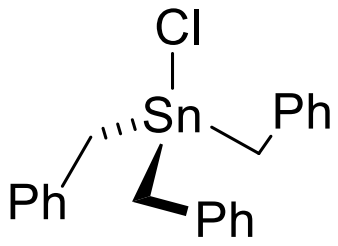
CCCW 2021

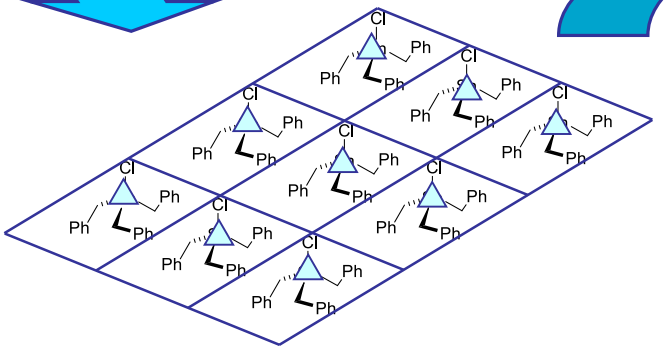
SHELX and the SHELX .ins file structure

Frank Schaper – Université de Montréal

Overview of a crystal structure determination

Crystallisation → **Crystal selection**





Molecular structure:
Atom positions

Crystal structure:
Unit cell and space group

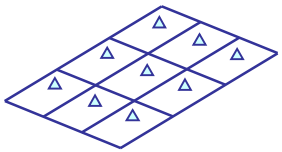
Crystal:
Macroscopic dimensions

Reflection intensities

H	K	L	I	σ
0	0	1	134.4	12.5
0	0	2	0.2	1.2
1	1	4	52.4	2.2

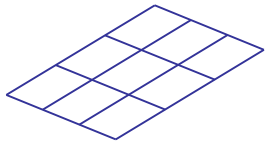
Solution : find a first estimate for the phases values
and refinement : refine your electron density to match the observed intensities

Data collection



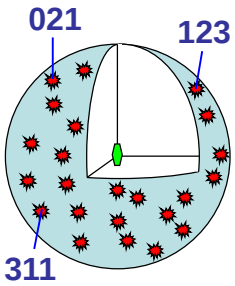
Space group
Translational symmetry and point group of the crystal

Space group determination

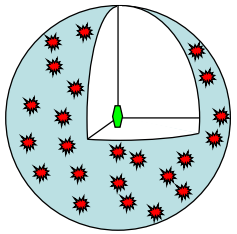


Unit cell

H	K	L	I	σ
0	0	1	134.4	12.5
0	0	2	0.2	1.2
1	1	4	52.4	2.2

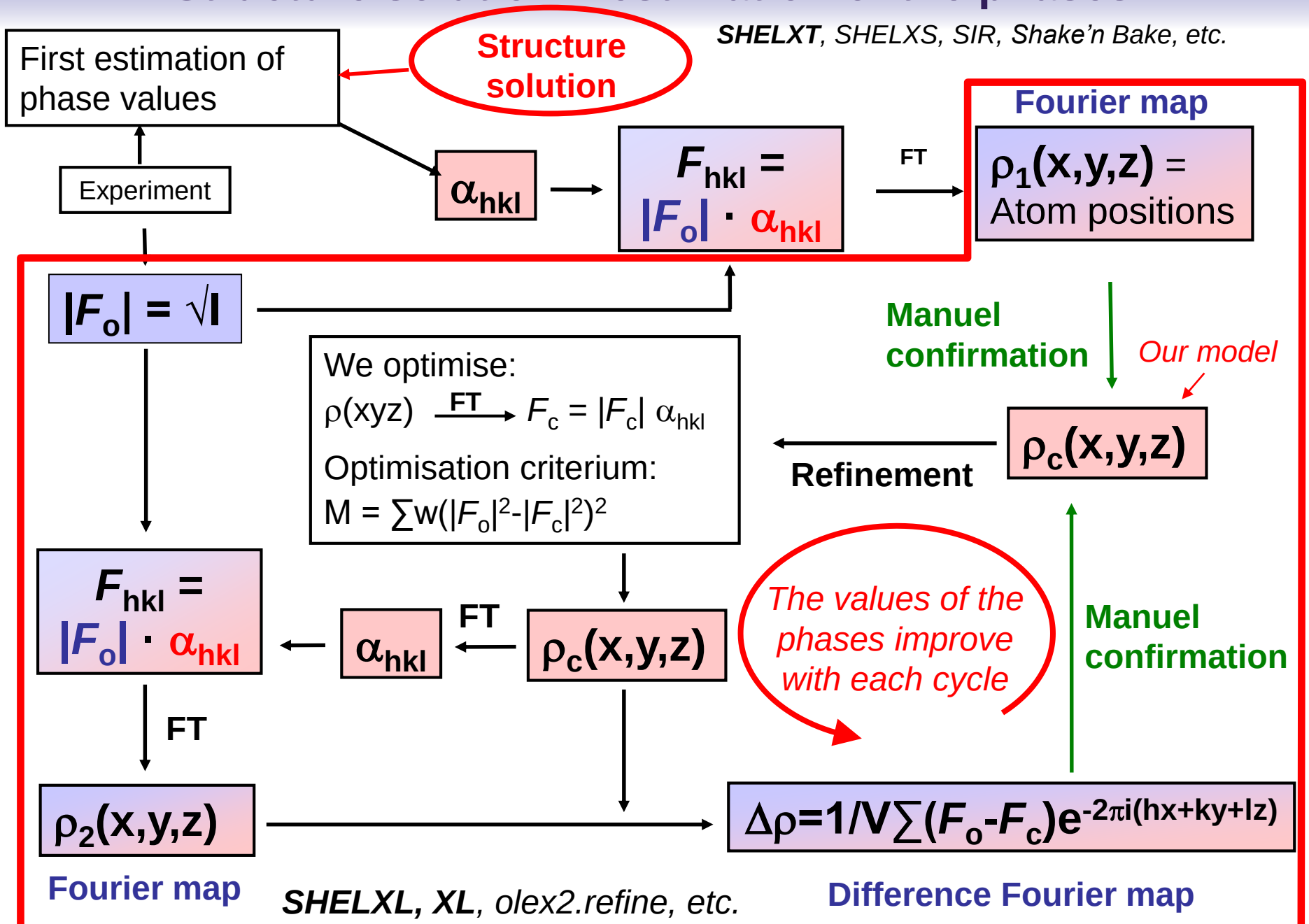


Unit cell determination



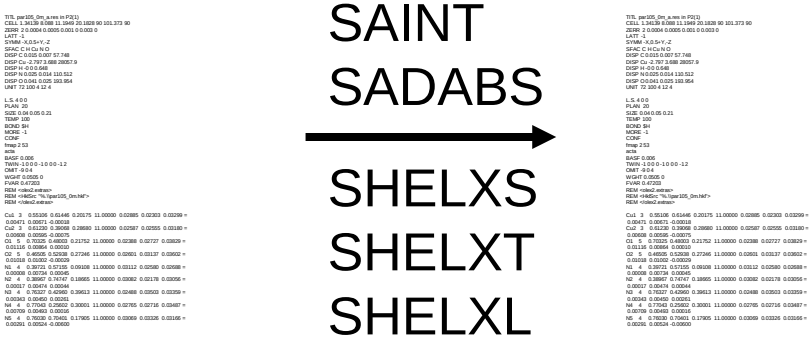
Raw data

Structure solution = estimation of the phases

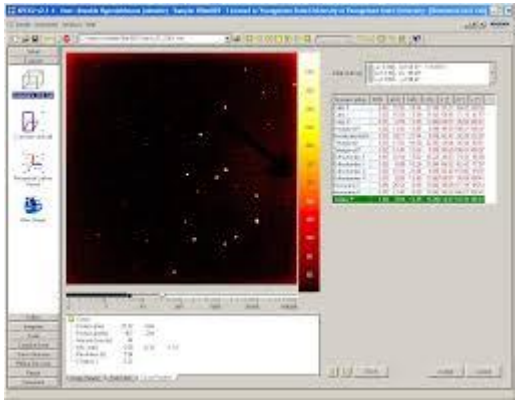


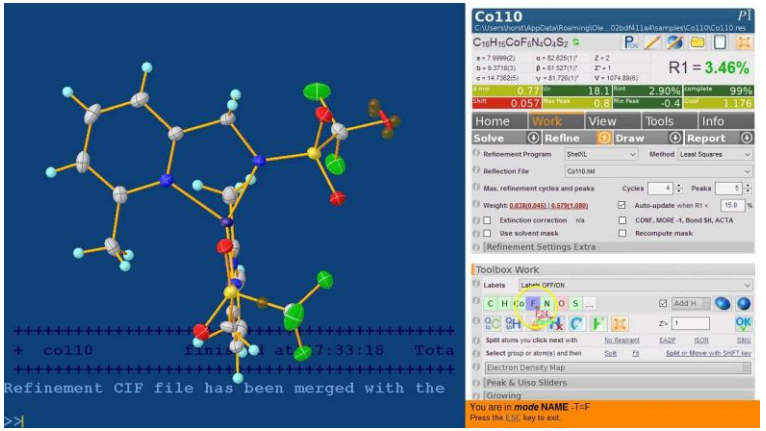
The difference between the GUI and the crystallographic programs

The **crystallographic programs** perform the necessary calculations using the reflections data. Typically they use **text files** as input or output.



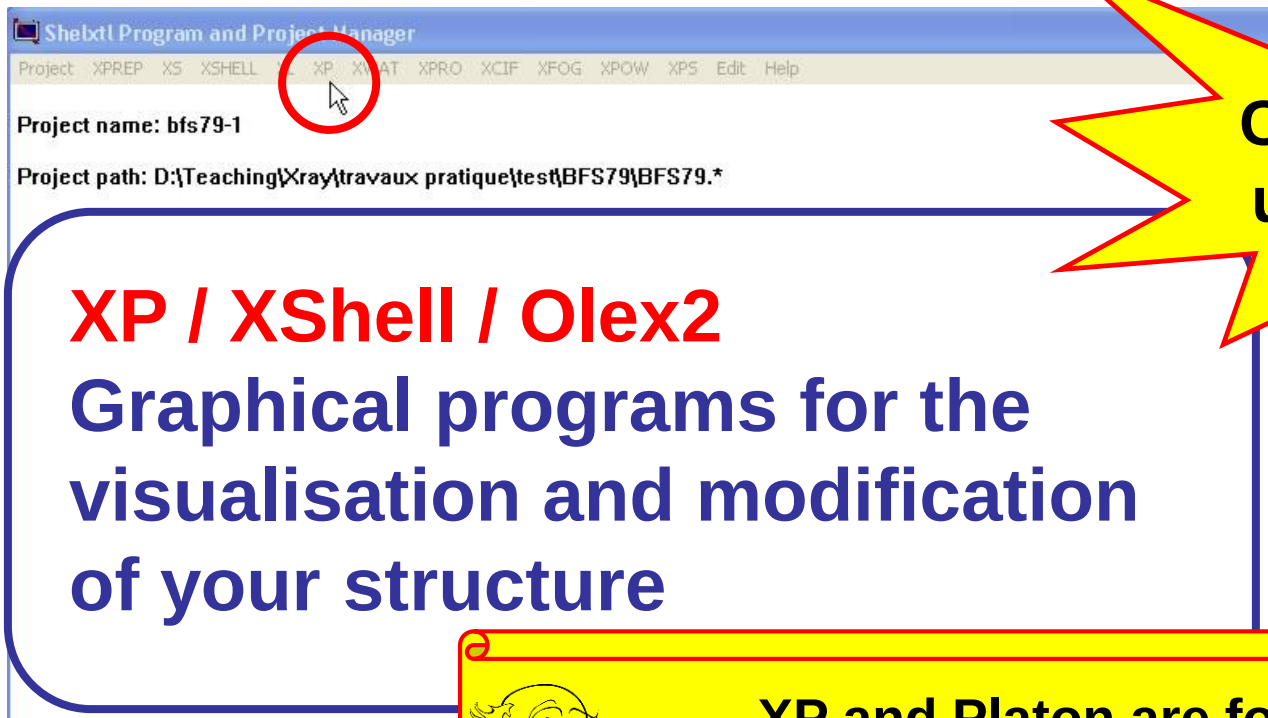
GUIs (graphical user interface) allow a visualisation of the structure and to enter commands in form of menus. It is important to realize that **GUIs simply change the input files** and start the respective crystallographic program in the background.



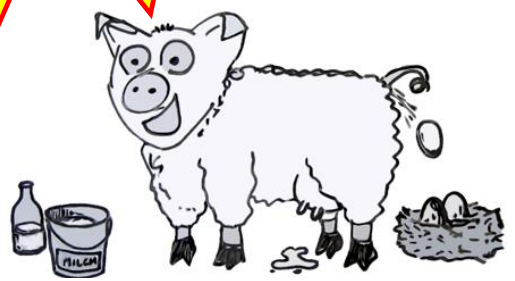


Choice of GUI

Cool people use OLEX2!



XP / XShell / Olex2
Graphical programs for the visualisation and modification of your structure



Egg-laying Woolmilkpig

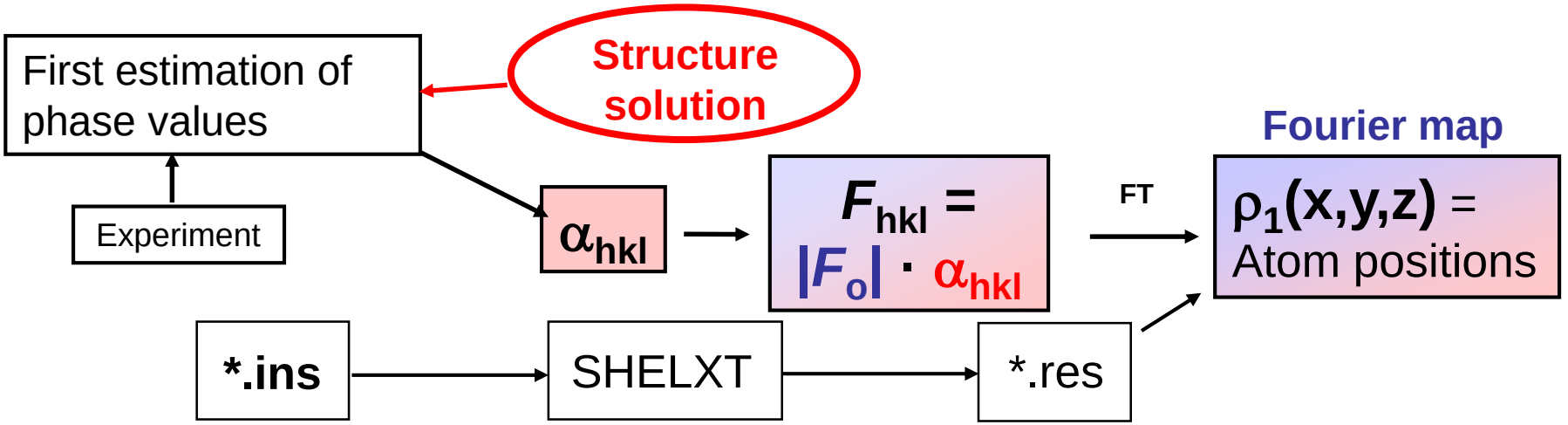
XP and Platon are for hardcore crystallographers

