Space Group & Structure Solution

Determine the Space Group

Space group determination can always be performed by examining the intensity data. A program that can facilitate this step is the command-prompt program xprep. Typical input is *project* name where the intensity data are in the file *project*.hkl and the cell parameters and chemical formula are in the file *project*.p4p.

To determine the space group by hand, perform the following steps:
- First be sure that the original unit cell is *reduced*.[1,2]
- From the reduced cell, search for possible higher symmetry, centered cells. This can be done using the Bravais routine in the APEX3 program[3] or using table 9.2.5.1 of the *Int'l Tables for Crystallography*, Vol. A.[4] Note that the cell parameters must plausibly match the values expected for the crystal system. For each of these plausible higher-symmetry cells, transform the (*hkl*) values and check the intensities to see if the data meets the Laue symmetry for that crystal system. These tests should begin with the plausible, highest-symmetry, centered cells. When the Laue symmetry is matched, you have found the appropriate crystal system and Laue symmetry for the sample.
- For this crystal system, check the intensities for possible reflection conditions (systematic absences), paying the greater attention to glide planes rather than screw axes because a glide plane affects an entire plane of the reciprocal lattice, while a screw axis affects only data on a single line. Once the reflection conditions are determined, the cell may need to be rotated using Table 4.3.2.1 of the *Int'l Tables for Crystallography*, Vol A,[5] to put the cell into a standard orientation.

Before running the xprep program, be sure that the *.p4p and the *.hkl or *.raw files have the same root file name.

Typical input is *project* name where the intensity data are in the file *project*.hkl and the cell parameters and chemical formula are in the file *project*.p4p. The program presents the results of different calculations asking the user to either accept their results or pick another option. For most selections, hit <enter> to accept the default.

```
XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014          _ □ X

++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
+  XPREP - Reciprocal space exploration  - Version 2014/2 for Windows  +
+  COPYRIGHT(c) 2014 Bruker-AXS                    All Rights Reserved +
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Screen size: 1920 x 1080
Window size:  640 x  979
Font size:      8 x   16  ( 125 x  168 )
Number of colors:    256

When xprep is started without a filename on the command line, the filename
is prompted for and then the type of data (SHELX, SCALEPACK, XDS etc.)
requested. XPREP without a filename is also required to read HKLF3 format data.
To generate ideal data, a SHELX .ins or .res file should be given. To determine
the Flack parameter by the Parsons' quotient method, these data are used to
find the space group, then the observed intensities are read in.

Alternatively xprep may be started by 'xprep name.xyz' where xyz is hkl, HKL
(for a XDS_ASCII.HKL format file), res, ins, fco (for a file written by XD),
or fcf for a CIF file written by SHELXL or from the IUCr or COD servers. If a
fco or fcf file contains both Fo^2 and Fc^2, both datasets will be read in:
first Fc^2, then Fo^2. After defining the space group in the usual way,
the Flack x may then be determined by the Parsons' quotient method.

'xprep name' reads a SHELX HKLF 4 format file name.hkl, then tries to find
name.spin or name.p4p to extract the cell dimensions and their esds.

'xprep name1 name2' reads name1.hkl and name2.p4p (or name2.spin).

-Ln on the command line allocates space for 1000000n data (default n=4).

  13986 reflections read from file ylid20.hkl
  Mean (I/sigma) =   23.73


Lattice exceptions: P      A      B      C      I      F      Obv     Rev    All

N (total) =         0   6953   6984   7025   6970  10481   9306    9304  13986
N (int>3sigma) =    0   6509   6484   6571   6437   9782   8640    8631  12991
Mean intensity =  0.0   27.3   26.1   25.7   27.3   26.4   26.9    27.0   26.7
Mean int/sigma =  0.0   23.9   22.9   23.8   23.7   23.5   23.6    23.7   23.7

Lattice type [P, A, B, C, I, F, O(obv.), R(rev. rhomb. on hex. axes)]

Select option [P]: █
```

