

## Exercise 1

# The RMCProfile working environment

The aim of this exercise is to setup the basic working environment for RMCProfile and to check that the main three programs are working correctly. RMCProfile does not have a GUI at this time so everything is driven from the command line in a terminal window. Finder can be used to set up and move or copy files, and a text editor such as `textedit` will be used to edit files where necessary. The commands will then be run from the terminal window called RMCProfile version 6.

1. First, go to the **/Applications/RMCProfile** folder—you should have created this folder when you installed RMCProfile. Then double click the file **rmcprofile\_setup.command**. This will launch a terminal window with the correct environment variables set. From here on in this will be referred to as the **rmcprofile** window.
2. Create a working directory on your machine to keep all the RMCProfile files together. This can be done with Finder or from the command line, something along the lines of **/myfolder** is fine.
3. Copy the folder **ex\_1** from the **rmcprofile/tutorial** folder to your new working directory.
4. In the **rmcprofile** window change the directory to this new folder; *i.e.* type

```
cd ~/myfolder/ex_1
```

5. Then type `'ls'` in the **rmcprofile** window. This will list all the files in the directory and it should look similar to the list below:

```

Directory of C:\temp\pdf_test\ex_1

31/03/2008  13:44    <DIR>          .
31/03/2008  13:44    <DIR>          ..
22/10/2007  22:56             40,495 sf6190k_all_rmc.gr
22/10/2007  22:56             138 rmcsf6_190k.back
22/10/2007  22:56            19,361 rmcsf6_190k.bragg
31/03/2008  13:49            14,089 rmcsf6_190k.cfg
31/03/2008  13:54             1,581 rmcsf6_190k.dat
22/10/2007  22:56             6,809 rmcsf6_190k.fs
22/10/2007  22:56              38 rmcsf6_190k.hkl
22/10/2007  22:56             471 rmcsf6_190k.inst
22/10/2007  22:56              47 rmcsf6_190k.poly
22/10/2007  22:56            2,758 rmcsf6_190k.sf
31/03/2008  13:59            19,704 rmcsf6_190keye.cfg
22/10/2007  22:56            32,104 sf6190k_all_conv8p82.dat
31/03/2008  13:49            30,907 rmcsf6_190k.his
31/03/2008  13:49            33,283 rmcsf6_190k.amp
31/03/2008  13:49           145,767 rmcsf6_190k.out
31/03/2008  13:49           87,320 rmcsf6_190k.braggout
                16 File(s)          434,872 bytes
                2 Dir(s)    13,156,237,312 bytes free

```

6. For now do not worry about what these file are—they are just a set of test files to ensure the programs work correctly on your machine. We will look at them in detail in the next exercise. For now, select the **rmcprofile** window, type

```
rmcprofile rmcsf6_190k
```

and hit return.

RMCPProfile should now start running. It may take a few minutes to initialise, but then it will display various messages explaining what it has been asked to do. It should finish with a set of parameters that indicate how well the data is being fitted. The final part of the output should appear similar to the text below:

```

Peaks calculated
scale 0.210434976038272
Bragg calculations initialised

----->
Time used 0.0 ,time limit 0.0 ,Last saved at 0.3

Bragg Chi**2/npts = 0.4830

10074 moves accepted 89690 generated and 85650 tested
Chi**2/dof 0.8880
Expt 1: Renorm 1.0000 Constant = 0.0000 Chi**2/nq = 2.320
Expt 2: Renorm 1.0000 Constant = 0.0000 Chi**2/nq =0.7584
Restraint energy = 3.465

Total time in loop so far : 4.6876073E-02s
Time per generated move is: 15.6253576278687 ms

Saving configuration and results to disk
*** DO NOT INTERRUPT UNTIL SAVING IS COMPLETED ***
Saving has been completed

Total time taken by program: 10.672

```

7. The next program to check is called **rmcplot**. This program displays the fits to data obtained with RMCProfile and also the partial structure factors calculated from the RMC configuration. Working again within the **rmcprofile** window, type `'rmcplot rmcsf6_190k'`. A new window should appear, called **PGPLOT Window1**—it may be hidden behind another window. This window is where the graphs will be displayed. In the **rmcprofile** window, enter `'5'` and hit return. The fit to the bragg profile should now be displayed. The fit will not look very good since the model being used to fit it is too small, but for now all that is important is that the fit is displayed at all.

You can use the left mouse click to zoom in on a region. Right-click will then restore the display and pressing `'q'` with the graphics window active will return control to the **rmcprofile** window, from where you can make a new selection or indeed quit. This is a very basic plotting program designed to provide a quick look at the fits. For more comprehensive functionality feel free to use the csv file RMCProfile produces with your favourite plotting program, such as excel, sigmaplot, origin, *etc.*

8. The final program to check at this stage is **atomeye**. This can be used to display a ball and stick model of the RMCProfile configuration. It is a quick way to check that the atoms are not in unphysical places and also whether the configuration is physically sensible or not. This program was written by Ju Li at Ohio state university, USA and is not developed by the RMCProfile team; it is just a very useful program, so we use it.

To check that the program is working, type

```
atomeye rmcsf6_190keye.cfg
```

in the **rmcprofile** window. You will have to quit **rmcplot** if you have not already done so—by entering `'0'`. If **atomeye** is working two new windows should appear: both are called **rmcsf6\_190k.cfg**. Don't worry if error messages are displayed in the **rmcprofile** window, so long as the new windows do in fact appear and they remain on the screen. Note that one or other or perhaps both of the new windows may initially be minimised. If this is the case, click on the new item in the task bar at the bottom of this screen.

The grey text window displays useful information, and is used to enter parameters. The new graphics window displays the image and is the main control window. With this new graphics window active the mouse can be used to move the view. Press `'F1'` to display a list of options in the grey window. To quit the programs press `'q'` with the graphics window active.

