P. Macchi, J. Hauser, L. Dos Santos, M. Fisch
Chemical Crystallography, www.chemcrys.unibe.ch
Department of Chemistry and Biochemistry, University of Bern, Switzerland

Urea Structure Solution Tutorial

Indexing, unit cell refinement, structure solution and subsequent refinement:

This tutorial demonstrates

(1) Indexing of a powder pattern of a substance,
(2) Unit cell refinement after successful indexing,
(3) Structure solution using the simulated annealing method,
(4) Subsequent structure refinement, next exercise.

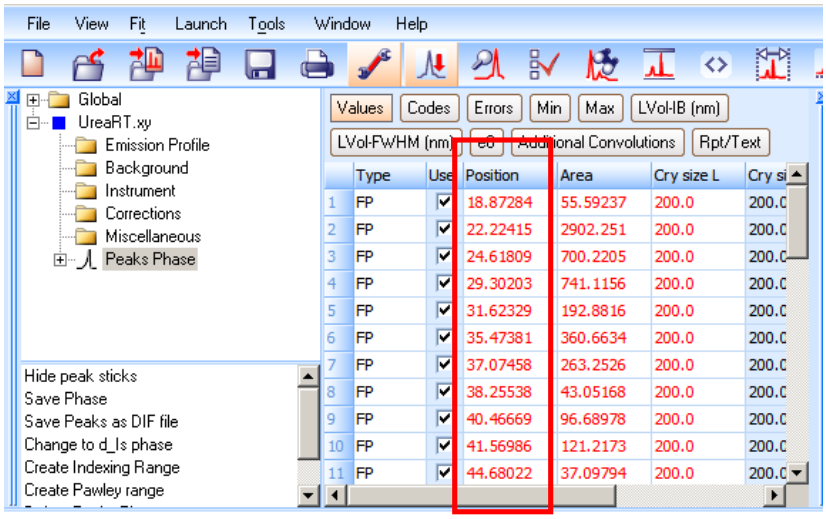
(1) Indexing

1.1 Create a list of observed peak positions (+2θ)

Load the diffraction pattern (*UreaRT.xy*) in TOPAS (*File → Load Scan Files*). Open the *Search Peaks* dialog (*View → Search Peaks*). Deselect the checkbox *Remove K-Alpha 2 Peaks* and adjust the *Peak Width* and *Noise Threshold* sliders to automatically identify peaks in the diffraction pattern. When finished, click *Add Peaks*.

A list of peak positions can be found in the *Peaks Phase* entry of the *Parameters* Window.

Peaks can be manually added to the list by zooming, and ctrl-click in the diffraction pattern or removed by selecting them in the list and pressing the delete key.



Right-click on the *Position* header and *copy all/selection*.

1.2 Set up an input file for indexing

In jEdit, create a new *index.inp* file using the following keywords from the TOPAS menu in jEdit

- **`seed`**, feeds a random number generator

- **`index_lam 1.540596`**, used for conversion of d and 2θ with Bragg's law

- **`index_zero_error`**, includes a zero shift

- Include all available Bravais lattices in the indexing procedure

- **`load index_th2`** and paste the peak position list copied from TOPAS between { }.

```
◇ Index.inp
 1 seed
 2 index_lam 1.540596
 3 index_zero_error
 4 Cubic_F
 5 ...
 6 Monoclinic_P
 7 Triclinic_P
 8
 9 load index_th2 {
10 18.87284
11 22.22415
12 24.61809
13 ...
14 101.0435
15 105.1219 }
16
```

Running this file with TOPAS will create an *index.ndx* file containing the indexing solutions:

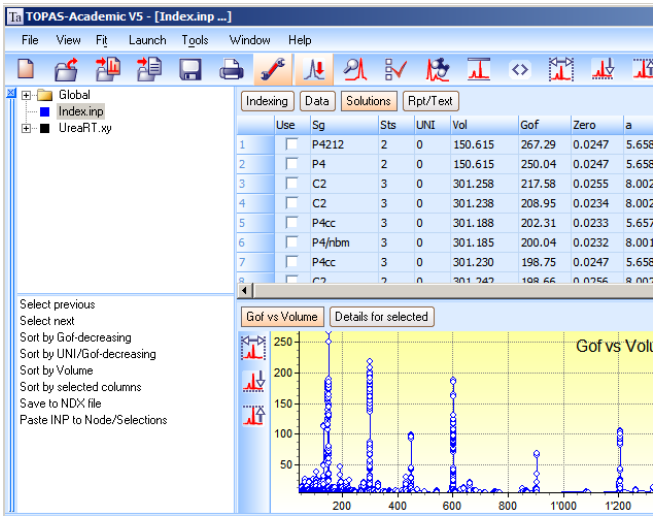
```
' Indexing method - Alan Coelho (2003), J. Appl. Cryst. 36, 86-95
' Time: 13.619 seconds
' Sg     Status UNI      Vol       Gof     Zero      Lps...
Indexing_Solutions_With_Zero_Error_2 {
   0) P4212      2   0     150.615   267.29   0.0247     5.6582     5.6582     4.7046     90.000     90.000     90.000 ' ===  24  20
   1) P4         2   0     150.615   250.04   0.0247     5.6582     5.6582     4.7046     90.000     90.000     90.000 ' ===  24  20
   2) C2         3   0     301.258   217.58   0.0255     8.0025     8.0018     4.7046     90.000     90.016     90.000 ' ===  20  21  20  15
   ...
```

1.3 Checking the solutions

Open the *index.inp* file in TOPAS (*File → Load Input (INP) files*) and load the scan file (*File → Load Scan files*).

