Model Building and Validation in CCP-EM Doppio

This tutorial will cover basic model building and validation in the CCP-EM Software Suite using the Doppio interface. We will cover:

- Fetching and docking alpha-fold 2 models into a 3D map
- Manual refinement using Moorhen
- Automated refinement with Refmac-Servalcat
- Adding and refining ligands
- Model validation

We will continue from the Relion SPA tutorial and will build an atomic model for beta-galactosidase using some of the model building tools in CCP-EM. **Because there may be variation in the way the reconstruction was calculated the job numbers used in this tutorial may vary.**

Fetch and dock an AF2 model

Click the **NEW JOB** tab and search for the **Fetch from AlphaFold DB** job from the **Fetch** section. We can get the appropriate model for *Escherichia coli (strain K12)* beta-galactosidase enter the following:

UNIPROT accession number: P00722

Also get FASTA sequence file? Yes

Click **RUN** to start the job. You can follow the progress in the **LOGS** tab and when complete you can look at the model in the **RESULTS** tab.

We'll use this as a starting model and first we need to fit this into the 3D electron density from the SPA tutorial in the correct orientation.

We will use a Molrep job to perform a 6D search (rotation and translation) to move the model into the density.

From **NEW JOB** find **Molrep** in the **Atomic Model Fit** section and enter the following (leave the other parameters with the default settings**:**

Input map: PostProcess/jobXXX/postprocess.mrc
Use your postprocessed map from the SPA tutorial

Input model: Fetch/jobXXX/AF-P00722-F1-model_v4.cif

Copies to find: 1
The AF2 model is a single monomer but for now we only want to dock a single copy.

Rotation peaks: 1

Translation peaks: 1
Should you want to dock multiple copies in the future Doppio will prompt you to increase the number peaks to search for such that Molrep can find additional copies.

Mode: RTF
Normally the standard Rotation Translation function works best but for some datasets the Spherically Averaged RTF (SAPTF) is better.

Click **RUN** to start the job. The job will take a few minutes to complete.

Click on the RESULTS tab, here you should be able to see the AF2 model successfully docking in the 3D map.

Checking the handedness

It is impossible to determine absolute handedness from a data set without tilting the microscopy stage. The SGD algorithm in the Initial Model job therefore has a 50 % chance of being in the opposite hand (and therefore every map that follows in subsequent jobs).

Examine the results of the Molrep job. If the model doesn't look like it fits well in the map the map's handedness may be incorrect (if using the precalculated results from the Relion 4.x tutorial it should be correct).

To help visualise the model in the map you can adjust the Opacity and Iso Value (a.k.a. contour value) from the Volume menu in the Mol* graphical viewer in the results tab.

The **`Flip map handedness`** job in **`Map Utilities`** can be used to flip the maps if this is the case. You will need to run this for multiple maps for downstream processing. Add additional maps to the input with the **+** button next to the file inputs

`Input map:` **`PostProcess/jobXXX/postprocess_masked.mrc`**

`PostProcess/jobXXX/postprocess.mrc`

`Refine3D/jobXXX/run_half1_class001_unfil.mrc`
This file should be from your final refine3D job

`Refine3D/jobXXX/run_half2_class001_unfil.mrc`
This file should be from your final refine3D job

`Refine3D/jobXXX/run_class001.mrc`
This file should be from your final refine3D job

`MaskCreate/jobXXX/mask.mrc`
This file will not be automatically suggested by the input field and will need to be typed in manually

Finally clone the previous Molrep job by right clicking on it and selecting **`clone job.`** Change **`Input map`** to the map generated from the flip map job and run it.

Initial model inspection in Moorhen

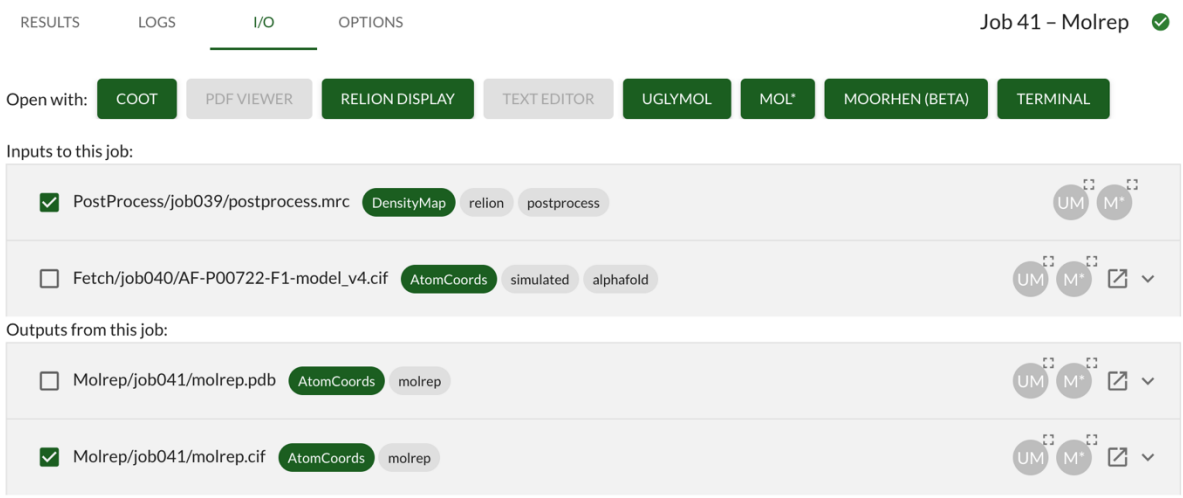
Go to the **`I/O`** tab in the **`Molrep`** job. You will see a list of input and output job nodes. You can select these and then launch additional viewing utilities which will be preloaded with your selections.

Some are external programs, e.g. Coot, Chimera/X, Relion Display and Terminal. These will only work at present when Doppio is running locally and the programs are available.

Others are internal web apps such as Moorhen and UglyMol which work in either local or remote modes of Doppio.

Here we will manually edit the model directly in Doppio using Moorhen. Moorhen is the new web version of Coot which uses Coot's processing tools in the background but with a new browser-based UI.

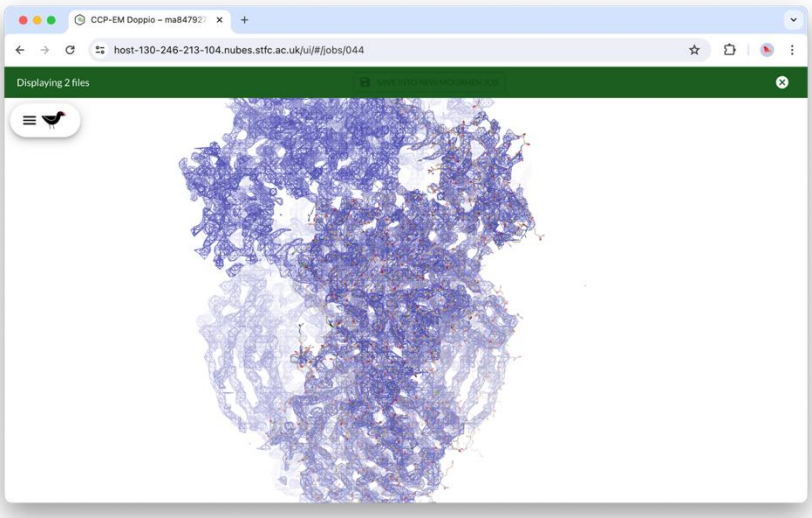
Select your postprocessed map and fit model from the I/O tab. (These may be different depending on if your map needed to be flipped)



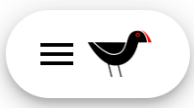
`Inputs: PostProcess/jobXXX/postprocess.mrc`
If you had to flip you map then this will be the flipped map instead.

`Outputs: Molrep/jobXXX/molrep.cif`

…and then click **`Moorhen(beta)`** to launch with the above files (your browser will automatically download the data so there may be a short delay depending on the size of the files and the connection speed).