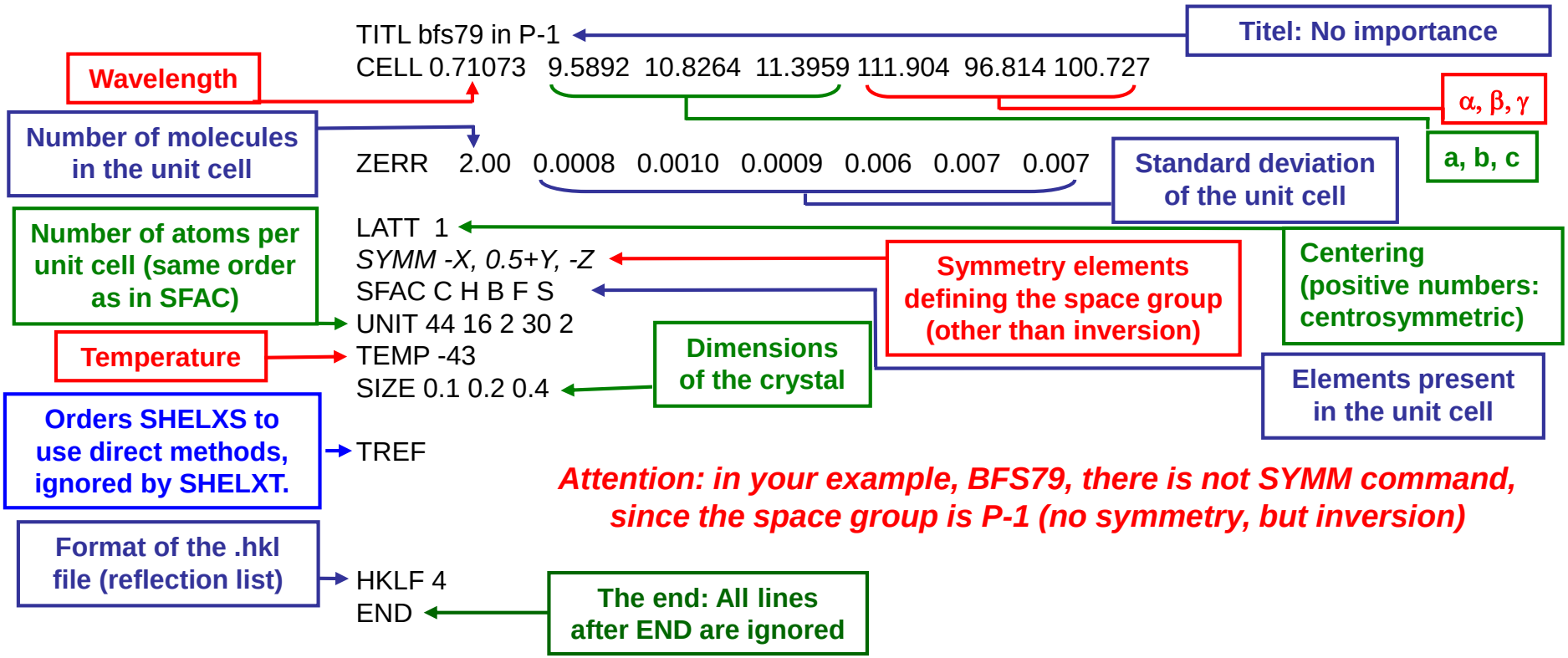
Nobody wants XShell

WinGX, Crystals, …

Image credits: Creative Commons

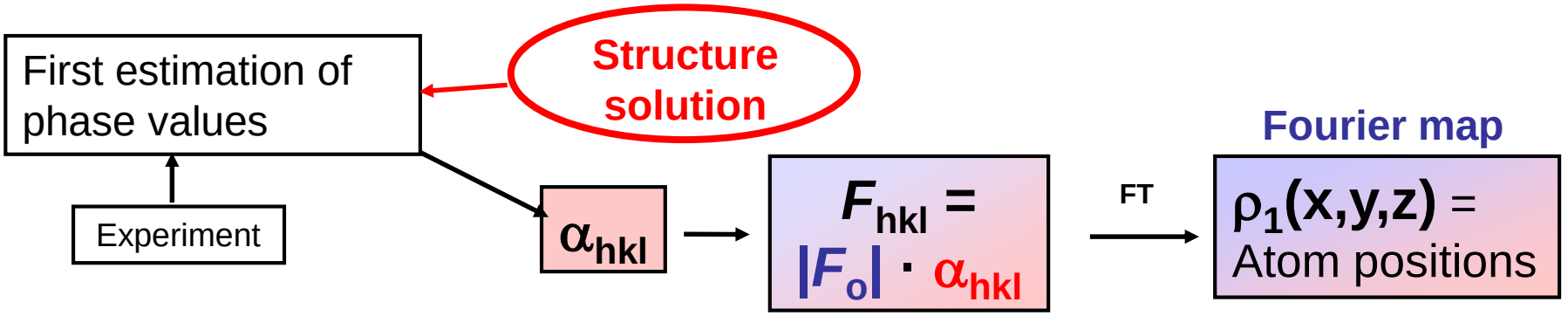
Structure solution : .ins and .res





Attention: in your example, BFS79, there is not SYMM command, since the space group is P-1 (no symmetry, but inversion)

Structure solution : .ins and .res




```
TITL bfs79 in P-1
CELL 0.71073   9.5892  10.8264  11.3959 111.904  96.814 100.727

ZERR   2.00  0.0008  0.0010  0.0009  0.006  0.007  0.007

LATT  1
SYMM -X, 0.5+Y, -Z
SFAC C H B F S
UNIT 44 16 2 30 2
TEMP -43
SIZE 0.1 0.2 0.4

TREF

HKLF 4
END
```

These commands have **no influence** on the actual refinement. But if they are missing or wrong, you end up with wrong information in your CIF. I. e., molecular weight, absorption coefficient, etc.

You should never ever change these commands by hand!

You might change LATT and SYMM… if you really know what you are doing.

Structure solution : .ins and .res

First estimation of phase values

Structure solution

Experiment

α_{hkl}

$F_{hkl} = |F_o| \cdot \alpha_{hkl}$

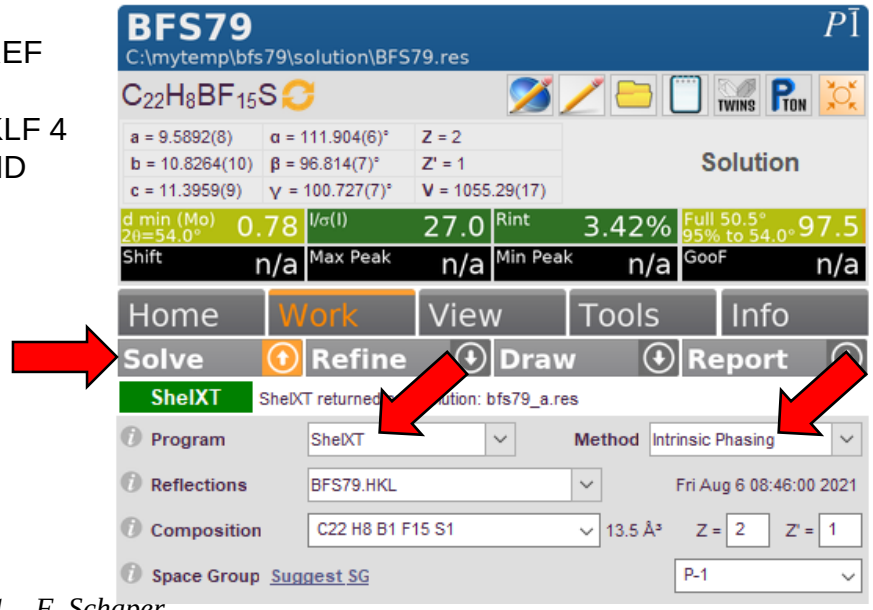
FT

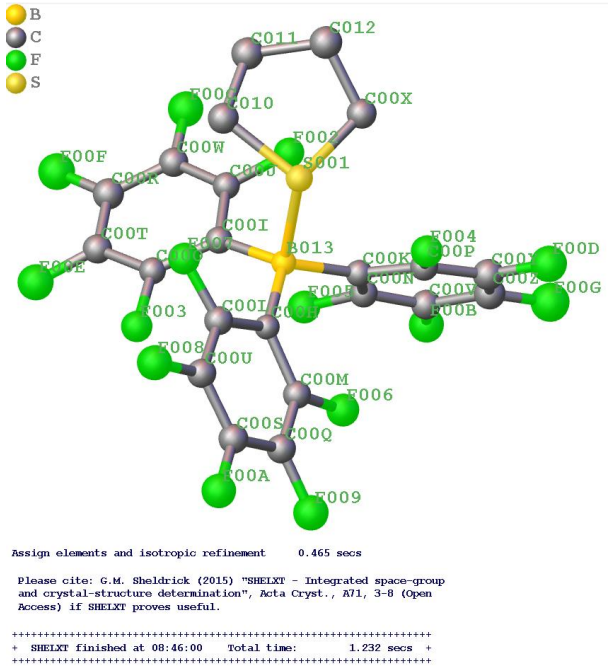
Fourier map

$\rho_1(x,y,z)$ = Atom positions

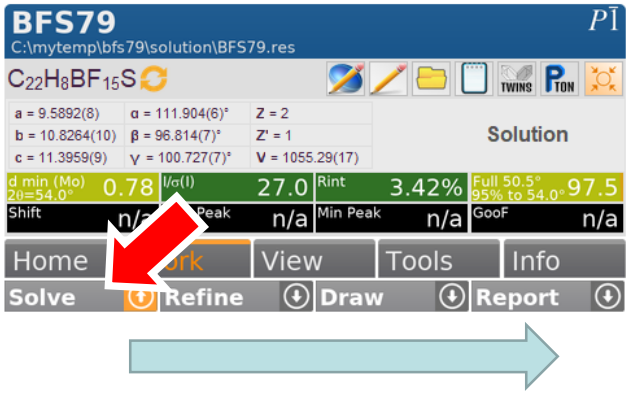
```
TITL bfs79 in P-1
CELL 0.71073   9.5892  10.8264  11.3959 111.904  96.814 100.727
ZERR    2.00   0.0008   0.0010   0.0009   0.006   0.007   0.007
LATT  1
SFAC C H B F S
UNIT 44 16 2 30 2
TEMP -43
SIZE 0.1 0.2 0.4

TREF

HKLF 4
END
```

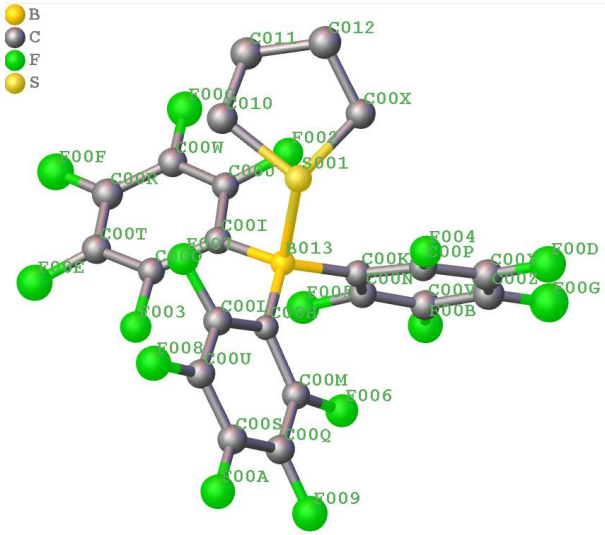
Right now we do not have any structural information!
We know only cell dimensions and space group.





A look behind the GUI





TITL bfs79 in P-1
CELL 0.71073 9.5892 10.8264 11.3959 111.904 96.814 100.727
ZERR 2.00 0.0008 0.0010 0.0009 0.006 0.007 0.007
LATT 1
SFAC C H B F S
UNIT 44 16 2 30 2
TEMP -43
SIZE 0.1 0.2 0.4

TREF

HKLF 4
END

******.INS**

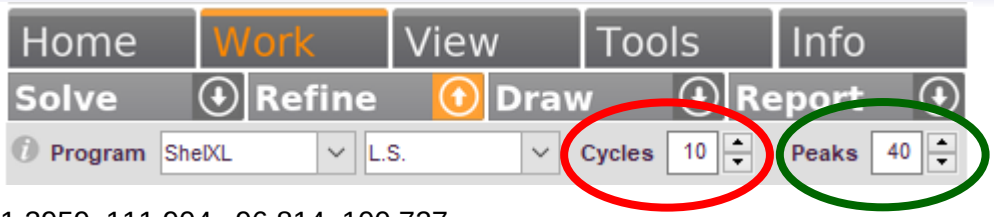
SHELXT

TITL bfs79_a.res in P-1
CELL 0.71073 9.5892 10.8264 11.3959 111.904 96.814 100.727
ZERR 2.000 0.0008 0.0010 0.0009 0.006 0.007 0.007
LATT 1
SFAC C H B F S
UNIT 44 16 2 30 2
TEMP -43
SIZE 0.1 0.2 0.4
L.S. 10
BOND
LIST 6
FMAP 2
PLAN 20
S001 5 0.72244 0.27584 0.78132 11.00000 0.02851 16.56
F002 4 0.44727 0.42232 0.86239 11.00000 0.03521 9.16
F003 4 0.43235 0.19947 0.41550 11.00000 0.03777 8.92
[…]
B013 3 0.63372 0.37475 0.67629 11.00000 0.02391 5.24
HKLF 4
END

******.RES**

atom positions

Our .res / .ins now contains refinement commands




```
TITL bfs79_a.res in P-1
CELL  0.71073   9.5892  10.8264  11.3959  111.904   96.814  100.727
ZERR    2.000   0.0008   0.0010   0.0009    0.006    0.007    0.007
LATT  1
SFAC C H B F S
UNIT 44 16 2 30 2
TEMP -43
SIZE 0.1 0.2 0.4
L.S. 10
BOND
LIST 6
FMAP 2

PLAN 40
WGHT    0.100000

FVAR       0.46325

S001  5  0.72244  0.27584  0.78132 11.00000  0.02851

F002  4  0.44727  0.42232  0.86239 11.00000  0.03521
F003  4  0.43235  0.19947  0.41550 11.00000  0.03777
[…]
B013  3  0.63372  0.37475  0.67629 11.00000  0.02391

HKLF 4
END
```

Elements present in the unit cell

Include distances and angles in .lst/.cif

How many refinement cycles

What type of output (lst, fcf) to produce (see SHELXL manual

What type of Fourier map to produce. No need to change this ever.

Number of electron densities (Q peaks) in difference Fourier map

Weighing scheme

Variables to refine Variable #1 is the OSF (overall scale factor)

Atom label

Atom position

Occupation factor

U_{iso} value of isotropic movement

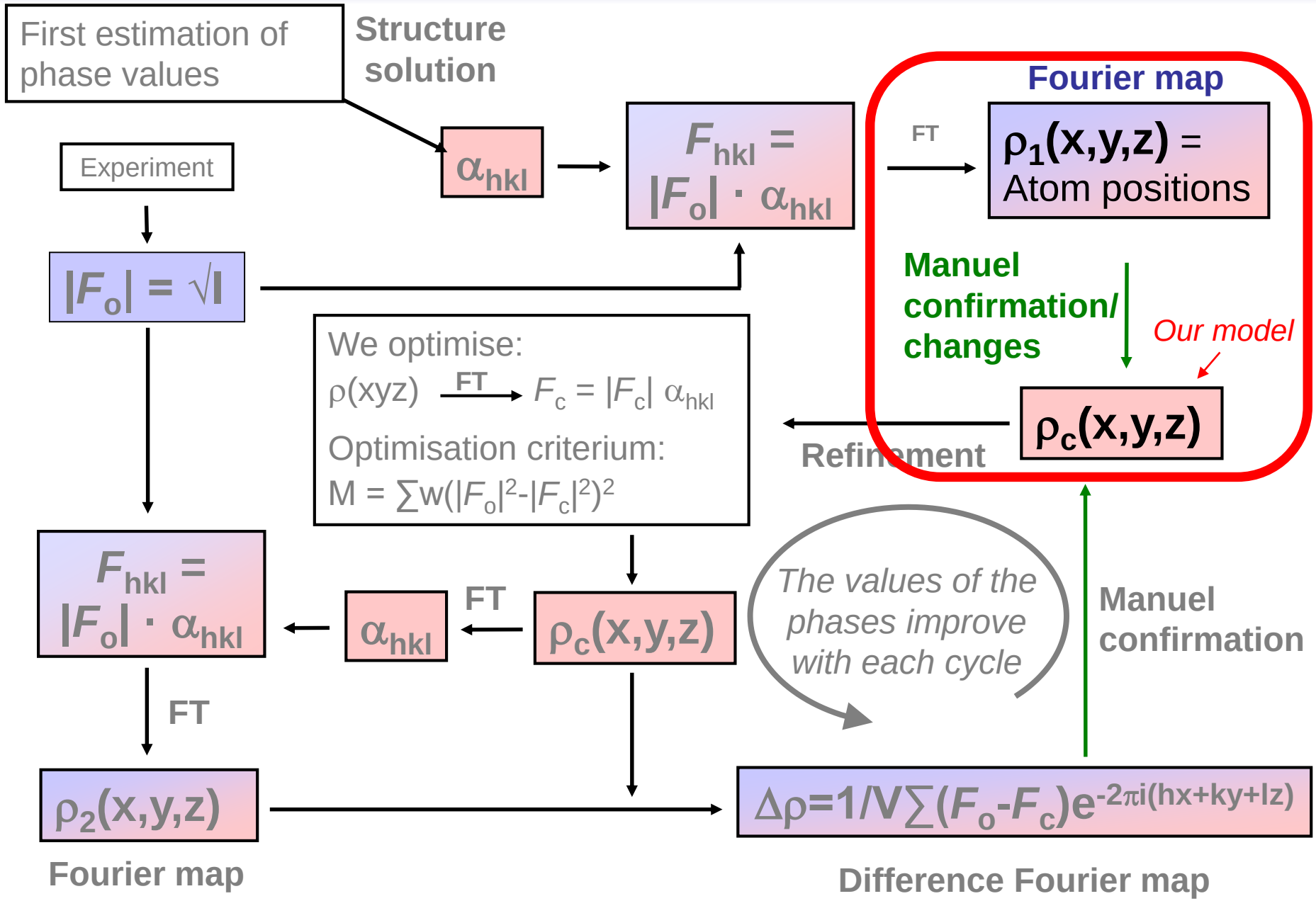
Which element? (number in SFAC)

Very important!

