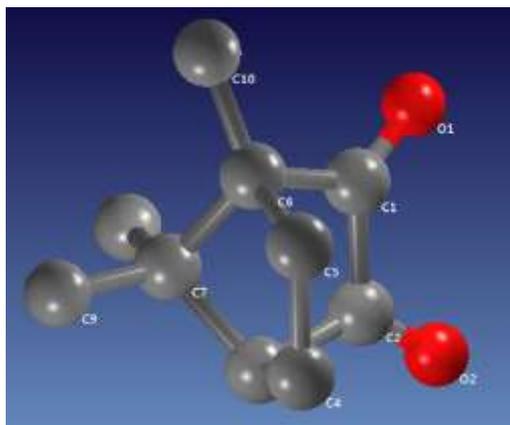


TUTORIAL EXPO: STRUCTURE MODEL OPTIMIZATION

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.

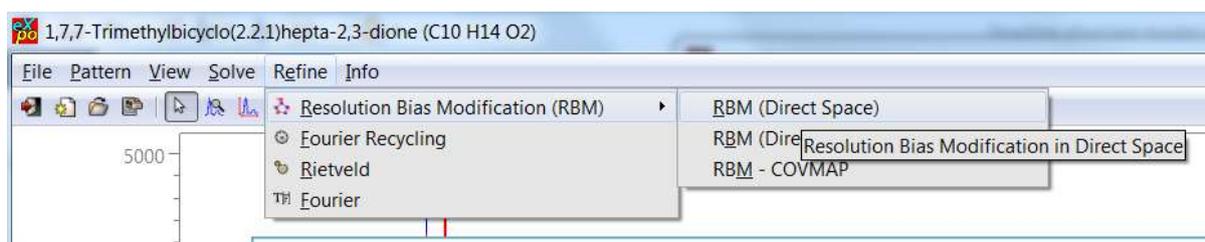
To run EXPO on camphor in default way:

- 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.

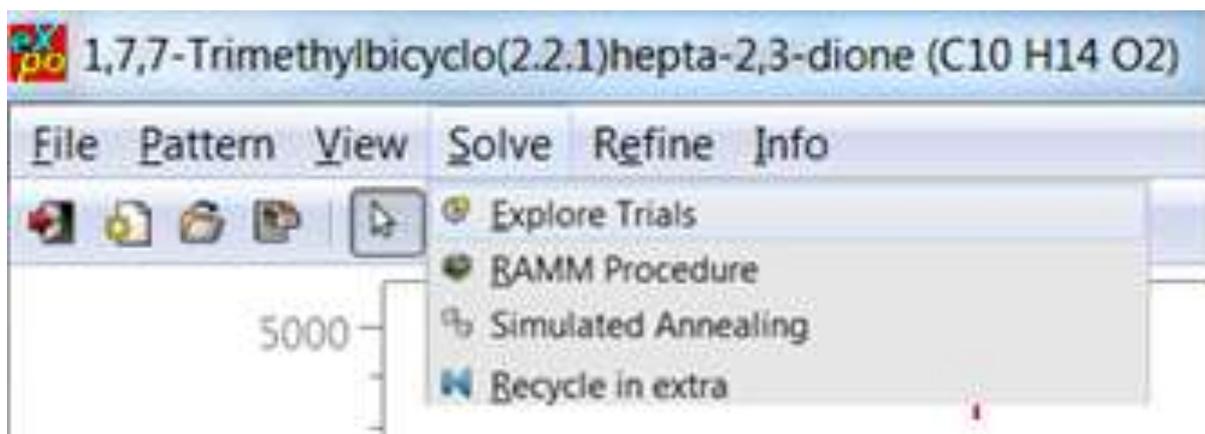
It is so rough and uninterpretable that is not advisable to try to improve it, for example, by cyclic application of RBM (RBM is advisable because the structure is organic). Indeed, by clicking on **Refine > Resolution Bias Modification (RBM) > RBM (Direct Space)** in the upper Menu