Here you can see the model and the density map. If the network connection is slow the graphics can be a laggy. You can try reducing the size of the window and/or reducing the displayed map radius:

Click on the Moorhen  and select **`Maps`** from the side panel.
Set the map radius to **`20`**

Getting started in Moorhen - Controls

A detailed tutorial on basic controls of Moorhen is available at:
https://moorhen-coot.github.io/wiki/2024/05/20/Moorhen-Cryo-EM-Tutorial.html

Full list of controls

From the Moorhen side panel:
Click **`Help -> Show controls...`**

Zoom in and out

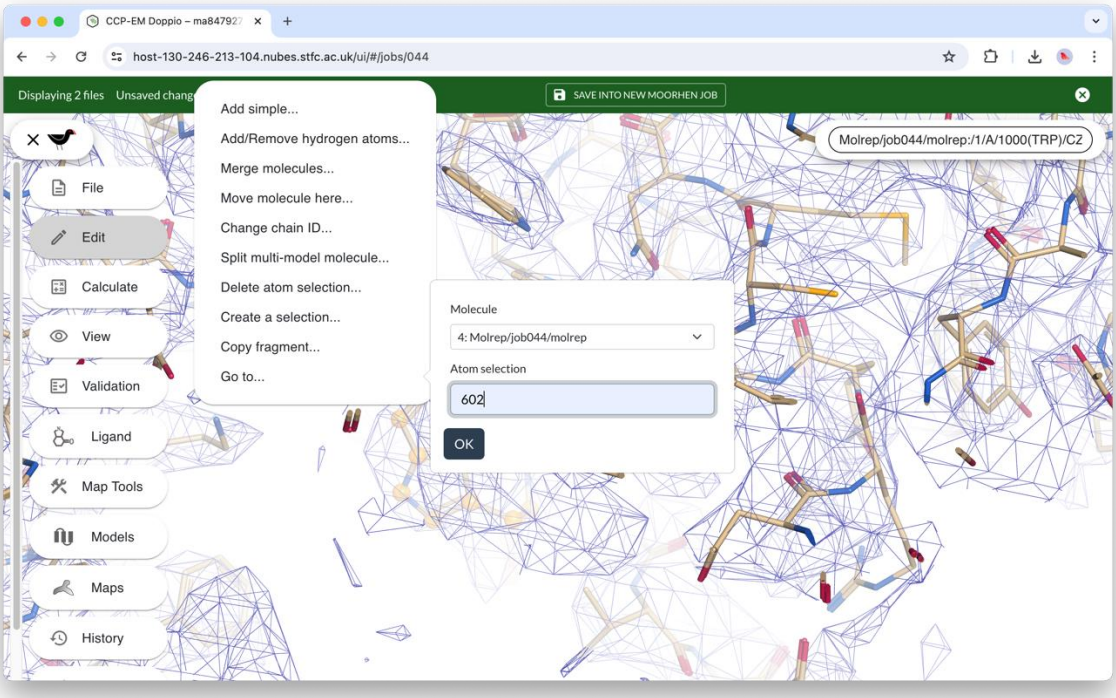
(1). Use the middle mouse wheel to Zoom in and out.
(2). **`Alt`** + Left mouse button diagonal drag.
(3). Zoom in and out using trackpad.

Recentre the display

There are multiple ways to centre the view on a specific residue:

(1) Define atom to recentre on

- **`Edit → Go to...`**



Use the notation **`/model/chain/residue`** e.g **`/1/A/723/`** for residue 723 on chain A of model 1.

(2) Visually select an atom to recentre on

Click middle-mouse (or option click on a Mac) on an atom in the graphics window

(3) Move forwards or backwards along chain with the keyboard

Left click on a residue then use **`Space`** to more forward (towards the c-termius) and **`Shift+Space`** to move backwards (towards the N-terminus).

(4) **`Shift+G`** will display a pop up box where you can type a residue to move to.

Change the Clipping (Slab)

Keyboard keys:

`1 - Increase front clip`
`2 - Decrease front clip`
`3 - Decrease back clip`
`4 - Increase back clip`

Recontour the Map

(1) Mouse: **`Ctrl + scroll wheel`**

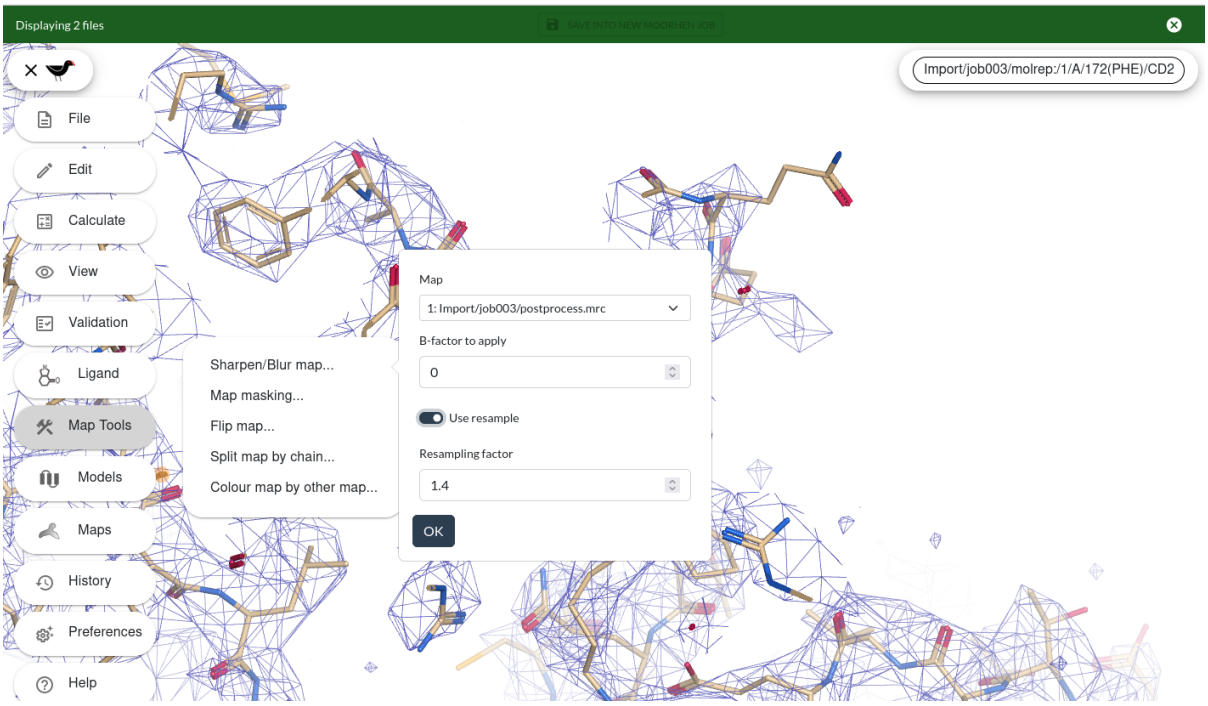
For EM maps it is important to set the Mouse wheel map contour sensitivity in the Preferences. 0.1 works best for this map (lowest sensitivity).

N.B. if you are displaying multiple maps the Scroll radio button in the Maps allows you to select which map is affected by this.

Map preparation

Often, it is easier to interpret an oversampled EM map. To oversample the **`postprocessed.mrc`**:

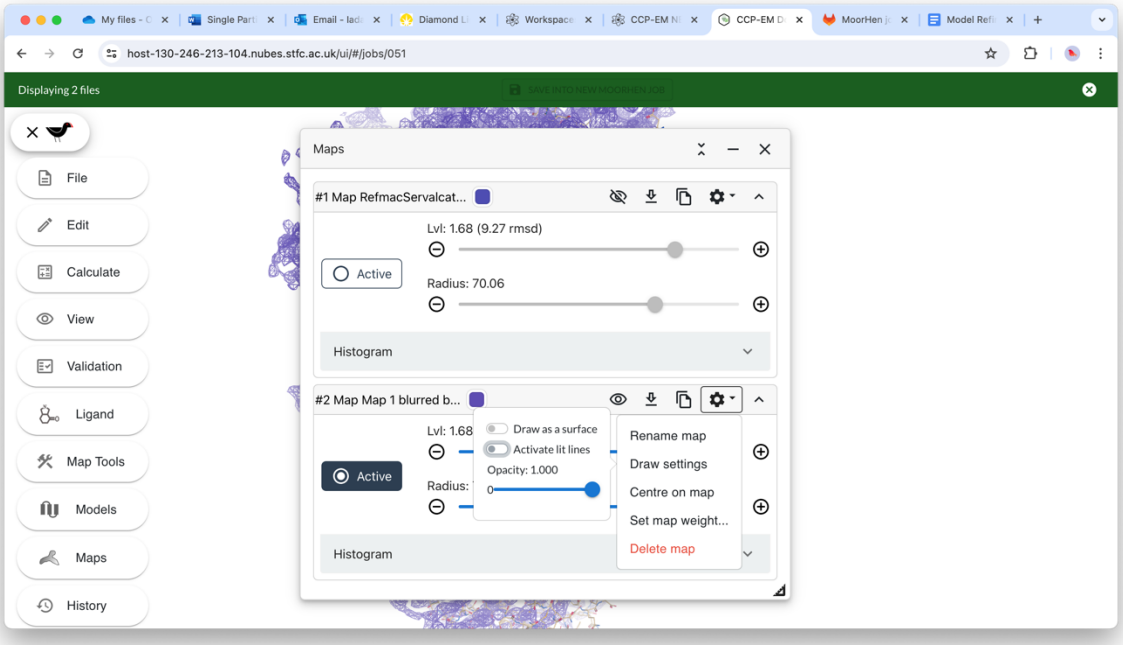
- **Map Tools → Sharpen/Blur map**
- "B-factor to apply" should be **`0`**
- Turn on **use resample** and use the resample factor **`1.4 → OK`**



Maps → Click on the "Gear" icon on the (newly-created) Masked map → Draw settings → Activate "**Lit lines**"

Make sure you choose the oversampled map for fixing model errors and real space refinement.

- **Maps** → Click on the "○ **`Active`**" button of the oversampled map. And hide the postprocessed map.



