

EXPO2014 TUTORIAL

EXPO2014 download and installation

1. Log in to <http://www.ba.ic.cnr.it/content/expo-downloads> if you already have an account, or register on <http://www.ba.ic.cnr.it/user/register> if you do not already have an account. After registration you will receive an e-mail confirmation and you will be allowed to download
2. Logged in users can freely download the latest version of *EXPO2014* for all supported operating systems: Windows XP and above, Linux, Mac OS X.
3. Install *EXPO2014* following the instructions on <http://www.ba.ic.cnr.it/softwareic/expo/expo2014-installation/>

The examples reported in this tutorial correspond to simple structures, which do not require a long execution time by *EXPO*.

The EXPO2014tutorial directory contains the following folders and files:

1) EXPOLiterature folder.

It contains the papers concerning the structures, which are mentioned in this tutorial.

WARNING: when you load a file for running *EXPO*, please, remember to choose the correct ‘Current Directory’.

2) EXPOdefault run folder.

It contains the following files: **and2.exp** [the input file for a default run of *EXPO*, from indexation to structure model optimization, in the case of 2,6-diamino-5-hydroxy-3-nitro-4*H*-pyrazolo[1,5-*a*]-pyrimidin-7-one monohydrate ($C_6H_6N_6O_4 \cdot H_2O$)]; **and2.pow** (the file containing the experimental profile counts); **and2.fra** (the file of the fractional coordinates and the isotropic thermal parameters of the true model).

To run *EXPO* on and2, you can create a new project or you can use the input ascii file ‘and2.exp’.

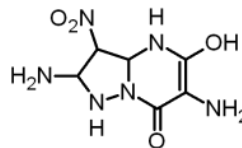
In the first case:

- i) Click on *EXPO* icon;
- ii) ‘File’ in the upper Menu;
- iii) ‘New’;
- iv) Fill in the boxes (Structure Name: and2; Profile Counts filename: select ‘and2.pow’ from the current directory);
- v) Select ‘Synchrotron’ and provide the Wavelength value in Å (0.950436);
- vi) ‘Save’ (the file ‘and2.exp’ has been created);
- vii) ‘Go’;
- viii) Click on ‘Next’ button to go on continuously;
- ix) Select ‘Yes’ for going on with the Indexing procedure;
- x) After that the cells have been provided and the most probable cell has been selected, click on ‘OK’;
- xi) Now you have to provide the cell content: (C6 H8 N6 O5)4;
- xii) ‘OK’;
- xiii) ‘Next’; Click on ‘Next’ button to go on continuously;
- xiv) Select Extinction Group: ‘P 1 21/n 1’ and Space Group Name: ‘p 21/n’;
- xv) ‘OK’ (the file ‘and21.exp’ has been created);
- xvi) ‘OK’;
- xvii) Click on ‘Next’ button to go on continuously.

In the second case:

The input file 'and2.exp' consists of the following lines:

```
%structure and2
%job and2 - C6H6N6O4 · H2O
%data
wavelength 0.950436
synchrotron
pattern and2.pow
%ntreor
%continue
```



(See the help on line for the meaning of each command line.)

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Use 'and2.exp' as *EXPO* Input Filename and give the output filename you like;
- v) 'Go';
- vi) Click on 'Next' button to go on;
- vii) 'Go';
- viii) Click on 'Next' button to go on continuously;
- ix) Select 'Yes' for going on with the Indexing procedure;
- x) After that the cells have been provided and the most probable cell has been selected, click on 'OK';
- xi) Now you have to provide the cell content: (C6 H8 N6 O5)4;
- xii) 'OK';
- xiii) 'Next'; Click on 'Next' button to go on continuously;
- xiv) Select Extinction Group: 'P 1 21/n 1' and Space Group Name: 'p 21/n';
- xv) 'OK' (the file 'and21.exp' has been created);
- xvi) 'OK';
- xvii) Click on 'Next' button to go on continuously.

