



# SHELXS is a very automatic Black Box. PATT solves a Patterson and is best for structures with a few heavy atoms in combination with many light atoms. Works very good in centrosymmetric space groups. TREF uses direct methods. You need atomic resolution (say 1.2 Å or better). Read: Sheldrick, G. M. Acta Cryst. Sect. A (1990), 46, 467. Direct methods have problems in the presence of inversion centers (use PATT or solve in noncentrosymmetric space group and transform by hand). Sometimes TREF 1000 (or 5000) helps.

#### **Structure Solution with SHELXS**

```
TITL b11 in P2(1)/c
CELL 0.71073 9.0890 16.7690 15.7000 90.000 99.310 90.000
ZERR 4.00 0.0018 0.0034 0.0031 0.000 0.030 0.000
LATT 1
SYMM -X, 0.5+Y, 0.5-Z
SFAC C H N SI CL
UNIT 96 140 4 4 8
TREF
HKLF 4
END
TITL bll in P2(1)/c
CELL 0.71073 9.0890 16.7690 15.7000 90.000 99.310 90.000
ZERR 4.00 0.0018 0.0034 0.0031 0.000 0.030 0.000
LATT 1
SYMM -X, 0.5+Y, 0.5-Z
SFAC C H N SI CL
UNIT 96 140 4 4 8
PATT
HKLF 4
END
```

#### **Structure Solution with SHELXS**

TITL bll in P2(1)/c CELL 0.71073 9.0890 16.7690 15.7000 90.000 99.310 90.000 ZERR 4.00 0.0018 0.0034 0.0031 0.000 0.030 0.000 LATT 1 SYMM -X, 0.5+Y, 0.5-Z SFAC C H N SI CL UNIT 96 140 4 4 8 TREF 1000 HKLF 4 END Force 1000 cycles of direct methods. More than 5000 is not really sensible.

In stubborn cases try **ESEL** or possibly **TEXP** (see manual) or, much easier, use SHELXD (XM) instead.

#### **Structure Solution with SHELXS**

```
TITL bl1 in P2(1)/c

CELL 0.71073 9.0890 16.7690 15.7000 90.000 99.310 90.000

ZERR 4.00 0.0018 0.0034 0.0031 0.000 0.030 0.000

LATT 1

SYMM -X, 0.5+Y, 0.5-Z

SFAC C H N SI CL

UNIT 96 140 4 4 8

TREF -363117

HKLF 4

END

TREF followed by a negative number generates only the

solution with the given code number (to be found in the .lst file

from the SHELXS run before!).

Sometimes it is possible to solve the structure in a space

group of lower symmetry (P1 instead of P-1).
```

# SHELXD is also a Black Box but a little less automatic.

**FIND n** Generates n *random* atoms. Go for ca. 25% fewer atoms than you expect to be in the asymmetric unit

**PLOP** Expands the list of atoms in cycles. Go for more atoms than you expect. Use three to five cycles.

MIND Specifies distances between atoms and symmetry equivalents (for example search for specific bond lenghts *i.e.* vectors in the Patterson possibly combined with the PATS or PSMF command). For small molecule structures go for: MIND 1.0 -0.1

Usón, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643-648. Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772-1779.

## **Structure Solution with SHELXD**

TITL	bll in P2	(1)/c								
CELL	0.71073	9.0890	16.7690	15.7000	90.000	99.310	90.000			
ZERR	4.00	0.0018	0.0034	0.0031	0.000	0.030	0.000			
LATT	1									
SYMM	-X, 0.5+Y	, 0.5-Z								
SFAC	CHNSI	CL								
UNIT	96 212 4	48								
FIND	22	Generate	e 22 random	atoms.						
PLOP	30 37 41	Expand p	beak list in th	nree rounds	to 30, 37, 4	1 atoms.				
MIND	1.0 -0.1	1.0 is mi	nimum allow	ed distance	between a	toms				
	-0.1 Å is maximum allowed distance for an atom from its own									
				(default is +	2.5), i.e. nc	special po	sition test.			
NTRY	1000	Stop afte	r 1000 FINE	) cycles.						
HKLF	4									
END										
SHELXD can do <u>much</u> more, but <b>FIND</b> , <b>PLOP</b> and <b>MIND</b> should be										
enough for small molecule structures and most of the other commands are										
	specific an									
	lfide bridges				ige about		uio (o.g.			
uisu	inde blidges	s ale 3.37	$\tau$ long, etc.	).						