Analysing the model

Go to **`Validation -> Ramachandran plot...`**

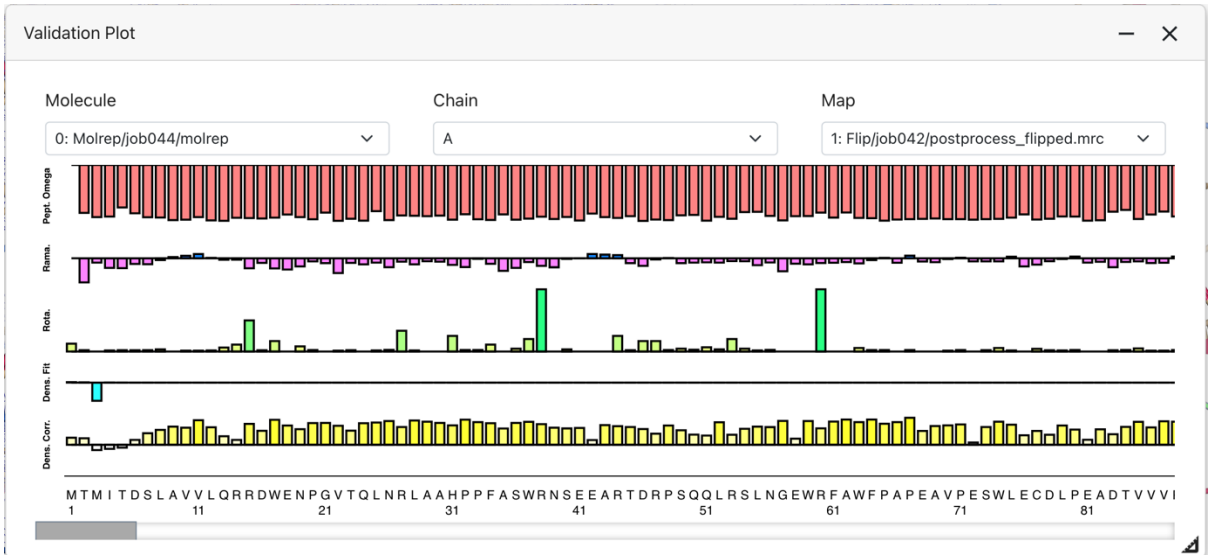
This displays the Ramachandran plot for the model. Right click on a point in the plot to centre the view on that residue.

Map-Model validation

How well does the AF2 model fit the experimental data? Does it fully explain the map density? To inspect this efficiently run:

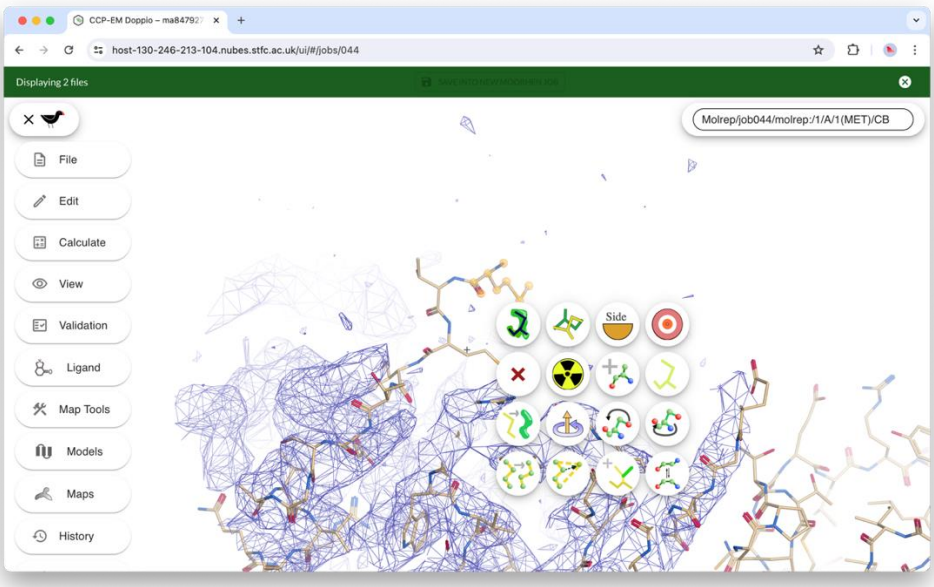
`Validation → Validation plot`

This will display a multiple validation plot. Look at the **`Density Fit`** plot (cyan).



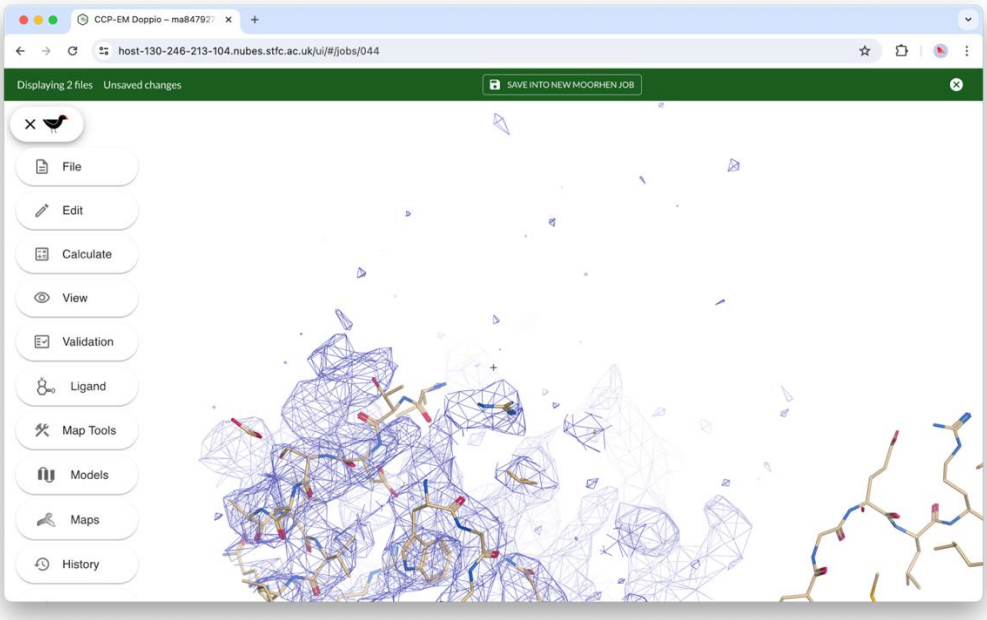
Something is wrong with Met[3]! Double click on the residue in the plot to go to this region of the protein.

There is no density for the three N-terminal amino acids Met[1]-Thr[2]-Met[3] of the alpha fold model in the map. Let's delete these three residues. Right click on Met[1] and select **`Delete residue`** (the red x).



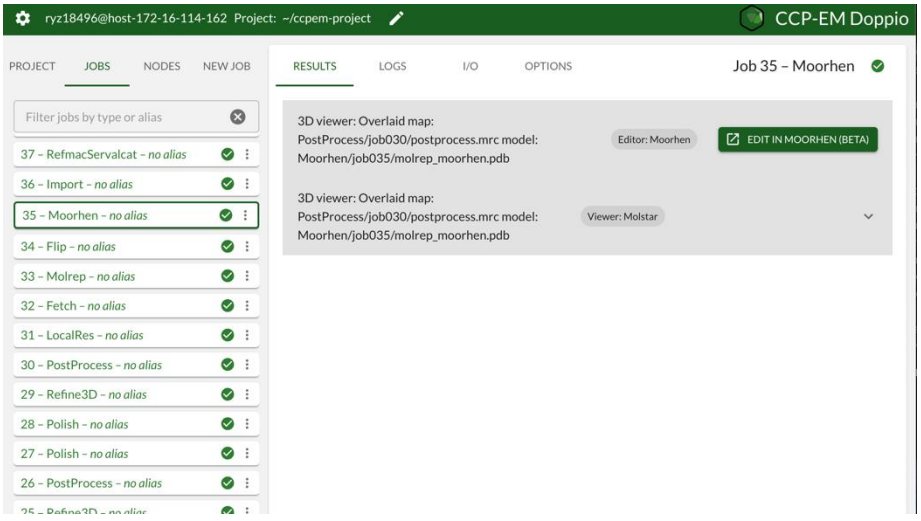
Now do this for the next two N-terminal residues Thr[2] and Met[3].

Next get a better fit for Ile[4]. Rotate the view so you can see the empty density where this residue goes.



Right Click on Ile[4] and select **`Refine Residue`** (the red bulls eye). Ile[4] and the adjacent residues should snap to fit the density.

Click the **`save as a new Moorhen job`** button to close the editor and save the model. The Moorhen session will be saved containing the changes to the model. The job will appear in your Doppio jobs list, and you can resume editing the Moorhen session from the **`results`** tab of that job. Any changes to the model will also be saved and new model node will be created for use in subsequent jobs of a different type e.g. Refmac-Servalcat.



Modelling a ligand

Click Edit in Moorhen to resume the Moorhen session we will now add a ligand to the structure.

