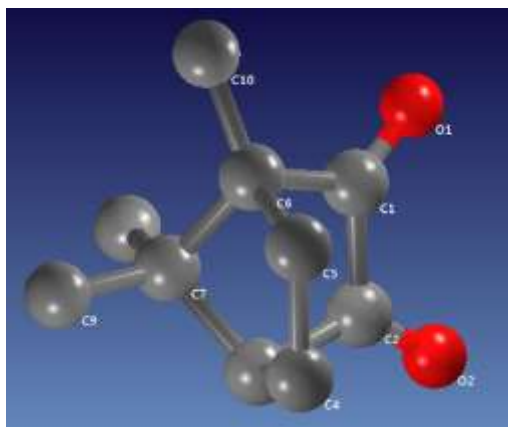


TUTORIAL EXPO: STRUCTURE MODEL OPTIMIZATION

The **Structure model optimization** folder contains:

- **COVMAP** folder

It contains: **camphor.exp** [the input file for the default run of *EXPO* in case of 1,7,7-Trimethylbicyclo(2.2.1) hepta-2,3-dione (C₁₀H₁₄O₂), after that the cell and the space group have been determined]; **camphor.pow** (the file containing the experimental profile counts); **camphor.fra** (the file of the fractional coordinates and the isotropic thermal parameters of the true model, hydrogen atoms excluded). The structure has not been published yet (courtesy by Dr. Michela Brunelli).



The input file ‘camphor.exp’ consists of the following lines:

```
%Structure camphor
%Job 1,7,7-Trimethylbicyclo(2.2.1)hepta-2,3-dione (C10 H14 O2)
%Data
Cell 12.008 11.488 6.631 90 91.613 90
SpaceGroup p 21/n
Content (C10H14O2) 4
Range 2.001 26.0
Pattern camphor.pow
filetype double
Wavelength 0.49002
Synchrotron
%continue
```

The range has been reduced to 26° because the signal is too noisy beyond that value. The directive ‘Filetype double’ has been introduced because of the format of camphor.pow file.

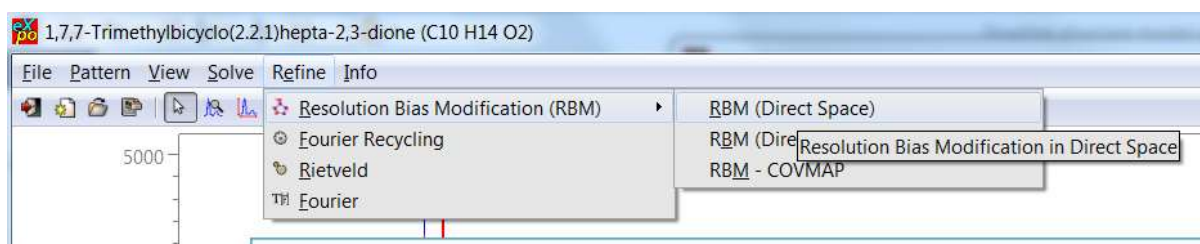
To run EXPO on camphor in default way:

- Click on EXPO icon
- **File** in the upper Menu
- **Load & Go**
- Use ‘camphor.exp’ as Input File and give the Output Filename you like (camphor.out is the default output file name)
- **Go**

- Click on **Next** to go on continuously until the end of the run.

The structure model obtained at the end of the Direct Methods procedure, executed on the first set of phases (default choice), is not interpretable.

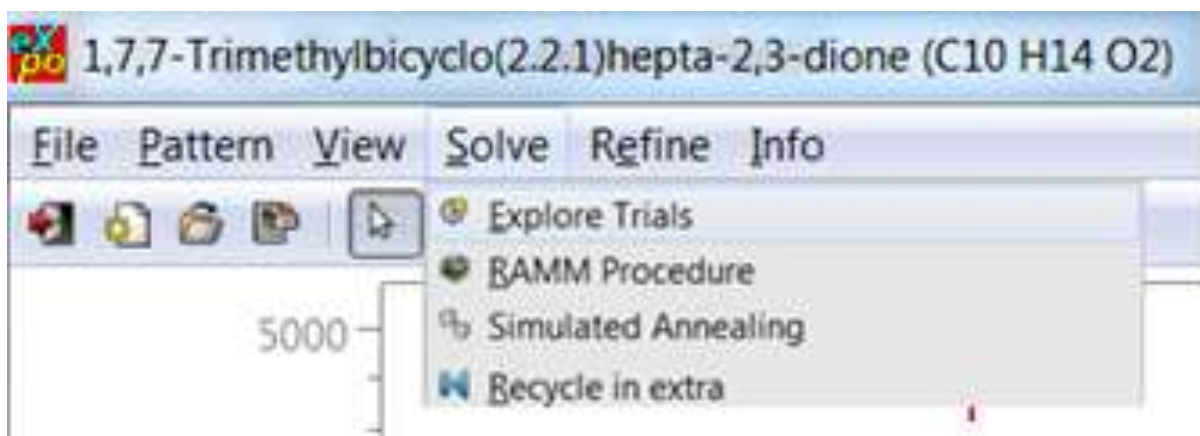
It is so rough and uninterpretable that is not advisable to try to improve it, for example, by cyclic application of RBM (RBM is advisable because the structure is organic). Indeed, by clicking on **Refine > Resolution Bias Modification (RBM) > RBM (Direct Space)** in the upper Menu



