Chapter 6.1.7.4

SHELXL-97

The Connectivity List

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

The connectivity list is a list of 'bonds' that is set up automatically, and may be edited using BIND and FREE. It is used to define idealized hydrogen atom positions, for the BOND and PLAN output of bond lengths and angles, and by the instructions DELU, CHIV, SAME and SIMU. Hydrogen atoms are excluded from the connectivity list (except when introduced by hand using BIND).

CONN bmax[12] r[#] atomnames or CONN bmax[12]
The CONN instruction fine-tunes the generation of the connectivity table and is particularly useful when π-bonded ligands or metal ions are present in the structure. For the purposes of the connectivity table (which is always generated), bonds are all distances between non-hydrogen atoms less than r1 + r2 + 0.5 Å, where r1 and r2 are the covalent radii of the atoms in question (taking PART into consideration as explained below). A shell of symmetry equivalent atoms is also generated, so that all unique bonds are represented at least once in the list. All bonds, including those to symmetry equivalent atoms, may be deleted or added using the FREE or BIND instructions.

Default values of r (identified by the scattering factor type) are stored in the program. These defaults may be changed (for both the connectivity table AND the PLAN -n output) by using the full form of the SFAC instruction. Alternatively the defaults may be overridden for the named atoms by specifying r on a CONN instruction, in which case r is used in the generation of the connectivity list but not by the PLAN instruction. '$' followed by an element name (the same as on a SFAC instruction) may also be employed on a CONN instruction (and also does not apply to PLAN). The second form of the CONN instruction may be used to change the maximum coordination number bmax for all atoms (which defaults to 12 if there is no CONN instruction).

If, after generating bonds as above and editing with FREE and BIND, there are more than bmax bonds to a given atom, the list is pruned so that only the bmax shortest are retained. A harmless side-effect of this pruning of the connectivity list is that symmetry operations may be stored and printed that are never actually used. Note that this option only removes one entry for a bond from the connectivity list, not both, except in the case of 'CONN 0' which ensures that there are no bonds to or from the named atoms. 'CONN 0' is frequently used to prevent the solvent water in macromolecular structures from making additional 'bonds' to the macromolecule which confuse the generation of idealized hydrogen atoms etc. In some cases it will be necessary to use FREE to remove a 'bond' from a light atom to an alkali metal atom (for example) in order to generate hydrogen atoms correctly. Refinements of macromolecules will often include BUMP and 'CONN 0 O_200 > LAST' (where the water happens to begin with residue 200). 'LAST' is used to indicate the last atom in the file, which saves trouble when adding extra waters.

The CONN instruction, like ANIS and HFIX, MUST precede the atoms to which it is to be applied. Repeated CONN instructions are allowed; the LAST relevant CONN preceding a particular atom is the one which is actually applied. CONN without atom names changes the default value of bmax for all following atoms. The following example illustrates the use of CONN:

CONN Fe 0
MPLA 5 C11 > C15 Fe
MPLA 5 C21 > C25 Fe
Fe .....
C11 .....
.....
C25 .....

which would prevent bonds being generated from the iron atom to all 10 carbons in ferrocene. In this example, the distances of the iron atom from the two ring planes would be calculated instead.

PART n sof
The following atoms belong to PART n of a disordered group. The automatic bond generation ignores bonds between atoms with different PART numbers, unless one of them is zero (the value before the first PART instruction). If a site occupation factor (sof) is specified on the PART instruction, it overrides the value on the following atom instructions (even if set via an AFIX instruction) until a further PART instruction, e.g. 'PART 0', is encountered).

If n is negative, the generation of special position constraints is suppressed and bonds to symmetry generated atoms with the same or a different non-zero PART number are excluded; this is suitable for a solvent molecule disordered on a special position of higher symmetry than the molecule can take (e.g. a toluene molecule on an inversion center). A PART instruction remains in force until a further PART instruction is read; 'PART 0' should be used to continue with the non-disordered part of the structure.

Some care is necessary in generating hydrogen atoms where disordered groups are involved. If the hydrogen atoms are assigned a PART number, then even if the atom to which they are attached has no part number (i.e. PART 0) the above rules may be used by the program to work out the correct connectivity for calculating the hydrogen atom positions. HFIX hydrogens are assigned the PART number of the atom to which they are attached. If the hydrogens and the atom to which they are attached belong to PART zero but the latter is bonded to atoms with non-zero PART, the LOWEST of these non-zero PART numbers is assumed to be the major component and is used to calculate the hydrogen positions. In general, if the same residue numbers and names and the same atom names but different PART numbers are used for different disorder components in a macromolecule, HFIX will generate

hydrogen atoms correctly without any special action being required. For example the use of HFIX with the following disordered serine residue:

```
HFIX_Ser 33 N
HFIX_Ser 13 CA
HFIX_Ser 23 CB
HFIX_Ser 83 CG
 :
RESI 32 Ser
N .....
CA .....
C .....
O .....
PART 1
CB   1 ... ... ... 21 ...
OG   4 ... ... ... 21 ...
PART 2
CB   1 ... ... ... -21 ...
OG   4 ... ... ... -21 ...
PART 0
```

would set up the AFIX hydrogens as if the following had been input. Note that only one, fully occupied, hydrogen is attached to CA; for this reason, and also to prevent small inconsistencies in the DFIX and DANG restraints, the disorder should be traced back one more atom than can be resolved (i.e. CB should be split even if it does not look as though this would be necessary in an electron density map):

```
RESI 32 Ser
N .....
AFIX 43
H0   2 ... ... ... 11 -1.2
AFIX 0
CA .....
AFIX 13
HA   2 ... ... ... 11 -1.2
AFIX 0
C .....
O .....
PART 1
CB   1 ... ... ... 21 ...
AFIX 23
HB1  2 ... ... ... 21 -1.2
HB2  2 ... ... ... 21 -1.2
AFIX 0
OG   4 ... ... ... 21 ...
```

```
AFIX 83
HG   2  ...  ...  ...  21  -1.5
AFIX 0
PART 2
CB   1  ...  ...  ... -21  ...
AFIX 13
HB1  2  ...  ...  ... -21  -1.2
HB2  2  ...  ...  ... -21  -1.2
AFIX 0
OG   4  ...  ...  ... -21  ...
AFIX 83
HG   2  ...  ...  ... -21  -1.5
AFIX 0
PART 0
```

where free variable 2 is the occupation factor for PART 1 (say 0.7) and the occupation factor of the second component is tied to 1-fv(2) (i.e. 0.3). The value for this free variable is set on the FVAR instruction and is free to refine. If there were more than two components, a linear free variable restraint (SUMP) could be used to restrain the sum of occupation factors to unity. The addition of disorder components after including hydrogen atoms will require some hand editing and so is less efficient, but the auxiliary program SHELXPRO can be persuaded to do most of the work

BIND atom1 atom2
The specified 'bond' (which may be of any length) is added to the connectivity list if it is not there already. Only one of the two atoms may be an equivalent atom (i.e. have the extension _$n).

FREE atom1 atom2
The specified 'bond' is deleted from the connectivity list (if present). Only one of the two atoms may be an equivalent atom (i.e. have the extension _$n).