

Ab initio model building with Buccaneer

Buccaneer is a program for *ab initio* model building. It automatically builds protein chains into cryoEM density maps using sequence information alone and doesn't require a homology model. Only two inputs are required: **(1) 3D map** (mrc format) and **(2) amino acid sequence** expected of the complex (typically fasta format but others supported). Buccaneer works very well with high resolution maps $\leq 3\text{\AA}$ but models can still be generated from maps up to $\sim 4\text{\AA}$. This tutorial will show you how to use it via the CCP-EM GUI.

In this example we will use the *E. coli* β -galactosidase map from the RELION 3.0 tutorial. To save time we have segmented the map in Chimera such that only one chain is present. You will use:

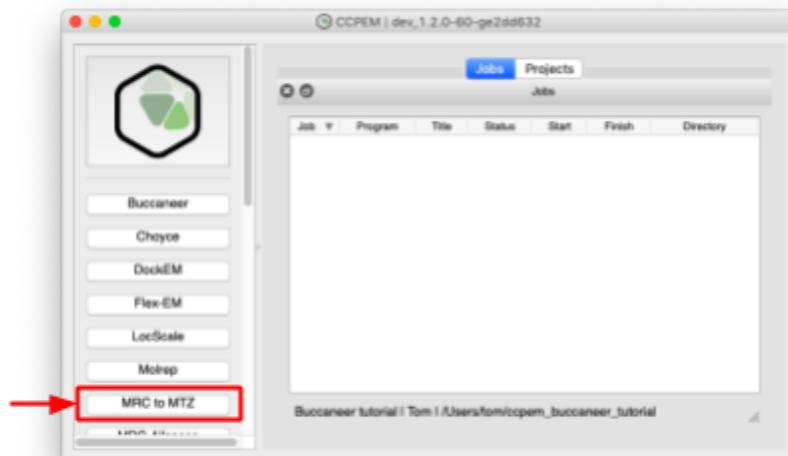
Map: run_class001_flip_seg_prep.mrc (2.9 \AA resolution)
Sequence: 5a1a.fasta.txt

You can find these in the Buccaneer tutorial data.

Part 1) Optimise global map sharpening: MRC to MTZ

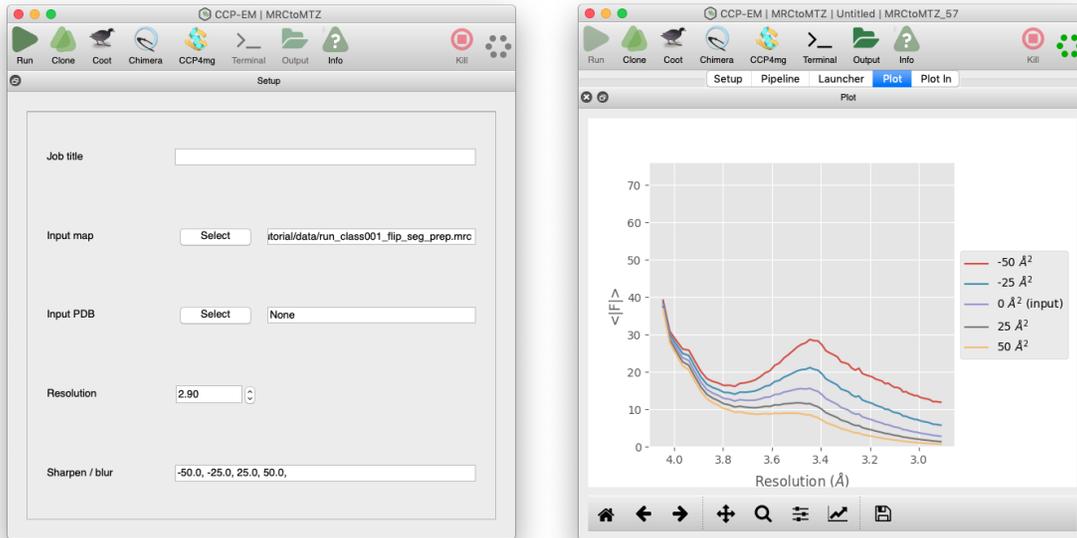
Before running Buccaneer, it's recommended to optimise the map sharpening. During the reconstruction process it is not possible to determine the absolute sharpening value of the map, in particular if your target is a multidomain protein. Thus it is recommended to apply an array of blurring and sharpening values and visually inspect your map prior to any model building.

1) Click 'MRC to MTZ' to launch task from the main GUI.



