Chapter 5.3.3

SHELXS-97

Patterson Interpretation

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

The Patterson superposition procedure in SHELXS was originally designed for the location of heavier atoms in small moiety structures, but it turns out that it can also be used to locate heavy atom sites for macromolecular ΔF data (see Chapter 15). For further details and examples see Sheldrick (1996) and Sheldrick, Dauter, Wilson, Hope & Sieker (1993).

5.3.3.1 Patterson interpretation algorithm

The algorithm used to interpret the Patterson to find the heavier atoms in the new version of SHELXS is totally different to that used in SHELXS-86; it may be summarized as follows:

1. One peak is selected from the sharpened Patterson (or input by means of a VECT instruction) and used as a superposition vector. This peak must correspond to a correct heavy-atom to heavy-atom vector otherwise the method will fail.   The entire procedure may be repeated any number of times with different superposition vectors by specifying 'PATT nv', with |nv| > 1, or by including more than one VECT instruction in the same job.

2. The Patterson function is calculated twice, displaced from the origin by +U and -U, where U is the superposition vector. At each grid point the lower of the two values is taken, and the resulting 'superposition minimum function' is interpolated to find the peak positions. This is a much cleaner map than the original Patterson and contains only 2N (or 4N etc. if the superposition vector was multiple) peaks rather than N^2. The superposition map should ideally consist of one image of the structure and its inverse; it has an effective 'space group' of $P\bar{1}$ (or $C\bar{1}$ for a centered lattice etc.).

3. Possible origin shifts are found which place one of the images correctly with respect to the cell origin, i.e. most of the symmetry equivalents can be found in the peak-list. The SYMFOM figure of merit (normalized so that the largest value for a given superposition vector is 99.9) indicates how well the space group symmetry is satisfied for this image.

4. For each acceptable origin shift, atomic numbers are assigned to the potential atoms based on average peak heights, and a 'crossword table' is generated. This gives the minimum distance and Patterson minimum function for each possible pair of unique atoms, taking symmetry into account. This table should be interpreted by hand to find a subset of the atoms making chemically sensible minimum interatomic distances linked by consistently large Patterson minimum function values.   The PATFOM figure of merit measures the internal consistency of these minimum function values and is also normalised to a maximum of 99.9 for a given superposition vector.   The Patterson values are recalculated from the original F_o

data, not from the peak-list.   For high symmetry space groups the minimum function is calculated as an average of the two (or more) smallest Patterson densities.

5. For each set of potential atoms a 'correlation coefficient' (Fujinaga and Read, 1987) is calculated as a measure of the agreement between E_o and E_c, and expressed as a percentage. This figure of merit may be used to compare solutions from different superposition vectors.

5.3.3.2 Instructions for Patterson Interpretation

PATT nv [#] dmin [#] resl [#] Nsup [#] Zmin [#] maxat [#]
nv is the number of superposition vectors to be tried; if it is negative the search for possible origin shifts is made more exhaustive by relaxing various tolerances etc. dmin is the minimum allowed length for a heavy-atom to heavy-atom vector; it affects ONLY the choice of superposition vector. If it is negative, the program does not generate any atoms on special positions in stage 4 (useful for some macromolecular problems). resl is the effective resolution in Å as deduced from the reflection data, and is used for setting various tolerances. If the data extend further than the crystal actually diffracted, or if the outer data are incomplete, it may well be worth increasing this number. This parameter can be relatively critical for macromolecular structures. Nsup is the number of unique peaks to be found by searching the superposition function.   Zmin is the minimum atomic number to be included as an atom in the crossword table etc. (if this is set too low, the calculation can take appreciably longer). maxat is the maximum number of potential atoms to be included in the crossword table, and can also appreciably affect the time required for PATT.

VECT X Y Z
A superposition vector (with coordinates taken from the Patterson peak-list) may be input by hand by a VECT instruction, in which case the first two numbers on the PATT instruction are ignored (except for their signs !), and a PATT instruction will be automatically generated if not present in the .ins file.   There may be any number of VECT instructions.

In the unlikely event of a routine PATT run failing to give an acceptable solution, the best approach - after checking the data reduction diagnostics carefully as explained above - is to select several potential heavy-atom to heavy-atom vectors by hand from the Patterson peak-list and specify them on VECT instructions (either in the same job or different jobs according to local circumstances) for use as superposition vectors. The exhaustiveness of the search can also be increased - at a significant cost in computer time - by making the first PATT parameter negative and/or by increasing the value of resl a little.   The sign of the second PATT parameter (a negative sign excludes atoms on special positions) and the list of

elements which might be present (SFAC/UNIT) should perhaps also be reconsidered.