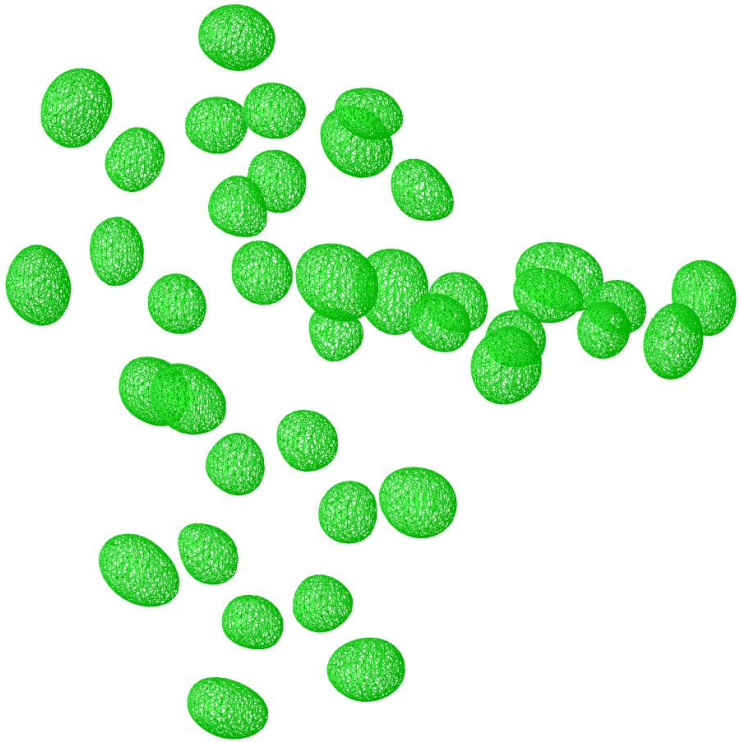
S001 **1** … is a carbon atom

C13 **5** … is a sulfur atom

Defining our model



We only work with models

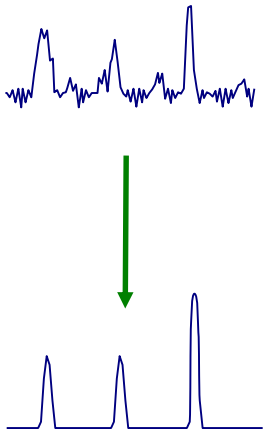


FT → $\rho_1(x,y,z)$ = Atomic coordinates

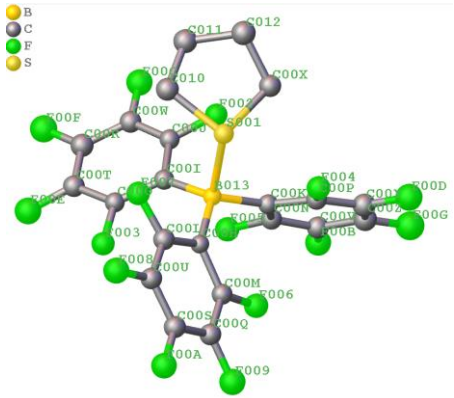
Manual confirmation ↓

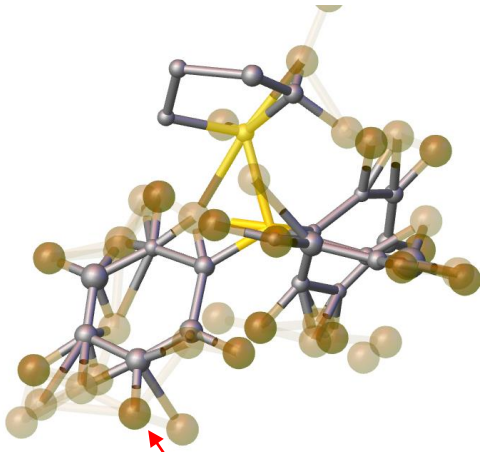
$\rho_c(x,y,z)$

Electron density map = Fourier map



Our model: .res, .ins, cif



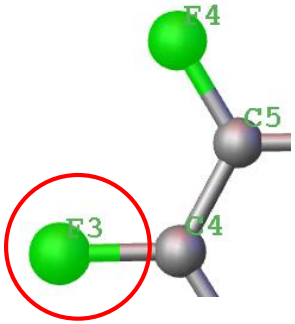


Q-peak

We never really work with the electron density map (calculated Fourier map)! We only see the position of the **maxima** in the electron density map (= Q peaks) and the placed atoms of our model (in *.res or *.ins).
(CTRL-Q if you do not see them)

OLEX and the ins file

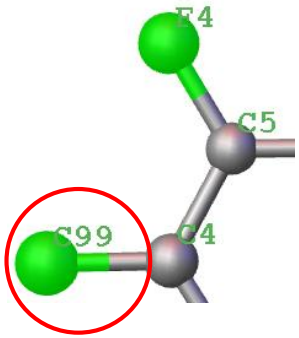
We can edit the ins file directly from OLEX2 via « edit ins ». When we save & close the ins file, the OLEX2 updates the structure. If we do changes in OLEX2, it updates the ins file.

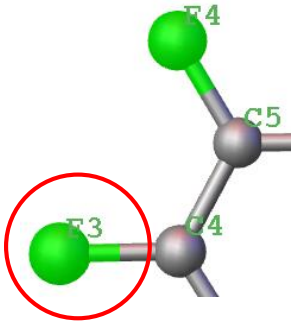


F3 4 0.01098 0.15455 0.57775 11.00000 0.05193

Atom type 4 identifies the atom as a fluorine atom. Changing the label to C99 does not change this.

C99 4 0.01098 0.15455 0.57775 11.00000 0.05193

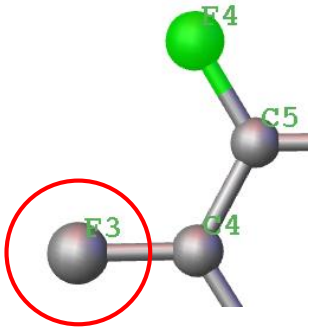


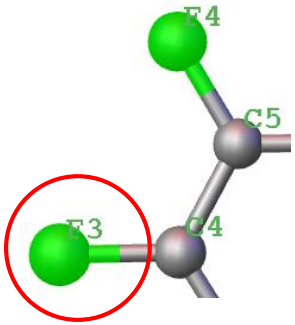


F3 **4** 0.01098 0.15455 0.57775 11.00000 0.05193

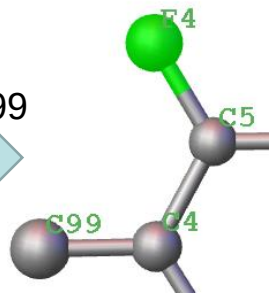
Changing the atom type to **1** makes this a carbon atom, even if the label does not change.

F3 **1** 0.01098 0.15455 0.57775 11.00000 0.05193



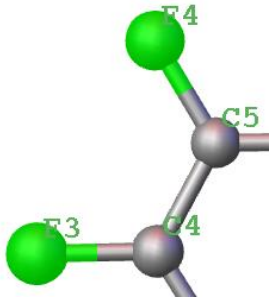


OLEX2:
Name F3 C99

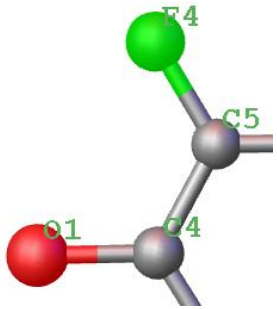


F3 **4** […]

C99 **1** […]



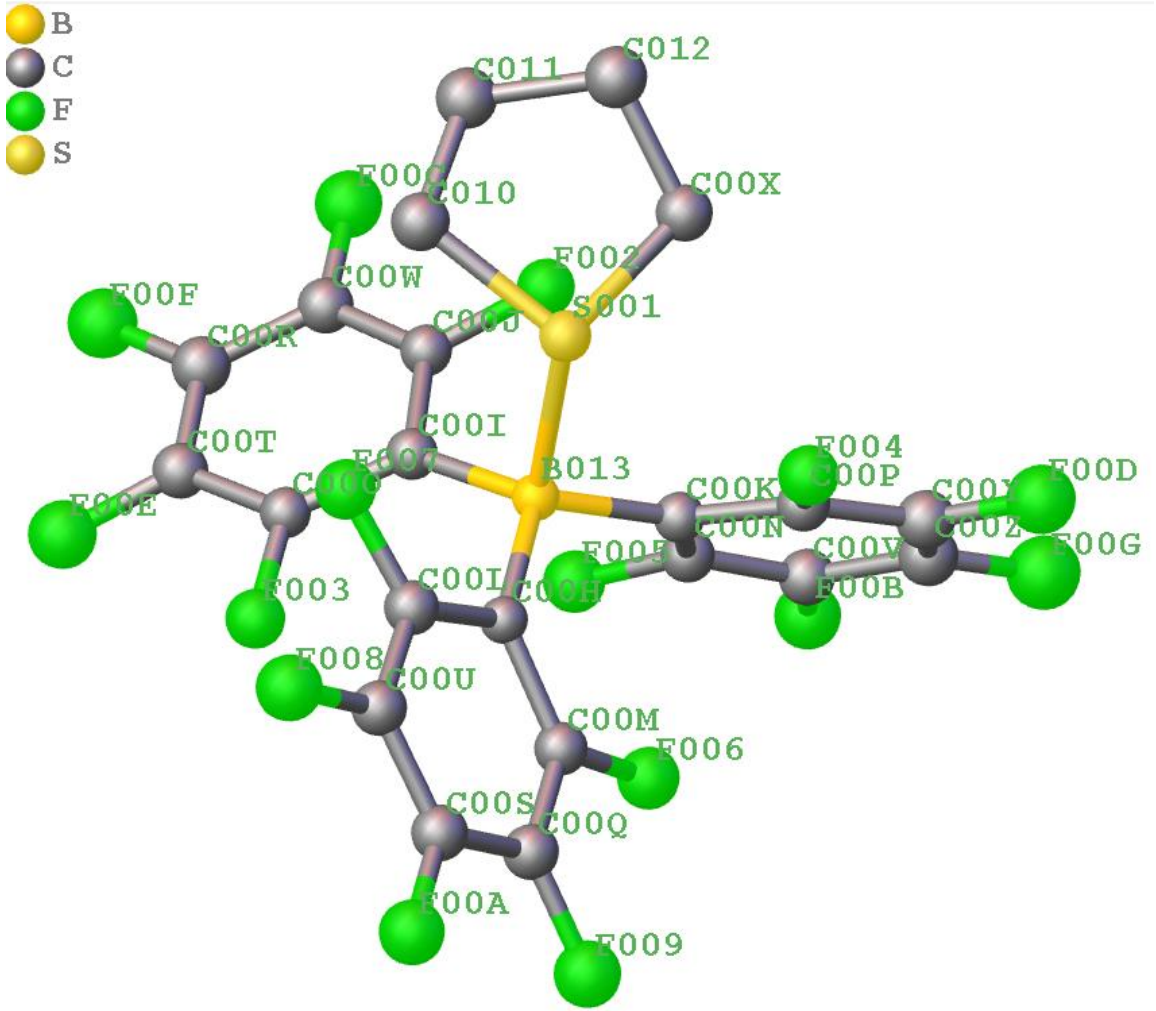
OLEX2:
Name F3 O1

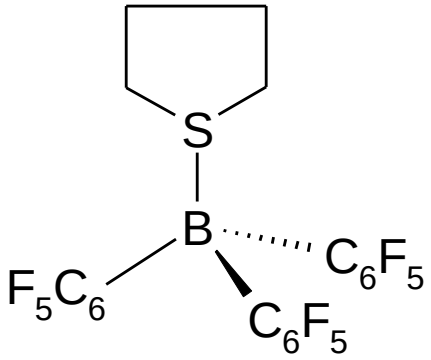


SFAC C H B F S
UNIT 44 16 2 30 2
[…]
F3 **4** 0.01098 […]

SFAC C H B F S **O**
UNIT 44 16 2 30 2 **1**
[…]
O1 **6** 0.01098 […]

Back to our model and start the refinement

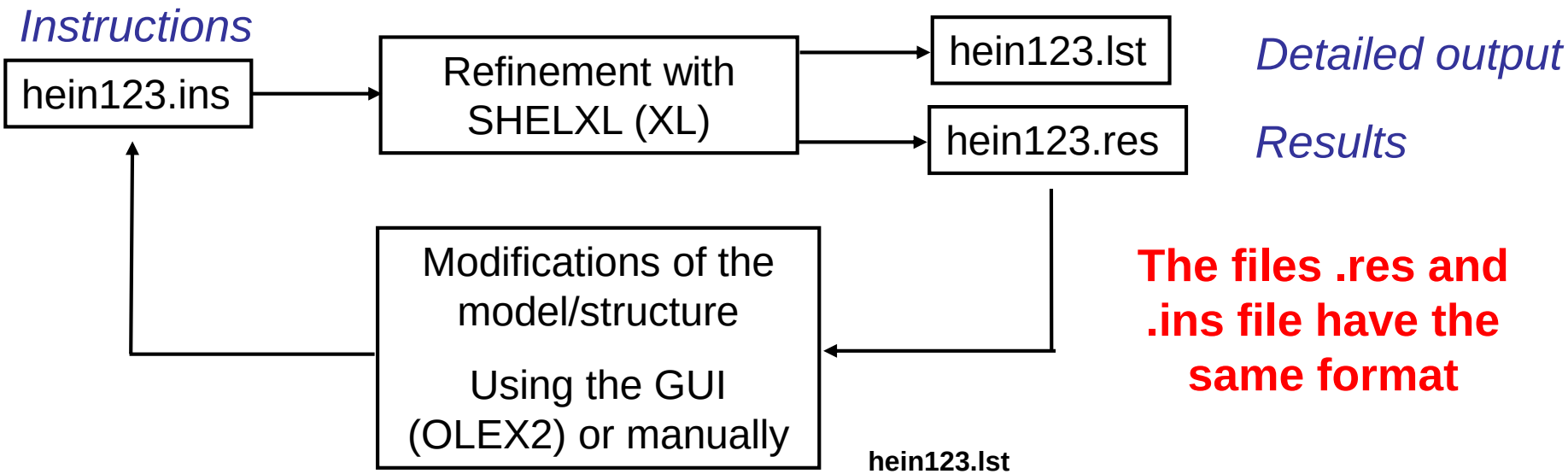




This was a very decent dataset and a simple structure (no complications). SHELXT was able to find all atoms and to assign them correctly.

Let's **refine** our model !

Refinement – ***.ins/***.res and ***.lst



The files .res and .ins file have the same format

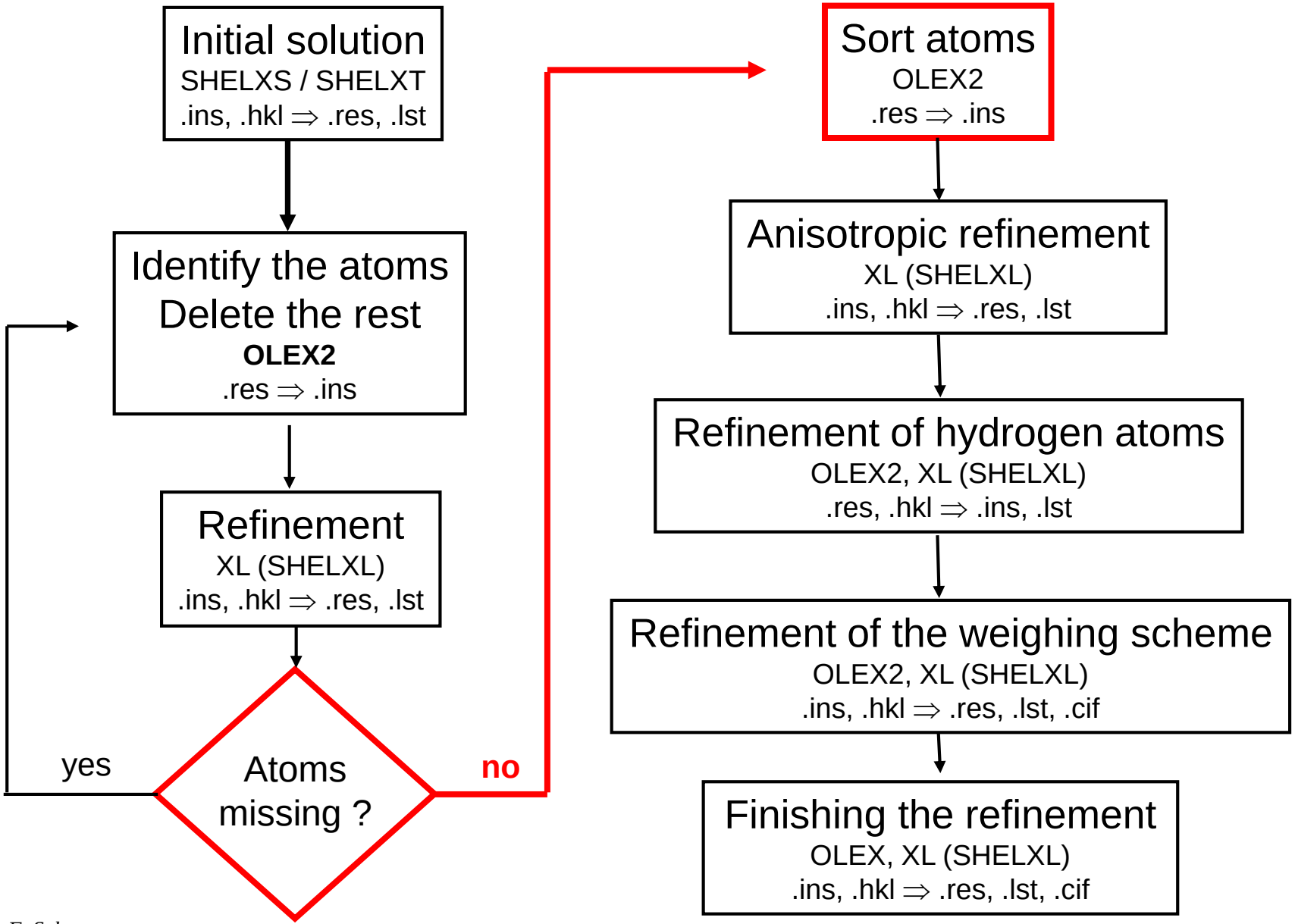
hein123.res/hein123.ins

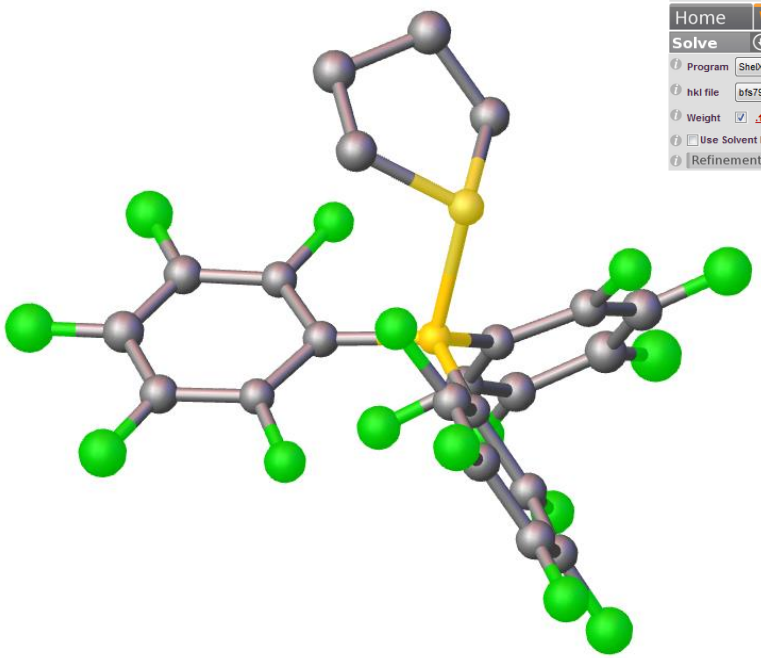
```
TITL Hein123
CELL 0.71073   8.1380  15.4444  15.1323  90.000  98.922  90.000
ZERR    2.00   0.0016   0.0031   0.0030   0.000   0.030   0.000
LATT  1
SYMM -X, 0.5+Y, -Z
SFAC C H N SI ZR
UNIT 70 108 4 8 4
L.S. 6
WGHT    0.100000
FVAR       0.09895
ZR1   5   0.42161   0.25000   0.91998  10.50000   0.02140
SI1   4   0.16560   0.14717   0.71608  11.00000   0.03452
N1    3   0.16921   0.25000   0.66987  10.50000   0.02912
C1    1   0.34492   0.13306   0.80705  11.00000   0.02231
…
HKLF 4
END
```