In the parameters window, select the *index.inp* entry and click *Solutions*. Copy the content of the *index.ndx* file and select *Paste INP to Node/Selections*.

Goodness of fit values of the indexing solutions are plotted against the unit cell volume.



Browse the solutions and compare their peak positions to those identified in the pattern.

Question: Which solution is most likely the best one? Why?

Question: Is the space group correct? If not, why? How can the correct space group be determined?

(2) Unit cell refinement

With known unit cell parameters and space group (from indexing), reflection intensities of a powder pattern can be extracted.

Different methods exist, of which the Pawley method allows refinement of each $M*I_{hkl}$ as independent parameter, combined with refinement of e.g. specimen displacement, lattice parameters, peak shape, background, etc. These parameters are later required for structure solution.

Set up a *Pawley.inp* file as for a Rietveld refinement but use **hkl_Is** instead of **str** for starting the phase section of the input file.

```
hkl_Is
phase_name "UreaRT_Pawley"
a .....
b .....
c .....
prm !eta  0.65921
prm !U    0.00295
prm !V   -0.00114
prm !W    0.00106
peak_type pv
pv_lor =  eta;
pv_fwhm = Sqrt(W + Tan(Th)*V + Tan(Th)*Tan(Th)*U);
CS_L(.....
CS_G(.....
space_group "....."
```

Question: Why are no atomic sites required for Pawley refinements? Why is no scale factor refined?

Question: What is the fundamental difference between a Rietveld and a Pawley refinement?

Question: What is the meaning of agreement factors (R-values) of a Pawley refinement? Why can the Pawley R_{wp} be relevant for structure solution?

(3) Structure Solution

Due to the restricted resolution in many powder XRD experiments and the "lost" information due to M and peak overlap, structure solution methods such as Patterson or direct methods usually fail.

Often successful for structure solution from powder data is the simulated annealing method. It can be thought of as a process, in which a crystalline material "solidifies" by slow cooling of a melt. For a given "temperature", a random model is refined. After convergence, parameters are randomized again, allowing to move out of local minima. "Temperature" decreases with increasing program cycles and thus randomization becomes smaller.

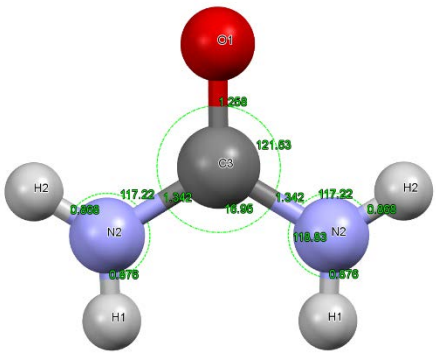
This method is especially powerful if the cell content (fragments or whole molecules of the structure) is roughly known, as in this case the urea molecule measured at low temperature. The molecule can therefore be defined as "rigid body" for which only rotation and translation parameters are refined in the simulated annealing process.

A rigid body can be defined using the z-matrix formalism. Each line of a z-matrix gives an internal coordinate for one of the atoms within the rigid body. The syntax is as follows:

z_matrix [Site1] [Site2] [bond-length] [Site3] [bond-angle] [Site4] [dihedral-angle]

With values measured in Mercury*, the z_matrix for urea at 90K is therefore:

```
rigid
z_matrix O1
z_matrix C3    O1    1.258
z_matrix N2_A C3    1.342 O1     121.53
z_matrix H1_A N2_A 0.876 C3     118.83 O1 180
z_matrix H2_A N2_A 0.868 C3    -117.22 O1 180
z_matrix N2_B C3    1.342 N2_A  116.95 O1 180
z_matrix H1_B N2_B 0.876 C3     118.83 O1 180
z_matrix H2_B N2_B 0.868 C3    -117.22 O1 180
```



The rigid body (*UreaLT.rgd*) can be viewed and modified in TOPAS in the rigid body editor (*Tools → New Rigid-body editor window*).