GENERAL

The structural model obtained at the end of a run of *EXPO* can be compared with another one (*f.e.*, the true structure contained in *.fra file) by following:

- a) if the Jav window, usually showing the obtained structure model, has been closed:
 - i) 'View' in the upper Menu;
 - ii) 'Jav Molecular Viewer';
 - iii) 'Tools' from the Jav Menu;
 - iv) 'Overlay Structures';
 - v) Select the file name containing the comparing model (and2.fra);
 - vi) 'OK';
 - vii) 'Tools';
 - viii) 'Overlay Info';
 - ix) To delete the imported model (and2.fra): 'Tools' and 'Delete Model';
- b) if the Jav window is available start from the point iii).

3) EXPOnodefaultrun folder.

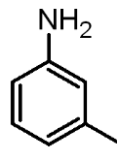
It contains:

a) EXPOIndexing folder.

It contains the following files: **lefebvre.exp** (an example of *EXPO* input file for a not default indexing by *N-TREOR09*; the non-default directive '*nix=3,*' has been introduced); **lefebvre.pow** (the file containing the experimental profile counts); **lefebvre.fra** (the file of the fractional coordinates of the true model).

The input file 'lefebvre.exp' consists of the following lines:

```
%structure lefebvre
%job m-toluidine (CH3C6H4NH2)
%initialize
%data
pattern lefebvre.pow
wavelength 1.54056
%ntreor
nix=3,
%continue
```



It can be easily verified that a default indexing process by *EXPO* is not able to find the correct monoclinic cell (published cell parameters: $a=24.8727$ Å, $b=5.8073$ Å, $c=8.7615$ Å, $\alpha=90.0^\circ$, $\beta=100.062^\circ$, $\gamma=90.0^\circ$). The introduction of the directive '*nix=3*' in the *EXPO* input file is successful.

The presence of impurity peaks is one of the possible reasons of indexing failure. In case of unsuccessful default indexation, if the presence of impurity peaks is hypothesized, a new indexing process can be performed by a no default run. In the *EXPO* input file, the directive '*nix=n,*' of the *%ntreor* command, can be used for increasing the number of allowed unindexed lines from 1 (*i.e.*, the default value for *N-TREOR09*) to *n*, with $n>1$.

To run EXPO on lefebvre:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Use 'lefebvre.exp' as *EXPO* Input Filename and give the output filename you like;
- v) 'Go';
- vi) Click on 'Next' button to go on continuously;
- vii) Click on 'Yes' each time a plausible cell is found;
- viii) After that the cells have been provided and the most probable cell has been selected, click on 'OK';
- ix) Provide the unit cell content '(C14 H18 N2)4';
- x) 'OK';
- xi) 'Next'; Click on 'Next' button to go on continuously;
- xii) Select Extinction Group: 'P 1 21/c 1' and Space Group Name 'p 21/c';
- xiii) 'OK' (the file 'lefebvre1.exp' has been created);
- xiv) 'OK';
- xv) Click on 'Next' button to go on continuously.

The first ranked cell is able to provide the correct solution.

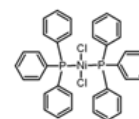
b) EXPOspacegroup folder.

It contains the following files: **nickel.exp** (the input file for space group determination once the pattern has been indexed. It contains the non-default directive '*findspace*');

nickel.pow (the file containing the experimental profile counts); **nickel.fra** (the file containing the fractional coordinates and the isotropic thermal parameters of the true model). The correct space group is P2/c, whose extinction group is 'P 1 c 1'. The largest probability value extinction group supplied by *EXPO* at the end of the space group determination step is 'P 1 21/c 1' (while the correct extinction group 'P 1 c 1' is the second one in the list of extinction groups ranked according to the probability value). In order to recognize the correct extinction group a visual inspection by graphic interface is strongly suggested by verifying if the extinction conditions, stated by the most probable extinction group, agree with the experimental pattern.