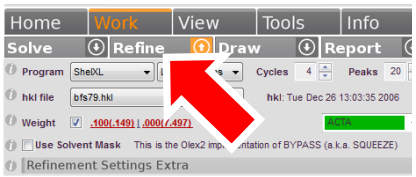
hein123.lst

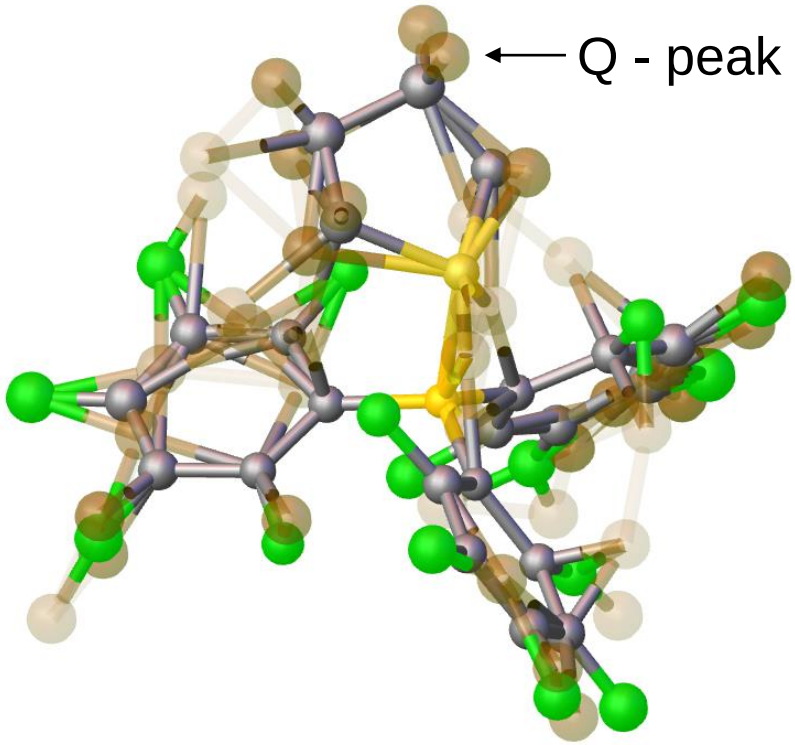
```
 50091  Reflections read, of which 26152  rejected

     39  Systematic absence violations
      0  Inconsistent equivalents
   4013  Unique reflections, of which      0  suppressed
R(int) = 0.0363     R(sigma) = 0.0207

Least-squares cycle   1
 wR2 =  0.4060 before cycle   1 for   4013 data and   189 /
189 parameters,  GooF = S =     7.187;     Restrained
GooF =      7.187  for      0 restraints

N      value       esd    shift/esd  parameter
    1     0.36649     0.00352  -210.419    OSF
    2     0.00205     0.00032     2.885   EXTI
    5     0.02660     0.00068     3.314    U22 Zr1
 Mean shift/esd =   1.629    Maximum =-210.419 for  OSF
 Max. shift = 0.029 A for C25      Max. dU = 0.003 for C14
```

Organigram of the refinement









BFS79.ins

TITL bfs79_a.res in P-1
[…]

S1	5	0.72250	0.2			
B1	3	0.63358	0.3			
F1	4	0.43229	0.1			
F2	4	0.15239	0.10107	0.37643	11.00000	0.04742
F3	4	0.01098	0.15455	0.57775	11.00000	0.05193
F4	4	0.16550	0.31956	0.82027	11.00000	0.04730
F5	4	0.44702	0.42274	0.86238	11.00000	0.03428
F6	4	0.77260	0.55856	0.55811	11.00000	0.04160
F7	4	0.89230	0.48432	0.35458	11.00000	0.04796
F8	4	0.90576	0.22039	0.23223	11.00000	0.04544

[…]

What is actually optimized?

Why do we still see Q peaks if all atoms are found?

BFS79.res

TITL bfs79_a.res in P-1
[…]

					11.00000	0.02810
					11.00000	0.02444
					11.00000	0.03683
F2	4	0.152391	0.101074	0.376430	11.00000	0.04742
F3	4	0.010979	0.154550	0.577751	11.00000	0.05193

[…]
END

Q1	1	0.7707	0.9637	0.9808	11.00000	0.05	1.65
Q2	1	0.7235	0.8144	0.8991	11.00000	0.05	1.10
Q3	1	0.6140	0.8724	0.8755	11.00000	0.05	1.10

[…]

Which parameters are refined ?

File .ins:

CELL 0.71073 8.1380 15.4444 15.1323 90 98.922 90

Si1 4 **0.16560 0.14717 0.71608** 11.00000 0.03452

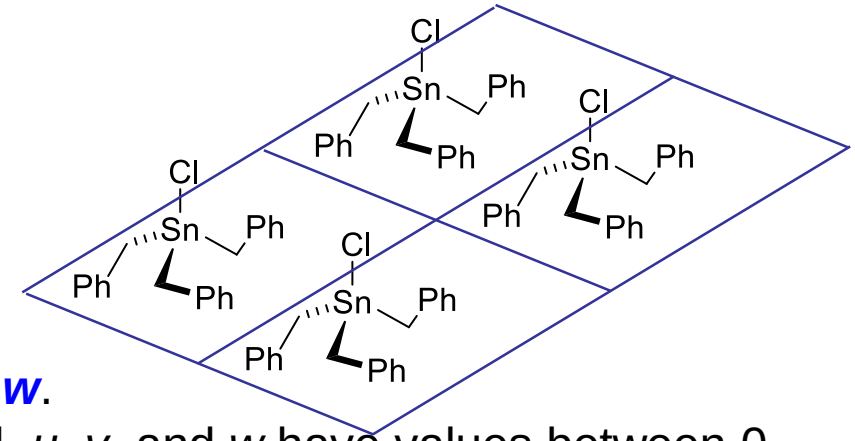
Atomic position

The atomic positions are provided as **fractional coordinates relative to the crystal system**, i. e. the unit cell dimensions. Thus, the values **are not in Å.** For example, the atome Si1 is found at :

x* = *u*·*a = 0.16560·8.1380 Å = 1.3476 Å

y* = *v*·*b = 0.14717·15.4444 Å = 2.2729 Å

z* = *w*·*c = 0.71608·15.1323 Å = 10.836 Å



The symbols for the **fractional atomic positions** are typically the letters ***u***, ***v***, and ***w***.
Given that each atoms is inside the unit cell, *u*, *v*, and *w* have values between 0 and 1. For practical reasons (connected set of coordinates), we find sometimes values <0 or >1, **but there is no justification for values <–1 or >2**.

Which parameters are refined ?

Sl1	4	0.16560	0.14717	0.71608	11.00000	0.03452

Occupation factor

The occupation factor indicates how many atoms occupy this position (or better: which percentage of these positions is occupied by an atom). Typically, this value is 1. It is not possible to have more than one atom per position, but it is possible to have smaller values.

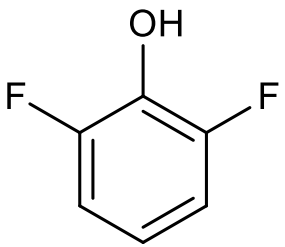
Disorder :

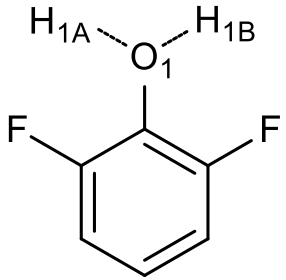
*The **1** indicates that the occupation factor is not free to refine (= fixed).*

Freely refining the occupation factor is (normally) nonsense.

***.ins:

O1	3	0.12560	0.23453	0.83456	11.00000	0.02932
H1A	2	0.12864	0.23364	0.80236	10.50000	0.04732
H2A	2	0.12853	0.23923	0.85935	10.50000	0.04593





You can freeze any variable by adding a 1 in front.

Which parameters are refined ?

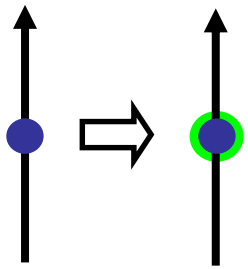
Sl1 4 0.16560 0.14717 0.71608 **11.00000** 0.03452

Occupation factor

The occupation factor indicates how many atoms occupy this position (or better: which percentage of these positions is occupied by an atom). Typically, this value is 1. It is not possible to have more than one atom per position, but it is possible to have smaller values.

Disorder :

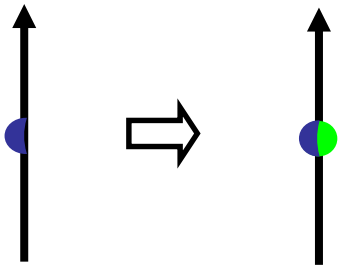
Special positions :



2 / C_2

An atom on a symmetry element, for example on a 2 axis, would be duplicated by the symmetry operation on the same position. To avoid this, the program would need to check for each atom whether it is present on a symmetry element or not.

Sl1 4 0.25000 0.14717 0.25000 1**0.50000** 0.03452



By applying an occupation factor of ½, the atom can be treated in the same way as other atoms in the unit cell.

Thermal movement / The temperature factor

		Atomic position			Occupation factor	U_{iso}
Sl1	4	0.16560	0.14717	0.71608	11.00000	0.03452

U_{iso} ?

- The spatial distribution of the electron density is influenced by the thermal motion of the atoms.
- Thermal motion is not identical (in amplitude and direction) for all atoms.
- The timescale of a diffraction experiment is much longer than thermal motion.

$\Rightarrow$ **We obtain a distribution of electron density, averaged over time and all unit cells**

Isotropic motion:

- The motion (vibration) of an atom is identical in all directions.
- It can be described by a Gaussian distribution:

$$\rho(r') = \sqrt{2\pi U}\, e^{\frac{-r'^2}{2U}}$$

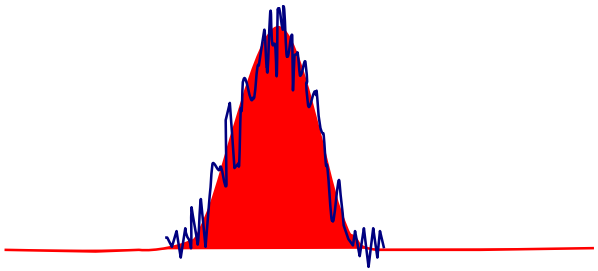
ρ : electron density

r' : distance of the atom from its equilibrium position

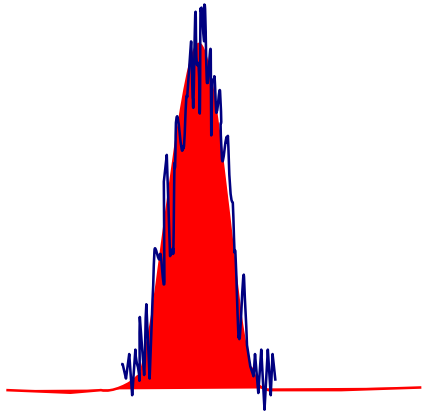
$U = \langle r'^2 \rangle$ Average of the squared displacement [Å²]

$B = 8\pi^2 U$ Atomic temperature factor (Debye-Waller factor) [Å²]

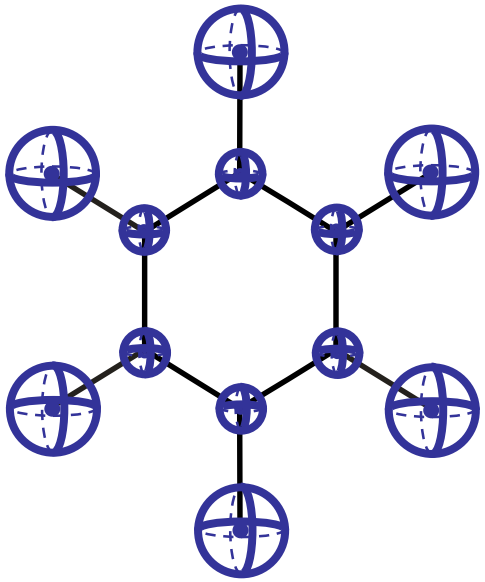
Thermal movement / The temperature factor

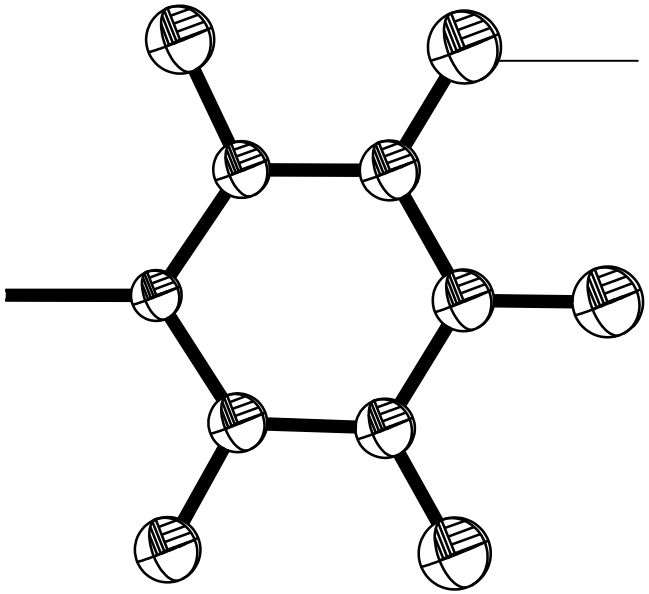


U_{iso} = 0.05 Å²

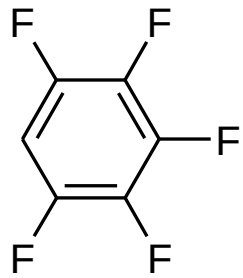


U_{iso} = 0.03 Å²



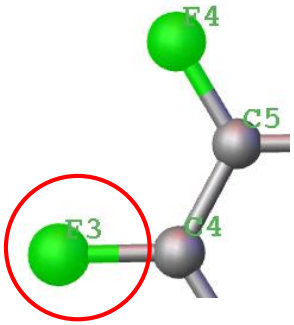


Probability of x% (typically 50%) that the atom is found inside a sphere with the radius indicated here.



The wrong atom type and U_{iso}

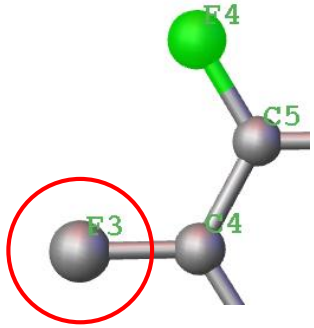
We can edit the ins file directly from OLEX2 via « edit ins ».
When we save & close the ins file, the OLEX2 updates the structure.

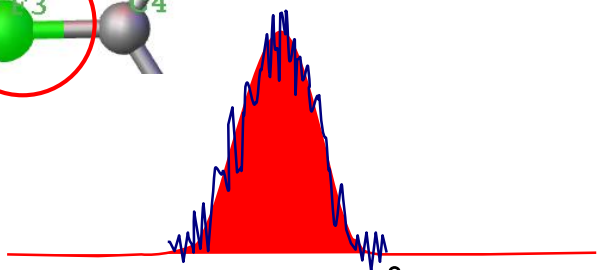


F3 **4** 0.01098 0.15455 0.57775 11.00000 0.05193

Changing the atom type to **1** makes this a carbon atom, even if the label does not change.

