# Crystal Structure Determination: A Critical View

By P. G. Jones
INSTITUT FÜR ANORGANISCHE CHEMIE DER UNIVERSITÄT GÖTTINGEN,
TAMMANNSTR. 4, D 3400 GÖTTINGEN, W. GERMANY

## 1 Introduction

Structural information provided by $X$-ray crystallography has been a cornerstone of chemical research in the last few decades. However, the very importance and widespread use of the technique has created a problem for the chemist, who may not be a specialist crystallographer; how should he critically evaluate published crystal structures?

There is an understandable tendency nowadays for crystal structures to be believed implicitly. As an act of faith this is touching, but, as any honest crystallographer would admit, many crystallographic papers contain infelicities ranging from tendentious statements through over-optimism to downright errors. This article is an attempt to provide the non-crystallographer with a guide to recognizing some limitations of published crystal structures, without involving him in detailed theory.

The following basic $X$-ray experiment is assumed: the *intensities* $I$ and associated standard deviations $\sigma(I)$ are measured with a diffractometer and mathematically reduced to *structure factors* $|F|$ [and associated $\sigma(F)$]; these are then used for least-squares refinement of a structural model of the *electron density* $\rho$, leading to refined values of the *atomic co-ordinates* and *thermal vibration parameters* (also known as *temperature factors*).

## 2 R Values: All Things to All Men?

Since crystal structure determinations are essentially quantitative, the question of criteria of accuracy arises. The two usual criteria are (a) $R$ values and (b) estimated standard deviations (e.s.d.'s) of derived parameters such as bond lengths. (These are normally expressed in the form 1.555(6) Å, meaning 1.555 Å with an e.s.d. of 0.006 Å).

$R$ values are an attempt to express the agreement of observed structure factors $F_{o}$ with $F_{c}$, those calculated for the refined model. The conventional $R$ value is given by

$$R = \sum |F_{o} - F_{c}| / \sum F_{o}$$

and it is thus clear that the lower the value of $R$, the better the structure—or is it? Unfortunately, there are several crystallographers' tricks for artificially reducing $R$ values. To see through these, other factors must be considered; weak reflections, weighting schemes, and data/parameter ratios ('data' here means the number of independent measured intensities—see below).

All crystals will show some weak reflections (especially at high diffraction angle $2\theta$), characterized by a low ratio of $I$ to $\sigma(I)$. It is general practice to omit from calculations those reflections with $I/\sigma(I)$ less than a certain threshold value, usually in the range 1.5 to 3 (corresponding to $F/\sigma(F)$ 3 to 6, since $I$ is proportional to $F^2$). This has three advantages: first, the computing time necessary for refinement is reduced; secondly, the $R$ value is reduced by elimination of those reflections associated with high $\sigma$; and thirdly, the problem of how to treat reflections with negative $I$ is avoided. The first is often necessary, since few researchers have access to unlimited computer time, but the second is purely cosmetic. Since $R$ can be reduced to an almost arbitrarily low value by use of a suitably high $\sigma$ threshold, more valid criteria of accuracy must be sought. One such is the 'weighted $R$ value', generally given the symbol $R'$ or $R_w$.

Weighted $R$ values represent an attempt to take the errors $\sigma(F)$ into account during refinement; reflections with higher $\sigma$ are given less weight. There are several different types of weighting scheme, but one of the commonest is

$$w^{-1} = \sigma^2(F) + gF^2$$

where $g$ is a small constant (typically 0 to 0.001); this takes into account that many common *systematic* errors depend approximately linearly on $I$. The weighted $R$ is then expressed as

$$R_w = \frac{\sum w^{1/2} |F_o - F_c|}{\sum w^{1/2} F_o}$$

Since $R_w$ values are based on the incorporation of more information ($\sigma$ values) into refinement, and are far less dependent on the $\sigma$ threshold, they may be regarded as more 'realistic', and $R_w$ is thus a better criterion of accuracy than $R$. Sensible weighting schemes generally lead to lower e.s.d.'s (see below) for refined atom co-ordinates, and it is thus rare to find a structure which has been refined without weights (*i.e.* with 'unit weights'); such structures should be viewed with some suspicion.

The reference to e.s.d.'s in the last paragraph reminds us that the structural information of an $X$-ray investigation is contained in refined parameters and their e.s.d.'s, rather than in $R$ values. There is thus a considerable argument in favour of judging the accuracy of a crystal structure in terms of the e.s.d.'s. The omission of too many reflections would tend to increase the e.s.d.'s, since they are to a good approximation inversely proportional to the square root of the number of reflections. A related effect is that higher e.s.d.'s are inevitable if too few reflections are measured and the data/parameter ratio is thus too low—perhaps to save diffractometer time, or to reduce $R$ by omitting higher angle (weaker) reflections. Measurements should normally be made to a diffraction angle $2\theta$ of at least $45°$ for Mo $K\alpha$ radiation, or $100°$ for Cu $K\alpha$. Structures with $2\theta_{max}$ lower than this, or where no $2\theta_{max}$ is quoted, may justifiably be regarded with suspicion.* The

---

*They are also far more difficult to solve; much of the effort expended in developing extremely powerful programs and using them to solve such structures could be avoided if proper data were collected in the first place!

data/parameter ratio should be at least 8 if at all possible (and may be *ca.* 20 for well-determined structures), yet some structures with a ratio of 5 still appear in the literature.

The close relationships between weighting schemes, e.s.d.'s, and weak reflection thresholds have been the subject of several articles in the crystallographic literature. It has been convincingly argued that *no* weak reflections should be omitted, as this may lead to systematic bias, albeit small, in refined parameters;[1] several test structures have shown conclusively that e.s.d.'s are increased by omission of weak reflections, and indeed that reasonable refinements are achieved by using only reflections *below* a weakness threshold.[2] However, the effect on e.s.d.'s has been shown to be small for a range of conventional thresholds,[3] and it may be assumed that the practice of omitting weak reflections will continue.

