EXPO&*more* International Workshop

30 September – 3 October 2019

CNR Istituto di Cristallografia

Indexing

Anna Moliterni

Institute of Crystallography (IC) – CNR
Bari - Italy

annagrazia.moliterni@ic.cnr.it

Outline

- Aim and features of powder pattern indexing;
- The indexing equation;
- Indexing by ***EXPO***: main features;
- Indexing by ***EXPO***: applications and suggestions in case of failures of a default run;
- Final remarks.

Powder pattern indexing

is the first step in the pathway of the solution process:

- INDEXING
- SPACE GROUP DETERMINATION
- PROFILE DECOMPOSITION AND INTENSITY EXTRACTION
- STRUCTURE SOLUTION AND MODEL OPTIMIZATION
- RIETVELD REFINEMENT

Indexing is a crucial step of the solution process:
its failure prevents from determining the crystal structure.

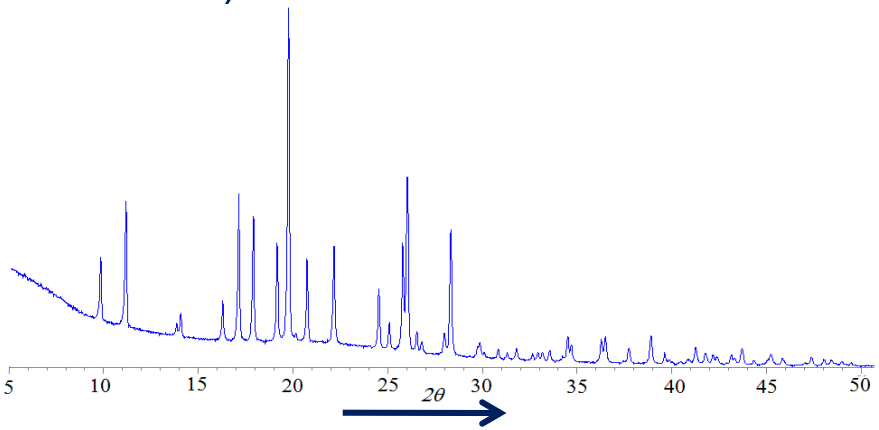
Aim of powder pattern indexing

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

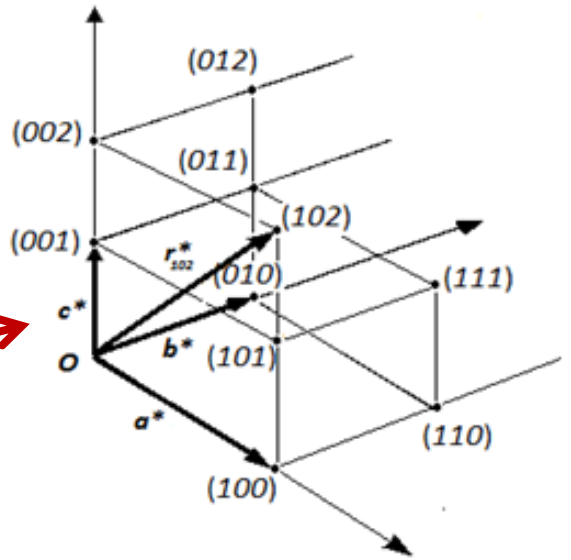
Indexing ⟷ Assignment of (*hkl*) **Miller indices** of a lattice plane to the diffraction intensity measured at the reciprocal lattice point

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

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

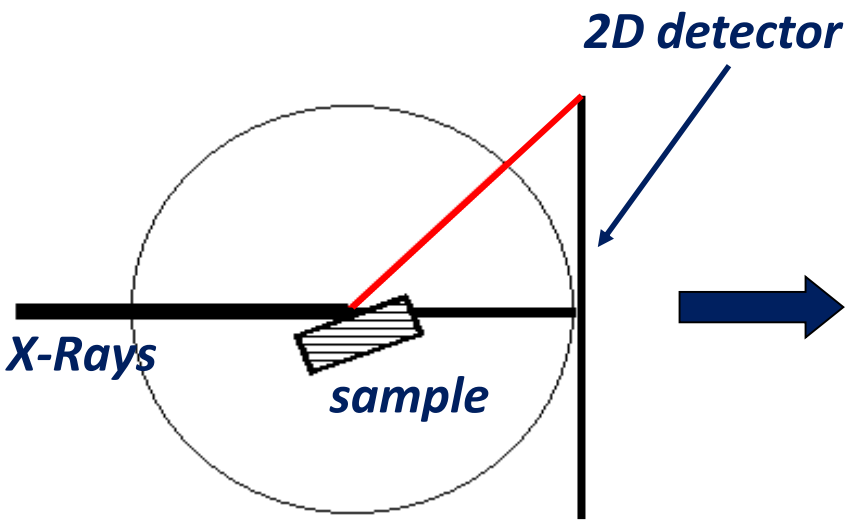


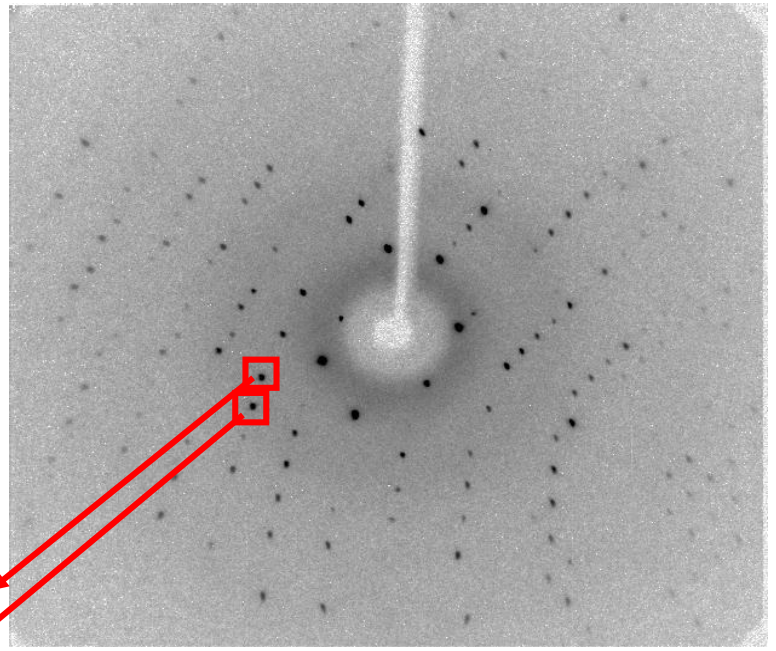
$$|\mathbf{r}^*_h| = |\mathbf{r}^*_{hkl}| = 1/d_h = 2\sin\theta_h/\lambda$$



Single crystal diffraction

(~ mm)



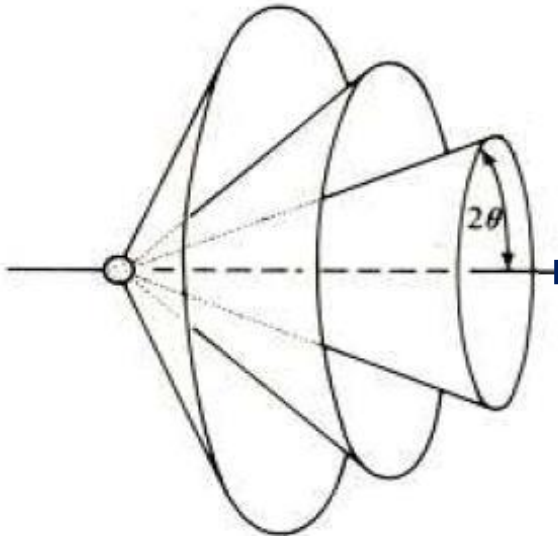


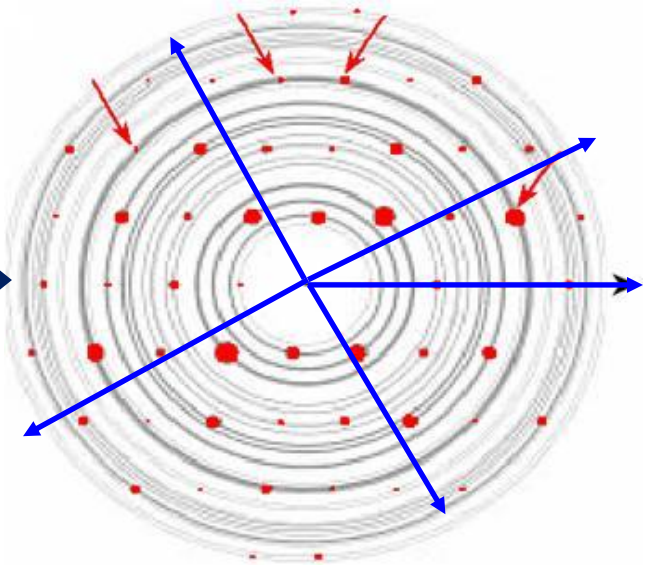
The diffraction effects are well separated.