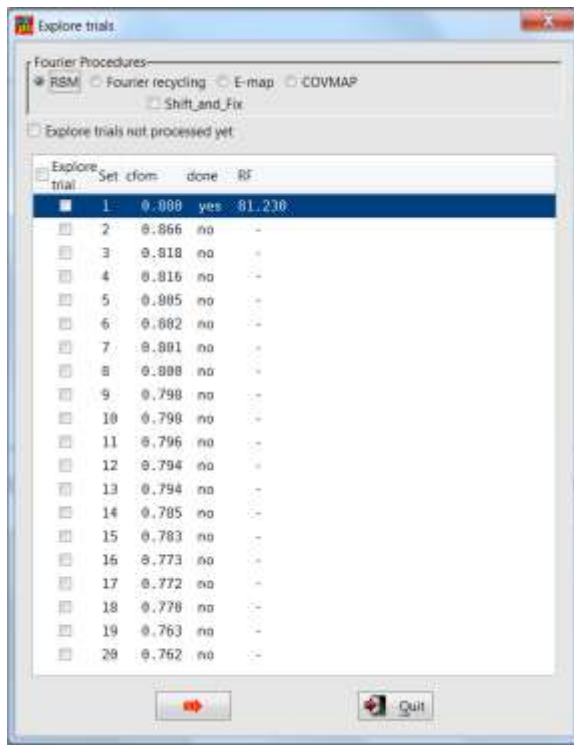
no improvement of the structure model is attained.

We can try to explore the other Direct Methods trials:

Solve > Explore Trials in the upper Menu



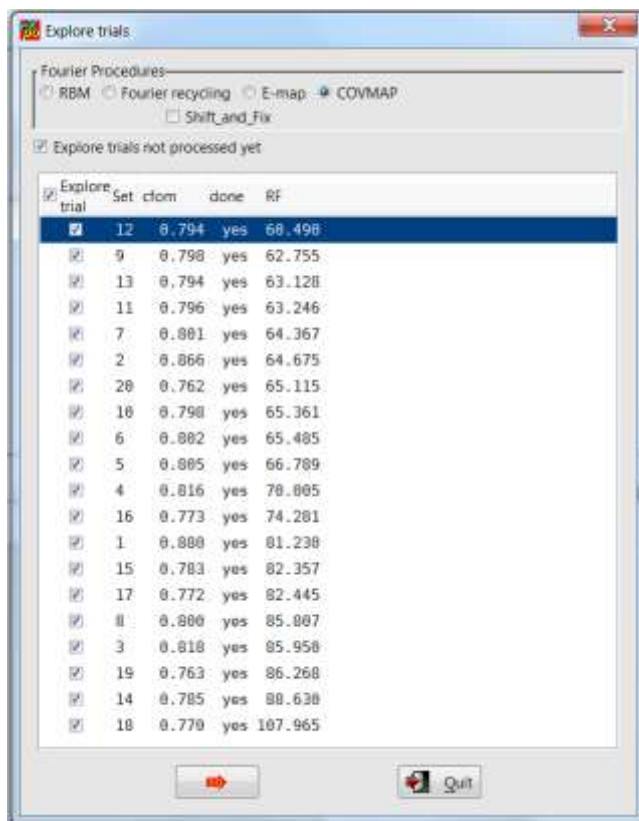
and exploring and ranking all the other 19 trials not processed in the standard run by Direct Methods (only the highest CFOM figure of merit phasing trial is automatically processed)



Click on **GO**.

The model first ranked by RF doesn't correspond to the correct solution.

The structure solution can be attempted by the optimization of **COVMAP** to be applied to the 20 Direct Methods phasing trials