The program calculates an average (mean) $I/\sigma(I)$ for the entire data set. Values of 10 or greater for the average $I/\sigma(I)$ suggest that the structure determination is very likely to succeed. The program then presents a table of statistics to determine the current lattice centering. The statistics displayed by the program are the total number of data that should systematically absent if that condition existed, the number of the systematically-absent data that are observed, an average intensity of the proposed systematically-absent data, and finally the $I/\sigma(I)$ of the systematically-absent data. The default choice is based on the $I/\sigma(I)$ results (> 3.87); however, be sure to look at the average intensity data as well. If the cell is found to be a centered lattice, the primitive cell is determined and the primitive cell is reduced.

```
XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014          _ □ ✕
Current dataset: ylid20.hkl                  Wavelength: 0.71073 Chiral: ?
---------------------------------------------------------------------------
Original cell:    5.966    9.045   18.408    90.00    90.00    90.00  Vol      993.3
          Esds:   0.000    0.000    0.001     0.00     0.00     0.00  Lattice: P
---------------------------------------------------------------------------
Current cell:     5.966    9.045   18.408    90.00    90.00    90.00  Vol      993.3
---------------------------------------------------------------------------
Matrix: 1.0000   0.0000   0.0000   0.0000   1.0000   0.0000   0.0000   0.0000   1.0000
---------------------------------------------------------------------------


   [D] Read, modify or merge DATASETS        [C] Define unit-cell CONTENTS
   [P] Contour PATTERSON sections            [F] Set up shelxtl FILES
   [H] Search for HIGHER metric symmetry     [R] RECIPROCAL space displays
   [S] Determine or input SPACE GROUP        [U] UNIT-CELL transformations
   [A] Absorption, powder, SIR, SAD, MAD etc. [T] Change TOLERANCES
   [M] Test for MEROHEDRAL TWINNING          [O] Self-rotation function
   [L] Reset LATTICE type of original cell   [Q] QUIT program

   Select option [H]:


   Determination of reduced (Niggli) cell

   Transformation from original cell (HKLF-matrix):
      1.0000   0.0000   0.0000    0.0000   1.0000   0.0000    0.0000   0.0000   1.0000

   Unitcell:       5.966    9.045   18.408    90.00    90.00    90.00

   Niggli form:    a.a =     35.59      b.b =     81.81      c.c =    338.85
                   b.c =      0.00      a.c =      0.00      a.b =      0.00

   Search for higher METRIC symmetry
   Identical indices and Friedel opposites combined before calculating R(sym)

   ---------------------------------------------------------------------------
   Option A: FOM = 0.000 deg.   ORTHORHOMBIC P-lattice   R(sym) = 0.011 [   3218]
   Cell:    5.966    9.045   18.408    90.00    90.00    90.00   Volume:      993.27
   Matrix: 1.0000   0.0000   0.0000   0.0000   1.0000   0.0000   0.0000   0.0000   1.0000
   ---------------------------------------------------------------------------
   Option B retains original cell

   Select option [A]: ▮
```

