



ModelZ Tutorial

Assessment of Structural Features in Cryo-EM Density Maps using Z-scores

Last updated: Sept 5, 2018 (ModelZ Version 1.0)

For questions, comments, suggestions, etc. please email gregp@slac.stanford.edu

Overview

In this tutorial, we will use the ModelZ tool to assess how well structural features are resolved in a density maps, using a fitted model. Structural features include:

1. secondary structure elements (SSEs) such as α -helices and β -sheets
2. side chains atoms

For further details, please see the following paper:

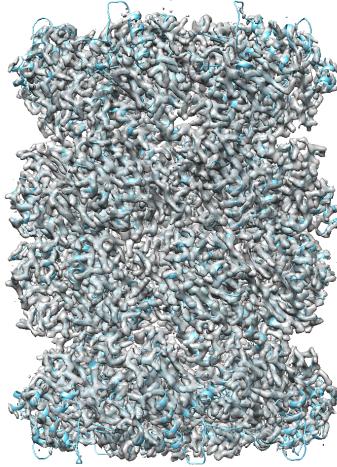
- o G. Pintilie and W. Chiu, "Assessment of Structural Features in Cryo-EM Density Maps using SSE and Side Chain Z-Scores," *J. Struct. Biol.*, Aug. 2018.
- o <https://www.ncbi.nlm.nih.gov/pubmed/30144506>

High Resolution Example (Side Chain Z-scores)

- Feel free to try this tutorial with your own map and model
- The first example, aimed at visualizing Side Chain Z-scores in a high-resolution map, the map and model are from the 2016 EMDB map challenge.
 - o To get the map:
 - curl ftp://public.emdbank.org/maps2016/finalmap/emcd103_Proteasome_fi_ltered.mrc.gz -o 103_fi.mrc.gz
 - gunzip 103_fi.mrc.gz

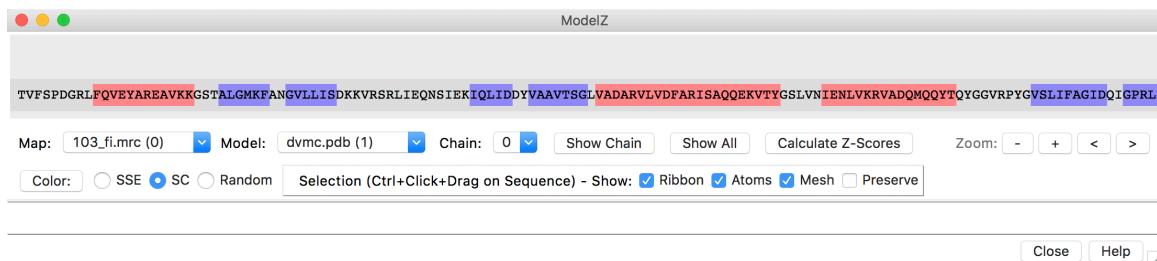
- To get the fitted model:

- curl ftp://public.emdbdatabase.org/maps2016/fitted-models/dvmc_fit_emcd103_Proteasome.pdb.gz -o dvmc.pdb.gz
- gunzip dvmc.pdb.gz



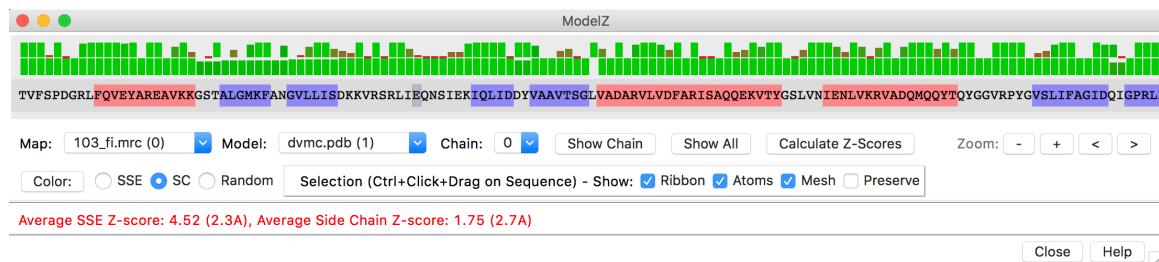
Map submitted to 2016 EMDB Challenge, Proteasome 103, with fitted model.