- *The experimental information is* ***three-dimensional****;*
- *There is a* ***biunivocal correspondence between diffraction effects*** *(measured at the reciprocal lattice point whose vector is* $\mathbf{r^*}_h$*)* ***and involved lattice planes****, with* $d_h=d_{hkl}$ *interplanar spacing, satisfying the* ***Bragg law*** $2\,d_h \sin\theta_h=\lambda$.

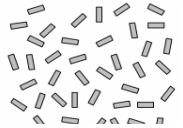
In case of single crystal data, indexing is usually a quite trivial task.

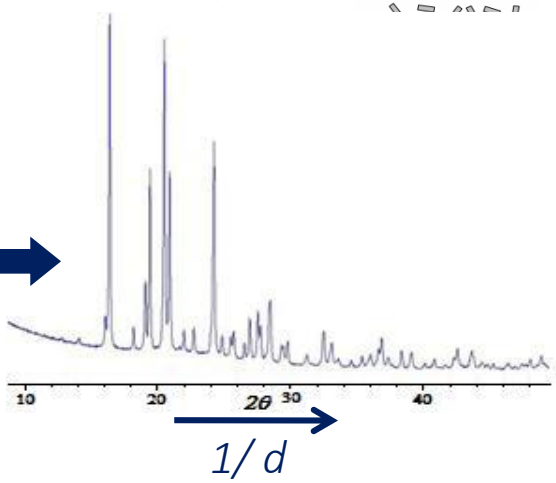
Features of powder pattern indexing





Randomly oriented crystallites (~ μm)





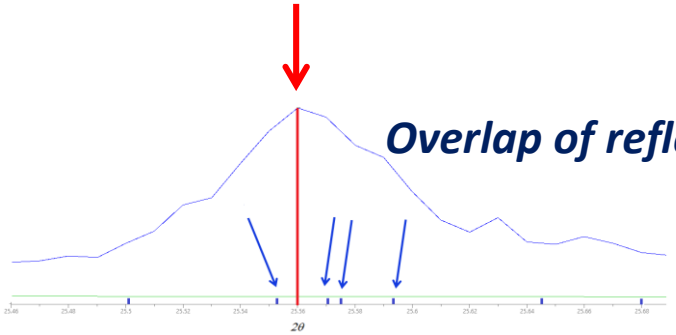
The corresponding single crystal pattern is superimposed (red spots)

$1/d_h = 2\sin\theta_h/\lambda$

- In case of powder data the three-dimensional information is lost.
  The experimental information is one-dimensional.
- The main limitation of powder diffraction technique with respect to single crystal method: *to reconstruct the three-dimensional reciprocal lattice from the one-dimensional powder diffraction profile.*

→ ***In case of powder data, indexing can be a challenge.***

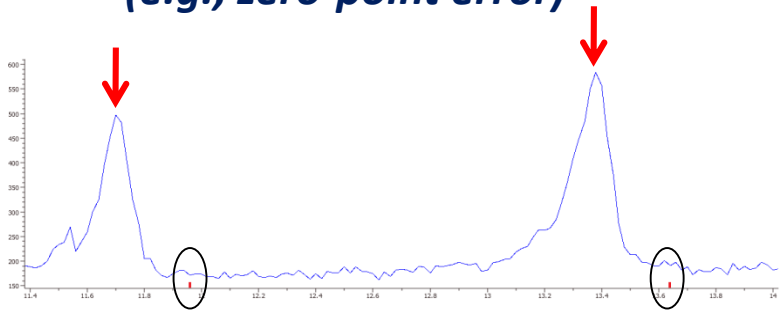
In case of powder data several problems can concur:



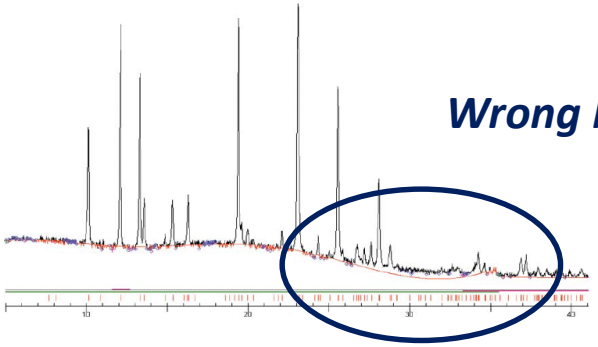
Overlap of reflections

and/or

Errors on experimental data (e.g., zero-point error)



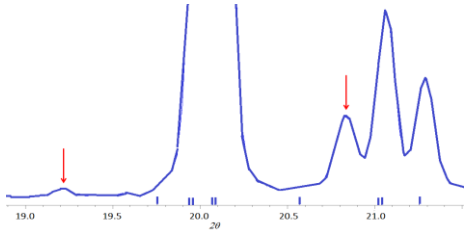
and/or

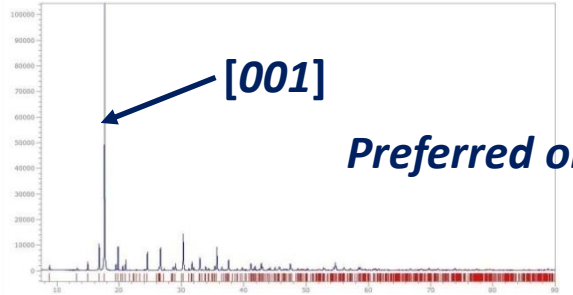


Wrong background definition

and/or

Presence of impurity peaks





Preferred orientation effects

and/or

In case of powder data, indexing can be not straightforward.

The indexing equation

*In order to determine the reciprocal and, consequently, the direct unit cell parameters the following basic equation has to be solved (**in the general triclinic case**):*

$$Q(hkl) = h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23}$$

$\downarrow$ $\{A_{ij}\}$

The solution is not unique.

The number of $\{A_{ij}\}$ to be determined depends on the investigated crystal system symmetry.

where

$Q(hkl) = 10^4\ |\mathbf{r^*}_h|^2 = 10^4\ 1/d_h^2$,

$d_h = f(a^*, b^*, c^*, \alpha^*, \beta^*, \gamma^*, h)$

$$A_{11} = 10^4\ a^{*2}, \quad A_{22} = 10^4\ b^{*2}, \quad A_{33} = 10^4\ c^{*2},$$
$$A_{12} = 10^4\ 2\ a^*b^*\cos\gamma^*, \quad A_{13} = 10^4\ 2\ a^*c^*\cos\beta^*,$$
$$A_{23} = 10^4\ 2\ b^*c^*\cos\alpha^*$$

Symmetry	$1/d_h$
Cubic	$(h^2 + k^2 + l^2)^{1/2}\ a^*$
Tetragonal	$[(h^2 + k^2)\ a^{*2} + l^2\ c^{*2}]^{1/2}$
Hexagonal	$[(h^2 + hk + k^2)\ a^{*2} + l^2\ c^{*2}]^{1/2}$
Orthorhombic	$[h^2\ a^{*2} + k^2\ b^{*2} + l^2\ c^{*2}]^{1/2}$
Monoclinic	$[h^2\ a^{*2} + k^2\ b^{*2} + l^2\ c^{*2} + 2hl\ a^*c^*\cos\beta^*]^{1/2}$
Triclinic	$[h^2\ a^{*2} + k^2\ b^{*2} + l^2\ c^{*2} + 2hka^*b^*\cos\gamma^* + 2hl\ a^*c^*\cos\beta^* + 2kl\ b^*c^*\cos\alpha^*]^{1/2}$

