

Crystal structure determination



Crystal structure determination process

Starting knowledge: **powder pattern & chemical formula**

- ☐ Indexing
- ☐ Space Group Identification
- ☐ Structure Solution



- ☐ Reciprocal Space Methods
- ☐ Direct Space Methods

- ☐ Rietveld Refinement

Powder Pattern Indexing

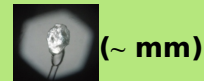
Aim ↔ A geometrical rebuilding of the three-dimensional reciprocal lattice from the experimental diffraction data

Indexing ↔ Assigning triplets of Miller indices (*h k l*) to every observed Bragg peak based on a selected unit cell

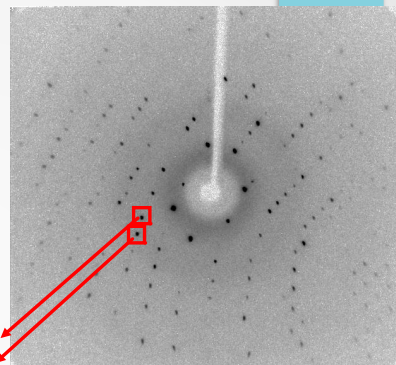
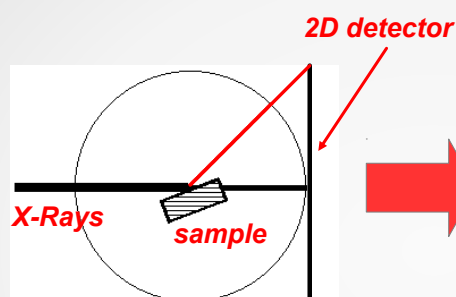
$$\mathbf{d}_{hkl}^* = h \mathbf{a}^* + k \mathbf{b}^* + l \mathbf{c}^*$$

where \mathbf{a}^* , \mathbf{b}^* and \mathbf{c}^* are the basis vectors defining the reciprocal unit cell

Powder vs single crystal methods



(~ mm)

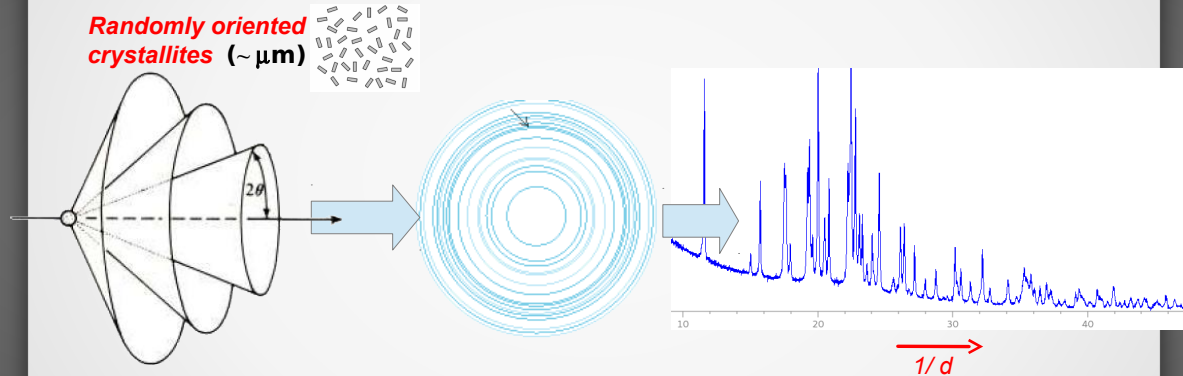


The diffraction effects are well separated,

- The experimental information is three-dimensional
- Lengths and directions of reciprocal vectors, \mathbf{d}_{hkl}^* , are available
- In case of single crystal, indexing is usually a quite trivial task:
 1. The unit cell (and symmetry) can be found unambiguously
 2. Each reflections can be assigned its correct Miller indices, *h k l*
 3. The diffracted intensity for each reflections can be measured with reasonable accuracy

Powder vs single crystal methods

The experimental data are a one-dimensional projection of the three-dimensional reciprocal lattice



Only the lengths, d_{hkl}^* , of the vectors in the reciprocal space are measurable

The indexing may become extremely complicated

The indexing equation

$$d_{hkl}^* = h a^* + k b^* + l c^*$$

$$d_{hkl}^* = \frac{1}{d_{hkl}} \quad a = \frac{b^* \wedge c^*}{V^*} \quad b = \frac{c^* \wedge a^*}{V^*} \quad c = \frac{a^* \wedge b^*}{V^*} \quad V^* = a^* \cdot b^* \wedge c^* = \frac{1}{V}$$



$$d_{hkl} = f(h, k, l, a, b, c, \alpha, \beta, \gamma)$$

Peak search \rightarrow θ_{hkl}^{obs} $\xrightarrow{\text{Bragg's equation}}$ $d_{hkl}^{obs} = \frac{\lambda}{2} \sin \theta_{hkl}^{obs}$

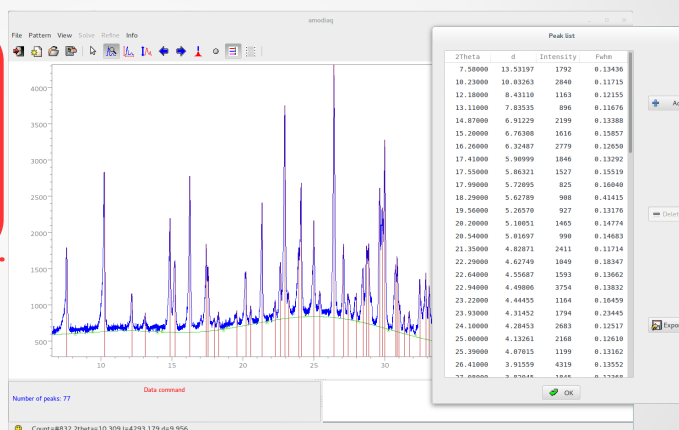
$$d_{hkl}^{obs} \simeq f(h, k, l, a, b, c, \alpha, \beta, \gamma)$$

Peak search

Peak search is needed in order to locate the peak positions

The method is a combinations of:

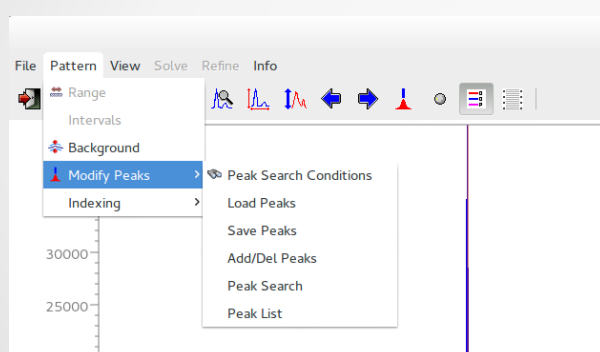
- background subtraction
- K α 2 stripping
- smoothing
- calculation of the derivatives



Peaks at low angle not clearly resolved and with an appreciable peak width should be not included in case they appear strongly uncertain