*Moral*: (i) judge accuracy by e.s.d.'s rather than $R$ values (but see next Section for a *caveat*); (ii) high $\sigma$ thresholds, low $2\theta_{max}$, or low data/parameter ratios may be symptoms of a poorly-determined structure.

### 3 Estimated Standard Deviations: Accuracy *versus* Optimism

An 'accurately determined' organic crystal structure (at room temperature) may claim bond length e.s.d.'s of $0.002 \text{Å}$. This Section will try to explain why the immediate reaction of the reader should be to reach for a pinch of salt.

The mathematics and computational aspects of least-squares refinement are well documented.[4] The e.s.d.'s obtained are *mathematically* sound, but are based solely on information fed to the computer [usually only $F$ and $\sigma(F)$] and, as such, must be lower bounds of a realistic error estimate. Sources of systematic error are ignored; the most serious is usually inaccuracy of cell constants. It is not always realised that the parameters derived from the measured intensities (*via* least-squares refinement) are merely *fractional co-ordinates* of the atoms, based on the unit cell axes. In order to obtain bond lengths (and other molecular dimensions) in absolute units (*e.g.* Å), it is necessary to perform appropriate vector calculations, taking into account the dimensions of the unit cell. The vector **r** between two atoms with fractional co-ordinates $x_1, y_1, z_1, x_2, y_2, z_2$ in a cell defined by the axis vectors **a, b, c** is given by

$$\mathbf{r} = (x_2 - x_1)\mathbf{a} + (y_2 - y_1)\mathbf{b} + (z_2 - z_1)\mathbf{c}$$

and the bond length is then $r = \sqrt{\mathbf{r}.\mathbf{r}}$. Any errors in the cell dimensions will thus give rise to additional errors in the bond lengths (over and above those calculated from least-squares refinement). In the case of heavy atom bond lengths, these extra systematic errors may be appreciably greater than the least-squares e.s.d.'s.

The conventional method of obtaining cell constants using a four-circle diffractometer is to determine accurate positions of several strong reflections in

[1] F. L. Hirshfeld and D. Rabinovich, *Acta Crystallogr.*, 1973, **A29**, 510.
[2] L. Arnberg, S. Hovmöller, and S. Westman, *Acta Crystallogr.*, 1979, **A35**, 497.
[3] R. E. Stenkamp and L. H. Jensen, *Acta Crystallogr.*, 1975, **B31**, 1507.
[4] For example, D. W. J. Cruickshank, 'Least-Squares Refinement of Atomic Parameters', in 'Crystallographic Computing', ed. F. R. Ahmed, Munksgaard, Copenhagen, 1970.

terms of the circle angles[5] $2\theta$, $\omega$, $\chi$, and $\phi$, and to refine the crude cell constants accordingly.[6] As with any refinement, e.s.d.'s may be estimated and would typically lie in the range 0.001 to 0.01 Å for a 10 Å axis (with equivalent fractional errors for other axis lengths). Unfortunately, the problem of systematic errors arises once again. Since axis lengths are inversely proportional to $\sin\theta$ (Bragg's law), a small error in the zero point of the $2\theta$ circle will cause systematic errors in the cell constants; for a typical case (Mo $K\alpha$ radiation, 10 Å axis, $2\theta_0$ 0.05°, $2\theta$ of reflections used *ca.* 20°) the calculated axis length would be 10.025 Å. This additional error is far in excess of the least-squares e.s.d.'s.* Since circle zeros, even on the best-constructed diffractometers, tend to wander, and recalibrating them is tedious, such sources of error are often ignored (until they become so large as to incommode data collection). In our example, this would lead to additional systematic errors of *ca.* 0.004 Å in bond lengths.†

An interesting example of the problems of cell constant measurement is provided by the compound 4-nitrophenyl-α-D-glucopyranoside. The structure was first reported in space group P2$_1$ with $a$ 28.810, $b$ 6.747, $c$ 6.729 Å, $\beta$ 103.68°.[8] A second, independent determination yielded cell constants $a$ 28.045, $b$ 6.767, $c$ 6.719 Å, $\beta$ 90.30°.[9] The apparent major discrepancies can be removed by an allowed axis transformation of the first cell to $a$ 27.993, $b$ 6.747, $c$ 6.729 Å, $\beta$ 90.17°, but the remaining differences, notably *ca.* $10\sigma$ in $b$, are difficult to explain.

*Moral*: (i) quoted cell constant e.s.d.'s may be wildly over-optimistic (as witness the occasional appearance of a fourth or even fifth place of decimals, at which level the thermal expansion over the laboratory temperature range may be significant); (ii) calculated least-squares e.s.d.'s of molecular dimensions should be increased somewhat to allow for errors in cell constants. The latter practice is already implemented in some program systems.

## 4 Incorrect Space Groups: The Case of the Missing Symmetry

It is a sad fact that crystallographers, being human, sometimes fail to assign the correct space group to a structure. In some cases this renders structure solution impossible and the error never reaches the literature; in others, structures are solved, refined, and published in space groups of unnecessarily low symmetry, one or more symmetry elements having gone unrecognized. This may lead to nothing more serious than, say, the refinement of two identical molecules rather than one,

---

*Errors arising from incorrect zeroing of the $\omega$ and $\chi$ circles, in contrast to $2\theta$, *are* reflected in increased least-squares e.s.d.'s.

†It should be pointed out that reliable cell constants *can* be obtained, *e.g.* from accurate $2\theta$ values free from circle zeroing errors.[7] These are, however, not straightforwardly obtainable on all makes of diffractometer.

