Tutorial on refinement of high pressure single-crystal X-ray diffraction data with SHELXL

1. Overview

After you get the lattice parameters using ATREX and RSV and solved your structure with programs such as Endeavour, Jana, Fox, SHELXT etc. the next step to do a refinement with the model you get.

SHELXL is a popular single-crystal diffraction refinement data. Tutorials and instructions can be found online or in the library:

http://shelx.uni-ac.gwdg.de/SHELX/shelxl_html.php

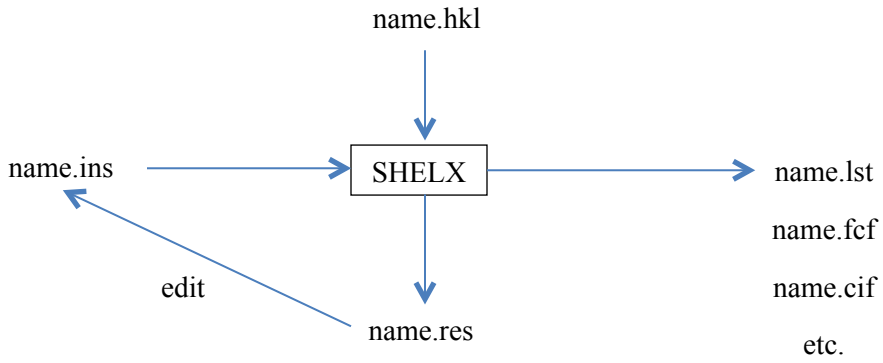
"Crystal structure refinement." By Muller, Peter, et al. *Internacional Union of Crystallography* (2006).

The concept behind SHELXL refinement program is that by means of Fourier transformation, a complete set of structure factor is calculated from the atomic model. The calculated intensities are then compared with the measured intensities. The calculated intensities are then compared with the measured intensities and the best model is that which minimizes M (Muller 2006):

Where w is weight and is derived from the standard deviation of that measurement, F_o is the observed structure factor and F_c is the calculated structure factor.

The quality of the model is judged with the help of various residual factors. The most popular residual factor is the unweighted residual factor based on F which is R1 in SHELXL:

SHELXL is organized as follows (Muller et al. 2006):



Instruction file (name.ins) and reflection file (name.hkl) are input. The way to generate .ins file will be explained in the following section. The reflection file is generated by RSV in a format: h k l . .res file includes the optimized parameters in the model given by SHELXL and in a similar format with .ins file and could be saved as .ins file and then put back into the program. .lst file lists some details about the

model such as most disagreeable peaks, bondlength and so on. .fcf file lists the calculated structure factor for each peak. A .cif file could be generated with the command ACTA.

There are many GUIs for SHELXL and wingx is a popular one. It can be downloaded from: http://www.chem.gla.ac.uk/~louis/software/wingx/

SHELX series programs can be downloaded from:

http://shelx.uni-ac.gwdg.de/SHELX/download.php

Please follow the instructions online and install the software.

2. Input file preparation

Most mineral structures could be found at the American mineralogist database: http://rruff.geo.arizona.edu/AMS/amcsd.php

In the case of the demo data, we download .cif file by Matsumoto et al. (1975).

Our omphacite sample is from the University of Arizona mineral collection R061129 (http://rruff.info/omphacite/display=default/R061129). The measured chemical composition is $(Ca_{0.51}Na_{0.48})_{\Sigma=0.99}(Mg_{0.44}Al_{0.44}Fe^{2+}_{0.14}Fe^{3+}_{0.02})_{\Sigma=1.04}Si_{2.00}O_6$.

Cif files can be explored in visualization software such as Endeavour (http://www.crystalimpact.com/endeavour/) and VESTA (http://jp-minerals.org/vesta/en/). Endeavour is commercial software and VESTA is free. In order to better explain the instructions in .ins file and avoid using commercial software, the preparation of the .ins file will be prepared by hand with the aid of VESTA in this tutorial. Endeavour can also generate SHELXL instruction files. The way to do it is by opening the cif file in Endeavour and then save it as .ins file. However, the .ins file generated by Endeavour usually needs modification and can be done by following the instructions below.

After reading .cif file into VESTA, structure and data sheet appears (figure 1). Then we open a plain text and name it with a suffix .ins. Some commands must be used in order to start a refinement. Usually, we start with the simplest model with only isotropic displacement factors and then complicated models with disorders and anisotropic factors will be added to the model step by step. Initial command will be explained one by one in the following text and the actually command will be in red and Arial type.

(1) TITL: The first command is title command.

