**MapQ & Q-scores Tutorial**

Last updated: Jan. 14, 2025

MapQ v2.9.7 Chimera Version 1.13+

Greg Pintilie

gregp@slac.stanford.edu

**Overview**

1. Use the MapQ Chimera plugin to visualize a map and model, to see how backbone and side chain residues fit the density.
2. Calculate Q-scores and visualize them on the protein ribbon along with color key.
3. Plot per-residue/nucleotide/molecule Q-scores, in context of Q\_peak, Q\_low\_95%, Q\_high\_95% values. These will be described briefly here. More details can be found in the biorXiv paper:
   * <https://doi.org/10.1101/2025.01.14.633006>

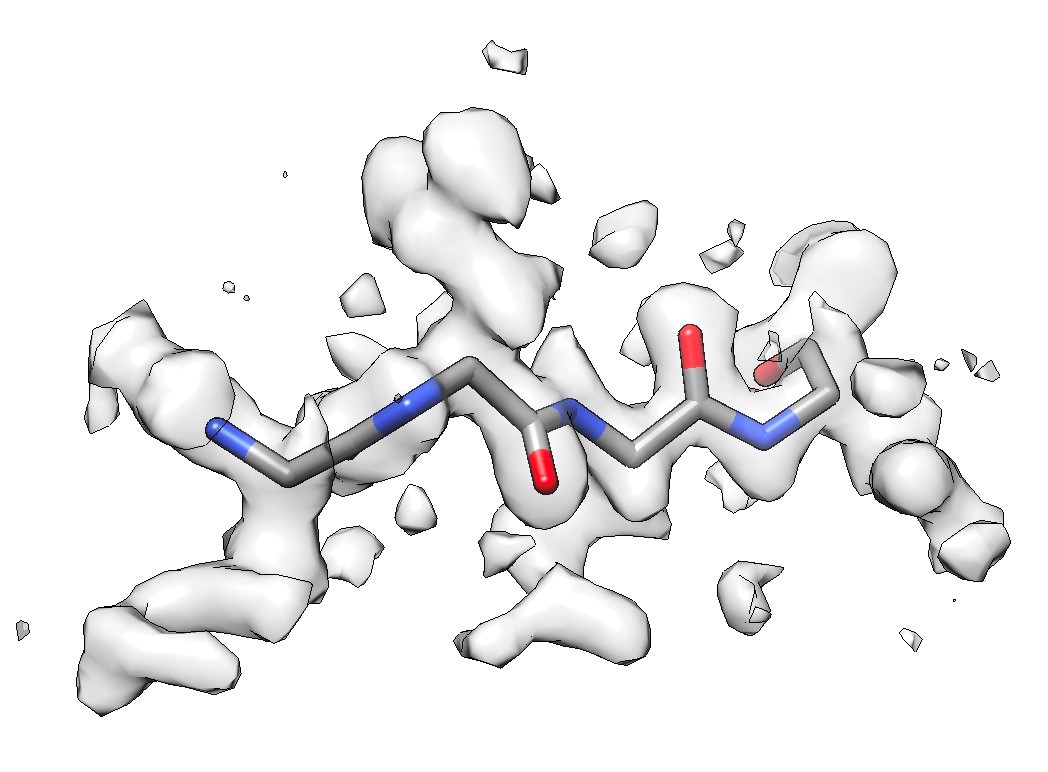
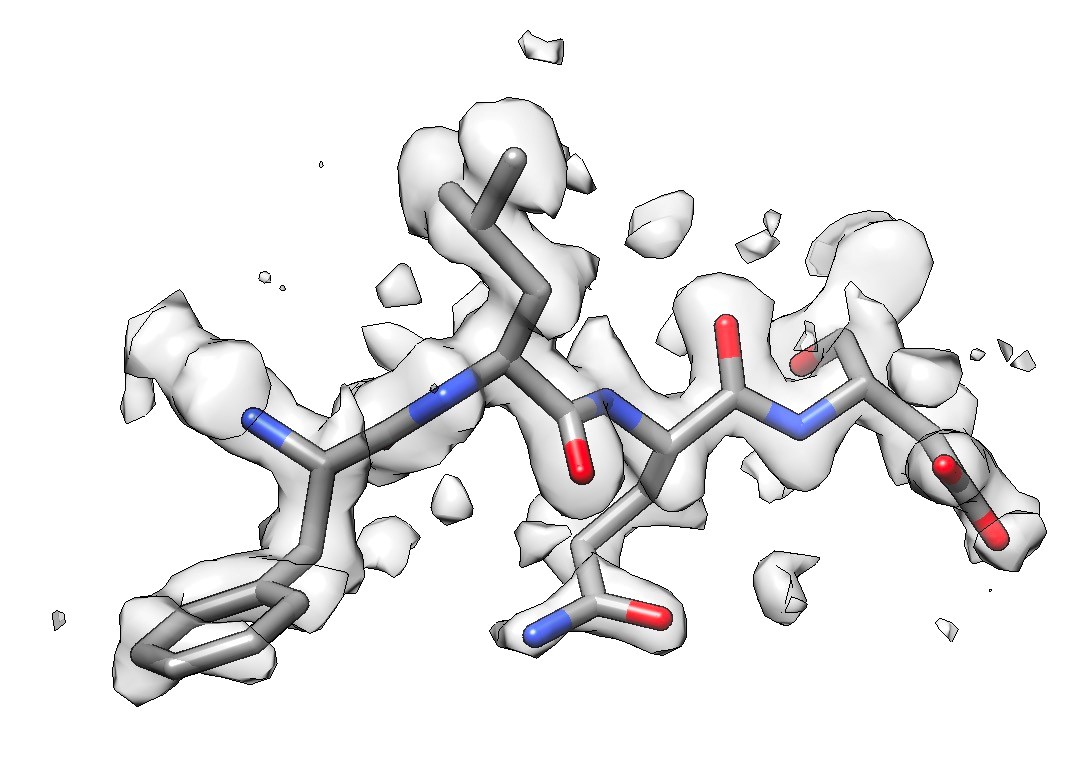