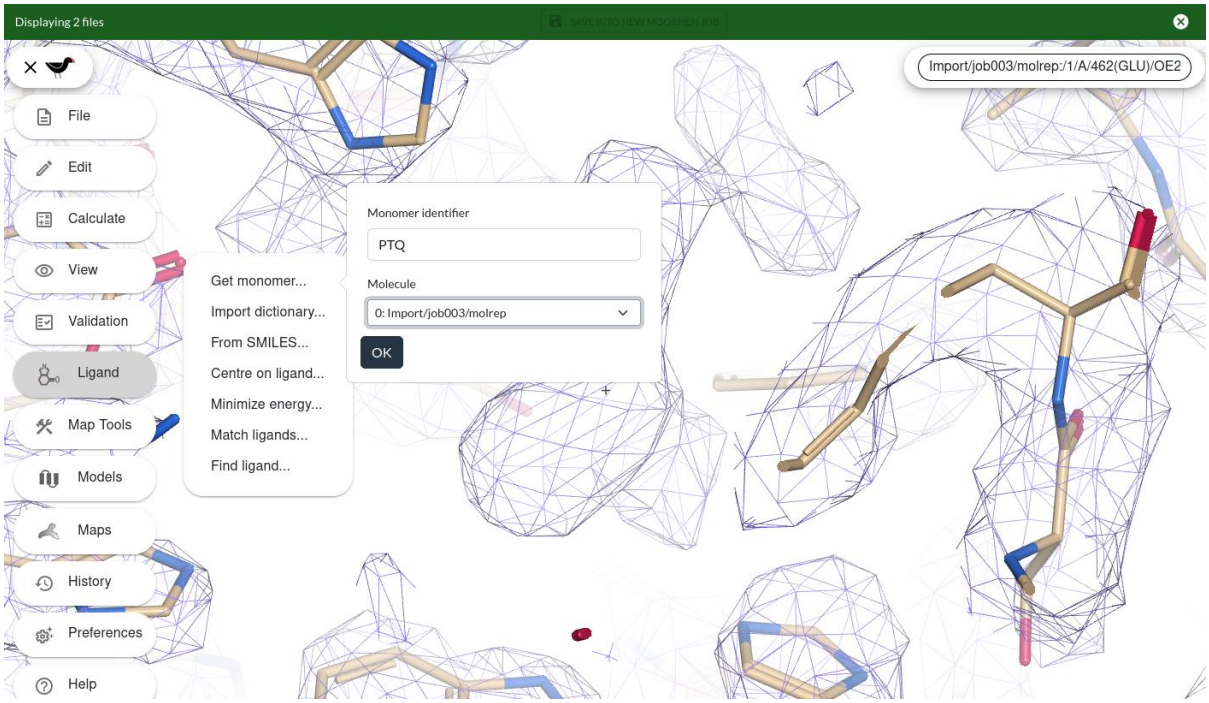
Modelling a known ligand

AF2 models do not contain waters or ligands. There should be a large unconnected blob of density between Phe^{602} and Trp^{569} . This ligand is modelled as inhibitor 2-phenylethyl 1-thio-beta-D-galactopyranoside in the original PDB:5a1a and its CCP4 Monomer library accession code is PTQ. Let's model this ligand in Moorhen.

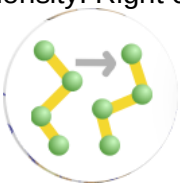
Go to residue Phe602 using **`Shift+G -> 602 -> Press Tick`** or **`Edit -> Go to`**. Use **`0/A/602`** to go to phe^{602}**.**

If the empty density between Phe^{602} and Trp^{569} is not in the centre of the screen centre the density using the or shift+alt (opt on mac) or holding middle mouse button. Use the central cross hair to guide you.

Insert the ligand using **`Ligand -> Get monomer with Monomer identifier: PTQ`**

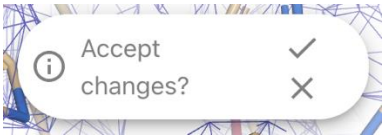


Now fit the ligand into the density. Right click on an atom in the ligand and select

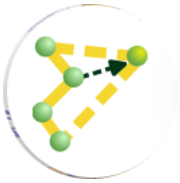


`rotate/translate zone`

Rotate (left mouse button) and Translate (middle mouse button) the ligand until you have roughly fit it in the density.

Make sure to click accept changes when finished. 

Next right click on an atom in the ligand and select

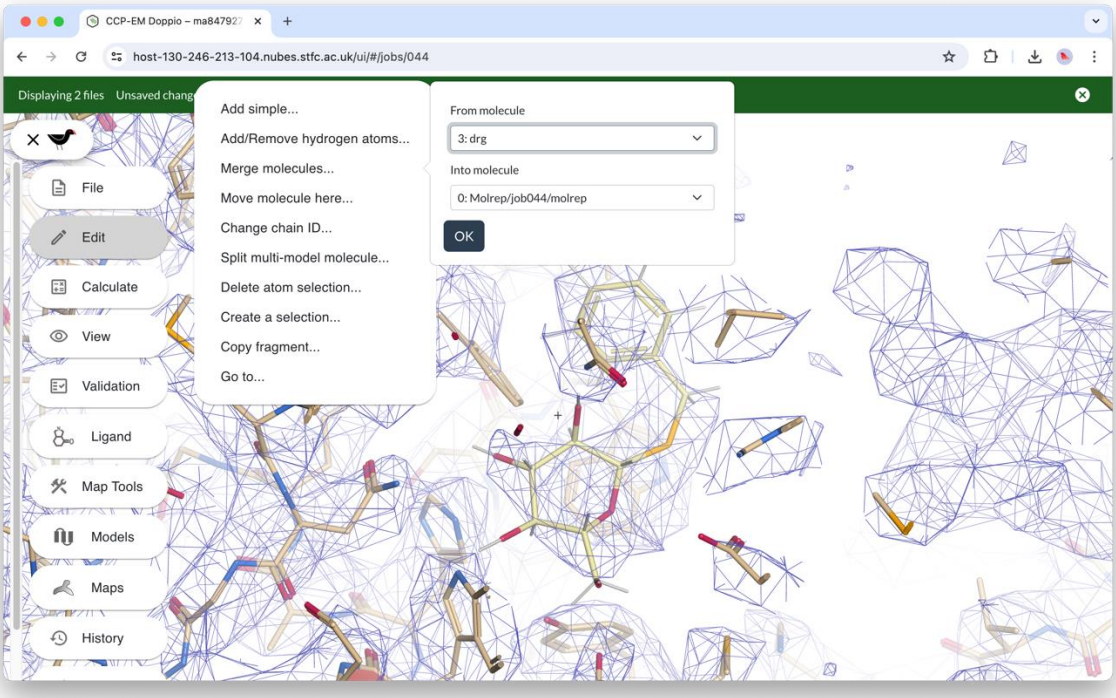


`drag atoms`.

Manipulate the ligand until you are happy with its fit. Again, remember to click **`accept changes`**

You can use Auto-fit Rotamer (**`right click menu -> black residue-green map icon`**) to improve the ft. N.B. the fit of the ligand does not need to be perfect; it will be refined in later steps.

Now merge the ligand model into the main model. Go to **`Edit -> Merge molecules...`**



Merge the ligand into the main model and then save the moorhen Job. If you are working on new or unusual ligand that is not in the monomer library, you will need to create a new structure and corresponding set of restraints using **AceDRG** in Doppio. See “Modelling a New Ligand” section at the end of this tutorial for further details.

Automated Refinement in Refmac-Servalcat Part 1

Here we will use Refmac-Servalcat to automatically refine the model against the experimental map and stereochemical restraints. Search (refmac) to find the **Refmac-Servalcat** job, create a new job and enter the following:

Input model: Moorhen/jobXXX/molrep_moorhen.pdb
Your Moorhen job number may vary depending on if the map was flipped.
Resolution: 2.8
Use the resolution of your final 3D refinement.
Input map 1 (half map 1 or…): Refine3D/jobXXX/run_half1_class001_unfil.mrc
Use the half map from your final 3D refinement. If the map had to be flipped use the flipped halfmaps from the flip job instead.
Half map refinement: Yes
We now always recommend using half maps as inputs if available. This ensures the input maps are unsharpened and unfiltered and Servalcat can generate automatically sharpened and weighted maps as per (Yamashita et al, 2021).
Input map 2 (half map 2): Refine3D/jobXXX/run_half2_class001_unfil.mrc

Input mask: MaskCreate/jobXXX/mask.mrc
Use the previously created mask. If the map was flipped make sure to use the flipped mask.
Masked Refinement: Yes
This will use the mask to cut-out a sub-volume of density around the given atomic model and only refine in this area. This speeds up refinement and means the fit-to-density statistics relate to this sub-volume not the whole map.

Leave the rest of the parameters as the defaults except:

Point group: D2
Servalcat will now enforce symmetry in the refinement i.e. it will refine against one copy of the protein and apply symmetry restraints. This is more efficient and enforces the symmetry which was applied in the 3D refinement of the map.
N.B. when using symmetry only one copy is allowed in the input model.

When the job has completed look at plots in the **RESULTS** tab.

FSC average should show improved fit-to-experimental-data as the average map-model FSC increases. You may also notice that the score is still improving and therefore refinement has not converged. Therefore, more cycles should run. The geometric plots also show improved agreement with less deviation from expected bond, angle and planarity values.