TITL DEMO

(2) CELL: the first number in this command is wavelength (Å) and is followed by the six lattice parameters a, b, c, α, β, γ. In the demo, the experiment is conducted with synchrotron source in APS and the wavelength is 0.3344 Å. There are different unit cell settings, you can choose whichever you want but the **IMPORTANT** thing is that the settings of the lattice parameters, the atomic positions and the reflections file generated by RSV be the same! The atomic position and the lattice parameters in .cif file agree with themselves. The way to check if the .hkl file generated by RSV agrees with the setting in .cif file is to check if the sequence of lattice parameters in RSV agrees with the sequence of lattice parameters in .cif file (figure2). If the setting in RSV is different from that in the .cif file, one can transform the setting in .hkl file by input the transformation matrix in RSV and then click 'Transform' (figure 2). By comparing figure 1 and figure 2, we know the setting in .cif file and in .hkl file is the same.

CELL 0.3344 9.38932 8.58425 5.15557 90 106.08 90

(3) ZERR: The first number in this command line is Z number (the number of formula in unit cell) and followed by the standard deviations with respect to CELL command line.

ZERR 2 0.00254 0.00136 0.00045 0 0.01329 0

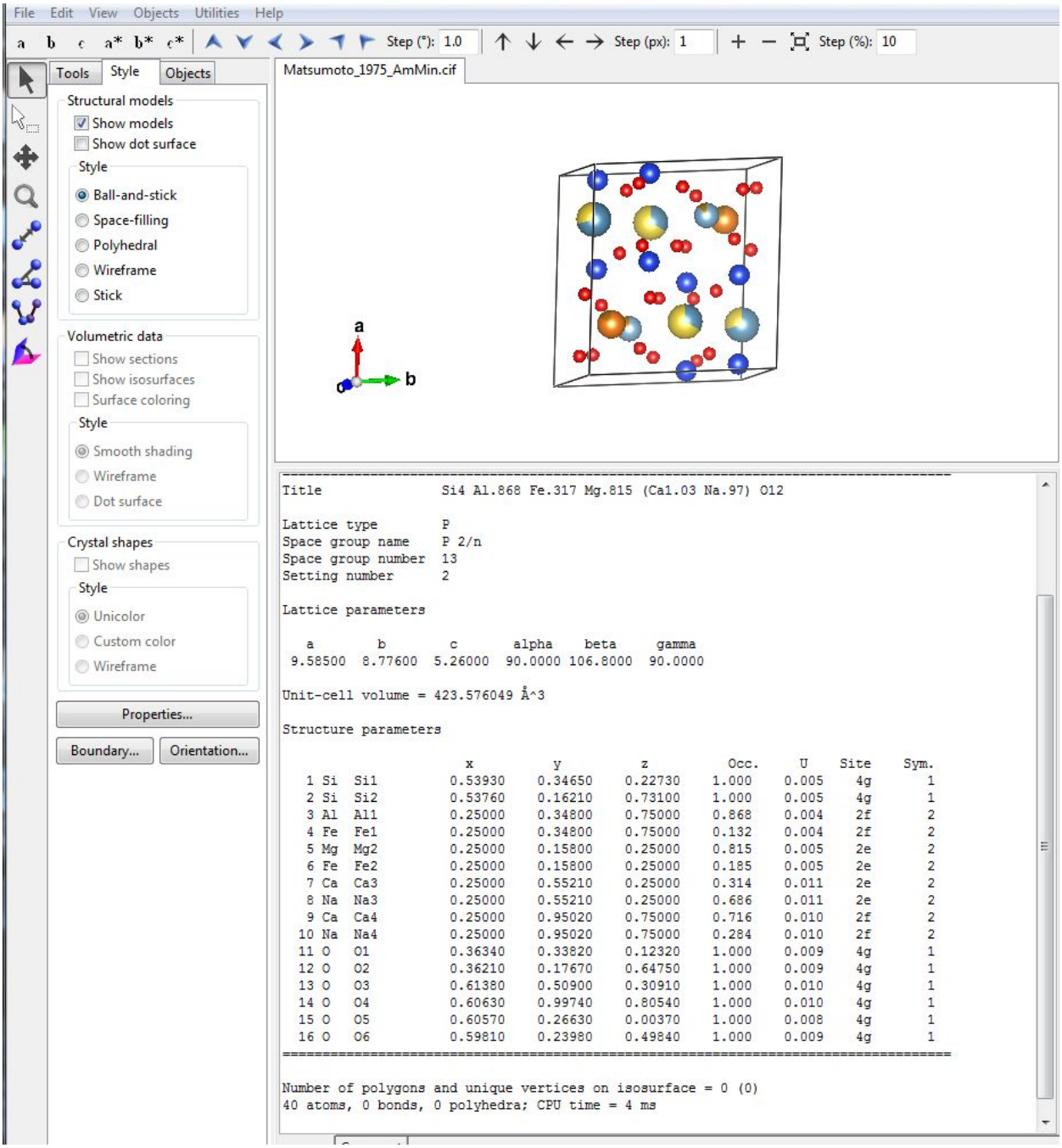


Figure 1 structure visualization and data sheet in VESTA.

