**You should try to complete AreTomo and IMOD processing for each of the datasets found under the folder: /mnt/hpc_users/share/resources/HPC_additional_sample_data/**

During the processing steps, take note any issues and write down questions you encounter during the processing for the datasets. We'll discuss these as a group on Friday February 21ˢᵗ, and review individual processing outputs from our $HOME folders.

The datasets provided include example tilt-series data collected for bacteria, milled lamella of HeLa cells, and tilt-series for DPS proteins provided by Dr. Jae Yang. Three datasets for RSV virus are provided within the RSV folder by Dr. Bryan Sibert. With these, AreTomo and IMOD will have different success or challenges that we will talk about further at the end of the week.

```
├── Bacteria
│   ├── Example_Aretomo.sh
│   ├── Box9_G2_26Kx_ts_1_aligned.st
│   └── Box9_G2_26Kx_ts_1_aligned.st.mdoc
├── Cell_Lamella
│   ├── Example_Aretomo.sh
│   ├── W1368_G2_26Kx_L2_ts_2_Alimc.mrc
│   └── W1368_G2_26Kx_L2_ts_2_Alimc.mrc.mdoc
├── DPS
│   ├── 3005—G2_105Kx_ts25_Alimc.mrc
│   ├── 3005—G2_105Kx_ts25_Alimc.mrc.mdoc
│   ├── Example_Aretomo.sh
├── imod_tutorial_data
│   └── tutorialData—1K.tar.gz
└── RSV
    ├── imod
    └── tiltseries
```

To begin, connect to the Open OnDemand desktop ([https://ondemandcryo.biochem.wisc.edu](https://ondemandcryo.biochem.wisc.edu)) and then open a terminal window. This command, *all one line*, will copy to your directory:
`` `rsync -avu /mnt/hpc_users/share/resources/HPC_additional_sample_data/ ~/HPC_additional_sample_data` ``

Change to the folder in your home directory with:`` `cd ~/HPC_additional_sample_data` ``

Change to a subdirectory, `` `cd Cell_Lamella` `` and then you can run an HPC job with AreTomo. Here there is an example script containing a command for AreTomo in the `` `Example_Aretomo.sh` ``.

You can use the command `` `submit_to_a5000.sh ./Example_Aretomo.sh` `` to submit the job on a GPU node. These jobs will take between 5-10 minutes to run and can be followed by checking `` `squeue` `` to see when they have completed. Then view the resulting tomogram.mrc that are created under the `` `output` `` folder with `` `3dmod output/tomogram.mrc` ``.

See the AreTomo manual ([https://gensoft.pasteur.fr/docs/AreTomo/1.3.4/AreTomoManual_1.3.0_09292022.pdf](https://gensoft.pasteur.fr/docs/AreTomo/1.3.4/AreTomoManual_1.3.0_09292022.pdf)) and try to change parameters in the script (ex. -VolZ 2500 -AlignZ 1800 control the volume thickness and region used in alignment). If you hit any issues in these steps, reach out and we can help you get

these completed. The RSV folders do not have an 'Example_Aretomo.sh', and you will need to create your own from the previous examples – make sure to change the stack name and pixel sizes.

Not all datasets will not generate ideal tomograms with the fiducial-free alignment of AreTomo. In those cases, working through Etomo for an alignment may be better. These datasets (**Bacteria, DPS, RSV**) include fiducials that can be used for an alignment using Etomo instead of AreTomo.

There are several excellent resources online for Etomo and tomogram reconstruction that you can follow along. If you haven't used Etomo, start by completing the tutorial dataset in (1).

1. A sample dataset that is included as 'imod_tutorial_data'. This can be a good starting point to learn the steps, with many example images to show what is being done at each stage. https://bio3d.colorado.edu/imod/doc/etomoTutorial.html
2. The second is specific for cryo-EM tilt-series data. It also describes steps for fitting CTF Correction and erasing gold beads. https://bio3d.colorado.edu/imod/doc/cryoExample.html
3. The IMOD Tomography Guide (https://bio3d.colorado.edu/imod/doc/tomoguide.html#TOP) is detailed reference for each of the steps in the Etomo workflow.

| Folder | Stack file | Pixel Size (Å) | Fiducial Diameter (nm) |
|---|---|---|---|
| Bacteria | Box9_G2_26Kx_ts_1_aligned.st | 4.727 | 10 |
| Cell_Lamella | W1368_G2_26Kx_L2_ts_2_Alimc.mrc | 4.727 | *NA* |
| DPS | 3005-G2_105Kx_ts25_Alimc.mrc | 1.22 | 6 |
| Imod data | BBa.st and BBb.st | 10.09 | 10 |
| RSV | w17_g1_ts_013_ali.st, w17_g1_ts_014_ali.st, w17_g5_ts_011_ali.st | 1.693 | 6 |

These datasets (**Bacteria, DPS, RSV**) include fiducials; for each, please try to complete a reconstruction using Etomo for a tilt-series in each on the Open OnDemand virtual desktop.

Some steps allow you to use parallel processing options. Here you should also be able to choose 'cluster' and that will allow you to submit these as CPU jobs on the HPC cluster. See example image below. You can increase the number of 'Used' CPUs higher, and 32-64 CPU may be sufficient for any jobs here.