We can inspect the model in more detail in **Moorhen** or **Coot** but before then let's set some more **Refmac-Servalcat** jobs running.

Automated Refinement in Refmac-Servalcat Part 2

In the previous **Refmac-Servalcat** job we used a small number of cycles which means the structure may not have converged and more cycles may be required. To repeat a job we can use the clone function in the **JOB** tabs click the triple dot menu next to the completed **Refmac-Servalcat** job and select **Clone** job. This will start to make a new job with the previous job's parameters. Keep them the same except:

Refmac cycles: 24

Click **RUN** to start the job.

We run several jobs in parallel so let's try optimising the refinement weight. This affects the balance between the experimental restraints and the stereochemical restraints and should be optimised for each dataset. The first job was run with **Auto weight** enabled where Servalcat tries to optimise the weight itself.

To find the weight used go to the completed **Refmac-Servalcat** job, go to the **LOGS** tab and scroll to the end and you'll see something like:

Weight used: 3.36999989
If you want to change the weight, give larger (looser restraints) or smaller (tighter) value to --weight_auto_scale.

Clone the second Refmac-Servalcat job (the one with 24 cycles) and change the following parameters:
Auto weight: No
Weight: 33
Increasing this puts greater emphasis on the experimental data and with respect to the stereochemical data.

Also it's a good idea to add an alias to help differentiate jobs e.g.
Job alias: WT_33

Click **RUN** to start the job and then clone this job and start another changing the following:

Auto weight: No
Weight: 0.33
Reducing this puts less emphasis on the experimental data and with respect to the stereochemical data.
Job alias: WT_0–33

These jobs will take ~15mins to run. If you are waiting, please continue to the next section.

Once the jobs have finished compare fit to data (average FSC) and the stereochemistry (geometric rmsd values for bonds, angles and chirality). Which weight was optimal for this dataset?

Inspecting Refmac-Servalcat Outputs in Moorhen

From a completed Refmac-Servalcat job go to the I/O tab, select the following output nodes and launch a new Moorhen job:

- **refined.pdb**
  This is the refined model (single chain).
- **refined_expanded.pdb**
  This is the symmetry expanded refined model (four chains).
- **sharpened_weighted.mrc**
  This is the sharpened and weighted map created by Refmac-Servalcat using the variance in the two half maps to calculate variance as described in Yamashita et al 2021.

To compare the single chain and symmetry expanded model click Models (from the Moorhen menu) and toggle displaying the models on and off by clicking the eye icon for each model.

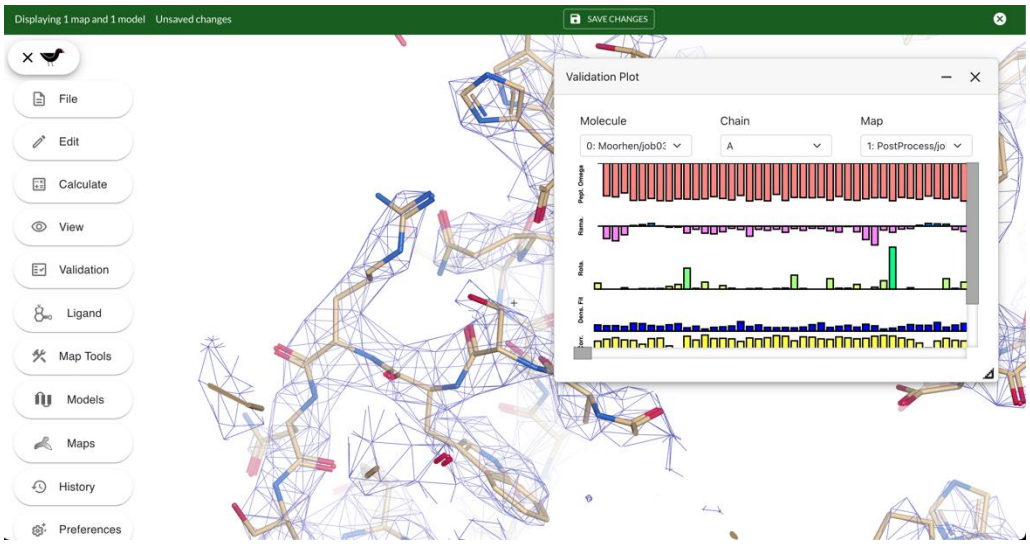
When making changes to the model below make sure you work on the single monomer copy (refined.pdb). Running Refmac-Servalcat with the edited model will apply changes to all symmetry related chains.

We can also find erroneous areas by using Coots validation analysis as before:
Validation → Validation plot

Using the validation plot look for any outstanding outliers that need to be inspected and manually refined. Some of the most obvious outliers are low probability rotamers e.g. Arg38, Arg60, Lys578 and Ser735 (N.B. depending on the previously steps you may see different outliers).

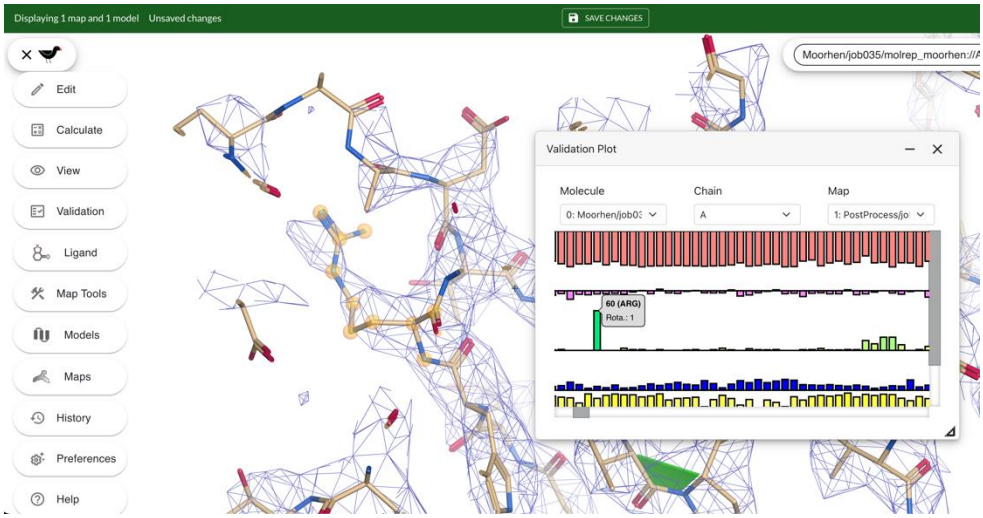
Clicking on the residue (e.g. Arg38 on the validation plot will centre the view on that residue. Right clicking on the residue itself will display the model tools panel:

Use the AutoFit rotamer (black residue-green map icon) and Refine Residues (red circles target icon) to improve the fit. (N.B. changing the map contour level You may need to change the map contour level to help

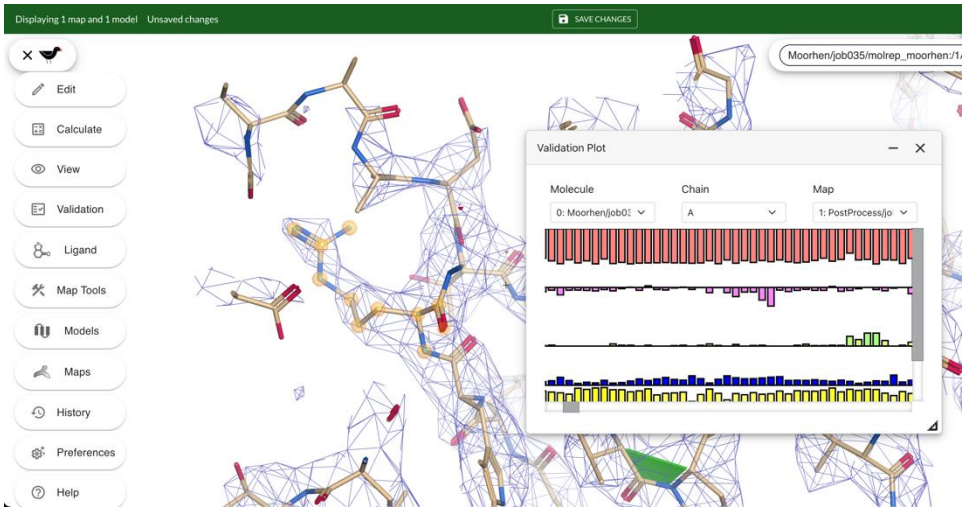


As you can see above the Arg fits the density better. It still has a relatively low probability rotamer but this is now supported by the experimental data.

For Arg60 we can repeat the process...



...and, in this case, both the fit to data and the rotamer probability improves.



Another interesting area to work on is the loop in the region spanning ~Ala^{726} to ~Ala^{737}. This is hard to model as the density is weaker here, most likely caused by local disorder in the flexible loop.
Use the modelling tools to improve the fit of the model in this region. For this region blurring the map removes high frequency noise and makes the main chain easier to model.