Peak search

Graphical tools are provided to improve the peak search



Set Search Conditions

Minimum 2-theta: 5.00

Maximum 2-theta: 64.96

% Int. Threshold: 1.50

Number of Peaks: 39

OK Cancel

Figures of merit

The most frequently used figures of merit are M_{20} and F_N

M_{20} has been introduced by de Wolff

$$M_{20} = \frac{Q_{20}}{2|\overline{\Delta Q}|N_{20}}$$

The more reliable indexing yields the higher M_{20}

- $Q = d^*{}^2 = 1/d^2$
- Q_{20} is the corresponding Q-value for the 20th observed Bragg peak
- $|\overline{\Delta Q}| = \frac{1}{20} \sum_{i=1}^{20} |\Delta Q_i| = \frac{1}{20} \sum_{i=1}^{20} |Q_i^{obs} - Q_i^{calc}|$ is the average absolute difference between the observed and calculated Q_{hkl} for the first 20 Bragg peaks
- N_{20} is the number of independent Bragg reflections possible up to the 20th observed diffraction peak

$M_{20} < 10$: it is likely that solution of the indexing problem is not correct

Figures of merit

F_N figure of merit has been introduced by Smith and Snyder.

$$F_N = \frac{1}{|\overline{\Delta 2\theta}|} \cdot \frac{N}{N_{poss}}$$

The best indexing result usually has the highest F_N

- N is the number of the observed Bragg peaks
- N_{poss} is the number of independent Bragg reflections possible up to the N th observed diffraction peak
- $|\overline{\Delta 2\theta}| = \frac{1}{N} \sum_{i=1}^N |\Delta 2\theta_i| = \frac{1}{N} \sum_{i=1}^N |2\theta_i^{obs} - 2\theta_i^{calc}|$ is the average absolute difference between the observed and calculated $2\theta_i$

the F_N usually should be greater than 10

Figures of merit

New figures of merit, exploiting the information contained in the full pattern, have been recently proposed, among them

- **M_{20}** (Le Bail, 2008)
- **WRIP20** (Altomare et al., 2009)

Indexing algorithms

- **Direct-space indexing:** unit cell dimensions are varied

Methods: Grid search method, Monte Carlo search, genetic algorithm, simulated annealing, ...

Applications: AUTOX, EFLECH, Hmap, McMaille, SVD-Index, X-Cell, FOX 34, GAIN.

- **Reciprocal space indexing:** indices h, k, l of low Bragg angle peaks are varied

Methods: trial-and-error, zone-search

Applications: TREOR, DICVOL and ITO

Indexing by Expo2014

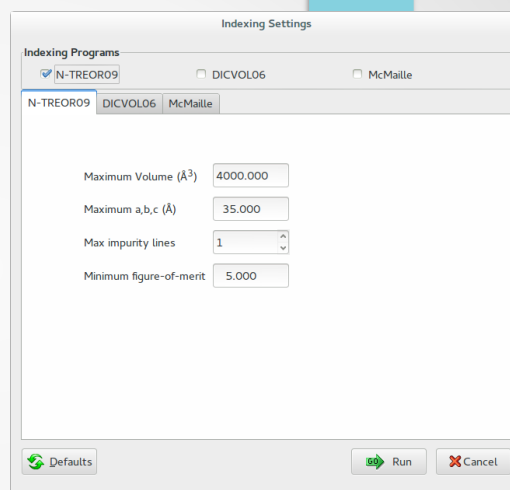
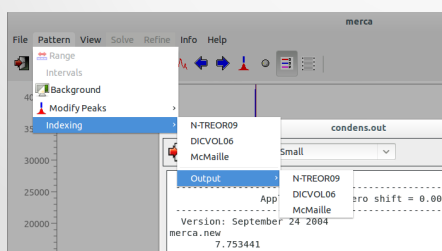
The indexing is performed by **EXPO2014** via the program **N-TREOR09**, the evolution of the **TREOR** software.

- **TREOR is a trial-and-error indexing program**, which is based on the permutation of indices in a selected basis set of lowest Bragg angles peaks.
- **Indexing starts from the cubic crystal system and may proceed through all crystal systems** down to triclinic. When an acceptable solution is found, the program terminates and therefore, indexing in lower symmetry unit cells will not be performed.
- **New features of N-TREOR09:**
 - ✿ **Correct the experimental data by zero-point error** (errors in 2θ peak positions can be responsible for indexing failure);
 - ✿ **Calculate an effective figure of merit (WRIP20) for identifying the correct cell among a set of plausible ones.**

Indexing by N-TREOR09

- Maximum volume and cell dimensions
- Maximum number of impurity lines
- Minimum figure of merit

Output files: condens.out, ntreor.out



DICVOL

DICVOL is an exhaustive trial-and-error indexing program with variation of parameters by successive dichotomy and partitioning of the unit cell volume.

- GUI may also be used to set maximum lengths of the unit cell edges, monoclinic angle and unit cell volume.
- Error in peak positions can be supplied and is assumed identical for every observed peak.
- It is possible to increase a minimum figure of merit, M_{20} (the default is 10)
- Any crystal system can be included or excluded from the indexing process
- Measured density and formula weight
- Maximum number of unindexed Bragg peaks
- Estimate and correct for the presence of zero-shift error

McMaille

- Treor and Dicvol run much faster than McMaille.
- The recommendation would be to start with TREOR or DICVOL, and finally McMaille if no convincing result is obtained.

Errors in powder indexing

The solution is not always straightforward because of several reasons:

- Inaccuracy in peak positions due to:
 1. zero-point error
 2. sample misplacement
 3. low resolution
 4. bad crystallinity
- The presence of impurities providing spurious additional peaks

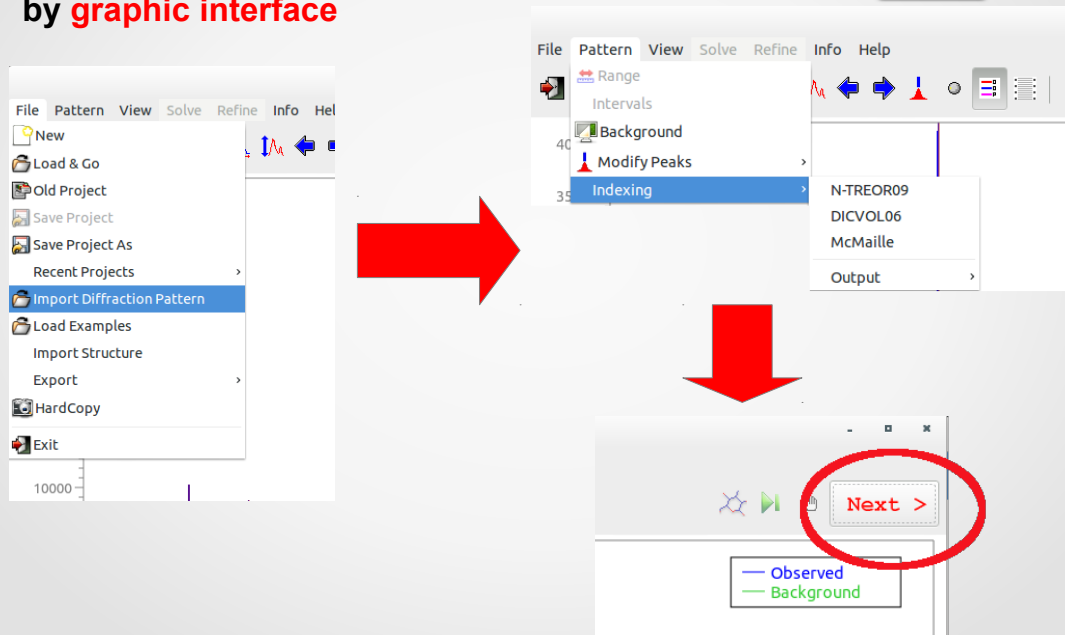
Data quality remains the most important factor in indexing

