Q&A Report September 10, 2021

Q For the models that fitted into low resolution or poor map quality maps e.g., because of protein flexibility, what ccc value could be meaningful or worth keeping the models? I assume ccc could vary depend on how you post-process the map.

A live answered

Q A tangential question - but are the CERES rerefinement phenix protocols available to the community?

Q A live answered. Hi Nilesh, the best resources are the publication and the about section. https://cci.lbl.gov/ceres/about https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7787109/pdf/d-77-00048.pdf. CERES aims to use the standard phenix tools. The automation scripts are not currently public, but we can discuss that possibility. A followup: If you are interested in the automation scripts, Dorothee suggested contacting her: dcliebschner@lbl.gov

Q I see, perhaps showing ccc AND visual fitting of the model in the map (with proper sharpening) mightbe helpful to validate the fitting?

A live answered

Q Are you comparing the CCC between one map and one model? But the map comes from averages of different atomic states, and theoretically can be be incompatible with any one model. Is there a principled way to compare the entire atomic ensemble with a map, and find the “best ensemble”?

A live answered

Q can you take a volume series from cryoSPARC and fit an ensemble to it?

A live answered

Q can you give a demo how to do the multiple fitting?

A live answered

Q My crude attempt was to fit a model to each of the volumes but yours could be more sophisticated.

A Yeah that’s most likely what we would do with TEMPy at the moment too. Although as discussed, more elaborate methods would be possible

Q Are virtual torsion outliers greater in lower resolution structures than in higher resolution structures (e.g. binned in 0.5 Ang. increments from 0.5 Ang. to 2.5 Ang. resolution)?

A live answered

Q Since the number of C3'-endo is significantly larger than C2'-endo, shouldn't the outlier thresholds different for these two? If I understand correctly, currently the density of 10 is used for both countours. It might lead to bias where C2'-endo is always having more outliers.

A live answered

Q What are practical applications of this analysis? Like in crystallography, when we see Ramachandran plot outliers, we fix orientations of neighboring carbonyls or Ca-Cb bond. Is there a rule to eliminate RNA outliers?

A live answered

Q No, I can’t

A live answered

Q If RNA/DNA is too complicated to be boiled down to 2D plots: Do you see a perspective for machine learning, e.g. a network that just tells you „this residue is good“ or „this residue requires attention“ during modeling/deposition?
Thanks for the great talk!

A live answered

Q Saying in crystallography I meant in proteins

A live answered

Q Thanks! What would you consider your main contribution? Are these familiar metrics that are now extended to lots of public data? Or are the metrics you showed novel to the RNA field?

A live answered

Q i would ask the same question about multiple fitting demo and also does any of these data are implemented in ccpem or other?

A live answered - Yes

Q I didn't mean to focus on RNA vs DNA. I just meant, did you invent any new statistical measures? Or did you use tried and true measures in the nucleic acid field?

A Hi Geoffrey, this was an attempt to develop a similar metric for RNA as is done for proteins with Rama plots.

Q Sorry, I didn't mean that by "statistical measure", I just mean the eta-theta, etc features. Are these are typical, or new?

A The virtual torsion angles have been described in previous papers (Wadley et al., 2007; Keating and Pyle, 2010) although they have not been used specifically for validation. We are implementing and analyzing these plots specifically for validating nucleic acids structures.

Q For map-model FSC, some people use square root of 0.175 (~0.35). No oversitting at 0.5 threshold would be good enough?

A No Answer

Q Are there any written workflows / guides for steps taken by the groups? Would be great to have that as a resouce to apply to your own structures.

A We collect workflow metadata for each submitted structure, and that is available in spreadsheet format for every challenge. Available via https://zenodo.org/communities/3dcryoemchallenges for previous challenges, https://challenges.emdataresource.org/?q=2021-model-challenge for the current challenge.

Q If I have mutiple localy refined map for different region a structure, do I submit all?

A You should deposit all maps that are described in your publication.

Q Is there a versioning similar to PDB model versioning we heard about planned for EMDB maps, since EMPIAR lets the community correct maps or even 3D classifications from raw image data?

A live answered

Q What is the lowest resolution accepted to perform this fitting into the map?

A live answered - We have gone up to a little over 5 A

Q Would the MDFF itself be sufficient for doing the fit to cryo-EM density, i.e. without the initial MELD simulation?

A live answered. It the topological information is accurate enough then you won't need to generate and ensemble with MELD.

Q Can be MDFF refined model from your software applied to Phenix validation? I had issue to submit MDFF refined model to Phenix due to presence of some water molecule and charged histidines. Is there some solution to remove water molecules and to change charged histidines back to uncharged?

A Yes, this is usually due to a residue naming convention that Phenix and VMD just do not agree on. Are your waters important? Are they coordinating or ordered?

Q I did not have ordered waters. Is there some alternative how to get validation of geometry parameters of model after MDFF?

A Yes, please mail dsarkar@asu.edu and me (asinghar@asu.edu) with the question, I will have to give you a long answer. Also, take a look at this paper: https://pubmed.ncbi.nlm.nih.gov/33028525/

Q Chris, could you describe the QM target data used to fit the NNP potentials a bit more?

A live answered