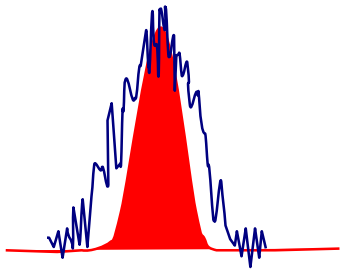
F3 **1** 0.01098 0.15455 0.57775 11.00000 **0.05193**



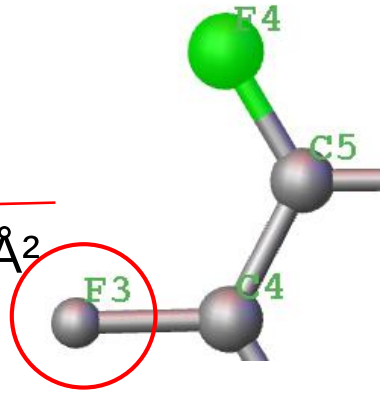


U_{iso} = 0.05 Å²

Refinement:



U_{iso} = 0.02 Å²



The isotropic parameter shrinks, since our model now has less electrons at this position.

F3 1 0.011198 0.154210 0.577322 11.00000 **0.02123**

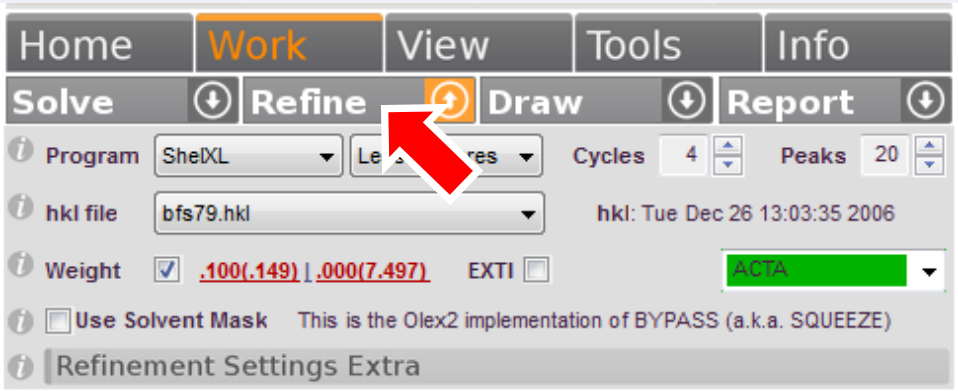
Structure refinement

We optimise:

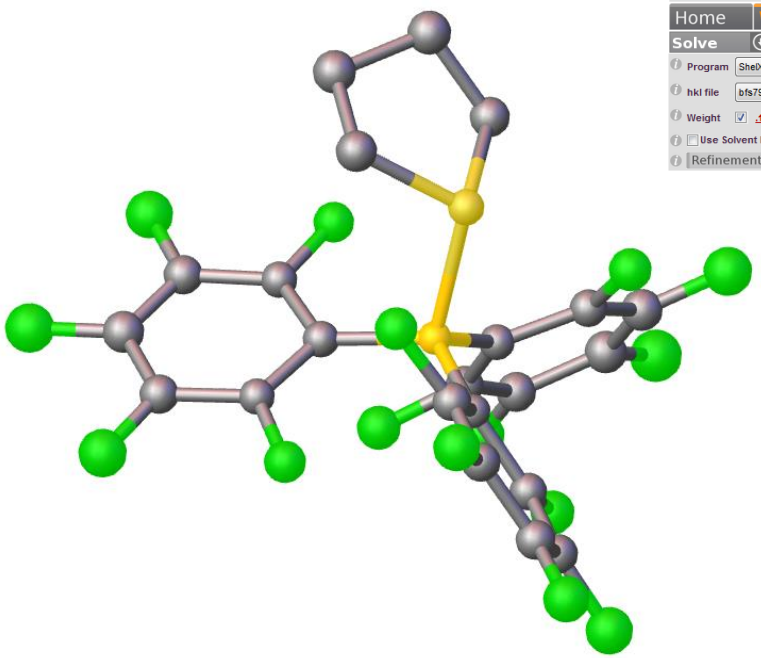
$\rho(xyz) \xrightarrow{FT} F_c = |F_c| \alpha_{hkl}$

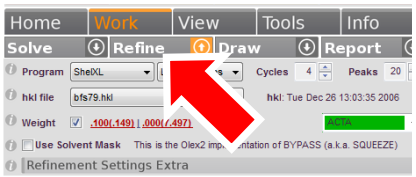
Optimization criterium:

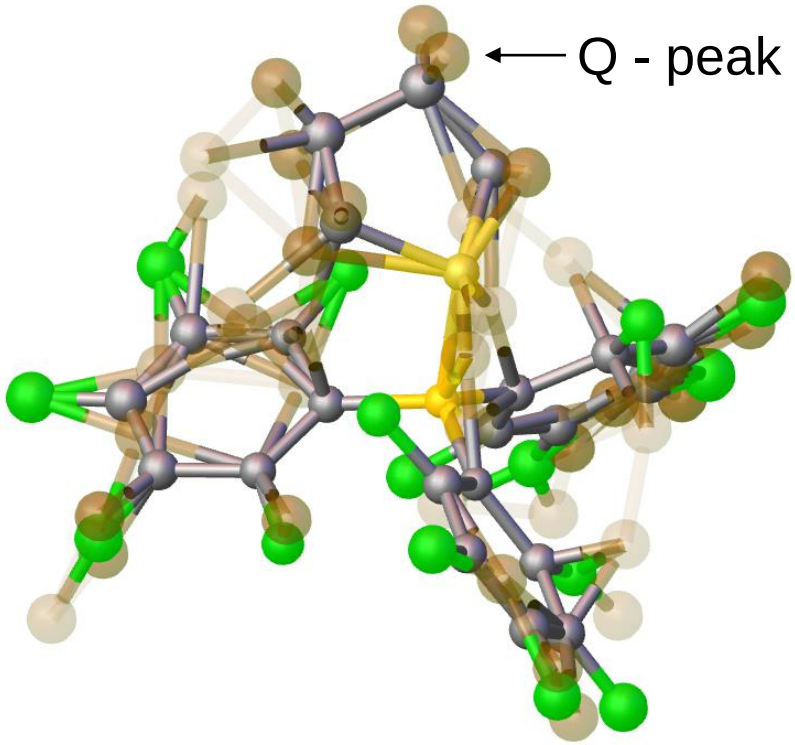
$M = \sum w(|F_o|^2 - |F_c|^2)^2$




```
Max. shift =  0.027 A for C9          Max. dU =-0.011 for B1
wR2 =  0.4010  before cycle   2 for    4359 data and     157 /    157 parameters
GooF = S =      1.311;      Restrained GooF =     1.311 for       0 restraints
Mean shift/esd =  1.662   Maximum =   -15.000 for  U11 S1        at 16:45:13
Max. shift =  0.027 A for C6          Max. dU =-0.010 for B1
wR2 =  0.3268  before cycle   3 for    4359 data and     157 /    157 parameters
GooF = S =      1.098;      Restrained GooF =     1.098 for       0 restraints
Mean shift/esd =  1.585   Maximum =   -12.926 for  U11 S1        at 16:45:13
Max. shift =  0.022 A for F9          Max. dU =-0.006 for C7
wR2 =  0.3030  before cycle   4 for    4359 data and     157 /    157 parameters
GooF = S =      1.039;      Restrained GooF =     1.039 for       0 restraints
Mean shift/esd =  0.217   Maximum =     1.508 for  U11 F3        at 16:45:13
Max. shift =  0.006 A for F9          Max. dU = 0.002 for F3
wR2 =  0.3021  before cycle   5 for    4359 data and       0 /    157 parameters
GooF = S =      1.034;      Restrained GooF =     1.034 for       0 restraints
wR2 = 0.3021, GooF = S = 1.034, Restrained GooF = 1.034 for all data
R1 = 0.1145 for 3937 Fo > 4sig(Fo) and 0.1221 for all 4359 data
    0  atoms may be split and     0  atoms NPD
R1 =  0.1219 for   4359 unique reflections after merging for Fourier
Highest peak    1.48  at  0.2336  0.0348  0.5215  [  0.61 A from F3 ]
Deepest hole   -1.69  at  0.1966  0.1020  0.5312  [  0.59 A from F3 ]
```







BFS79.ins

TITL bfs79_a.res in P-1
[…]

S1	5	0.72250	0.27589	0.78126	11.00000	0.02810
B1	3	0.63358	0.3			
F1	4	0.43229	0.1			
F2	4	0.15239	0.10107	0.37643	11.00000	0.04742
F3	4	0.01098	0.15455	0.57775	11.00000	0.05193
F4	4	0.16550	0.31956	0.82027	11.00000	0.04730
F5	4	0.44702	0.42274	0.86238	11.00000	0.03428
F6	4	0.77260	0.55856	0.55811	11.00000	0.04160
F7	4	0.89230	0.48432	0.35458	11.00000	0.04796
F8	4	0.90576	0.22039	0.23223	11.00000	0.04544

[…]

Why do we still see Q peaks if all atoms are found?

BFS79.res

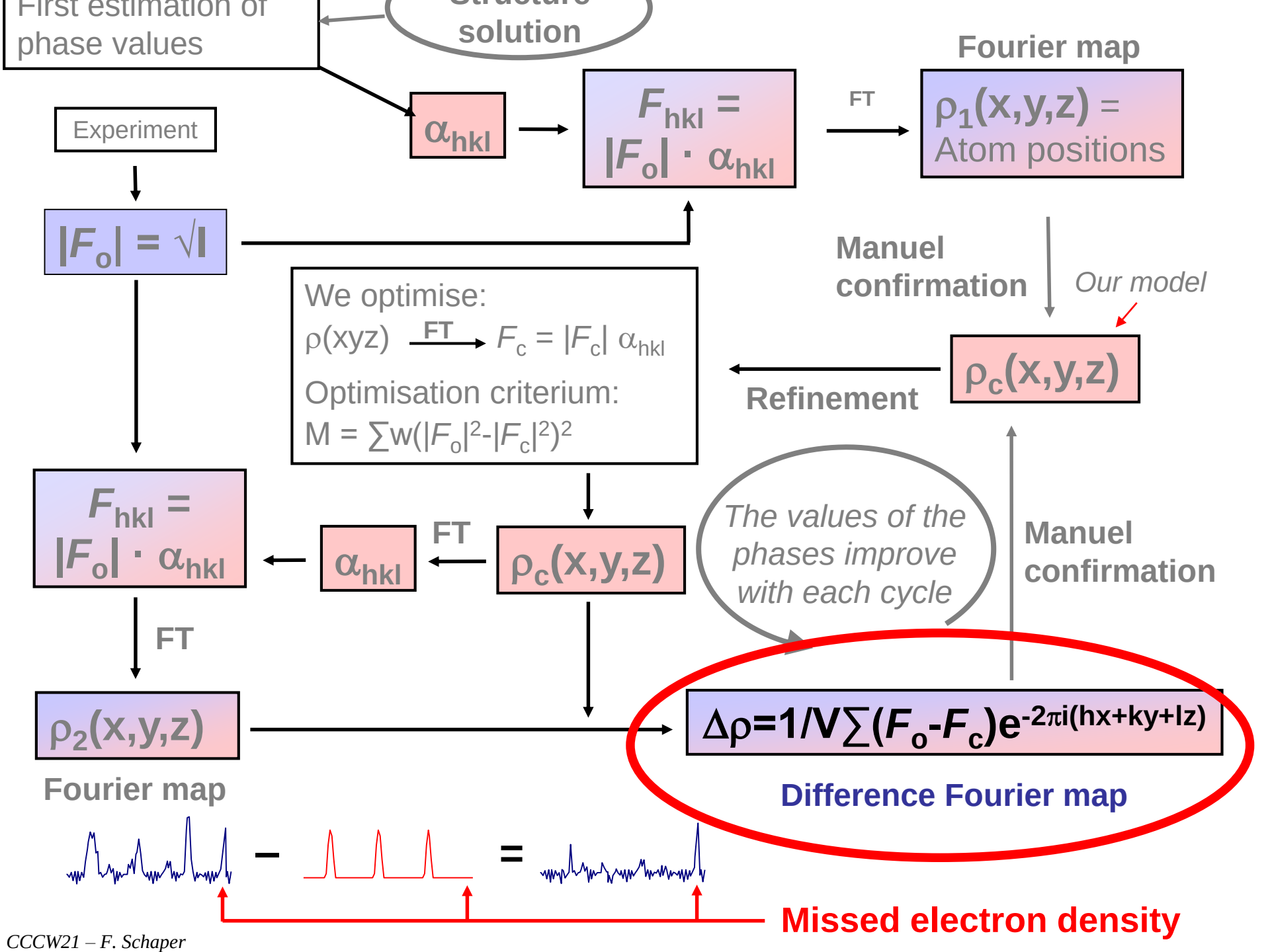
TITL bfs79_a.res in P-1
[…]

S1	5	0.722498	0.275894	0.781256	11.00000	0.02810
					11.00000	0.02444
					11.00000	0.03683
F2	4	0.152391	0.101074	0.376430	11.00000	0.04742
F3	4	0.010979	0.154550	0.577751	11.00000	0.05193

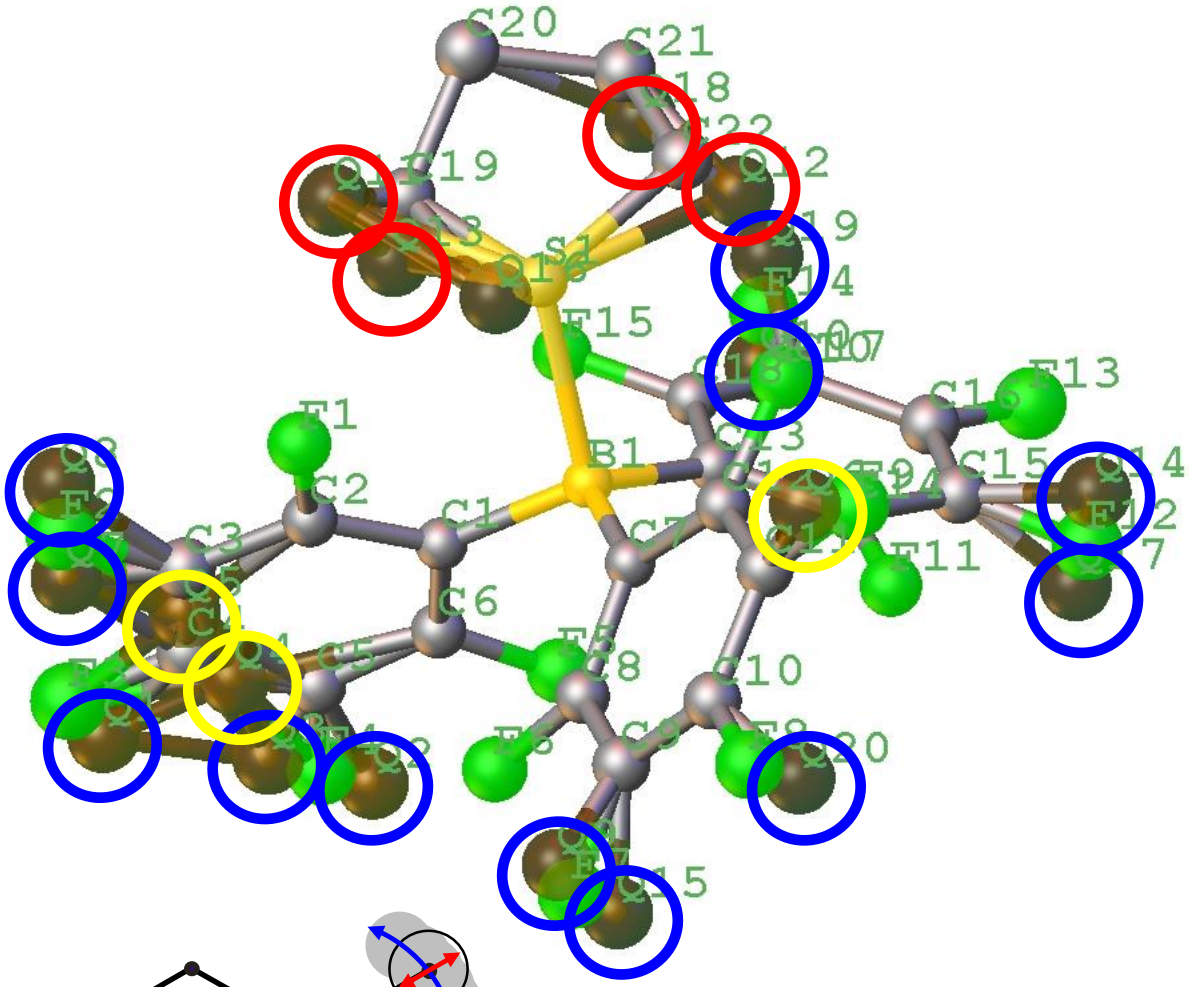
[…]
END

Q1	1	0.7707	0.9637	0.9808	11.00000	0.05	1.65
Q2	1	0.7235	0.8144	0.8991	11.00000	0.05	1.10
Q3	1	0.6140	0.8724	0.8755	11.00000	0.05	1.10

[…]


First estimation of phase values
Structure solution
Fourier map
Experiment
α_{hkl}
$F_{hkl} = |F_o| \cdot \alpha_{hkl}$
FT
$\rho_1(x,y,z)$ = Atom positions
$|F_o| = \sqrt{I}$
Manuel confirmation
Our model
We optimise:
$\rho(xyz) \xrightarrow{FT} F_c = |F_c|\ \alpha_{hkl}$
Optimisation criterium:
$M = \sum w(|F_o|^2 - |F_c|^2)^2$
$\rho_c(x,y,z)$
Refinement
$F_{hkl} = |F_o| \cdot \alpha_{hkl}$
α_{hkl}
FT
$\rho_c(x,y,z)$
The values of the phases improve with each cycle
Manuel confirmation
FT
$\rho_2(x,y,z)$
$\Delta\rho = 1/V \sum (F_o - F_c) e^{-2\pi i(hx+ky+lz)}$
Fourier map
Difference Fourier map
–
=
Missed electron density

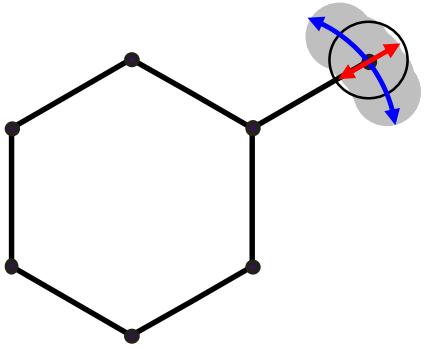
Remaining maxima in the difference Fourier map

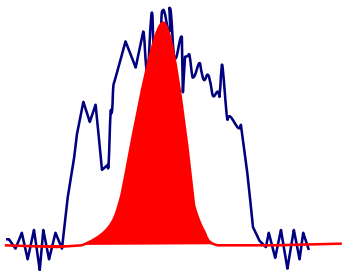


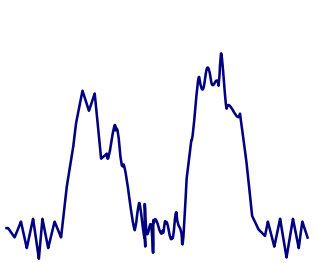
All atoms of the structure have been found and remaining maxima are not any longer associated with additional atoms.