Some advice

- Do not accept unindexed peaks, unless you are able to explain them
- To obtain a proper solution, it is necessary to eliminate the impurity Bragg peaks from the indexing process
- Every observed peak must correspond to a calculated reflection
- Finally, in the case of a new material, the correctness of the *ab initio* indexing is generally ensured by solving and refining the crystal structure
- Do not waste computer time on bad data

Run Expo2014 for indexing

by **graphic interface**

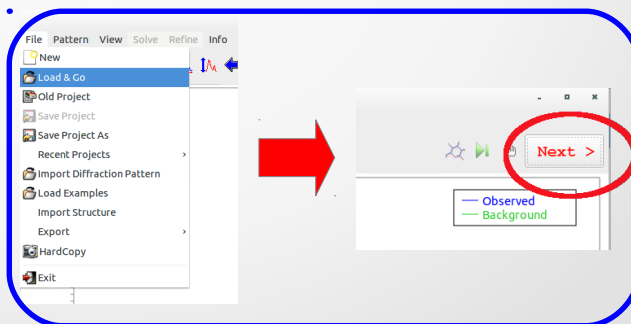


Run Expo2014 for indexing

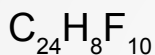
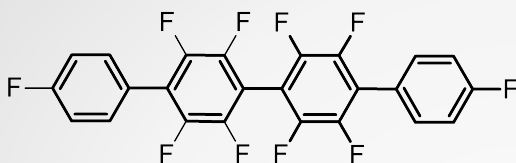
via an **ASCII input file (*.exp)** requiring minimal information and consisting of:

- **commands** (the first character in the line must be '%')
- **directives** (sub-commands following the related command)

```
%structure dfqp
%initialize
%job decafluoroquarterphenyl (C24 H8 F 10)
%data
    pattern dfqp.dat
    range 7.000 70.000
    wave 1.790000
%ntreor
%continue
```



Indexing of decafluoroquarterphenyl



decafluoroquarterphenyl

```
%structure dfqp
%job decafluoroquarterphenyl
%data
  pattern dfqp.dat
  range 7.000 70.000
  wave 1.790000
%ntreor
%dicvol
```

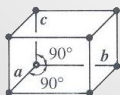
Published cell ^{*}:

a: 24.0519

b: 6.1529

c: 12.4207

β : 102.755

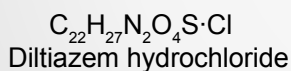


Plausible cell parameters

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	FOMnew	H20	shift	NIX	Symmetry	Info
1	N	24.08344	6.15797	12.43329	90.000	102.698	90.000	1798.8	0.966	20.00	0.000	0	I 1 a 1	
2	N	24.10843	6.16559	12.44866	90.000	102.749	90.000	1804.8	0.759	12.00	0.040	0	I 1 a 1	
3	N	24.06071	6.14751	12.42618	90.000	102.737	90.000	1792.8	0.703	10.00	-0.040	0	I 1 a 1	
4	D	24.08300	6.15900	12.43290	90.000	102.678	90.000	1799.2	-	16.80	0.004	2	Mono	

^{*}Nowell, H., Attfield, J. P., Cole, J. C., Cox, P. J., Shankland, K., Maginn, S. J. & Motherwell, W. D. S. (2002). New J. Chem. 26, 469–472.

Indexing of diltiazem hydrochloride

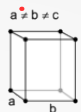


Published cell ^{*}:

a: 42.190

b: 9.075

c: 6.037



Default maximum cell axis value = 35 Å

Change default settings by GUI or adding the directive CEM in the input file

Indexing Settings

Indexing Programs: ☒ N-TREOR9 ☐ DICVOL06 ☐ McMaille

N-TREOR9 | DICVOL06 | McMaille

Maximum Volume (Å³): 4000.000

Maximum a,b,c (Å): **50.0**

Max impurity lines: 1

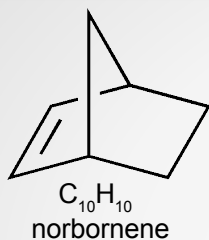
Minimum figure-of-merit: 5.000

Buttons: Run, Cancel

```
%structure diltia
%initialize
%job diltia
%data
  pattern pd_0029.pow
  wave 1.540560
%ntreor
CEM=50,
```

^{*}Kojic Prodic, B., Ruzic Toros, Z., Sunjic, V., Decorte, E. & Moimas, F. (1984). Helv. Chim. Acta, 67, 916–925.

Indexing of norbornene



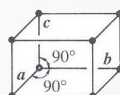
Published cell *

a: 7.6063

b: 8.6220

c: 8.749

β : 97.24



Two unindexed lines from small amount of the hexagonal phase

Indexing Settings

Indexing Programs
☒ N-TREOR09
☐ DICVOL06
☐ McMaille

Maximum Volume (\AA^3) 4000.000
 Maximum a,b,c (\AA) 35.000
 Max input lines 2
 Minimum figure-of-merit 5.000

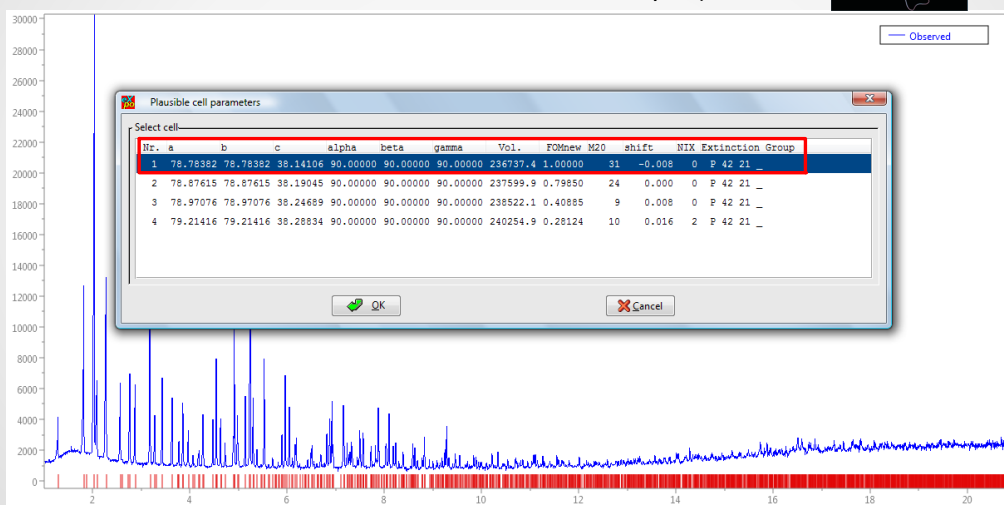
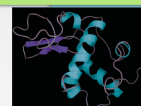
Defaults Run Cancel

*M. Brunelli et al. "Crystal and Molecular Structures of Norbornene." Zeitschrift für Kristallographie, 2001 ; 51-55.