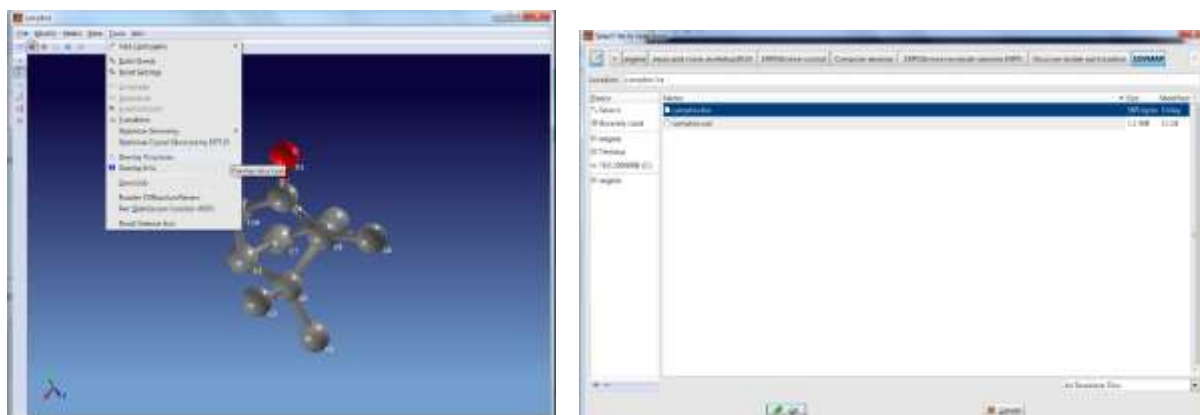


Click on **GO**. The execution of COVMAP requires a time longer than the standard RBM process.

Now the COVMAP model first ranked by RF corresponds to the correct solution.

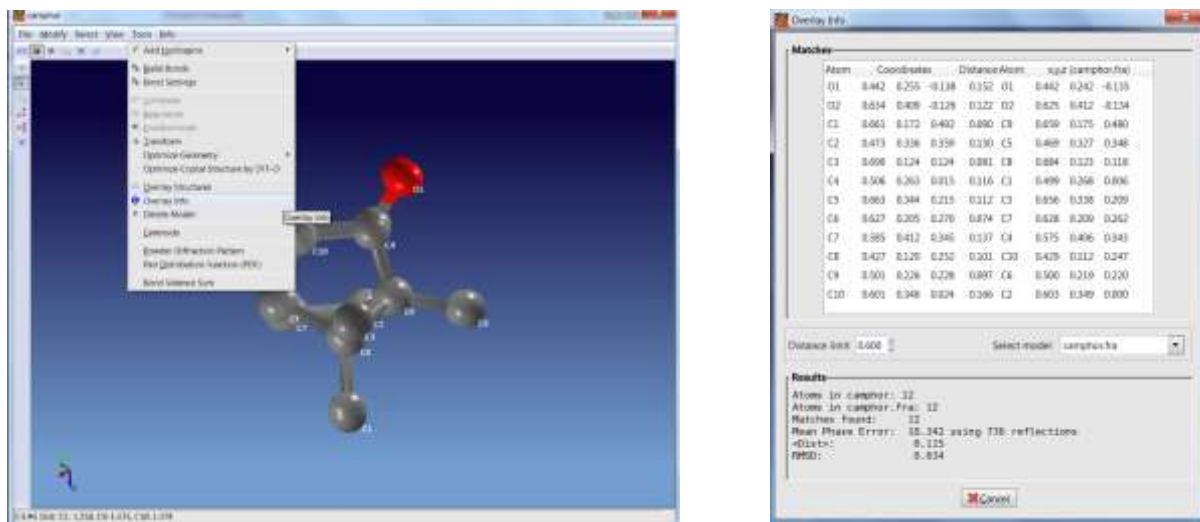
The obtained solution can be compared with the published fractional coordinates contained in the camphor.fra file. It can be done by the following graphic pathway:

Tools > Overlay structures in the upper Menu and select camphor.fra and **OK**



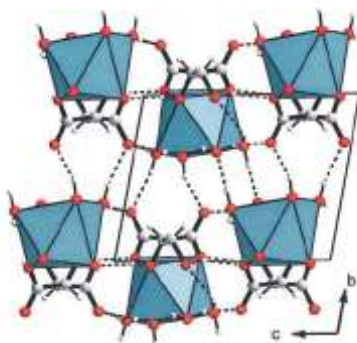
The two models are superimposed and information on comparison can be output:

Tools > Overlay Info in the upper Menu



- **SHIFT_AND_FIX** folder

It contains: **tartrate.exp** [the input file for the default run of *EXPO* in case of Calcium tartrate tetrahydrate ($\text{CaC}_4\text{H}_{12}\text{O}_{10}$), after that the cell and the space group have been determined]; **tartrate.pow** (the file containing the experimental profile counts); **tartrate.fra** (the file of the fractional coordinates and the isotropic thermal parameters of the true model, hydrogen atoms excluded); **tartrate.pdf** (the article in which the structure is cited).



