

## SCXRD: Producing the crystal structure

Software: Olex2<sup>1</sup>

ENSURE YOU HAVE PROCESSED THE DATA USING XPLAIN ON THE DIFFRACTOMETER COMPUTER BEFORE YOU START THIS.

There are four computers set up in 30:1095 with the appropriate software needed for structure solution and refinement. These are primarily for solution and refinement as these are the only available computers with the necessary software on. Other computers (university-based) are available for viewing structures, analysing data, database work and literature searching (the latter can also be done at home via the VPN). It is also possible to download this from the internet for free (<http://www.olex2.org/download/>).

When solving the structure you should note down what YOU did - the parameters and commands used and the outcome of these showing how the structure was solved and refined to give an accurate representation of the chemical. It is not crucial that you get the 'perfect' structure, but if the data is not as good as you may wish to get the best structure that you understand where your model falls down and be critical of this.

You should attach your final .cif file for each crystal structure, and write a short description of the structure including reasons why you believe, and can show, it is a true and accurate 3D representation of the compound. Furthermore you should include the unit cell and space group for each polymorph and what these tell you about the two structures.

You will need to download your .hkl and .ins files from LabTrove. To do this, click on the attached file to open a pop-up window. Right click on 'HKL' under 'Download as' then select 'save as'. You can then save the file in the location of your choice and it will be in the correct format. You must download and save both the .hkl and .ins files, ensure they have the same file name stem and are in the same (new) folder.

### **SOLVING AND REFINING**

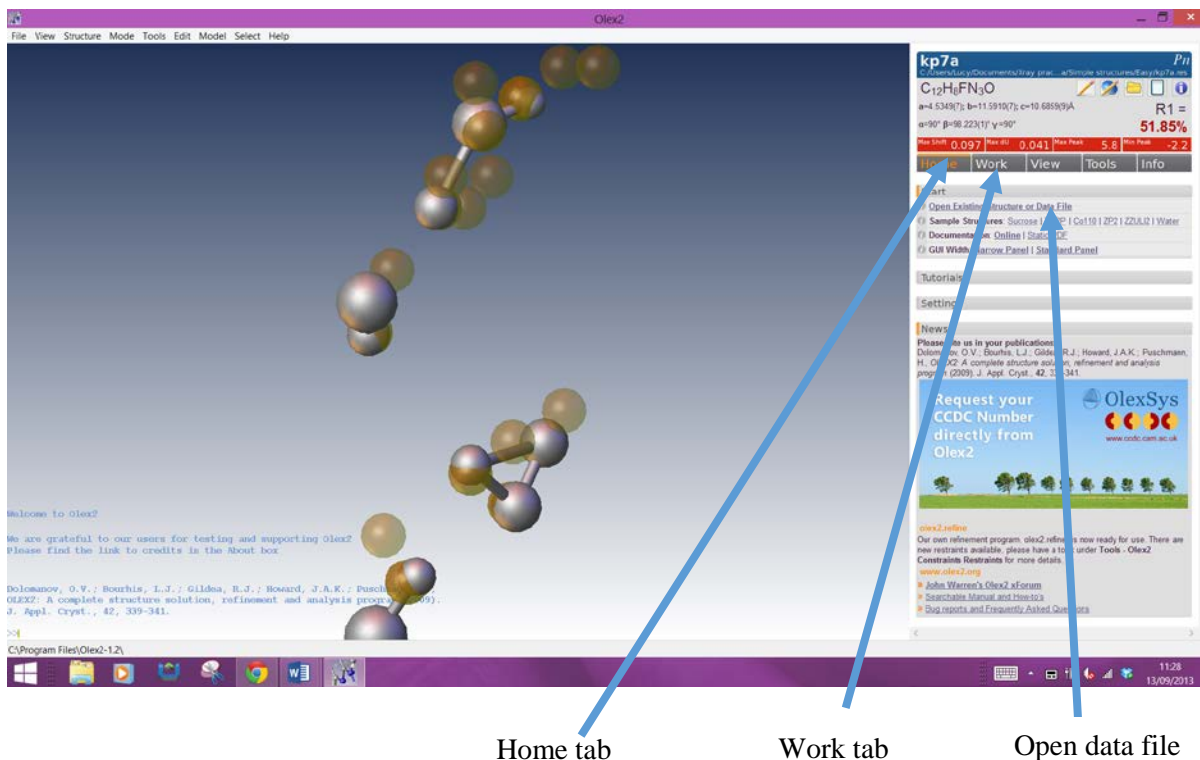
The basic sequence once you have corrected data from the diffractometer is:



This is elaborated step by step below. Note that this is an iterative process and it is not essential to get everything assigned first time round, nor for assignments to be correct. You will complete a number of refinement cycles before you obtain anything similar to your final model.