Expo is able to index small proteins also

HEWL (HEN EGG-WHITE LYSOZYME)*

Published Cell : a= 78.8 \AA , c= 38.2 \AA $\alpha = \beta = \gamma = 90.0^\circ$



* Data kindly provided by Irene Margiolaki

Space group determination

Analysis of the systematically absent reflections



Extinction symbol (ES)

Cryst. system	Ext. symb	Space groups
Mon.	$P1 - 1$	$P2, Pm, P2/m$
Mon.	$P1 2_1 1$	$P2_1, P2_1/m$
Mon.	$P1 2_1/c 1$	$P2_1/c$
Orth.	$P - -$	$P222, Pm2m, P2mm, Pmm2, Pmmm$
Tetr.	$P - -$	$P4, P\bar{4}, P4/m, P422, P4mm, P\bar{4}2m, P\bar{4}m2, P4_2m, P4/mmm$

The ES does not unambiguously define the space group

Extracting integrated intensities

Relatively straightforward stage in the structure solution process

Accidental or systematic **peak overlap** can clearly cause difficulties

Two principal techniques have been developed

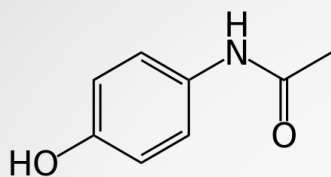
- The iterative **Le Bail method** (Le Bail et al. 1988) based upon Rietveld's original method (Rietveld 1969) for determining observed structure-factor magnitudes.
- **The Pawley method** (1981) is a least-squares analysis of a powder diffraction pattern.

The Expo procedure for space group determination

This may be synthesized into the following steps:

1. The experimental powder diffraction diagram is decomposed *via* the **Le Bail algorithm** into single diffraction intensities in the space group having the largest Laue symmetry and no extinction conditions (e.g., P2/m for monoclinic, P2/m2/m2/m for orthorhombic, P4/mmm for tetragonal, P6/mmm for trigonal-hexagonal systems, and Pm3m for the cubic system).
2. The normalized intensities $z_h = |E_h|^2$ are submitted to **statistical analysis** for the determination of the space group symmetry.
3. The algorithm provides a probability value for each **extinction symbol** compatible with the lattice symmetry established by the indexing procedure.

Indexing and space group determination of paracetamol



N-(4-hydroxyphenyl)ethanamide
(paracetamol)
 $C_8H_9NO_2$

Plausible cell parameters

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	F00new	M20	shift	NIX	Symmetry Info
1	H	11.70585	9.35867	7.09980	90.000	97.413	90.000	773.3	1.000	53.00	0.000	0	P 1 21/m 1
2	E	11.70640	9.37910	7.09980	90.000	97.415	90.000	773.0		61.50	0.005	0	None

Missing Information

Cell Parameters

a: 11.706400 b: 9.379100 c: 7.099800 α : 90.0000 β : 97.41500 γ : 90.0000

Volume: 773.007

Space Group

☒ Find Space Group

Space Group Symbol: P 2/m

Cell Content:

Content Volume: Volume per Atom:

OK

Cell content

Cell content = $Z \cdot$ molecular formula

$$Z = \frac{V_{\text{cell}}}{V_{\text{mol}}} \quad V_{\text{mol}} = \text{volume of molecule}$$

18 Å³ rule: $V_{\text{mol}} = M \cdot 18$

M = number of non-hydrogen atoms

$$Z = \frac{V_{\text{cell}}}{18 \cdot M}$$

Accurate estimation of molecular volume

$$V_{\text{mol}} = \sum n_i v_i$$

n_i = number of atoms of the i^{th} type in the structure

v_i = volume contribution (in Å³) for the i^{th} atom type

Some approximate atomic volumes:

$v_{\text{H}} = 5 \text{ Å}^3$	$v_{\text{O}} = 11 \text{ Å}^3$
$v_{\text{C}} = 14 \text{ Å}^3$	$v_{\text{S}} = 25 \text{ Å}^3$
$v_{\text{N}} = 12 \text{ Å}^3$	$v_{\text{F}} = 11 \text{ Å}^3$

Cell content of paracetamol

Chemical formula: $C_8H_9NO_2$

Volume of unit cell = $V_{\text{cell}} = 773 \text{ \AA}^3$

18 \AA^3 rule

$$Z = \frac{V_{\text{cell}}}{18 \cdot M} = \frac{773}{18 \cdot 11} = 3.90 \sim 4$$

Accurate method

$$V_{\text{mol}} = \sum n_i v_i = 8 v_C + 9 v_H + v_N + 2 v_O = 8 \cdot 14 + 9 \cdot 5 + 12 + 2 \cdot 11 = 191 \text{ \AA}^3$$

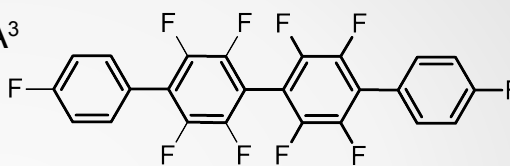
$$Z = \frac{V_{\text{cell}}}{V_{\text{mol}}} = \frac{773}{191} = 4.05$$

Cell content: $(C_8H_9NO_2)_4$ or $C_{32}H_{36}N_4O_8$

Cell content of decafluoroquarterphenyl

Chemical formula: $C_{24}H_8F_{10}$

Volume of unit cell = $V_{\text{cell}} = 1799 \text{ \AA}^3$



18 \AA^3 rule

$$Z = \frac{V_{\text{cell}}}{18 \cdot M} = \frac{1799}{18 \cdot 34} = 2.90 \sim 2 \text{ or } 4$$