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

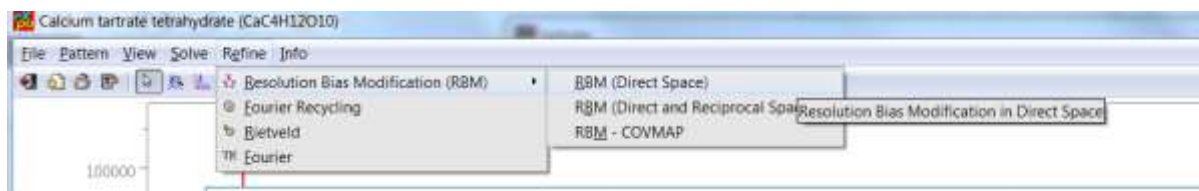
```
%Structure tartrate
%Job Calcium tartrate tetrahydrate (CaC4H12O10)
%Data
Cell 8.222 10.473 6.249 105.97 107.51 94.94
SpaceGroup p -1
Content (CaC4H12O10) 2
Pattern tartrate.pow
Wavelength 1.5418
Synchrotron
%continue
```

To run EXPO on tartrate in default way:

- Click on EXPO icon
- **File** in the upper Menu
- **Load & Go**
- Use 'tartrate.exp' as Input File and give the Output Filename you like (tartrate.out is the default output file name)
- **Go**
- Click on **Next** to go on continuously until the end of the run.

The structure model obtained at the end of the Direct Methods procedure, executed on the first set of phases (default choice), is not interpretable.

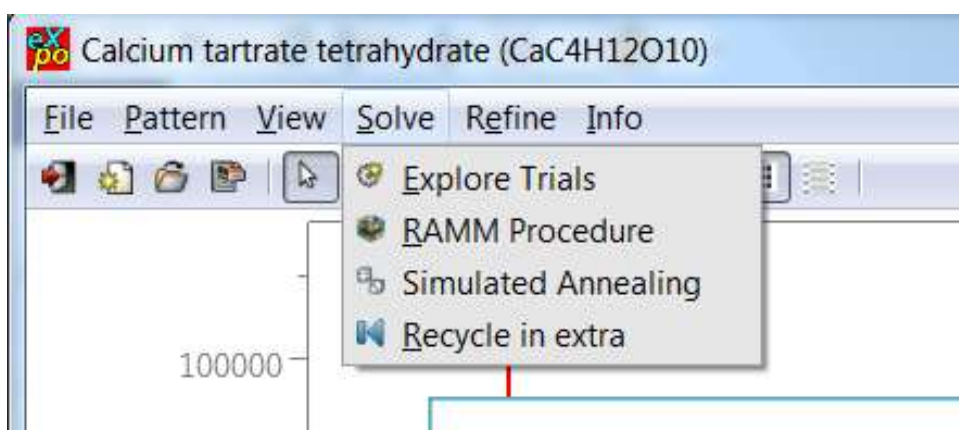
It is so rough and uninterpretable that is not advisable to try to improve it, for example, by cyclic application of RBM (RBM is advisable because the structure is metal-organic). Indeed, by clicking on **Refine > Resolution Bias Modification (RBM) > RBM (Direct Space)** in the upper Menu



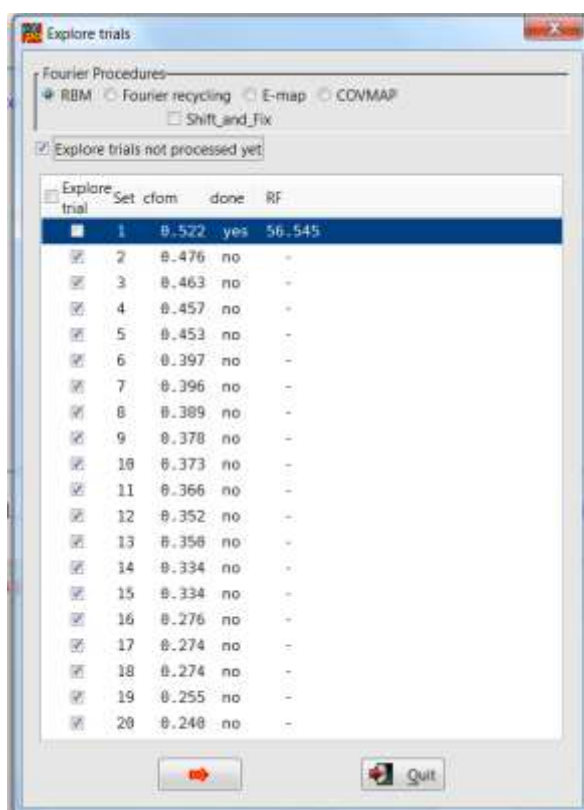
no improvement of the structure model is attained.