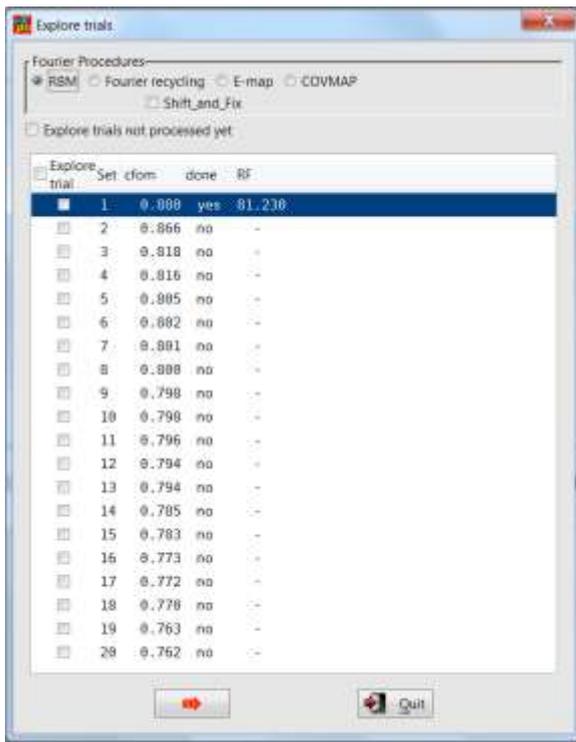
no improvement of the structure model is attained.

We can try to explore the other Direct Methods trials:

Solve > Explore Trials in the upper Menu



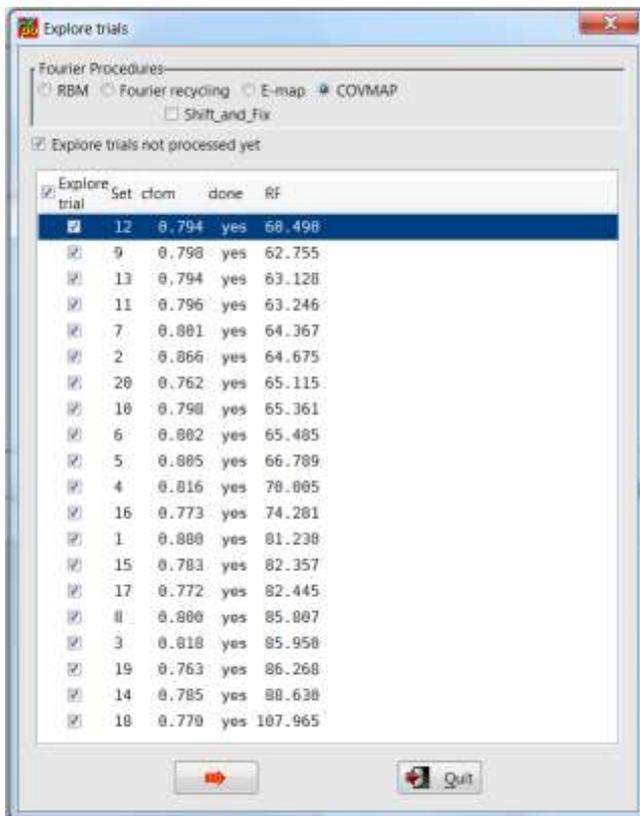
and exploring and ranking all the other 19 trials not processed in the standard run by Direct Methods (only the highest CFOM figure of merit phasing trial is automatically processed)



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.