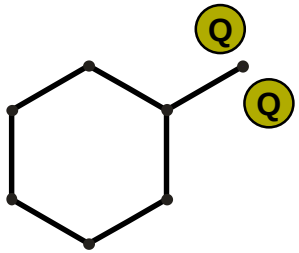
The remaining maxima of electron density are due to (in decreasing order of intensity) :

- **Thermal movement of atoms**
- **Hydrogen atoms**
- **Noise and artefacts**







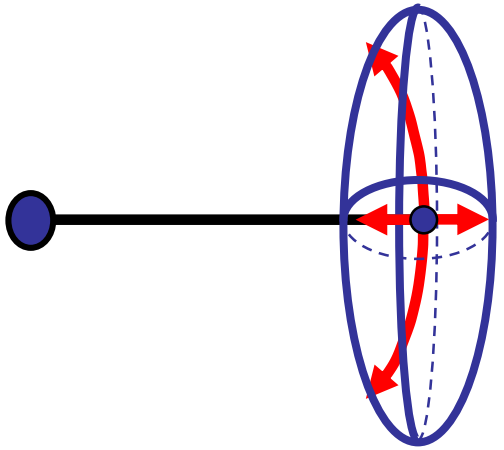


Thermal movement / The temperature factor

Anisotropic thermal motion :

- In reality, thermal movement of atoms is not isotropic. Deformation of bond angles are typically less energetic than stretching vibration parallel to the bond.

  Compare: C-H stretch: 3000 cm^{-1}, bending: 600-1500 cm^{-1}.



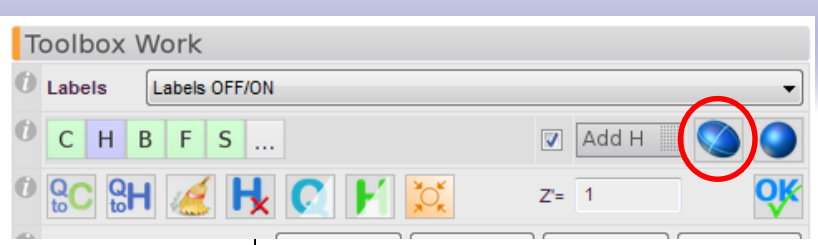
- In our structural model, thermal movement is described by an ellipsoid (three Gaussian functions in three dimensions). The mathematic form is that of a symmetric tensor:

$$U = \begin{pmatrix} \langle x'^2 \rangle & \langle x'y' \rangle & \langle x'z' \rangle \\ \langle x'y' \rangle & \langle y'^2 \rangle & \langle y'z' \rangle \\ \langle x'z' \rangle & \langle y'z' \rangle & \langle z'^2 \rangle \end{pmatrix}$$

$$q(r^*) = e^{-2\pi^2(U_{11}x'^2 + U_{22}y'^2 + U_{33}z'^2 + 2U_{12}x'y' + 2U_{13}x'z' + 2U_{23}y'z')}$$

- Instead of **1 parameter (U_{iso})**, we now work with **6 parameters** (three values for the three diameters of the ellipsoid, three for its orientation in space).
- Anisotropic refinement thus becomes challenging if the ratio of data/parameters is low.

Anisotropic refinement



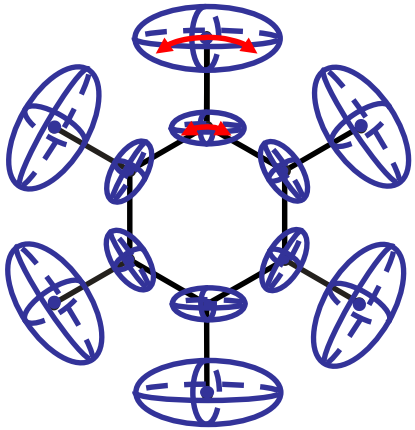
OLEX2

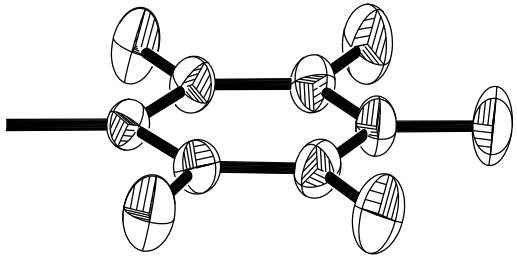
***.ins:**

```
ANIS
WGHT     0.100000
FVAR        0.71880
S1     5    0.277586     0.724064     0.718741     11.00000     0.02858
B1     3    0.366643     0.625301     0.823380     11.00000     0.02507
F1     4    0.079477     0.474487     0.653317     11.00000     0.03965
```

***.res:**

```
WGHT     0.100000
FVAR        0.73527
S1     5    0.277751     0.724339     0.718759     11.00000     0.03304     0.02854 =
          0.02460     0.01056     0.00663     0.01163
B1     3    0.366302     0.625285     0.823440     11.00000     0.02772     0.02171 =
          0.02300     0.00834     0.00485     0.00625
F1     4    0.079086     0.473686     0.653072     11.00000     0.02849     0.03982 =
          0.04679     0.01366     0.00319     0.00667
```





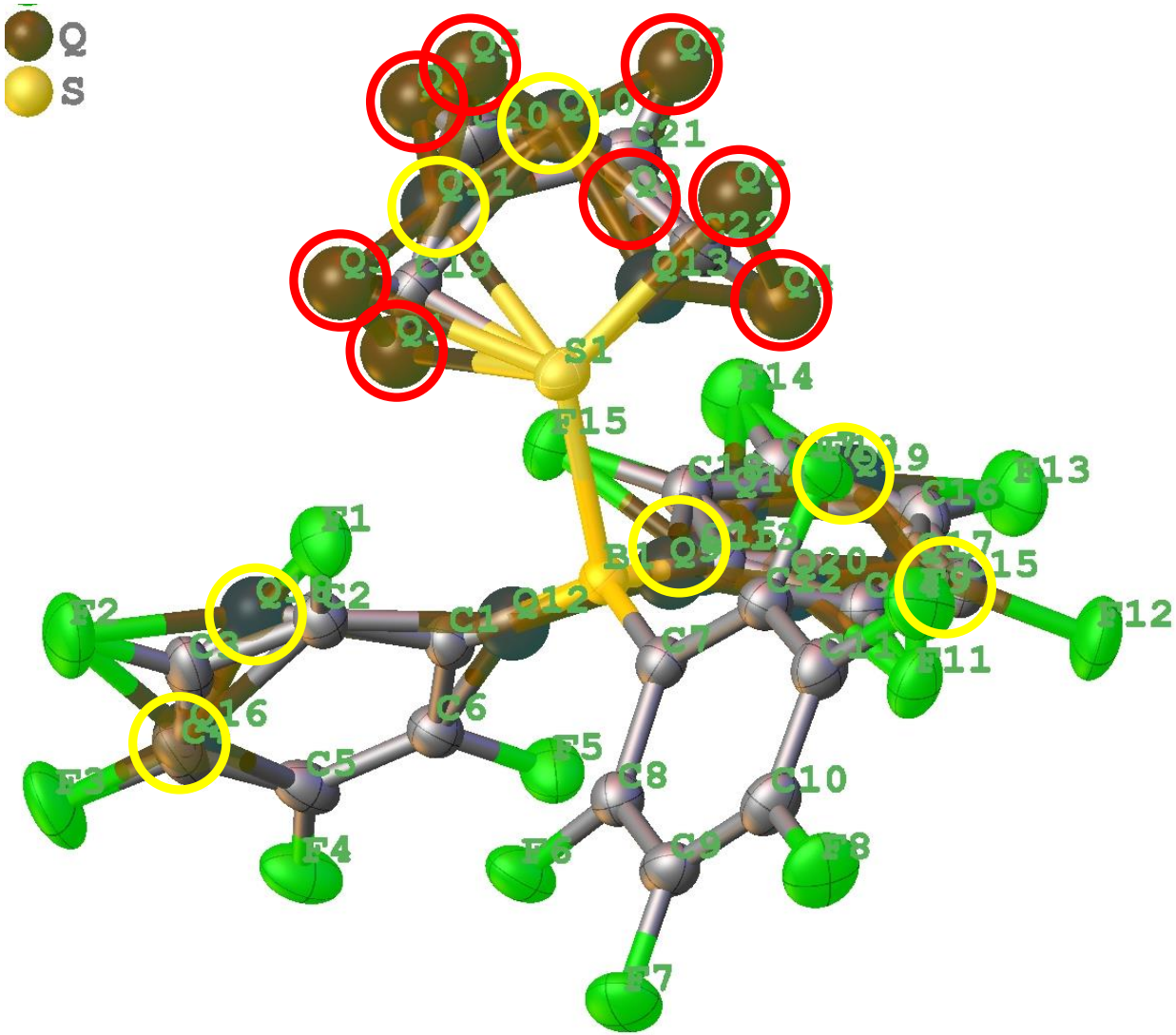
U_{eq} is the equivalent isotropic value, calculated from :

$$U_{eq} = \frac{1}{3}\sum_{i=1}^{3}\sum_{j=1}^{3} U_{ij} a_i^* a_j^* \overrightarrow{a_i} \cdot \overrightarrow{a_j} = \sqrt[3]{U_1 U_2 U_3}$$

Typical values for U_{eq}:

Heavy atoms:	0.005-0.02 Å^2
Light atoms (H-F):	0.01-0.06 Å^2
Terminal atoms:	0.03-0.2 Å^2

Difference Fourier map after anisotropic refinement

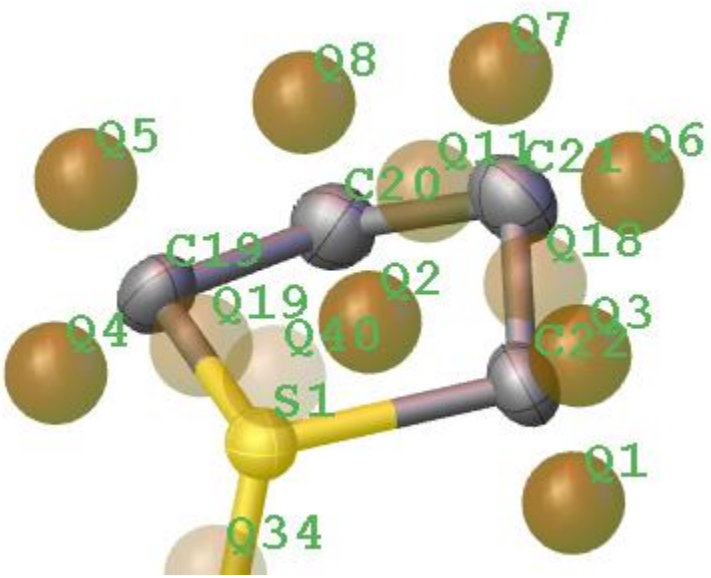


After anisotropic refinement, the remaining maxima can be attributed to:

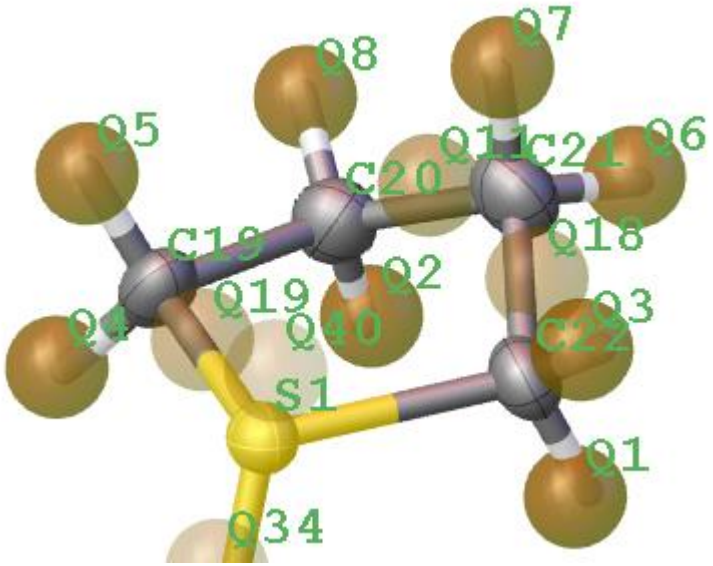
- ~~**Thermal movement of atoms**~~
- **Hydrogen atoms**
- **Noise and artefacts**

The 8 most intense maxima (Q peaks with the lowest numbers, i. e. Q1-Q8) are found where we would expect hydrogen atoms!

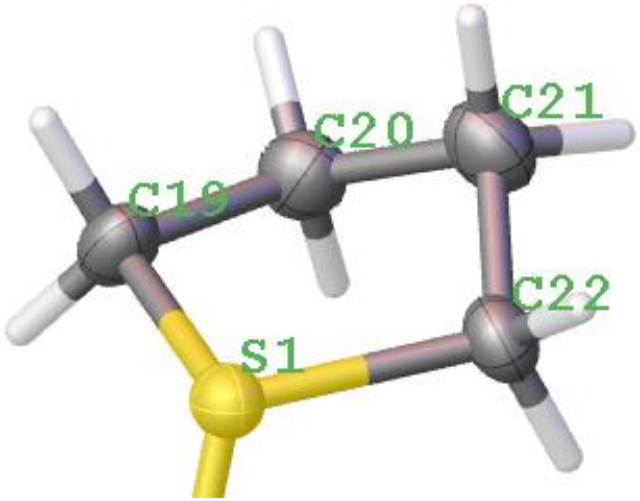
Adding hydrogen atoms



OLEX2:
HADD



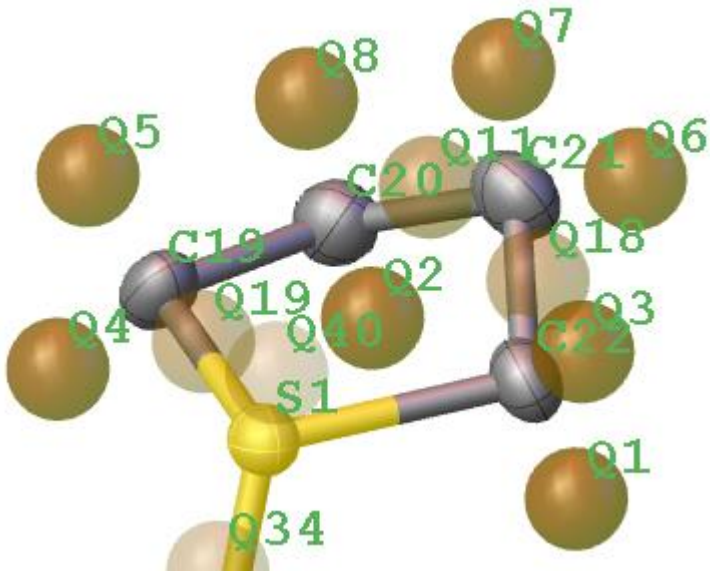
Attention:
HADD places hydrogen atoms in calculated positions. This data set is good enough that the observed maxima actually agree with the calculated positions.