Accurate method

$$V_{\text{mol}} = \sum n_i v_i = 24 v_C + 8 v_H + 10 v_F = 24 \cdot 14 + 8 \cdot 5 + 10 \cdot 11 = 486 \text{ \AA}^3$$

$$Z = \frac{V_{\text{cell}}}{V_{\text{mol}}} = \frac{1799}{486} = 3.70 \sim 4$$

Cell content: $(C_{24}H_8F_{10})_4$ or $C_{96}H_{32}F_{40}$

Space group determination of paracetamol

Missing Information

Cell Parameters

a: 11.706400 b: 9.379100 c: 7.099800 α : 90.0000 β : 97.41500 γ : 90.0000

Volume: 773.007

Space Group

☒ Find Space Group

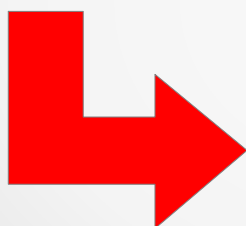
Cell Content: C8H9NO2

Content Volume: 765.040 Volume per Atom: 17.568

Space Group Symbol:

P 2/m

OK



Find space group

Select extinction group

Extinction Group	Fig.Mer
1) P 1 21/n 1	0.585
2) P 1 n 1	0.282
3) P 1 21 1	0.089
4) P 1 _ 1	0.043
5) P 1 21/c 1	0.025
6) P 1 c 1	0.012

Select space group

Space Group Name	No. in CSD
p 21/n	163311 (35.05 %)

☐ User's space group p-1

OK Cancel

Space group determination of paracetamol

Find space group

Select extinction group

Extinction Group	Fig.Mer
1) P 1 21/n 1	0.585
2) P 1 n 1	0.282
3) P 1 21 1	0.089
4) P 1 _ 1	0.043
5) P 1 21/c 1	0.025
6) P 1 c 1	0.012

Select space group

Space Group Name	No. in CSD
p 21/n	163311 (35.05 %)

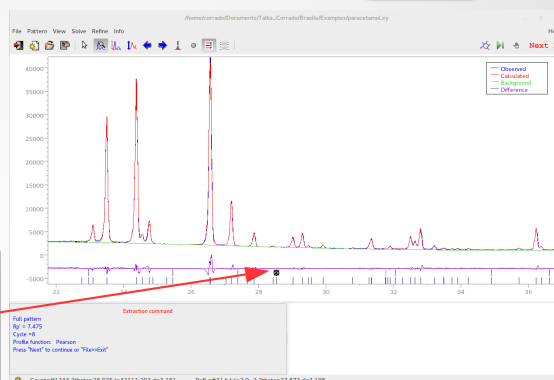
☐ User's space group p-1

OK Cancel

List of systematically absent reflections

Num	h	k	l	Extinction condition	Probability for extinction	Type
33	0	3	0	21 (0 k 0 : k)	1.000	Single
143	0	5	0	21 (0 k 0 : k)	0.766	OverLapped
29	2	0	1	n (h 0 l : h + l)	1.000	Single
80	3	0	-2	n (h 0 l : h + l)	0.998	Single
106	3	0	2	n (h 0 l : h + l)	0.996	Single
154	1	0	6	n (h 0 l : h + l)	0.995	Single
271	4	0	-5	n (h 0 l : h + l)	0.992	OverLapped
14	1	0	2	n (h 0 l : h + l)	0.940	Single
10	1	0	-2	n (h 0 l : h + l)	0.984	Single
44	1	0	-4	n (h 0 l : h + l)	0.972	OverLapped
71	3	0	0	n (h 0 l : h + l)	0.965	OverLapped
75	0	0	5	n (h 0 l : h + l)	0.956	OverLapped
185	4	0	1	n (h 0 l : h + l)	0.948	OverLapped
122	3	0	-4	n (h 0 l : h + l)	0.947	OverLapped
190	4	0	-3	n (h 0 l : h + l)	0.944	OverLapped
2	1	0	0	n (h 0 l : h + l)	0.992	Single
201	0	0	7	n (h 0 l : h + l)	0.940	OverLapped
47	2	0	-3	n (h 0 l : h + l)	0.928	Single
16	0	0	3	n (h 0 l : h + l)	0.926	Single

Close



Space group determination of decafluoroquarterphenyl

```

%structure dfqp
%initialize
%job decafluoroquarterphenyl (C24 H8 F 10)
%data
pattern dfqp.dat
range 7.000 70.000
wave 1.790000
cont (C24 H8 F10) 4
cell 24.085501 6.155077 12.435071 90.00000 102.68274 90.00000
%continue
  
```

Find space group

Select extinction group		Select extinction group (pseudo applied)		Select space group	
Extinction Group	Fig.Mer	Extinction Group	Fig.Mer	Space Group Name	No. in CSD
1) I 1 a 1	0.463	1) P 1 21/a 1	0.417	i a	5006 (1.07 %)
2) I 1 _ 1	0.364	2) P 1 21/c 1	0.380	i 2/a	37507 (8.05 %)
3) P 1 21/a 1	0.103	3) P 1 a 1	0.178		
4) P 1 21/c 1	0.094	4) P 1 21 1	0.164		
5) P 1 21/n 1	0.094	5) P 1 c 1	0.162		
6) P 1 a 1	0.044	6) I 1 a 1	0.100		
7) P 1 21 1	0.041	7) P 1 21/n 1	0.082		

☐ User's space group
 p-1
OK Cancel

Structure solution by Direct Methods

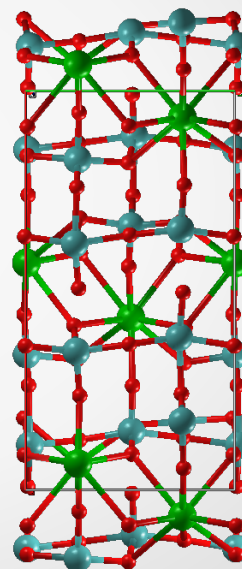
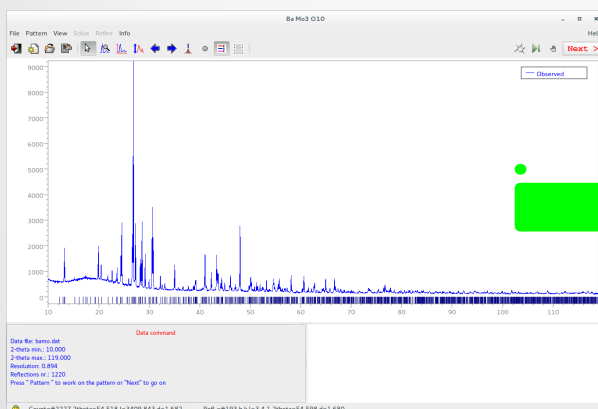
- a) Full pattern decomposition process
- b) Normalization of integrated intensities
- c) Phasing process
- d) E-map calculation
- e) Model optimization



Crystal structure determination of $\text{BaMo}_3\text{O}_{10}$ *

Space group: $P2_1$; $Z = 4$

$a = 14.695(2) \text{ \AA}$, $b = 7.5704(7) \text{ \AA}$, $c = 6.9618(6) \text{ \AA}$, and $\beta = 100.381(8)^\circ$

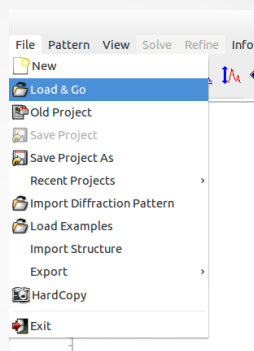


* Werner, P.-E., Moustiakimov, M., Marinder, B.-O. & Knight, K. S. (1997). Z. Kristallogr. 212, 665–670.

