

MapQ Tutorial

Last Updated: Aug. 9, 2019

Greg Pintilie

gregp@slac.stanford.edu

Overview

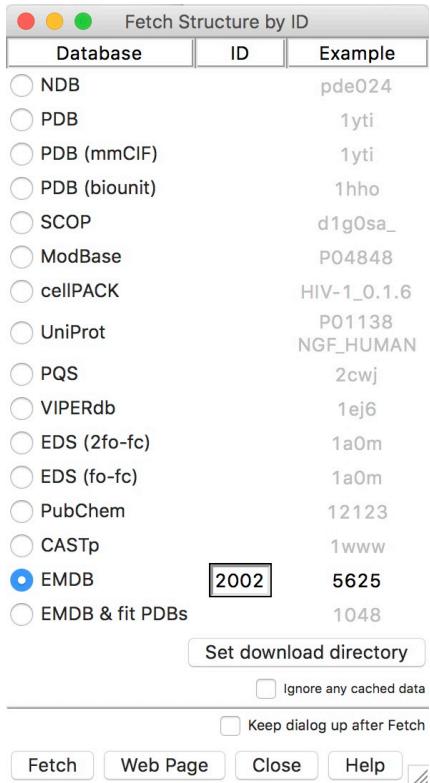
- MapQ is a UCSF Chimera plugin that allows calculation and visualization of Q-scores, which measure resolvability of atomic features in cryoEM maps using a fitted model.
- This tutorial covers the following topics:
 0. Get the software (UCSF Chimera and the MapQ plugin)
 1. Download cryoEM map and X-ray model
 - The model is not initially fitted to the map, so we will cover segmenting the map and rigidly fitting the Xray model.
 - Feel free to skip this part and go to the next section if you already have a map and model that is fitted to it.
 2. Use the MapQ interface to calculate and visualize Q-scores for amino acid residues and solvent atoms.

0. Getting the Software

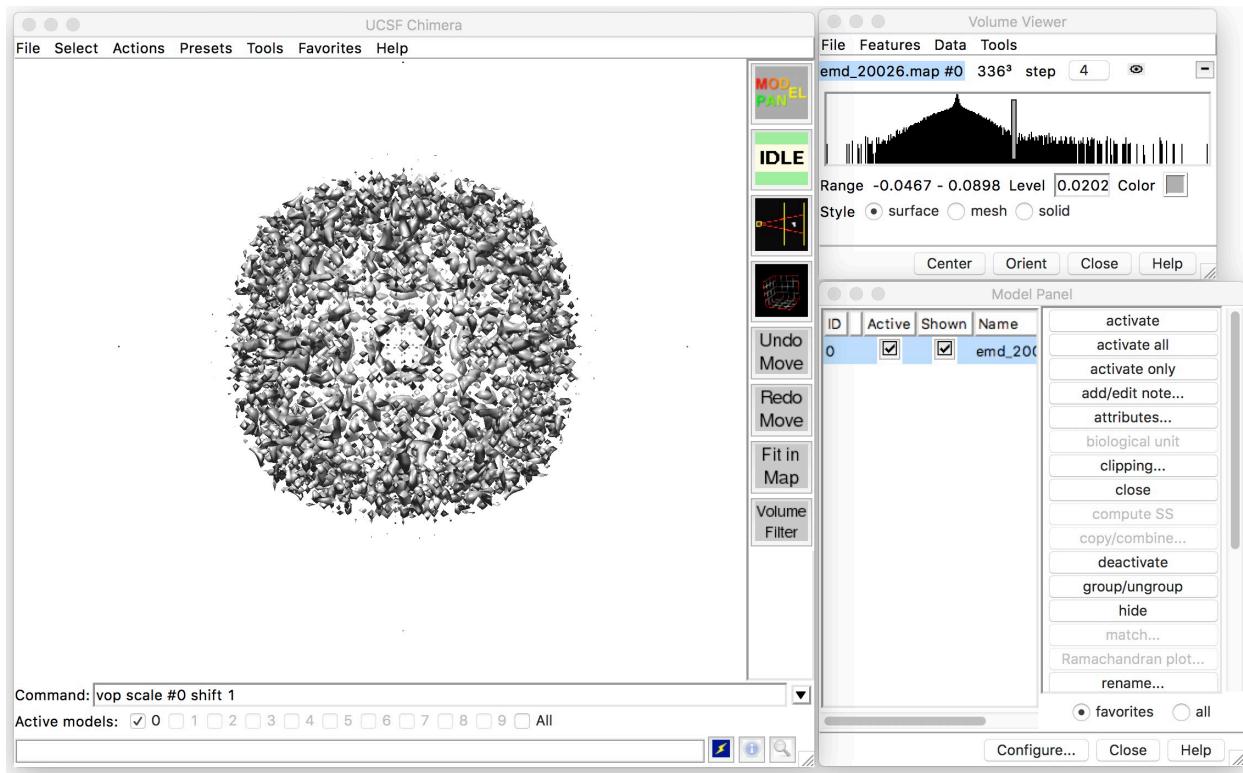
- Download Chimera for your system from
 - <https://www.cgl.ucsf.edu/chimera/download.html>
 - Either production or daily release should be ok
- Download and install the MapQ plugin
 - <https://github.com/gregdp/mapq>
 - Follow the Install instructions at the above link

