Model Building and Refinement Practical: CD44

Ed Lowe (modified)

Introduction

In this practical we will continue working with the CD44 experimental phases we determined in the MAD/SAD phasing practical. We will begin by inspecting a CD44 model which has been automatically built by the program Buccaneer.

1. **Inspect Output from Automatic Model Building for Errors and Make Corrections**

    1.1.1. From “File -> Open Coordinates” open the file cd44_buccaneer.pdb

    1.1.2. From “File > Auto Open MTZ...” open the file cd44_buccaneer.mtz

    1.1.3. Use the scroll wheel of the mouse to change the contour level of the electron density map. Scroll to a value near 1.0 rmsd (this value is displayed in the top right corner of the graphics window).

    1.1.4. Click on the “Map” button in the top right-hand corner of the graphics window and select the “FWT PHWT” map for use in refinement.

    1.1.5.1.1.5 Click on the “R/RC” button in the top right-hand corner of the graphics window to open the Refinement and Regularization control panel. Under “Weight Matrix”, set the Refinement Weight to 20.

1.2. Automatic model building will only rarely produce a model which is both complete and correct. It is helpful to compare the known sequence of CD44 with the model

    1.2.1. Select “Validate > Alignment vs. PIR...” and choose cd44_buccaneer.pdb as the model. Then choose to link chain A and the file cd44.seq (you may need to browse to /home/crystal/cd44). A “Residue Mismatches” panel showing residues present in the sequence file but different or absent from the current model will be generated.

    1.2.2. From the “Residue Mismatches” panel, select “Mutate A 2 UNK to Ala”

        1.2.2.1. Inspect the map at this point. Since Buccaneer has built an Ala residue for any unknown residues not docked with the protein sequence during model building (and marked them UNK) there is no need to add any atoms at this point. We can use the “Simple Mutate” tool (the  icon on the toolbar) to tell coot that this residue is in fact an Ala.

    1.2.3. From the “Residue Mismatches” panel, select “Mutate A 21 UNK to Asn”

        1.2.3.1. Inspect the map at this point. It should be quite clear where the side chain of Asn 21 should be placed. Use the “Mutate & AutoFit” tool () to mutate UNK 21 to Asn.

1.2.4. The small loop from A 22 to A 25 has proved more difficult to build automatically (it looks like auto-tracing has followed a side-chain rather than the main-chain at one point). Fortunately, the map in this region look quite good so we will try to complete this region of the model using the loop fitting tools in coot.

1.2.4.1. Use the “Delete Item...” tool () to remove the range of residues from A 22 to A 25. You can do this either by deleting one residue at a time by deleting “Residue/Monomer” or remove them all at once using “Delete Zone” and clicking on the first and last residues to be deleted.

1.2.4.2. Check the fit of residue Asn 21, paying special attention to the position of its carbonyl oxygen.

1.2.4.3. From the “Calculate” menu, select “Fit Loop... > Fit Loop by Rama Search…”. Make sure that molecule “cd44_buccaneer.pdb” and chain “A” are selected. Enter residue numbers 22 and 25 as the beginning and end of the region to be built. The sequence for the loop to be built is “GRYS”. Click the “Fit Loop” button and watch as coot build your loop for you.

1.2.4.4. The model now looks like it fits the map a great deal better, but there is still some room for improvement so we will carry interactively refine the region we have just built. Use the “Real Space Refine Zone” () tool and click on residues immediately before and after the loop we have just built - residues Asn 21 and Ile 26 would be good ones to select.

1.2.4.5. At this point a putative refined model will be displayed with carbon atoms shown in white and a panel of ‘traffic light’ indicators will appear indicating the quality of model geometry in the refined region. Are you happy with the position the putative refined residues have adopted? Do the traffic light indicators all show green, signifying good model geometry? If so, accept the refinement and continue. If not, try to improve the model, perhaps with help from a demonstrator.

1.2.5. From the “Residue Mismatches” panel, select “Mutate A 93 UNK to THR”

1.2.5.1. Inspect the map at this point. To me, it looks like residue A 93 has been somewhat misplaced but residues A 94-96 are placed well in the map. We can trust their placement in this density sufficiently to trust that there are no residues incorrectly missing or added and we can therefore extrapolate the sequence back from A 97 Asp.

1.2.5.2. Use the “Delete” tool to remove residue A 93.

1.2.5.3. Coot allows us to mutate a range of residues at once, and can attempt to fit the newly placed sidechains in density automatically. From the “Calculate” menu, select “Mutate Residue Range”.

1.2.5.4. Make sure that molecule “cd44_buccaneer.pdb” and chain “A” are selected. Enter 94 and 96 as the beginning and end of the range and the sequence “SQY”. Tick the box to Autofit mutated residues. Click the “Mutate” button. Once again, this could use a little improvement so use the “Real Space Refine Zone” tool on the range you have just mutated. It is usually a good idea to extend the refinement one or two residues beyond the region you have just built, so in this case residues 94 and 97 would be good beginning and ending points.

1.2.5.5. We are left with a model where residues A 92 and A 93 have not been built. Looking at the map in this region it is very difficult to see where the main chain should be traced, so we are better off leaving the residues absent. With luck, future refinement will improve the map sufficiently to allow these residues to be built.