Higher-symmetry, centered cells are considered by calculating the cell parameter fit to the transformed cells and then calculating the merging R for the crystal system. The program stops looking for lower crystal symmetry groups when the merging R is found to be reasonably low. The combination of lattice type and crystal class is selected.

```
XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014          _ □ X
Current dataset: ylid20.hkl                  Wavelength: 0.71073 Chiral: ?
-------------------------------------------------------------------------
Original cell:    5.966    9.045  18.408    90.00    90.00    90.00  Vol      993.3
        Esds:     0.000    0.000   0.001     0.00     0.00     0.00  Lattice: P
-------------------------------------------------------------------------
Current cell:     5.966    9.045  18.408    90.00    90.00    90.00  Vol      993.3
-------------------------------------------------------------------------
Matrix: 1.0000  0.0000  0.0000  0.0000  1.0000  0.0000  0.0000  0.0000  1.0000
-------------------------------------------------------------------------
Crystal system: Orthorhombic                 Lattice: P
-------------------------------------------------------------------------

 [S] Determine SPACE GROUP
 [C] Must be CHIRAL (sample is optically active)
 [N] NOT NECESSARILY chiral (eg. may be racemate)
 [I] INPUT known space group
 [E] EXIT to main menu  or  [Q] QUIT program

 Select option [S]:


 [A] Triclinic,  [M] Monoclinic,  [O] Orthorhombic,  [T] Tetragonal,
 [H] Trigonal/Hexagonal,  [C] Cubic  or  [E] EXIT

 Select option [O]:


 Lattice exceptions: P      A      B      C      I      F     Obv    Rev    All

 N (total) =        0   6953   6984   7025   6970  10481   9306   9304  13986
 N (int>3sigma) =   0   6509   6484   6571   6437   9782   8640   8631  12991
 Mean intensity =  0.0   27.3   26.1   25.7   27.3   26.4   26.9   27.0   26.7
 Mean int/sigma =  0.0   23.9   22.9   23.8   23.7   23.5   23.6   23.7   23.7

 Lattice type [P, A, B, C, I, F, O(obv.), R(rev. rhomb. on hex. axes)]

 Select option [P]:


 Mean |E*E-1| = 0.696 [expected .968 centrosym and .736 non-centrosym]


 Systematic absence exceptions:

          b--    c--    n--   21--    -c-    -a-    -n-   -21-    --a    --b    --n   --21

 N       723    706    697     12    464    480    464     21    239    238    229     37
 N I>3s  636    615    593      2    379    385    328      2    205    212    183      4
 <I>    31.1   32.8   38.6    0.1   53.3   48.7   27.0    0.0   39.1   41.1   33.9    0.1
 <I/s>  27.0   26.4   27.7    1.5   29.8   28.0   20.4    1.1   26.9   28.7   26.6    1.6


 Identical indices and Friedel opposites combined before calculating R(sym)

 Option  Space Group  No.  Type  Axes  CSD  R(sym) N(eq)  Syst. Abs.   CFOM