## **Structure Refinement**

The solution from SHELXS is frequently already very good. However, the coordinates are not quite accurate, the atom types of some or all atoms have been assigned incorrectly (if at all), and details of the structure are missing (H-atoms, disorders, solvent molecules, *etc.*).

The atomic positions in the first . res file are not the direct result of the diffraction experiment, but an interpretation of the electron density function calculated from the measured intensities and the "somehow determined" phase angles.

Better phases can be calculated from the atomic positions, which allow re-determining of the electron density function with a higher precision. From the new electron density map, more accurate atomic positions can be derived, which lead to even better phase angles, and so forth.

#### **Structure Refinement**

Close examination of the  $F_o$ - $F_c$  map helps to introduce new atoms and remove "bad" old ones.

Once all non-hydrogen atoms are found, the atoms can be refined anisotropically.

Once the model is anisotropic, the hydrogen atom positions can be determined or calculated.



Refinement against  $F^2$  or F? Past: F Advantage: Faster computing. Problems:  $I \sim F^2$ . That means extraction of a root! Difficult for very weak reflections. Negative reflections need to be ignored or arbitrarily set to a small positive number. Estimation of  $\sigma(F)$  from  $\sigma(F^2)$  is very difficult. The least squares method is very sensitive to the weights, which are calculated from the standard uncertainties. Refinement against F results in inaccuracies in the refinement. Now:  $F^2$ Advantages: none of the problems mentioned arise. Disadvantage: A little slower. or  $M = \sum w (|F_o| - |F_c|)^2$  $M = \sum w \left( F_o^2 - F_c^2 \right)^2$ *F*: structure factor; *o*: observed; *c*: calculated; *w* weighting factor (derived from  $\sigma$ ).





# **Constraints and Restraints**

Both improve the data-to-parameter-ratio: Constraints remove parameters, restraints add data.

#### Constraints

Constraints are mathematical equations, relating two or more parameters or assigning fixed numerical values to certain parameters, hence reducing the number of independent parameters to be refined.

Site occupation factors are constraints present in every structure. Even for disordered atoms the sum of the occupancies is constrained to add up to 1.0.

Atoms on special position require constraints for their coordinates, occupancies and sometimes also their ADPs:

Special Position Constraints								
Special position	Constraints on coordinates	Constraints on U <sup>ij</sup> values	Constraints on occupancies					
Inversion Centre	<i>x</i> , <i>y</i> , <i>z</i> fixed to lie on inversion centre	None	0.5					
Mirror plane $\ \ \ $ to <i>b</i>	<i>y</i> is fixed to lie on the mirror plane	$U^{12} = U^{23}$	0.5					
Twofold axis parallel to b	x and z fixed to lie on the twofold axis	$U^{23} = U^{12} = 0$	0.5					
Tetragonal fourfold	y and z fixed to lie on the fourfold axis	$U^{11} = U^{22}$ and $U^{12} = U^{13} = U^{23} = 0$	0.25					



# **Rigid Group Constraints**

A group of atoms with known (or assumed) geometry: Refine six parameters (translation and rotation) rather than 3N parameters for the individual atoms (9N for anisotropic). A seventh parameter can be refined as scale factor.

Typical examples: Cp or Cp\* ligands, phenyl rings,  $SO_4^-$ , perchlorate ions, *etc*.

In SHELXL: **AFIX mn** / **AFIX 0** command **m** describes the geometry of the group and **n** the mathematical treatment.

## **Hydrogen Atoms**

Only 1 delocalized electron.

Hydrogen atoms can be "placed on mathematically calculated positions and refined using a riding model".

That means X-H distances and H-X-H or H-X-Y angles are constraint to certain values. Not the hydrogen *positions*!

In SHELXL:

**HFIX** mn generates the appropriate **AFIX** commands Again m describes the geometry of the group and n the mathematical treatment.