[5] 'International Tables for *X*-Ray Crystallography', Vol. 4, ed. J. A. Ibers and W. C. Hamilton, Kynoch Press, Birmingham, 1974.

[6] For a brief review of methods, see *Ref.* 7.

[7] W. Clegg, *Fresenius' Z. Anal. Chem.*, 1982, **312**, 22.

[8] P. Swaminathan, *Acta Crystallogr.*, 1982, **B38**, 184.

[9] P. G. Jones, G. M. Sheldrick, A. J. Kirby, and W. B. T. Cruse, *Z. Kristallogr.*, 1982, **161**, 69.

and, despite the waste of computer time, the molecular dimensions will still be reliable. Should the omitted symmetry element be a centre of symmetry, the consequences are more serious; refinement will be slow and erratic, and derived parameters unreliable. This is inevitable, and has its mathematical basis in the matrix algebra of refinement.

I shall describe here two typical problems of this type and some simple, though not infallible, methods whereby the non-specialist can detect similar cases in the literature. Little familiarity with space groups is necessary; after all, few chemists care to venture into the recondite depths of space group tables,[10] the symmetry operators relating molecules often being chemically unimportant in comparison with the dimensions of a single molecule. Some recent detailed reviews of incorrectly assigned space groups are highly recommended reading.[11-14]

The first structure is a ternary oxide $X_4YO_3$.[15] There are four features which raise suspicion; each is in itself quite possible, but the combination of improbabilities is impressive. (a) The space group is given as P1 (synthetic materials sometimes crystallize in non-centrosymmetric space groups, but P1 is most unusual); (b) One cell angle is 89.99°; (c) there are two independent formula units in the asymmetric unit; (d) there are simple mathematical relationships between co-ordinates of pairs of atoms (see Table 1); thus each pair shows $\Delta x\ 1/2, \Delta y\ 0$. This last is the single strongest indication of 'missing' symmetry.

The unit cell $a\ 8.144, b\ 6.220, c\ 5.758$ Å, $\alpha\ 117.54, \beta\ 89.99, \gamma\ 111.24°$ can be transformed to a new cell $a'\ 11.030, b'\ 5.758, c'\ 8.144$ Å, $\alpha'\ 90, \beta'\ 114.10, \gamma'\ 90°$ by the matrix $0\ -2\ -1/0\ 0\ 1/-1\ 0\ 0$.[16] The same matrix transforms the reflection indices, whereupon the systematic absences $h0l, l$ odd and $hkl, h + k$ odd become apparent (corresponding to monoclinic space groups Cc or C2/c). The atom co-ordinates may be transformed with the transposed inverse matrix $(0\ -0.50/0\ -0.51/-1\ 0\ 0)$; with an allowed origin shift, the pairs of atoms are then related by the operator $x, -y, 0.5 + z$. The space group is thus Cc, and the refinement proceeds to $R$ 0.056 (see Table 1).*

The second structure, a hydrated oxide $H_3M_5O_{14}$,[15] will be discussed more briefly. It was published in the monoclinic space group $P2_1/m$, with $a\ 5.518, b\ 16.50, c\ 5.519$ Å, $\beta\ 107.0°$. In such cases (monoclinic with $a$ and $c$ equal, or triclinic with any two axes equal), a cell with an additional right angle may be constructed; here

---

*The crystallographic reasoning and matrix methods described here are probably unfamiliar to the non-specialist (for a fuller description, see *Ref*. 17—which is probably the best book on X-ray methods for the beginner). However, several programs are available which use matrix algebra to test for 'missing' symmetry, and which present the results in a form which is easy to interpret.[16] See also the 'Morals' below.

[10] 'International Tables for X-Ray Crystallography', Vol. A, ed. T. Hahn, D. Reidel Publishing Co., Dordrecht, Holland, 1983.
[11] R. E. Marsh and V. Schomaker, *Inorg. Chem.*, 1979, **18**, 2331.
[12] R. E. Marsh and V. Schomaker, *Inorg. Chem.*, 1981, **20**, 299.
[13] F. H. Herbstein and R. E. Marsh, *Acta Crystallogr.*, 1982, **B38**, 1051.
[14] R. E. Marsh and F. H. Herbstein, *Acta Crystallogr.*, 1983, **B39**, 280.
[15] Further details from the author on request.
[16] W. Clegg, *Acta Crystallogr.*, 1981, **A37**, 913.
[17] P. Luger, 'Modern X-Ray Analysis on Single Crystals', de Gruyter, Berlin, 1980.

**Table 1**  *Atom co-ordinates of the ternary oxide* $X_4YO_3$

| | P1 *cell* | | | Cc *cell* | | |
|---|---|---|---|---|---|---|
| | $x$ | $y$ | $z$ | $x$ | $y$ | $z$ |
| X1 | .600 | .659 | .769 | .671 | .288 | .399 |
| X2 | .102 | .658 | .193 | | | |
| X3 | .790 | .823 | .396 | .588 | .169 | .709 |
| X4 | .291 | .825 | .735 | | | |
| X5 | .378 | .460 | .225 | .771 | .158 | .120 |
| X6 | .881 | .457 | .540 | | | |
| X7 | .698 | .268 | .948 | .866 | .340 | .802 |
| X8 | .198 | .268 | .628 | | | |
| Y1 | .000 | .000 | .000[a] | .000 | .849 | .000[a] |
| Y2 | .500 | .000 | .302 | | | |
| O1 | .859 | .669 | .986 | .165 | .003 | .640 |
| O2 | .361 | .669 | .986 | | | |
| O3 | .546 | .911 | .559 | .340 | .489 | .918 |
| O4 | .047 | .909 | .659 | | | |
| O5 | .583 | .322 | .326 | .045 | .551 | .953 |
| O6 | .082 | .320 | .304 | | | |

[a] Fixed zero co-ordinates to define the origin

the matrix $1\ 0\ 1\ /\ 1\ 0\ -1\ /\ 0\ 1\ 0$ forms a cell with $a\ 6.565$, $b\ 8.872$, $c\ 16.50$ Å, all angles $90°$.[16] A detailed analysis (as above for $X_4YO_3$) shows that the structure can be described in the orthorhombic space group $Cmcm$. This can be seen from a diagram of the structure (Figure 1).