 [A] P2(1)2(1)2(1)  # 19  chiral   1  5917  0.011  3218  1.6 / 20.4   0.70

 Select option [A]: █
```

The $<|E^2 - 1|>$ for all data is calculated to test for a center of symmetry. From the lattice and crystal system the reflection conditions (systematic absences) for the glide planes and screw axes

of the possible space groups are tested. Note that the tests for glide planes and screw axes are not perfect. These tests can make mistakes, so carefully review the statistics. The program then lists either no space group, a unique space group, or a small list of possible space groups. If a more likely space group is not listed, then end this routine and select Input Space Group to put in your preferred choice.

```
 XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014        _ □ ×

 Resolution  #Data #Theory %Complete Redundancy Mean I Mean I/s Rmerge  Rsigma
  Inf - 3.64    22     23     95.7      7.39    200.36  151.12  0.0200  0.0046
 3.64 - 2.31    47     47    100.0     11.53    133.50  189.60  0.0152  0.0043
 2.31 - 1.78    69     69    100.0     11.78     59.75  153.97  0.0153  0.0052
 1.78 - 1.51    68     68    100.0     12.94     63.39  158.83  0.0156  0.0051
 1.51 - 1.38    68     68    100.0     12.47     35.87  125.61  0.0166  0.0063
 1.38 - 1.26    74     74    100.0     12.30     25.60  114.27  0.0192  0.0073
 1.26 - 1.18    66     66    100.0     11.65     27.54  104.67  0.0192  0.0077
 1.18 - 1.12    65     65    100.0     11.34     21.25   87.69  0.0226  0.0092
 1.12 - 1.06    81     81    100.0     11.04     12.51   65.82  0.0294  0.0114
 1.06 - 1.02    58     58    100.0     11.10     11.19   62.29  0.0344  0.0133
 1.02 - 0.98    73     73    100.0      9.90     12.54   58.07  0.0363  0.0143
 0.98 - 0.95    69     69    100.0      9.86      9.11   49.73  0.0374  0.0166
 0.95 - 0.92    71     71    100.0      9.94      7.44   43.14  0.0448  0.0194
 0.92 - 0.89    81     81    100.0      9.36      6.50   35.68  0.0482  0.0214
 0.89 - 0.87    63     63    100.0      8.98      5.69   33.43  0.0553  0.0252
 0.87 - 0.85    56     56    100.0      8.73      4.63   29.85  0.0635  0.0288
 0.85 - 0.83    72     72    100.0      8.50      4.50   27.51  0.0638  0.0315
 0.83 - 0.81    89     89    100.0      8.42      4.59   25.61  0.0685  0.0310
 0.81 - 0.80    40     40    100.0      8.32      3.64   22.54  0.0722  0.0356
 0.80 - 0.78    88     88    100.0      8.23      3.46   21.23  0.0789  0.0390
 0.78 - 0.77    58     58    100.0      7.57      3.51   20.36  0.0755  0.0419
 ------------------------------------------------------------------------------
 0.87 - 0.77   403    403    100.0      8.30      4.09   24.52  0.0696  0.0340
  Inf - 0.77  1378   1379     99.9     10.14     23.89   70.99  0.0216  0.0084

 Merged [A],  lowest resolution =  9.20 Angstroms

 Graphical output: 1=<I/s>, 2=Rmerge, 3=Rsigma, <Enter>=none: █
```