We can try to explore the other Direct Methods trials:

Solve > Explore Trials in the upper Menu



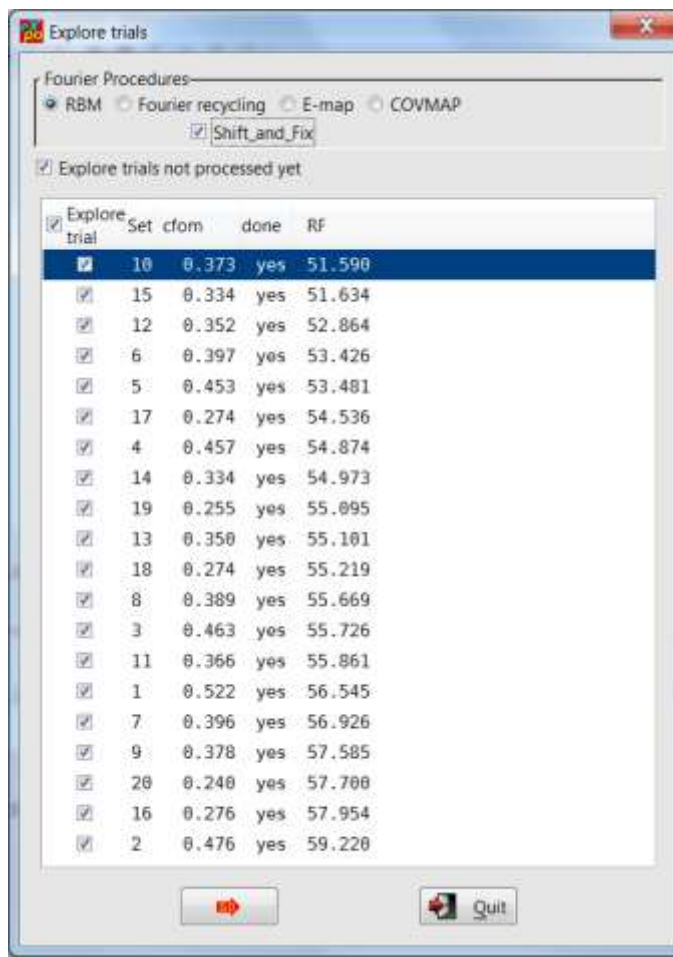
and exploring and ranking all the other 19 trials not processed in the standard run by Direct Methods (only the highest CFOM figure of merit phasing trial is automatically processed)



Click on **GO**.

The model first ranked by RF doesn't correspond to the correct solution.

The structure solution can be attempted by the optimization of **SHIFT_AND_FIX** to be applied to the 20 Direct Methods phasing trials

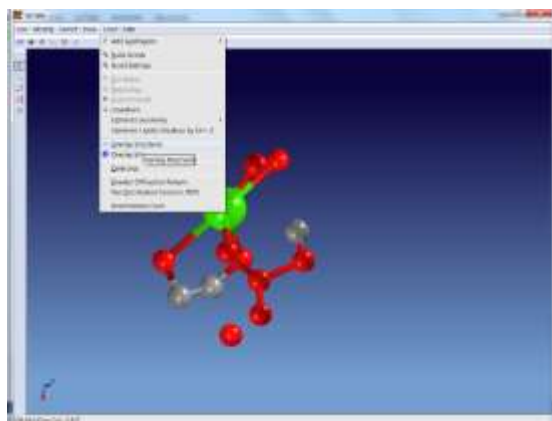


Click on **GO**. The execution of SHIFT_AND_FIX requires a time longer than the standard RBM process.

Now the model first ranked by RF corresponds to the correct solution (the chemical label should be corrected: left-click on the wrongly labelled atom position, right-click, Change Species, select the label).

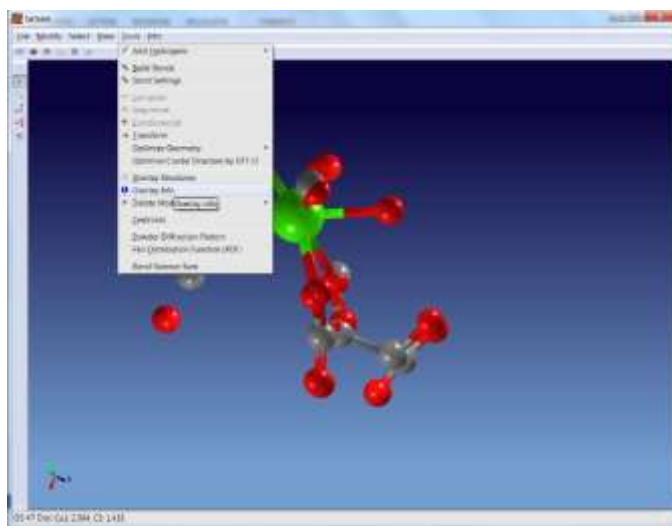
The obtained solution can be compared with the published fractional coordinates contained in the tartrate.fra file. It can be done by the following graphic pathway:

Tools > Overlay structures in the upper Menu and select tartrate.fra and **OK**



Atom	Coordinates	Distance Atom	Atom Distance (Å)

The two models are superimposed and information on comparison can be output:
Tools > Overlay Info in the upper Menu