2) Enter map and resolution and set the Sharpen / blur values (try -50, -25, 25, 100). N.B. negative values sharpen, positive blur (hover the mouse over the input box to see the tooltip).

You may also find it useful to add a job title. Then press “Run”.



3) Map inspection: power spectrum

When the job is finished the power spectrum is displayed in the ‘Plot’ and ‘Plot In’ tabs.

What is the effect of blurring and sharpening? What do you expect the power spectrum to look like?

4) Map inspection: visual in Coot

Press the ‘**Coot**’ button in the toolbar. In this workshop we will use interactive sharpening in Coot.

Click ‘Display Manager’ in the toolbar to toggle the display of the different maps.

For ‘starting_map.mtz FoutBlur_50.00 Pout0’ and ‘starting_map.mtz FoutBlur_25.00 Pout0’ click ‘**Delete Map**’ to remove these.

To increase the viewable area of the map set the map radius to 25Å:

Edit → **Map Parameters** → **Map radius** → **25**

To show the box outline:

Draw → **Cell & Symmetry** → **Show Unit Cells?** → **Yes**

To move around the maps without an atomic model hold ‘**ctrl**’, hold **left click**.

Use the mouse scroll to set the contour level

Interactive map sharpening:

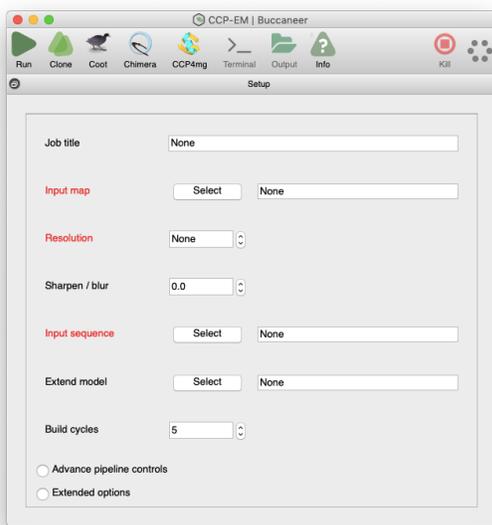
Calculate → **Map Sharpening/Blurring** → **Move the slider**

What is the effect of blurring and sharpening?

Which value gives the most interpretable map and what features are desirable?

Part 2) *Ab initio* model building: Buccaneer

- 1) From the main GUI launch the Buccaneer task window.
- 2) Enter the following parameters and hit run:
Input map: run_class001_flip_seg_prep.mrc
Resolution: 2.9
Input sequence: 5a1a.fasta.txt
Sharpen / blur: <value decided above>
Build cycles: 2
Advanced pipeline controls -> CPU: 2



The Buccaneer pipeline takes ~1hr (on MacBook Air) for this dataset.

5 build cycles is the default value, 2 is used here to save time.

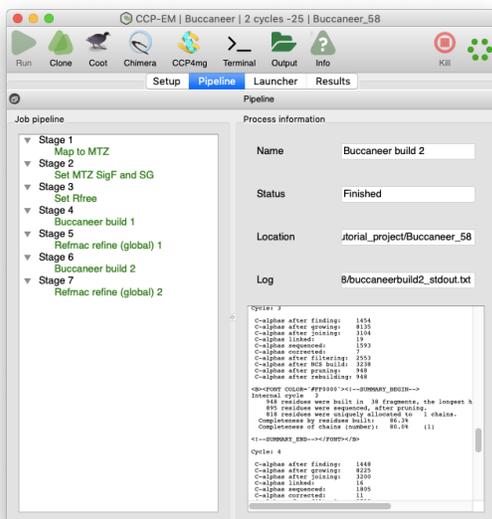
If you are unsure which sharpen / blur value to use try -25 or use the 'Clone' button to generate another task with the same parameters. You can then run multiple jobs in parallel with different values.

The original fasta sequence file can be found from the downloads link in PDB entry:

<https://www.ebi.ac.uk/pdbe/entry/pdb/5a1a>

Chains B,C,D were manually removed in a text editor.