The program then asks if you want to display statistics for the whole data set; select the defaults, unless you want more details about the data. An option to display layers of the diffraction data is available and is often checked when the subsequent refinement has problems.

```
XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014          _ □ x

Current formula is:
C11H10O2S


Tentative Z (number of formula units/cell) =   4.0  giving rho = 1.379,
non-H atomic volume =  17.7  and following cell contents and analysis:

C      44.00     64.05 %              H      40.00      4.89 %
O       8.00     15.51 %              S       4.00     15.55 %

[Z] change Z,  [F] new FORMULA,    [R] change RADIATION,
[E] EXIT to main menu   or   [Q] QUIT program

Select option [E]: ▊
```

The default path through the program asks about the chemical formula and calculates a possible Z for the given formula. If the radiation used is not the default Mo $K\alpha$, then select the appropriate radiation.

```
XPREP Version 2014/2 for Windows Copyright(C) Bruker-AXS 2014          _ □ X

Current dataset: ylid20.hkl                    Wavelength: 0.71073 Chiral: ?
-------------------------------------------------------------------------
Original cell:   5.966   9.045  18.408   90.00   90.00   90.00  Vol    993.3
        Esds:   0.000   0.000   0.001    0.00    0.00    0.00  Lattice: P
-------------------------------------------------------------------------
Current cell:    5.966   9.045  18.408   90.00   90.00   90.00  Vol    993.3
-------------------------------------------------------------------------
Matrix: 1.0000  0.0000  0.0000  0.0000  1.0000  0.0000  0.0000  0.0000  1.0000
-------------------------------------------------------------------------
Crystal system: Orthorhombic  Space group: P2(1)2(1)2(1)  # 19 [chi]  Laue: 3
-------------------------------------------------------------------------
Formula: C11H10O2S                              Formula wt:     206.26
Z:  4.00   Density:  1.379  At.vol:  17.7  F(000):    432.00  Mu[mm-1]:   0.29
-------------------------------------------------------------------------


[D] Read, modify or merge DATASETS        [C] Define unit-cell CONTENTS
[P] Contour PATTERSON sections            [F] Set up shelxtl FILES
[H] Search for HIGHER metric symmetry     [R] RECIPROCAL space displays
[S] Determine or input SPACE GROUP        [U] UNIT-CELL transformations
[A] Absorption, powder, SIR, SAD, MAD etc. [T] Change TOLERANCES
[M] Test for MEROHEDRAL TWINNING          [O] Self-rotation function
[L] Reset LATTICE type of original cell   [Q] QUIT program

Select option [Q]: f

Output file name (without extension) [ylid20]:

XM/SHELXD (M) or XS/SHELXS (S) format [S]:


File ylid20.ins set up as follows:

TITL ylid20 in P2(1)2(1)2(1)
CELL 0.71073   5.96570   9.04490  18.40780  90.0000  90.0000  90.0000
ZERR   4.00   0.00020   0.00030   0.00060   0.0000   0.0000   0.0000
LATT -1
SYMM 0.5-X, -Y, 0.5+Z
SYMM -X, 0.5+Y, 0.5-Z
SYMM 0.5+X, 0.5-Y, -Z
SFAC C H O S
UNIT 44 40 8 4
TEMP 0.000
SIZE 0.40 0.40 0.40
TREF
HKLF 4
END

Do you wish to (over)write the intensity data file ylid20.hkl ? [Y]: █
```