Once you have improved several areas of the model, save the coordinates and run another Refmac-Servalcat job to see if the statistics have improved and to generate a new symmetry expanded model. Use the clone job function and update the input model to the output from the corresponding Moorhen job.

Model refinement and validation is a lengthy process with multiple rounds of iteration between manual refinement in Coot and automated refinement in Refmac-Servalcat. Building models correctly is a time-consuming process but it is necessary to give you and any others who may use it in the future the best possible structure to work with.

Detailed model validation in Doppio

Further validation of models is available from the **`Atomic model validation`** job in Doppio (search model validation). This runs multiple validation tools including MolProbility, Refmac, TEMPy and CheckMySequence to extensively validate your model.

In the Atomic model validation job enter your current model and the following options and then click **`run`**:

`Input model: RefmacServalcat/jobXXXX/refined.pdb`
`Servalcat FSC (Model-map FSC): Yes`
`CheckMySequence (Sequence agreement): Yes`

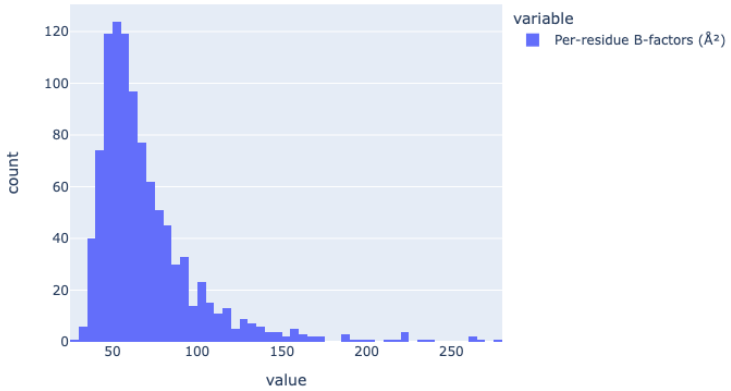
`Input map: PostProcess/jobXXX/postprocess.mrc`
It is best to use the unmasked version as metrics are calculated with the supplied mask.

`Resolution: 2.8`
`Input halfmap 1: Refine3D/jobXXX/run_half1_class001_unfil.mrc`
`Input halfmap 2: Refine3D/jobXXX/run_half2_class001_unfil.mrc`
`Input map mask: MaskCreate/jobXXX/mask.mrc`
`Input map mask: MaskCreate/jobXXX/mask.mrc`
`Input sample sequence: Fetch/jobXXX/P00722.fasta`

Analysing the results

When the job is complete click the **`results`** tab.

B-factor distribution



Typical B-factor distribution showing positive skew without excessive low values.

We expect a positively skewed distribution (i.e. greater number of low B-factors) with some large B-factor outliers which either represent ill-modelled resides or highly mobile regions. A large number of very low B-factors and/or B-factors close to 0^2 Å demonstrate an over-sharpened map however a non-skewed distribution (guassian or normal distribution) or a negatively skewed distribution shows an over-blurred map.

Global map-fit scores

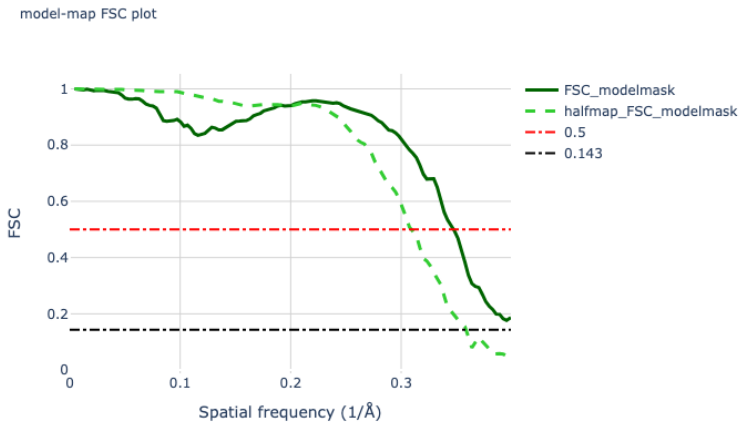
Global map-fit scores

Metric	Score
FSCavg_modelmask_FSC0.5	0.827
FSCavg_modelmask	0.631
FSCavg_h1_modelmask	0.592
FSCavg_h2_modelmask	0.596
CCC overlap mask	0.709
Model overlap fraction	0.699
CCC contoured map	0.311
MI overlap mask	0.015

It is difficult to give precise expected values for these metrics due to the lower number of structures, range of resolutions, etc. however we can provide a guide what the minimum expected values should be. During the refinement process you see these improve and make best efforts maximise these scores to get the best model possible to explain your data.

FSCavg_modelmask_FSC0.5 shows agreement of the whole map with the model upto the 0.5 cutoff and should be >0.7.
FSCavg_modelmask is the same calculation as FSC0.5 but includes higher frequency terms and therefore will be slight lower.
CCC overlap mask (cross-correlation within the model-map region as defined by the mask). This should be >0.6.

Map-model FSC



The FSC_modelmask is the FSC between the model and the full map, the halfmap_FSC_modelmask is the FSC between the two half maps. We expect the resolution (spatial frequency) of the FSC_modelmask at 0.5 cutoff to be similar to the halfmap_FSC_modelmask at 0.143 cutoff. If the 0.5 cutoff is significantly worse this shows the model is poorly fitted.

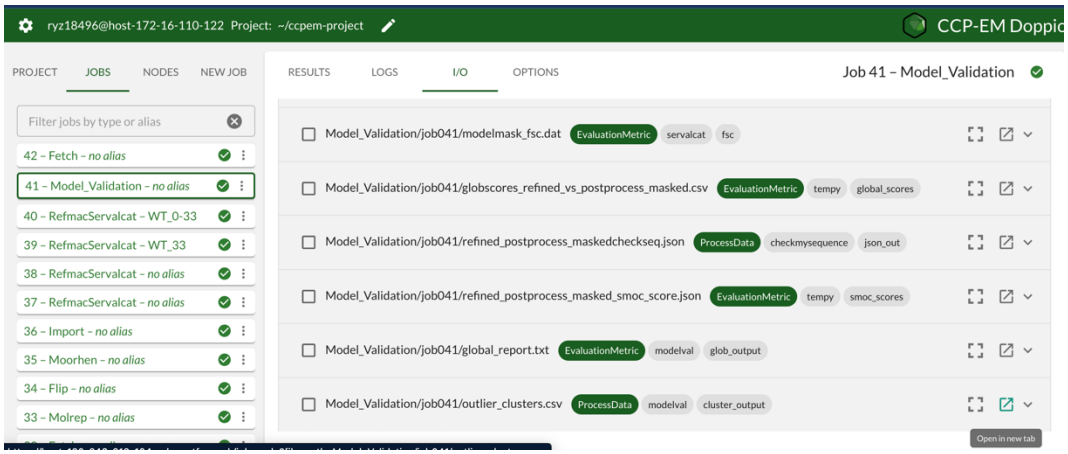
Outlier clusters

To help identify areas that need manual inspection outliers from the various metrics are spacially clustered and grouped together in the outlier cluster table.

Outlier clusters

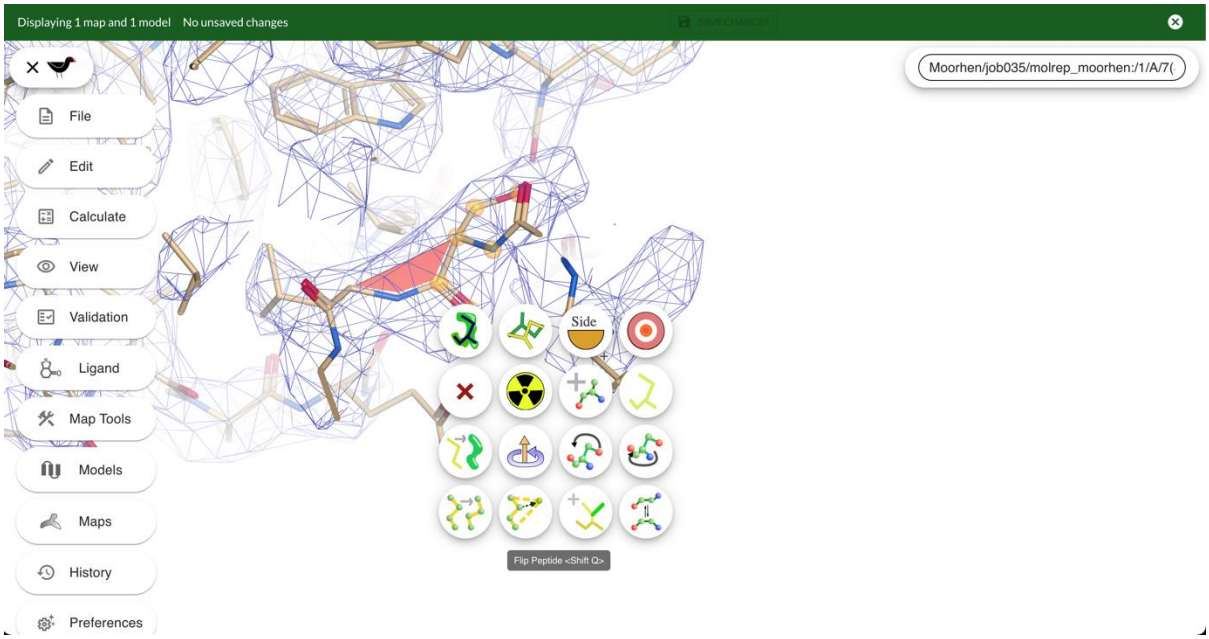
Cluster	Chain_Residue	Outlier Type(s)
1	A_7	ramachandran_outlier | rotamer_outlier | bad_clash (atom:O...
1	A_8	omega_outlier | bond_angles (1_A_7_C-1_A_8_N-1_A_8_CA) | ...
1	A_6	bad_clash (atom:O - A_7_OG) | bad_clash (atom:C - A_7_OG) | ...
1	A_5	bond_angles (1_A_5_CA-1_A_5_CB-1_A_5_OG1)
2	A_462	ramachandran_outlier
2	A_423	omega_outlier | bad_clash (atom:HG3 - D_280_HD11)
2	A_424	bad_clash (atom:HE3 - D_283_HG2) | bad_clash (atom:N - D_...
2	A_463	bad_clash (atom:HB3 - A_428_HG21)
3	A_149	omega_outlier | bad_clash (atom:HA - A_148_HA)
3	A_148	bad_clash (atom:HA - A_149_HA)
3	A_150	smoc (raw:0.704;Z:-2.412)