The indexing equation

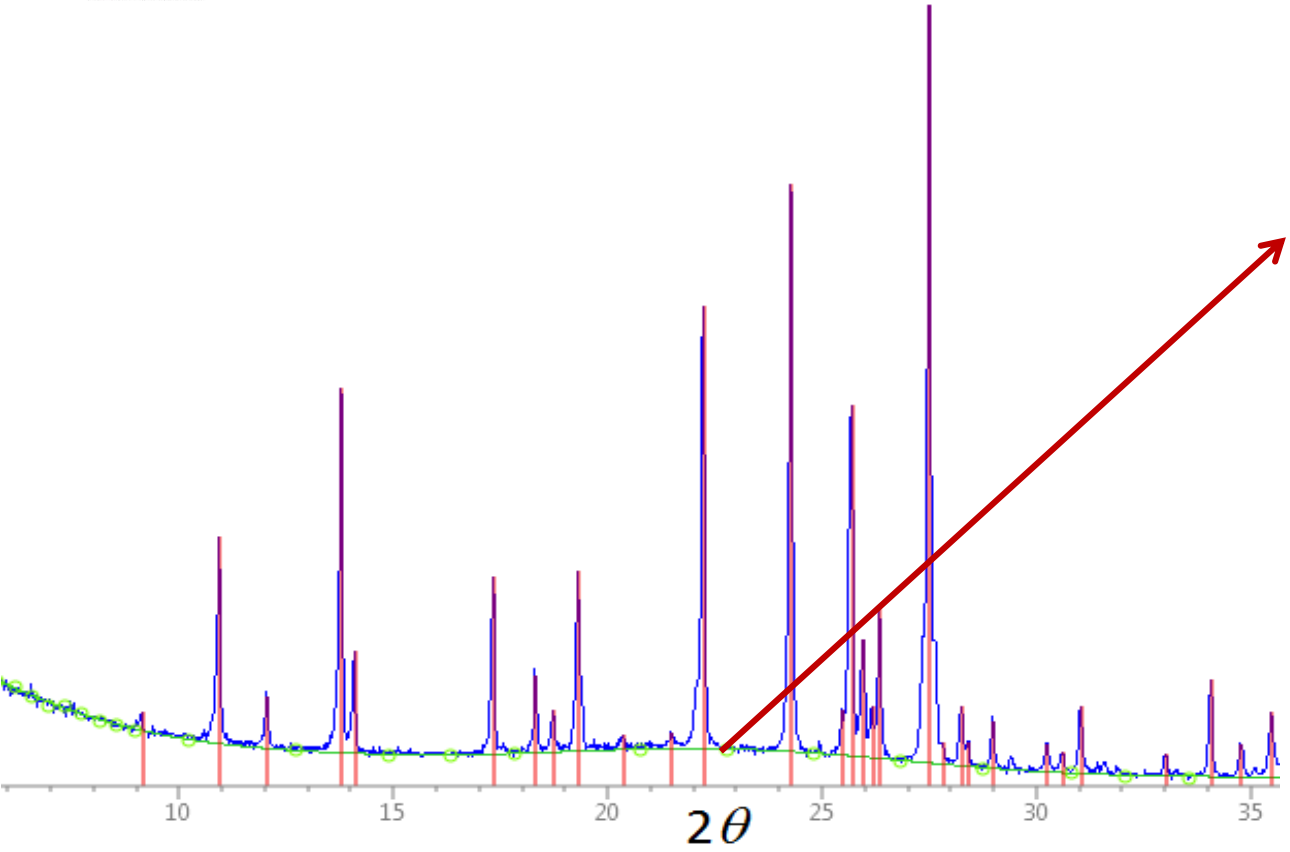
$$Q(hkl) = h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23}$$

Unavoidable errors on the experimental $Q(hkl)$ → $\{A_{ij}\}$ *should satisfy the indexing equation by taking into account suitable tolerance* ε *on* $Q_{obs}(hkl)$ *:*

$$Q(hkl) - \varepsilon \; < \; h^2A_{11} + k^2A_{22} + l^2A_{33} + hkA_{12} + hlA_{13} + klA_{23} \; < \; Q(hkl) + \varepsilon$$

If $\{A_{ij}\}$ *are known* → $(a^*, b^*, c^*, \alpha^*, \beta^*, \gamma^*)$ *are easily determined* → $(a, b, c, \alpha, \beta, \gamma)$ *are derived*

$$\begin{cases} \boldsymbol{a} = \dfrac{\boldsymbol{b}^* \wedge \boldsymbol{c}^*}{V^*}, \quad \boldsymbol{b} = \dfrac{\boldsymbol{c}^* \wedge \boldsymbol{a}^*}{V^*}, \\ \boldsymbol{c} = \dfrac{\boldsymbol{a}^* \wedge \boldsymbol{b}^*}{V^*} \\ V^* = \boldsymbol{a}^* \cdot \boldsymbol{b}^* \wedge \boldsymbol{c}^* = 1/V \end{cases}$$



2θ	d
9.142967	9.664400
10.954328	8.070067
12.070744	7.326062
13.795424	6.413804
14.098940	6.276405
17.344723	5.108500
18.330000	4.836068
18.736002	4.732177
19.315796	4.591421
20.371922	4.355713
21.503235	4.129041
22.241919	3.993552
24.285576	3.661919
25.502274	3.489902
25.704069	3.462959
25.979597	3.426854
26.182821	3.400715
26.356972	3.378639
27.516184	3.238878
27.817354	3.204492
28.268555	3.154361
28.415022	3.138434
29.007622	3.075652
30.268778	2.950315
30.617796	2.917475
31.050356	2.877810
33.035824	2.709247
34.079815	2.628603
34.789608	2.576581
35.498699	2.526721
35.804897	2.505812
36.439728	2.463601

$d = \lambda / 2\sin\theta$

$\lambda = 1.54056$ Å

$a = 10.2513$ Å $\quad b = 14.6411$ Å $\quad c = 7.5588$ Å

$\alpha = \gamma = 90.0°$ $\quad \beta = 109.43°$

Figures of merit

The solution of the indexing equation is not unique

Special mathematical tools, able to evaluate the reliability of a possible cell, can be used for recognizing the most plausible cell.

The most used figures of merit are
M_{20} (de Wolff, 1968) and *F_N (Smith & Snyder, 1979).*

- ***M_{20}***

$$M_{20} = \frac{Q_{20}}{2 \langle \varepsilon \rangle N_{20}}$$

where

$Q_{20}=10^4/d_{20}^2$, $d_{20} = d$ in case of the 20th experimental and indexed line (line=d value),

N_{20} = number of generated lines up to the d_{20} value,

$\langle \varepsilon \rangle$ = average value of the discrepancy $|Q_{obs} - Q_{calc}|$ calculated for the 20 observed and indexed lines.

The smaller the $\langle \varepsilon \rangle$ (*i.e.*, the larger the agreement between observed and calculated Q values of the lines) and the smaller the N_{20} (*i.e.*, the smaller the cell size), the larger M_{20}.

M_{20} can be overestimated in case of a large number of unindexed lines.

If the number of unindexed lines with $Q \leq Q_{20}$ is not larger than two and $M_{20} > 10$ the indexing can be considered physically reasonable.

Figures of merit

- F_N *(Smith & Snyder, 1979)*

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

where

$<|\Delta 2\theta|>$ = *average value of the discrepancy between the observed and calculated* 2θ *values,*

N_{poss} = *number of possible lines up to the* N^{th} *observed line.*

The pair $(<|\Delta 2\theta|>, N_{poss})$ *is usually provided together with the* F_N *value.*

Figures of merit

New figures of merit, exploiting the information contained in the full pattern, have been recently proposed, among them
McM_{20} (Le Bail, 2008) and WRIP20 (Altomare et al., 2009).

- ***McM_{20}***

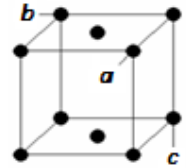
$$McM_{20} = [100/(R_p\ N_{20})]\ \ Br\ \ Sy$$

where

N_{20} = number of generated lines up to the 20th experimental line (for a primitive cell),

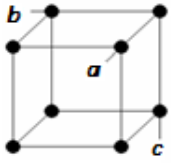
R_p = profile R factor,

Br
- *> 1 in case of centred cell,* → *i.e., A-, B-, C-, I-, R- or F-centred cell. It contains more than one lattice point, e.g.,*



C-centred cell

- *= 1 in case of primitive cell,* → *it contains one lattice point*



Primitive cell

Sy
- *> 1 in case of high symmetry cell,*
- *= 1 for monoclinic or triclinic cell*

Figures of merit

- WRIP20

$$WRIP20 = RAT_{Rp}^{2} \; RAT_{Ind} \; RAT_{Pres} \; w_u \; RAT_{M20}^{1/2}$$

It takes account of:

- RAT_{Rp}^{2}: *Agreement between observed and calculated powder pattern*
- RAT_{Ind}: *Overlap in the powder pattern*
- RAT_{Pres}: *Symmetry*
- w_u: *Number of unindexed lines*
- $RAT_{M20}^{1/2}$: *M_{20} value*

WRIP20 *can be more effective than the traditional M_{20} figure of merit.*

Indexing methods and software

Traditional indexing methods are based on the indexing equation and are applied by the most widely used indexing computer programs; among them:

- *DICVOL* (Boultif and Louër, 1991) and its updated versions: *DICVOL04* (Boultif and Louër, 2004), *DICVOL06* (Louër and Boultif, 2006, 2007) and *DICVOL14* (Louër and Boultif, 2014);
- *TREOR* (Werner *et al.*, 1985) and its evolution: *N-TREOR* (Altomare *et al.*, 2000) and *N-TREOR09* (Altomare *et al.*, 2009);
- ITO (Visser, 1969).

Approaches alternative to the traditional methods and based on global optimization methods (*e.g.*, Monte Carlo, Genetic Algorithm) have been proposed and implemented in software [see, *f.e.*, McMaille (Le Bail, 2004) and GAIN (Harris *et al.*, 2000)].

They search in direct space for the unit cell parameters (a, b, c, α, β, γ) providing the best agreement between the experimental and the calculated pattern in terms of a suitable figure of merit (*e.g*, the agreement factor R_p), in correspondence with the global minimum of the hypersurface $R_p[a, b, c, \alpha, \beta, \gamma]$.