What happens in the ins file ?

```
C19   1    0.58371  0.12711  0.76751  11.00000  0.05447  0.02329  0.03258 =
 0.01158  0.01211  0.00636
C20   1    0.53585  0.16502  0.89601  11.00000  0.04641  0.03599  0.03512 =
 0.01616  0.01325  0.01003
```

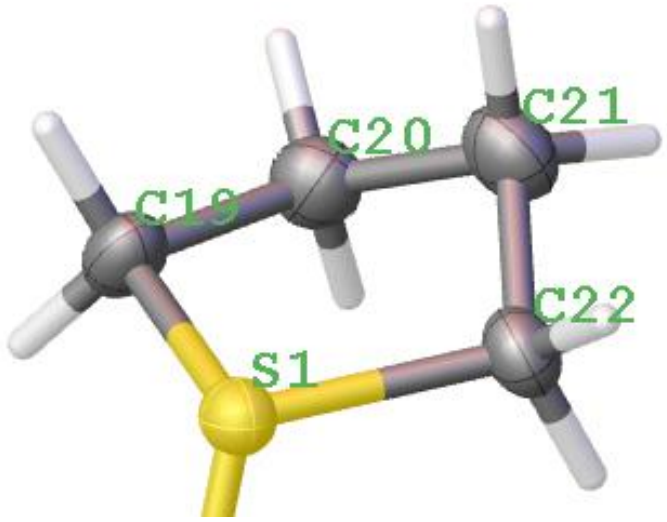
OLEX2:
HADD




```
C19   1    0.58371  0.12711  0.76751  11.00000  0.05447  0.02329  0.03258 =
 0.01158  0.01211  0.00636
AFIX 23
```

AFIX indicates the start of a « rigid group ». The internal geometry of a rigid group does not change.

```
H19a  2    0.50108  0.10640  0.69784  11.00000 -1.20000

H19b  2    0.62427  0.04684  0.74958  11.00000 -1.20000
AFIX 0

C20   1    0.53585  0.16502  0.89601  11.00000  0.04641  0.03599  0.03512 =
 0.01616  0.01325  0.01003
AFIX 23
H20a  2    0.45720  0.21130  0.89603  11.00000 -1.20000
H20b  2    0.50235  0.08214  0.91123  11.00000 -1.20000
AFIX 0
```



Rigid groups in SHELXL

For non-H atoms:

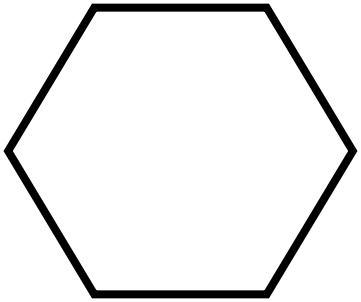
In rigid groups, the atom positions (3*n*) are replaced by the position and orientation (6 variables) of a group of atoms in **an idealized geometry**. The geometry of this group is **invariable** and the atoms cannot move independent during the refinement.

type of rigid group

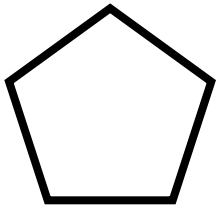
AFIX mn ← *refinement of the rigid group*

AFIX m6 : the group is completely rigid

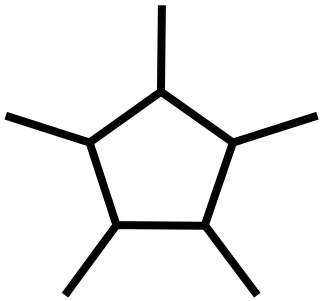
AFIX m9 : the group can expand or contract, but keeps its relative geometry.



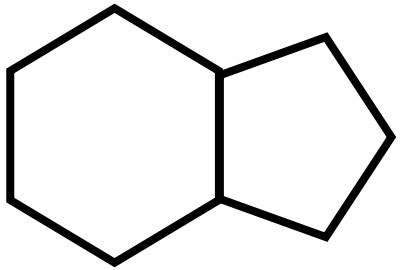
AFIX 66
C1 x y z
:
C6 x y z
AFIX 0



AFIX 56
C1 x y z
:
C5 x y z
AFIX 0



AFIX 106
C1 x y z
:
C10 x y z
AFIX 0

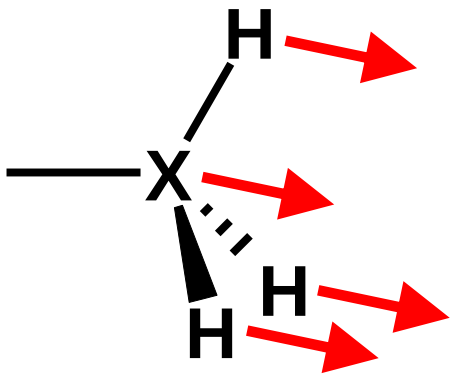


AFIX 116
C1 x y z
:
C11 x y z
AFIX 0

Riding model for hydrogen atoms

The hydrogen atoms “ride” on the atom to which they are bound with a fixed, idealized geometry. During the refinement, the same displacement are applied to the whole group.

$$x_H = x_X + \Delta x$$
$$y_H = y_X + \Delta y$$
$$z_H = z_X + \Delta z$$
$$U_{iso,H} = 1.2/1.5 \cdot U_{eq,X}$$



If hydrogen atoms are refined using the “riding model”, **no additional parameter** is added to the refinement.

Methyl group

- with freely refined hydrogens : 9 (C) + 3·4 (H isotropic) = 21 parameter
- riding model : 9 parameters

Hydrogen atoms: 4. Constraints / “riding model”

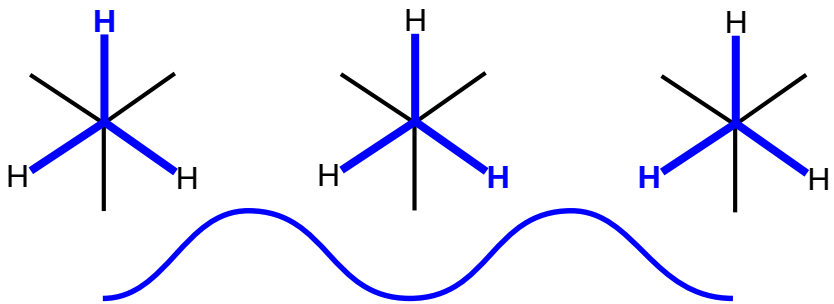
AFIX mn: m = type of the rigid group; n = refinement mode (for more details : see the SHELXL manual)

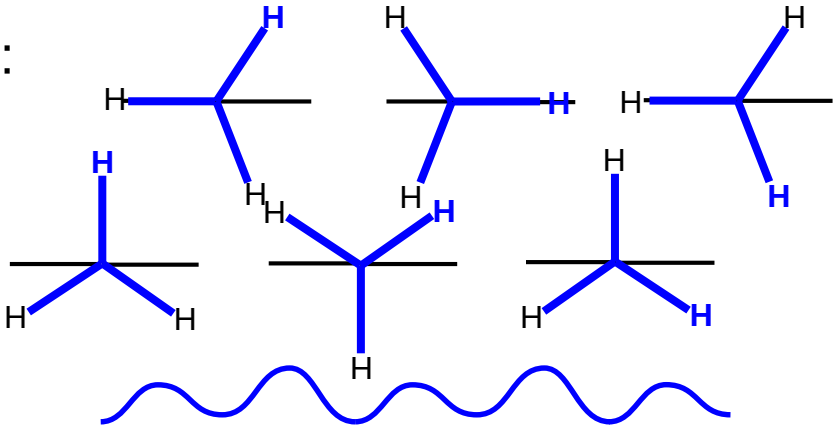
m =
- 1 : Y-**CH**(-Y)$_2$
- 2 : Y-**CH$_2$**-Y
- 3 : Y-**CH$_3$**
- 4 : **C(sp^2/arom.)-H**
- 8 : X-**OH**
- 9 : X=**CH$_2$**
- 12 : Y-**CH$_3$** (désordonné)
- 13 : Y-**CH$_3$** (angle dièdre optimisé)
- 14 : X-**OH** (angle dièdre optimisé)
- 15 : X$_{4/5}$**BH**
- 16 : X≡**CH**

n =
- **3** : No refinement.
- **4** : As n=3, but the C-H distance is refined (same for all C-H of that group.)
- **7** : Only for CH_3: same as n=3, but rotation around the Y-CH_3 is allowed.
- **8** : same as n=7 with refinement of the C-H distance.

We typically use: AFIX 33, **AFIX 23, AFIX 13, AFIX 43, AFIX 137 (or 37)**, AFIX 127

AFIX **37**, 133 or **137** important for C_{sp2}-Me:



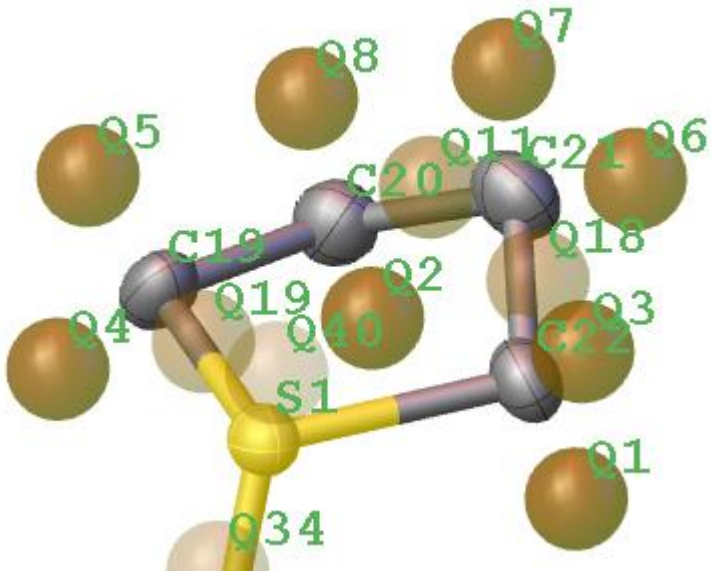


What happens in the ins file ?

```
C19   1    0.58371  0.12711  0.76751  11.00000  0.05447  0.02329  0.03258 =
 0.01158  0.01211  0.00636
C20   1    0.53585  0.16502  0.89601  11.00000  0.04641  0.03599  0.03512 =
 0.01616  0.01325  0.01003
```

*HADD is very convenient, but **always** verify what it did! In particular CH_2 vs. $C_{Ar}H$ errors*

OLEX2: HADD




```
C19   1    0.58371  0.12711  0.76751  11.00000  0.05447  0.02329  0.03258 =
 0.01158  0.01211  0.00636
AFIX 23
H19a  2    0.50108  0.10640  0.69784  11.00000 -1.20000
H19b  2    0.62427  0.04684  0.74958  11.00000 -1.20000
AFIX 0
C20   1    0.53585  0.16502  0.89601  11.00000  0.04641  0.03599  0.03512 =
 0.01616  0.01325  0.01003
AFIX 23
H20a  2    0.45720  0.21130  0.89603  11.00000 -1.20000
H20b  2    0.50235  0.08214  0.91123  11.00000 -1.20000
AFIX 0
```

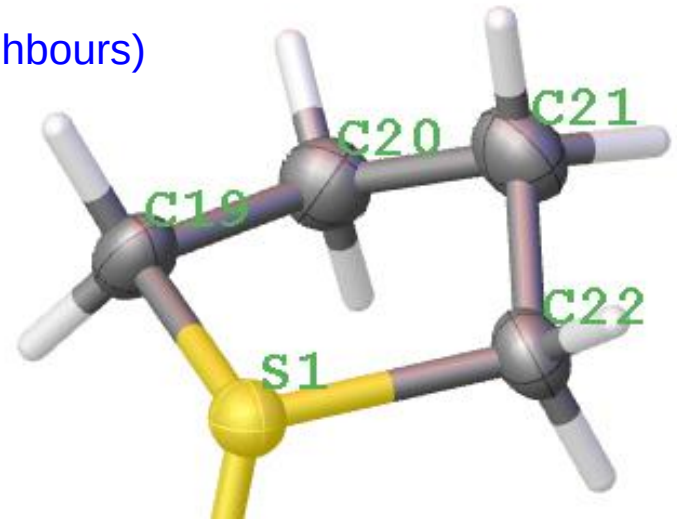
copied from C19

Calculated : $1.2 \cdot U_{eq}$ (C19)

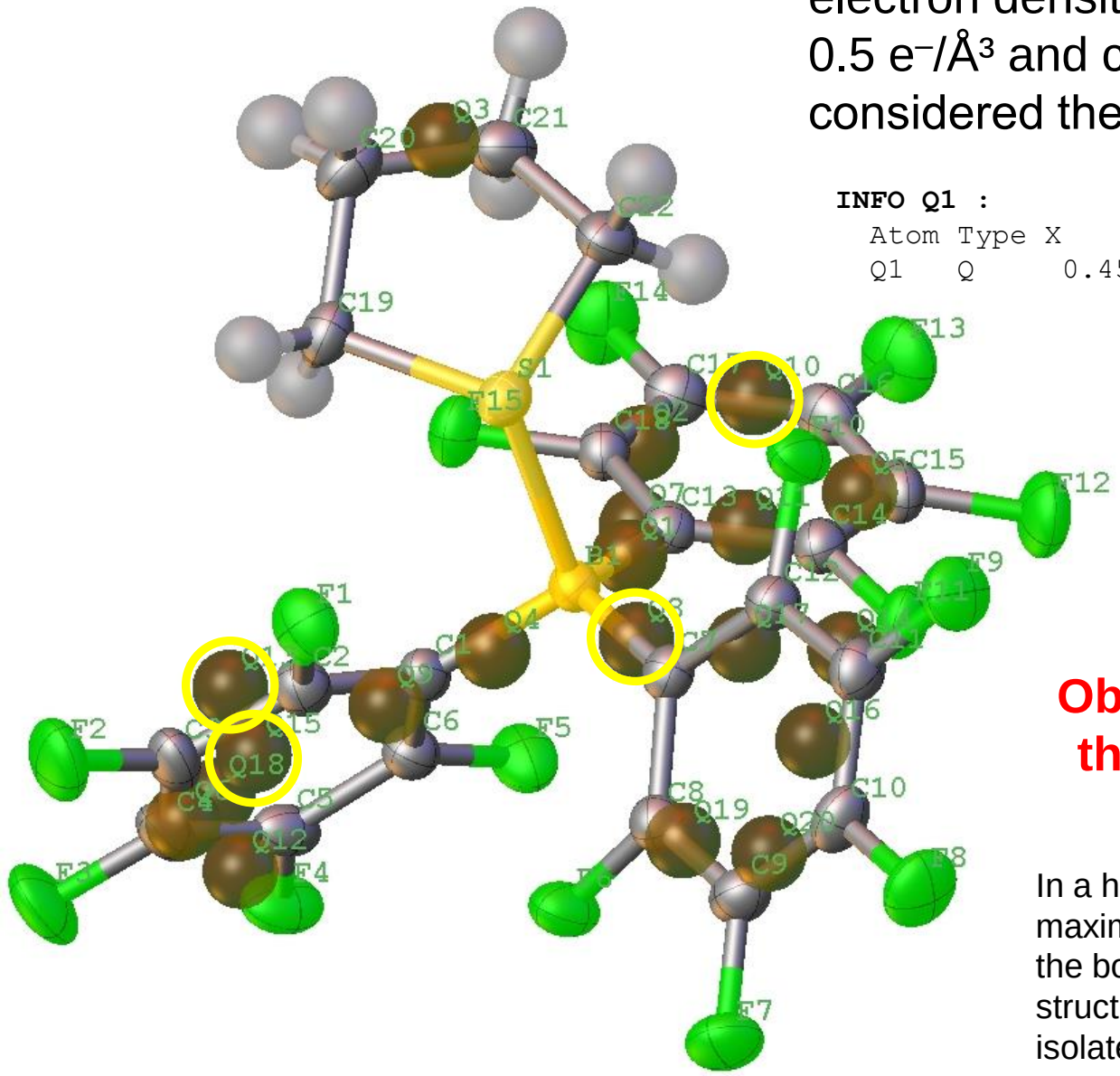
rigid group starts

calculated from C19 (and neighbours)

rigid group ends



And what's left after that…



After refinement of hydrogen atoms, all electron density maxima are smaller than 0.5 $e^-/Å^3$ and can be ignored. (0.5 $e^-/Å^3$ is considered the "allowable noise level.)

INFO Q1 :

Atom	Type	X	Y	Z	Ueq	Peak
Q1	Q	0.458	0.660	0.845	0.050	0.280

- ~~Thermal movement of atoms~~
- ~~Hydrogen atoms~~
- Noise and artefacts

Observe the positions of the maxima, though! Is this really noise?

In a high-quality dataset, the remaining maxima correspond to electron density in the bonds. This is an artefact of our structural model: we assume a model of isolated, non-interacting atoms.

Refinement of the weighing scheme

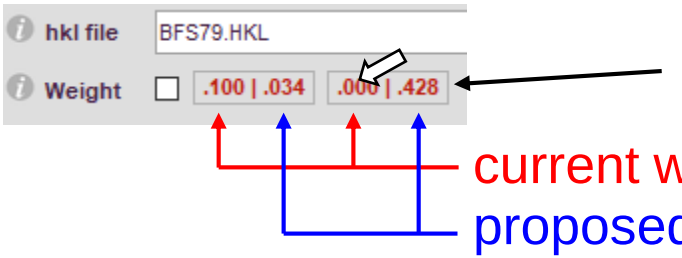
*.lst: *weak reflections* *intense reflections*

Fc/Fc(max)	0.000	0.012	0.023	0.034	0.047	0.061	0.079	0.101	0.135	0.201	1.000
Number in group	463	423	422	475	397	439	438	433	430	439	
GooF	**1.406**	**1.493**	**1.414**	**1.317**	**1.076**	**1.029**	**0.813**	**0.734**	**0.633**	**0.453**	
K	1.690	1.053	1.022	1.020	1.008	1.011	1.012	0.997	1.002	0.990	

We us an appropriate weighing scheme so that the GooF («Goodness of Fit») is independent from the intensity.