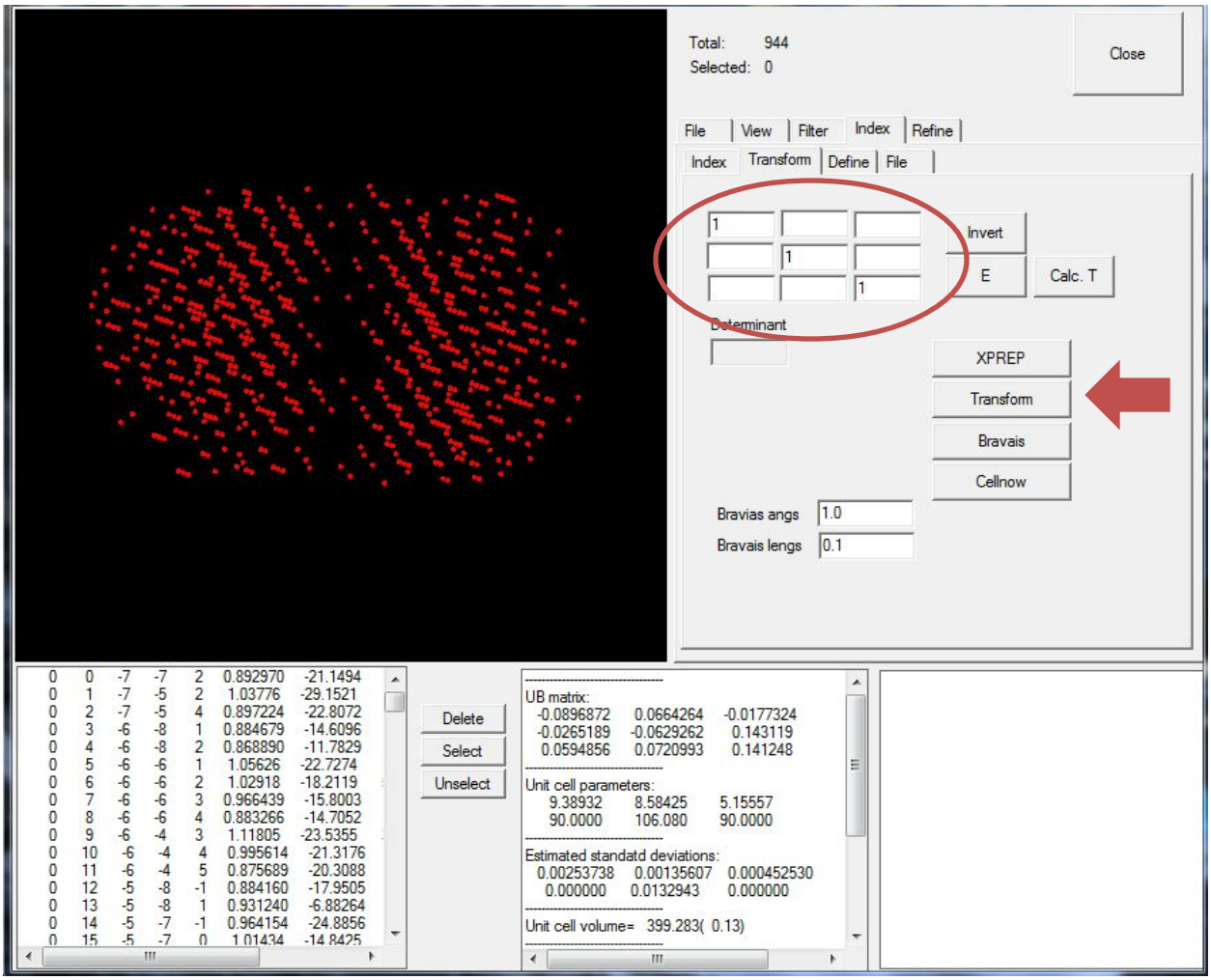


Figure 2 Lattice parameters and transformation matrix in RSV.

(4) LATT: indicates the lattice type. 1=P, 2=I, 3=rhombohedral obverse on hexagonal axes, 4=F, 5=A, 6=B, 7=C. N must be made negative if the structure is non-centrosymmetric. In our case, it is P lattice.

LATT 1

(5) SYMM: shows the symmetry operator. The operator x, y, z is always assumed and may not be input. The presence of an inversion center should be indicated by LATT instead of SYMM. In this case, the international tables give four symmetries: <1> x, y, z; <2> ½-x, ½-y, z; <3> -x, -y, -z; <4> ½+x, ½+y, -z. To follow the rules above, only ½-x, ½-y, z should be written down.