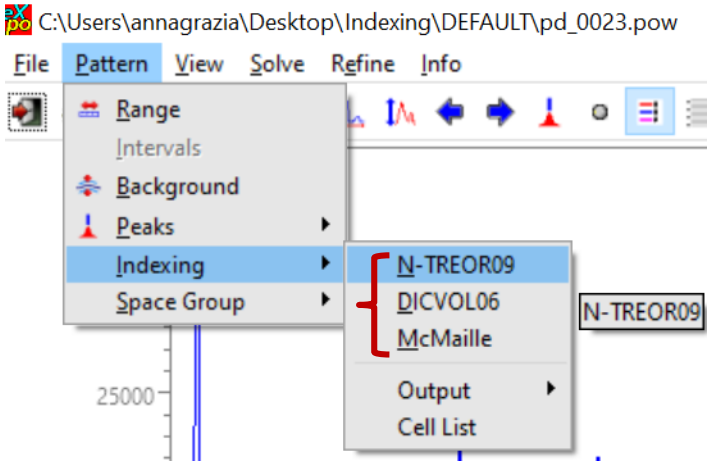
Indexing by EXPO: N-TREOR09

N-TREOR09 (Altomare et al., 2009) is the default indexing program implemented in *EXPO (Altomare et al., 2013).*

N-TREOR09 is the updated version of *N-TREOR* (Altomare *et al.*, 2000), both the programs are the evolution of the indexing program *TREOR90* (Werner *et al.*, 1985).

Recently, additional indexing programs have been implemented in EXPO:

- *DICVOL06* (Louër and Boultif, 2006, 2007);
- *McMaille* (Le Bail, 2004).



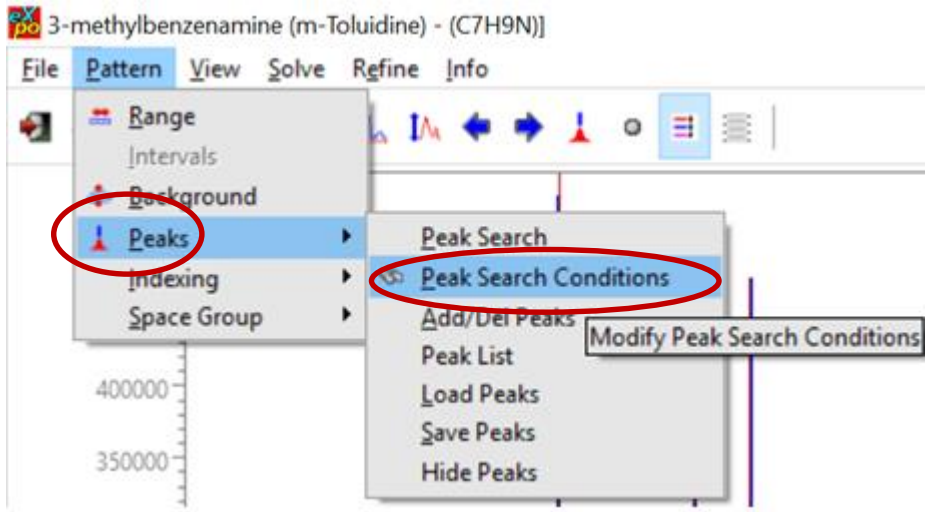
Default indexing by EXPO: main features

The method

The default indexing program *N-TREOR09* is based on an 'index heuristics' strategy, working in 'index space', a 3*N*-dimensional Miller index space (*N* is the number of experimental lines).

This approach searches for the correct cell by varying trial Miller indices assigned to some low resolution experimental lines (basis lines).

A peak-search procedure has to be applied to provide a list of experimental *d* values to the indexing step.



Indexing via the 'index heuristics' strategy

Observed $\{d_{hkl}\}$ → Selection of basis lines (among the set of low resolution experimental d_{hkl}) to which trial values of the *(hkl)* Miller indices are assigned.

The $\{Q(hkl)\}$ of the basis lines and the assigned *(hkl)* values are known →

In order to find the correct cell the $\{A_{ij}\}$ are determined by solving a system of indexing equations involving the set of observed *Q(hkl)*, the tentative *(hkl)* Miller indices and the unknown $\{A_{ij}\}$:

$$MA=Q$$

where

M is the Miller indices matrix,

Q is the set of the observed *Q(hkl)* values for the selected basis lines,

A is the set of unknown $\{A_{ij}\}$ parameters to be determined.

Indexing via the 'index heuristics' strategy

The number of the basis lines depends on the investigated crystal system symmetry.

Different combinations of the basis lines are also tried, in order to reduce the risk of involving peaks due to an impurity.

The search for the correct cell is carried out from cubic up to lower symmetry crystal systems.

The reliability of each candidate cell is evaluated by the M_{20} and F_N figures of merit.

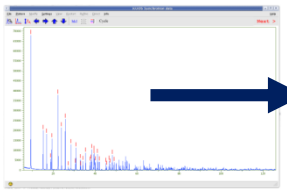
A cell is assumed to be plausible if $M_{20} \geq 10$ and the number of unindexed lines is *not greater than 1.*

Indexing via the 'index heuristics' strategy

Advantages of this approach:

- The search in the index space is fast in all the crystal systems (except for the triclinic case).

- A user defined number of unindexed lines is allowed (only one unindexed line is tolerated in a default run).



N-TREOR

Peak search

Indexing

Cell reduction and conventional cell calculation

Unit cell parameters refinement

Go on?

NO

YES

2θ-zero shift

STOP

List of NSOL plausible cells in function of M_{20}

N-TREOR09

Triclinic search improvement

WRIP20, the new figure of merit for identifying the correct cell

Indexing by N-TREOR09: main features

N-TREOR09 is able to

- Correct the experimental data by zero-point error (errors in 2θ peak positions can be responsible for indexing failure);
- Carry out a more exhaustive search in monoclinic and orthorhombic systems, in case of failure of a default run;
- Perform an exhaustive search in triclinic system;
- Calculate an effective figure of merit (WRIP20) for identifying the correct cell among a set of plausible ones.

- Successfully index a large set of powder diffraction patterns of various complexity (also small protein powder diffraction patterns).

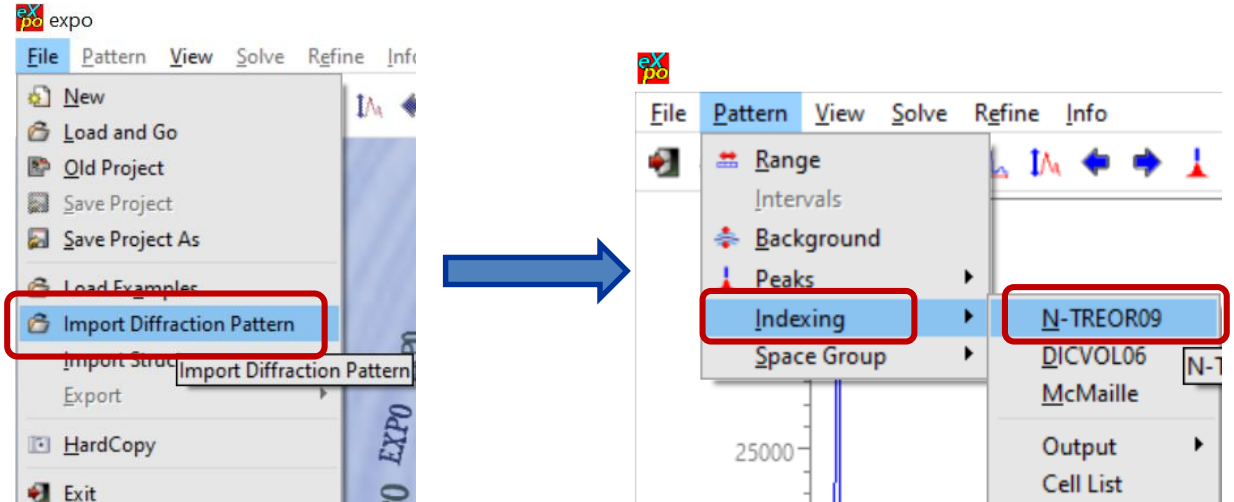
How to carry out the indexing step by EXPO:

- By loading an external input file, requiring minimal information and consisting of commands (the first character in the line must be '%') and directives (sub-commands following the related command):

```
%Structure clomipra
%Job Clomipramine hydrochloride (C19H24ClN2Cl)
%Data
Pattern    pd_0023.pow
Wavelength 1.54056
%ntreor
%continue
```

Minimal information

- By graphic interface:



Indexing by EXPO: some applications

- **A simple example:** successful indexing by default in case of ***Cerium dioxide***
- **An example of more than one plausible cell supplied by default: *KUOS***
- **An example of successful indexing in case of small protein powder diffraction data: *HEWL***
- **Some examples of indexing failure by a default run: *LEFEBVRE, LaTi, DILTIA;*** suggestions to overcome the pitfalls by a not default run of ***EXPO.***

In a default Indexing run by EXPO

The plausible cells searched by *N-TREOR09* satisfy some conditions, *among them*:

- ***The maximum value of the cell volume should be* 4000 Å^3**
- ***The maximum value of cell axis should be* 35 Å**
- ***The maximum allowed number of unindexed lines should be* 1**
- ***The minumum value of the de Wolff figure of merit M_{20} should be* 10**

In case of indexing failure it could be useful to change one (of more than one) of the above conditions in order to increase the probability of finding the correct cell.