The input file 'nickel.exp' consists of the following lines:

```
%Structure nickel
%Job Dichlorobis(triphenylphosphine)
nickel(II) (C36 H30 Cl2 Ni P2)
%Data
Cell 11.638 8.197 17.388 90 107.03 90
Content C 72 Cl 4 Ni 2 P 4 H 60
Pattern nickel.pow
Wavelength 1.54056
findspace
%continue
```



The asymmetric unit is one half of the molecule.

The directive findspace of the %Data command activates the space group determination procedure.

To run *EXPO* on nickel:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Use 'nickel.exp' as *EXPO* Input Filename and give the output filename you like;
- v) 'Go';
- vi) 'Next' until the automatic space group determination procedure is performed and a window shows the extinction groups listed according to the calculated probability;
- vii) Click on the button located above the 'OK' and 'Cancel' buttons, in order to access to the list of systematically absent reflections for the most probable extinction group;
- viii) Click on 'Num' column in order to sort the absent reflections in terms of increasing sequential numbers, corresponding to increasing 2θ values;
- ix) Zoom the pattern in the low 2θ region 10° - 11° (i.e., by using the 'Zoom' button from the main window) and check if some reflections belonging to this region are also in the list of absent reflections (by marking the reflection in the list it is shown in the pattern). Reflections n. 5 (101) and n. 6 (010) are systematically absent due to the presence of the *c* glide and of the 2_1 axis, respectively. But their absence doesn't agree with the experimental pattern, due to the presence of a corresponding evident observed peak;
- x) Click on 'Close' in order to check the second extinction group in the list;
- xi) Select by mouse the second extinction group (P 1 c 1) and repeat the steps vii)-ix). It is easily seen that only the reflection n. 5 (101) is now absent (due to the *c* glide presence) while the reflection n. 6 (010) is present, in agreement with the experimental pattern. Assuming that the pattern doesn't contain impurity peaks, this visual inspection suggests that the correct extinction group isn't the first one in the list but the second one. By selecting the P 1 c 1 extinction group, two space groups become available: P c and P 2/c;

- xii) Select the space group P 2/c having a highest frequency in the Cambridge Structural Database (CSD);
- xiii) 'OK'; (the file 'nickell.exp' has been created);
- xiv) Click on 'Next' button to go on continuously.

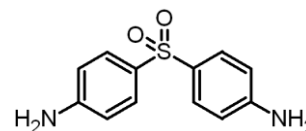
The final structure model obtained by *EXPO* at the first set of phases (default choice), is not interpretable. You have to try different phasing trials by applying the procedure described at the next point c).

c) EXPOAlltrials folder.

It contains the following files: **dapsone.exp** [the input file for the default run of *EXPO* in case of dapsone ($C_{12}H_{12}N_2O_2S$), after that the cell and the space group have been determined]; **pd_0005.xye** (the file containing the experimental profile counts); **dapsone_true.cif** (the file of the fractional coordinates of the true model).

The input file 'dapsone.exp' consists of the following lines:

```
Structure dapsone
%Job dapsone C12 H12 N2 O2 S
%Data
Cell 25.538 8.061 5.762 90 90 90
SpaceGroup p 21 21 21
Content (C12 H12 N2 O2 S) 4
Pattern pd_0005.xye
Wavelength 1.54056
%continue
```



To run EXPO on dapsone:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Use 'dapsone.exp' as *EXPO* Input Filename and give the output filename you like;
- v) 'Go';
- vi) Click on 'Next' button to go on continuously.

The structure model obtained at the end of Direct Methods procedure, executed on the first set of phases (default choice), is not interpretable. You have to try different phasing trials. The following steps automatically carry out the structure solution process for each stored phasing trial and finally the best structure model is selected:

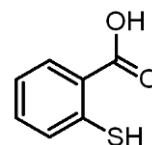
- vii) 'Solve' in the upper Menu;
- viii) 'Explore trials';
- ix) 'Select all new trial';
- x) 'Go'.

d) EXPORAMM folder.