Two points should be stressed here. First, in both these cases the overlooked symmetry element was *not* a centre of symmetry, and the molecular dimensions of the published structures are thus reliable (for cases where this was not so, see below and *Refs.* 11—14). Secondly, the mere existence of a possible cell with two or three right angles (*i.e.* with metric symmetry monoclinic or orthorhombic, respectively) does not necessarily mean that the structure belongs to this crystal class. A good example is $\delta$-$S_7$,[18] which crystallizes with two independent molecules in a monoclinic cell with $a \simeq c$. This allows the construction of a metrically orthorhombic cell, but neither the atom co-ordinates nor the intensity data are consistent with such a transformation. Indeed, an attempt to obtain the structure in an orthorhombic space group had proved unsuccessful.[19]

The problem of deciding if a centre of symmetry is present or not (*i.e.* dis-

[18] R. Steudel, J. Steidel, J. Pickardt, F. Schuster, and R. Reinhardt, *Z. Naturforsch., Teil B*, 1980, **35**, 1378.

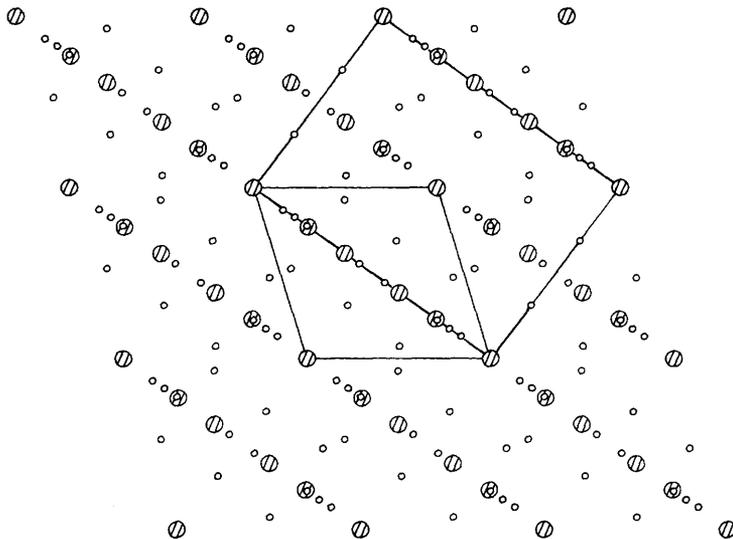[19] I. Kawada and E. Hellner, *Angew. Chem.*, 1970, **82**, 390.

**Figure 1** *The structure of* $H_3M_5O_{14}$ *projected down the monoclinic* b *axis. The orthorhombic symmetry (larger cell) is clearly visible. (Hatched circles* M, *small circles* O; H *atoms not determined)*

tinguishing between pairs of space groups such as P1, P$\bar{1}$; P*nma*, P*na*2$_1$) may present some difficulties. Describing a centrosymmetric structure in a non-centrosymmetric space group may well lead to a lower $R$ factor, since the number of parameters is approximately doubled (a modified version of the principle 'give me enough parameters and I can fit an elephant'); this is therefore no criterion. Statistical tests[20] can, when carefully used, resolve the ambiguity; unfortunately, careless use may lead to the wrong conclusion and thus to gross errors in published molecular dimensions.[21] A particularly unfortunate example (potassium tetroxalate, published in P1 instead of P$\bar{1}$) has been discussed instructively by Gilmore and Speakman.[22]

*Moral*: Overlooked crystallographic symmetry may lead to any (or several) of the following symptoms: synthetic compound in non-centrosymmetric space group; more than one formula unit in the asymmetric unit; accidental exact right angles in the cell; accidentally equal axes; simple mathematical relationships between co-ordinates of pairs of atoms. If the 'missing' symmetry is a centre of symmetry, serious systematic errors will be present.

[20] W. C. Hamilton, *Acta Crystallogr.*, 1965, **18**, 502.
[21] R. E. Marsh, *Acta Crystallogr.*, 1981, **B37**, 1985.
[22] C. J. Gilmore and J. C. Speakman, *Acta Crystallogr.*, 1982, **B38**, 2809.

### 5 Absorption Corrections: More Honoured in the Breach...
The absorption of $X$-rays by crystals is governed by the equation

$$I/I_0 = \exp(-\mu t)$$

where $I/I_0$ is the fractional loss of intensity, $t$ the path length, and $\mu$ the absorption coefficient. For Mo $K\alpha$ radiation and organic crystals, $\mu$ is of the order $0.1$ mm$^{-1}$ and $t < 0.8$ mm (the maximum $X$-ray beam diameter usually available), so absorption effects may safely be neglected. The introduction of heavy atoms increases $\mu$, and in an extreme case such as $Au_2O_3$,[23] $\mu t$ may be as high as 10 (crystal size $0.1$ mm, $\mu$ 110 mm$^{-1}$) and $I/I_0$ $5 \times 10^{-5}$. It is clear that in such cases severe systematic errors arise because of the different path lengths through the crystal for different diffraction geometries (*i.e.* different reflections). What is often *not* appreciated, because the necessary integral calculations are complicated, is that serious systematic errors also occur in the case of exactly spherical crystals.