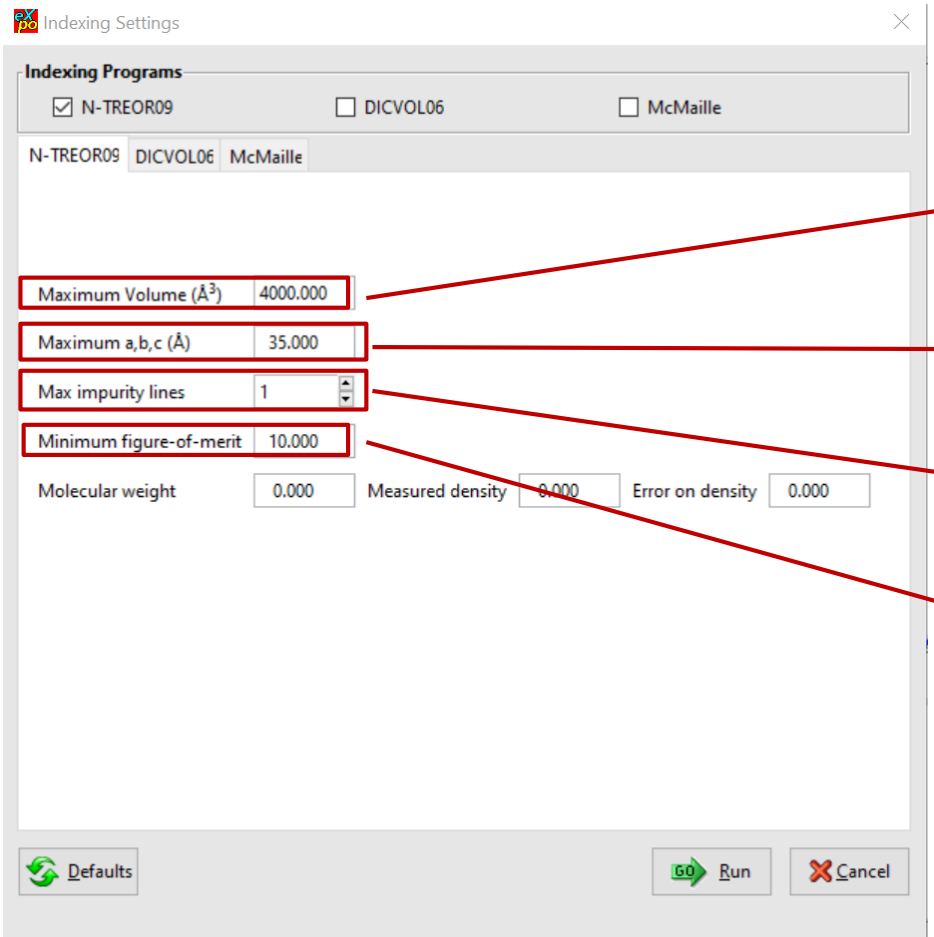
It can be done:

- ***By graphic interface***
- ***By introducing in the input file one (or more than one) N-TREOR09 directive* (see the *EXPO* Documentation for the syntax and meaning of all *N-TREOR09* directives).**

Not default Indexing run by EXPO:

By graphic interface

By introducing in the input file one (or more than one) *N-TREOR09* directive (see the *EXPO* Documentation)



It corresponds to the directive **VOL= 4000,**

It corresponds to directive **CEM=35,**

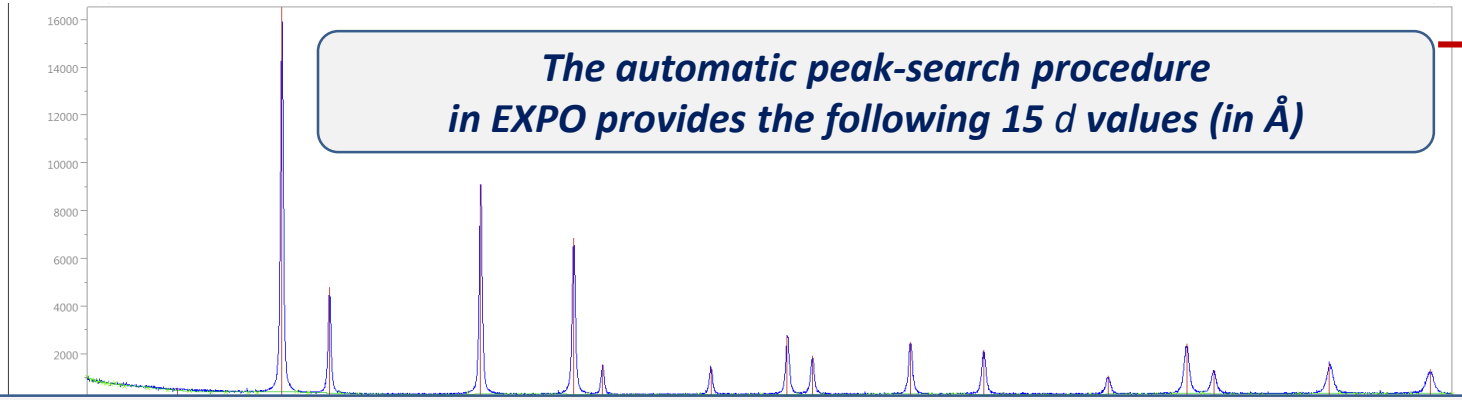
It corresponds to directive **NIX = 1,**

It corresponds to directive **MERIT=10,**

Indexing by EXPO

A simple indexing case: Cerium dioxide - Conventional X-ray data

Published cell: a= 5.40972 Å – cubic system (CIF file No. 4343161 of COD database)



The automatic peak-search procedure in EXPO provides the following 15 *d* values (in Å)

3.125367
2.706134
1.914097
1.631856
1.562573
1.352750
1.241911
1.210190
1.104908
1.041464
0.956799
0.914638
0.901849
0.855865
0.825362

The indexing program starts its run by assuming the largest symmetry (the cubic one). In this case the indexing equation is simple:

$$Q(hkl) = 10^4\ (1/d_{hkl})^2 = (h^2 + k^2 + l^2)\ A_{11}, \quad \text{where } A_{11} = 10^4\ a^{*2}$$

The first experimental line is selected as basis line to which trial (hkl) values are assigned, so that: $(hkl)_{max} = (6\ 6\ 6)$ and $(h + k + l)_{max} = 8$;

$$1/d_{hkl} = (h^2+k^2+l^2)^{1/2}\ a^*$$

$$a^*=1/a$$

$$\Rightarrow\ a = (h^2 + k^2 + l^2)^{1/2}\ d_{hkl}$$

The correct cell if obtained when the trial Miller indices (1 1 1) are assigned to the first line (d_{111}= 3.125367 Å)

$$\Rightarrow\ a = 3^{1/2}\ d_{111} = 5.41329\ Å$$

A simple indexing case:

Cerium dioxide- Conventional X-ray data

Published cell: a= 5.40972 Å – cubic system (CIF file No. 4343161 of COD database).

The trial cell parameter a= 5.41329 Å is used for indexing the rest of lines.

The cell is accepted and refined if all the experimental lines are indexed (only one unindexed line is tolerated in a default run of N-TREOR09).

3.125367
2.706134
1.914097
1.631856
1.562573
1.352750
1.241911
1.210190
1.104908
1.041464
0.956799
0.914638
0.901849
0.855865
0.825362

The final refined cell parameter is a = 5.41216 Å
M_{20} = 276, no unindexed lines.

For each observed line the corresponding Miller indices are given

d_{hkl}	h k l	d_{hkl}	h k l	d_{hkl}	h k l
3.1254	1 1 1	1.3528	4 0 0	0.9568	4 4 0
2.7061	2 0 0	1.2419	3 3 1	0.9146	5 3 1
1.9141	2 2 0	1.2102	4 2 0	0.9018	6 0 0
1.6319	3 1 1	1.1049	4 2 2	0.8559	6 2 0
1.5626	2 2 2	1.0415	5 1 1	0.8254	5 3 3

An example of more than one plausible cell supplied by *EXPO*: *KUOS*

*KUOS** – Deuterated potassium uranyl phosphate trihydrate – $KUO_2PO_4 \cdot 3D_2O$

Synchrotron data - Published cell : ***6.99379 6.99379 17.78397 90.00 90.00 90.00***

```
%Structure kuos
%job Deuterated potassium uranyl phosphate trihydrate -  KUO2PO4*3D2O
%data
pattern kuos.pow
wavelength 1.45072
Synchrotron
filetype double
%ntreor
%continue
```

**Fitch, A.N. & Cole M. (1991). Mater. Res. Bull. 26, 407-414.*

An example of more than one plausible cell supplied by *EXPO*: *KUOS*

KUOS – Deuterated potassium uranyl phosphate trihydrate – $KUO_2PO_4 \cdot 3D_2O$

N-TREOR09 supplies more than one plausible cell:

Plausible cell parameters

Select cell

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	M20	FOMnew	Mc20	shift	NIX	Symmetry Info
1	N	6.99361	6.99361	17.78342	90.000	90.000	90.000	869.8	107.00	8.164	-	0.000	0	P n c c
2	N	9.89023	9.89023	8.89187	90.000	90.000	90.000	869.8	116.00	7.249	-	0.000	1	P n b _
3	N	6.99742	6.99742	8.89775	90.000	90.000	90.000	435.7	29.00	3.604	-	0.038	1	P n _ _