Finally, an option to set up data files for structure solution is presented.

Structure Solution

The recommended program for structure solution is xt or shelxt (these are the same program). This program uses a combination of Patterson map and direct methods with some quick structure factor calculations to produce a model. Note that several space groups may be considered, although only the space groups with the same crystal system and cell centering are attempted.
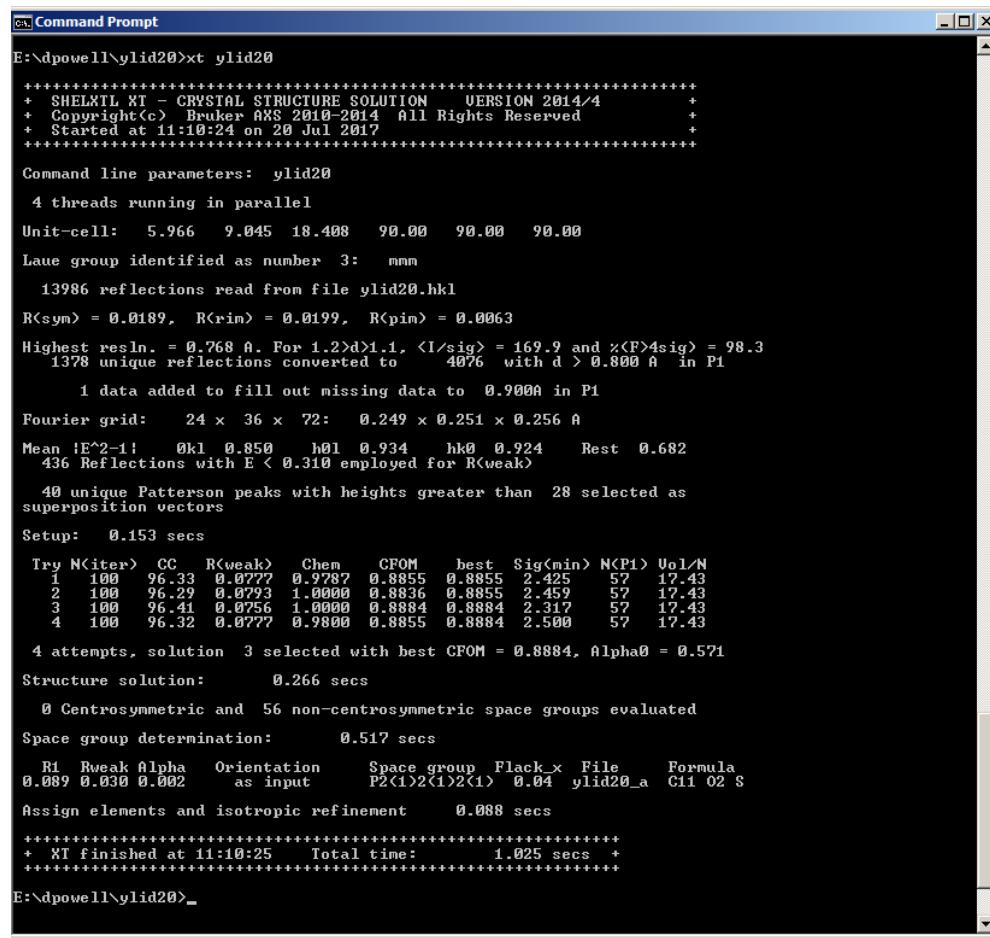
```
Command Prompt                                                          _ | □ | X

E:\dpowell\ylid20>xt ylid20

++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
+   SHELXTL XT - CRYSTAL STRUCTURE SOLUTION    VERSION 2014/4          +
+   Copyright(c)  Bruker AXS 2010-2014  All Rights Reserved            +
+   Started at 11:10:24 on 20 Jul 2017                                 +
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Command line parameters:  ylid20

 4 threads running in parallel

Unit-cell:   5.966    9.045   18.408    90.00    90.00    90.00

Laue group identified as number  3:    mmm

  13986 reflections read from file ylid20.hkl

R(sym) = 0.0189,  R(rim) = 0.0199,  R(pim) = 0.0063

Highest resln. = 0.768 A. For 1.2>d>1.1, <I/sig> = 169.9 and %<F>4sig) = 98.3
   1378 unique reflections converted to    4076  with d > 0.800 A  in P1

      1 data added to fill out missing data to  0.900A in P1

Fourier grid:    24 x  36 x  72:   0.249 x 0.251 x 0.256 A

Mean :E^2-1:    0kl  0.850     h0l  0.934     hk0  0.924     Rest  0.682
   436 Reflections with E < 0.310 employed for R(weak)

   40 unique Patterson peaks with heights greater than  28 selected as
superposition vectors

Setup:   0.153 secs

 Try N(iter)  CC   R(weak)   Chem    CFOM    best  Sig(min) N(P1) Vol/N
   1   100   96.33  0.0777  0.9787  0.8855  0.8855  2.425    57   17.43
   2   100   96.29  0.0793  1.0000  0.8836  0.8855  2.459    57   17.43
   3   100   96.41  0.0756  1.0000  0.8884  0.8884  2.317    57   17.43
   4   100   96.32  0.0777  0.9800  0.8855  0.8884  2.500    57   17.43

 4 attempts, solution  3 selected with best CFOM = 0.8884, Alpha0 = 0.571

Structure solution:       0.266 secs

  0 Centrosymmetric and  56 non-centrosymmetric space groups evaluated

Space group determination:       0.517 secs

  R1   Rweak Alpha  Orientation    Space group  Flack_x  File     Formula
0.089 0.030 0.002     as input     P2(1)2(1)2(1)  0.04  ylid20_a  C11 O2 S

Assign elements and isotropic refinement      0.088 secs

++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
+   XT finished at 11:10:25    Total time:       1.025 secs  +
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

E:\dpowell\ylid20>_
```