The questions thus arise (i) how severe must absorption effects be before a correction is necessary, and (ii) how can a correction be applied. To answer (i) it is helpful to consider the consequences of absorption errors if no correction is applied. It can be shown that absorption is more severe at low $2\theta$; the general decrease of intensity with increasing $2\theta$ (independent of absorption) is thus to some extent cancelled out. Since this general decrease is caused by thermal motion of the atoms, the apparent thermal motion is artificially decreased by absorption effects. In severe cases this leads to the (physically impossible) phenomenon of negative thermal motion or 'non positive definite' atoms, in crystallographers' jargon. Related effects include the apparent highly anisotropic thermal motion of atoms if the absorption itself is highly anistropic (*e.g.* for needle- or plate-shaped crystals; see Figure 2), and the appearance of large artefacts in the electron density near heavy atoms. The atomic *positions* are not severely affected, although the location of light atoms may become difficult in the presence of too many spurious peaks.

It is generally agreed that such absorption effects become appreciable when $\mu t$ is *ca.* 0.5. Unfortunately, the application of adequate absorption corrections was, until recently, a far from trivial exercise. It involved first the calculation of crystal shape and size by assuming an idealized geometry (*e.g.* a cylinder) or by indexing the crystal faces and measuring the distances between them, and secondly, calculating path lengths and performing complex integrals for *each* reflection. If the crystal did not exhibit well-developed faces and did not approximate to a simple shape, the problem was intractable. Nowadays more general and convenient methods are available (*e.g.* the so-called $\psi$-scans, involving measuring a reflection and its equivalents at different diffraction geometries and applying a correction which minimizes the intensity differences; for a brief review of methods see *Ref.* 24). Regrettably, more effort is often expended in inventing excuses for lack of absorption corrections than would have been necessary to perform them in the first place. Here are some examples from recent publications:[15] 'The expense to information ratio was deemed too high'; 'with the absorption coefficient as *high* (my italics) as

[23] P. G. Jones, H. Rumpel, E. Schwarzmann, and G. M. Sheldrick, *Acta Crystallogr.*, 1979, **B35**, 1435.
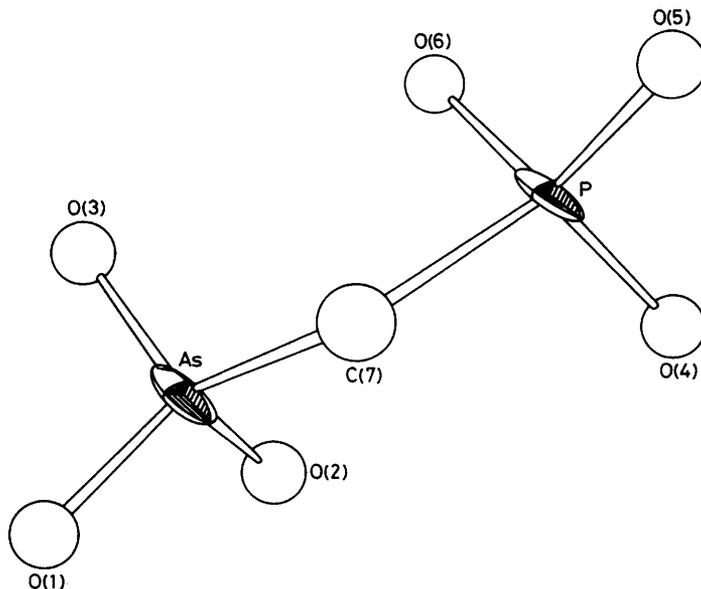[24] N. Walker and D. Stuart, *Acta Crystallogr.*, 1983, **A39**, 158.

**Figure 2** *The effects of neglecting absorption. The ion shown is* $(HO_3P \cdot CH_2 \cdot AsO_3H)^{2-}$; *note the spuriously elongated ellipsoids of the heavy atoms*[25]

$761\,cm^{-1}$, no absorption correction was attempted... The anisotropic thermal parameters are not given, since they did not stay positive definite'; 'in view of the very large absorption coefficient ($258\,cm^{-1}$) and the brittleness of the crystals, an absorption correction was out of the question'. Or, freely translated, 'we knew our data contained severe systematic errors but we didn't do anything about it'.

*Moral*: Structures with $\mu t$ 0.5 or greater need an absorption correction. (Since $\mu$ and the crystal size are generally amongst the published crystal data, the reader can calculate $\mu t$ for himself). The absence of absorption corrections in such cases reduces the accuracy of the structure and the meaningfulness of the thermal parameters.

**6 Thermal Ellipsoids: The Error Dustbin**
The conventional representation of the thermal motion of an atom by an ellipsoid (inside which the atom has a given chance, usually 50%, of being) is familiar to most chemists. A typical thermal ellipsoid plot of a well-behaved organic crystal structure is given in a related article in *Chemistry in Britain*.[26] Unusually large or anisotropic ellipsoids may genuinely represent unusual thermal motion, *e.g.* of a long side chain not stabilized by secondary interactions such as hydrogen bonds.

[25] L. Falvello, P. G. Jones, O. Kennard, and G. M. Sheldrick, *Acta Crystallogr.*, 1977, **B33**, 3207.
[26] P. G. Jones, *Chem. Brit.*, 1981, **17**, 222.

However, if no such obvious cause can be found, the ellipsoids should be regarded with some suspicion; they are a remarkably good way of mopping up systematic errors. One such error, absorption, was discussed in the last Section; if all the ellipsoids in a structure point the same way, the cause may well be absorption (other possibilities include crystal decay, crystal larger than $X$-ray beam . . .). An alternative possibility is that the model being refined is in error.