5.3.3.3 Instructions for partial structure expansion

TEXP na [#] nH [0] Ek [1.5]
na PHAS reflections with E_o > Ek and the largest values of E_c/E_o are generated for use in partial structure expansion or direct methods. The first nH atoms (heavy atoms) in the atom list are retained during partial structure expansion, the rest are thrown away after calculating phases. At least one atom MUST be given! TEXP automatically generates appropriate FMAP, GRID and PLAN instructions.

TEXP (and/or PHAS) may be used in conjunction with TREF to generate fixed phases for use in direct methods; the special TEXP option na = 0 provides point atom phases for ALL reflections, which are then refined during the phase annealing and tangent expansion stages of direct methods (as specified on the PHAN and TREF instructions). It is not necessary to use different starting phases for the different phase sets, because the phase annealing stage itself introduces (statistically distributed) random phase shifts! This is a powerful method of partial structure expansion for cases when the phasing power of the partial structure is not quite adequate, e.g. when it consists of only one atom (say P or S in a large organic structure). If at least 5 atoms have been correctly located then TEXP alone should suffice.

When TEXP is used without TREF a tangent formula expansion (to all reflections with E > Emin as specified on the ESEL instruction) is first performed, followed by several cycles (see FMAP) of E-Fouriers and peak-list optimization. TEXP is particularly useful for cases in which several not very heavy atoms (e.g. P, S) have been located by PATT followed by hand interpretation of the resulting 'crossword table'. In such cases nH should be set to the number of such atoms and na to about half the number of reflections with E > 1.5 (see the first page of the SHELXS-96 output).

PHAS h k l phi
A fixed phase for structure expansion or direct methods. PHAS may be used to fix single phase seminvariants that have been obtained from other programs or derived by examination of the best TREF solutions. The phase angle phi must be present, and should be given in degrees.

atomname sfac x y z sof [1] U (or U11 U22 U33 U23 U13 U12)
Atom instructions begin with an atom name (up to 4 characters which do not correspond to any of the SHELXS command names, and terminated by at least one blank) followed by a scattering factor number (which refers to the list defined by the SFAC instruction(s)), x, y, and z in fractional coordinates, and (optionally) a

site occupation factor (s.o.f.) and an isotropic U or six anisotropic U_{ij} components (both in $Å^{-2}$). The U or U_{ij} values are ignored by SHELXS but may be included for compatibility with SHELXL.

When SHELXS writes the *.res* output file, a dummy U value is followed by a peak height (unless an atom type has been assigned by the program before the *E*-Fourier recycling). Both the dummy U and the peak height are ignored if the atom is read back into SHELXS (e.g. for partial structure expansion). SHELXL also ignores the peak height if found in the .ins file. In contrast to SHELX-76 it is not necessary to pad out the atom name to 4 characters with blanks, but it should be followed by at least one blank. References to 'free variables' and fixing of atom parameters by adding 10 as in SHELX-76 and SHELXL will be interpreted correctly, but SHELXL AFIX, RESI and PART instructions are simply ignored (so idealized hydrogen atoms etc. are NOT generated). The site occupation factor for an atom in a special position should be divided by the number of atoms in the general position that have coalesced to give the special position. It may also be found by dividing the multiplicity of the special position (as as given in International Tables) by the multiplicity of the general position. Thus an atom on a fourfold axis will usually have s.o.f. = 10.25 (i.e. 0.25, fixed by adding 10).

MOVE dx [0] dy [0] dz [0] sign [1]
The coordinates of the following atoms are changed to: x = dx + sign * x, y = dy + sign * y, z = dz + sign * z (after applying FRAG and SPIN - if present - according to PATSEE conventions); MOVE applies to all following atoms until superseded by a further MOVE.   MOVE is normally used in conjunction with SPIN and FRAG (see below) but is also useful on its own for applying origin shifts.

TEXP may be used in conjunction with ESEL -1 for a partial structure expansion in the effective space group P1 (C1 etc. if the lattice is centered). This can be very effective if it is suspected that a fragment is correctly oriented but translated from its real position, or if the space group cannot be unambiguously assigned. Hand interpretation of the resulting *E*-map is then however necessary to locate the positions of the crystallographic symmetry elements.