9. If everything worked correctly you are now ready to move on to exercise 2.

## Exercise 2

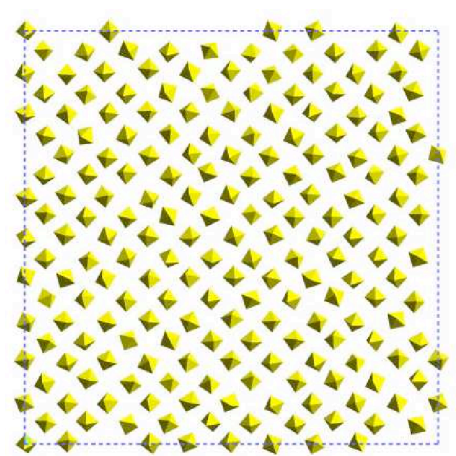
# Setting up an RMCProfile refinement

This exercise will guide you through the steps required to set up an RMCProfile refinement from scratch. It is assumed that you already have a GSAS refinement of your data and have produced corrected total scattering data. The **ex\_2** folder contains all the data you need, so please copy this to your working area. The **gsas** folder has the GSAS refinement and all data needed to provide bragg profile information and a starting configuration for RMCProfile.

The **data** folder contains time of flight neutron structure factor  $F(Q)$  data and the radial distribution  $g(r)$ . Together with the **gsas** folder, these two folders provide all the information required to run RMCProfile and produce a configuration of the system.

The **rmc** folder contains a completed run together with all the relevant output files; it is included to illustrate what the results will (hopefully) look like.

The example files supplied are for the molecular crystal  $\text{SF}_6$  and the following paragraph gives you some background information to help understand the aim of the RMCProfile refinement.



$\text{SF}_6$  is a disordered molecular crystal made up of  $\text{SF}_6$  octahedra arranged over the crystal lattice. Molecular crystals are often the most disordered crystalline systems, giving rise to large amounts of diffuse scattering (this is easily seen in a plot of the example data). The example data and files supplied are taken from neutron scattering experiments carried out at 190 K using the GEM diffractometer at ISIS. At this temperature the crystal structure has body centred cubic symmetry, with one molecule on the corner of the unit cell and one in the centre. However this average structure places the fluorine atoms of neighbouring molecules too close to one another, and so on a local scale the system tries to maximise the distance between these contacts by rotating the octahedra in different directions. The system is said to be frustrated, with fluorine–fluorine repulsion driving the motion of the molecules.

With the data supplied, you should be able to start with the “ideal” ordered structure and use RMCProfile to model these local deviations from the average bcc arrangement.

The preparation of the files for an RMCProfile refinement can be split into three stages: first, the information from a GSAS refinement is obtained, next the total scattering data is prepared,

and finally the control files for RMCProfile are produced. The next three sections take you through each of these stages in turn.

1. As with most RMC modelling, the process actually starts with a GSAS refinement. A GSAS refinement of SF<sub>6</sub> at 190 K is contained in the **gsas** folder and can be used to supply RMCProfile with all the information required to model the Bragg profile. It also contains the information need to produce a starting configuration of the bcc cell written out in *P1* symmetry. The program **data2config** should produce all the files you need. If however it doesn't work for your structure, GSAS supplies all the information required to build a starting configuration and it is a good exercise to try with your own system; at least it ensures that you start with a good understanding for the long-range average structure.

We will now extract the required information from the GSAS files using the **data2config** program.

In the **rmcprofile** window change the directory to the **gsas** folder; *i.e.*

```
cd ~\myfolder\ex_2\gsas
```

or just

```
cd gsas
```

if you are already in the **ex\_2** directory.

Now type

```
'data2config
```

```
-noannotate -rmc6f -order [S F] -supercell[3 3 3] SF6_190K.TBL '
```

and hit enter. The program should write out '1' and end. The command tells data2config to extract the information from the file SF6\_190K.TBL<sup>1</sup> with the following options: '-noannotate' tell the program to not request any metadata; '-rmc6f' will write out the starting configuration in the new rmc6f formation; '-order [S F]' means the file will have the atoms ordered with the sulphur atoms first and then fluorine atoms; '-supercell [3 3 3]' means generate a supercell 3×3×3 (in future if you wish to produce a larger supercell just change the numbers to the required size).

At this point you should have four file types for RMCProfile located in the **gsas** folder: **.bragg**, **.back**, **.inst** and **.rmc6f**. These are the bragg profile data, gsas background parameters, instrument peak-shape profile parameters and the starting configuration respectively. We will copy these to the correct location a bit later on.

In the **rmcprofile** window, move to the top level of the **ex\_2** folder by entering 'cd ..'

You have now produced all the files required from the GSAS refinement.

2. The next stage is to prepare the total scattering data for RMCProfile. How to collect and produce total scattering data is beyond the scope of this simple tutorial and so will not be covered here.

The data folder contains an  $F(Q)$  file (*sf6\_190k\_fq.dat*) produced from three of the detector banks of the GEM diffractometer at ISIS. It is quite common to combine several

---

<sup>1</sup>The **.TBL** file is created with GSAS and to find out more about producing the GSAS files required please see the appendix

banks of data into one data set, although RMCPProfile can be used to fit many  $F(Q)$  data sets simultaneously.

Take a look at the  $F(Q)$  file, using excel or any other plotting program available on your system. A basic alternative, supplied with RMCPProfile, is called **datamplot** that plots multiple files in rmc data format. To use this program to plot the  $F(Q)$  data, select the **rmcprofile** window and move into the **data** folder by entering 'cd data'. Enter 'datamplot sf6\_190k\_fq.dat sf6\_190k\_fq.dat' (the filename is entered twice since the program expects two or more files). This should then bring up a plot of the data file. To exit the plot, press 'q' with the plot window selected, then select option '5' in the **rmcprofile** window.

The data contained in the  $F(Q)$  file are scaled from  $-0.27593$  to  $0$ ; the  $0.27593$  factor is the sum of the neutron partial scattering factors and it needs to have this scale in order to be suitable for RMCPProfile (although there are now a number of scale options as detailed in the RMCPProfile manual). The data also need to be convolved with the RMCPProfile configuration box function to enable a fair comparison, for a  $3 \times 3 \times 3$  supercell, the appropriate truncation distance is  $8.82 \text{ \AA}$ . However, this version of RMCPProfile will do this convolution for you so nothing needs to be done at this stage.