One source of error in the model is disorder; one or more atoms are 'averaged' over several positions, either by free rotation (*e.g.* of spherical ions such as $PF_6^-$)[26] or statically (two or more possible positions for the atom(s) involved). In both cases the 'average' model may be of a single site smeared out anisotropically. A typical example is furnished by the structures of some $MSb_5O_{13}$ salts.[27] The $M^+$ cations appear to lie on special positions with symmetry *mm*, but with extremely high temperature factor components in one direction. This model is almost certainly an approximation to a static disorder of $M^+$ over two or more sites, although, as is often the case, such a model could not be refined. Without a full chemical analysis these materials could not have been characterized.

The possibility of disorder should be borne in mind if unusual temperature factors are encountered. The converse, that normal temperature factors rule out disorder, is not necessarily true, as was recently shown in spectacular fashion by the structure of the secododecahedrene (1).[28] The structure solution appeared, most unexpectedly, to show a molecule of dodecahedrane (2), with normal thermal parameters and a good $R$ value. When the crystal was redissolved, n.m.r. spectroscopy confirmed the structure as (1). The explanation is that the molecules of (1) in the crystal adopt no fewer than 120 different orientations, the average of which is a good approximation to (2).



(1)  (2)

The representation of thermal motion by ellipsoids is only an approximation, and as such has a significant drawback; since the motion of, for instance, a terminal group may be nearer a 'banana' than an ellipsoid, the use of the latter as a model causes a systematic shortening of apparent bond lengths. This effect is known as 'libration'. As an example, consider the structure of carbonyl gold(I) chloride, $(OC)AuCl$,[29] in which the molecules are constrained by symmetry to be exactly

[27] D. Bodenstein, W. Clegg, G. Jäger, P. G. Jones, H. Rumpel, E. Schwarzmann, and G. M. Sheldrick, *Z. Naturforsch., Teil B*, 1983, **38**, 172.

[28] O. Ermer, *Angew. Chem.*, 1983, **95**, 251.

[29] P. G. Jones, *Z. Naturforsch., Teil B*, 1982, **37**, 823.

linear. The bond lengths Au—C 1.93, C—O 1.11 Å are unexpectedly short; this is almost certainly due to libration, and thus no great significance should be attached to these unreliable results. In general, if structural correlations are to be derived from small differences (*ca.* 0.01 Å) in bond lengths, the systematic errors introduced by libration may be highly significant, and in severe cases invalidate the discussion.*

In some cases a correction for libration may be applied (and where extreme accuracy is sought, such a correction should always be attempted) by a detailed analysis of the anistropic thermal parameters.[30, 31] The relative advantages of the two commonest methods have recently been reviewed.[32] The structure of (OC)AuCl is regrettably a good example of a case where a libration correction is impossible; the light atom parameters are imprecisely determined in the presence of gold, and, despite the application of an absorption correction, the extremely severe absorption would almost certainly lead to systematic errors in the thermal parameters. (In fact no anisotropic refinement of C and O was attempted).

*Moral*: (i) unusual temperature factors may reflect systematic errors in intensity data or errors in the structural model (in particular, disorder); (ii) appreciable thermal motion, especially perpendicular to a bond, causes an apparent shortening of a crystallographically determined bond length. For the most accurate structures, a libration correction should be applied; if this is not done, discussions based on small differences on bond lengths may be invalid.

### 7 Problems with Light Atoms: Small is Beautiful?

Since crystallographic methods provide a picture of the electron density, the positions of atoms with more electrons will be more accurately determined (this has been implicitly assumed above). However, given a reasonable absorption correction, it is usually possible to locate non-hydrogen atoms without difficulty, even in the presence of heavy transition metals (hydrogen atoms are a special case—see below). A more serious problem may be to distinguish similar light atoms (C, N, O, F), especially if the exact chemical constitution of the compound is not known. A case in point is the $SO_3F^-$ ion, for which the O and F sites are often impossible to distinguish (and may indeed be disordered).[33] More serious chemically is the inability to distinguish between alternatives such as M—NCO, M—OCN or M—NSO, M—OSN.[15]

Hydrogen atoms present a special, extreme case of the problems of light atom location. The single electron is difficult to find, and the apparent position will be shifted towards the atom X of the X—H bond (because the H electron spends a

*One extenuating circumstance is that a series of related compounds may well exhibit similar thermal motion, and thus similar librational effects; a discussion of *trends* in bond length would thus remain valid.

[30] W. R. Busing and H. A. Levy, *Acta Crystallogr.*, 1964, **17**, 142.
[31] V. Schomaker and K. N. Trueblood, *Acta Crystallogr.*, 1968, **B24**, 63.
[32] R. Srinivasan and N. R. Jagannathan, *Acta Crystallogr.*, 1982, **B38**, 2093.
[33] P. G. Jones and O. Kennard, *Acta Crystallogr.*, 1978, **B34**, 335.

finite time away from the H nucleus); accurate neutron diffraction data and spectroscopic data establish the internuclear length of a C—H bond as *ca.* 1.08 Å, whereas *X*-ray data generally give about 0.96 Å* [34]

The treatment of H atoms in *X*-ray structure determination may be considered under two headings.