Atom	Coordinates	Distance Atom	Atom Distance (Å)
C1	0.834 0.226 0.660	0.012 C1	0.811 0.228 0.659
C1	0.662 0.119 1.089	0.158 C2	0.579 0.137 1.062
O1	1.138 0.450 1.706	0.287 O1	1.183 0.463 1.693
O2	0.612 0.288 0.121	0.119 O3	0.625 0.281 0.101
O4	1.297 0.122 1.089	0.117 O4	1.283 0.120 1.069
O4	1.017 0.411 1.712	0.253 O5	1.009 0.405 1.668
O5	0.980 0.153 1.573	0.113 O5	0.972 0.161 1.572
O6	0.642 0.486 0.805	0.171 O7	0.657 0.470 0.823
O7	1.010 0.182 0.611	0.184 O8	1.031 0.171 0.615
O8	0.671 0.123 0.675	0.041 O8	0.674 0.127 0.679
O9	1.150 0.165 1.002	0.041 C4	1.152 0.163 1.007
O10	1.095 0.296 1.185	0.136 C1	1.131 0.283 1.195
C2	0.683 -0.001 0.475	0.278 O8	0.692 -0.012 0.463
C3	1.122 0.256 1.451	0.201 C3	1.128 0.251 1.418
C4	1.108 0.188 1.587	0.118 C2	1.142 0.184 1.620

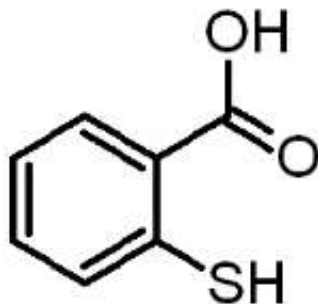
Distance (Å): 0.600 Select model: tartrate.frz

Results

Atom in tartrate: 15
 Atom in tartrate.frz: 25
 Matches found: 15
 Mean Phase Error: 24.436 using Y01 reflections
 <Dist>: 0.175
 RMSD: 0.052

- **RAMM** folder.

It contains: **merca.exp** [the input file for the default run of *EXPO* in case of 2-Mercaptobenzoic acid (C₇H₆O₂S), 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, hydrogen atoms excluded); **merca.pdf** (the structure publication).



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

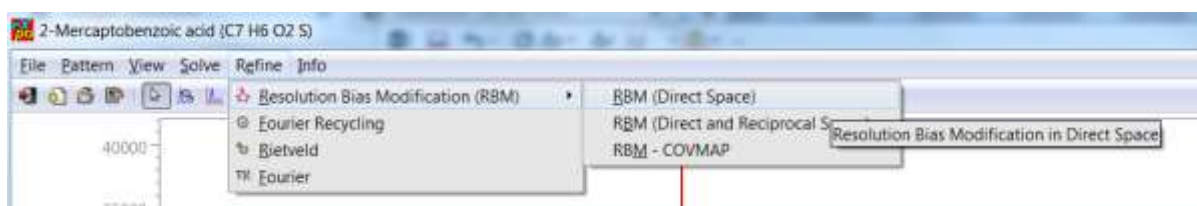
```
%Structure merca
%Job 2-Mercaptobenzoic acid (C7 H6 O2 S)
%Data
Cell 7.885 5.976 14.949 90.0 100.48 90
SpaceGroup p 21/c
Content (C7 H6 O2 S) 4
Pattern merca.pow
Wavelength 1.54056
%continue
```

To run EXPO on merca in default way:

- Click on EXPO icon
- **File** in the upper Menu
- **Load & Go**
- Use 'merca.exp' as Input File and give the Output Filename you like (merca.out is the default output file name)
- **Go**
- Click on **Next** to go on continuously until the end of the run.

The structure model obtained at the end of the Direct Methods procedure, executed on the first set of phases (default choice), is not interpretable.

You can apply RBM cycling (RBM is advisable because the structure is metal-organic) by iterated clicking on **Refine > Resolution Bias Modification (RBM) > RBM (Direct Space)** in the upper Menu



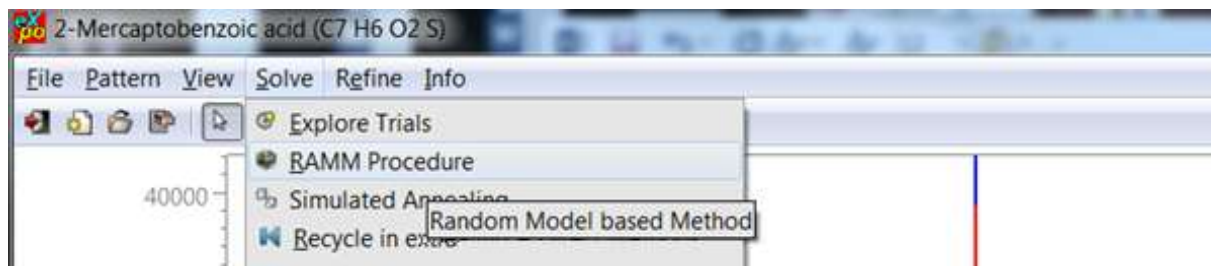
At the end of the RBM application, the model is improved in some way but the complete and correct solution is not attained.

