Chapter 6.1.7.7

SHELXL-97

Lists & Tables

George W. Sheldrick
Dept. of Structural Chemistry
Tammannstraße 4
37077 Göttingen, Germany
email: gsheldr@shelx.uni-ac.gwdg.de

The esds in bond lengths, angles and torsion angles, chiral volumes, Ueq, and coefficients of least-squares planes and deviation of atoms from them, are estimated rigorously from the full correlation matrix (an approximate treatment is used for the angles between least-squares planes). The errors in the unit-cell dimensions (specified on the ZERR instruction) are taken into account exactly in estimating the esds in bond lengths, bond angles, torsion angles and chiral volumes. Correlation coefficients between the unit-cell dimensions are ignored except when determined by crystal symmetry (so that for a cubic crystal the cell esds contribute to errors in bond lengths and chiral volumes but not to the errors in bond angles or torsion angles). The (rather small) contributions of the unit-cell errors to the esds of quantities involving least-squares planes are estimated using an isotropic approximation.

For full-matrix refinement, the esds are calculated after the final refinement cycle. In the case of BLOC'ed refinement, the esds are calculated after every cycle (except that esds in geometric parameters are not calculated after pure Uij/sof cycles etc.), and the maximum estimate of each esd is printed in the final tables. This prevents some esds being underestimated because not all of the relevant atoms were refined in the last cycle, but at the cost of overestimating all the esds if the R-factor drops appreciably during the refinement. Thus large structures should first be refined almost to convergence (either by CGLS or L.S./BLOC), and then a separate final blocked refinement job performed to obtain the final parameters and their esds. It is important that there is sufficient overlap between the blocks to enable every esd to be estimated with all contributing atoms refining in at least one of the refinement cycles.

BOND atomnames
BOND outputs bond lengths for all bonds (defined in the connectivity list) that involve two atoms named on the same BOND instruction. Angles are output for all pairs of such bonds involving a common atom. Numerical parameters on a BOND instruction are ignored, but not treated as errors (for compatibility with SHELX-76). A BOND instruction with no parameters outputs bond lengths (and the corresponding angles) for ALL bonds in the connectivity table, and 'BOND $H' on its own includes all bonds to hydrogens as well (but since the hydrogens are not included in the connectivity table, bonds involving symmetry equivalent hydrogens are not included). Other element names may also be referenced globally by preceding them with a '$' on a BOND instruction. BOND is set automatically by ACTA, and the bond lengths and angles are written to the *.cif* file. Note that the best way to calculate B-H-B angles is with RTAB !

CONF atomnames
The named atoms define a chain of at least four atoms. CONF generates a list of torsion angles with esd's for all torsion angles defined by this chain. CONF is often used to specify an n-membered ring, in which case the first three atoms must be

named twice (n+3 names in all). If no atoms are specified, all possible torsion angles not involving hydrogen are generated from the connectivity array. The torsion angles generated by CONF are also written to the .cif file if an ACTA instruction is present. All torsion angles calculated by SHELXL follow the conventions defined by Allen & Rogers (1969).

MPLA na atomnames
A least-squares plane is calculated through the first na of the named atoms, and the equation of the plane and the deviations of all the named atoms from the plane are listed with estimated standard deviations (from the full covariance matrix). The angle to the previous least-squares plane (if any) is also calculated, but some approximations are involved in estimating its esd. na must be at least 3. If na is omitted the plane is fitted to all the atoms specified.

RTAB codename atomnames
Chiral volumes (one atomname), bonds (two), angles (three) and torsion angles (four atomnames) are tabulated compactly against residue name and number. codename is used to identify the quantity being printed; it must begin with a letter and not be longer than 4 characters (e.g. 'Psi' or 'omeg'). There may not be more than 4 atom names. It is assumed that the atoms have the same names in all the required residues. For chiral volumes only, the necessary bonds must be present in the connectivity list (the same conventions are employed as for CHIV). Since the atoms do not themselves have to be in the same residue (it is sufficient that the names match), the residue name (if any) is printed as that of the first named atom for distances, the second for angles, and the third in the case of torsion angles. The latter should be consistent with generally accepted conventions for proteins. A typical application of RTAB for small-molecule structures is the tabulation of hydrogen-bonded distances and angles (with esd's) since these will not usually appear in the tables created automatically by BOND. For an example of this see the 'sigi' test job in chapter 3.

If RTAB refers to more than one residue (e.g. RTAB_*), it is ignored for those residues in which not all the required atoms can be found (e.g. some of the main chain torsional angles for the terminal residues in a protein).

HTAB dh[2.0]
The new HTAB instruction provides an analysis of the hydrogen bonds. A search is made over all polar hydrogens (i.e. hydrogen bonded to electronegative elements) present in the structure, and hydrogen bonds printed for which: **H•••A < r(A)+dh and <DHA > 110°**. If it appears likely that the hydrogens have been assigned wrongly (e.g. two -OH groups have been assigned to the same O•••O vector) a suitable warning message appears. This output should be checked carefully, since the algorithms used by HFIX/AFIX to place hydrogens are by no means infallible! To obtain esd's on the distances and angles involved in the hydrogen bond, the second

form of the HTAB instruction (and if necessary EQIV) should be used (see below); HTAB without atom names is used first to find the necessary symmetry transformations for EQIV..

HTAB donor-atom acceptor-atom

The second form of the HTAB instruction is required to generate the esds and the CIF output records. The donor atom D and acceptor A should be specified; the program decides which of the hydrogen atoms (if any) makes the most suitable hydrogen bond linking them. Only the acceptor atom may specify a symmetry operation (_$n) because this standard CIF entry for publication in Acta Crystallographica requires this.