This is ordered by the size of the cluster (number of residues). We fine the most effective way of inspecting and fixed these errors is to download the outlier_clusters.cvs file by clicking **`open in a new tab icon`** from the **`I/O`** panel:

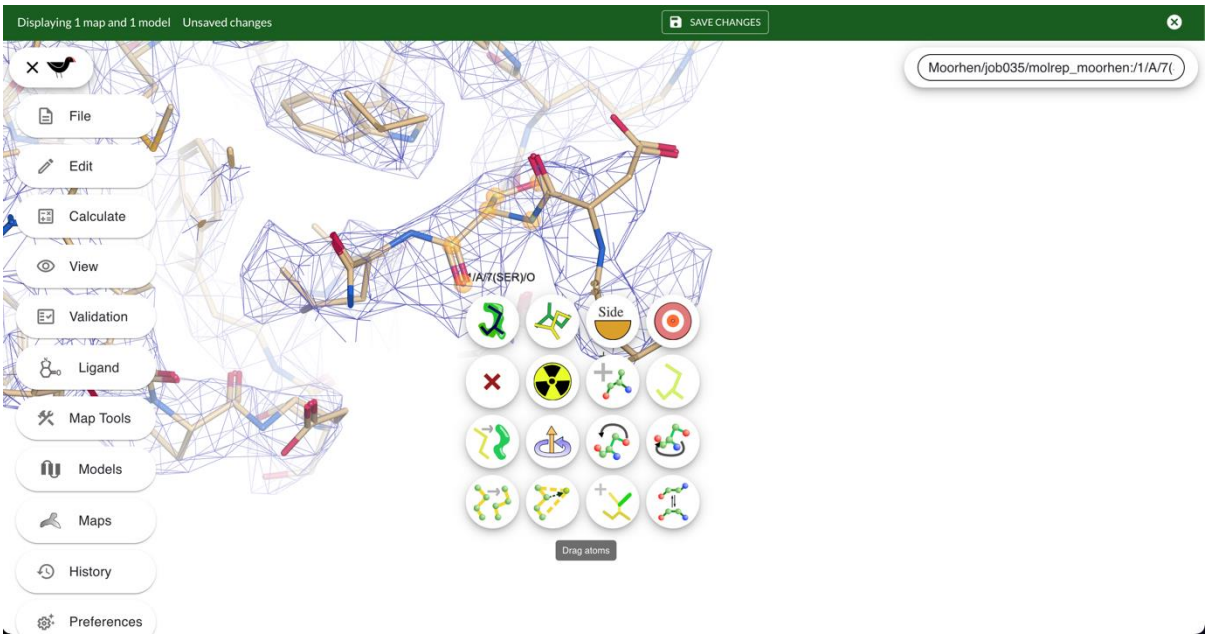


You can view this in a spreadsheet or text editor whilst then running Moorhen to inspect these regions.

In this case Cluster A_5-7 reveals an incorrectly modelled cis-peptide (illustrated by red triangle).



use Flip Peptide and Refine Residues to fix this area.



You may also need to use Drag atoms, followed Refine Region to place the backbone oxygen into density.

Work through the other clusters to make your model as best as possible and then re-run **`Refmac-Servalcat.`**

Modelling a novel ligand

If the ligand is new or it does not exist in the CCP4 Monomer library, then it necessary to create a description for the new ligand. This is done using the program AceDRG. For the

purpose of the tutorial let's pretend that PTQ is a new ligand, and we give it a generic name LIG.

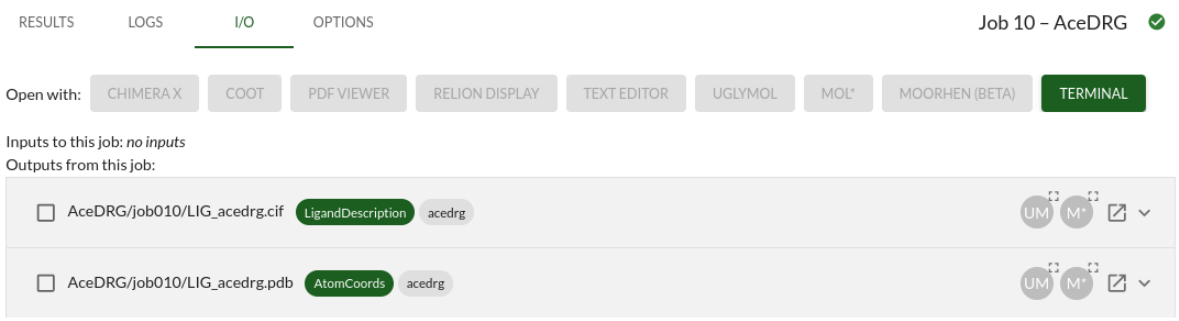
Select **AceDRG** job from the **Create Ligand** category. Create the restraints file with the same SMILES string as before:

SMILES String: c1ccc(cc1)CCS[C@H]2[C@@H]([C@H]([C@H]([C@H](O2)CO)O)O)O

Monomer code: LIG
This is just an arbitrary name for the ligand.

Press **run** to create the ligand restraints file. We must now download the ligand description to the local machine. Go to the job's **I/O** tab.

N.B. AceDRG generates two outputs: atomic coordinates (pdb) and corresponding set of stereochemical restraints (cif).

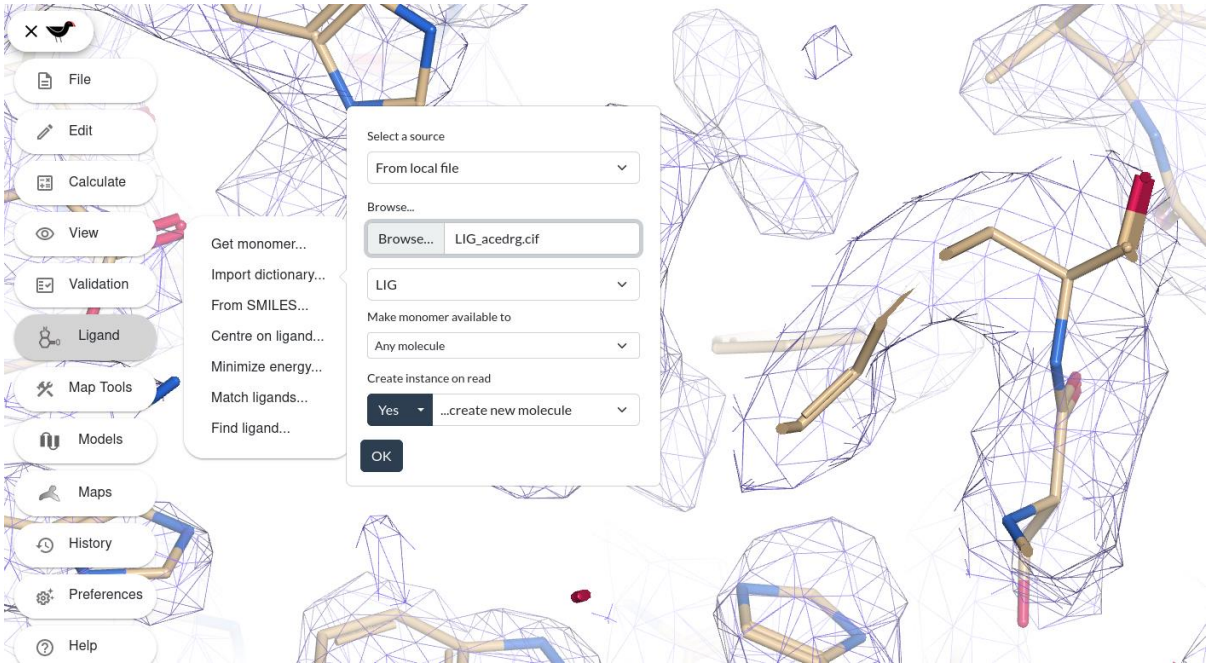


Click on the **open in new window** icon  for **AceDRG/jobxxx/LIG_acedrg.cif**. This will download the file onto your local machine.

