Vanhove, M. A. & Somorjai, G. A. (1997). *Surf. Sci.* **391**, 176–182.
- Görbitz, C. H. & Etter, M. C. (1992). *J. Am. Chem. Soc.* **114**, 627–631.
- Kitaigorodskii, A. I. (1970). *Advances in Structure Research by Diffraction Methods*, Vol. 3, edited by R. Brill & R. Mason, pp. 173–247. Oxford: Pergamon Press.
- Martín, A. & Orpen, A. G. (1996). *J. Am. Chem. Soc.* **118**, 1464–1470.
- 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. & Connelly, N. G. (1990). *Organometallics*, **9**, 1206–1210.
- Orpen, A. G. & Quayle, M. J. (1998). Unpublished results.
- Pidun, U. & Frenking, G. (1995). *Organometallics*, **14**, 5325–5336.
- Wilson, A. J. C. (1992). Editor. *International Tables for Crystallography*, Vol. C, ch. 9. Dordrecht: Kluwer Academic Publishers.