- Select on the top menu Tools -> Volume Data -> ModelZ



- Upon opening, the dialog automatically selects one of the open maps (above it is 103_fi.mrc), and model (dvmc.pdb). It also select a chain (0). These are shown and can be changed in the drop-down menus next to **Map:**, **Model:**, and **Chain:** respectively.
 - Note that selecting another Model or Chain in these drop-down menus will also automatically show that model and only the selected chain in the model (the other chains are still in the model but hidden).
 - Clicking the **Show All** button will show all chains in the model
 - Clicking the **Show Chain** will show only the selected chain.
- At the top of the dialog, the **sequence** of the selected chain is shown, with secondary structures highlighted, **α -helices** in red and **β -sheets** in blue.

- The sequence can be moved left or right, using **middle mouse button** drag, or **scroll wheel** (the pointer has to be positioned on the sequence while using the scroll wheel).
- The sequence can also be ‘zoomed’, i.e. have the text size increased or decreased, using the – and + buttons to the right of the **Zoom:** label. The < and > buttons will go to the beginning or end of the sequence.
- To calculate Z-scores, press the **Calculate Z-Scores** button.
 - The dialog will now show a bar-graph representation of the Z-scores at each residue:



- After the calculations are done, note the status message at the bottom of the dialog:
 - **Average SSE Z-score: 4.52 (2.3A), Average Side Chain Z-score: 1.75 (2.7A)**
 - The **average SSE Z-score** is **4.52**. This score is expected* in maps with $\sim 2.3\text{\AA}$ resolution.
 - The **average Side Chain Z-score** is **1.75**. This score is expected* in maps with $\sim 2.7\text{\AA}$ resolution.
 - * These statistics are based on all the maps submitted to the 2016 EMDB Map challenge and their reported resolutions (see publication mentioned above for more details).
 - The reported resolution for this map is 2.8\AA .
- The bar-graph region has two vertical halves:
 - The top half shows **Side Chain** Z-scores
 - Low scores have short/red bars (Gly and Ala have no scores).
 - High scores have tall/green bars.
 - The tallest/most green bar represents a Z-score of 2 or higher.
 - Short/red bar represents a Z-score of 0 or lower.
 - The bottom half shows **SSE** Z-scores
 - A single SSE Z-score represents all residues in the SSE, so each residue within the SSE is given the same Z-score.

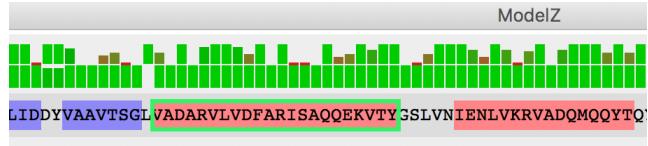
- Color coding and bar length are as above, with Z-scores of 0 and below being short red bars, and Z-scores of 4 and above being tall green bars.
- In the above, the map is of quite high resolution so SSEs are well defined – most SSE bars are tall and green. However side chains have more varied Z-scores.
- Z-scores can be used to color each residue to give an idea of how well features are resolved throughout the map. To do so, use the following radio button group:

Color: SSE SC Random

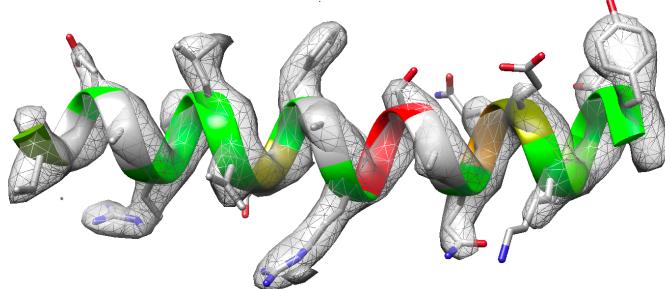
- With **SSE** selected, the SSE Z-scores are used to color the ribbon of the selected chain.
- With **SC** selected, the Side Chain Z-scores are used.
- With **Random** selected, each chain is colored a random color. (This is useful when showing all chains, i.e. after pressing the **Show All** button).
- Small segments of the model and map can be extracted for better visualization of the side chains. As shown on the dialog, this can be done by holding the **Ctrl** key on the keyboard, and **Click+Drag** with the **left mouse button** on the sequence itself.