Input file for crystal structure determination of $\text{BaMo}_3\text{O}_{10}$

File **bamo.exp**

```
%struct bamo
%job Ba Mo3 O10
%init
%data
  pattern bamo.dat
  Cell 14.695 7.5704 6.9618 90.000 100.381 90.000
  wave 1.5406
  space p 21
  cont (Ba Mo3 O10) 4
%continue
```



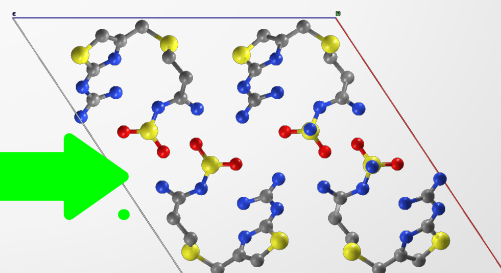
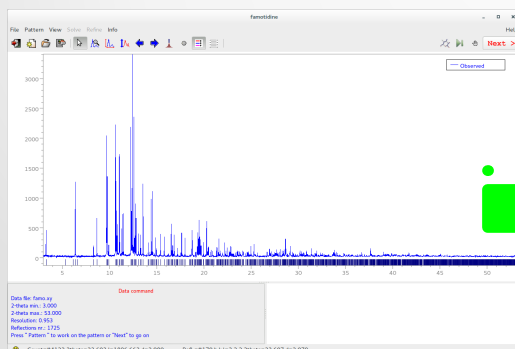
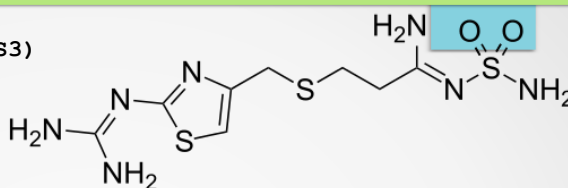


Crystal structure determination of famotidine

```
%job famotidine (C8 H15 N7 O2 S3)
%structure famo
%init
%data
```

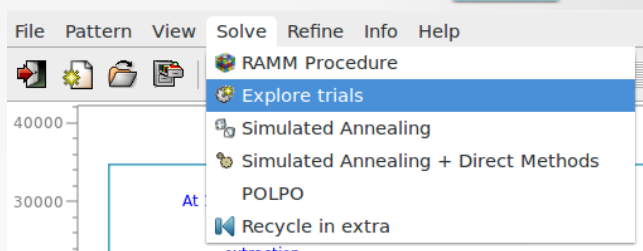
```
pattern famo.xy
cont (C8 H15 N7 O2 S3) 4
wave 0.8507473
cell 17.6547 5.29320 18.2590 90.0 123.5580 90.0
space p 21/c
synchrotron
```

```
%continue
```



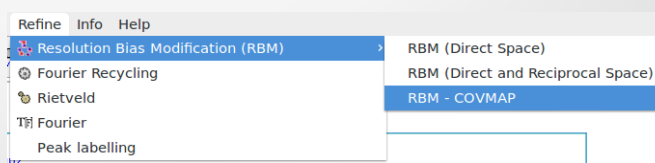
When default Expo2013 fails: strategies

- Explore trials



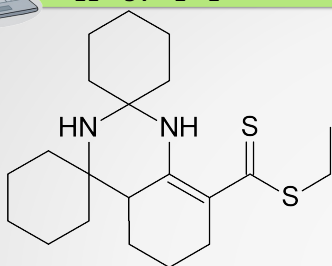
- Apply the RAMM (Random based Method)

- Structure model optimization

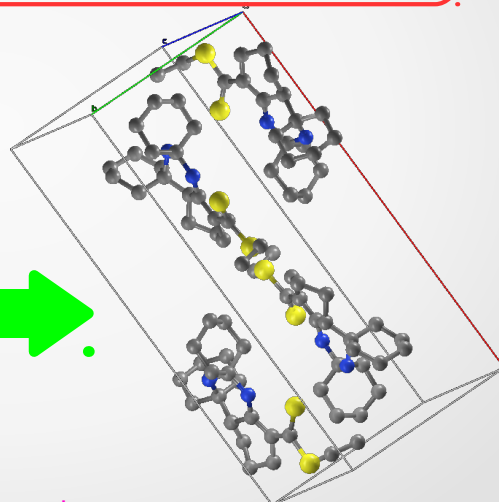
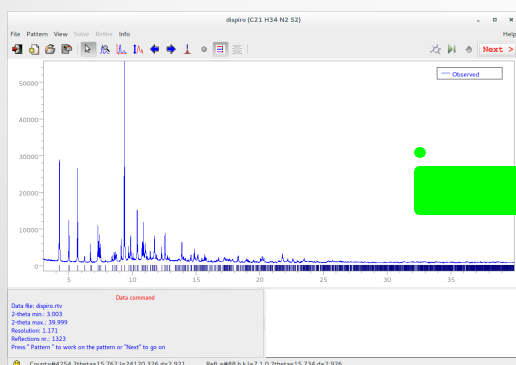




Crystal structure solution of a bidentate pro-ligand, $C_{21}H_{34}N_2S_2$ from powder synchrotron diffraction data*



All trials were explored during the Direct Methods procedure



*Ávila, E. E. (2009). Acta Cryst. B65, 639–646

Exploring trials

Solve Refine Info

Explore trials

- Simulated Annealing
- Simulated Annealing + Direct Methods
- POLPO
- Recycle in extra

Explore trials

Fourier Procedures

☒ RBM ☐ Fourier recycling ☐ E-map ☐ COVMAP

☐ Select all new trials

Develop trial	Set	cform	done	RF
<input checked="" type="checkbox"/> 20	0.866	yes	46.688	
<input type="checkbox"/> 9	0.885	yes	46.855	
<input type="checkbox"/> 18	0.874	yes	47.061	
<input type="checkbox"/> 15	0.879	yes	47.619	
<input type="checkbox"/> 4	0.891	yes	48.847	
<input type="checkbox"/> 17	0.876	yes	54.167	
<input type="checkbox"/> 10	0.884	yes	58.672	
<input type="checkbox"/> 16	0.878	yes	59.413	
<input type="checkbox"/> 2	0.902	yes	59.703	
<input type="checkbox"/> 3	0.902	yes	60.070	
<input type="checkbox"/> 5	0.888	yes	60.082	
<input type="checkbox"/> 12	0.882	yes	60.485	
<input type="checkbox"/> 19	0.869	yes	60.952	
<input type="checkbox"/> 8	0.886	yes	61.373	
<input type="checkbox"/> 6	0.888	yes	61.710	
<input type="checkbox"/> 13	0.882	yes	61.901	

