Cryo EM - Data Processing

Scope

The data processing step in the cryo-EM workflow is highly dependent on the target protein. However, some steps, especially in the preprocessing part, are fairly common. Here, we present the simplified version of a cryo-EM flowchart with useful external links toward more comprehensive resources. These main steps are part of many cryo-EM data processing bundle programs, such as RELION, cryoSPARC, cisTEM, and EMAN2.

Procedures

A. Preprocessing

In this step, raw cryo-EM movies collected on a transmission electron microscope are corrected for image shift between subframes, with subsequent contrast transfer function (CTF) correction.

- Motion correction (MotionCor2, RELION's own implementation)
- CTF estimation (CTFFIND4, Gctf)

B. Particle Picking

Automated particle picking is available in all cryo-EM processing bundles, as well as in some standalone programs.

- crYOLO
- TOPAZ
- Gautomatch

C. Particle Classification

Once the particle picking process is done, the next step involves 2D classification, which allows further division between "good" particles and contaminations (or broken particles). This step can be performed several times until the resulting 2D classes look satisfactory.

The next step is usually 3D classification, for which an initial model is needed. At this step, many parameters can be optimized, including a number of classes or regularization parameters.

D. Refinement

The next step usually includes refinement of a particular, homogeneous class. To address small heterogeneities within the structures, perform more focused refinement with masking and/or subtraction of certain parts of the structures.

- Focused Refinement
- Multi-Body Refinement

Bayesian polishing is performed to account for radiation damage and any residual particle motion.

Ewald sphere correction improves resolution of large particles.