MapQ Tutorial

Last updated: Jan. 31, 2021 (MapQ v2.5.4, Chimera Version 1.13)

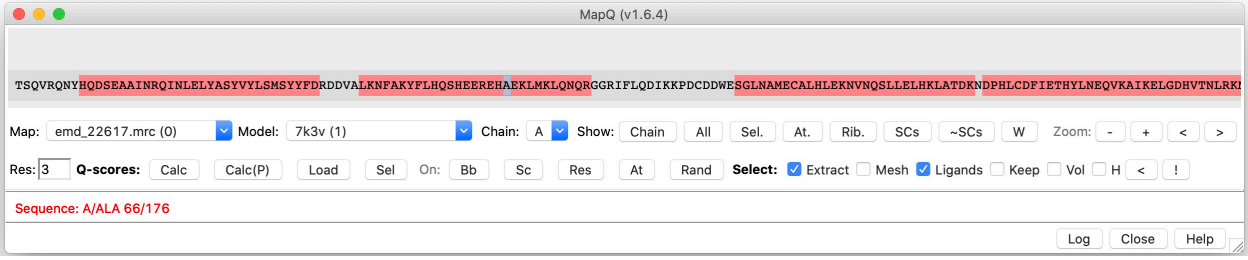
Greg Pintilie
gregp@slac.stanford.edu

Steps

1. Download and open the map EMD:22657, and the model, PDB:7k3v.

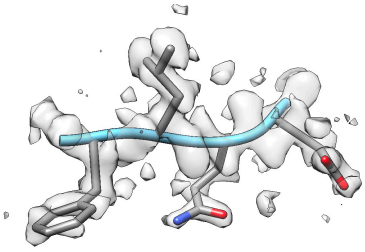
2. Open the MapQ dialog:
   - From Segger -> Options Panel -> MapQ (press the Options button at the bottom of the Segger dialog).
   - From the Chimera window, Tools -> Volume Data -> MapQ.
   - If neither of these are present there, install the latest version from www.github.com/gregdp/mapq

3. In the MapQ dialog, select the map and model in the Map: and Model: fields:

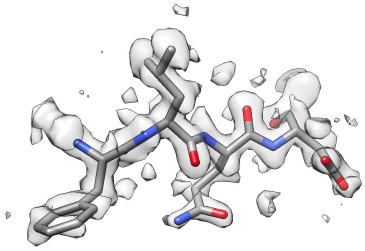


- Chain A is selected by default, and its sequence will be shown in the dialog.
  - The sequence is made up of amino acid residues (e.g. Tyrosine, T, Serine, S, Glutamine, Q, etc.)
  - Residues that are in 'loops' are shown with gray background.
  - Residues that are in 'helices' are shown with reddish background.
- The buttons to the right of 'Chain:' are:
  - '**Chain**' shows only the selected chain
  - '**All**' shows all chains
  - '**Sel**.' makes what is selected visible.
- 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.

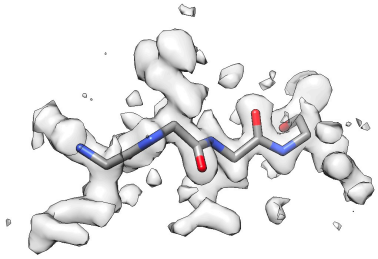
- With each part selected, click '**At.**', '**Rib.**', '**SCs**', '**~SCs**', '**W**' to show/switch between seeing Atoms, Ribbon, Side Chains, No Side Chains, Wire-bonds. This will show the protein in different ways, as illustrated below.
- The buttons to the right of Zoom: '**-**' and '**+**' decrease the sequence text size, '**<**' and '**>**' go to the beginning and end of sequence, respectively.