SYMM 0.5-x, 0.5-y, z

(6) SFAC and UNIT: element should be specified in SFAC and the number of atoms of each type should be specified in UNIT in the order used by SFAC. Sometimes with disordered model it is difficult to specify number of atoms. The number specified in UNIT is used to calculated density and is not used for refinement, so putting ones in UNIT command line doesn't matter. XRD doesn't resolve atomic elements as good as electron microprobe, so only the major elements may be refined, so the following elements are put into the model.

SFAC Ca Na Mg Al Si O

UNIT 1 1 1 1 1 1

(7) LIST: shows the format write to .fcf file. .fcf file can be put into ATREX IDL again for systematic correction. In order to match the format used in ATREX IDL, the number 4 option should be used.

LIST 4

(8) L.S. : list the number of cycles of refinements performs. Usually 10 cycles are used.

L.S. 10

(9) WGHT: it is followed by these numbers: a b c c e f. The weighting scheme is defined as follows:

$$w = q / [\sigma^2(F_o^2) + (a*P)^2 + b*P + d + e*\sin(\theta)/\lambda]$$

where P = f * Maximum of (0 or F_o^2) + (1-f) * F_c^2. Usually, a default number of 0.1 is used.

WGHT 0.1

(10) FVAR and atomic instructions: FVAR are followed by the overall scale factor (osf) and free variables. The free variables are usually combined with site occupancy factors used in atomic instructions. Atom instructions begin with an atom name (up to 4 characters, of which the first must be a letter) and then follows with the sequence number of the atom listed in the SFAC command. Atomic positions x, y, z follow. The **important** thing is that the atomic positions must be in the same setting with the lattice parameters and the reflection file as mentioned above in CELL. Site occupancy factor (sof) then follows. The sof is given as 10*m+p and m is an integer and abs(p) is less than 5. M=1 means no refinement of sof and m>1 means refine sof. The refined sof corresponds to the m^{th} number following FVAR. Displacement factor follows sof and an initial value of 0.005 is usually used.

In our case, it is an ordered omphacite phase obtained from slow cooling (disordered phase has space group C2/c and order phase has space group P2/n).

From the cif file, we see that majority of Ca occupy position 0.25000, 0.95020, 0.75000 and sits on the 2e special position so p is set as 0.5 (general position has a multiplicity of 4). We want to refine the occupation so we set m=2 so that it will correspond to the second number in FVAR. Other atomic positons are set in a similar way.

(11) HKLF and END: this instruction indicates the type of .hkl file used. RSV gives type 4 so will put 4 here. END command is need. Anything after END will not be put into the program.

Our start model is this:

TITL DEMO

CELL 0.3344 9.38932 8.58425 5.15557 90 106.08 90

ZERR 2 0.00254 0.00136 0.00045 0 0.01329 0

LATT 1

SYMM 0.500-x, y, 0.500-z

SFAC Ca Na Mg Al Si O

```
UNIT 1 1 1 1 1 1
LIST 4
L.S. 10
WGHT    0.1
FVAR    1 1 1 1 1
Ca1  1 0.25000 0.95020 0.75000 21 0.005
Na1  2 0.25000 0.55210 0.25000 31 0.005
Mg1  3 0.25000 0.34800 0.75000 41 0.005
Al1  4 0.25000 0.15800 0.25000 51 0.005
Si1  5 0.53930 0.34650 0.22730 11 0.005
Si2  5 0.53760 0.16210 0.73100 11 0.005
O1   6 0.36340 0.33820 0.12320 11 0.005
O2   6 0.36210 0.17670 0.64750 11 0.005
O3   6 0.61380 0.50900 0.30910 11 0.005
O4   6 0.60630 0.99740 0.80540 11 0.005
O5   6 0.60570 0.26630 0.00370 11 0.005
O6   6 0.59810 0.23980 0.49840 11 0.005
HKLF 4
END
```

3. First refinement

With .ins and .hkl file in hand, we can start our first refinement. Please put them into same directory. Open WINGX and open project by click File->project->start/select new project. Then a dialog will pop up and ask you to select the .ins file (figure 3).

Refinement can be done by clicking Refine-> SHELXL. All the input and output files can be opened by clicking the Open … buttons below.

The first refinement as shown in figure 5 gives a very good R1 ~ 7% and this indicates that the input model is correct in general.

Results are given in the .res file. Atomic positions, displacement factors and FVAR are updated in the file and a suggested weighting scheme is also suggested after the END instruction (figure 6). Weighting scheme should be retained at the beginning of refinement and only be adjusted in at the end of refinement.