#### **Open Olex<sup>2</sup> software**

Click the 'Home' tab on the right hand side and under 'Start' open the required file navigating to the .ins file written from Xplain in the previous step.




When the .ins file is open (the display screen (blue background) will be blank) click the work tab then the arrow next to solve to show the options for solving the structure. Click the 'solve' button and this will then show brown spheres indicating electron density peaks (Q Peaks) as seen below.




You can see the peak height of each peak by either hovering the mouse pointer over the sphere or by using the 'info' tab and selecting electron density peaks. Peak height is related to the atomic mass relative to the other atoms in the structure ie larger Q-peaks relate to heavier atoms with higher atomic masses.

Under this menu you can also use the sliders to view only the strongest or weakest peaks, filtering out others (left side filters out most intense peaks, right side filters out weakest peaks). Commands will be applied only to those selected, eg delete, select and name \$Q C.


To view the intensities of the Q peaks, using the 'Toolbox work' tab halfway down the work panel, select 'Q Peak intensities' from the labels drop down list. Also in the toolbox work area, the 'Q'

symbol  can be used to change the display between the Q peaks, no Q peaks and bonded Q peaks (or press Ctrl+Q).

Note – not all Q peaks are necessarily real atoms therefore do not worry if some bonds do not seem to make chemical sense. If the Q peaks do not form a complete structure ie some are isolated or not sensibly connected, type “compaq –a” or use the  button in the right hand panel.

You then need to assign atom types to the Q peaks. Select atoms you believe to be of the same atom type (eg Carbon) which will turn green and then either type “name C” (for Carbon) or click the C/H/F/N... buttons on the right hand side. Try to assign as many atom types as possible – it should be possible to locate all C and O atoms-look at the overall shape of the Q peaks rather than focussing on their intensities. Note also that there will be some noise in the data which will appear as small, feint Q peaks. These will remain throughout structure solution and should not be assigned to any atom type. Trust your chemical judgement and instinct!!


Once you have done the initial assignment, refine the structure to see how well the assignment fits the data and improve the model. Click the arrow next to the refine button to show the options, then click ‘refine’ when you are happy. Refining calculates structure factors based on the current structural model and then compares these to the experimentally measured values. You should refine every time you make a change (a bit like saving a document) as this updates the structure and ensures you are working with the most up to date model.

If a peak, after refinement, appears comparatively a lot larger than the others then this is not real and can be deleted. Any peaks now greater than ~3 units on the difference map are missing C or O atoms. Assign these then refine again. Continue this until there are no more large Q peaks left. Once all atoms are assigned to an atom type (and no large Q peaks left unsigned), the Q peaks can be hidden using Ctrl+Q or the Q peak icon . It should be relatively easy to locate all the heavy atoms in this structure. Remember, the R factor is not the only indication of the structure’s quality. Your structure should make chemical sense, ie the configuration and connectivity should be chemically correct (atom valence etc) with reasonable geometry (bond angles and lengths).

Rotate the structure and compare the size of the spheres for the atoms. They should all be approximately the same size – if not, it is possible that you have incorrectly assigned an atom type. These are currently modelled isotropically with one parameter defining the displacement parameter. The atom spheres should be approximately equal although exceptions do exist, for example in long alkyl chains such as butyl – the sphere size will generally increase along the chain. Disorder can also affect the size of the spheres.



If a sphere appears too small in comparison to others in the structure it is a heavier atom (more electron density) than has been assigned (eg an oxygen assigned as a carbon) and if a sphere appears too large the opposite is true, the atom assigned is too heavy for the atom actually present. If spheres have become too small to see, draw a box (Shift+left mouse) around an atom to select it. Repeat the cycles of altering atom types and refining until all spheres are a similar size. At this point the R1 factor (top right) should be about 8-9% the wR2 ~25%, difference map max ~0.7 e Å<sup>-3</sup> and min ~ -0.6 e Å<sup>-3</sup> (these values may be higher if your data is not the best quality but this is not a problem, equally they may be lower if you have good quality data). Ensure all the atom types are correct and continue.

When you have got to this stage, you then need to model the atoms anisotropically (as opposed to isotropically). This will convert the spheres into ‘rugby ball shapes’ allowing for atomic displacement and give 6 parameters for atom positioning rather than one. This should further reduce the R factor

and increase the accuracy of the model. Click the rugby ball-shaped icon on the right hand side  and Olex will automatically model the atoms anisotropically.