1.2.6. From the “Residue Mismatches” panel, select “Insert A 152”

1.2.6.1. Inspect the map at this point. There does not seem to be any density to support adding any more residues to the C-terminus of the current model.

1.2.7. Use the “Go To Atom” tool () to return to residue 2 of chain A.

1.2.7.1. You can now progress along the polypeptide chain by pressing “space” to move to the next residue (N to C) and “shift+space” to return to the previous residue (C to N).

1.2.7.2. Work your way along the polypeptide backbone inspecting the fit between the model and the electron density map. You may attempt to correct any errors in the model using the “Auto Fit Rotamer” ()and “Realspace Refine Zone” tools.

1.2.7.3. When you reach residue Asp 5 you will notice that the model fits the map very poorly. Use the Auto-Rotamer tool to correct the orientation of Asp 5.

1.2.7.4. When you reach residue Arg 11 you will notice that the model fits the map poorly. Try using the “Auto Fit Rotamer” and “Real Space Refine Zone” to fix this error. You will notice that the resultant model is still a rather poor fit with the electron density map.

1.2.7.5. Fortunately it is possible to intervene manually in cases such as this where the refinement has become trapped in a false minimum. Use the mouse to drag the refined model (the one with the carbon atoms displayed in white) into the electron density. You should find that the Arg sidechain will snap neatly into the map once you have dragged it in the right direction. As long as you are happy with the geometry of this refined model, accept the refinement. NB. *It is also possible to drag individual atoms by dragging with Ctrl+left mouse button. In some cases this can be very helpful.*

1.2.7.6. His 17 has been built without a side chain, but there is clear electron density present, so you can make use of the “Mutate & AutoFit” tool to place the sidechain.

1.2.7.7. Continue to work your way around chain A, fixing errors where you find them. You may find at least one error that the flip peptide () tool will help you fix.

1.2.7.8. When you have reached the end of chain A (or when the demonstrators tell you that you have used enough time on this part of the practical) continue to section 2.2.8.

1.2.8. Buccaneer did not do quite so well building the second copy of cd44 and has split it into two separate chains (B and C). Chain B contains most of the model, consisting of residues 22-151. Although they may contain some small differences, at this early stage of refinement it is reasonable to assume that the two chains are at least similar to each other, so we can use coot to copy our edited and improved chain A to provide a good approximation of the second molecule.

1.2.8.1. First we want to remove chain C, since it will be in the way.

1.2.8.2. Open the “Display Manager” and change the display of your model from “Bonds (Colour by Atom)” to “C-alphas/Backbone”.

1.2.8.3. Zoom out by dragging upwards with the right mouse button until you can see all of chain C (it consists of two beta-strands joined by a hairpin). Shit-left clicking on a residue will identify it, so you can make sure that you have correctly identified chain C.

1.2.8.4. Open the “Delete Item...” tool, select “Delete Zone” and click on both ends of chain C.

1.2.8.5. Now you can copy chain A onto chain B. From the “Extensions” menu, select “NCS > Copy NCS Chain...”

1.2.8.6. Make sure that molecule “cd44_buccaneer.pdb” and chain “A” are selected. Click “OK”.

1.2.8.7. Now refine that model with Refmac.

2. **More, After Refinement**

2.1.1. Open “Validate > Difference Map Peaks” The correct map and model should be selected by default. The default sigma level (5.0) is also sensible, so click “Find Peaks”. A list of peaks will then be generated - work your way down them, correcting problems as you find them. Don’t worry about adding solvent molecules at this point - we’ll cover that in section 3.3

2.1.2. Open “Validate > Ramachandran plot” and select the current model. An interactive Ramachandran plot will be displayed, with any outliers shown

in red. Click on any outliers you find - are there any problems with the model that you can fix? A hint - the flip peptide () tool may be useful.

2.1.3. Open “Validate > Geometry Analysis” and select the current model. A histogram plotting geometry for each chain in the model will be displayed, with small green bars indicating good geometry and large orange/red bars suggesting problems that need to be fixed. Are there any problems in need of attention?

2.2. Whilst examining the structure, you will have noticed that there are numerous peaks in the electron density map that would be well explained by ordered water molecules. Coot contains tools to help you add waters to the model.

2.2.1. Open “Calculate > Other Modelling Tools > Find Waters”

2.2.2. The default values in the find waters dialog are acceptable for our purposes, so click “find Waters” to proceed.

2.2.3. The new waters have been added to chain D, and it is important to inspect them and make sure that they have all been added in appropriate places. Use the “Go To Atom” tool and select residue 1 in chain D.

2.2.4. Does this water molecule look correct? If so, hit the ‘space’ bar to move onto the next water. If not, either move or delete the water molecule as required.

2.3. Upon correcting as many errors with the model as possible, you would once again save your coordinate model and use it as input for a further round of refinement, repeating successive rounds of rebuilding and refinement until no further errors need to be corrected. If you have sufficient time, carry out another round of refinement as described in sections 2.3 to 3.4. Have the R and Rfree values improved relative to those observed in 3.4.3 as a result of your editing?