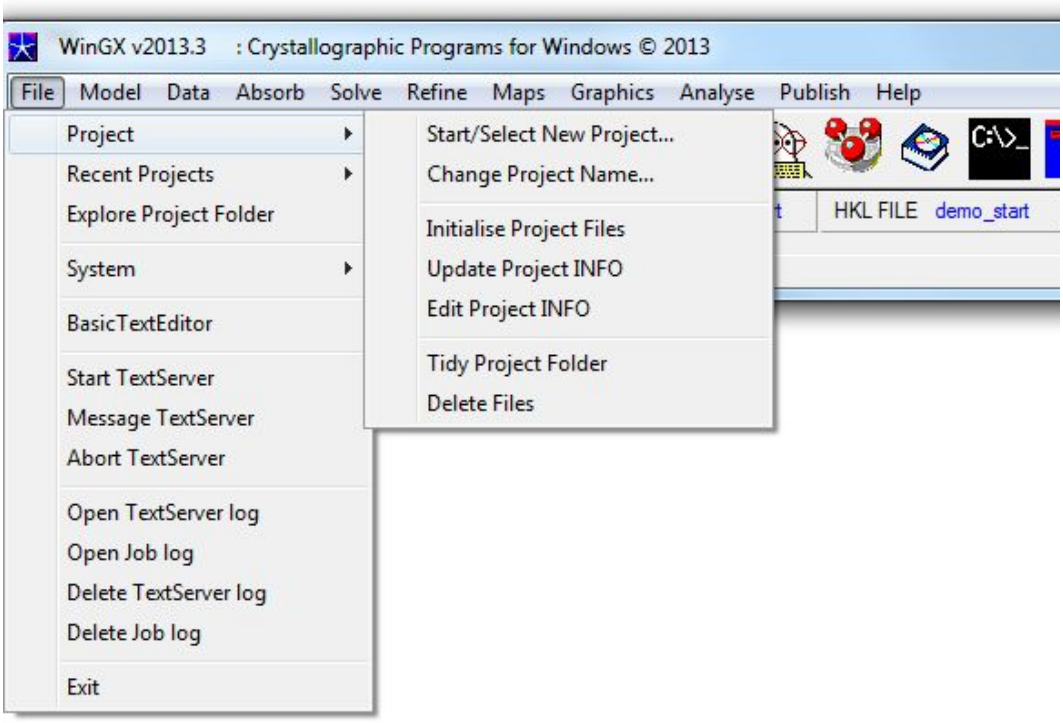


Figure 3 start new project

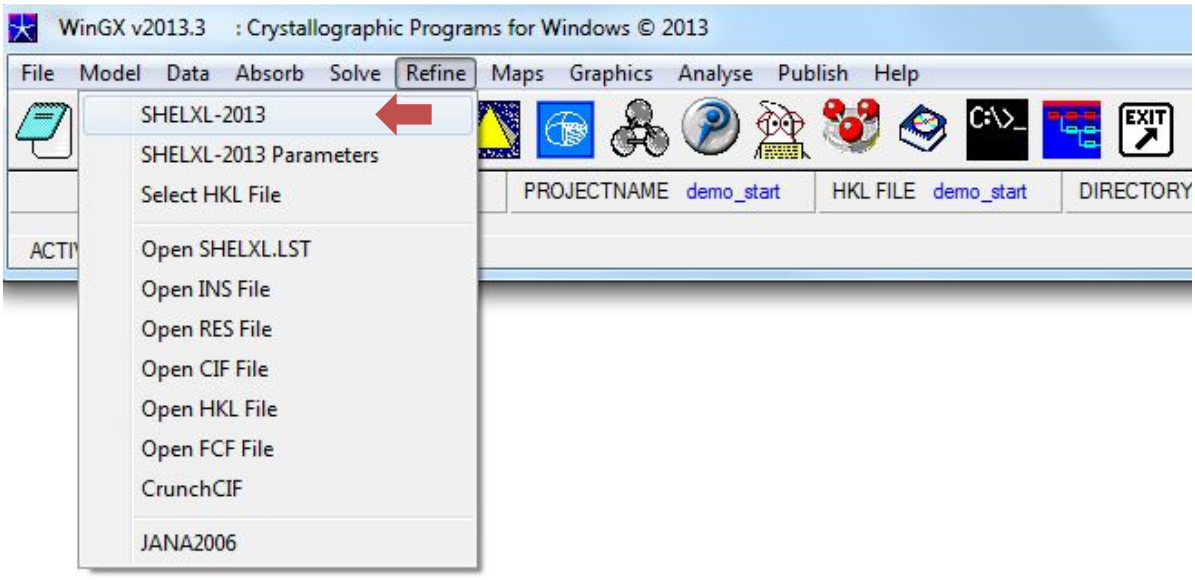


Figure 4 start refinement

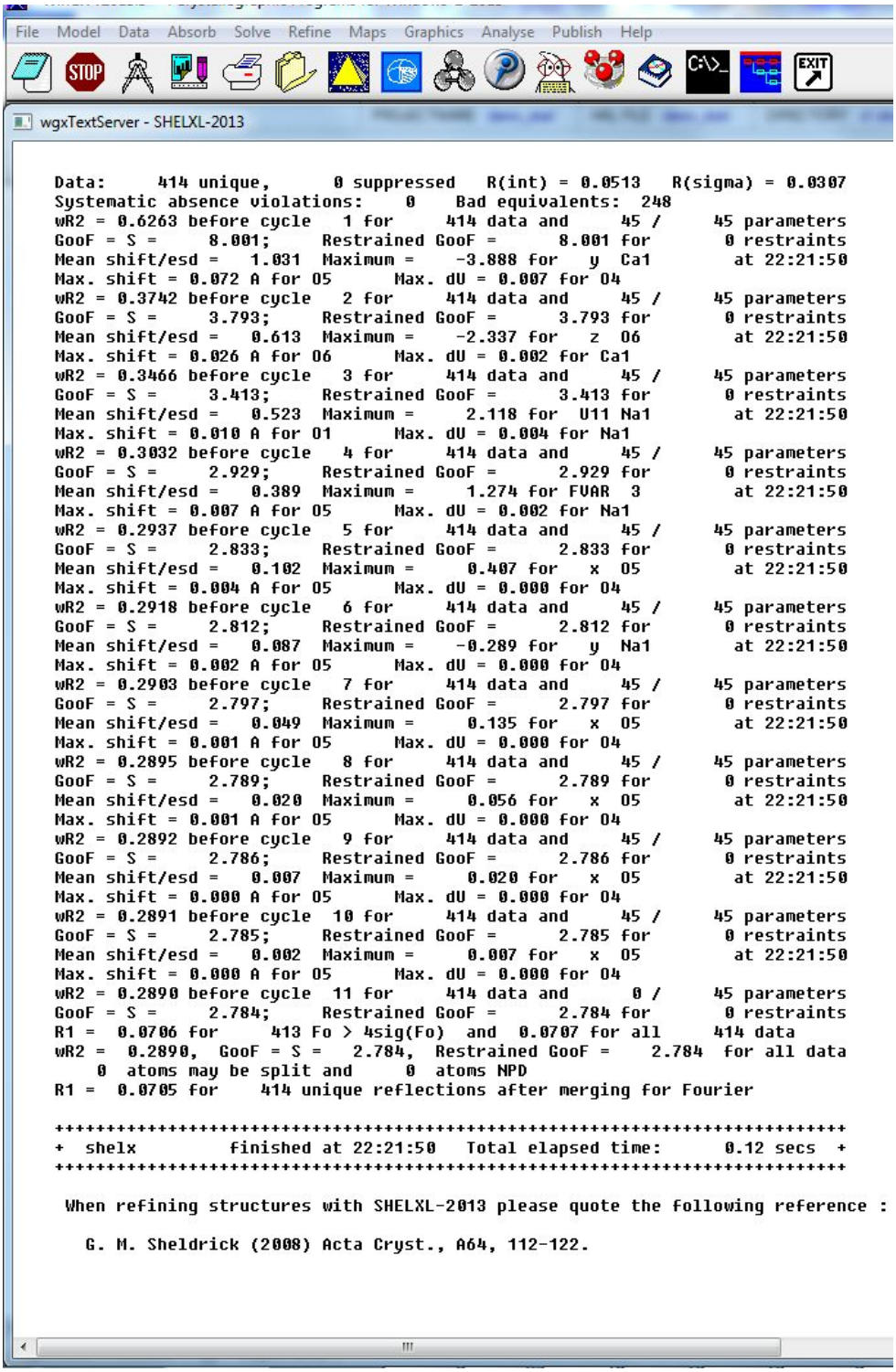


Figure 5 first refinement results

File Edit Search Help

```
TITL DEMO
CELL 0.3344 9.38932 8.58425 5.15557 90 106.08 90
ZERR 2 0.00254 0.00136 0.00045 0 0.01329 0
LATT 1
SYMM 0.500-x, y, 0.500-z
SFAC Ca Na Mg Al Si O
UNIT 1 1 1 1 1 1
LIST 4
L.S. 10
WGHT    0.100000
FVAR       0.82011   0.93193   1.19789   1.23101   1.22208
CA1   1    0.250000    0.946145    0.750000    20.50000    0.01097
NA1   2    0.250000    0.555738    0.250000    30.50000    0.01059
MG1   3    0.250000    0.347119    0.750000    40.50000    0.00440
AL1   4    0.250000    0.160686    0.250000    50.50000    0.00657
SI1   5    0.539536    0.348244    0.228025    11.00000    0.00397
SI2   5    0.537606    0.160363    0.729081    11.00000    0.00606
O1    6    0.360906    0.340987    0.123408    11.00000    0.00854
O2    6    0.363221    0.174422    0.650429    11.00000    0.00480
O3    6    0.609305    0.511967    0.313982    11.00000    0.00270
O4    6    0.607455    0.991833    0.811267    11.00000    0.01551
O5    6    0.610432    0.272406    0.000845    11.00000    0.00779
O6    6    0.601580    0.233208    0.489528    11.00000    0.00483
HKLF 4

REM  DEMO
REM R1 =  0.0706 for     413 Fo > 4sig(Fo)  and  0.0707 for all     414 data
REM     45 parameters refined using      0 restraints

END

WGHT      0.1378      6.2944
```