#### **Other Constraints**

EADP atom1 atom2 Forces the two atoms to have identical ADPs

EXYZ atom1 atom2 Force the two atoms to have identical oordinates.

**EADP** and **EXYZ** can be useful for disorders or the refinement of mixed crystals (*e.g.* zeolithes).



#### **Geometrical Restraints**

Besides a restraint on chiral volumes (CHIV) and a restraint for atoms that are supposed to lie on a common plane (FLAT), SHELXL has two kinds of distance restraints: <u>direct</u> and <u>relative</u> distance restraints. The former restrain distances to a given target value (DFIX, DANG), the latter restrain equivalent distances to be equal (SADI, SAME).

Advantage of relative distance restraints: No need for "outside" information, refinement converges well (*esp.* for Z'>1).

Disadvantage: underestimated standard uncertainties of bond lengths and angles. And it is too easy to refine a structure in a space group with too low symmetry ( $\rightarrow$  you'll be Marshed).

DFIX d s atomnames syntax wrong in book!!! The distance between the atom-<u>pairs</u> named in atomnames is restrained to posses the value d within the standard uncertainty s (default 0.02).

#### DANG

Just as **DFIX**, but default standard uncertainty is 0.04. Use **DFIX** for 1,2-distances ad **DANG** for 1,3-distances.

#### SADI s atomnames

Restrains the distance between two or more <u>pairs</u> of atoms named in **atomnames** to be equal within the standard uncertainty **s** (default value 0.02).

#### SAME s1 s2 atomnames

The command **SAME**, followed by a list of atom names, must be located at the correct position within the .ins file. **SAME** makes the first atom in the list of atom names equivalent to the first atom immediately following the **SAME** command, the second atom equivalent to the second following, *etc.* 

"Equivalent" means here that the 1,2- and 1,3-distances of corresponding atoms are restrained to be equal within the standard deviations s1 or s2 (default values are 0.02 for 1,2-and 0.04 for 1,3-distances).

The program automatically sets up the n(n-1)/2 restraint equations that are required when n atoms should be equal.





#### FLAT s atomnames

The atoms named in **atomnames** are restrained to lie on a common plane within the standard uncertainty **s** (default value  $0.1 \text{ Å}^3$ ).

#### CHIV V s atomnames

The chiral volumes of the named atoms are restraint to the value of v within the standard uncertainty of s (default value: 0.1 Å<sup>3</sup>. The default value for v is 0.

The chiral volume is defined as the volume of the tetrahedron formed by the three bonds to an atom. The sign of the chiral volume is determined by the alphabetical order of the atoms forming the three bonds.

*E.g.* the chiral volume of the alpha carbon in an L-amino acid is ca. 2.5 Å.

#### **Restraints on Displacement Parameters**

**SIMU** and **DELU** take into account that atoms, which are bound to one another, move similarly, both in direction and amount.

**ISOR** assumes approximate isotropic behaviour for otherwise anisotropically refined atoms.

Both **SIMU** and **DELU** are based on physically very sensible assumptions and can be used on almost all atoms in a model when the data-to-parameter-ratio is low or other problems with the refinement make this desirable.

**SIMU** should not be applied uncritically to very small ions and atoms that are part of freely rotation groups.

#### DELU s1 s2 atomnames

This rigid bond restraint is applied to all bonds connecting to atoms mentioned in atomnames. It restrains the ADPs of two atoms <u>in the direction of the bond between them</u> to be equal within the standard uncertainty s1 (default 0.01). If no atomnames are given, all atoms are understood.

#### SIMU s st dmax atomnames

Atoms closer to one another than  $\operatorname{dmax}$  (default: 1.7 Å) are restraint to have the same  $U^{ij}$  components within the standard uncertainty of **s** (default value: 0.04). For terminal atoms **st** is assumed (default: 0.08). If no **atomnames** are given, all atoms are assumed.

**SIMU** is much bolder an assumption than **DELU** (hence the much larger standard uncertainty).



The U<sup>ij</sup> values of the atoms mentioned in atomnames are refined to behave approximately isotropic within the standard uncertainty s, or st for terminal atoms (default 0.1 and 0.2). If no atomnames are given, all atoms are understood.