The rugby balls (like the spheres before) should be of equal size and shape. If there are any incorrect atom assignments, they will be displayed as tetrahedral-shaped atoms if the atom should be heavier than the assignment and larger ellipsoids if the atom should be lighter than the assignment given. If reassignment is required it is best to reset them to spheres (select the sphere next to rugby ball icon) to give the isotropic atoms again, make the adjustments and refine before returning to the anisotropic rugby ball atoms. After this the R1 should be ~7-8% (this may be lower for a good data set), wR2 ~20-21% and difference map max  $\sim 0.7 \text{ e } \text{\AA}^{-3}$ , min  $\sim -0.4 \text{ e } \text{\AA}^{-3}$ . Note these are guide values and every data set is different. The lower your values the better, and as long as they keep decreasing with each step of the refinement you are on the right tracks.

Next, the hydrogen atoms should be added by clicking on 'Add H' next to the anisotropic and isotropic icons. Hydrogen atoms will automatically be added and included in subsequent refinements.

The H button  can be used to toggle between including hydrogen atoms, showing hydrogen atoms with hydrogen bonding (red dashed line) and excluding hydrogen atoms. Ensure that no extra hydrogen atoms have been added. If so, delete all hydrogen atoms on the parent atom by selecting atoms and 'delete'. Refine the structure again. Note, if some hydrogen atoms are selected and the delete all H atoms button  or command 'kill \$H' is typed, only the selected hydrogen atoms will be deleted. Repeat the hydrogen atom addition as before and only the missing hydrogen atoms will be added. If there are still problems, or hydrogen atoms missing after using the Add H command these must be added manually.

Some hydrogen atoms may not be correctly placed (especially OH groups) and therefore it is necessary to look at the Q peaks to see where the electron density is located. To optimise the positions either multiple refinement cycles can be conducted and the H atoms may rotate onto the Q peaks. Alternatively, the hydrogen in question can be selected then Toolbox work  $\rightarrow$  Select group or atoms(s) and then...  $\rightarrow$  Fit. Once in the 'Fit' mode, hold Shift down and this allows you to click and drag the hydrogen atoms to the desired position using the left mouse button. When finished press Esc button to leave the 'Fit' mode. After this, refine the structure again as this will adjust hydrogen atom positions to maintain the geometric constraints. To check the positions of the hydrogen atoms, the structure can be rotated and viewed from all angles to ensure the positions all look sensible.

Next, use the 'Weights' option under refinement options (ensure the arrow next to 'refine' is in orange and pointing upwards to be able to see the refinement options). Tick the 'Auto update weights' box and Refine several times until the weights do not change any more, and should turn green. This is the solved and refined structure.

Ensure the 'CONF, MORE -1, Bond \$H, ACTA' box is selected near the top of the refinement display menu as this will produce your final .cif file.

An extinction correction may be required (the lines of text in the refinement cycles may indicate this). To do this, tick the box next to extinction correction in the top section of the refine menu (under Weights), then do a refinement cycle.

Note: each time you make adjustments and refine, the quality and accuracy of the structure should be increasing and the R factor decreasing. By the end of the above cycles the R factor should be between 4-5% (or below for a good structure), wR2 around 10% and the largest Q peak in the difference map

about 0.3 units. However, if the data are not good enough, these values will not be possible. Your final R factor will be determined by the Rmerge in data processing by the computer. Everything in the top right of the toolbar should turn green by the end of a good structure solution and refinement. Increasing the number of cycles in each refinement can speed up this and ensure everything is set to the correct position and values before finishing the structure.

### TIDYING THE STRUCTURE

To get the best structure, and most accurate representation of the compound there are certain values which should be monitored throughout the solving and refining process and can be found under the 'Info' tab and 'refinement indicators' header.

These are:

R1

wR2

Difference map max and min

GooF

Hooft y

Flack x

R1 and wR2 are an indication of the agreement between the observed and calculated structure factors ( $F_o$  and  $F_c$ ). These will reduce as the refinement progresses and the agreement improves. High values (>40/70 respectively) indicate that the solution is incorrect.

Difference map max/min gives information on the highest and lowest peaks and holes in the electron density map.

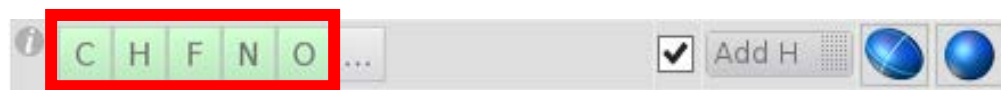
GooF is the goodness of fit and should converge towards 1 towards the end of the refinement.