Figure 6 .res file

4. Refinement of anisotropy

In the starting model, an isotropic displacement factor is use. However, the more realistic model should be an anisotropic one. ANIS instruction followed by the species names adds anisotropic displacement into the model. Adding anisotropic factor increases the freedom of the model and sometimes negative displacement will result from it. So we always start from refining anisotropic for cations first and then the anions. High pressure data usually has a worse quality and oxygen may not be able to be refined anisotropically.

ANIS CA1 NA1 MG1 AL1

5. Peaking omitting

Sometimes sample peaks are contaminated by diamond or gasket peaks and omitting those peaks gives better results. Most disagreeable peaks can be found in .lst file. The instruction for omitting peaks is:
OMIT h k l

6. Systematic error correction

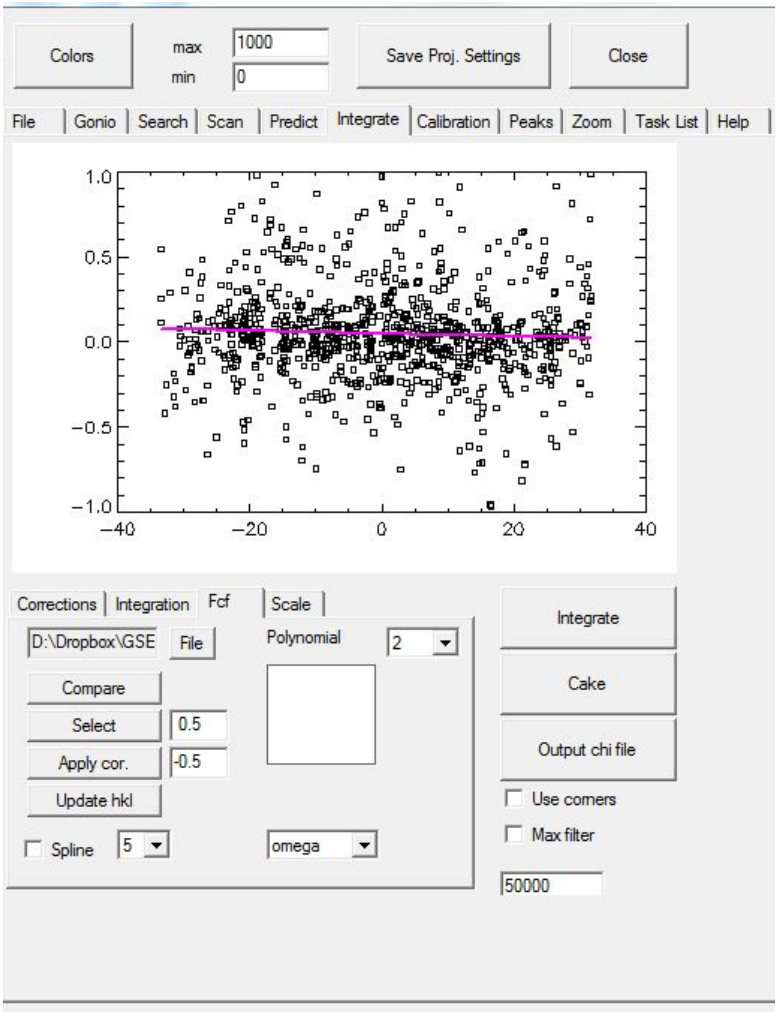


Figure 7 systematic error correction with the use of .fcf file in ATREX IDL

Calculated diffraction data (.fcf) can be put into ATREX IDL and then compared with the observed reflection data and systematic error can be corrected this way. The way to do it is first put a MERG 0 instruction in .ins file and then run the refinement. In this way, symmetry dependent peaks will not be averaged out. Then open the observed peaks from 'Peaks' tab in ATREX and open the .fcf file from SHELXL. Click compare and the results will show. Similar to the scale step in ATREX IDL, you can apply correction by clicking 'Apply cor.' and select outliers within the range shown in the box and then click 'select' and then delete them. After the correction is done, you can click 'update hkl' and then use it for refinement.

7. More complicated refinement

Disorder usually exists in mineral sample and this can be refined with FVAR and SUMP instructions. For further information, please refer to the SHELXL help.

Twinning can also be refined. Different twinning has different refinement strategy and it is out of the scope of this tutorial. If interested in the topic please refer to TWIN and HKLF 5 instructions online.