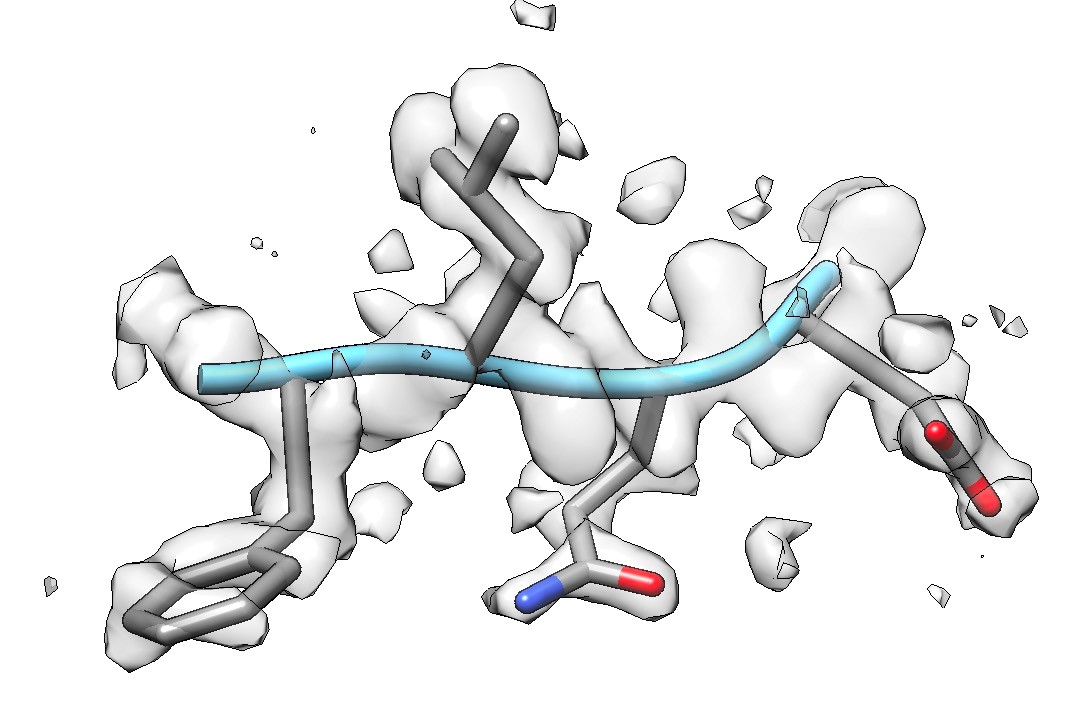
**1. Install plugin, visualize map and model**

* Download and install the latest version of the plugin
  + [www.github.com/gregdp/mapq](http://www.github.com/gregdp/mapq)
* Download and open the map EMD:22657, and the model, PDB:7k3v
  + <https://www.emdataresource.org/EMD-22657>
  + <https://www.rcsb.org/structure/7k3v>
* Open the MapQ dialog:
* From the Chimera window, Tools -> Volume Data -> MapQ.
* In the MapQ dialog, select the map and model in the Map: and Model: drop-down menus
  + Open them with File -> Open, or with the … buttons in the MapQ dialog
  + They may already be selected by default if they were already open.

Graphical user interface, application

Description automatically generated

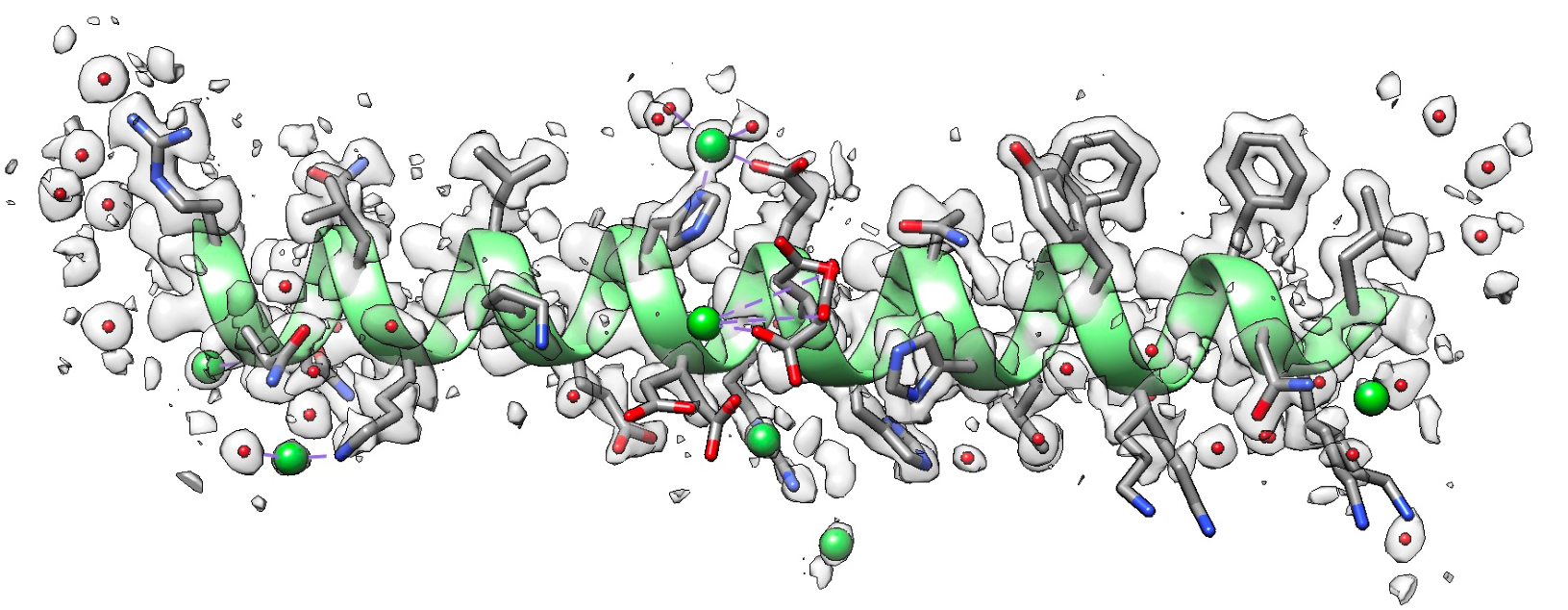
* + With Chain A selected, its sequence will be shown in the dialog.
  + The sequence is made up of amino acid residues (e.g. Tyrosine, T, Serine, S, Glutamine, Q, etc.)
  + Residues that are in ‘loops’ are shown with gray background.
  + Residues that are in ‘alpha-helices’ are shown with reddish background.
  + Residues that are in ‘beta-sheets’ are shown with blue background
  + The “…” button can be used to open new maps and models using the usual Chimera file selection dialog.
  + The buttons to the right of ‘Show:’ 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 (you should see a green outline around it)
    - The map around the selection will be extracted and shown.
  + With some part of the protein selected, click ‘**At.**’, ‘**Rib.**’, ‘**SCs**’, ‘**~SCs**’, ‘**W**’ to show/switch between seeing Atoms, Ribbon, Side Chains, No Side Chains, Wire-bonds, respectively. This will show the protein in different ways, as illustrated below.
  + The blank “ “ button hides residues that are currently selected.