The programs xt or shelxt require no additional input from the user other than a *project*.hkl file and a *project*.ins file that contains the wavelength of the radiation, the cell parameters, space group symmetry, and atom types and numbers. This program will also handle non-merohedrally twinned data sets, where the multiple domains are still retained in the data set. Because different space groups are tried the output files are renamed as *project*_a.res (and *project*_a.hkl), *project*_b.res (and *project*_b.hkl) … , where the _a, _b, … refer to different space groups.

Other programs for solving structures include the older xs or shelxs, that solves structures by either direct methods or Patterson (single superposition) methods, patsee a molecular replacement program. The program olex2 contains a link to a charge-flipping routine. The program shelxd uses a dual space method to solve macromolecular structures or structures with poor resolution.

Model Building

This is most easily accomplished using a graphics program such as shelxle, olex2, or ortep3. The shelxle offers simple options to not only build the molecules, but also modify the current output data file, and refine the structure, so this program will be described.

The input file is read in through the ".res" button. Some atom types may be incorrect. Most of the atoms in the structure should be shown.
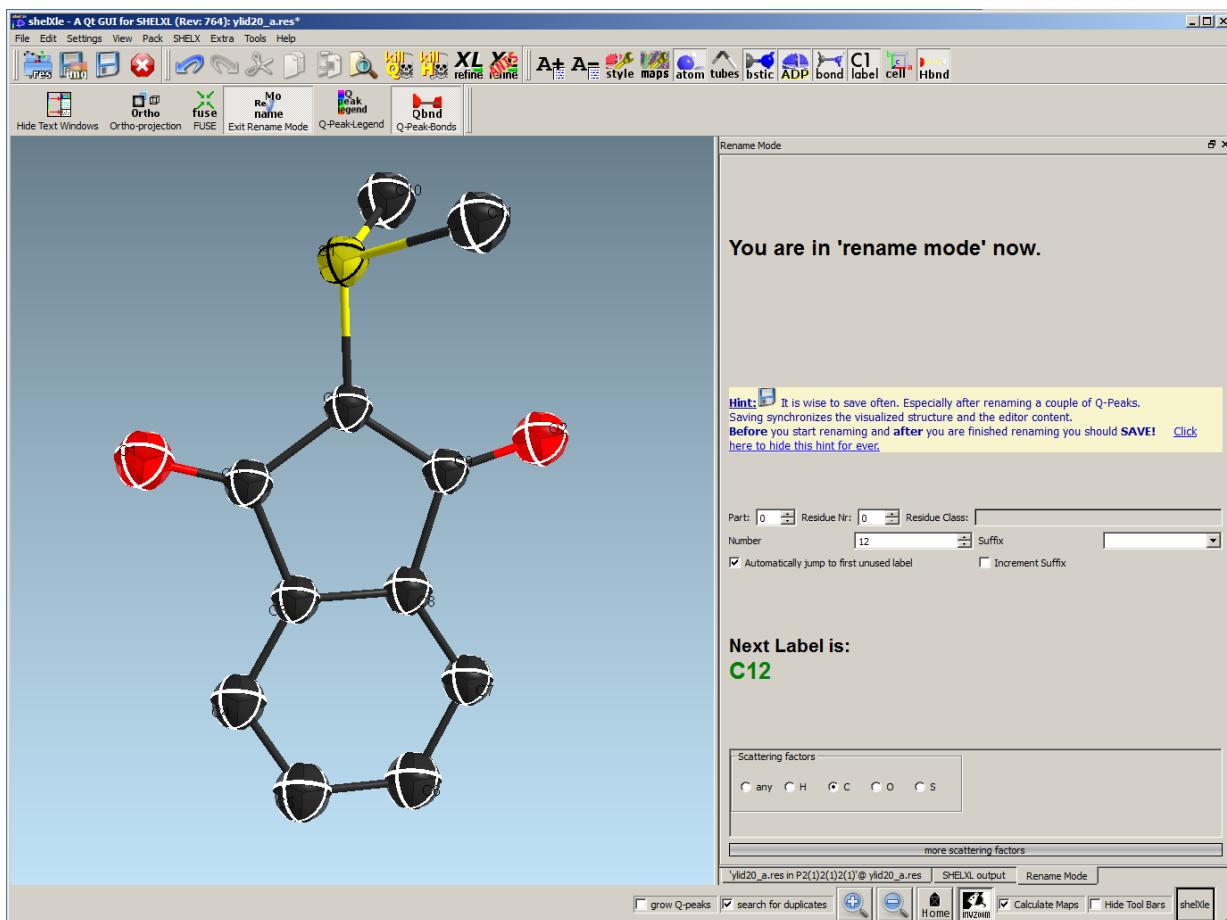
Select the rename mode and in the right-side menu select the "Automatically jump to first unused label" check box. Then select the atom type to label and, if needed, select a letter label to append to the atom number, and begin selecting atoms to change their labels. Repeat this process for all atom types. The orientation of the view can be changed by moving the mouse to any point on the display away from an atom, holding down the left mouse button, and moving the mouse to rotate the view.
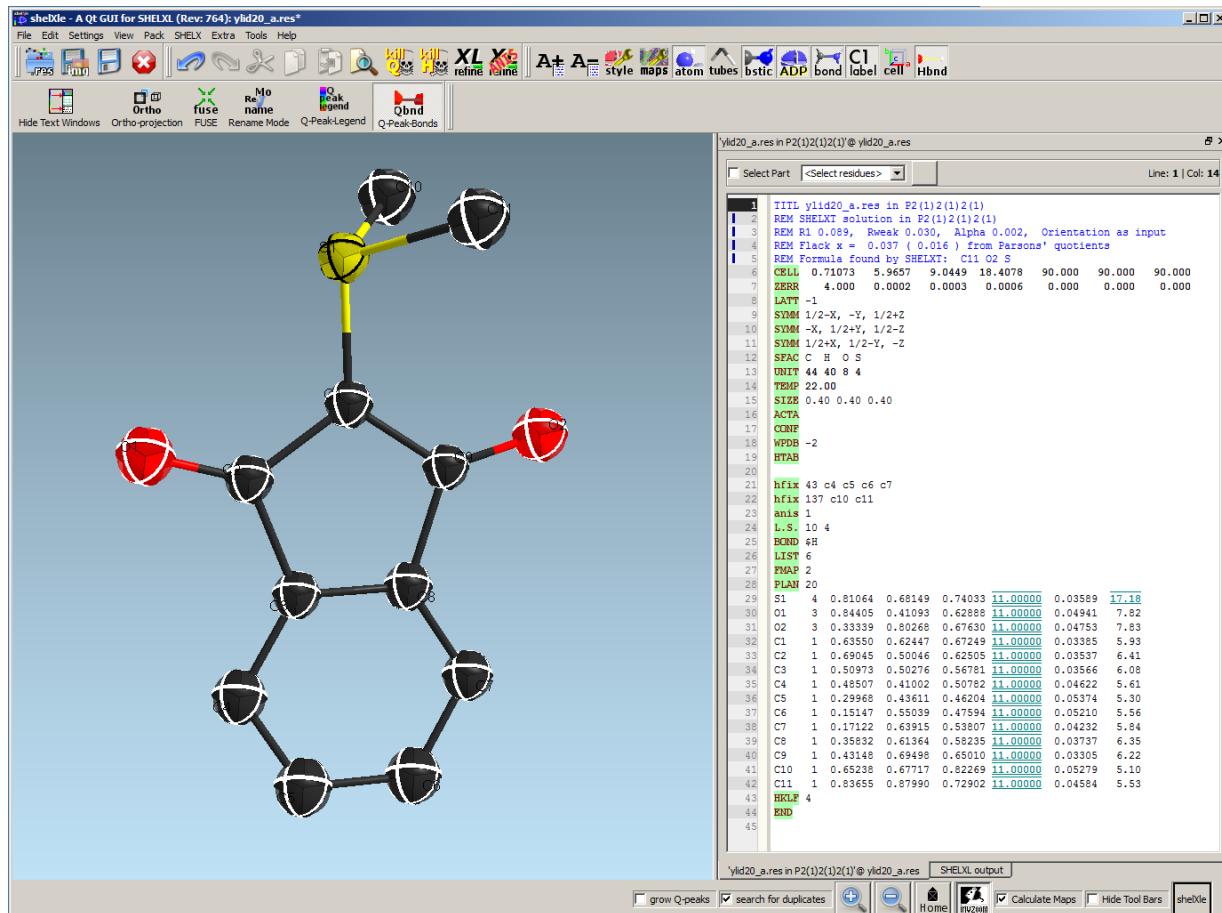
After labeling all atoms, select SHELX > Sort atoms. Add the following lines to the *.res file between the Unit instruction and the atoms:

ACTA
CONF
WPDB -2
HTAB

Then select the SHELX > Copy current file and refine button. The refinement run is shown in the right-side panel. Click the "Load refinement results" button.

Instructions for finishing refinement are in the next document.

References

1. A. D. Mighell, *J. Appl. Cryst.*, **1976**, *9*, 491-498.
2. B. Gruber, *Acta Cryst.*, **1989**, A*45*, 123-131.
3. APEX3, **2007**, Bruker AXS Inc., Madison, Wisconsin, U.S.A.
4. Table 9.2.5.1, *Int'l Tables for Crystallography*, Vol A, **2005**, 5th Ed., Springer, 753.
5. Table 4.3.2.1, *Int'l Tables for Crystallography*, Vol A, **2005**, 5th Ed., Springer, 63-69.