The next file we need from this folder is the  $G(r)$  file (`sf6_190k_gr.dat`). This file contains the fourier transform of the  $F(Q)$  data. Once again the preparation of the  $G(r)$  data themselves will not be described here; the data are normally obtained during processing of the total scattering data. Take a look at the file using either a linux plotting program or **datamplot**. Again the  $G(r)$  file is scaled from  $-0.27593$  to  $0$  and not  $0$  to  $1$  as is perhaps more standard. RMCPProfile comes with a basic program for rescaling data called **data\_rescale** if you have data that needs rescaling when trying to use your own data, although as mentioned above most formats can now be read directly by RMCPProfile.

Next we need to copy the  $F(Q)$  and  $G(r)$  files into the top level of the **ex\_2** folder. In the **rmcprofile** window type

```
cp * ..
```

this copies all the files in the current directory to the one above.

3. The next file to prepare is the main **.dat** control file for RMCPProfile. This can be copied from the **rmc** folder to the top level of the **ex\_2** folder and is called `rmcsf6_190k.dat`. The 'rmcsf6\_190k' is called the "stem name" and this needs to be the same for all the files RMCPProfile uses except the experimental data files.

Assuming you are still in the **data** directory, In the **rmcprofile** window type:

```
cd ..  
cp rmc/rmcsf6_190k.dat .
```

You should open the `rmcsf6_190k.dat` control file with a text editor (*e.g.*, `textedit`) and check that the names of the data files in the **.dat** file match the names of the files from the **data** directory. A description of the **.dat** file format and contents is given in the RMCPProfile manual.

Next we need to copy the starting configuration and bragg files from the **gsas** folder as mentioned above these need to have the same "stem name" as the main **.dat** file.

In the **rmcprofile** window type

```
ls
```

This should produce a list of the files in the ex\_2 folder and should look as follows:

```
RMCPprofile: ex_2 >ls
data gsas rmc rmcsf6_190k.dat sf6_190k_fq.dat sf6_190k_gr.dat
RMCPprofile: ex_2 >
```

to copy the required file from the gsas folder enter the following commands:

```
cp gsas/SF6_190K.rmc6f rmcsf6_190k.rmc6f
cp gsas/SF6_190K.inst rmcsf6_190k.inst
cp gsas/SF6_190K_3.bragg rmcsf6_190k.bragg
cp gsas/SF6_190K_3.back rmcsf6_190k.back
```

Here we have copied the '\_3.back' and '\_3.bragg' files, as at present you can only fit one bragg dataset and so here we have chosen to fit the 3rd histogram in the gsas refinement.

4. Finally, since the SF<sub>6</sub> molecules form an octahedron, it make sense to use a constraint to hold these units together. There are a number of restraints available for this such as polyhedral restraints, potential restraints and distance window constraints, here we will use the latter. The distance window constraint as the name implies defines distance windows in which pairs of atoms are allowed to move. Initially RMCPprofile searches for atom pairs within the ranges defined and produces a neighbours list. Then as the refinement process proceeds these neighbours are constrained to remain within the set distance windows. To turn on this constraint a **.dw** file needs to be present. This file contains two rows of numbers, the first row defines the shortest distance pairs of atoms can have and the second line defines the longest distances. A suitable file for SF<sub>6</sub> would be:

```
0 1.37 2.00
0 1.74 2.42
```

Here the '0's are the shortest and longest distances for 'S-S' pairs and are a special case where we do not want to constrain the sulphur atoms, so in this way no neighbours will be found and nothing is constrained. The '1.37' and '1.74' are the shortest and longest distances for 'S-F' atom pairs and the '2.0' and 2.42 are for the 'F-F' atom pairs. The number are derived by looking at the peaks in the  $G(r)$  file and choosing values encompassing the relevant peak without cutting into it. So if you plot the sf6\_190k\_gr.dat using **datamplot** by typing 'datamplot sf6\_190k\_gr.dat sf6\_190k\_gr.dat' you should see that the first two peaks which are for 'S-F' and 'F-F' fit within the windows defined above (you can zoom in on these peaks using the left mouse button and unzoom by clicking the right mouse button).

Please now create a `rmcsf6_190k.dw` file containing the numbers above either by using a text editor or copying the file from the **rmc** directory.

To check you have copied and created all the files you need type `ls` and the output should be similar to:

```
RMCPprofile: ex_2 >ls
data rmcsf6_190k.back rmcsf6_190k.inst
gsas rmcsf6_190k.bragg rmcsf6_190k.rmc6f
rmc rmcsf6_190k.dat sf6_190k_fq.dat
rmcsf6_190k.dw sf6_190k_gr.dat
```

5. You should now have all the files you need to run RMCPprofile on this system. Both the data and the size of configuration (a  $3 \times 3 \times 3$  supercell) have been chosen to run quickly. A reasonable fit should be achievable within about 10 minutes. However, the configuration is too small to ever give a useful model or fit. Doing so would normally require a large configuration and 10–20 hours of run time. The next exercise will describe how to run RMCPprofile on these files.



## Exercise 3

# Basic RMC refinement

This exercise will guide you through the steps required to run an RMCProfile refinement and perform some basic analysis of the results. There tends to be two stages to the analysis. The first is to check that the minimisation is proceeding usefully, and that the model being produced is sensible—this is our main aim here. The second type of analysis tends to be sample specific, starting once a final configuration has been reached and aiming to decide what the model tells us about the system being studied. Typically this may involve writing custom programs or scripts, and as such the process is very hard to generalise.

1. To start, copy the **ex\_3** directory into your work area, and make a directory called **start** in it (*i.e.* `'mkdir start'`). Then copy the files you created from the top level of you **ex\_2** folder into the **start** folder (*i.e.* `'cp ../ex_2/* start'`). Next, create a copy of the **start** folder, calling it **initial\_fit** (*i.e.* `'cp -r start initial_fit'`).