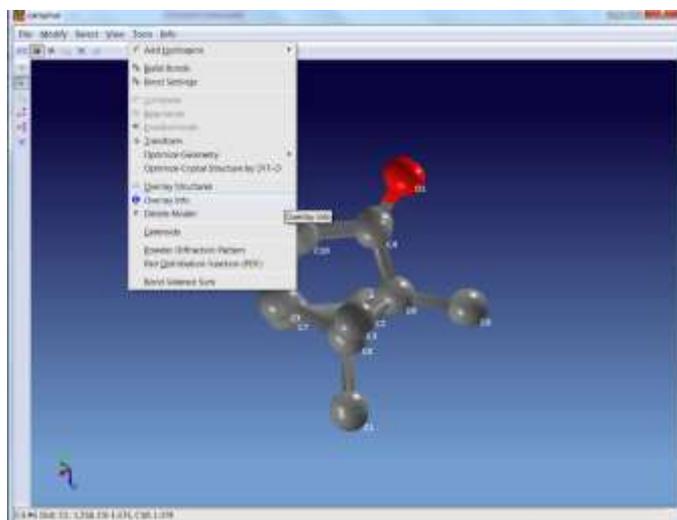
The obtained solution can be compared with the published fractional coordinates contained in the camphor.fra file. It can be done by the following graphic pathway:

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



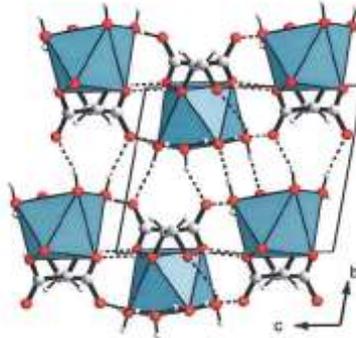
Atom	Coordinates	Distance	Atom	xyz (camphor.fra)
O1	0.442 0.255 -0.138	0.152	O1	0.442 0.242 -0.135
O2	0.634 0.409 -0.129	0.222	O2	0.625 0.412 -0.134
C1	0.663 0.172 0.440	0.090	C8	0.650 0.175 0.440
C2	0.413 0.328 0.358	0.330	C5	0.460 0.327 0.348
C3	0.698 0.129 0.129	0.081	C8	0.684 0.123 0.118
C4	0.506 0.263 0.015	0.116	C3	0.499 0.258 0.006
C5	0.063 0.384 0.215	0.112	C3	0.056 0.338 0.209
C6	0.627 0.205 0.270	0.074	C7	0.628 0.200 0.262
C7	0.385 0.412 0.346	0.137	C4	0.375 0.406 0.343
C8	0.427 0.120 0.250	0.201	C20	0.420 0.112 0.247
C9	0.500 0.226 0.228	0.097	C4	0.500 0.219 0.220
C10	0.601 0.348 0.024	0.206	C2	0.603 0.340 0.000

Distance limit: 0.600 Selected model: camphor.fra

Results
 Atoms in camphor: 12
 Atoms in camphor.fra: 12
 Matches found: 12
 Mean Phase Error: 15.342 using 738 reflections
 -Delta: 0.125
 RMSD: 0.034

- **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:

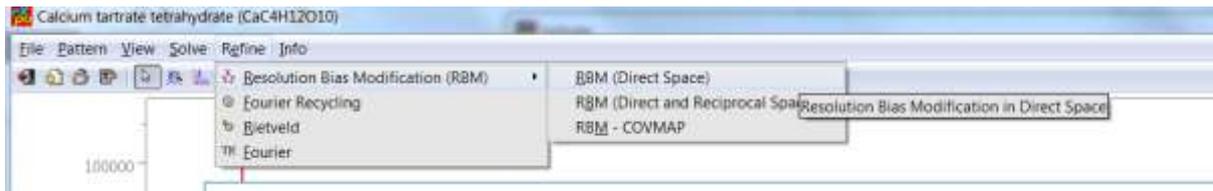
```
%Structure tartrate
%Job Calcium tartrate tetrahydrate (CaC4H12O10)
%Data
Cell 8.222 10.473 6.249 105.97 107.51 94.94
SpaceGroup p -1
Content (CaC4H12O10) 2
Pattern tartrate.pow
Wavelength 1.5418
Synchrotron
%continue
```

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.