Selection (Ctrl+Click+Drag on Sequence) - Show: Ribbon Atoms Mesh Preserve

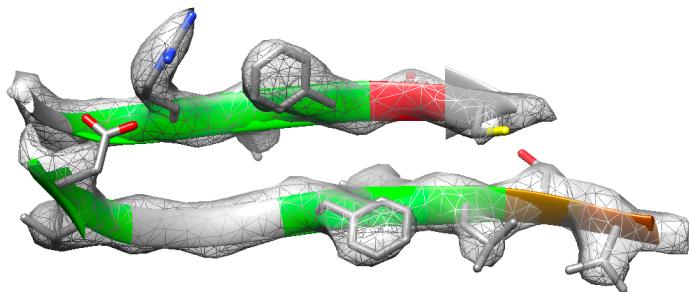
- Several toggles affect the result as follows:
 - ‘Ribbon’ toggles whether a ribbon is shown for the backbone in the selection. Otherwise, the backbone atoms are shown.
 - ‘Atoms’ toggles whether atoms are shown. This is useful for showing the side chain atoms when looking at side chain Z-scores. (Note that if Ribbon is off and Atoms is off, nothing will be shown, so at least one of these should be on.)
 - ‘Mesh’ toggles whether two new extracted maps are created, one in mesh mode as above, otherwise a single extracted map is created and shown with a transparent gray surface at the same threshold as the map selected in the **Map:** Menu.
- For example, to select one of the helices in the example above, hold the **Ctrl** key, and Click+Drag starting on the V all the way to the Y. As shown below, this will create a bright green box around the selected residues.



- With all options set as above, the following should be shown in the main Chimera window:



- 'Preserve' toggles whether the previous selection is kept - the extracted model and map will include previous selections until this is toggled to off. This is useful to extract several β -strands together, for example:
 - With 'Preserve' off, Ctrl+Click+Drag the first strand.
 - Then, with 'Preserve' on, Ctrl+Click+Drag the second strand. The result will extract both strands:

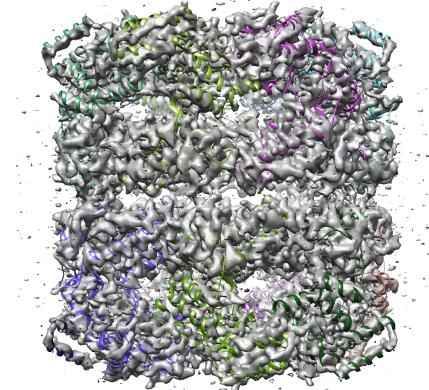


- At any point, to show the entire chain again, press the 'Show Chain' button. When extracting, new sub-maps will appear as new maps in the Model Panel, and they can be closed or hidden just as any other maps. Note that when selecting a new part of the sequence, the extracted maps will replace any previously extracted maps.

Medium-resolution example (SSE Z-scores)

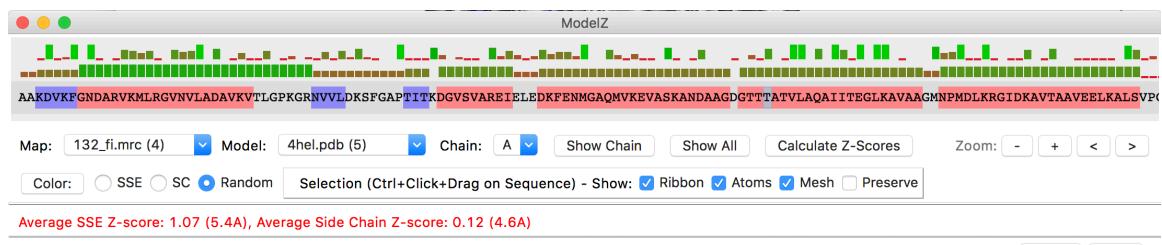
- This example shows how SSE Z-scores can be useful for medium resolution maps

- Again feel free to try with your own map and fitted model.
- The images below were generated with another map submitted to the 2016 EMDB map challenge:
 - To get the map:
 - curl ftp://public.emdatabank.org/maps2016/finalmap/emcd132_GroEL_filtere.d.mrc.gz -o 132_fi.mrc.gz
 - gunzip 132_fi.mrc.gz
 - To get the fitted model:
 - curl ftp://public.emdatabank.org/maps2016/fitted-models/4hel_fit_emcd132_GroEL.pdb.gz -o 4hel.pdb.gz
 - gunzip 4hel.pdb.gz