In the **rmcprofile** window change the directory to **ex\_3\initial\_fit**; *i.e.*, enter

```
cd ~/myfolder/ex_3/initial_fit
```

and then `'ls'`. The listing that appears should be similar to the one below. Check that all the “stem names” have been kept updated and are the same:

```
RMCPProfile: initial_fit >ls
rmcsf6_190k.back rmcsf6_190k.dw sf6_190k_fq.dat
rmcsf6_190k.bragg rmcsf6_190k.inst sf6_190k_gr.dat
rmcsf6_190k.dat rmcsf6_190k.rmc6f
```

Open the RMCProfile **.dat** file (probably called **rmcsf6\_190k.dat**) with a text editor and change the line

```
TIME_LIMIT ::      10.00 MINUTES
SAVE_PERIOD ::     0.20 MINUTES
```

to

```
TIME_LIMIT ::      0.00 MINUTES
SAVE_PERIOD ::     0.00 MINUTES
```

and save the file. This change will tell RMCPProfile to initialise and generate an initial fit and then stop.

2. Next type 'rmcprofile [stem name]' to run RMCPProfile; e.g., 'rmcprofile rmcsf6\_190k'. RMCPProfile should now run for a little while, and if all seems OK then move on to the next step. However, if you see an error message as below:

```
Using files: rmcsf6_190k
```

```
-----  
  
XML/CML output requested  
Info> Read history file if present  
Reading configuration file rmcsf6_190k.his6f  
Info> ntypes,ni: 2  
          54          324  
Info> Elements in configuration:  
S  
F  
Info> Number of atom types = 2  
Info> Number of partials = 3  
Material: SF6  
Phase: BCC  
Sample temperature: 190K  
Investigator: Matt  
Info> Numbers of data sets = 1 1 0 0  
Status> Calculating partial weighting coefficients.  
Info> Assuming no renormalisation of neutron PDF data required  
At line 75 of file ../main/rmcprofile_procedures.f90  
Fortran runtime error: No such file or directory
```

then the name of the  $G(r)$  data file in the main **.dat** file is not correct or the file has not been copied. You would need to edit the **.dat** file and change the name to the correct one or copy the missing file into the current directory. Once you have saved the new file, rerun RMCPProfile as above.

3. After a short while (being nice and vague as it depends on the machine being used), RMCPProfile should finish initialising and then display the level of the current fit, as below.

```
Chi**2/dof 1145.  
Expt 1: Renorm 1.0000 Constant = 0.0000 Chi**2/nq = 565.2  
Expt 2: Renorm 1.0000 Constant = 0.0000 Chi**2/nq = 114.7  
denergy: 0.0000000000000000  
Distance window summary  
ipar n(lt) n(ok) n(gt)  
1 0 0 0  
2 0 324 0  
3 0 648 0  
Number of moves allowed by DW constraints: 0  
Number of moves rejected by DW constraints: 0
```

```
Total time in loop so far : 4.5910478E-03 s
Time per generated move is: 1.53034925460815 ms
```

```
Saving configuration and results to disk
*** DO NOT INTERRUPT UNTIL SAVING IS COMPLETED ***
Updating neighbours file
Writing mamp file
Saving has been completed
```

```
Total time taken by program: 8.011
```

In the **rmcprofile** window, type 'ls' and hit enter to produce a list of the files in the directory. You should see that several new files have been produced. The **.his6f** and **.amp** files are just history files for restarting a run. The results of the refinement are contained in the **.rmc6f** file (which now contains the new atom locations and the original information will have been replaced), the **.out** file and also the **.braggout** file. These last two files contain the fits to the data, together with other useful information such as the partial radial distribution functions. They can be plotted using the **rmcplot** program. The fits are also available in the comma separated files **.csv** for plotting in other programs

Now type 'rmcplot [stem name]'—e.g., 'rmcplot rmcsf6\_190k'—and select '1' from the menu when it appears. A plot of the partial  $g(r)$  functions should appear in the graphics window (you may need to find it if it is hidden behind other windows). In this case, the plot should be a series of delta functions, since all the atoms are on the ideal positions given by the average structure. Press 'q' (with the graphics window selected) to return control to the menu.

Next, select '5' from the menu, and a plot of the Bragg profile fit should appear. The position of the peaks in the RMC fit should be correct, as should the peak shape and background, since these all come directly from the GSAS data. The only aspect that will be wrong is the intensity of the peaks. Again, this is because all the atoms are on their "ideal" positions, and hence there is no thermal motion or static disorder in the configuration. Consequently, the first thing RMCProfile will do is to displace the atoms in order to reduce the intensity of these peaks. The distance window constraint will help maintain the SF<sub>6</sub> octahedral connectivity during this process, and the total  $G(r)$  will define the final local structure of these polyhedra. Again, press 'q' with the graphics window selected in order to return control to the menu.

Finally, select '3' from the menu (answering '1' when asked which  $T(r)$  to plot), and a plot of the fit to the total  $G(r)$  will be displayed. As with the partial  $g(r)$  functions, the total  $G(r)$  function consists of a series of delta functions. The actual data (the red curve) is lurking somewhere at the bottom of these. Use the left-click of the mouse to zoom in on the low- $r$  region and produce a plot that spans roughly from 0 to 3.5 in  $x$  and  $-0.5$  to 6 in  $y$ . You should now be able to see the data more clearly and it will be easier to see that—unlike the Bragg profile fit—the positions of the peaks are not correct. This tells us straight away that the real local structure of the SF<sub>6</sub> octahedra and their local environment is not the same as that derived from the ideal average structure. We can say this even though we haven't yet refined anything. Press 'q' to return to the menu and select '0' to exit.

To finish our look at the initial model we will now produce a 3D model of the starting configuration in **atomeye**. Type 'get\_atomeye\_sf6 rmcsf6\_190k' in the **rmcprofile** window, this will create an **atomeye** file called `rmcsf6_190keye.cfg`. To plot this file, type 'atomeye rmcsf6\_190keye.cfg' in the **rmcprofile** window. With the atomeye graphics window active, pressing 'tab' will toggle use of the perspective view. If you then press 'page down' the atom spheres will shrink. Next, press 'b', and the S–F bonds will be displayed; this will make it easier to visualise the SF<sub>6</sub> octahedra. This configuration should look like a bcc array of octahedra, as defined by the average structure. Examine the structure and the **atomeye** function for a while, then when you are happy to move on, press 'q' whilst the graphics window is active.

4. The next step is to use RMCProfile to refine this model. Make a copy of the initial fit directory called **first.run**. Open this directory and edit the **[stem name].dat** file (e.g., `rmcsf6_190k.dat`), changing the line

```
TIME_LIMIT ::      0.00 MINUTES
SAVE_PERIOD ::      0.00 MINUTES
```

back to

```
TIME_LIMIT ::      10.00 MINUTES
SAVE_PERIOD ::      0.20 MINUTES
```

and save the file. This will tell RMCProfile to refine the structure for 10 minutes. The small configuration used here will not produce a very good fit but it should illustrate the minimisation process.

In the **rmcprofile** window, change to the **first.run** directory and type 'rmcprofile [stem name]' (e.g., 'rmcprofile rmcsf6\_190k'). At this point text should start to appear indicating what RMCProfile is doing.. The run will take 10 minutes and during this time you can also follow the progress of the fits in the another **rmcprofile** window using **rmcplot**. Every time you choose a different plot, **rmcplot** reloads the file, so in this way you can follow the progress by re-plotting the various fits. RMCProfile will save the current fits every 12 second (0.2 minutes).

Initially the fit to the Bragg profile and the  $F(Q)$  improves at the expense of the  $G(r)$  fit and so, as predicted, the octahedra are distorted to reduce the Bragg intensities. Eventually the fit to the Bragg data get close enough that the  $G(r)$  fit starts to dominate and RMCProfile starts to regularize the octahedra (whose connectivity has been maintained by the distance window constraint). As the refinement progresses the general trend is for all the  $\chi^2$  to reduce; however, if you look closely you will notice that some will increase at various times. This shows that the weighting of the data is correct and that local minima are being avoided.

Once the refinement has finished, examine the fitting parameters in the **RMCProfile** window, the Chi\*\*2 values should be below 1.

If you haven't done so already, take a look at the final fits using **rmcplot**. The fits to the Bragg profile and  $F(Q)$  should look reasonable, although there are extra peaks in the Bragg profile due to the small configuration size. The fit to the  $G(r)$  should also be fairly reasonable although the fit may not quite have finished its minimisation process, and also the noise is caused by the limited number of atoms in the configuration.

The only analysis we will do here is to produce a new atomeye file (as above) and to look to see that the atoms are more disordered.

If you would like to see the fits produced by a quick-ish run with a  $10 \times 10 \times 10$  supercell then take a look at the files in the **ex\_3\sf6** folder. Here the Bragg profile fit is probably as good as it can be. The  $G(r)$  fit is better, but still not perfect.

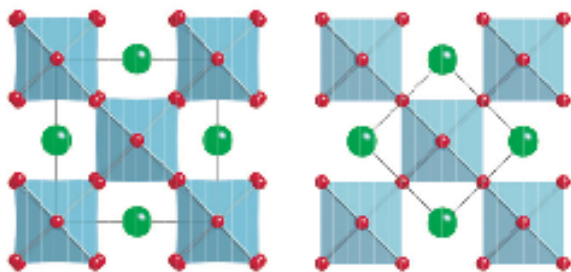
## Exercise 4

# Case Study: Using RMCProfile to study the 105K phase transition in strontium titanate

In this exercise you will use what you have learnt in the previous exercises to investigate a phase transition in the perovskite material strontium titanate ( $\text{SrTiO}_3$ ). Perovskites are made up from a network of corner sharing octahedra ( $\text{BO}_6$ ) with cations in the interstices (A) and so have the general formula  $\text{ABO}_3$ . The ideal structure is cubic with space group  $\text{Pm}\bar{3}\text{m}$ , where the A atoms are at the corners, the B atoms at the body centre and the O atoms on the cube faces. Perovskite materials have many useful physical properties, such as ferroelectricity and piezoelectricity, and are widely used in industry.

Perovskites have two main mechanisms to change their structure, the B atom can displace from the centre of the octahedra or the octahedra can tilt as a unit. These changes are usually driven by a size mismatch of the A and B atoms and often result in a distortion of the polyhedra, displacements of the A atoms and a change in space group and unit cell metric. These changes then produce the ferroelectric, piezoelectric and other useful properties these materials have.