**A. H Atom Determination in the presence of Heavy Elements.**—Small features in the residual electron density are examined when all other atoms have been found and satisfactorily refined. With luck, judgement, and good data, those in chemically sensible positions can be accepted as H atoms, although their refinement will probably be extremely imprecise unless certain constraints are used (*e.g.* fix the C—H bond at 0.96 Å*). It should now go without saying that adequate absorption corrections are necessary to minimise artefacts in the electron density. A far more reliable method of locating H atoms is neutron diffraction, where the H atom contributions to the diffraction are not overwhelmed by those of the heavier elements; this, however, requires very large crystals and access to an atomic reactor. To obviate this problem to some extent, Orpen[35] has developed an indirect method; possible H sites ('holes' in the *X*-ray structure) are analysed in respect of their potential energy, based on interaction curves between H and other atoms. High energy sites are ruled out. Although the energies calculated are very approximate, the success rate is impressive. Agreement with known neutron structures is achieved in virtually every case, thus lending confidence in calculated H atom positions when no neutron data are available. The method has also shown many published H atom positions to be false.[35]

**B. H Atom Location in Organic Structures.**—Given moderately accurate data and no disorder, H atoms are usually obvious features in the electron density when all other atoms are accounted for. The trend nowadays, in view of the imprecision of refined H atom positions, is to include the H atoms in the refinement using a 'riding model' incorporating various constraints (*e.g.* C—H 0.96 Å, as mentioned above, H—C—H 109.5° for *sp*$^3$ C, fixed H atom temperature factors). It therefore seems unnecessary to identify H atoms in the electron density, since their idealized positions are generated by the program. The consequence of this attitude is clear; if the atoms of the C/N/O framework have been wrongly identified (perhaps on the basis of a preconceived structure), the H atoms will also be wrongly placed. This error should be detected from several symptoms; the temperature factors of the wrongly assigned C/N/O atoms may be unusually high or low (another aspect of the 'error dustbin'—but who checks temperature factors carefully, especially now that many journals do not publish them); bond lengths and, in particular, non-bonded X · · · H contacts may be anomalous. Nevertheless, it is a safe assumption

---

*This value is incorporated into some program systems in order to include H atoms at geometrically calculated, rather than *X*-ray determined, sites.

[34] M. R. Churchill, *Inorg. Chem.*, 1973, **12**, 1213.
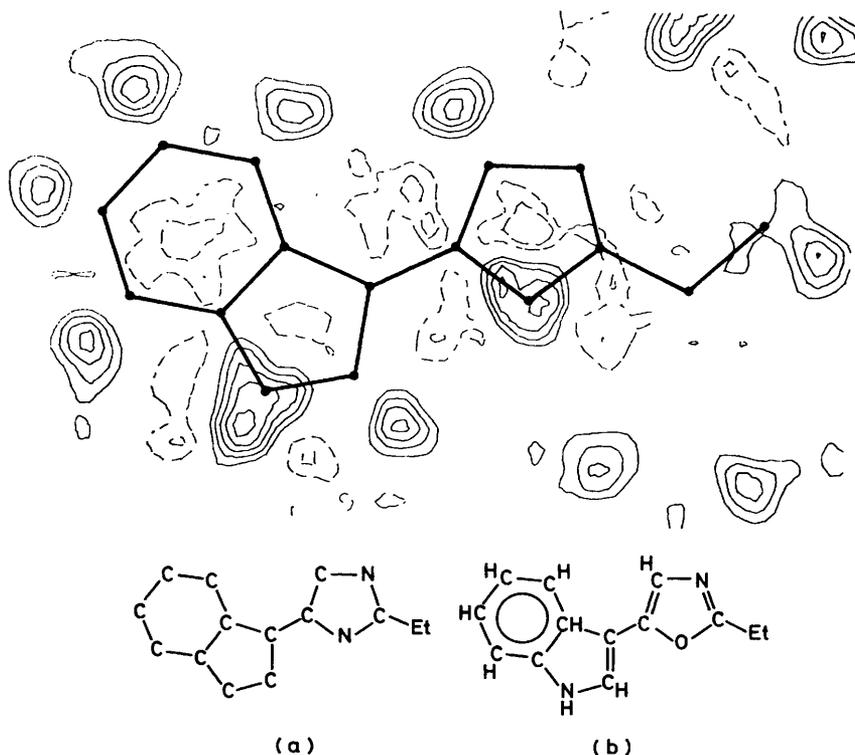[35] A. G. Orpen, *J. Chem. Soc., Dalton Trans.*, 1980, 2509.

**Figure 3** *Problems in light atom assignment : difference electron density of the largely planar molecule $C_{13}H_{12}N_2O$. Formulae: (a), as initially refined; (b), corrected version. The electron density clearly shows all hydrogen atoms of the rings and excess density at two atoms; the assigned atom types were amended accordingly*
(Reproduced by permission from *Ref.* 36)

that several such incorrect structures have been published; obvious candidates are large natural products, which are often of unknown structure and which may yield poor quality *X*-ray data.

An example of this problem is given in Figure 3. The structure $C_{13}H_{12}N_2O$ was solved without difficulty, but it was not immediately clear which atom types were which. A careful analysis of the difference electron density after initial refinement allowed the C/N/O assignement to be corrected and all H atoms to be located.

*Moral*: (i) H atom positions in heavy atom structures (*e.g.* cluster hydrides) may be erroneous if no supporting data, such as potential energy calculations, are available; (ii) H atom positions in organic structures are probably correct if the difference electron-density has been critically inspected; (iii) bonding modes of ions such as NCO⁻ are difficult to determine with *X*-ray methods alone.

[36] M. Noltemeyer, G. M. Sheldrick, H.-U. Hoppe, and A. Zeeck, *J. Antibiotics*, 1982, **35**, 549.

## 8 Acentric Structures: Through the Looking-glass†

Acentric (non-centrosymmetric) structures provide their own peculiar pitfalls. These arise mostly from the breakdown of Friedel's law (which states that the intensity of a reflection *hkl* is equal to that of *hkl*; the breakdown is caused by the phenomenon known as anomalous scattering of *X*-rays). The differences in intensity between *hkl* and *hkl* are generally immeasurably small for light atom structures; if heavier atoms are present, however, an accurate analysis of the measured differences can distinguish between a given structural model and its mirror image. For chiral materials, this corresponds to the determination of the absolute configuration (as in the pioneering work of Bijvoet[17]).