LIST m[#] mult[1]

m = 0:   No action.

m = 1:   Write h,k,l, F_o, F_c and phase (in degrees) to .fcf in X-PLOR format. Only unique reflections after removing systematic absences, scaling [to an absolute scale of F(calc)], applying dispersion and extinction or SWAT corrections (if any), and merging equivalents including Friedel opposites are included. If F_o^2 was negative, F_o is set to zero. Reflections suppressed by OMIT or SHEL [or reserved for R(free)] are not included.

m = 2:   List h,k,l, F_o, σ (F_o) and phase angle in degrees in FORMAT(3I4,2F8.2,I4) for the reflection list as defined for m = 1.

m = 3:   List h,k,l, F_o, σ (F_o), A(real) and B(imag) in FORMAT(3I4,4F8.2), the reflections being processed exactly as for m = 2.

m = 4:   List h,k,l, F_c^2, F_o^2, $\sigma(F_o^2)$ and a one-character status flag. F_o^2 are scaled to F_c^2 and possibly corrected for extinction, but no corrections have been made for dispersion and no further merging has been performed. FORMAT (3I4,2F12.2,F10.2,1X,A1) is employed. The status flag is 'o' (observed), 'x' [observed but suppressed using 'OMIT $h\ k\ l$', SHEL or reserved for R(free)], or '<' (F_o^2 is less than t.$\sigma(F_o^2)$), where t is one half of the F-threshold s specified on an OMIT instruction).

m = 5:   Write h,k,l, F_o, F_c,   and ϕ (phase   angle   in   degrees)   in FORMAT(3I4,2F10.2,F7.2) for the reflection list as defined for m = 1. Like the m = 1 option, this is intended for input to somestandard macromolecular FFT programs (such as W. Furey's PHASES program), thereby providing a possible route to a graphical display of the electron density.

m = 6:   Write a free-format CIF file containing h,k,l, F_o^2, $\sigma(F_o^2)$, F_c and ϕ (phase angle in degrees) for the reflection list as defined for m = 1. This is the recommended format for the deposition of reflection data with the PDB, and is also the format required for the generation of refinement statistics and electron density maps using SHELXPRO.

For m = 4 only, mult is a constant multiplicative factor applied to all the quantities output (except the reflection indices!), and may be used if there are scaling problems. For other m options mult is ignored. For m = 2,3 or 4 only a blank line is included at the end of the file as a terminator. The reflection list is written to the file *name.fcf*, which is in CIF format for n = 3, 4 or 6; however the actual reflections are always in fixed format except for n = 1 or 6. The program CIFTAB can - amongst other options - read the m = 4 output and print F_o/F_c/T(F) tables in compact form on an HP-compatible laser printer. n = 4 is the standard archive format for small-molecule structures, n = 6 for macromolecules (with Friedel opposites averaged). Since the final refinement is normally performed on all data (including the R_{free} reference set) the LIST 6 output is not able to flag the R_{free} reflections.

ACTA 2thetafull[#]
A 'Crystallographic Information File' file *name.cif* is created in self-defining STAR format. This ASCII file is suitable for data archiving, network transmission, and (with suitable additions) for direct submission for publication. ACTA automatically sets the BOND, FMAP 2, PLAN and LIST 4 instructions, and may not be used with other FMAP or LIST instructions or with a positive OMIT s threshold. A warning message appears if the cell contents on the UNIT instruction are not consistent with the atom list, because they are used to calculate the density etc. which appears in the *.cif* output file.

2thetafull is used to specify the value of 2θ for which the program calculates the completeness of the data for the CIF output file as required by Acta Crystallographica. If no value is given, the program uses the maximum value of 2θ for the reflection data. If the data were collected to a specific limiting 2θ, or if a limit was imposed using SHEL, this would be a good choice. Otherwise the choice of 2thetafull is a difficult compromise; if it is too low, the paper will be rejected because the resolution of the data is not good enough; if it is higher, the lower completeness might lead to rejection by the automatic Acta rejection software! SHELXL calculates the completeness by counting reflections after merging Friedel opposites and eliminating systematic absences (and the reflection 0,0,0).

SIZE dx dy dz
dx, dy and dz are the three principal dimensions of the crystal in mm, as usually quoted in publications. This information is written to the *.cif* file. If a SIZE instruction is present in the *.ins* file, SHELXL uses it to write the estimated minimum and maximum transmission to the *.cif* file. This should give order of

magnitude estimates that should be replaced by the values from the actual absorption correction if these were applied. The empirical SHELXL estimates take into account that most of the diffraction from strongly absorbing crystals takes place at the edges and corners; these estimates of the actual absorption of the crystal may be a little smaller than those from psi-scan and other semi-empirical routines that include absorption by the mounting fibre and glue or oil.

TEMP T[20]

Sets the temperature T of the data collection in degrees Celsius. This is reported to the *.cif* file and used to set the default isotropic U values for all atoms. TEMP must come before all atoms in the *.ins* file. TEMP also sets the default X-H bond lengths (see AFIX) which depend slightly on the temperature because of librational effects. The default C-H bond lengths and default U-values are rounded to two decimal places so that they may be quoted more easily.

WPDB n[1]

Writes the refined coordinates to a *.pdb* file. If n is positive hydrogen atoms are omitted; if |n| is 1 all atoms are converted to isotropic and ATOM statements generated, and if |n| is 2 ANISOU statements are also generated (but the equivalent B value is still used on the ATOM statement). The atom names and residue classes and numbers should conform to PDB conventions. This provides a direct link to X-PLOR and other programs which use (more of less) the official (Brookhaven) dialect of the PDB format. Note that SHELXPRO can be used to extend the PDB output file to include refinement details etc. (from the .lst file) for deposition with the PDB, and also to modify disordered residues so that they can be interpreted by programs such as O that cannot read the full standard PDB format.