1a. Loading and fitting the model

- Open Chimera
- Load the file, in one of two ways:
 - From File menu, select Fetch by ID...
 - Select EMDB, enter 20026, click ‘Fetch’ button at the bottom of the dialog

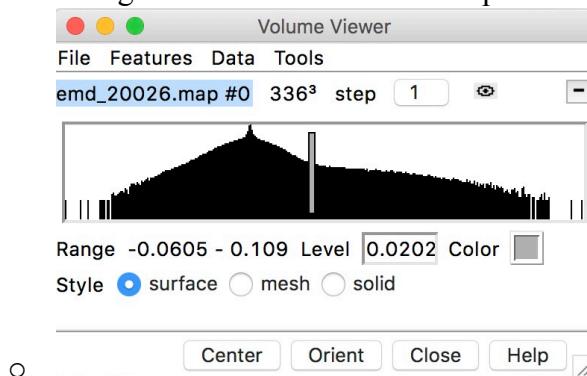


- Go to:
 - <https://www.emdataresource.org/EMD-20026>
- Select Download tab
- Click on 'emd_20026.map.gz' to the right of 'Map'
- After it downloads, double click or uncompress it in a terminal with the command
 - gunzip emd_20026.map.gz
- In Chimera, from File menu, select Open, navigate and select the map, click Open
- You should see something like this in the Chimera window, and two dialogs should pop up, 'Model Panel, and 'Volume Viewer'
 - If these dialogs go away, they can be opened again from the top menu:
 - Tools -> General Controls -> Model Panel
 - Tools -> Volume Data -> Volume Viewer



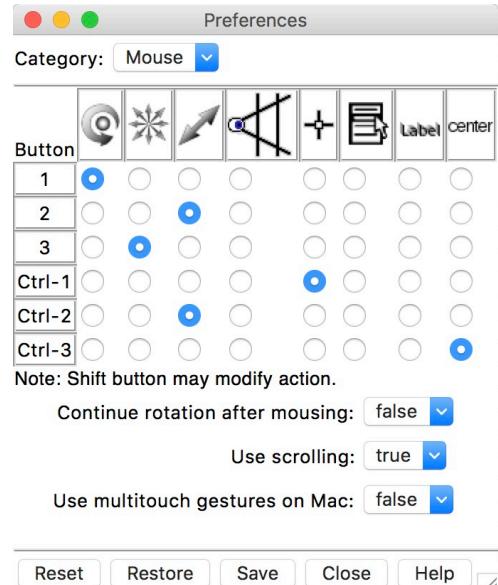
- Volume Viewer

- This shows the Cryo-EM maps that are open (emd_20026.map), maps size (336^3), and ‘step’ (4).
- For efficiency, the ‘step’ is 4 initially, however this shows less detail. Select 1 to see the full detail.
- The map is shown in the main window (titled UCSF Chimera) as an ‘Iso-surface’, i.e. a surface that has the same map value at all the points on it. The range and histogram of the values in the map are shown in the Volume Viewer dialog:

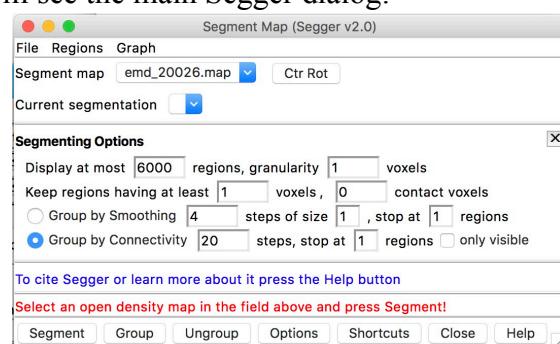


- The range is -0.0605 to 0.109. The histogram shows how many grid points (relatively) have a map value between these two endpoints. The gray bar on the histogram represents the map value on the surface, also commonly called the ‘threshold’ or ‘contour level’.
- Drag the bar on the histogram from left to right (with the mouse left button), and see what happens to the surface shown in the main window.

- You can click the ‘eye’ icon to show/hide the map in the main window (useful when dealing with multiple maps).
- Use the mouse in the main window to move the map by click+drag the mouse with different mouse buttons:
 - Rotate
 - Shift left/right up/down
 - Zoom in out
 - To see/set what each mouse button does, select from the menu bar ‘Favorites’ -> Preferences. Then select ‘Mouse’ in the Categories field:

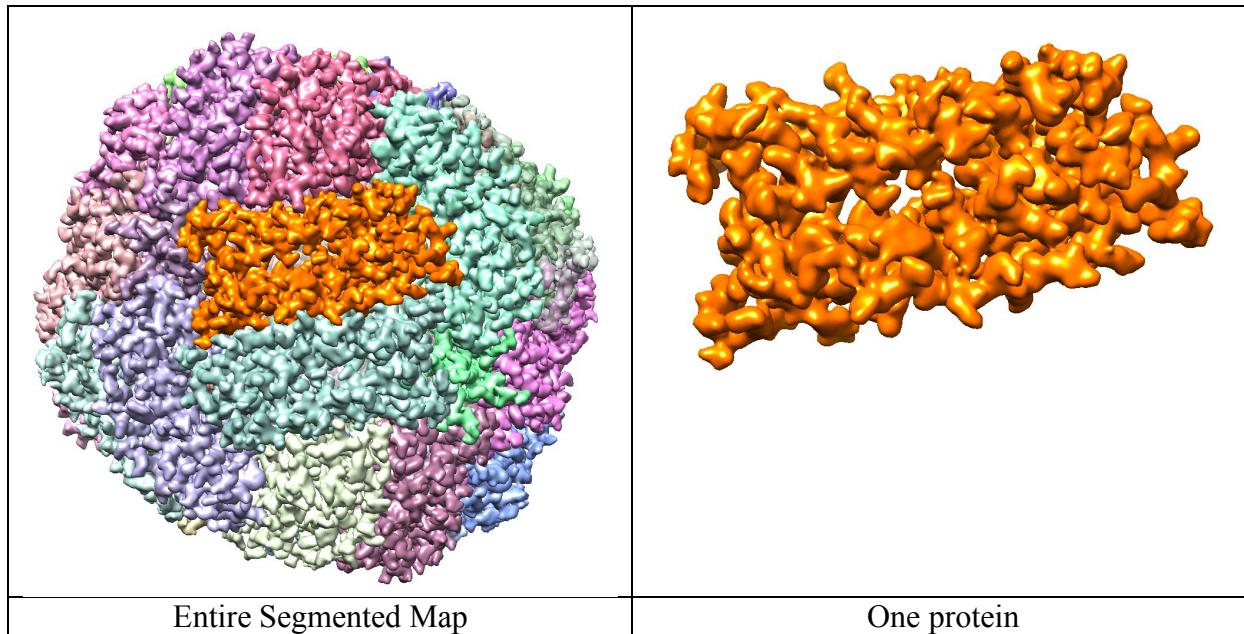


- Segmenting the map (to help with fitting the model quickly):
 - For this we will use a new version of Segger which is not bundled with Chimera:
 - <https://cryoem.slac.stanford.edu/ncmi/resources/software/segger>
 - From the main menu, select Tools -> Volume Data -> Segment Map
 - You will see the main Segger dialog:



- Click ‘Options’ button at the bottom to show the options...
- The map is segmented at the currently set threshold, hence, on the Histogram in the Volume Viewer, find a threshold where there is not too much background noise, but the surface is enclosing as much volume as possible
- Then, select the map in the ‘Segment map’ field in the Segment Map dialog, and press ‘Segment’ button.

- Set a threshold 0.015 in the Volume Viewer to get a good segmentation here.
- Select Group by Connectivity (should be already by default), and press Segment at the bottom of the Segment Map dialog. You should see a result as below. You can select a region (Click+Ctrl on it), and then press “Only Sel.” in the Shortcuts panel to see it by itself.

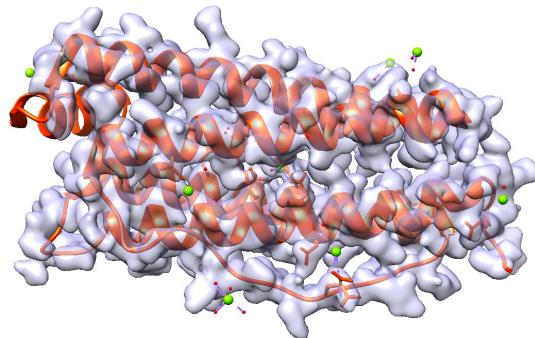


1b. Loading and fitting the model

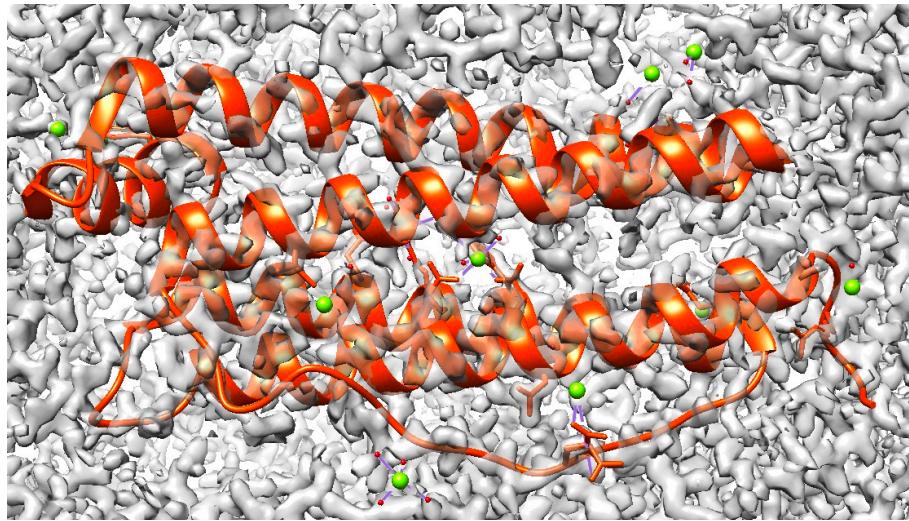
- Open the model PDB:3ajo, two options:
 - File -> Fetch by ID..., PDB: 3ajo
 - Download from:
 - <https://www.rcsb.org/structure/3ajo>
 - Download Files -> PDB Format
 - In Chimera, File -> Open..., navigate and select 3ajo.pdb, click Open
- Initially, you will see the model in a different position from the map.
- Zoom in on the model using the mouse:
 - In Chimera, ‘Models’ or Structures of proteins are shown using ‘Ribbons’
 - Ligands, e.g. Metal atoms, are shown as spheres, and nearby ‘side chains’ from protein amino acids are shown using ‘sticks’.

Map and model	Zoomed in model, showing side chains (sticks), Ligands (green spheres), and water molecules (red spheres).