- 3) Monitoring job progress in 'Pipeline' tab



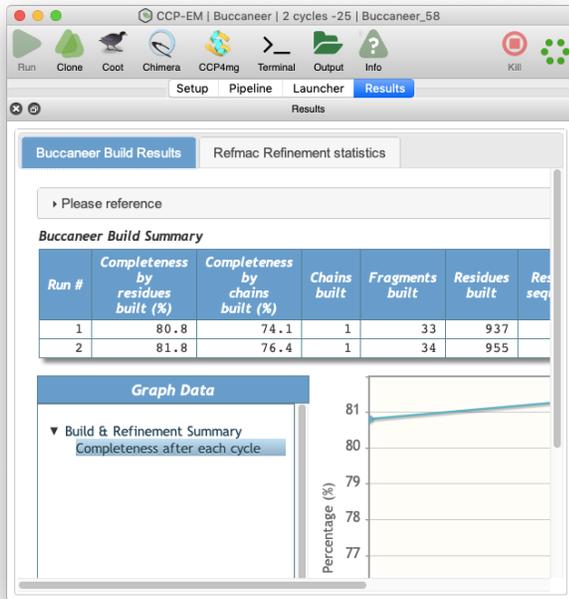
Whilst the job is running the green hexagon in the top right will rotate.

You can see the progression through the various stages in the pipeline on the left.

Green is finished, blue running, grey yet to start (red failed).

Clicking on the stage will display log file for that stage. Double clicking on the log file window will launch your preferred text editor with the log file.

4) Results inspection. Click the 'Results' and the 'Launcher' to see the output from Buccaneer.



When the job is complete (hexagon solid green) the 'Results' tab shows the number of residues built.

You can see the completeness of the model in the 'Buccaneer build results' tab and 'Refmac refinement statistics' tab shows an overview of the refinement statistics after initial automatic refinement in Refmac.

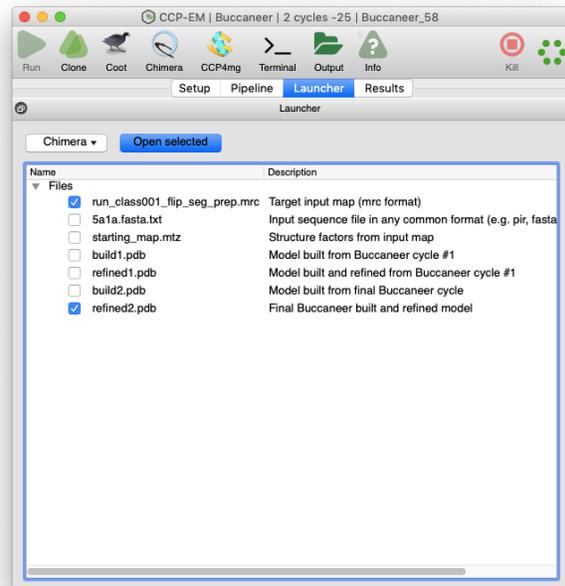
'Completeness by residue' refers to the total number of residues built from the input sequence.

'Completeness by chain' refers to the percentage of residues built and assigned to expected chain.

The 'Launcher' tab allow you to select which files your interested in viewing for visual inspection.

Use the post refinement models i.e. 'refined<cycle>.pdb'.

N.B. Model building is stochastic so you may find models from intermediate steps are better than the final model if you run multiple cycles.



If you ran multiple cycles which build cycle worked best for you? If you ran parallel runs with different input parameters (e.g. sharpening, resolution cutoff, etc) what effect did these have?

5) Validating the results in Coot. Select the map and coordinates in the 'Launcher' tab and press the 'Coot' button.

In Coot you use the 'Display Manager' to select the maps and model. Use the sequence viewer (Draw → Sequence View) to see the sequence of the built fragments. Clicking on the residue will orientate the view.

Do the fragments look correct? Can you see any errors in the structure, e.g. sequencing errors or register errors and if so can you manually correct them in Coot? Are there incorrectly built sequences that should be removed? At the end of the fragments can you manually built additional residues? If you have time you can 'walk' through the structure by pressing space and this moves the focus by one residue at a time. You can then optimise each residues fit to density and stereochemistry (see Coot tutorial for further details).

Cheat hint: load additional file 5a1a_chainA_molrep.pdb to show a complete model. This was generated by docking chain A from the PDB entry into the map via Molrep so it's not a perfect fit but shows a fully sequenced model which you can compare with the ab initio model generated by Buccaneer.

When you are ready you can save the edited coordinates (File → Save Coordinates) and run Refmac (see Refmac tutorial).

Buccaneer references:

Cowtan, K. The Buccaneer software for automated model building. *Acta Cryst D* 62, 1002-101, 2006

Cowtan, K. Fitting molecular fragments into electron density. *Acta Cryst D* 64, 83-89, 2008.

CCP-EM reference:

Burnley, T., Palmer, C.M. & Winn, M. Recent developments in the CCP-EM software suite. *Acta Cryst D* 73, 469-47, 2017.

Contact:

Do please report any issues or bugs.... it's much appreciated and helps us make the software better:

ccpem@stfc.ac.uk