8. Final refinements

With the model almost settled, we can refine the weighting scheme. One can just copy the suggested WGHT and replace the old one in .res file and then save it as .ins and run refinement again. Usually a publishable refinement result should give a R1 ~ 5%.

In order to generate .cif file and include bond length, ACTA and BOND instructions must be included in the .ins file.

9. Table generation

Tables that are easy for viewing can be generated by using ciftab (http://shelx.uni-ac.gwdg.de/SHELX/cif.php) provided by SHELX suite. Template can also be downloaded. It works in Windows and Linux environment. The tutorial illustrates how to use the tool in Ubuntu Linux.

Unzip the CIFTAB program and then the templates in the same directory. Start the program by typing ./ciftab command and then follow the questions. Then a table in .rtf format will be generated (figure 8). Usually a default value is given in '[]' and typing enter means default value.

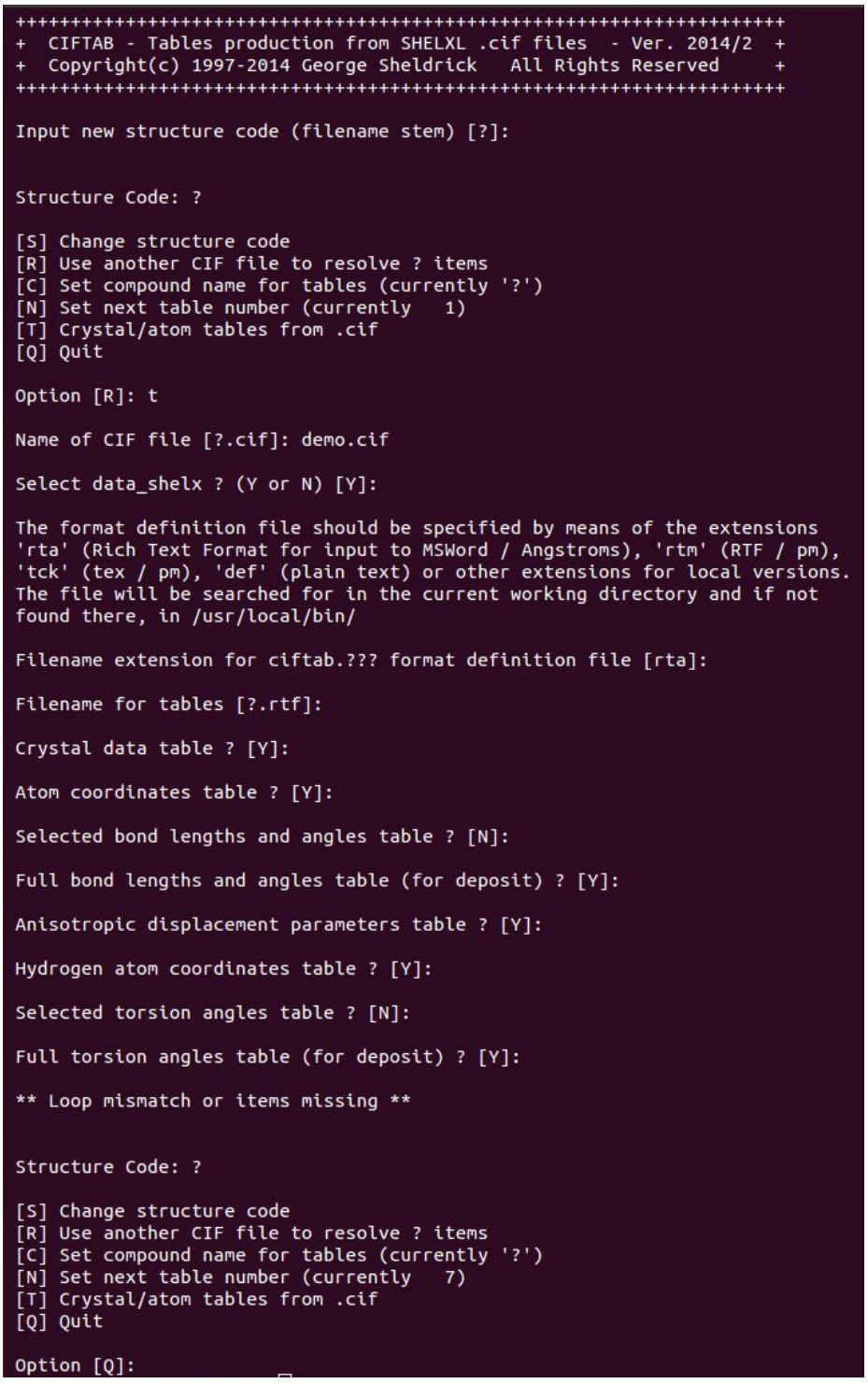


Figure 8 Table generation using ciftab.

Reference

Muller, Peter, et al. "Crystal structure refinement." *Internacional Union of Crystallography* (2006): 1-95.