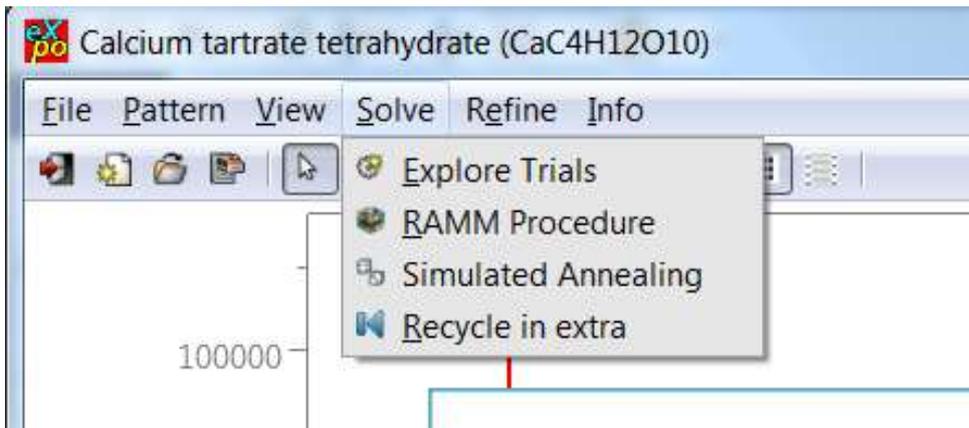
It is so rough and uninterpretable that is not advisable to try to improve it, for example, by cyclic application of RBM (RBM is advisable because the structure is metal-organic). Indeed, by clicking on **Refine > Resolution Bias Modification (RBM) > RBM (Direct Space)** in the upper Menu



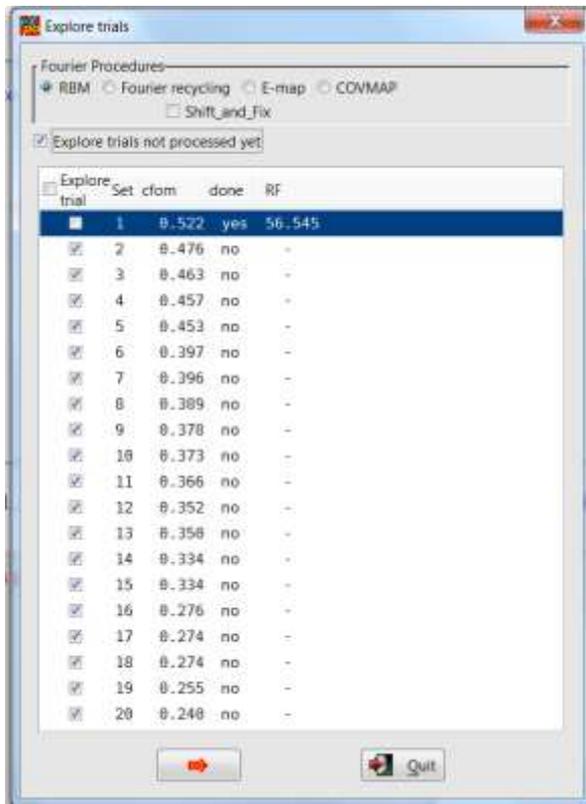
no improvement of the structure model is attained.

We can try to explore the other Direct Methods trials:

Solve > Explore Trials in the upper Menu



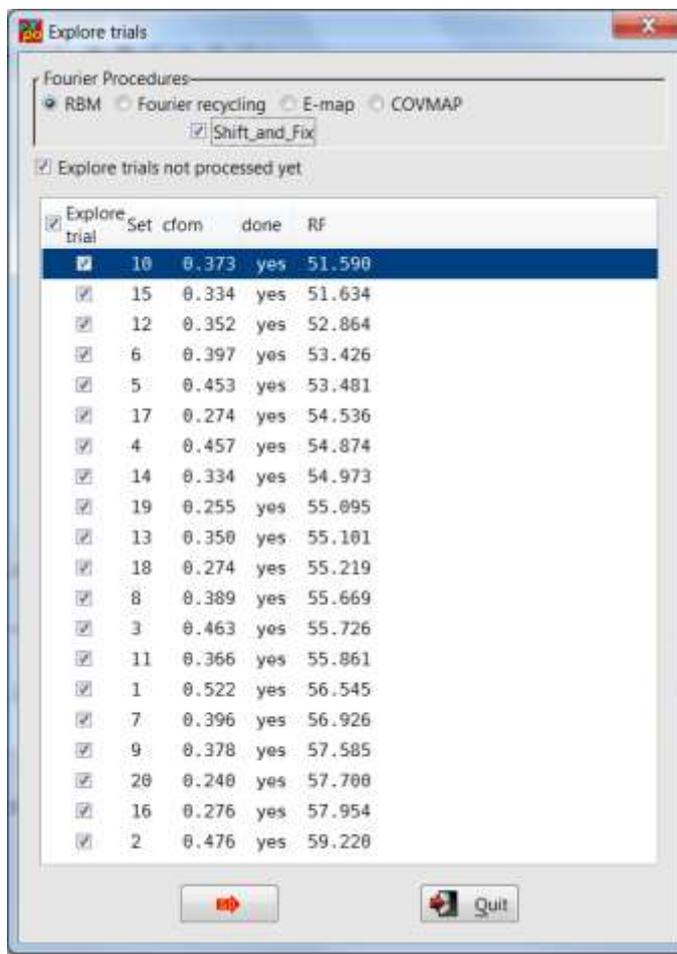
and exploring and ranking all the other 19 trials not processed in the standard run by Direct Methods (only the highest CFOM figure of merit phasing trial is automatically processed)



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



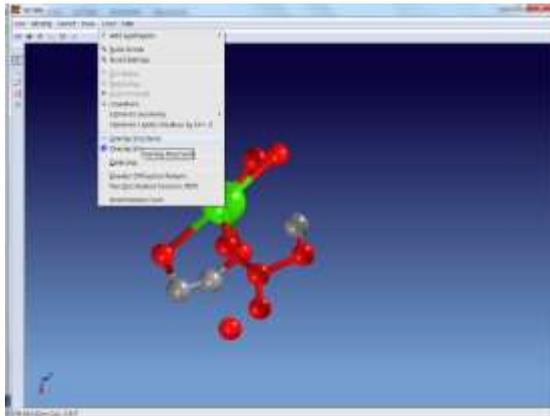
Explore trial	Set	cfom	done	RF
<input checked="" type="checkbox"/>	10	0.373	yes	51.590
<input checked="" type="checkbox"/>	15	0.334	yes	51.634
<input checked="" type="checkbox"/>	12	0.352	yes	52.864
<input checked="" type="checkbox"/>	6	0.397	yes	53.426
<input checked="" type="checkbox"/>	5	0.453	yes	53.481
<input checked="" type="checkbox"/>	17	0.274	yes	54.536
<input checked="" type="checkbox"/>	4	0.457	yes	54.874
<input checked="" type="checkbox"/>	14	0.334	yes	54.973
<input checked="" type="checkbox"/>	19	0.255	yes	55.095
<input checked="" type="checkbox"/>	13	0.350	yes	55.101
<input checked="" type="checkbox"/>	18	0.274	yes	55.219
<input checked="" type="checkbox"/>	8	0.389	yes	55.669
<input checked="" type="checkbox"/>	3	0.463	yes	55.726
<input checked="" type="checkbox"/>	11	0.366	yes	55.861
<input checked="" type="checkbox"/>	1	0.522	yes	56.545
<input checked="" type="checkbox"/>	7	0.396	yes	56.926
<input checked="" type="checkbox"/>	9	0.378	yes	57.585
<input checked="" type="checkbox"/>	20	0.240	yes	57.700
<input checked="" type="checkbox"/>	16	0.276	yes	57.954
<input checked="" type="checkbox"/>	2	0.476	yes	59.220

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).