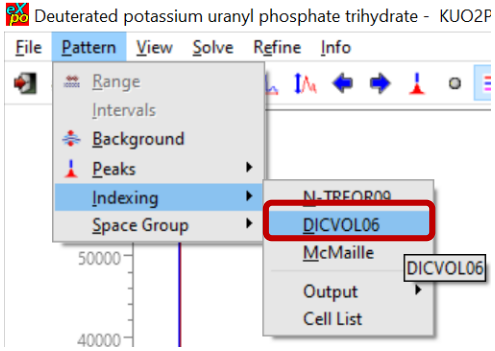
Published cell : 6.99379 6.99379 17.78397 90.00 90.00 90.00

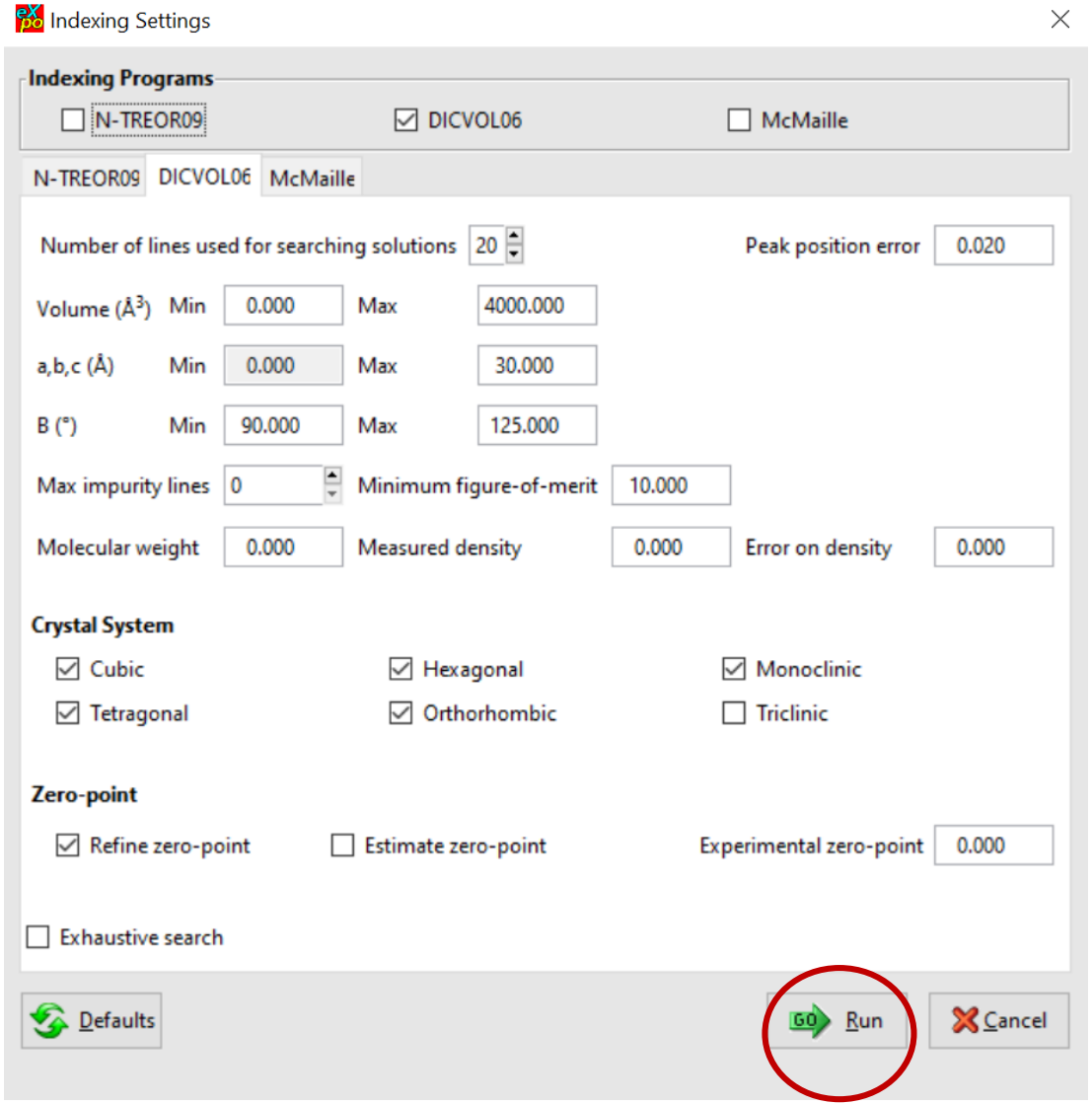
To run *DICVOL06*

OK | Cancel | Export

WRIP20 (FOMnew in the window) is able to rightly recognize the correct cell, on the contrary of M_{20}.

An example of more than one plausible cell supplied by *EXPO*: *KUOS*





An example of multiple cells supplied by default: *KUOS*

Published cell : 6.99379 6.99379 17.78397 90.00 90.00 90.00

Plausible cell parameters

Select cell

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	M20	FOMnew	Mc20	shift	NIX	Symmetry Info
1	N	6.99361	6.99361	17.78342	90.000	90.000	90.000	869.8	107.00	8.164	-	0.000	0	P n c c
2	N	9.89023	9.89023	8.89187	90.000	90.000	90.000	869.8	116.00	7.249	-	0.000	1	P n b _
3	N	6.99742	6.99742	8.89775	90.000	90.000	90.000	435.7	29.00	3.604	-	0.038	1	P n _ _
4	D	6.99350	6.99350	17.78360	90.000	90.000	90.000	869.8	109.90	-	-	0.000	0	Tetra
5	D	7.10530	17.78290	5.02430	90.000	100.181	90.000	624.8	69.00	-	-	0.000	10	Mono
6	D	7.03710	17.78270	6.29460	90.000	96.417	90.000	782.8	64.30	-	-	0.002	4	Mono
7	D	14.05100	8.89120	5.12050	90.000	95.515	90.000	636.8	34.10	-	-	0.002	8	Mono
8	D	9.09240	8.89220	8.13690	90.000	109.589	90.000	619.8	27.20	-	-	-0.009	6	Mono
9	D	9.96000	8.89170	7.03980	90.000	96.650	90.000	619.3	22.80	-	-	0.001	6	Mono
10	D	9.40040	8.88940	7.93030	90.000	109.804	90.000	623.5	22.00	-	-	0.016	6	Mono

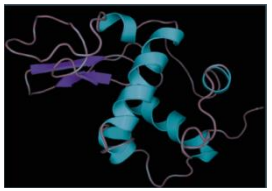
OK Cancel Export

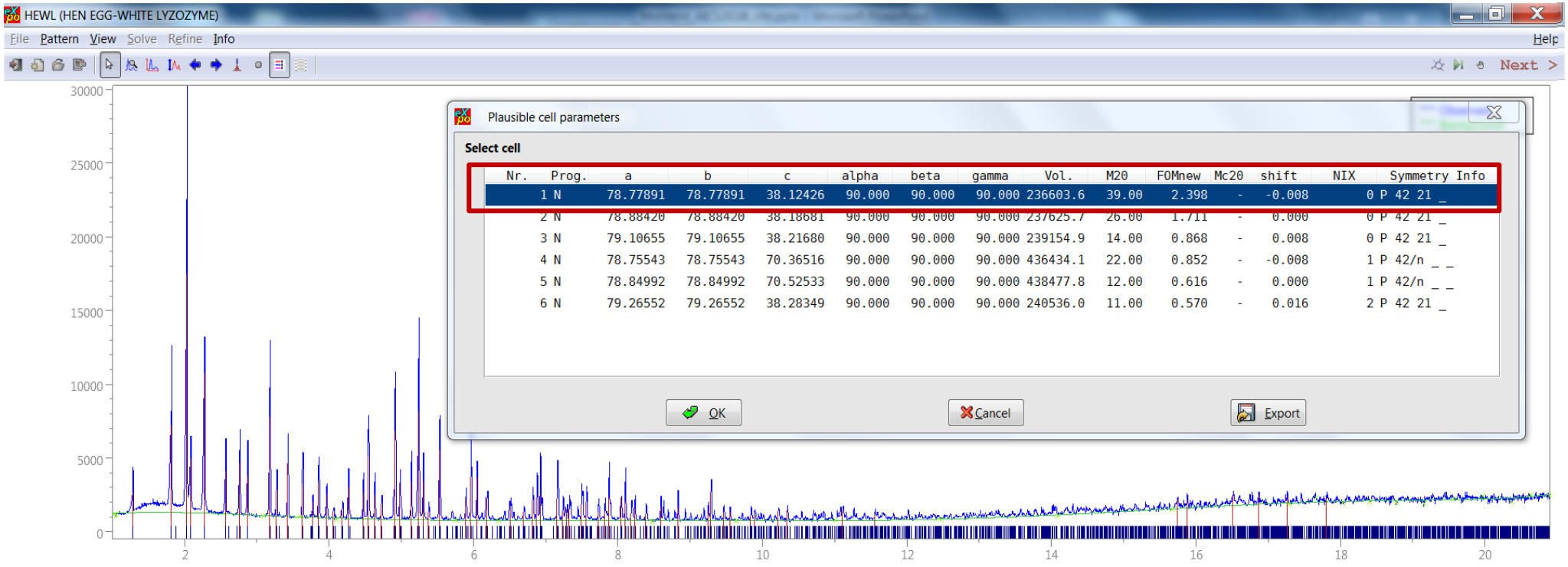
Similar results obtained by more than one indexing program → Increase of the confidence in the indexing results

An example of successful indexing in case of small protein powder diffraction data: *HEWL*

HEWL (HEN EGG-WHITE LYSOZYME)*

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





* Synchrotron data kindly provided by Dr. Irene Margiolaki, Patras, Greece.