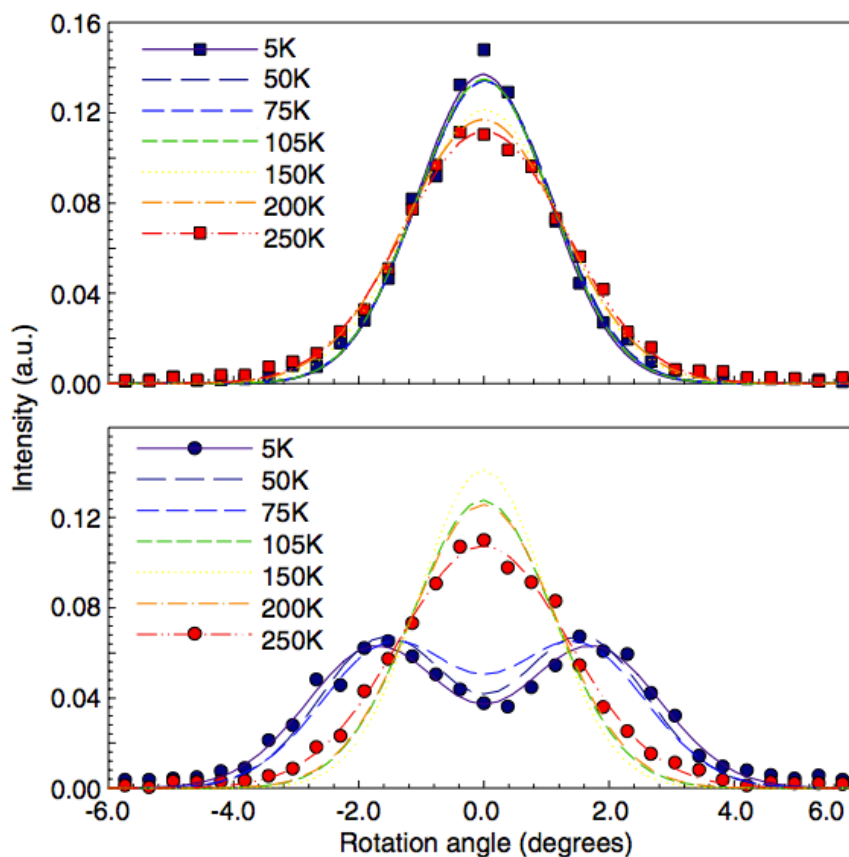
One of the simplest transition in a perovskite is the displacive transition in  $\text{SrTiO}_3$  at 105 K.  $\text{SrTiO}_3$  has the cubic  $\text{Pm}\bar{3}\text{m}$  structure at room temperature and on cooling below 105 K the octahedra begin to tilt to form a tetragonal structure, with space group  $\text{I4/mcm}$ .



The figure shows the low-temperature structure on the left and the high-temperature cubic structure on the right. If you look carefully at the low-temperature structure you should be able to see the slight displacement of the oxygen atoms (red) caused by the rotation of the octahedra. The tilting is very subtle and at 5K it is only of the order of 2 degrees from the ideal arrangement.

It is quite possible that the phase transition in  $\text{SrTiO}_3$ , together with that in quartz, is one of the best studied of all displacive phase transitions. Because of this, it is interesting to apply new techniques to this system as they become available, partly to provide a useful benchmark, and partly to see if the new technique can provide new insights. So it was natural to try running RMCProfile on this systems to see if

the method and the data were sensitive to such a small effect on a local level. The results were published in our paper (Hui et al JPCM **17** (2005) S111-S124) and the aim here is to recreate part of this analysis. The figure below shows the tilt angle distribution extracted from RMCPProfile models produced by fitting neutron total scattering data. It illustrates one of the main advantage of large box modelling in that you obtain a distribution of values rather than just a single average value. The aim of this exercise is to recreate this plot for data at 5 K and 293 K.



The **ex\_5** folder contains all the files you will need for this exercise. In the top level folder there are two further folders **5K** and **293K** and these contain **data**, **gsas**, **rmc** and **gasp** folder that contain the files needed for each temperature.

1. Start by moving to the **5K** folder and using the files in the **gsas** folder to extract the bragg data and a starting configuration of a  $6 \times 6 \times 4$  supercell, with the atoms ordered as [Sr Ti O].
2. Next move back to the top level of the **5K** folder and prepare the main RMCPProfile .dat control file, you can copy the file from the **RMC** folder if you don't want to start from scratch. I suggest you call it `srtio3_5k.dat` and set it to fit the neutron  $G(r)$  file `srtio3_wk_5k_rmc.gr` and the neutron  $F(Q)$  file `srtio3_wk_5k_rmc.fq` and the bragg data with a DMIN of 0.5. Set the `TIME_LIMIT` to 10.0 minutes and the `SAVE_PERIOD` to 0.2 minutes. The .dat file in the rmc folder should be setup this way.
3. Now you should copy the files you need from the **data** and **gsas** folders to the top level of the **5K** folder and you should be almost ready to run RMCPProfile. Remember the files from

the **gsas** folder need to be copied to have the same “stem names” as the main `.dat` file e.g. `'srtio3_5k'`. Also check that the names of the data files copied from the **data** folder are the same as those listed in the control file.

4. Again we will use the distance window constraint to maintain the connectivity of the structure during the rmc refinement. Plot the `.gr` file with `datampplot` or another plotting program. Since this is neutron data and Ti has a negative neutron scattering length (-3.438 fm), peaks involving Ti and another atom will have a negative amplitude, such as the first peak in the data which is due to the Ti-O bonds. So the `.dw` file needs to have a first line defining the minimum distances and a second line defining the maximum distances for each pair, they are in the order (Sr-Sr, Sr-Ti, Sr-O, Ti-Ti, Ti-O and O-O) and in the data these are the 4th, 3rd, 2nd, 4th, 1st and 2nd peaks in the `.gr` file with increasing  $r$  respectively. Note that due to the symmetry of the structure some peaks are due to more than one correlation. So by examining where these peaks start and finish you should be able to decide in suitable values for the distance windows, such as:

```
3.7 3.1 2.50 3.7 1.68 2.50
4.1 3.6 3.05 4.1 2.21 3.05
```

A suitable `.dw` file containing these values is located in the `rmc` folder, so either copy this to top level of the **5K** folder once you have checked you are happy with the values (NB if you press 'c' with the graphics window of `datampplot` active a crosshairs will appear that you can use to read positions from the graph and the middle mouse button will exit this mode, or just `ctrl-C` out of the program) or produce your own files with similar values.

5. You should now have all the files you need to run `RMCPProfile` and a listing of the directory should look similar to this:

```
data srtio3_5k.back srtio3_5k.inst
gasp srtio3_5k.bragg srtio3_5k.rmc6f
gsas srtio3_5k.dat srtio3_wk_5k_rmc.gr
rmc srtio3_5k.dw srtio3_wk_5k_rmc.fq
```

Make a directory called `first_run` and copy the `RMCPProfile` files to this directory e.g. `'cp sr* first_run'`. Now run `RMCPProfile` in this new folder and follow its progress in another `RMCPProfile` window with `rmcplot`. After 10 minutes the `Chi**2` values should be between 20 to 30. If you plot the PDF fit you will see, that while it is fairly good, it is also clear that in the `rmc` fit (the black line) the peaks are too intense as their areas are larger than those of the data (the red line). The area in the `rmc` fit is defined purely by the structure supplied (mainly the density of atoms in the configuration) so unless this is wrong the problem is that the data are not scaled correctly. This is a fairly common problem since when preparing the data is often quite hard to be sure it is on the correct absolute scale and unfortunately, since the number of atoms in the `rmc` configuration is fixed, `RMCPProfile` is fairly intolerant to this scale issue and it will introduce extra disorder to attempt to reduce the peak areas. However, `RMCPProfile` can be used to indicated the correct scaling required and then the data can be corrected if possible.