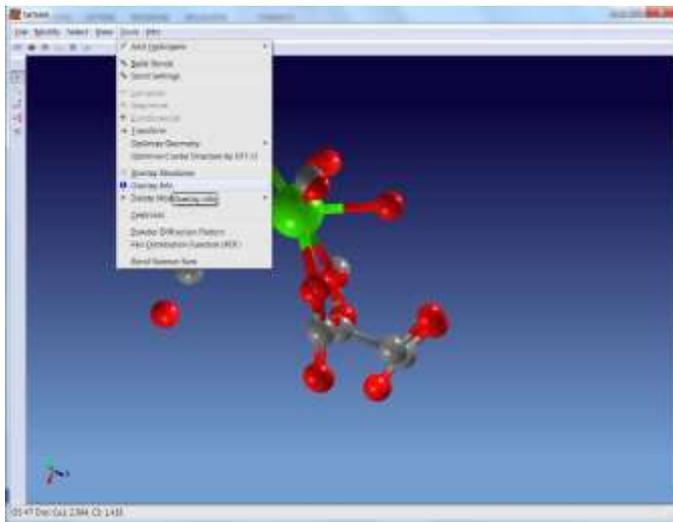
The obtained solution can be compared with the published fractional coordinates contained in the `tartrate.fra` file. It can be done by the following graphic pathway:

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



Atom	Coordinates	Distance Atom	Atom in tartrate
Ca1	0.834 0.726 0.660	0.032	Ca1
Cl	0.642 0.338 1.688	0.158	O2
O1	1.338 0.458 1.706	0.267	O1
O2	0.642 0.288 0.321	0.118	O2
O3	0.625 0.281 0.201		
O4	1.297 0.122 1.888	0.117	O4
O5	1.017 0.421 1.712	0.253	O5
O6	0.980 0.253 1.573	0.113	O3
O7	0.642 0.486 0.805	0.171	O7
O8	1.020 0.182 0.631	0.184	O6
O9	0.671 0.323 0.675	0.041	O8
O10	1.150 0.165 1.002	0.041	O4
O11	1.095 0.296 1.185	0.336	Cl
C1	0.683 -0.301 0.475	0.239	O8
C2	1.322 0.256 1.453	0.201	C3
C4	1.308 0.388 1.567	0.318	C2

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



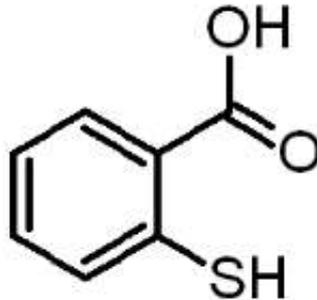
Atom	Coordinates	Distance Atom	Atom in tartrate
Ca1	0.834 0.726 0.660	0.032	Ca1
Cl	0.642 0.338 1.688	0.158	O2
O1	1.338 0.458 1.706	0.267	O1
O2	0.642 0.288 0.321	0.118	O2
O3	0.625 0.281 0.201		
O4	1.297 0.122 1.888	0.117	O4
O5	1.017 0.421 1.712	0.253	O5
O6	0.980 0.253 1.573	0.113	O3
O7	0.642 0.486 0.805	0.171	O7
O8	1.020 0.182 0.631	0.184	O6
O9	0.671 0.323 0.675	0.041	O8
O10	1.150 0.165 1.002	0.041	O4
O11	1.095 0.296 1.185	0.336	Cl
C1	0.683 -0.301 0.475	0.239	O8
C2	1.322 0.256 1.453	0.201	C3
C4	1.308 0.388 1.567	0.318	C2

Distance limit: 0.600 Select model: tartrate

Results
 Atoms in tartrate: 15
 Atoms in tartrate fra: 25
 Matches found: 15
 Mean Phase Error: 24.438 using 702 reflections
 $\langle \text{dIct} \rangle$: 8.175
 RMSD: 0.952

- **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:

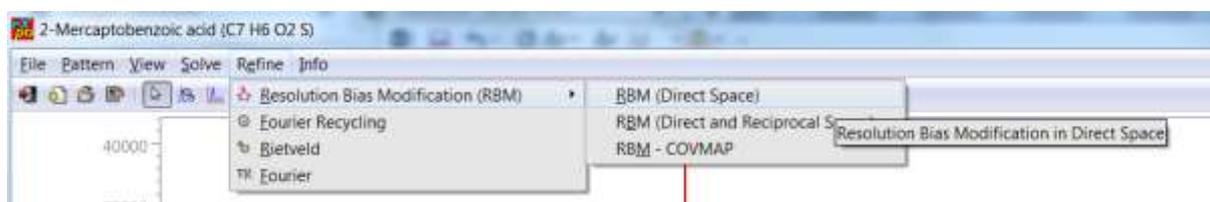
```
%Structure merca
%Job 2-Mercaptobenzoic acid (C7 H6 O2 S)
%Data
Cell 7.885 5.976 14.949 90.0 100.48 90
SpaceGroup p 21/c
Content (C7 H6 O2 S) 4
Pattern merca.pow
Wavelength 1.54056
%continue
```