Basso, S., Fitch, A. N., Fox, G. C., Margiolaki, I. & Wright, J. P. (2005). *Acta Cryst.* D**61**, 1612-1625.

Indexing by EXPO: some suggestions

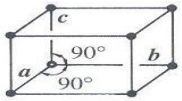
In case of indexing failure of a default run of *EXPO:*

- Try a different selection of the experimental peaks used for the indexing process (the failure could be due to the presence of impurity peaks and/or to the absence of one or more than one useful peak);
- Use one or more than one *N-TREOR09* directive (see the *EXPO* Documentation) in order to modify the default choices of the indexing program;
- Try to find the correct cell by applying alternative indexing programs (*e. g.*, *DICVOL06*, *McMaille*).

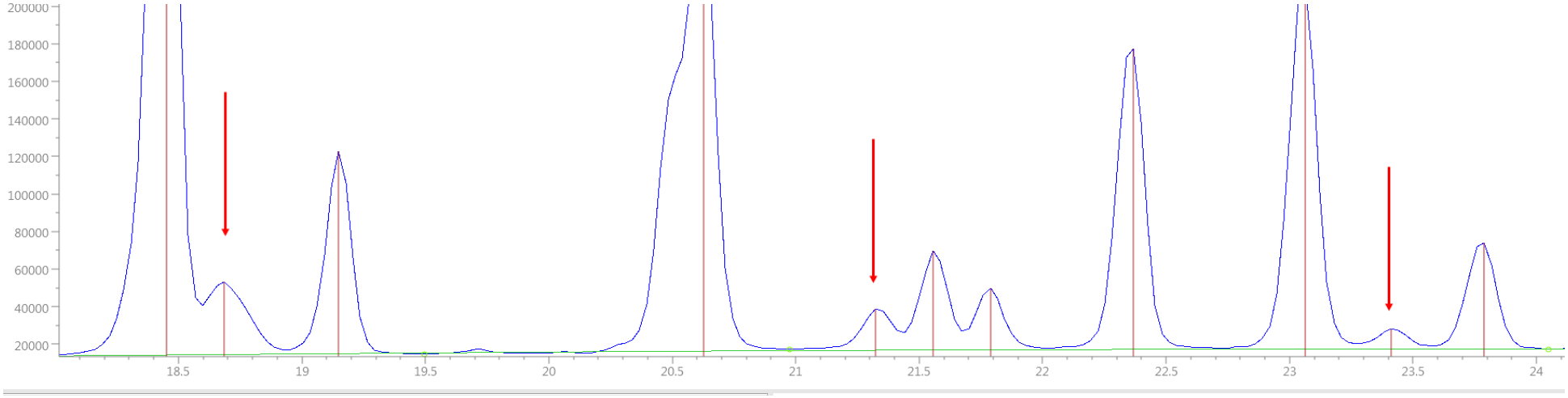
Indexing failure of a default run of EXPO:

LEFEBVRE – m-toluidine – C_7H_9N - (Conventional X-ray data)*

Published Cell 24.8727 5.8073 8.7615 90. 100.062 90.0



The set of selected peaks used for indexing contains more than one impurity peak belonging to the low-2θ region.



In a default run the maximum allowed number of unindexed lines is 1.

→ Indexing failure

**Rukiah, M., Lefebvre, J., Descamps, M., Hemon, S. , Dzyabchenko, A. (2004). J. Appl. Cryst. 37, 464-471.*

How to avoid the indexing failure by EXPO:

LEFEBVRE – m-toluidine – C_7H_9N - (Conventional X-ray data)

Published Cell 24.8727 5.8073 8.7615 90. 100.062 90.0

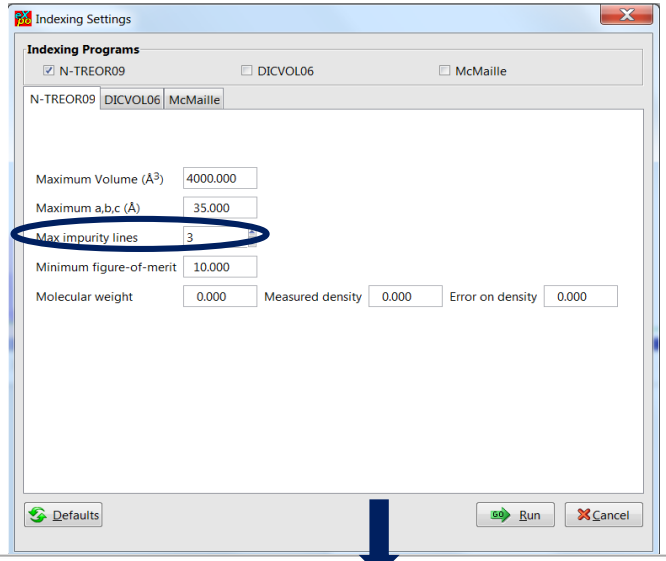
Non-default input file:

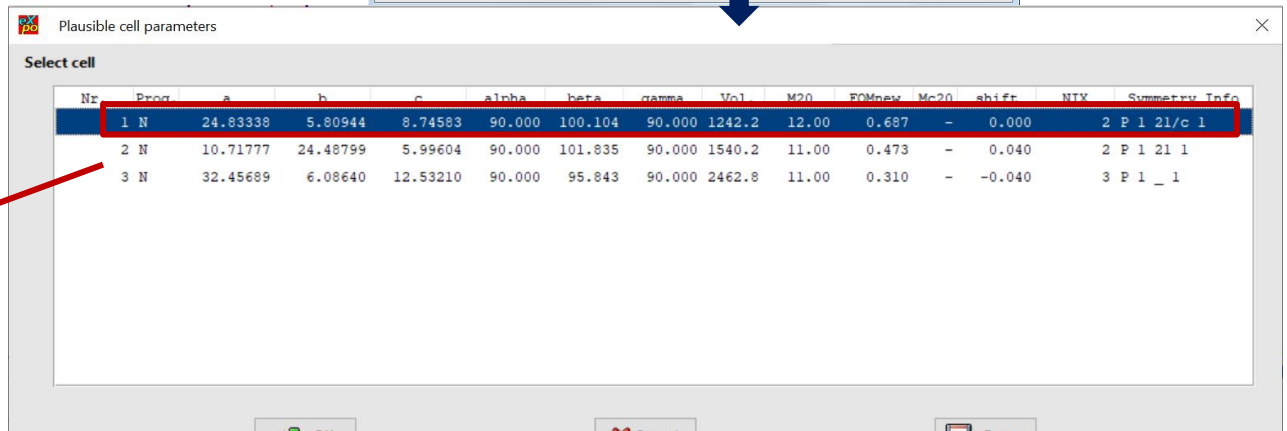
```
%Structure lefebvre
%Job 3-methylbenzenamine (m-Toluidine) - (C7H9N)
%Data
Pattern lefebvre.pow
Wavelength 1.54056
%Ntreor
nix=3,
%Continue
```

Maximum number of allowed unindexed lines: 3

Correct indexing by *EXPO*

The default conditions can be changed also by graphic interface





How to avoid the indexing failure by EXPO:

LEFEBVRE – m-toluidine – C_7H_9N - (Conventional X-ray data)

Published Cell 24.8727 5.8073 8.7615 90. 100.062 90.0

default input file:

Non-default peak-search

Plausible cell parameters

Select cell

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	M20	FOMnew	Mc20	shift	NIX	Symmetry Info
1	N	24.85401	5.81185	8.75840	90.000	100.065	90.000	1245.7	18.00	0.893	-	0.040	0	P 1 21/c 1
2	N	12.24354	6.93470	5.94527	98.232	99.790	91.657	491.6	16.00	0.571	-	-0.040	1	P 1
3	N	8.13361	24.50180	5.69665	90.000	97.163	90.000	1126.4	19.00	0.570	-	0.040	3	P 1 21/a 1
4	N	10.71134	24.48181	6.00968	90.000	101.751	90.000	1542.9	9.00	0.429	-	0.040	1	P 1 21 1
5	N	13.12868	12.20253	11.30472	90.000	112.358	90.000	1674.9	11.00	0.423	-	0.000	1	P 1 _ 1
6	N	11.69633	24.57158	4.97516	90.000	96.080	90.000	1421.8	13.00	0.405	-	0.080	1	P 1 21/n 1
7	N	18.73197	5.62725	12.27630	90.000	101.837	90.000	1266.5	9.00	0.367	-	-0.040	1	P 1 _ 1