It contains: **merca.exp** [the input file for the default run of *EXPO* in case of 2-Mercaptobenzoic acid ($C_7H_6O_2S$), after that the cell and the space group have been determined]; **merca.pow** (the file containing the experimental profile counts); **merca.fra** (the file of the fractional coordinates and the isotropic thermal parameters of the true model).

The input file 'merca.exp' consists of the following lines:

```
%Structure merca
%Job 2-Mercaptobenzoic acid (C7 H6 O2 S)
%Data
Cell 7.885 5.976 14.949 90.0 100.48 90
SpaceGroup p 21/c
Content (C7 H6 O2 S) 4
Pattern merca.pow
Wavelength 1.54056
%extraction
%continue
```



To run *EXPO* on merca:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Use 'merca.exp' as *EXPO* Input Filename and give the output filename you like;
- v) 'Go';
- vi) Click on 'Next' button to go on continuously.

The structure model obtained at the end of Direct Methods procedure, executed on the first set of phases (default choice), is not interpretable. The structural solution can be obtained by applying the RAMM procedure by choosing:

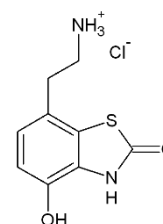
- vii) 'Solve' in the upper Menu;
- viii) Select 'RAMM Procedure'.

4) EXPORietveld folder.

It contains the following files: **ammonium.exp** [the *EXPO* input file, containing cell and space group information, for Rietveld refinement of 2-(4-Hydroxy-2-oxo-2,3-dihydro-1,3-benzothiazol-7-yl)ethylammonium chloride ($C_9H_{11}N_2O_2S \cdot Cl$)]; **ammonium.pow** [the file containing the experimental profile counts]; **ammonium.cif** (the CIF file of the structure model obtained by a solution process); **ammonium.fra** (the file of the fractional coordinates and the isotropic thermal parameters of the true model).

The input file 'ammonium.exp' consists of the following lines:

```
%Structure ammonium
%Job ethylammonium chloride (C9H11N2O2SCl)
%Data
Content c 36 n 8 o 8 s 4 cl 4 h 44
SpaceGroup p 21/c
Pattern ammonium.pow
Wavelength 1.54056
Cell 10.24252 14.63629 7.55266 90.000 109.407 90.000
%fragment ammonium.cif
%rietveld
```



To run *EXPO* on ammonium for refining a structure model saved in an external file (e.g., ammonium.cif):

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';

- iv) Use 'ammonium.exp' as *EXPO* Input Filename and give the output filename you like;
- v) Select 'Automatic refinement of profile' and 'Automatic refinement of structure' in the Rietveld Refinement' window;
- vi) 'Refine';
- vii) At the end of the automatic refinement process click on 'Quit' to leave the 'Rietveld Refinement' window.

To save the refined model in a CIF file:

- viii) 'File' in the Jav Menu;
- ix) 'Export Structure' and give the CIF filename you like.

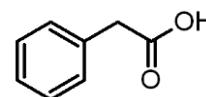
The refinement process by the Rietveld method can be carried out directly on the structure model provided by *EXPO* without importing the CIF file.

5) **EXPOdirectspace** folder contains:

- a) **phenyla** folder: **phenyla.exp** [the *EXPO* input file for structure solution of phenylacetic acid ($C_8H_8O_2$) by direct space methods, once cell parameters and space group have been determined]; **pd_0009.xye** (the file containing the experimental profile counts); **Structure3D_CID_999.sdf** (the file containing the starting structure model); **phenyla_true.cif** (the file of the fractional coordinates of the true model).