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



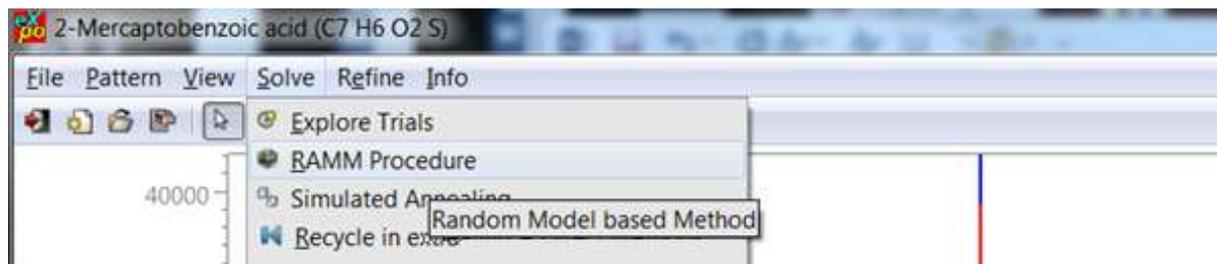
At the end of the RBM application, the model is improved in some way but the complete and correct solution is not attained.

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).

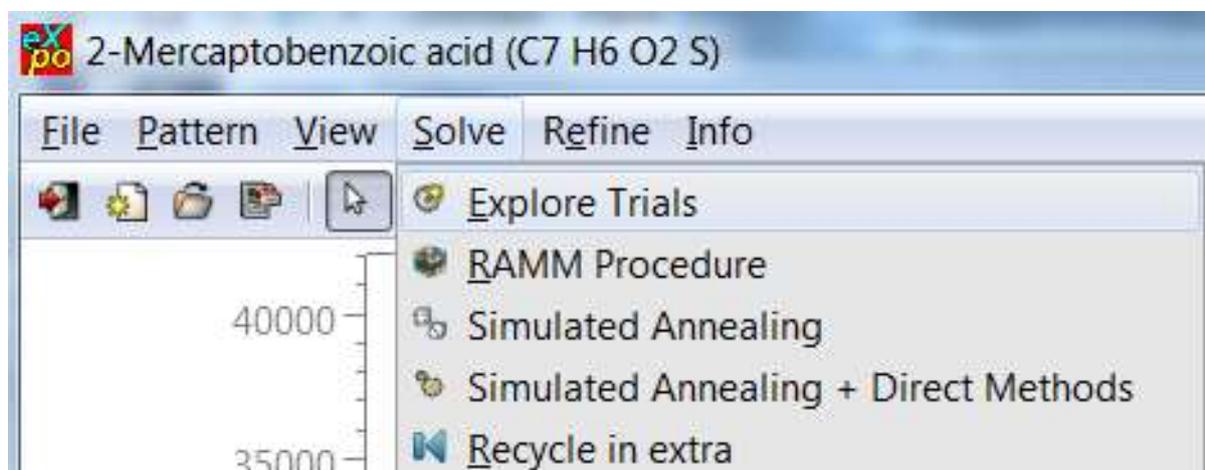
Solve > RAMM Procedure in the upper Menu