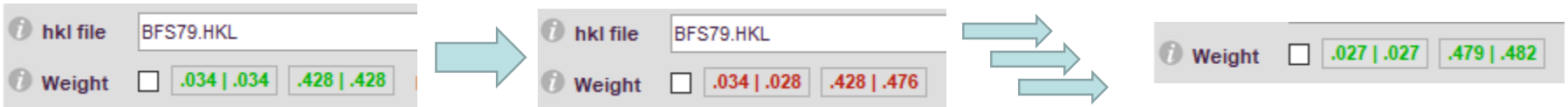
$$w = 1/\sigma^2(F_o^2) + (aP)^2 + bP \qquad P = 1/3\left(F_o^2 + 2F_c^2\right)$$

How do we refine the weighing scheme in OLEX2 :



red color indicates non-refined weighing scheme

Click on the numbers to copy proposed to current weighing scheme



Refine structure

Repeat until consistent

Refinement of the weighing scheme

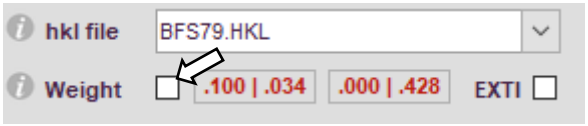
***.lst:** *weak reflections* *intense reflections*

Fc/Fc(max)	0.000	0.012	0.023	0.034	0.047	0.061	0.079	0.101	0.135	0.201	1.000
Number in group	463	423	422	475	397	439	438	433	430	439	
GooF	**1.406**	**1.493**	**1.414**	**1.317**	**1.076**	**1.029**	**0.813**	**0.734**	**0.633**	**0.453**	
K	1.690	1.053	1.022	1.020	1.008	1.011	1.012	0.997	1.002	0.990	

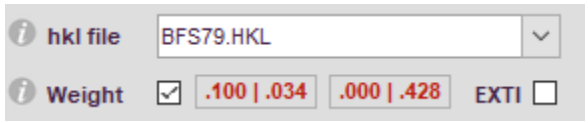
We us an appropriate weighing scheme so that the GooF («Goodness of Fit») is independent from the intensity.

$$w = 1/\sigma^2_{(F_o^2)} + (aP)^2 + bP \qquad P = \tfrac{1}{3}\left(F_o^2 + 2F_c^2\right)$$

Faster :



Activate « Automatic Update » of the weighing scheme after each refinement. (Do that only after you are done with the hydrogen atoms).



Refine until consistent

***.lst:**

Fc/Fc(max)	0.000	0.012	0.023	0.035	0.047	0.062	0.079	0.101	0.135	0.202	1.000
Number in group	466.	418.	459.	438.	419.	417.	442.	435.	427.	438.	
GooF	**1.117**	**1.023**	**1.008**	**1.059**	**0.974**	**1.102**	**0.990**	**1.064**	**1.056**	**0.993**	
K	**1.732**	1.049	1.028	1.018	1.008	1.011	1.013	0.997	1.004	0.992	

What happens behind the GUI ?

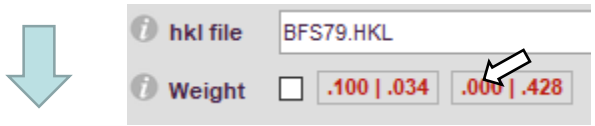
***.res:**

```
…
WGHT    0.1000
FVAR 0.3234
…
HKLF 4
END
WGHT    0.0342    0.4326
```

current weighing scheme

proposed weighing scheme

The end: All lines after END are ignored



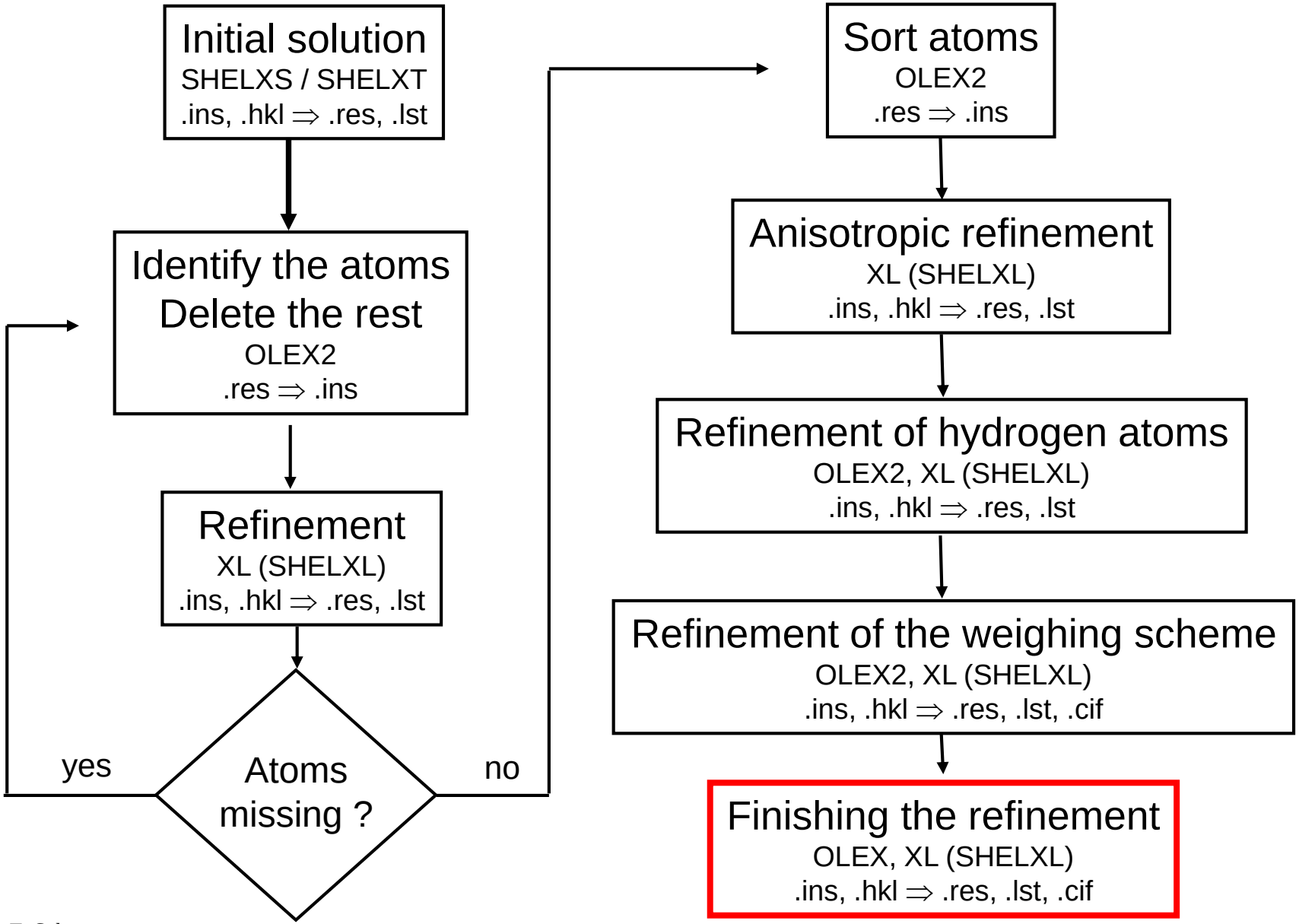
Click on the numbers to copy proposed to current weighing scheme

***.ins:**

```
…
WGHT 0.0342    0.4326
FVAR 0.3234
…
HKLF 4
END
WGHT    0.0342    0.4326
```

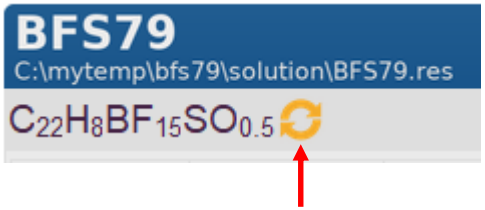
Of course… we *could* do this by hand in the ins file ourselves instead of using OLEX2. ☺

Organigram of the refinement



Finishing the refinement

Make sure that the formula is correct



Cycle indicates disagreement of UNIT instruction and atom list

***.lst:**

…

** Cell contents from UNIT instruction and atom list do not agree **

Unit-cell contents from UNIT instruction and atom list resp.

C	44.00	44.00
[...]		
O	**1.00**	**0.00**

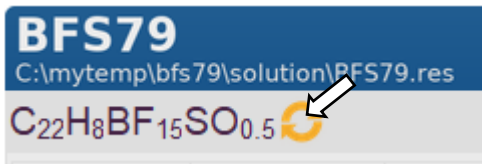
Attention: you **have to** verify if the error is in the **model** or in the **UNIT** command!

Typical causes:

- Extra molecules found after unit instruction was given (solvant, etc.) : **UNIT is wrong**
- HADD added Hs incorrectly : **model is wrong** → do something
- Errors in disorder treatment (particularly if numbers are non integers) : **model is wrong** → do something

Only if UNIT is wrong:

Clicking on the cycle adapts UNIT command to atom list



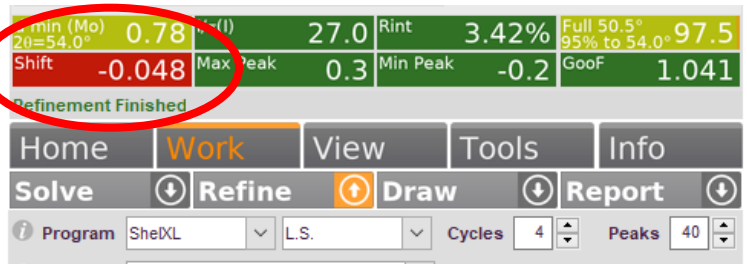
*.res
SFAC C H B F S **O**
UNIT 44 16 2 30 2 **1**

*.ins
SFAC C H B F S
UNIT 44 16 2 30 2

Refine

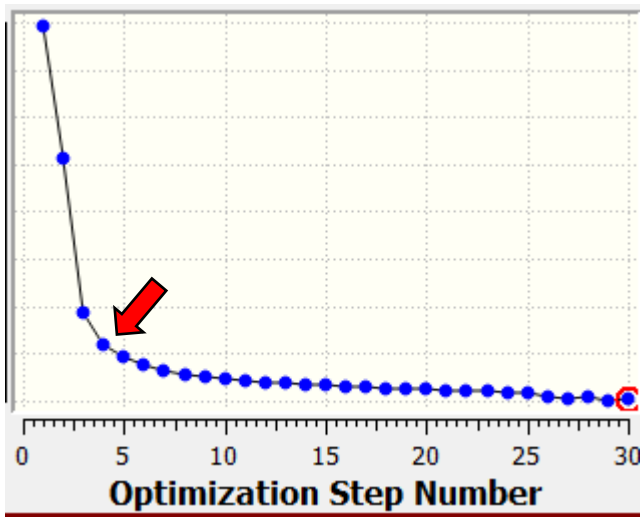
Finishing the refinement

Let the refinement converge completely



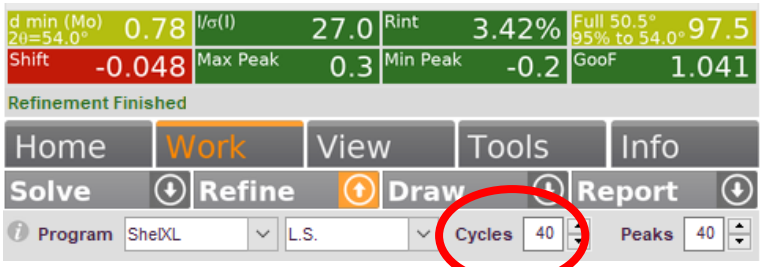
.lst:

Least-squares cycle 4
wR2 = 0.0836 before cycle 2 for 4359 data and 352 / 352 parameters
[…]
Mean shift/esd = 0.013 Maximum = -0.048 for y F13
Max. shift = 0.000 A for C4 Max. dU = 0.000 for F2



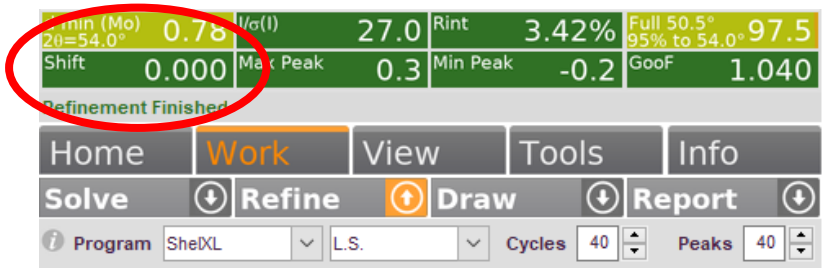
Once we are done with our changes to the model, there is simply no reason, other than laziness, not to let the refinement reach the minimum!

«max. shift» has to be smaller than < 0.003 Å



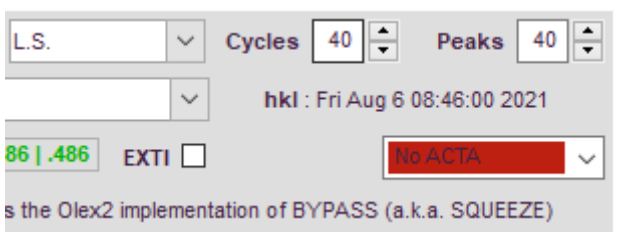
***.ins**: L.S. 40

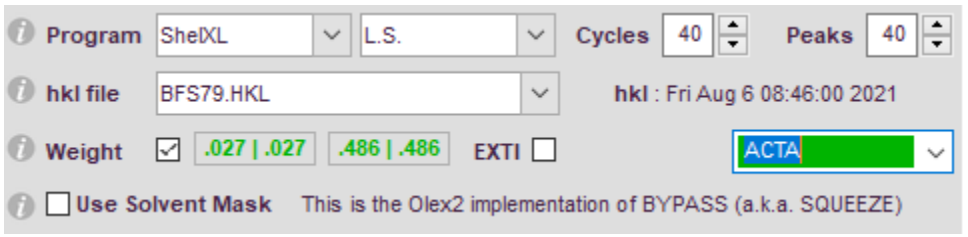
refine



Finishing the refinement

Generate *.cif (and *.fcf) with the ACTA command





***.ins:**

…
SFAC C H N F P
UNIT 56 60 16 24 4
L.S. 40
FMAP 2
…

When generating the CIF, ACTA imposes some "quality control"

- OMIT >0 is not permitted
- FMAP 2 is enforced
- LIST 4 is enforced

***.ins:**

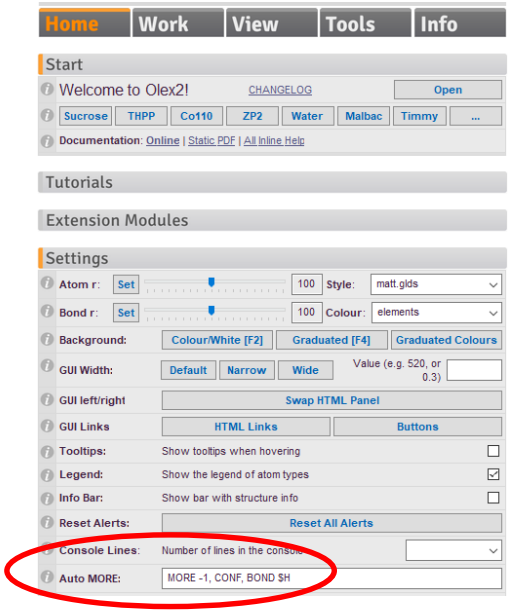
…
SFAC C H N F P
UNIT 56 60 16 24 4
L.S. 40
ACTA
FMAP 2
…

BOND $H and CONF

***.ins**

BOND $H : Includes bond information (including H atoms in the CIF. Required by CHECKCIF.
CONF : Includes torsion angles with esd in CIF

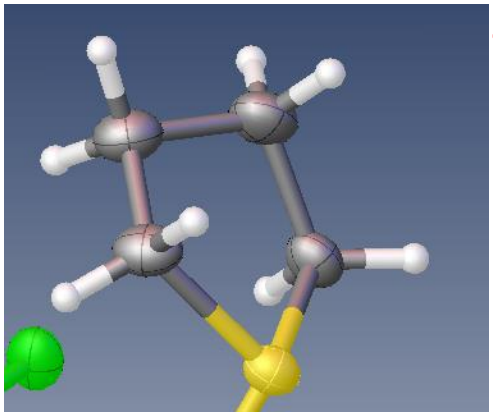
OLEX2 always includes these commands automatically. This is governed in HOME - SETTINGS

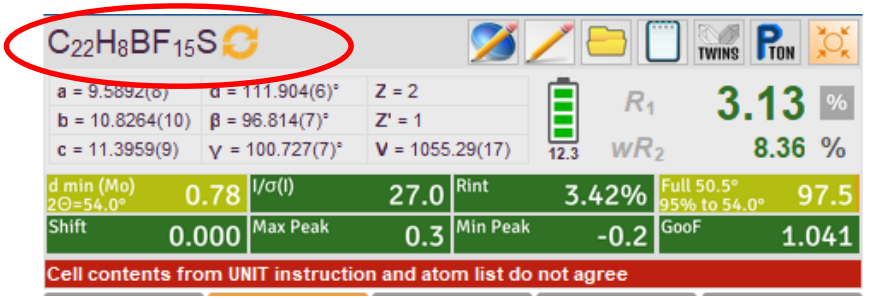


Take-home message

Many/most standard refinement features can be (easier) addressed via the GUI. When should we bother opening the .res or .ins file?

- Whenever you find it more convenient (e. g. disorder refinements)
- **At least once** to double-check for anything unusual before making the final CIF: Unsorted atoms, etc.
- **Every time** when you run into problems





Checkcif:

PLAT041 Calc. and Reported SumFormula Strings Differ
Sum formula C22 **H7.95** B F15 S C22 H8 B F15 S

OLEX2 : edit res

[...]
AFIX 23
H19A 2 0.50169 0.10539 0.69766 **0.98700** -1.20000
H19B 2 0.62439 0.04681 0.74976 **0.95578** -1.20000
AFIX 0