**ISOR** can be useful for solvent molecules, esp. water, for which **SIMU** and **DELU** are ineffective.



# **Other Restraints**

The **SUMP** command allows to linearly relate several free variables:

```
SUMP c sigma c1 m1 c2 m2 ...
```

applies the following linear equation to the specified free variables:

 $c = c1 \cdot fv(m1) + c2 \cdot fv(m2) + ...$ 

where c is the target value for the restraint and sigma the standard uncertainty. c1, c2, etc. are weighting factors and frequently 1; m1, m2, etc. refer to the individual free variables.

#### **Other Restraints**

DEFS globally changes the default standard uncertainties for the following restraints: CHIV, DANG, DELU, DFIX, FLAT, SADI, SAME and SIMU using the following syntax:

DEFS sd[0.02] sf[0.1] su[0.01] ss[0.04] maxsof[1]

In parentheses are the default values.

## **Other Restraints**

DEFS sd[0.02] sf[0.1] su[0.01] ss[0.04] maxsof[1]

sd is the default for s in DFIX and SADI, and for s1 in the SAME instruction; for DANG twice the value of sd is applied. sf is the default standard uncertainty for CHIV and FLAT, su is the default value for s1 and s2 in DELU, and ss is the default value for s in SIMU.

The default value for st in SIMU and ISOR, as well as s2 in SAME are calculated from the respective s or s1 values (unless specified differently by the user).

**maxsof** specifies the maximum value up to which a site occupation factor is allowed to refine to. Fixed site occupation factors and sofs linked to free variable are not restricted by **maxsof**.

#### **Free Variables**

In general, any parameter P or any **DFIX**, **DANG**, or **CHIV** restraint can be defined in the .ins file as

#### 10 · m + p

There are four different cases:

where fv(m) is the value of the m<sup>th</sup> free variable.

## **Free Variables**

#### 10 · m + p

m = 0: the parameter P with the starting value p is refined freely.

Trivial: describes a refinable parameter P as possessing the starting value p.

## **Free Variables**

#### 10 · m + p

m = 1: the value of p is fixed and not refined at all.

Assume you want to <u>constrain</u> an atom to lie on a mirror plane parallel to the *a*-*b* plane at  $c = -\frac{1}{4}$ . The task is to fix the value for the *z* coordinate to -0.25.

According to the above, this can be done by giving m the value of 1, and the value for p should be the atomic parameter of z (*i.e.* -0.25).

Hence, the atomic parameter for z in the .ins file for this atom reads 9.75.

## **Free Variables**

#### 10 · m + p

m = 1: the value of p is fixed and not refined at all.

To give a second example: Sometimes it can be helpful to fix the isotropic displacement parameter of an atom, U, at a certain value, for example 0.05.

As always when parameters are fixed: m = 1; and p is the desired value for *U*: 0.05.

The site occupation factor for the atom in question is then given as 10.05.

#### **Free Variables**

#### 10 · m + p

 $m > 1: P = p \cdot fv(m)$   $m < -1: P = p \cdot [fv(-m)-1]$ where fv(m) is the value of the m<sup>th</sup> free variable.

This involves additional free variables. Most common case: disorder.

Linking the occupancy of an atom to the <u>second</u> free variable, instead of the first one:

sof: 11.0000 → 21.0000 11.0000 → -21.0000

#### **Free Variables**

#### 10 · m + p

 $m > 1: P = p \cdot fv(m)$  $m <-1: P = p \cdot [fv(-m)-1]$ where fv(m) is the value of the m<sup>th</sup> free variable.

**CHIV** and the distance restraints **DFIX** and **DANG** can also be combined with free variables:

#### **Free Variables**

**CHIV** and the distance restraints **DFIX** and **DANG** can also be combined with free variables. *E.g.* restraining a  $ClO_4^-$  ion to be tetrahedral: Assuming the atoms in the ion are named Cl(1) and O(1) to O(4), the restraints using **SADI** are as follows:

SADI C11 01 C11 02 C11 03 C11 04 SADI 01 02 01 03 01 04 02 03 02 04 03 04

Using DFIX and the second free variable in the same scenario:

DFIX 21 Cl1 01 Cl1 02 Cl1 03 Cl1 04 DFIX 21.633 01 02 01 03 01 04 02 03 02 04 03 04

This corresponds to m = 2 for the second free variable and p is 1.0 in the first line and 1.633 in the second line (taking into account that the 1,3-distances in a regular tetrahedron are 1.633 times as long as the 1,2-distances).

