

## EMAN2.1 – A New Generation of Software for Validated Single Particle Analysis and Single Particle Tomography

SC Murray<sup>1,2</sup>, JG Galaz-Montoya<sup>2,3</sup>, G Tang<sup>2,3</sup>, JF Flanagan IV<sup>2,3</sup>, SJ Ludtke<sup>2,3</sup>

<sup>1</sup>. Graduate Program in Structural and Computational Biology & Molecular Biophysics, Baylor College of Medicine, One Baylor Plaza, Houston TX 77025

<sup>2</sup>. National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza Houston TX 77025

<sup>3</sup>. Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston TX 77025

EMAN is a general-purpose scientific image-processing suite developed primarily for the TEM community[1]. With over ½ million lines of Python and C++ code, hundreds of image processing algorithms and a cross-platform graphical interface, it is a capable tool for quantitative greyscale image analysis in 2-D or 3-D.

Over the last three years, the Cryo-EM community has begun shifting focus from improving resolution, to validation of results. EMAN2.1 represents the results of major new developments in EMAN2's single particle reconstruction and single particle tomography workflows. The new tools such as e2refine-easy, integrate “gold standard” resolution assessment into the refinement process along with a number of new optimizations, both speeding the refinement process, and eliminating the need for empirical filtration of reconstructions by users. EMAN2.1 also integrates support for tilt-pair validation and “true resolution” testing to insure self-consistency among data and final 3-D maps. This is one of the few methods which can identify incorrect maps at low resolution.

Another approach often applied to increase confidence in a 3-D structure is to reprocess the same data using multiple algorithms, preferably based on different mathematical methods. EMAN2.1 includes an interface for converting data and metadata into the appropriate format for reprocessing in Relion[2] or FreAlign[3], two alternative single particle reconstruction packages. Once these packages complete their refinements, the results can be imported back into EMAN2.1 for comparison and analysis. As an alternative, EMAN2.1 can perform the opposite process as well. A refinement completed originally in one of these other packages can be converted into an EMAN2.1 project, which can then be used to re-refine the data from scratch.

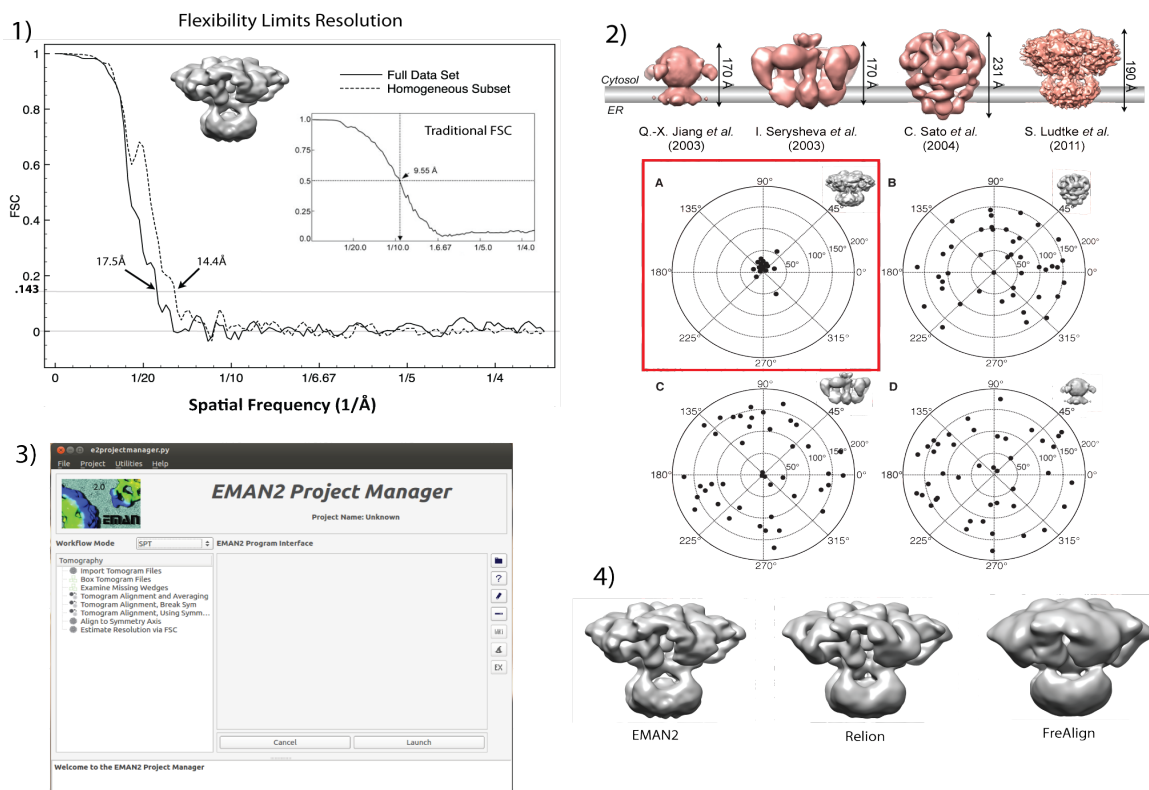
Another important area of improvement is single particle tomography. Rather than the traditional approach of reconstructing large numbers of 2-D images of identical particles in random orientations, in single particle tomography, tomographic data is collected for fields of particles, producing a low resolution and incomplete, but 3-D reconstruction for each individual molecule. EMAN2.1 now incorporates tools for subtomogram extraction, and a variety of different approaches for alignment and averaging of particles. This approach is a powerful alternative to single particle analysis particularly in cases where the particles are flexible or heterogeneous in solution.

EMAN2.1 also incorporates several important ease-of-use improvements. At the request of our users, the EMAN2.0 strategy for storing image data and metadata has been retired in favor of flat HDF and JSON formatted files. The project manager and file browser were both rewritten for speed and

flexibility. The new 3D viewer now allows scenes to be constructed from multiple objects including 3-D annotations, and supports much more flexible visualization options. Overall, the new version represents a substantial step forward in producing accurate and reliable 3-D structures, with less effort on the part of the scientist end-user.

References:

[1] G Tang et al. *Journal of Structural Biology* **157** (2007), p.38-46  
 [2] S Scheres. *Journal of Structural Biology* **180** (2012), p.519-530  
 [3] N Grigorieff. *Journal of Structural Biology* **157** (2007), p.117-125



**Figure 1.** An example of “gold standard” resolution assessment in EMAN2. In this example, the gold standard resolution is substantially lower than the resolution found using traditional measures, however we clearly see that this is due to flexibility in the specimen rather than noise bias, as using a more homogeneous subset of data produces a significant resolution improvement.

**Figure 2.** An example of tilt-validation in EMAN2 showing 4 published structures of IP3R which are completely different from each other. Using tilt-validation we can determine that only one structure is consistent with the data.

**Figure 3.** The single-particle tomography workflow showing a variety of new and useful programs

**Figure 4.** Three Structures generated using the comparative refinement ability of EMAN2 showing similarity in final reconstruction across several different reconstruction programs.