Clearly it is a great advantage to be able to determine absolute configuration by *X*-ray methods. The converse, though equally clear, is sometimes ignored; for any acentric structure containing heavier atoms, the refinement must take the anomalous scattering into account. This may be done either by (a) refining an alternative model with all co-ordinates changed in sign and accepting the model with the lower *R* value (valid for most space groups, but the difference in *R* may be very small—see below) or (b) refining a factor $\eta$ multiplying the imaginary components of the anomalous scattering contributions of all atoms; a value of $+1$ indicates a correct model, $-1$ that a change of sign of co-ordinates is necessary. The $\eta$ method, first suggested by Rogers,[37] is valid for all space groups and appears to be the more powerful of the two.

Bearing these comments in mind, it is distressing to find many published acentric structures for which no mention of anomalous scattering (or absolute configuration) is made, even with very heavy atoms present (*e.g.* Au[15]). It was long ago demonstrated that the neglect of anomalous scattering can lead to systematic errors in bond lengths.[38] A good example of this is shown by the gold(I) complex $(Ph_3Sb)_4Au^+ (C_6F_5)_2Au^-$,[39] which crystallizes in the acentric space group P3c1.* The refinement was initially unsatisfactory in that the six independent Au—Sb bond lengths showed a considerable scatter (2.52—2.72 Å). Changing the sign of the co-ordinates reduced this scatter to 2.59—2.67 Å (typical e.s.d.'s 0.01 Å) although having little effect on the *R* value. The latter model was accepted on the basis of this reduced scatter (the $\eta$ refinement was at that time unavailable). The former model, with its inappropriate treatment of anomalous scattering, would have led to systematic errors of up to 0.17 Å in bond lengths.

---

*Experienced crystallographers will have noticed that in P3c1, as in all acentric space groups with glide planes, there is no 'absolute configuration'; here the analysis of anomalous scattering determines instead the 'polar axis direction'. This is one of several oversimplifications in this Section, which do not, however, invalidate its general principles.

†Note added in proof: A greatly expanded version of Section 8 has been submitted to *Acta Crystallographica*. Further recommended reading: D. Rogers and F. H. Allen, *Acta Crystallogr.*, 1979, **B35**, 2823 (critique of published absolute configurations): H. D. Flack, *Acta Crystallogr.*, 1983, **A39**, 876 (a suggested alternative to the $\eta$ method).

[37] D. Rogers, *Acta Crystallogr.*, 1981, **A37**, 734.

[38] D. W. J. Cruickshank and W. S. McDonald, *Acta Crystallogr.*, 1967, **23**, 9.

[39] P. G. Jones, *Z. Naturforsch., Teil B*, 1982, **37**, 937.

*Moral:* All published acentric structures containing heavy atoms (how heavy is a heavy atom? Not very: the presence of one P atom in an organic structure can allow the determination of its absolute configuration) should also contain some mention of the treatment of anomalous scattering effects. If these effects were ignored, the bond lengths may be associated with systematic errors many times larger than the e.s.d.'s quoted.

## 9 Miscellany

This last section presents a brief selection of errors not easily classifiable under the headings 2—8. The experienced reader will doubtless be able to add to the list.

**A. Gross Errors in Published Data.**—Since crystallographic publications contain lists of co-ordinates, bond lengths, *etc.*, it is not surprising that numerical errors creep in from time to time. This can prove very confusing for the reader who tries to perform calculations with (for example) permuted axes. Fortunately, the Cambridge Crystallographic Data Centre checks all structures destined for its files, and erring authors (as the present author knows only too well) are asked to provide corrected values.

**B. Insufficient Data.**—Apart from the problem of unsuitably low $2\theta$ cutoff (Section 2), close examination of crystallographic datasets may reveal that not all independent data are present. In some high symmetry space groups, the crystallographer may not have been familiar with the index range of the unique data (*e.g.* Laue group $6/m$, data collection appropriate to $6/mmm$[13]). In acentric space groups, it is often merely a matter of taste (and machine time) whether Friedel opposites are collected, especially for light atom structures (see Section 8). However, the determination of absolute configuration without Friedel opposites may not be 100% reliable.

**C. Mis-interpretation of Peaks.**—Crystallographic methods enable us to locate peaks in the electron density. The assignment of atom types to these peaks is more subjective, and care should be taken not to confuse atoms of similar atomic number (see Section 7). A striking example of such a problem is the structure of the (alleged) novel chlorine(VII) compound $[ClF_6]^+ [CuF_4]^-$. It has recently been suggested[40] that traces of water and the use of silica vessels had led instead to $[Cu(H_2O)_4]^{2+} [SiF_6]^{2-}$, the cell of which is almost identical to that reported for the 'chlorine(VII)' compound.

[40] H. G. von Schnering and D. Vu, *Angew. Chem.*, 1983, **95**, 421.

[41] J. A. Ibers, 'Problem Crystal Structures', and J. Donohue, 'Incorrect Crystal Structures: Can They Be Avoided?', in 'Critical Evaluation of Chemical and Physical Structural Information', ed. D. R. Lide jr. and M. A. Paul, Nat. Acad. Sci., Washington D.C., 1974

It is appropriate in these concluding paragraphs to explain why a crystallographer should seek to sow the seeds of doubt in the minds of those who are, after all, his customers. The answer is very simple: only by making the chemist more familiar with the problems involved, and constructively critical of the results obtained, can consistently high standards of published crystal structures be achieved.

I have tried, wherever possible, to give examples of problems and errors taken from my own work. The diligent reader could certainly find some that I have overlooked.

Finally, it is hoped that this article will stimulate the chemist to find out more about crystallographic methods, both by reading some of the excellent general texts available[17, 42] and, much better, by active collaboration with departmental crystallographers.

[42] J. D. Dunitz, '*X*-Ray Analysis and the Structure of Organic Molecules', Cornell U.P., New York, 1979.