The value of the second free variable is refined freely and will converge at the mean Cl-O-distance.

The advantage of the second way is that the average CI-O distance will be calculated with a standard uncertainty (in addition to the individual CI-O distances with their standard uncertainties). The disadvantage is that one additional least squares parameter is to be refined (the second free variable).

# **Crystallographic Results**

The final .res file contains the complete anisotropic model with all hydrogen atoms, which can be used to generate attractive figures for scientific publications (or grant proposals) and gain several kinds of information about a molecule. The most obvious are bond lengths and angles, but numerous other quantities can be calculated from the atomic coordinates, such as torsion angles or hydrogen bonds.

## **Bond Lengths and Angles**

If the command **BOND** appears in the header of the .ins file, SHELXL writes into the .lst file a table of all bond lengths and angles in the connectivity table.

**BOND** \$H expands this table to include all distances and angles involving hydrogen atoms as well.

# **Torsion Angles**

If the crystallographer or chemist wishes certain torsion angles to appear in a separate table in the .lst file, each of these torsion angles can be specified in a **CONF** command:

#### CONF atomnames

where **atomnames** defines a covalent chain of at least four atoms. If no atom names are specified, SHELXL generates all possible torsion angles.

#### **Planes**

#### MPLA na atomnames

SHELXL calculates a least-squares plane through the first **na** atoms of the named atoms. The equation of this plane, together with the deviations of all named atoms from the plane and the angle to the previous least-squares plane (if present) is written into the .lst file.

If **na** is not specified, the program fits the plane through all named atoms

# Hydrogen Bonds

If the command **HTAB** appears in the header of the .ins file, SHELXL performs a search over all polar hydrogen atoms present in the structure and examines hydrogen bonding.

The bonds listed in the .lst file are those for which the distance between acceptor and hydrogen atom are smaller than the radius of the acceptor atom plus 2.0 Å, and the angle between the donor atom, the hydrogen and the acceptor atom is larger than  $110^{\circ}$ .

## Hydrogen Bonds

HTAB donor-atom acceptor-atom

SHELXL generates hydrogen bonds with standard uncertainties and, in combination with **ACTA**, the appropriate table in the .cif file.

**EQIV** can be used to specify a symmetry equivalent of the acceptor atom.

# RTAB

#### RTAB codename atomnames

The command RTAB allows to compile a variety of structural quantities. Depending on how many atoms are specified in the qualifier **atomnames** the following is calculated (and tabulated):

- chiral volumes (one atom specified)
- distances (two atoms)
- angles (three atoms)
- torsion angles (four atoms specified)

**codename** must be specified and serves as an aid to identify the tabulated quantity in the .lst or .cif file. It must begin with a letter and cannot be longer than four characters.

# ACTA

If the command **ACTA** appears in the header of an .ins file, SHELXL generates a .cif file.

ACTA automatically sets the BOND, FMAP 2, PLAN and LIST 4 instructions.

ACTA cannot be combined with other FMAP or LIST commands.

Torsion angles defined by **CONF** and hydrogen bonds defined by **HTAB** are also written into the .cif file.

The quantities defined by **RTAB** and **MPLA** commands do not appear in the .cif file. They can be found in the .lst file only.

#### XP

XP predates Windows, which explains a lot... XP is best learned by using it. You have the printout of the help file. We'll take the plunge on Friday.

Important commands:

HELP

READ, REAP, FMOL, INFO, PROJ, TELP, UNIQ, GROW, FUSE, NAME, PICK, SAVE, SORT, FILE

BANG, ENVI, SYMM, SGEN, PUSH, CENT

EDEN, MPLN, MATR, DIAG, PERS, ORTH, OFIT, PACK, MGEN, INVT, PRUN

MIT OpenCourseWare http://ocw.mit.edu

5.067 Crystal Structure Refinement Fall 2009

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.