The structure solution can be obtained by one of the two following strategies:

- **the RAMM procedure:**

(The RAMM method which usually requires quite long execution time can be attempted as first non-default choice because the structure is small and RAMM execution time is not expected to be long).

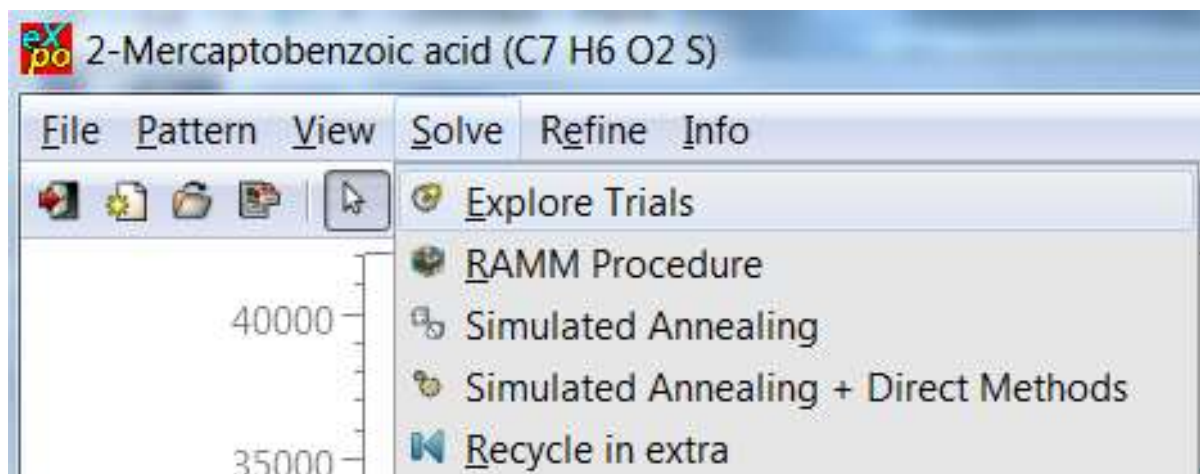
Solve > RAMM Procedure in the upper Menu



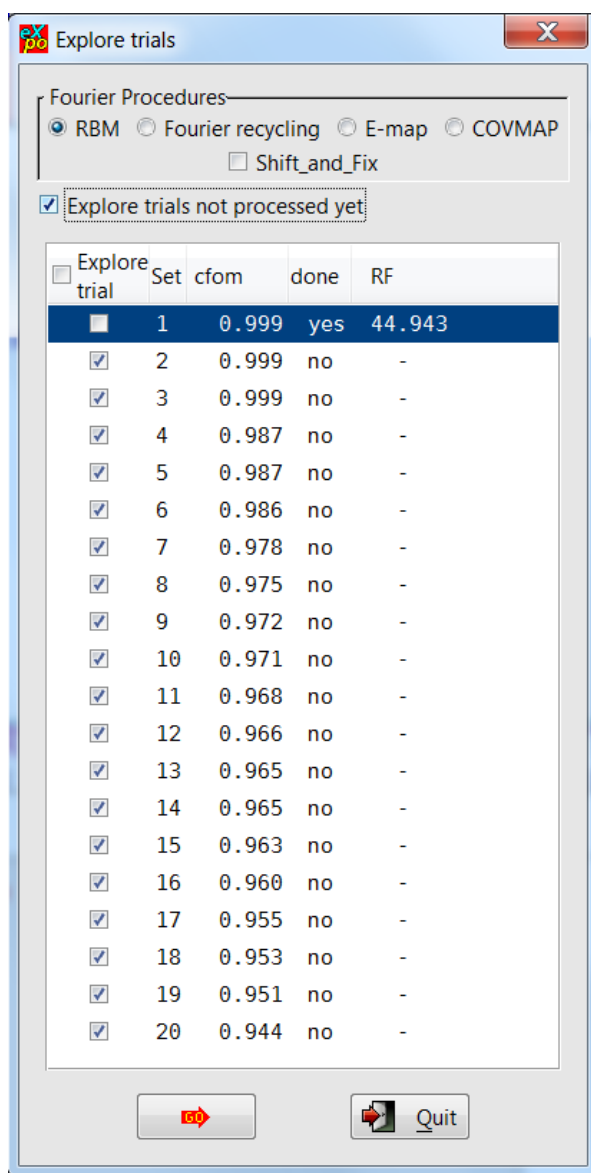
The procedure provides only one model that corresponds to the correct solution.