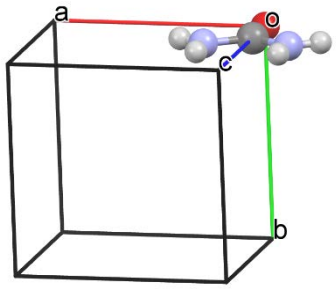
O1 is the origin of the rigid body, C3–O1 is along the c-axis, and C3–N2 lies in the ac-plane. Therefore, in an orthogonal system,

```
rotate rot_a 0 qa 1 qb 0 qc 0
rotate rot_b 0 qa 0 qb 1 qc 0
rotate rot_c 0 qa 0 qb 0 qc 1
```

allows rotation of the molecule with O1 as center of rotation around the a-axis (rot_a), b-axis (rot_b) and c-axis (rot_c).

O1 (and consequently the whole rigid body) can be translated along a (t_a), b (t_b) and c (t_c) with:

```
translate
ta t_a 0.01
tb t_b 0.01
tc t_c 0.01
```



As a z-matrix only defines the intramolecular geometry, each site requires an entry in the conventional list of sites:

```
site O1   x 0.01 y 0.01 z 0.01 occ O 1 beq 4
site C3   x 0.02 y 0.02 z 0.02 occ C 1 beq 4
site N2_A x 0.03 y 0.03 z 0.03 occ N 1 beq 4
site H1_A x 0.04 y 0.04 z 0.04 occ H 1 beq 4
site H2_A x 0.05 y 0.05 z 0.05 occ H 1 beq 4
site N2_B x 0.06 y 0.06 z 0.06 occ N 1 beq 4
site H1_B x 0.07 y 0.07 z 0.07 occ H 1 beq 4
site H2_B x 0.08 y 0.08 z 0.08 occ H 1 beq 4
```

Fractional coordinates of the sites are defined by the rigid body (and its translation and rotation). **Note that in TOPAS V4 and V5, start coordinates should not be set to 0.**

Due to the twofold symmetry of the molecule and the space group symmetry, it is likely that sites become mapped on themselves. Their occupancy, when initially set to 100% will therefore be overestimated. With

```
occ_merge X*  occ_merge_radius 1
```

occupancies of all X* sites (X = site, X* includes all X1, X2, X3, etc.) will be merged if they are close within a defined radius [Å]. Occupancies below 100% may therefore indicate that a site lies close to or at a special position.

Input file

An input file for simulated annealing is similar to a Rietveld refinement input file. Known parameters such as background, specimen displacement, lattice parameters and peak shape should be used from the Pawley refinement and subsequently be fixed.

The only refined parameters are therefore

rot_a, **rot_b**, **rot_c**, **t_a**, **t_b**, **t_c** and **scale**.

The simulated annealing procedure can be initiated using the `Auto_T(2)` macro at the beginning of the input file.

At the end of the input file, the following commands can be added:

`Out_CIF_STR("Urea_RT_Solved.cif")` outputs a .cif file of the accepted solution,

view_structure shows the structure during the simulated annealing process.

Question: Is there any way to judge whether a good solution may have been found before the simulated annealing procedure is finished (which can take some time).

Preparing the solution for refinement

To identify special positions, open the obtained *UreaRT_Solved.cif* file in Mercury* and switch on the symmetry elements. Duplicated sites (symmetry of the molecule!) have to be manually removed.

The simulated annealing input file can be switched to Rietveld refinement by removing `Auto_T(2)` and the model manually improved. Monitoring the occupancy factors can be informative.

Final refinement of the structure is analogous to the next exercise.

P. Macchi, J. Hauser, L. Dos Santos, M. Fisch
Chemical Crystallography, www.chemcrys.unibe.ch
Department of Chemistry and Biochemistry, University of Bern, Switzerland

Urea Structure Refinement

Structure refinement:

You have solved the structure of urea from single-crystal XRD data measured at 90 K. *UreaRT.xy* is a powder XRD pattern of urea measured at room temperature. Refine the structure of urea at room temperature by starting from the 90 K structure.

During the refinement, the structure can be visualized in TOPAS by adding `view_structure` to the `str` part. Finally, a .cif file can be written with `Out_CIF_STR("UreaRT.cif")`.

Comment:

Some sites lie on special positions, which might not be recognized at first (e.g. sites on mirror planes but not with obvious coordinates like ½, ½, ½). In contrast to ShelX, TOPAS does not automatically recognize sites on special positions. When refining their positions, symmetry operators must be introduced manually, or the site will slip off the symmetry element. Check the positions e.g. with Mercury* or the International Tables for Crystallography.

Example:

For the following site:

```
site O1 x 0.5000 y 0.0000 z 0.9021 occ O 1.0 beq 4.17
```

a symmetry operator 'x, x+½, z' is introduced:

```
site O1 x = !a1b2c3 0.0000 y = a1b2c3 + 0.5; z 0.9021 occ O 1.0 beq 4.17
```

x is fixed in this case and y is = x + 0.5 but not refined. If you want to refine x, remove the ! before the parameter name (in the example "**!a1b2c3**", can be any word).
The value of an equation can be output with: `= y + 0.5;:0` where 0 is updated at the end of the refinement.

*Mercury can be found here:

http://www.ccdc.cam.ac.uk/solutions/csd-system/components/mercury/