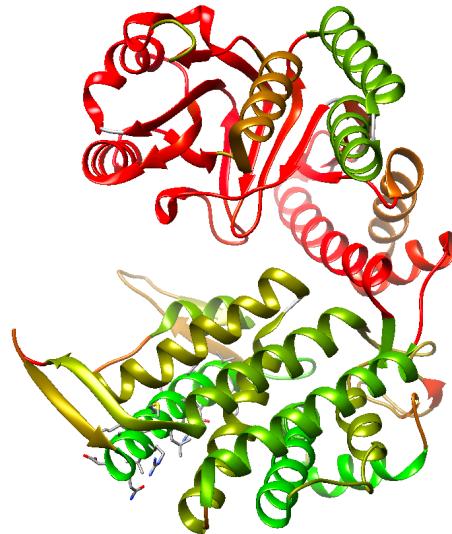
Map from 2016 EMDB Map challenge, GroEL 132, with fitted model.

- With the map and model, chain A selected, press the Calculate Z-Scores button. The result in the ModelZ dialog should look like this:

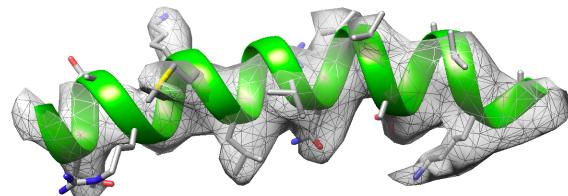


- Note that side chains are not resolved as well in this map, and only sporadic green bars are seen for the top half of the bar graph.
- The SSE bar graph now has more variety, i.e. some SSEs are well-resolved (taller green bars) and some are not resolved as well (shorter red bars).
- Note also the result text displayed in red at the bottom of the dialog:
 - **Average SSE Z-score: 1.07 (5.4A), Average Side Chain Z-score: 0.12 (4.6A)**
 - The **average SSE Z-score** is **1.07**. This score is expected* in maps with **~5.4 Å** resolution.
 - The **average Side Chain Z-score** is **0.12**. This score is expected* in maps with **~4.6 Å** resolution.

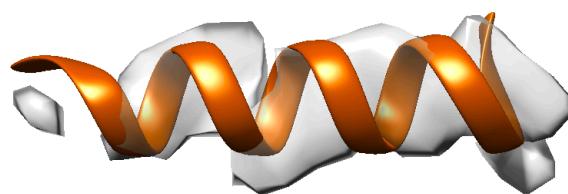
- * These statistics are based on all the maps submitted to the 2016 EMDB Map challenge and their reported resolutions (see publication mentioned above for more details).
- The reported resolution for this map is 4.1Å.
- With “SSE” selected, press the **Color:** button. The chain will be colored as such:



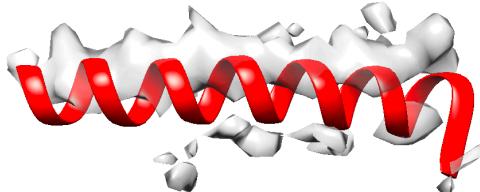
- In the above image, we can see that SSEs are resolved better in the equatorial domain at the bottom, and not as well in the apical domain at the top.
- For example, a helix in the equatorial domain:



- In the above, the helix is well defined, with helix pitch slightly visible, and some side chains even having corresponding high density.
- Another helix in the middle domain:



- The above helix is a bit less defined in the density, and hence has a more reddish color.
- A helix in the apical domain:



- In the above, the density map seems to have helical density close to the helix, but the helix may not be well-fitted to the map, i.e. flexible fitting or refinement may be needed in this case.

Conclusions

- Z-scores are useful for evaluating how well structural features are resolved in a density map.
- A well fitted model is required for the analysis.
- Z-scores can indicate which features throughout the model are well resolved.
- Z-scores are good when they are positive, the more positive the better. In short, this means that the feature at its placed location matches the map better than at slightly displaced locations; see publication for more details.

Other Considerations

- While SSE and Side Chain Z-scores are meant to measure how well certain features are resolved in a density map, the model has to be fitted to the map as well as possible.
- When working with a model that is not fitted to the map:
 - The model can be fitted rigidly, e.g. with:
 - Fit in Map from the Tools -> Volume Data menu
 - fitmap command
 - Fit to Segments also from Tools -> Volume Data menu (See the Segger / Fit to Segments tutorial)
 - Situs
 - etc.
 - The model could also be flexibly fitted or refined with:
 - Molecular Dynamics Flexible Fitting
 - Phenix
 - Refmac
 - FlexEM
 - etc.