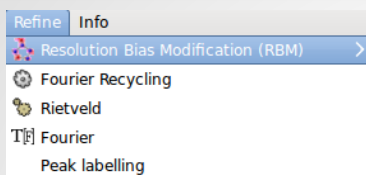
Quit

Command %alltrials

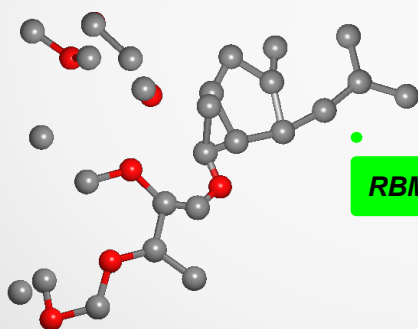
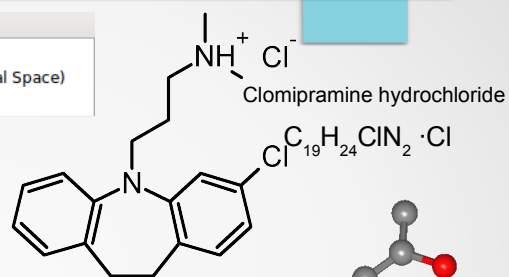
The use of the command `%alltrials` in the input file automatically activate the 'Explore trials' approach

```
%struct dispiro
%job dispiro (C21 H34 N2 S2)
%init
%data
  pattern dispiro.rtv
  cont (C21 H34 N2 S2) 4
  wave 0.80098
  range 3.003 40
  cell 21.7356 10.0565 9.4510 90.000 99.9602 90.000
  space p 21/n
  synchrotron
%extra
%normal
%invar
%phase
%alltrials
```

Applying RBM for model completion

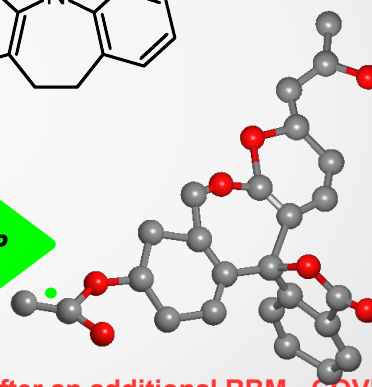


RBM (Direct Space)
RBM (Direct and Reciprocal Space)
RBM - COVMAP



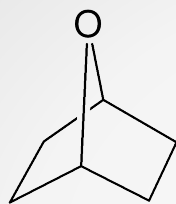
Model after alltrials

RBM - COVMAP

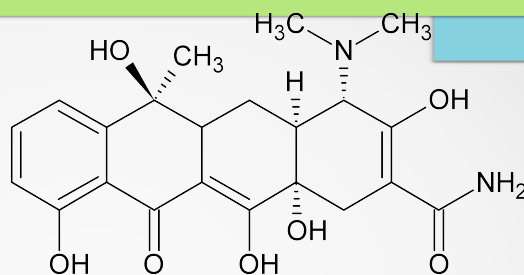


Model after an additional RBM - COVMAP run, graphically requested

Limits of Direct Methods

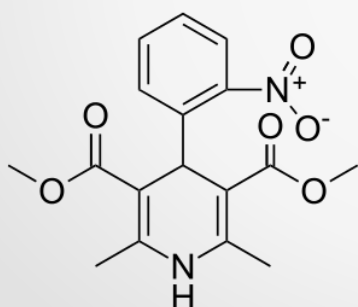


phase III of oxanorbornane
 $Z'=4$, synchrotron data

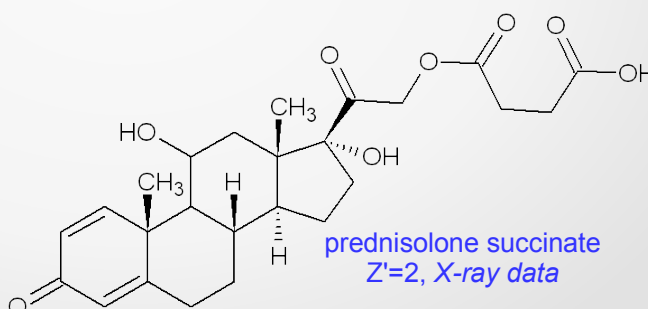


tetracycline hydrochloride
 $Z'=1$, synchrotron data

HCl



C polymorph of nifedipine
 $Z'=2$, synchrotron data



prednisolone succinate
 $Z'=2$, X-ray data

The Rietveld method

Based on the idea suggested in the middle **1960s** by Rietveld

Aim of the method

All structural and instrumental parameters are refined by fitting a calculated profile to the observed data without extraction of the individual integrated intensities

- Requires a model of a crystal structure
- Nonlinear least squares method

Fundamentals of the Rietveld Method

The minimized function is given by:

$$\Phi = \sum_i w_i \cdot (y_{oi} - y_{ci})^2$$

y_{oi} is the observed intensity at the i th data point

y_{ci} is the calculated intensity at the i th data point

w_i is a weight $1/\sigma^2(y_{oi})$.

For mixture of several phases, the contribution from every crystalline phase is accounted in the expression of y_{ci} .



Quantitative analysis of a multiple phase crystalline material

Parameters to be refined:

$$y_{ci} = f(x_1, x_2, \dots, x_j, \dots, x_n)$$

x_j parameters to be refined

- atomic positions
- atomic site occupancies
- atomic thermal factors
- lattice parameters
- background coefficients
- profile function parameters
- preferred orientation
- 2θ zero correction

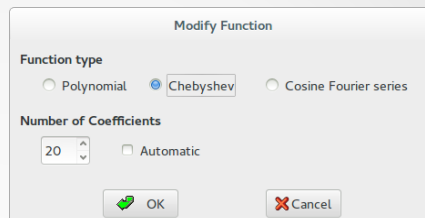
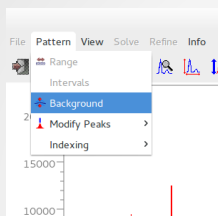
Some of them are correlated
(e.g., occupancy factors
and thermal factors)



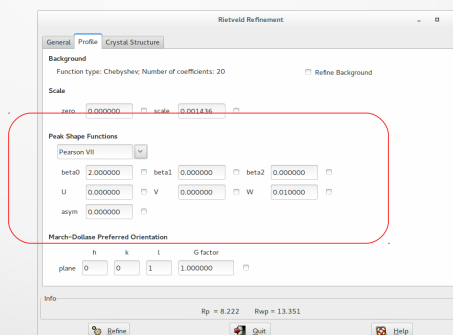
**They must not be
refined all together**

