I ran an ab initio refinement

Tue, 12/05/2017 - 11:46 #1

michael.cianfrocco

I ran an ab initio refinement

I ran an ab initio refinement and, while it generally worked, the later iterations clearly got over-refined. So, I'm wondering how to select an intermediate volume after ab initio has completed? It seems that I probably can't retrieve this after it finishes, since I did see that I could stop the run and take the current volume while it was running.

Otherwise, I can just rerun the calculation.

Thanks!

PS - Congrats on the release, very exciting :)

timgrant

Hi Michael,

Hi Michael,

Thanks for the congratulations! Yes, the ab-inito often ends in something that is quite overfit - this is because the FSC is mostly ignored. The overfitting has not really been a problem for us, as long as the map is correct at lower resolution it should be fine for further processing (e.g. auto refinement) as long as you specify a resolution that the map is not really overfit at (e.g. start from 20A) to start from. The auto refinement then does take the FSC into account, and so the result shouldn't have so much over fitting.

To answer your original question, there is no easy way to take an earlier model, this will likely be added in a future release. If you really want to do it, you can get access to the intermediate files, they are stored in Scratch/Startup. They will be deleted when you click the Finish button, or close the project, so you need to get to them before that. They will also be binned, so you will need to resample them to the correct pixel size (you can use the command line resample program that comes with cisTEM) and then import them as a volume asset if you would like to use them.

Cheers!

Yue Liu

Hi Tim.

Hi Tim.

I would be grateful if you could help me with two related questions.

1.Is there a way to know at which resolution there is basically no overfitting after running the ab-inito 3D refinement? In the tutorial, the final resolution limit in the ab-inito run was 8A, which is also the starting point of auto-refinement as recommended. Just wonder if there might be any concern of introducing overfitting into the final reconstruction.

2. For classfication and 3D refinement actions, when a job is killed (walltime exceeds a time limit on cluster), is it possible to restart from where it ends (e.g., in the middle of a run)?

Thanks.

Yue

timgrant

Hi Yue,

Hi Yue,

1. The information a bit beyond the resolution limit is free from refinement bias, as it is not included in the refinement. That means that ideally you would have high confidence in your structure up to at least this resolution. The 8A start after the first auto refine is probably too high resolution for most samples, but Apoferritin is a bit of a weird case. At any rate because the final resolution of 3A is considerably beyond the 8A you can a degree of confidence that this is ok. In a normal scheme you would do your first auto refine from 20A or even 30A. If this goes well and you end up with a high resolution structure, things can probably be improved by doing another auto-refine from 8A, but this is only safe to do if your resolution already extends beyond 8A. The short answer is that you need to be confident that the "true" resolution of your structure significantly exceeds the refinement resolution, if it doesn't then the resolution estimation is highly likely to be wrong.

2. For 2D classification you should be able to continue from the round that you got to by selecting it as the input starting references. This is also true of the manual 3D refinement. For the auto refinement it is not currently possible to continue from a crashed job. In this case you either need to rerun the auto refinement from the start, or switch to manual refinement.

Cheers!