The last two, Hooft y and Flack x are applicable only for non-centrosymmetric space group structures (with no inversion centre). They are displayed with an error and should be close to 0 if the solution is correct. A value ~1 indicates that the structure should be inverted and those very different to 0 or 1 show a racemic twin.

Check the chemical composition is correct with no missing or extra atoms. To amend the chemical formula to concur with that in the structure displayed click the 'OK' icon and refine again.



When the structure and chemical formula are in agreement, the atom type buttons will turn green



Next, go to the 'Info' tab and to the 'Bad Reflections' header tab. If any reflections have large  $|F_o^2 - F_c^2| / \text{esd}$  values they may need to be omitted and will be obvious as they will be very different to the rest of the values. If more than several are needing omitting they may require investigating.

### CHECKING THE STRUCTURE

General checks that should be made include:

Visual check of the ellipsoids to ensure they are of similar size and proportions and look reasonable.

Check the number of hydrogen atoms present overall and on each parent atom.

Under the 'Info' tab, go to the 'Refinement Indicators' header and any errors in the parameters will be displayed in orange or red (depending upon the severity).

Check bond lengths and angles are sensible

Checking bond lengths and angles:

To view the bond lengths –

- 1) Hover mouse pointer over a bond and the length will be displayed or select the bond then right click to display the properties.
- 2) Select the involved atoms (two for bond length, three for bonded angle or four for torsion angle) then use the 'View' tab and 'Geometry' header tab to show the 'Distance and Angles' of the selection.
- 3) Right click on the molecule, Fragment → Select bonds with select all bonds in the molecule. Use the 'Tools' tab and Images header to select 'Label selected atoms'. To remove these labels you must be in Tools→Images. Use Ctrl+A to select all labels then delete. If atoms are selected, Ctrl+I will invert the selection and the labels can then be deleted.

Comparisons and checks of bond lengths and angles can be made using the Cambridge Structural Database to ensure the values are sensible and within expected ranges.

Uncheck the 'merge CIF' box in the drop down report menu (arrow next to the title).

Once finished, you should check your .cif file to ensure there are no glaring errors with any part. To do this, go to <http://checkcif.iucr.org/> and upload your .cif file for checking. Choose the 'validation of CIF only (no structure factors)'. Alternatively, in olex2 you can click the 'checkcif report' button (report drown down menu) and this will automatically send the cif to checkcif.

The screenshot shows the checkCIF website interface. On the left, there is a form with the following fields and options:

- checkCIF is sponsored by Crystallography Journals Online (with logos for IUCr, WILEY, and ROYAL SOCIETY OF CHEMISTRY).
- A service of the International Union of Crystallography.
- checkCIF reports on the consistency and integrity of crystal structure determinations reported in CIF format.
- Please upload your CIF using the form below.
- File name: Choose File (No file chosen)
- Select form of checkCIF report:  HTML  PDF
- Select validation type:  Full validation of CIF and structure factors  Validation of CIF only (no structure factors)
- Output Validation Response Form:  Level A alerts only  Level A and B alerts  None
- Send CIF for checking
- Information about this version of checkCIF ...
- Useful links: Prepublication check for submissions to IUCr journals, Details of checkCIF/PLATON tests, CIF dictionary, Download CIF editor (pubCIF) from the IUCr, Download CIF editor (enCIFer) from the CCDC.

Two blue arrows point from the right side of the image to the form fields:

- One arrow points to the 'Choose File' button, with the text 'Upload file here' next to it.
- Another arrow points to the 'Validation of CIF only (no structure factors)' radio button, with the text 'Validation type' next to it.

This will produce a guidance report with alerts of varying levels indicating areas that may need looking at or reviewing. Alert level A's should always be resolved or have good reason behind their occurrence. B and C levels alerts should also be resolved if possible.

Information about the crystal and data collection can be added via olex2 under the various headings in the report tab.

The final checkcif report can be downloaded as a pdf file and uploaded to LabTrove along with your .cif file, ensuring you comment on anything you believe to be significant or explain why there are certain alerts.

### Viewing the Crystal structure:

Mercury<sup>2</sup> (from CCDC, on all university computers) is sufficient to view, compare and analyse crystal structures (opens the .cif file). There are other software packages available which have additional/alternative features should you wish to use them, such as Diamond, available under 'Your School Software' on university computers.

See the 'Using Mercury' guide for features such as overlaying structures, viewing the packing and interactions in a crystal structure, measuring distances and bond lengths as well as how to generate a calculated powder pattern from the crystal structure.

1. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
2. C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466-470.