# Phasing, phase improvement and model building in CCP4

Kevin Cowtan, April 2008

**Introduction:**

In this tutorial we will take the SIRAS data for the RNAse test data at 3A, and use it to obtain an atomic model of the structure, using CCP4 programs for phasing, phase improvement, and model building. Many of the techniques described here may also be employed when solving structures by molecular replacement.

**Overview:**

The tutorial will involve the following steps, which will demonstrate some of the problems you might encounter when solving a real structure:

1. Perform automated heavy atom location and phasing, using CRANK.
2. Perform density modification to improve the electron density, using DM.
3. Perform automated model building using BUCCANEER.
4. Examine the model using COOT. The electron density is poor and the model is incomplete, so we will go back and try and get a better map by using NCS.
5. Determine the non-crystallographic symmetry and perform NCS averaging, using DM.
6. Perform automated model building using BUCCANEER using the improved map.
7. Examine the model in COOT. Use the validation tools to identify problems and fix them.

This document provides the basic crystallographic information you will need to solve this structure – information which you would have for a real problem. However the detailed instructions for each step are given on the new CCP4 documentation website: www.ccp4wiki.org.

Open this site in your web browser now. Select the link "Using the CCP4 software". This page gives information about the basic steps of the structure solution process. We will be looking at 3 of these steps: "Experimental phasing", "Phase improvement", and "Model building".

**Getting started:**

Launch CCP4i and create a new project for this tutorial. (Top right of the window: Change project/Add project). Pick a suitable name for your project (e.g. PhasingTutorial), and a directory for the files.

You will need two files for this tutorial: an MTZ file containing the observed structure factor data, and a sequence file containing the protein sequence: `rnase30.mtz` and `rnase.pir`. The locations of these files will be provided.

**Step 1: Automated heavy atom location and phasing using CRANK.**

Go to the "Experimental phasing" page on ccp4wiki, and select "Automated experimental phasing with Crank". This page will tell you how to run Crank.

You have the following information:

> *The data for the RNAse structure has been collected on 2 crystals, a native, and a Platinum derivative, for which anomalous data has been collected using CuKa radiation. There are estimated to be around 5 Pt atoms in the asymmetric unit.*

The MTZ file contains structure factor *amplitudes* (not intensities) for both datasets.

The native data is in the columns labelled:

`FNAT,SIGFNAT`

The derivative data is in the columns labelled:

`FPTNCD25(+),SIGFPTNCD25(+),FPTNCD25(-)SIGFPTNCD25(-)`

With the help of the Crank documentation, set up a phasing calculation for this data.

Before you click run, we will change some of the default options to save time. At the bottom of the Crank window, there is a section labelled "Experimental pipeline".

- Change the first two menus to read "Start the pipeline with substructure detection and end with

substructure phasing".

- Change the pipeline menu to read "SHELXCDE/BP3/SOLOMON/ARPWARP"

Now click "Run now".

**Step 2: Density modification using DM.**

Go to the "Phase improvement" page on ccp4wiki, and select "Phase improvement with DM".
Follow the instructions on this page to first determine the solvent content, and then to run DM using the MTZ file which was created by CRANK.

**Step 3: Automated model building using BUCCANEER.**

Go to the "Model building" page on ccp4wiki, and select "Automated model building with Buccaneer". This page will tell you how to run Buccaneer.
Follow the instructions on this page to run Buccaneer using the MTZ file which was created by DM.
Note: the current version of Crank does not leave a Free-R flag in the MTZ file – for now just turn off this option in the buccaneer task window.
While Buccaneer is running, start looking at the map in Coot (next section).

**Step 4: Examine the model using COOT.**

Launch Coot and read in the model from Buccaneer.
Read in the MTZ file from DM, using the columns `FDM` and `PHIDM`, and get an idea of the quality of the map.
Look at the chains in the model using the "Draw/Go to atom" tool. The chain trace is broken into several chain fragments, and some chains are unsequenced (UNK).
The cell content analysis indicated 2 mols/ASU. You may now be able to identify NCS from the partial model. From the sequence in the "Draw/Go to atom" you may fine some sections of sequence which have been built more than once. Or you may be able to identify related chains by shape, by viewing the Ca-trace. Make a note of the chain Ids and the range of overlapped sequence numbers (excluding UNKs unless the conformations look similar). If you can't find any NCS, you may need to run buccaneer for a few more cycles.

**Step 5: NCS averaging using DM.**

Go to the "Phase improvement" page on ccp4wiki, and select "Phase improvement by NCS averaging with DM".
Follow the instructions in the section "Determining the NCS operators using a partial model" to determine an NCS operator using the duplicated residue ranges you found earlier. Check the RMS difference between the superposed chains is reasonable (<1.5A).
In the CCP4i task list, select your previous DM job, and then click "Rerun job" from the right-mouse context menu. Set up an averaging calculation using the NCS operator you determined previously.
Give the output MTZ file a new name. Run the job and examine the output carefully. Follow the instructions in the section "Program output" to determine if the NCS calculation has worked.

**Step 6: Automated model building using BUCCANEER.**

In the CCP4i task list, select your previous buccaneer job, and then click "Rerun job" from the right-mouse context menu.
Change the input MTZ file to the one you created with your DM averaging job and run it.
Compare the output log file, and in particular the summary sections, to see if the results have improved.
While Buccaneer is running, compare the un-averaged and averaged maps in Coot.

**Step 7: Complete the model in COOT.**

Launch Coot and read in the model from Buccaneer.
Read in the MTZ file from DM and get an idea of the quality of the map.
Use the "Model/fit/refine" tools to correct any breaks in the chain.
Use the "Validate" tools to identify any problems in the model and correct them.
Add waters to the model. Re-refine your model in Refmac.