- Select one segmented region in the main window (Ctrl+Click on the surface)
- Open the ‘Fit to Segments’ dialog:
 - From Tools -> Volume Data -> Fit To Segments
 - From Segment Map dialog, Options, Other Tools -> Fit
- Make sure 3ajo.pdb is selected in ‘Structure of Map to Fit:’
- Click ‘Fit’ at the bottom right in
- You should see something like this:

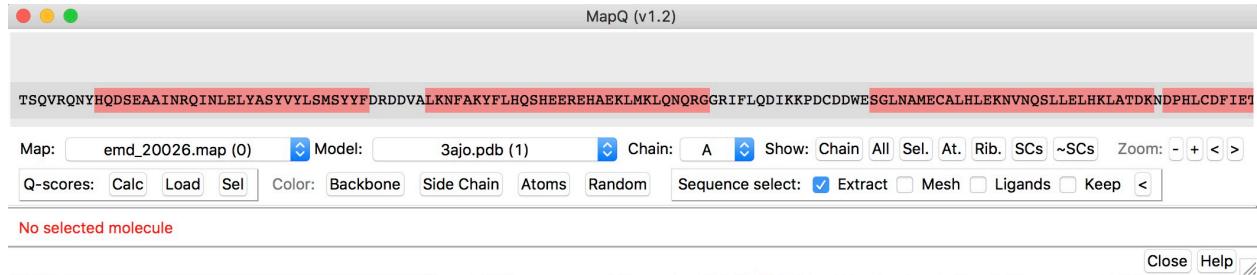


- Show the map if necessary and hide the segmentation file (with .seg) in the Models Panel:



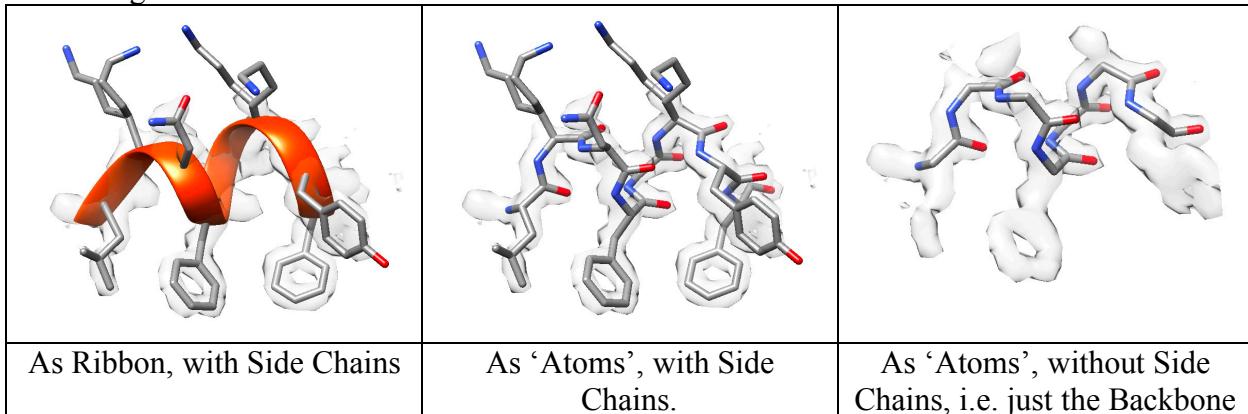
3. Using MapQ to calculate and visualize Q-scores

- For this part, we will use the ‘MapQ’ tools/plugin
- Open ModelZ dialog:
 - Tools menu -> Volume Data -> MapQ
- The ModelZ Dialog looks like this:

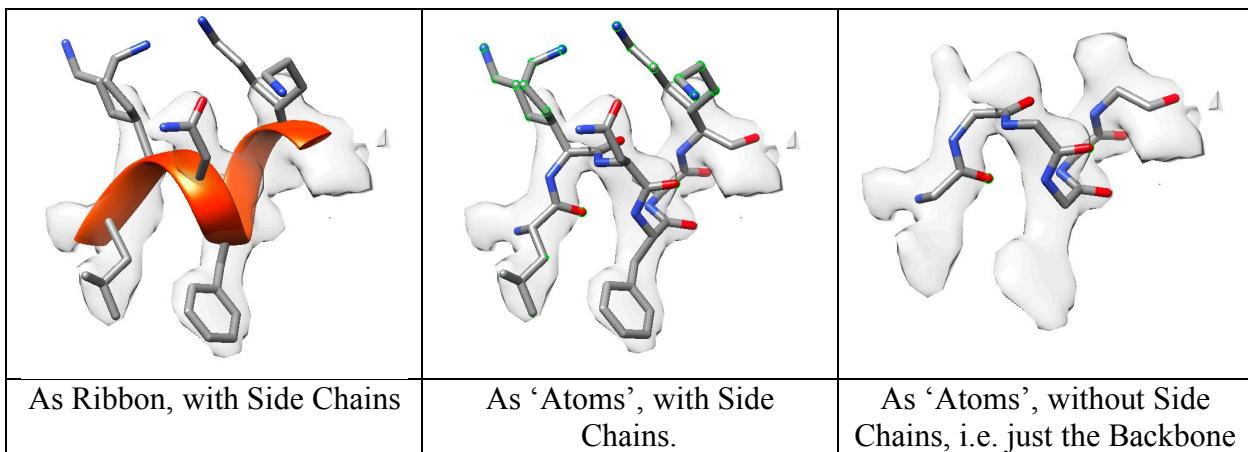


- A map, model, and chain are already selected.
 - If not, or to select a new map/model/chain, click on the respective fields just under the sequence.
- The sequence for the selected chain is shown
 - The sequence is made up of amino acid residues (e.g. Tyrosine, T, Serine, S, Glutamine, Q, etc.)
- Residues that are in ‘loops’ are shown with gray background.
- Residues that are in ‘helices’ are shown with reddish background.
- Select an entire loop/helix (Ctrl+Click+Drag on the sequence).
 - This selects just that part of the protein, and
 - The map around the selection will be extracted and shown.
- With each part selected, click ‘At.’, ‘Rib.’, ‘SCs’, ‘~SCs’ to show/switch between seeing Atoms, Ribbon, Side Chains, No-Chains. This will show the protein in different ways.

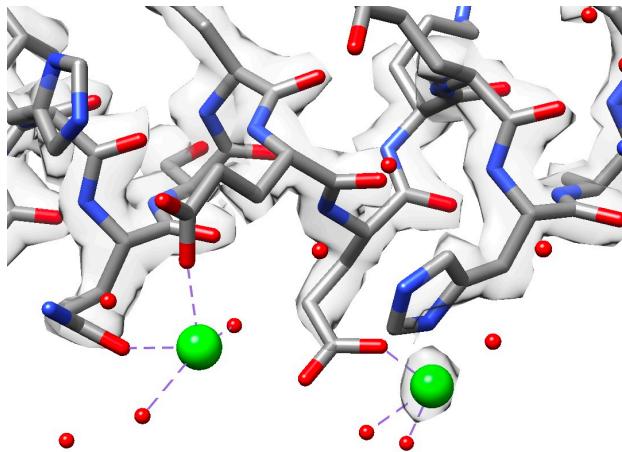
- E.g. a helix:



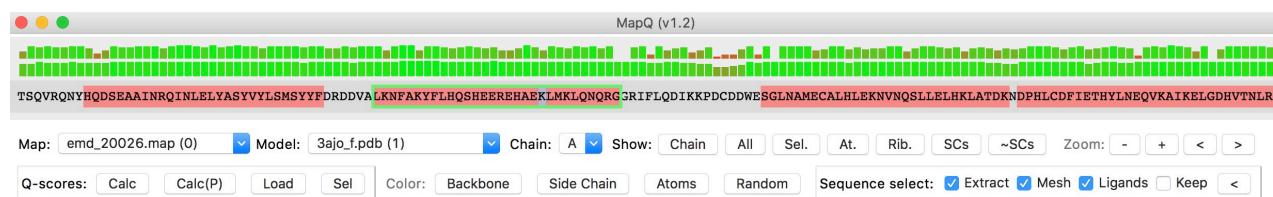
- Select the map '20027' in the Map: field and repeat the above with the same segment. The same part now looks more like this:



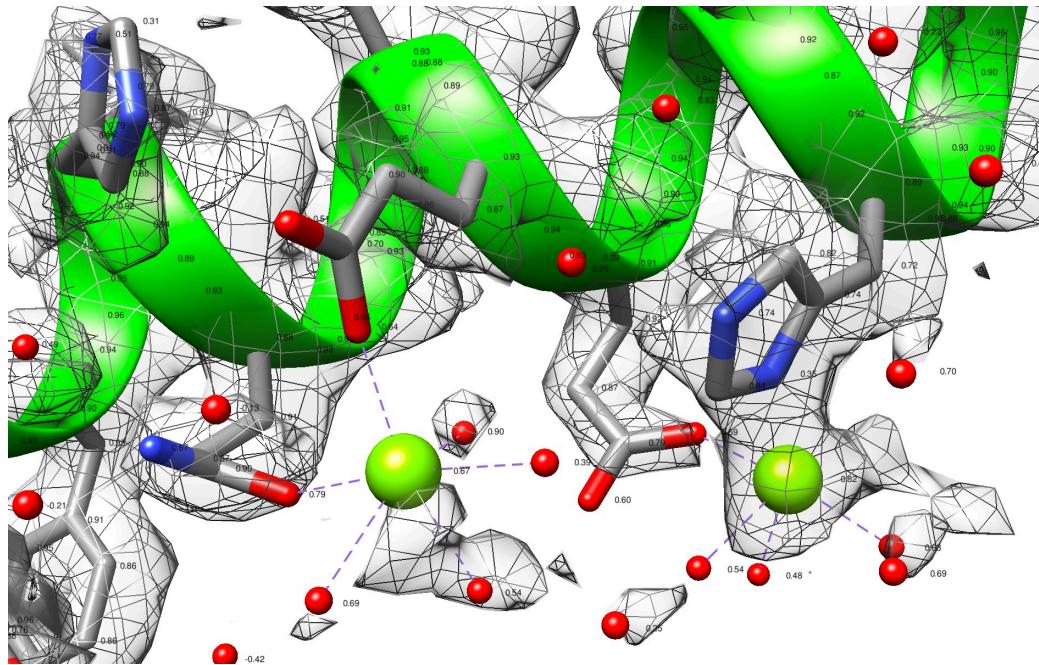
- Make sure 'Ligands' is checked in the MapQ dialog (to the right of 'Sequence Select:'). Then, select the entire second helix. The selection and extracted map will now include nearby 'ligand' atoms, which in this case are water and Mg atoms as below:



- Note how at lower thresholds you can see contours around some of the Mg and water (O) atoms. These atoms were placed in the X-ray model, and are also resolved in the cryoEM map.
- To calculate Q-scores, you can press either:
 - “Calc” button – this will calculate Q-scores for all atoms in the selected model and chain. Select ‘All’ in the Chain field to calculate for all atoms in the model.
 - “Sel” button – this will calculate Q-scores only for the selected atoms.
- After calculating Q-scores, you will see the area above the sequence show a bar graph of Q-scores for each residue. On the bottom are backbone Q-scores, on the top are side chain Q-scores:



- You will see most residues in this case have good Q-scores (green). A couple have lower (red).
- To color or show Q-scores on the model, you can press a button to the right of the Color: label:
 - **Backbone** – colors the model ribbon by backbone score (green is high, red is lower Q-score).
 - **Side Chain** – colors ribbon by side chain score.
 - **Atoms** – shows a label with the Q-score next to each atom, as below
 - **Random** – colors each chain a random color



- If you select the ‘**Keep**’ checkbox, when selecting a new part of the sequence, the previous selection is also included in what is shown. So for example, you can extract/show two parts of the model at the same time. Press the ‘<’ button to undo the last selection (helpful if the new selection was not what you were looking for). Using this, we can show, for example, the GLU 27 residue which is close to the Mg on the right:

