Segger Tutorial – SWIM for RNA
(Segment-guided Water and Ion Modeling)

Last updated: Dec 17, 2024 (Segger v2.9.7)

Greg Pintilie
gregdp@stanford.edu

Overview

- In this tutorial, we will use the **SWIM** tool in **Segger** to add Ions and Water molecules to a cryoEM map of an RNA-only molecule.
  - To launch this tool, in the Segger dialog, make sure "Tools" are shown by pressing the **Tools** button at the bottom of the dialog. Then you will see a **SWIM** button. Press this button to launch the SWIM dialog.
- Roughly:
  - Waters molecules can typically be resolved at ~2.5Å and higher resolutions
  - Ions can be seen at 3.0Å and higher (even if they are not fully separable from densities of nearby atoms).

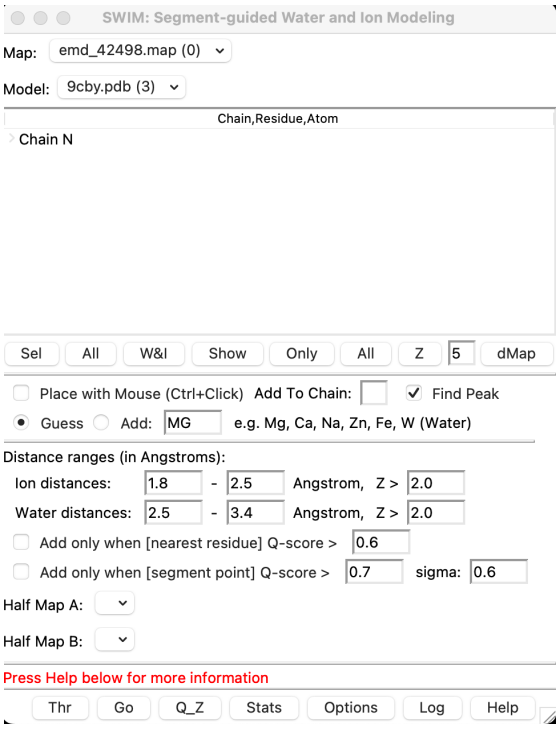
Steps

0. Download or fetch the map EMD:42498 (also the half maps), and the model, PDB:9cbw
- https://www.emdataresource.org/EMD-42498
- https://www.rcsb.org/structure/9CBY

1. This model already has some waters and ions in it.
- To look through it, feel free to zoom in to some waters and ions.
- You can check their Q-scores and Z-values (see below) by selecting an atom at a time, and then press the "Q_Z" button at the bottom of the SWIM dialog.
- If you press the "Options" button and select Half Map A and Half Map B in the panel, this will also report the Q-scores of the selected atom in these respective maps.

2. Select and delete all existing ions and water molecules:
- From the Chimera Select menu, select Residue, all non-standard.
- From the Chimera Actions menu, select Atoms/Bonds, delete

3. Open the SWIM dialog from Segger Tools panel. It should look like this:

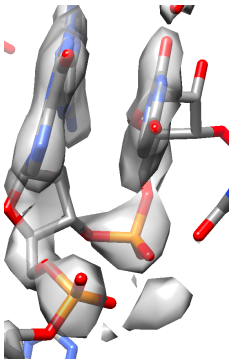


In the SWIM dialog, you can select a map and model to work on at the top. The list of chains are all the chains in the model. It is a tree list, so you can expand to see residues and atoms.

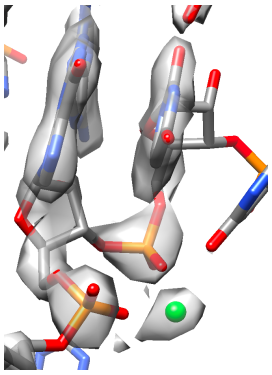
4. Select Chain A in the list, and press the **Sel** button, then **Only** button under the list to show only what is selected. The buttons have the following functionality:

- **Sel**: selects the parts of the model selected in the list
- **All**: selects the entire model
- **W&I**: selects all the water and ions in the model
- **Show**: shows what is selected in the list
- **Only**: shows only what is selected in the list
- **All**: shows the entire model
- **Z [ ]**: creates a masked map ‘zoned’ with the part of the model that is selected

6. Now we can look up close for waters and ions. Show atoms in the protein using Actions menu on the Chimera window, Atoms/Bonds -> show. Look around residue 257, and set the contour level of the zoned map to ~10 in the Volume Viewer dialog.



Before placing with Ctrl+Click



After

7. In the SWIM dialog, check '**Place with Mouse (Ctrl+Click)**', then Ctrl+Click on the blob near the residue. On the bigger blob, it will place an Ion, if the '**Guess**' option is checked.
   - To see the rationale based on nearby atoms, see the IDLE log (press Log at the bottom of the dialog to open it before placing the Ion/Water).
   - The main criteria are the distances from the peak density in the blob to the nearest atoms, and the type of atom (polar or charged).
   - If you press the **Options** button at the bottom of the SWIM dialog, the following panel is shown, where the distances and minimum Z-scores can be adjusted
     - The Z-score of a water/ion is the map value at its position, in multiples of sigma above the average. The average is the average map value at all grid points in the map, and the sigma is the standard deviation. Note this only makes sense if the map is not masked in any way.
     - A minimum Z of 2 is reasonable for placing waters and ions in most cases. If you want to be more selective, it can help to increase these values.

Distance ranges (in Angstroms):

Ion distances:	1.8	-	2.5	Angstrom, Z >	2.0
Water distances:	2.5	-	3.4	Angstrom, Z >	2.0

8. When the '**Find Peak**' option is checked, the highest value in the 'blob', or the true peak of the blob is found (by uphill climb). If the option is not checked, the water/ion is placed on the highest value under where the Click is performed.

9. If '**Add: [_]**' is checked, the atom entered in the text field is added instead of the '**Guess**' mechanism. Use Mg, Ca, Na, etc. for ions, and W for water.

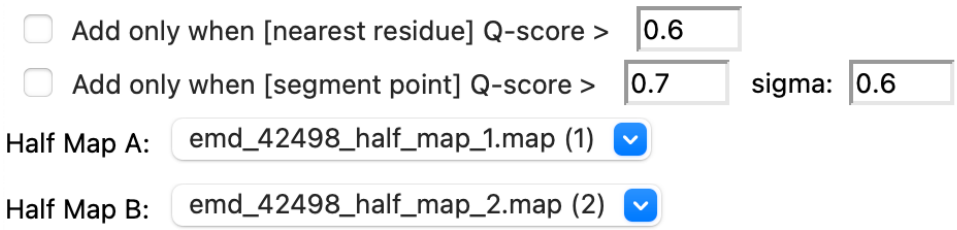
10. '**Add to Chain [_]**' may be changed to another letter or number. If left empty, the water/ion is placed as a residue in the same chain as the nearest atom.

11. When done adding manually, uncheck the '**Place with Mouse**' in the SWIM dialog. Note that Ctrl+Click normally is used to select atoms/maps, etc., so this brings back this standard functionality.

12. Now let's try placing more ions/waters **automatically** using the **Go** button.
    - One consideration is what contour level to use in the map. With the full map selected in the SWIM dialog, press the **Thresholds** button at the bottom. The contour levels are shown in the status line (in red) in the SWIM dialog and in IDLE, as follows:

      - 1 sdev above avg: 0.9919
      - 2 sdev above avg: 1.9830
      - 3 sdev above avg: 2.9741

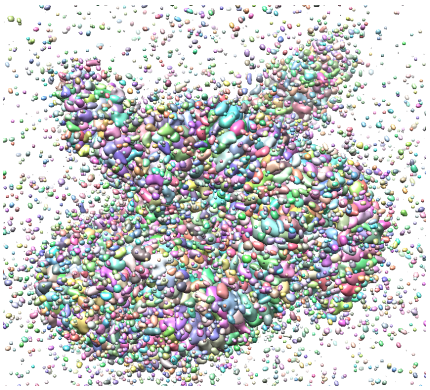
    - 1-sigma represents 1 standard deviation above the average of all values in the map, 2-sigma 2 standard deviations, and 3-sigma 3 standard deviations.
    - A sensible value may be 2-sigma above the mean, to ensure that waters/ions are placed on peaks that are significantly above the background noise.
    - A second consideration is whether the peak is well-resolved, which typically means its profile should be closer to a Gaussian function. To avoid placing waters/ions on less resolved peaks, the following option can also be checked in the **Options** panel in the SWIM dialog:



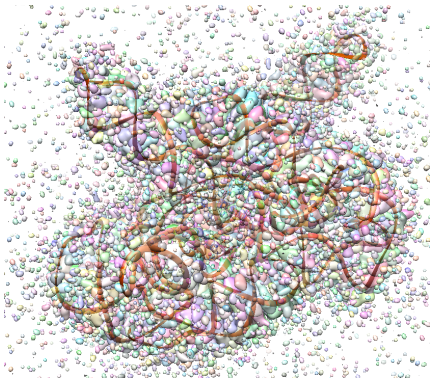
    - The higher the Q-score, the more resolved the peak should be. The sigma parameter is the width of the Gaussian the peak is compared to. The value 0.6 is typically seen in maps at ~1.5Å resolution.
    - If half-maps are available, they can be selected in the fields shown above. The Q-score in these maps will also have to be above the indicate value when placing waters and ions for extra validation.
    - The Q-score of the nearest residue can also be considered, only placing waters and ions when it's Q-score is above a given value (here, 0.6 is the default. To use this value, make sure to calculate Q-scores for the map and model using the MapQ dialog.

13. In the Volume Viewer dialog, enter the contour level for the full map, emd_42498.

14. In the Segger dialog, select the emd_42498 in the 'Map:' field, enter 0 for the number of steps in either 'Group by Connectivity' or 'Group by Smoothing' (we want to keep the original watershed regions without grouping), and press the **Segment** button. You should see something like this (you can select all the regions and press the **Tr.** button in the Segger dialog to make them transparent):



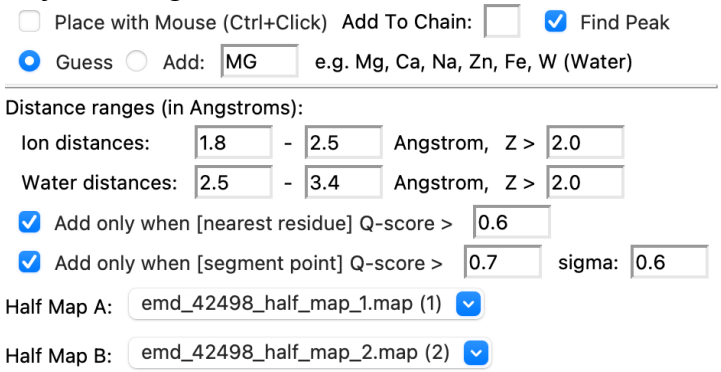
Regions after segmenting



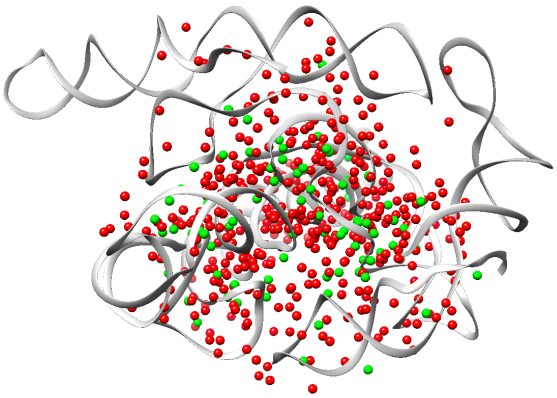
Regions after making them transparent with the Tr. button

15. Back in the SWIM dialog, select Chain A in the list, and press the **Select** button. SWIM only looks for peaks near selected atoms.

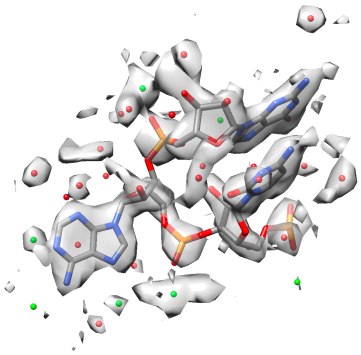
- Adjust settings as desired. Here we will use the following settings



- Again, make sure to calculate Q-scores for the structure first with MapQ – in this example a sigma of 0.6 was used in the MapQ dialog.
- Then press the **Go** button.
    - You will see a progress in the SWIM dialog status and the Chimera window status. After about a minute, you should see the results, with 412 waters and 81 ions.
    - Note that SWIM will create a new file in the same folder as the model file, with a name including the map name, the number of waters, ions, and parameters used, in this case:
    - 9cby__emd_42498__thr1.983_wZ2.0_iZ2.0_Qpeak0.70_sig0.60_Qres0.60__hA__hB__412-water__81-ion_.pdb
    - At this point, the entire model can be selected again, and the process repeated by pressing **Go** again. This may find more waters in the second shell, i.e. near other freshly placed waters or ions.
- Note that the type of ions placed is taken from the **Add: [ ]** field to the right of the **Guess** option in the SWIM dialog. By default, it is set to MG.

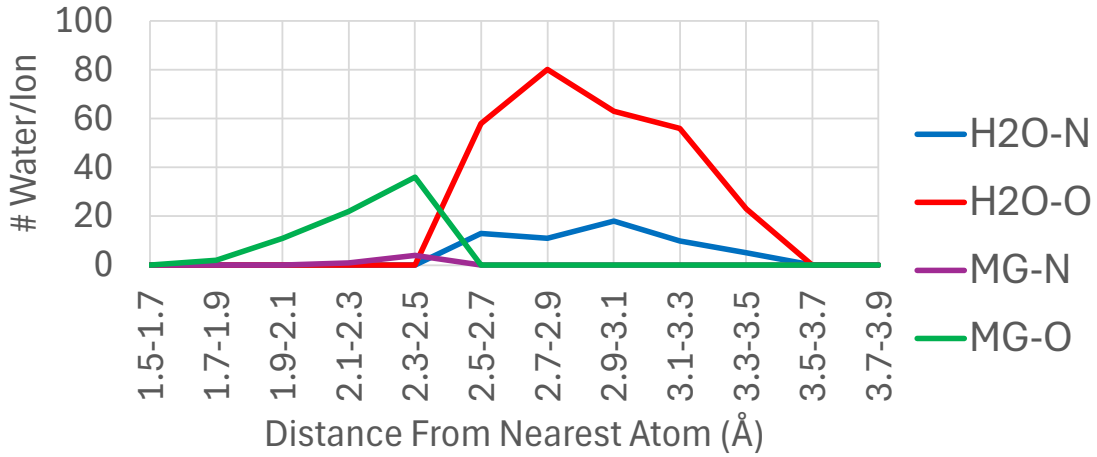


All waters and ions



A close-up

16. Finally, press the **Stats** button at the bottom of the SWIM dialog. This will collect some statistics on the distances between the waters/ions and the nearest atoms in the model, and also on Q-score distributions. The output is written to the IDLE window (make sure to open the **Log** before pressing Stats). This output can be copied and pasted to Excel, for example, to create plots like this:



- The above shows a radial-distance plot for placed water atoms to nearby atoms of type N or O. Peaks for H2O-O can be seen at ~2.8Å, which is the typical distance for hydrogen bonds.