Go back to the **I/O** your Molrep job and launch a new Moorhen job with the postprocessed map and AF model. a
Inputs: PostProcess/jobXXX/postprocess.mrc
If you had to flip you map then this will be the flipped map instead.

Outputs: Molrep/jobXXX/molrep.cif

Now to insert the ligand into the map **Ligand -> Import dictionary...**



Once the ligand is placed in the density, fit, refine and merge the ligand into the model as described above.

When running **`Refmac-Servalcat`** with a new ligand use the merged model from Moorhen as well as also providing the corresponding stereochemical restraints:

`Input model: Moorhen/jobXXX/molrep_moorhen.pdb`
`Input ligands: AceDRG/jobxxx/LIG_acedrg.cif`

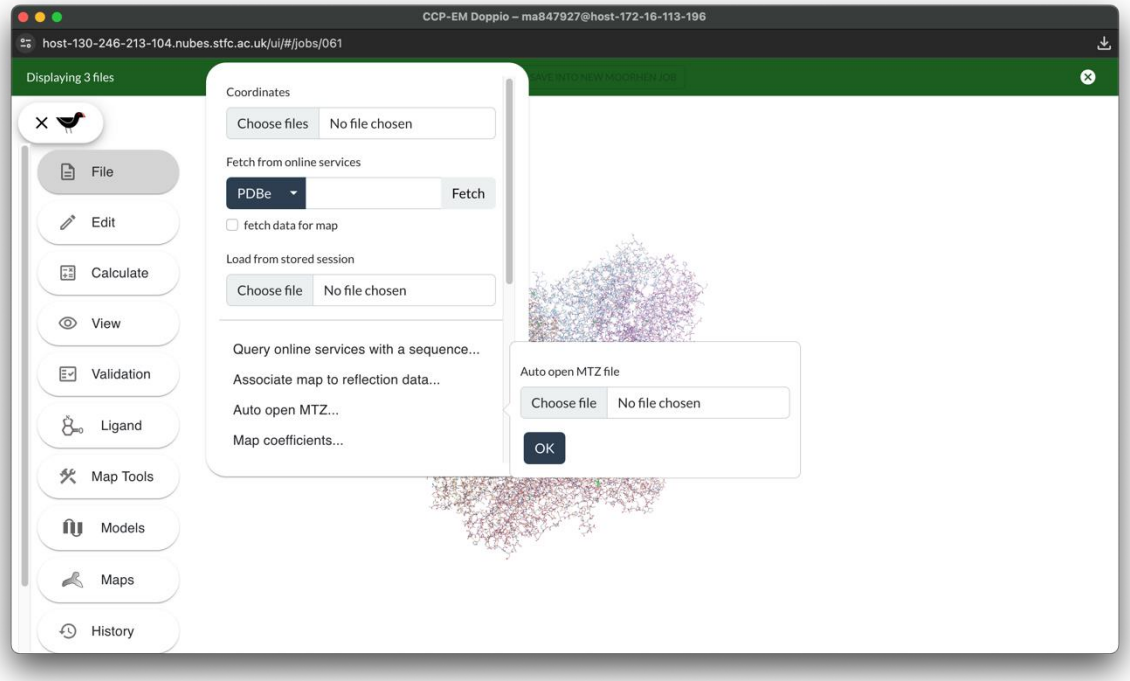
Beta: Difference map

Please note: The Moorhen job is still in beta and currently cannot handle loading a mtz difference map from the GUI so we must download this file to the local machine first. This will be fully integrated in a later version of Doppio and Moorhen but can be useful especially when working with ligands.

In the **`I/O`** tab of the RefmacServalcat job download **`RefmacServalcat/jobxxx/diffmap.mtz`**

by clicking on the open in new window icon 

The difference map is useful for finding areas of the model that do not agree with the experimental map. Let's load it from the file we just downloaded. Go to **`File -> Auto open MTZ...`** to select the file and open it.

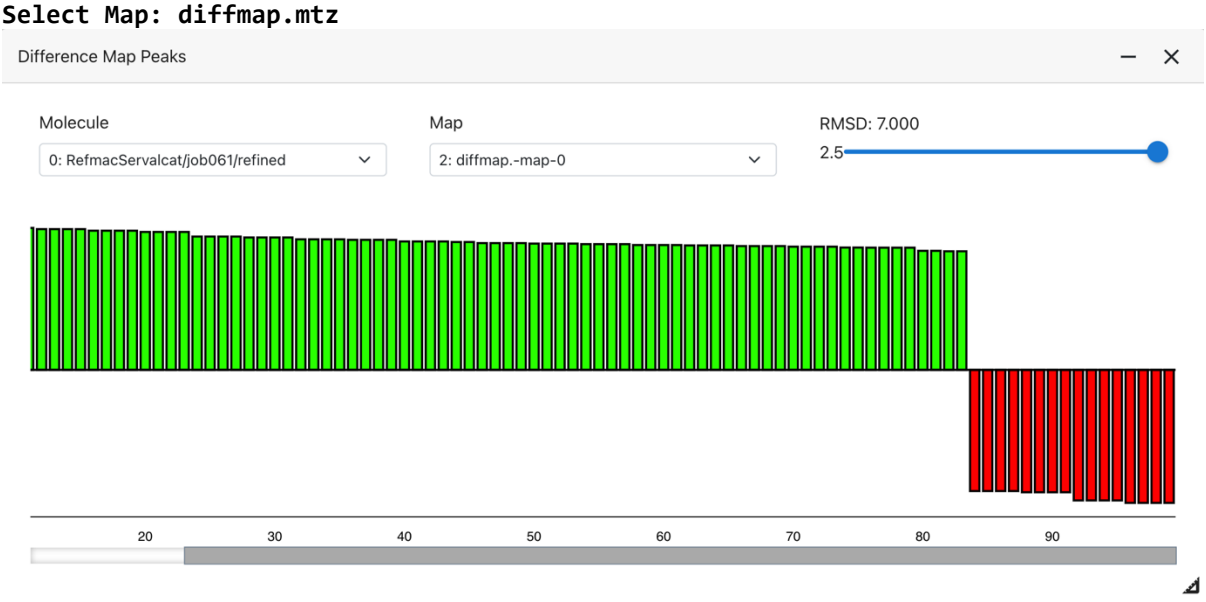


Now use the validate tool to search for such discrepancies:

Validate → Difference Map Peaks

Select Model: RefmacServalCat/jobxxx/refined.pdb

Select Map: diffmap.mtz



Click through the peaks and use Real Space Refine, Regularize Zone and other tools to improve the model. Examples of areas that may need investigating include the N-terminal, Arg^{882}, Arg^{600}, Arg^{443}, Phe^{602}, etc. Also use the difference map to check the fit of the PTQ ligands.

Relion References:

Rosenthal P and Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. Journal of Molecular Biology,333(4):721–745, October 2003.

Scheres SHW. Classification of Structural Heterogeneity by Maximum-Likelihood Methods. In Cryo-EM, Part B: 3-D Reconstruction, volume 482 of Methods in Enzymology, pages 295–320. 2010.

Scheres SHW. A Bayesian view on cryo-EM structure determination. Journal of Molecular Biology, 415(2):406–418, January 2012. doi:10.1016/j.jmb.2011.11.010.

Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology, 180(3):519–530, December 2012. doi:10.1016/j.jsb.2012.09.006.

Scheres SHW and Chen S. Prevention of overfitting in cryo-EM structure determination. Nature methods, 9(9):853–854, September 2012. doi:10.1038/nmeth.2115.

Shaoxia Chen, Greg McMullan, Abdul R. Faruqi, Garib N. Murshudov, Judith M. Short, Sjors H. W. Scheres, and Richard Henderson. High-resolution noise substitution to measure overfitting and validate resolution in 3d structure determination by single particle electron cryomicroscopy. Ultramicroscopy, 135:24–35, December 2013. doi:10.1016/j.ultramic.2013.06.004

Refmac-Servalcat references:

Yamashita, K, Palmer, C M, Burnley, T, Murshudov, G. N. Cryo-EM single particle structure refinement and map calculation using Servalcat. Acta Cryst D77, 1282-129, 2021.

Current approaches for the fitting and refinement of atomic models into cryo-EM maps using CCP-EM. Nicholls, RA, Tykac M, Kovalevskiy, O, & Murshudov, GN Acta Cryst D74, 492-505, 2018.

Moorhen and Coot references:

https://github.com/moorhen-coot/Moorhen
Casañal, A, Lohkamp, B and Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. Protein Sci. 29(4): 1069–1078. 2020.

CCP-EM, Doppio/ccpem-pipeliner references:

Burnley, T, Palmer, CM & Winn M. Recent developments in the CCP-EM software suite. Acta Cryst D73, 469-47, 2017.
https://gitlab.com/ccpem/ccpem-pipeliner