OK Cancel Export

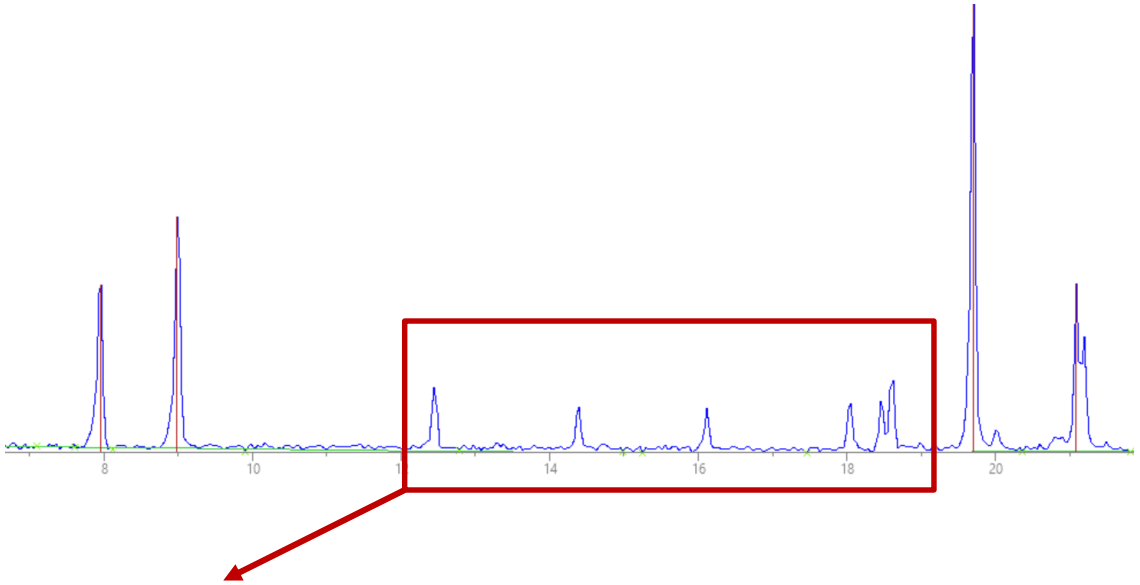
Correct indexing by *EXPO*

Indexing failure of a default run of EXPO:

LaTi* – non-perovskite compound $LaTi_2Al_9O_{19}$ – (Conventional X-ray data)

Published Cell 22.59355 10.99919 9.72968 90.0 98.5634 90.0

%structure LaTi
%job LaTi
%data
pattern LaTi.rtv
%ntreor
%continue



Due to a large number of experimental diffraction peaks, the automatic peak search by ***EXPO*** neglects some small intensity peaks belonging to the low 2θ region.

→ Indexing failure

**Kasunic, M., Meden, A., Skapin, S. D., Suvorov, D., Golobic , A. (2011). Acta Cryst. B67, 455–460.*

How to avoid the indexing failure by EXPO:

LaTi – non-perovskite compound $LaTi_2Al_9O_{19}$ – (Conventional X-ray data)

By modifying the Peak Search conditions to select the small intensity peaks belonging to the low 2θ region

Plausible cell parameters

Select cell

Nr.	Prog.	a	b	c	alpha	beta	gamma	Vol.	M20	FOMnew	Mc20	shift	NIX	Symmetry Info
1	N	22.60105	11.00379	9.73009	90.000	98.588	90.000	2392.7	13.00	1.430	-	0.040	0	C 1 c 1
2	N	10.97106	12.52613	9.71741	97.747	90.012	64.149	1188.2	14.00	0.669	-	0.000	0	P 1
3	N	12.57451	11.03190	9.74620	90.000	97.652	64.210	1204.0	12.00	0.512	-	0.080	0	P 1
4	N	18.61060	9.85008	12.71730	90.000	96.268	90.000	2317.3	9.00	0.409	-	0.000	0	P 1 _ 1

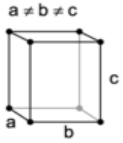
Published Cell *22.59355* *10.99919* *9.72968* *90.0* *98.5634* *90.0*

OK Cancel Export

Correct indexing by *EXPO*

Indexing failure of a default run of EXPO:

DILTIA – Diltiazem hydrochloride – $C_{22}H_{27}N_2O_4S{\cdot}Cl$*



Conventional X-ray data - Published cell: 42.190 9.075 6.037 90. 90. 90.

Default input file:

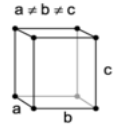
%Structure diltia
%Job Diltiazem hydrochloride
%Data
Pattern diltia.pow
Wavelength 1.54056
%ntreor
%continue

Default maximum cell axis value = 35 Å → Indexing failure

** Florence et al., (2005). J. Appl. Cryst. 38, 249–259.*

How to avoid the indexing failure by EXPO:

DILTIA – Published cell: 42.190 9.075 6.037 90. 90. 90.



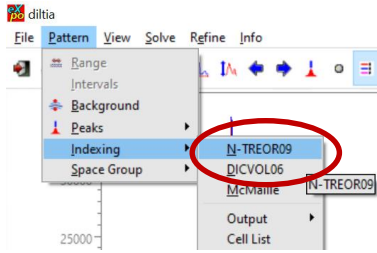
Non-default input file:

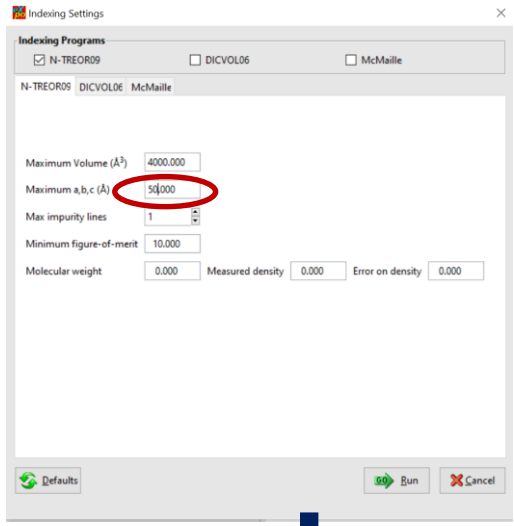
%Structure diltia
%Job diltia
%Data
Pattern diltia.pow
Wavelength 1.54056
%ntreor
CEM=50,
%continue

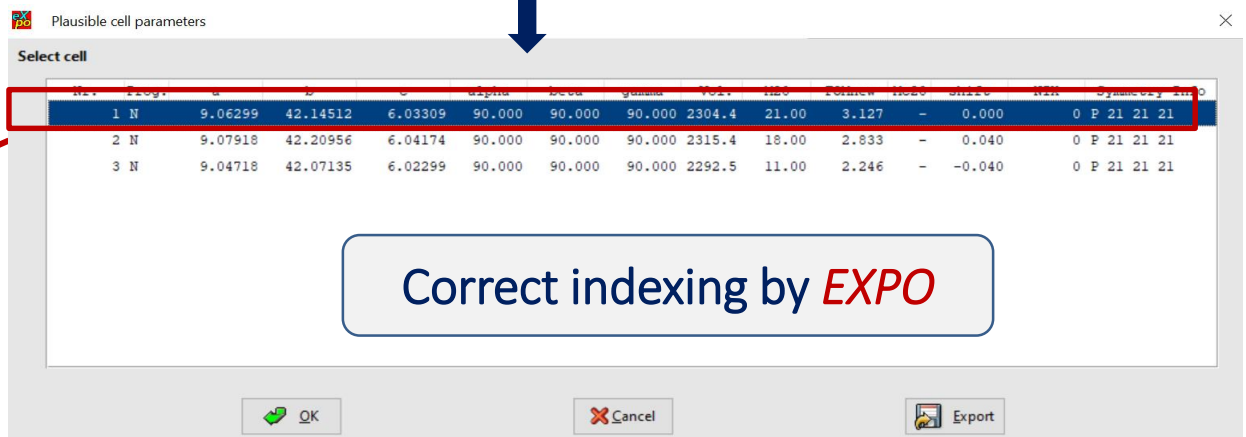
Maximum cell axis value = 50 Å

Correct indexing by *EXPO*

The default conditions can be changed also by graphic interface







Correct indexing by *EXPO*

Final remarks

The basic recommendation is to collect high quality data with no impurity peaks.

'The "indexing problem" is essentially a puzzle: it cannot be stated in a rigorous term (…). It would be quite an easy puzzle if errors of measurements did not exist'
(de Wolff, 1957);

- Check the selected peaks and in case of indexing failure, try with a different peak selection;
- In case of indexing failure, try a not default indexing process by suitable directives;
- Even if at least one plausible cell has been found, it can be useful to use more than one indexing program (especially if the different programs are based on different and/or complementary indexing procedures);

↓

similar results obtained by different programs strengthen the confidence in the correctness of the cell.

The certainty that the cell is the true one is attained only at the end of the crystal structure solution, once the correct crystal structure has been determined.

FOR YOUR ATTENTION