The input file 'phenyla.exp' consists of the following lines:

```
%Structure phenyla
%Job Phenylacetic acid (C8 H8 O2)
%Data
Cell 10.226 4.967 14.467 90 99.25 90
SpaceGroup p 21/a
Pattern pd_0009.xye
%fragment Structure3D_CID_999.sdf
%sannel
```



%fragment Structure3D_CID_999.sdf is the command used to import the 3D starting structure model. The file Structure3D_CID_999.sdf has been downloaded from the pubchem database:

<https://pubchem.ncbi.nlm.nih.gov/compound/999#section=Top>

%sannel is the command enabling the access to the graphic interface of direct space methods (DSM).

The directive wavelength is not necessary because the information is provided in the first line of the 'pd_0009.xye' file.

To run *EXPO* on phenyla:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Import the input file 'phenyla.exp';
- v) Modify the resolution in the dialog window: 2.5 Å is enough to obtain accurate atomic positions.
- vi) Press the button 'Execute' in the dialog window to run DSM. 10 runs of the DSM procedure will be performed. For each *i-th* DSM run a CIF file ('structure_name_best*i*.cif') containing the best solution is created.

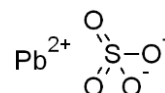
At the end of the 10 runs, the solution corresponding to the lowest cost function is automatically selected.

- vii) Click on the button 'Solutions' to explore the 10 best structural models, ranked according to the cost function. Click on 'Save', to select one of the ten best solutions;
- viii) To compare the selected best solution with the published model phenyla_true.cif click on 'Quit' (to leave the 'Global optimization dialog' window) and on 'Tools' > 'Overlay Structures' to load the phenyla_true.cif file.

b) PbSO₄ folder: pbso4.exp [the *EXPO* input file for structure solution of lead(II) sulfate (PbSO₄) by direct space methods, once cell parameters and space group have been determined]; **pbso4.dat** (the file containing the experimental profile counts); **pbso4.mol** (the file containing the starting structure model); **9015524.cif** (the file of the fractional coordinates and the isotropic thermal parameters of the true model, downloaded from the COD database <http://www.crystallography.net/cod/9015524.html>).

The input file 'pbso4.exp' consists of the following lines:

```
%structure pbso4
%job Lead(II) sulfate (PbSO4)
%data
Pattern pbso4.dat
Wavelength 1.54056
Cell 6.95802 8.48024 5.39754 90 90 90
SpaceGroup p b n m
%fragment pbso4.mol
%sannel
```



The starting structural model in the file **pbso4.mol** contains randomly located SO₄ tetrahedron and Pb isolated atom.

To run *EXPO* on pbso4:

- i) Click on *EXPO* icon;
- ii) 'File' in the upper Menu;
- iii) 'Load & Go';
- iv) Import the input file 'pbso4.exp';

Since atoms in special positions are expected, it is advisable to carry out the structure solution process applying the dynamical occupancy correction (DOC).

- v) To activate DOC click on the tab 'Internal DOF' button, then on the button 'Atomic Parameters and Dynamical Occupancy Correction', and finally on the check button 'D.O.C.', that enables to apply DOC to all the atoms. Alternatively, instead of activating DOC by graphic interface, you can modify the input file by adding the directive 'doc' after the command %sannel;
- vi) 'OK';
- vii) Click on the button 'Execute', in the dialog window, to run DSM. 10 runs of the DSM procedure will be performed. For each *i-th* DSM run a CIF file ('structure_name_best*i*.cif') containing the best solution is created.

At the end of the 10 runs, the solution corresponding to the lowest cost function is automatically selected.

- viii) Click on the button 'Solutions' to explore the 10 best structural models, ranked according to the cost function. Click on 'Save', to select one of the 10 best solutions;
- ix) Click on 'Quit' to leave the 'Global optimization dialog' window;
- x) On the Jav molecular viewer, click on 'Modify' > 'Delete Duplicate Atoms' and then on 'OK' to remove the oxygen atom in excess.

The correctness of the obtained final model provided by DSM can be verified comparing it with the true model in the file 9015524.cif.