Profile parameters

Background function

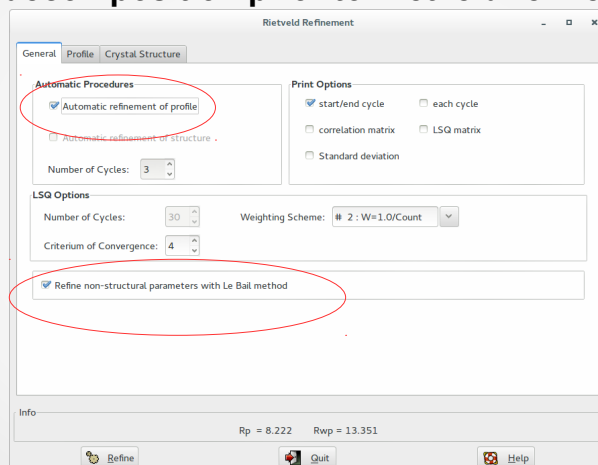


Peak shape functions



Profile parameters

The Le Bail technique can be adopted to perform a full pattern decomposition prior to Rietveld refinement



This strategy is suggested especially if the available structure model is not completed (Rietveld refinement guidelines, L.B. McCusker, R.B. Von Dreele, D.E. Cox, D. Louer, P. Scardi, *J. Appl. Cryst.* **32** (1999) 36)

Refinement strategies

The refinement can be carried out by following two alternative approaches:

- The user can decide the refinement strategy via graphic interface
- An automatic refinement schedule can be applied
 - Scale
 - 2θ correction
 - Background coefficients
 - W
 - U , V , other profile parameters
 - Coordinates of atoms
 - Isotropic displacements

Restraints

$$\Phi = \sum_i w_i \cdot (y_{oi} - y_{ci})^2 + w_{dist} \sum_i w_i^{dist} \cdot (dist_i^{exp} - dist_i^{calc})^2 + w_{dist} \sum_i w_i^{ang} \cdot (a_i^{exp} - a_i^{calc})^2 + w_{plane} \sum_i w_i^{plane} \cdot (p_i^{exp} - p_i^{calc})^2$$

Each type of restraints is included in the refinement as a set of observations, in addition to the main set

Manage restraints

Select Restraint Type and weight: Distance restraints Weight on distances: 10000.00 Save Load

List of restraints:

Active	Atoms	Current	Target	Weight
<input checked="" type="checkbox"/>	C1-N1	1.355	1.321	0.030
<input type="checkbox"/>	C1-O1	1.233	1.221	0.030
<input type="checkbox"/>	C1-S1	1.785	1.751	0.030
<input type="checkbox"/>	C2-C3	1.388	1.380	0.030
<input type="checkbox"/>	C2-C7	1.401	1.380	0.030
<input type="checkbox"/>	C2-S1	1.758	1.751	0.030
<input type="checkbox"/>	C3-C4	1.398	1.380	0.030
<input type="checkbox"/>	C3-N1	1.397	1.443	0.030
<input type="checkbox"/>	C4-C5	1.388	1.380	0.030
<input type="checkbox"/>	C4-O2	1.372	1.395	0.030
<input type="checkbox"/>	C5-C6	1.399	1.380	0.030
<input type="checkbox"/>	C6-C7	1.393	1.380	0.030
<input type="checkbox"/>	C7-C8	1.512	1.530	0.030
<input type="checkbox"/>	C8-C9	1.521	1.530	0.030
<input type="checkbox"/>	C9-N2	1.498	1.443	0.030

Active restraints:

OK Cancel

Constraints

Constraints are mathematical relationships between parameters

Symmetry constraints are mandatory and automatically imposed by the program

- *Special position*
 e.g., atom on special position $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ should not be refined,
 atom on special position (x, x, x) in space group P23 should have equal shift on x,y,z
- *Unit cell dimension*
 e.g., $a=b=c$ and $\alpha=\beta=\gamma$ in cubic crystal system
- *Occupation factor*
 e.g., A,B atoms in same site: $occA + occB = 1$

Constraints

Constraints imposed by the user to reduce the number of parameters

- Riding model (move H atoms synchronously with the C atoms)
- Rigid body (molecules or groups that are treated as a whole)
- Constraints on ADPs (ADPs are made to shift synchronously)

Statistical measures of a refinement

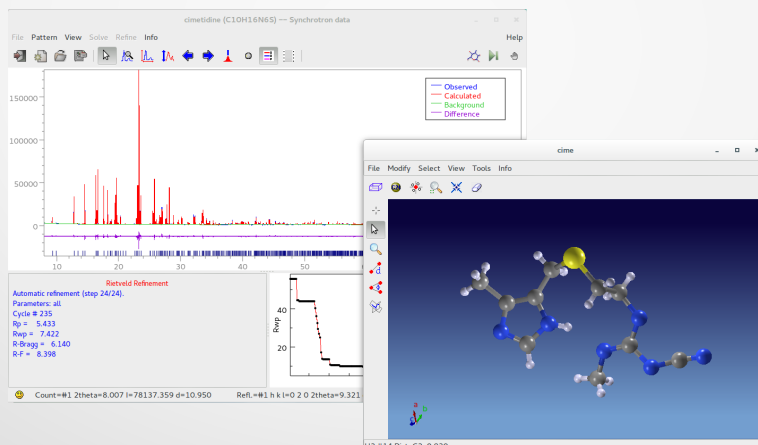
- Weighted profile R-factor
$$R_{wp} = \frac{\sum_i w_i \cdot (y_{oi} - y_{ci})^2}{\sum_i w_i y_{oi}^2}$$
- Unweighted profile R-factor
$$R_p = \frac{\sum_i |(y_{oi} - y_{ci})|}{\sum_i y_{oi}}$$
- Expected R value
$$R_{exp} = \sqrt{\frac{(N - P)}{\sum_i w_i y_{oi}^2}}$$
- Goodness-of-fit
$$\chi^2 = \frac{R_{wp}}{R_{exp}}$$
- Other residual on F or F^2 :
$$R_F = \frac{\sum_{hkl} |(F_{hkl}^{obs} - F_{hkl}^{calc})|}{\sum_{hkl} F_{hkl}^{obs}} \quad R_B = \frac{\sum_{hkl} |(I_{hkl}^{obs} - I_{hkl}^{calc})|}{\sum_{hkl} I_{hkl}^{obs}}$$

For the full pattern and the Background-subtracted pattern
 R_{wp}, R_p

Quality of refinement

Important criteria for the quality of the refinement:

- the fit of the calculated pattern to the observed data
- the chemical sense of the structural model

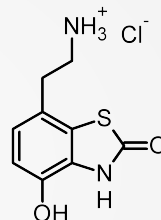


Structure refinement of $C_9H_{11}N_2O_2S \cdot Cl$

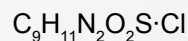
Input file for Rietveld refinement:

```

%Structure ammonium
%Job ethylammonium chloride (C9H11N2O2SCl)
%Data
  Cell 7.555 14.640 10.246 90 109.30 90
  SpaceGroup p 21/a
  Pattern ammonium.xy
  Wavelength 1.54056
%fragment ammonium_riet.cif
%rietveld
  
```



2-(4-Hydroxy-2-oxo-2,3-dihydro-1,3-benzothiazol-7-yl)
ethylammonium chloride



From graphical interface:

- File > Import Diffraction Pattern
- File > Import Structure
- Refine > Rietveld

Thank you
for your attention