- **Exploring the other Direct Methods trials:**

Solve > Explore Trials in the upper Menu



and exploring and ranking all the other 19 trials not processed in the standard run by Direct Methods (only the highest CFOM figure of merit phasing trial is automatically processed)

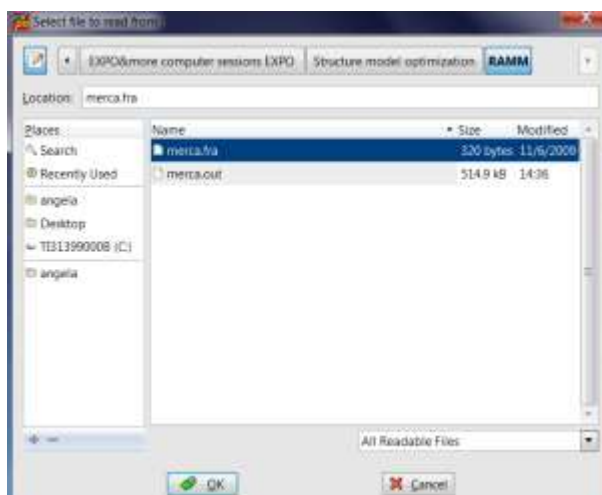
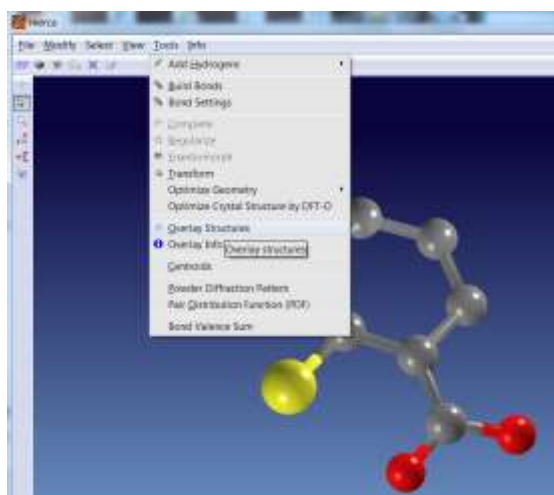


Click on **GO**.

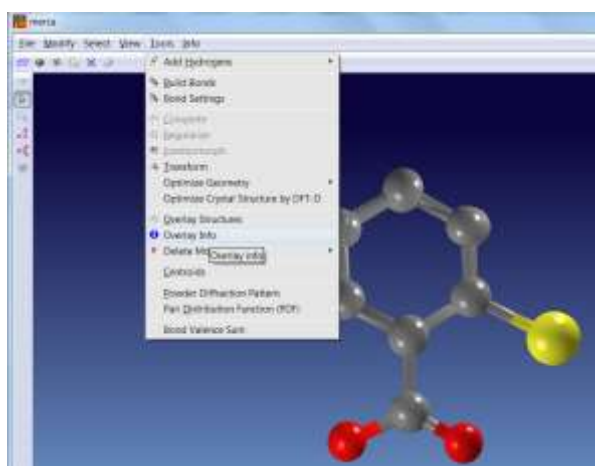
The model first ranked by RF corresponds to the correct solution.

The obtained solution can be compared with the published fractional coordinates contained in the merca.fra file. It can be done by the following graphic pathway:

Tools > Overlay structures in the upper Menu and select merca.fra and **OK**



The two models are superimposed and information on comparison can be output:
Tools > Overlay Info in the upper Menu



The screenshot shows the 'Overlay Info' dialog box. It contains a table of matches between two models, 'merca.fra' and 'merca.out'. The table lists atoms, their coordinates, and the distance between them. Below the table, there are fields for 'Distance limit' and 'Select model'.

Atom	Coordinates	Distance	Atom	Coordinates
S1	1.056 1.116 0.682	0.000	S1	1.017 1.117 0.694
O1	0.777 0.518 0.489	0.000	O2	0.776 0.541 0.502
O2	1.010 0.718 0.529	0.000	O1	1.012 0.713 0.578
C1	0.796 1.018 0.646	0.025	C2	0.797 1.036 0.645
C2	0.505 1.118 0.675	0.136	C4	0.506 1.112 0.696
C3	0.417 0.913 0.620	0.337	C5	0.407 0.922 0.619
C4	0.546 0.779 0.578	0.261	C6	0.562 0.788 0.585
C5	0.738 0.836 0.389	0.052	C1	0.738 0.841 0.397
C6	0.680 1.187 0.678	0.145	C3	0.678 1.186 0.680
C7	0.853 0.708 0.551	0.118	C7	0.855 0.694 0.557

Distance limit: 0.600
 Select model: merca.fra

Results:
 Atoms in merca.fra: 18
 Atoms in merca.out: 18
 Matches found: 18
 Root Mean Square Error: 0.052 using 244 reflections
 RMSD: 0.058
 RMSD: 0.032