*As Ribbon, with Side Chains As ‘Atoms’, with Side As ‘Atoms’, without Side*

*Chains. Chains, i.e. just the Backbone*

• Selecting the entire second helix:



# • Options

* The “Options” button at the bottom of the dialog will show another line in the dialog:



* Resolution – this is the estimated resolution of the map used to calculate Q-scores. The resolution is used to calculate Q\_peak, Q\_low\_95% and Q\_high\_95% lines, as described below.
* Sigma – by default, 0.4 is used by EMDB validation reports and is appropriate for maps up to 1Å resolution. Another possible value is 0.6 (the old default), which generates slightly higher Q-scores, but is only appropriate for maps up to 1.5Å resolution.
* Gaps – show gaps in the sequence with “.” if checked; otherwise gaps are not shown
* Extract – when selecting part of the sequence, the selected residues will be shown only, hiding all other residues, along with masked density. If not selected, the residues become the current selection, leaving everything as before.
* Mesh – also adds a mesh map when creating the zoned map for the selection
* Ligands – show any small molecules or ions next to selected residues
* Keep – keeps previously shown selection, otherwise it is removed
* H – when selected, hydrogen atoms are shown if present, otherwise they will not
* < - when using Keep, this reverts the last selection
* Zone – Zones selected map with selected atoms – creates a new smaller map enclosing the selection only
* Near – show residues near selection
* Zoom: ‘**-**‘ and ‘**+**’ decrease the sequence text size, ‘**<**’ and ‘**>**’ go to the beginning and end of sequence, respectively.

**2. Calculating and visualizing Q-scores**

* To calculate Q-scores, press the Q-scores … button, and select one of the options:

A screenshot of a cell phone

Description automatically generated

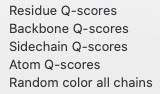
* + Press the second option for single processor (will take about 1min), or one of the multi-processing options which will be faster. Note that for the multi-processing option, the map and model selected have to have been opened from a local folder for the option to work. This may fail if the files are on Google Drive due to how it handles files.
  + When using a single process, you will see the progress and ETA at the bottom of the main Chimera window. You can cancel the process by pressing the X.
  + When using multiple processes, you can keep track of the progress and ETA by navigating to the folder where the map and model are, going into the temp folder (has \_\_Q-scores\_\_mp\_\_calculation\_\_files\_\_ in the name), and looking at the \_\_stat.txt files (there is one for each process).
  + It can help to see status messages by opening the IDLE window (Tools -> General Controls -> IDLE), or press Log at the bottom of the MapQ window.

* After calculating Q-scores using the **Calc** button (or loading them with the **Load** option), 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:

A screenshot of a phone

Description automatically generated

* + The resolution entered in the “Resolution:” field is used to calculate what the height of the bars should be, with Q\_peak being the highest value.
  + To visualize Q-scores on the model, press “Visualize…” button and select one of the options:



* + **Residue** – colors ribbon by entire residue Q-score
  + **Backbone** – colors the model ribbon by backbone score (green is high, red is lower Q-score).
  + **Sidechain** – colors ribbon by side chain score
  + **Atom** – shows a label with the Q-score next to each atom currently selected. When no atoms are selected, all previously shown labels are hidden.
  + **Random color all chains** – colors each chain a random color



*Q-scores are visualized on the protein ribbon, as per color key*

*Left:Q\_BackBone – mostly all resolved*

*Middle:Q\_SideChain – most are resolved*

*Right: Color key*

* 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/nt/molecule 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), but mostly we use the non-smoothed column.
    - Three columns come from statistical analysis of many maps and models in the EMDB:
    - Q\_peak: The most commonly observed Q-score value for the resolution entered in the MapQ dialog in the Resolution field.
    - Q\_low\_95%: only ~5% of the maps and models have Q-score value below this line for this resolution
    - Q\_high\_95%: only ~5% of maps and models have Q-score value above this line for this resolution
* 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, Qscores 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.

In the plot below, we see that the backbone atoms are resolved as commonly observed. A few side chains are not resolved, and have correspondingly lower Q-scores.