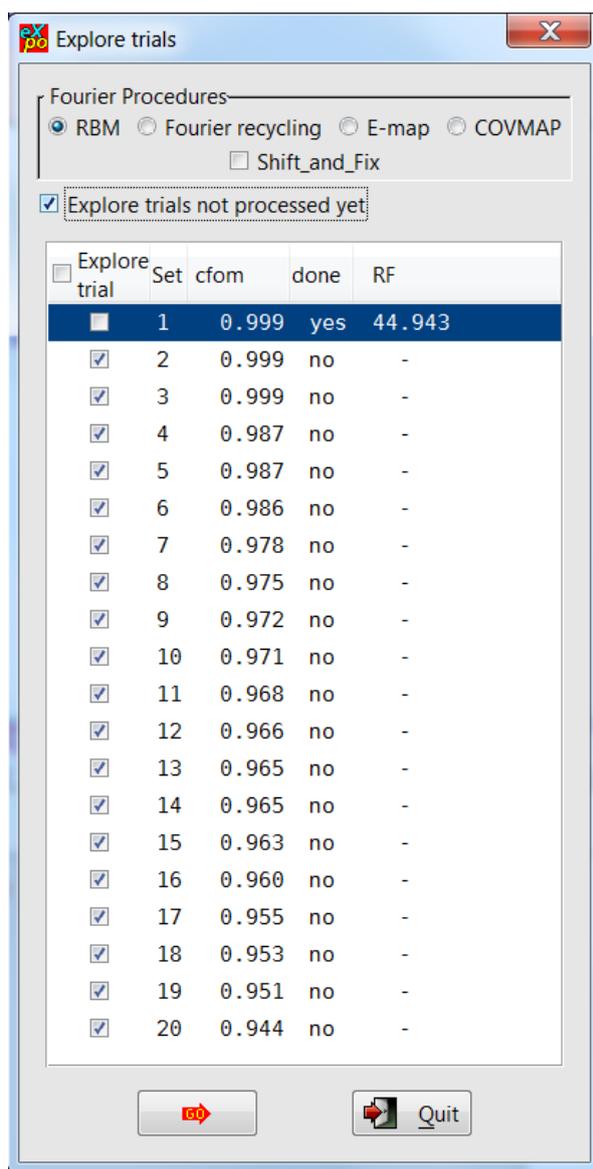
The procedure provides only one model that corresponds to the correct solution.

- **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)

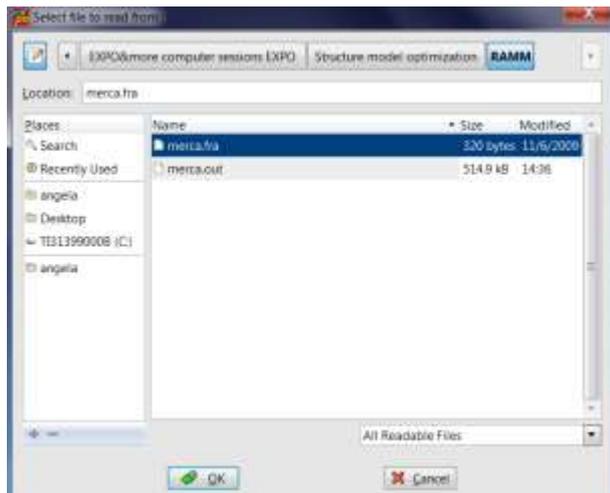
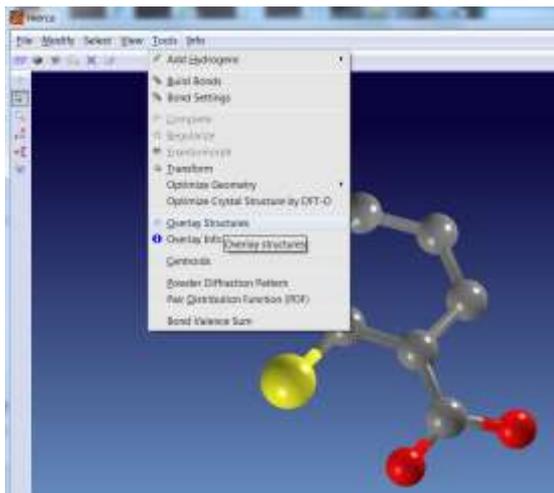


Click on **GO**.

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