Edit the main control file so the `TIME_LIMIT` and the `SAVE_PERIOD` are set to 0.0 minutes. Then in the two data sections change the keyword `NO_FITTED_SCALE` to `FITTED_SCALE` then save the file. This will ask `RMCPProfile` to run through one cycle but scaling the data to give the best fit, so run `RMCPProfile` again with these settings.



After a short while RMCPProfile should end and the  $\chi^2$  values for the  $G(r)$  and  $F(Q)$  data (EXPT 1 and 2) should lower. The `Renorm` value, which was 1 before, should now be between  $\sim 1.2$  and  $\sim 1.1$ , this gives an indication of how much the data needs to be rescaled.

Since data processing is outside the scope of this tutorial we have prepared some rescaled data and put it in the data folder called `srtio3_wk_5k_sc_rmc.gr` and `srtio3_wk_5k_sc_rmc.fq`. (If of course you have already been fitting these files by mistake then you are wondering what I am wittering on about; in this case please try again fitting the correct data without the `_sc` in it.) Assuming everything so far makes sense and has worked please make a new directory called `second_run`, then copy the fresh RMCPProfile files from the top of the **5K** folder into it and the rescaled data from the **data** folder into it. Then edit the main control file so RMCPProfile will fit this new data and run RMCPProfile. If everything is correct the initial  $\chi^2$  values should be large ( e.g. 'EXPT1 > 500'), otherwise the files being used are not "fresh" and the final result may not be as good.

After the 10 minute run the Bragg  $\chi^2$  should be less than 20 and the others for the  $G(r)$  and  $F(Q)$  data should be less than 7 and the fits to the data should look good. If this is not the case but they are close to these values then RMCPProfile may just need to run for a bit longer, so set the time limit to 5 minutes and rerun RMCPProfile. Once you are happy with the fits you are ready to analyse the configuration.

6. The first analysis basic analysis, as with  $\text{SF}_6$ , is to look at the configuration visually as a ball and stick model with a suitable program such as `atomeye`. To produce the correct input file, enter `get_atomeye_srtio3 [stem name]` this should produce a file called `[stem name]eye.cfg`/ Open this file in a text editor and change the line `A = 1.0 Angstrom` to `A = 10.0 Angstrom` then save the file. This is just to overcome a problem `atomeye` has with atoms that have more than 24 neighbours, you should be able to view the file now by entering `atomeye [stem name]eye.cfg`. You should be able to see the atoms have jiggled about a bit but they are basically very ordered.

Next we will analyse how the octahedra have rotated using a program called GASP, this stands for Geometric Analysis of Structural Polyhedra and as this suggests it uses geometric algebra to analyse whether the polyhedra have rotated or distorted and by how much. To do this GASP needs two configurations of atoms to compare and it will then see if you can match the atom positions be just rotating the polyhedral or if it needs to distort the polyhedra. The manual for GASP is in the **tutorial** folder of the main RMCPProfile distribution, so please look at this if you would like to know more about the program.

So to run GASP on the  $\text{SrTiO}_3$  results we need two configurations, the first is the `.rmc6f` we have just produced, so please copy this to the `gasp` folder in the **5K** folder, then move to the **gasp** folder. Next we need to create an ideal configuration with everything the same as the 5K structure apart from the tilt angle should be zero. In the **gasp** folder you should find two crystallographic information files (`.cif`, if you compare these in a text editor you will see that the only difference is that the oxygen x and y positions in the `srtio3_5k_cubic.cif` are all exactly 0.25 or 0.75, whereas they should be slightly offset from this in the tetragonal structure due to the tilting, hence the `'_cubic'` as it is now more like the untilted cubic structure. So we can use this file with `data2config` to produce the ideal configuration for the GASP analysis by entering `data2config -noannotate -order [Sr Ti O] -supercell [6 6 4] -rmc6f srtio3_5k_cubic.cif`.

Next we need to convert the two `.rmc6f` file into the correct format for `gasp`, so enter `get_gasp_srtio3 srtio3_5k` and `get_gasp_srtio3 srtio3_5k_cubic`, this will produce two new `.dat` file that `GASP` needs.

The main `GASP` control file is called `srtio3_gasp.cfg` and should contain:

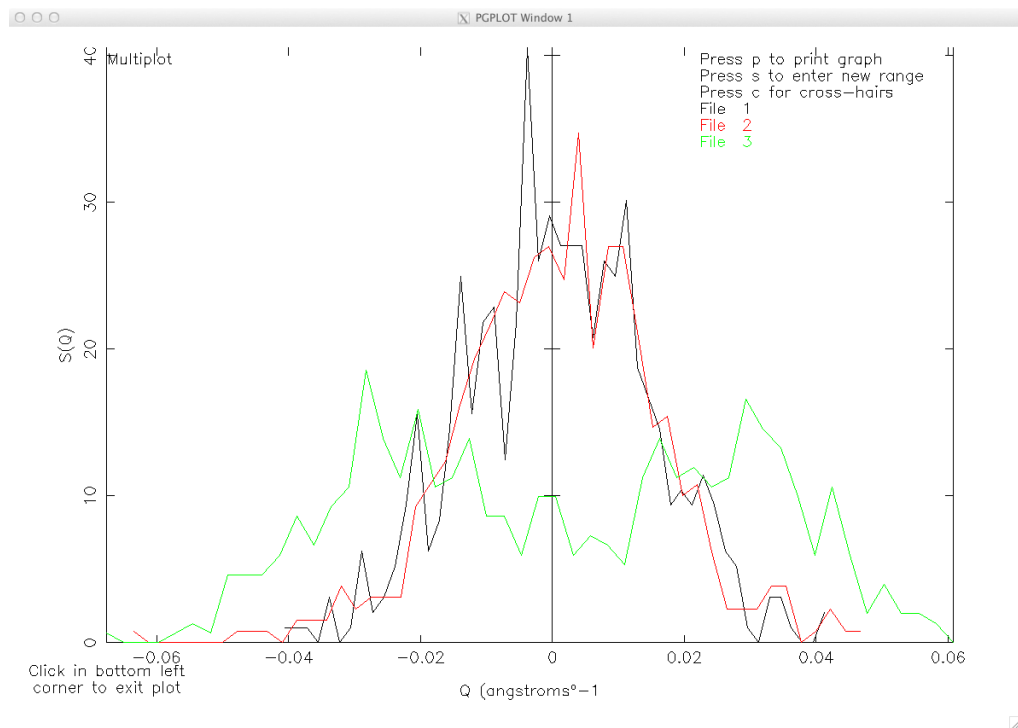
```
SrTiO3
Sr Ti O
OCT Ti O 2.21

PRE srtio3_5k_cubic
fast srtio3_5k srtio3_5k_cubic

com srtio3_5k srtio3_5k_cubic
```

This file format is described in the `GASP` manual and here it is setup to compare the files `srtio3_5k` and `srtio3_5k_cubic`. so if your 'stem name' is not `srtio3_5k` you will need to change this in the file to the correct name.

To run this file with `GASP` just enter `gasp` and enter `srtio3_gasp` when asked Please give the name of the configuration file. Don't worry if it ends with some `malloc` error and/or a strange phrase possibly about `Elvis`, the first is just a system message and the second is from the author Steve's sense of humour! If everything has worked there would be some new `.con`, `.pol` and `.com` files in the folder, the `.com` file has the rotor comparisons we are after.



To extra the information about the octahedral tilt enter `get_gasp_rotors srtio3_5k_srtio3_5k_cubic.com srtio3_5k_rot`, again this assumes you have used the 'stem name' suggested and so the `.com` file may have a different name, in this case change the command as needed. If everything runs well this should produce three `.rot` files for the x, y and z rotation direction. These are just text files containing a list of values for each of the 576 octahedra. To be able to plot these we need to histogram the data, so enter `histogram_data` followed by the name of a `.rot` file for each of the files, e.g. `'srtio3_5k_rot_x.rot'`. All being well you now have three `.bin.dat` files and these can be plotted with **datamplot** or some other plotting program. Entering `datamplot srtio3_5k_rot_xbin.dat srtio3_5k_rot_ybin.dat srtio3_5k_rot_zbin.dat` (assuming the file are called this) should produce a plot very similar to the one in the figure above. Clearly the distributions in files 1 ad 2 are centred about zero, whereas the file 3 distribution has a maximum at about 0.03 and -0.03, these are the rotations for the z rotor in radians, so about 1.7 degrees. This is very close to the expected  $\sim 2$  degrees.

7. So the obvious next question is does RMCProfile always produce this double maximum distribution in one direction or is it coming from the data. To answer this question we need to analysis some data from a sample in the cubic phase. The **293K** folder contains all the data we need to do this, so please repeat the process above but with this data. This time start with the `_sc_rmc.gr` and `_sc_rmc.sq` (unless you want to repeat the rescaling part as well) and use a supercell of  $8 \times 8 \times 8$ . Again the **rmc** folder contains a control file you can start with and since the start configuration has zero tilt, the unrefined configuration can be used as the second configuration in the GASP analysis.

## Exercise 5

### Other things to try

So well done if you have completed everything and reached this point! By now you should be able to happily run RMCPProfile and analyse the data to some extent at least. However, you have not used many of the features of RMCPProfile and we have skipped over quite a few things by supplying starting files. So please consult the manual for anything you don't understand (please be aware it is still a work in progress) and/or perhaps try out some of the following:

1. What effect does using the wrongly scaled data have on the final octahedral tilt results?
2. What is the effect of just fitting one or two of the different data types i.e just Bragg or just Bragg and the  $G(r)$ ?
3. If you rerun RMCPProfile with the same starting files you will get a slightly different final configuration. This can and should be used to improve the statistics of your results. So rerun RMCPProfile several time and then combine the .rot files before histogram-ing them. This should produce less noisy distributions. You can also use the runs from other people or run a single large configuration.
4. What does the **CONVOLVE ::** option do when fitting reciprocal space data and what does the fit look like if you turn it off?
5. You can fit different functions by specifying the **FIT\_TYPE ::**. Try fitting a  $D(r)$  or  $T(r)$  from the  $G(r)$  data or a  $QF(Q)$  from the  $F(Q)$  data. What effect does this have on the final fits and results?
6. Using the **CML ::** keyword and **ccviz** command can be using to produce a .xhtml interactive log file that can be read by modern web browsers. This feature may have been on during your runs so you can run **ccviz** on your .cml file or if not you could produce some now.

## Appendix A

# GSAS refinements

As mentioned above most RMCPprofile refinements start by producing a configuration from a Rietveld refined average structure. For historical reasons the program we use for our Rietveld refinements is GSAS and so data2config is designed to read the required information from the GSAS files. However, for data2config to be able to acquire the information it needs GSAS must be run in the following way:

1. It is best that the refinement only contains one phase, since additional phase information in the files may confuse things.
2. At present the only background function supported is the Chebyshev polynomial.
3. The d spacing range of the refinement should be set to the same range as you intend to fit in RMCPprofile.
4. Once you are happy with the refinement a `.TBL` file needs to be produced using the “pubtable” function.
5. The histogram you would like to fit in RMCPprofile needs to be written to the `.LST` file using the “hstdmp” function.