As Ribbon, with Side Chains

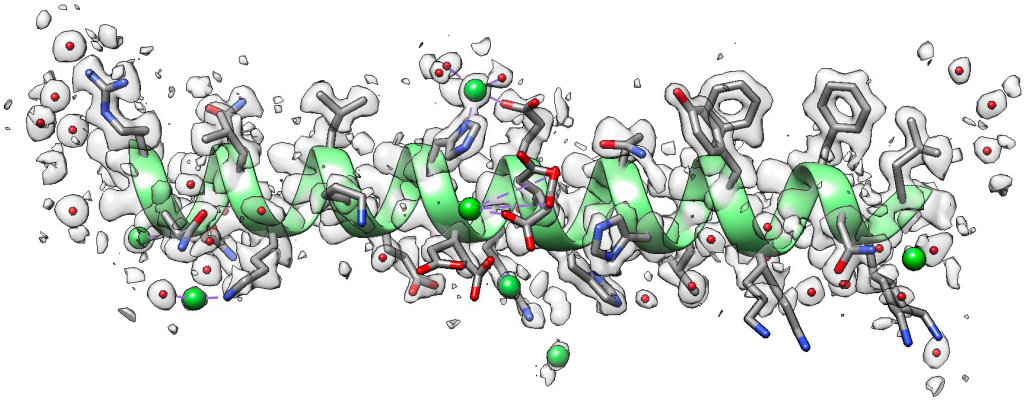


As 'Atoms', with Side Chains.



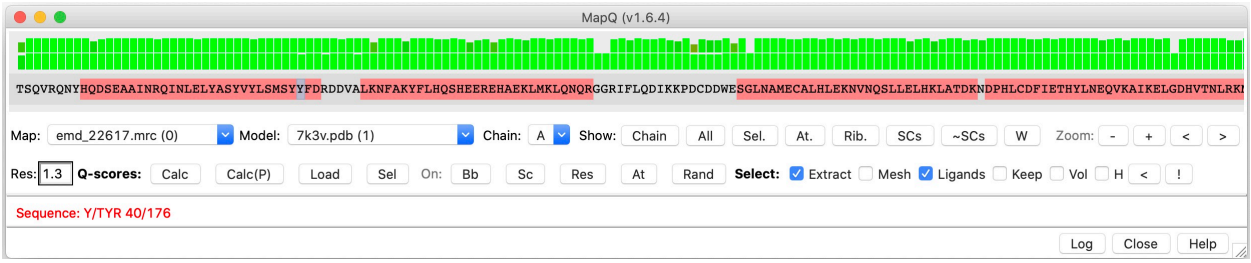
As 'Atoms', without Side Chains, i.e. just the Backbone

- Making sure '**Ligands'** is checked, 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:



- To calculate Q-scores, you can press either:
  - "**Calc**" button
    - This will calculate Q-scores for all atoms in the selected model and chain. Only Chain A for this model.
    - For this protein, it will take about 4 min to calculate the Q-scores for each atom. You can see the progress and ETA at the bottom of the main Chimera window. To cancel, you can press the 'X' button there.
  - **Calc(MP)** button
    - This will use multiple Chimera processes to make the calculation faster.

        - This will not show the progress, though it will create some temporary files in the folder where the map is placed; you can look in the _stat.txt file for an ETA.
        - For this function to work, the map and model must have been loaded from a folder accessible by Chimera, in order to make these temporary files. Otherwise, it will throw an exception.
    - "**Load**" button
        - Once the Q-scores are calculated, a new file will appear in the folder of the map, which is the same as the original model, but will have Q-scores stored in the B-factor column. It will have the name of the map and model together, e.g." 7k3v__Q__emd_22617.pdb"
        - Pressing this button will look for this file and load it instead of calculating Q-scores again. It will also:
            - Recreate the .txt file with per-residue Q-scores, using the Res: [xx] value for the expected Q-score (see section 4 below).
            - Refresh the bar plots in the MapQ dialog showing per-residue Q-scores, with the height and color of the bars normalized to this expected Q-score. So for example, for a map at 3Å resolution, the expected Q-score is 0.59, so this value will correspond to the maximum height of the bar and the most green color.
    - "**Sel**" button
        - This will calculate Q-scores only for atoms that are selected (with Ctrl+Left Click – press the up arrow to expand selection to residue, secondary structure, or entire chain).
        - It typically takes just a few seconds to calculate the Q-score for each atom, so this is faster than calculating Q-scores for an entire model.
- After calculating Q-scores using the **Calc** button (or loading them with the **Load** button), 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:



    - The resolution of the map should be entered in the Res:[_] field. This value is used to calculate what the height of the bars should be, i.e. the expected Q-scores for a residue should be.
    - To visualize Q-scores on the model, you can press a button to the right of the Color: label:
        - **BB** – colors the model ribbon by backbone score (green is high, red is lower Q-score).
        - **Sc** – colors ribbon by side chain score

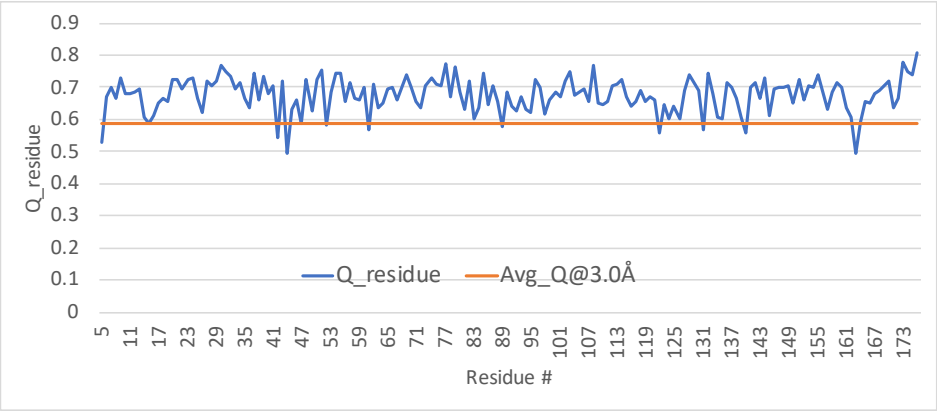
    - **Res** – colors ribbon by entire residue Q-score
    - **At** – shows a label with the Q-score next to each atom, as below
    - **Rand** – colors each chain a random color

- The checkboxes next to Select: are:
    - **Extract**: Will create a smaller masked map of the selection
    - **Mesh**: Will also add a mesh of the masked map
    - **Ligands**: will add nearby ligands to the selection
    - **Keep**: when selecting a new part of the sequence, the previous selection is also included in what is shown.
        - Press the '<' button to undo the last selection (helpful if the new selection was not what you were looking for).
        - Using this mode, we can show, for example, the GLU 27 residue which is close to the Mg on the right, while keeping the previous selection:

4. Per-residue Q-scores

- After Q-scores are calculated you will see two new files in the folder where your model is saved:
    - [model_name]_Q_[map_name].pdb
        - This file stores Q-scores in the occupancy column
        - Q-scores are loaded from this file when you press the **Load** button, if it exists
    - [model_name]_Q_[map_name]_[chain].txt
        - This is a text file which saves per-residue Q-scores
        - You can load it in Excel or another graphing program to create per-residue Q-score plots as shown below.
        - At each residue, the average Q score over all atoms in the residue are in the Q_residue column. Average Q-scores for backbone and side chain atoms are in the Q_backBone and Q_sideChain columns.
        - There are a few columns in which Q-scores from nearby residues are averaged (with 1,2,3 and 5 nearby residues). This makes the plot a bit nicer if the Q-scores fluctuate a lot.
        - In the last column in each section, the average Q-score at a resolution of xxÅ is shown, where xx is the value entered in the Res:[_] field in the MapQ dialog. This value is calculated from the equations in the plots below, where Q-scores are correlated to the reported resolutions of maps in the EMDB.
- Important notes:
    - These files are created only if you loaded your model from a file on disk; if you loaded it with Fetch or copied it from another model somehow, you will not see these files
    - The files will only have Q-scores for the chain selected; if 'All' is selected, Q-scores for all chains will be calculated and included in the files above

- If you want Q-scores for all chains in the .txt file, make sure to select 'All' next to chains before hitting the **Calc** button. Otherwise, only the Q-scores for the selected chain will be calculated.



Per-residue Q-scores example