# ResLog Plots: A New Metric for the Quality of Cryo-EM Reconstructions

Scott M. Stagg[1,2], Alex J. Noble[3], Michael Spilman[2], and Michael Chapman[4]

[1] Department of Chemistry and Biochemistry, 95 Chieftain Way, Florida State University, Tallahassee, FL 32306.
[2] Institute of Molecular Biophysics, 91 Chieftain Way, Florida State University, Tallahassee, FL 32306.
[3.] Department of Physics, 77 Chieftan Way, Florida State University, Tallahassee, FL 32306-4350
[4.] Department of Biochemistry & Molecular Biology, School of Medicine, Oregon Health & Science University, Portland, OR 97239-3098

As the field of three-dimensional electron microscopy becomes increasingly mainstream there is a growing need for reliable metrics for assessing the quality of the 3D reconstructions. The most commonly used metric for the quality of a 3D reconstruction is the resolution calculated from the Fourier Shell Correlation (FSC) curve. In this method, the FSC curve is determined by splitting a given single particle dataset into two halves, reconstructing 3D volumes, and calculating the correlation of Fourier coefficients of increasing spatial frequency shells. Resolution is then estimated by reporting the resolution at a point on the FSC curve, usually at FSC0.5 or FSC0.143 [1]. One source of controversy regarding FSC is that it is possible to produce a single particle refinement that converges on an incorrect 3D model that reasonable resolution. This problem arises from the fact that generating accurate 3D reconstructions depends on the ability to accurately classify particles relative to each other. Accurate classification depends on a variety of factors such as the initial model used for refinement, the shape of the molecule, the amount of noise in the data, etc. If the orientations of too many particles are inaccurately determined, this will distort the corresponding 3D reconstruction and can result in structures with artifactual features that could be interpreted as biologically relevant. What is needed is a metric that can be used to distinguish correct and incorrect 3D reconstructions.

It has been shown that there is a linear dependence of spatial frequency (inverse resolution) on the logarithm of number of particles that go into a reconstruction [1-3]. Here we show that the y-intercept of a plot of the log of the number of particles versus the spatial frequency corresponds to the average quality of the alignment of the particles and the slope of the plot corresponds to the overall quality of the data going into a reconstruction. These parameters can be used to find the optimal set of particles for a given single particle dataset (Fig. 1). Finally we show that the value at the y-intercept can be used to distinguish correct from incorrect 3D reconstructions (Fig. 2).

References:
[1] Rosenthal, P.B., Henderson, R., J Mol Biol **333** (2003), p. 721.
[2] LeBarron, J., Grassucci, R.A., Shaikh, T.R., Baxter, W.T., Sengupta, J., Frank, J., J Struct Biol **164** (2008), p. 24.
[3] Stagg, S.M., Lander, G.C., Quispe, J., Voss, N.R., Cheng, A., et al., J Struct Biol **163** (2008), p. 29.
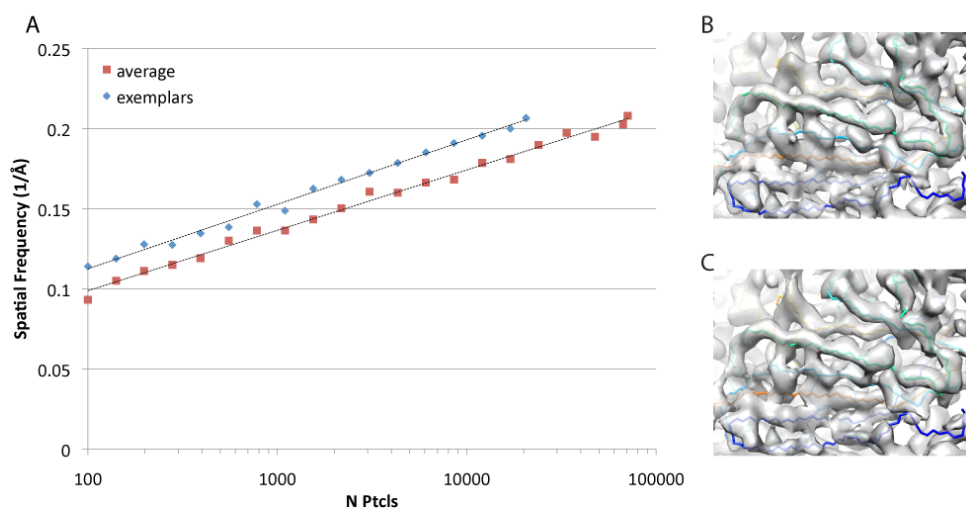
Figure 1



Fig. 1. Comparison of optimal particles with randomly selected particles for an adeno-associated virus dataset. A) Plot of spatial frequency (1/resolution) of a reconstruction versus the number of particles contributing to the reconstruction (a ResLog plot). The plot compares reconstructions including all particles (red squares) with reconstructions with subpar particles removed (blue diamonds). B) 4.8 Å resolution reconstruction from 70,725 particles. C) 4.8 Å reconstruction from 20,580 exemplar particles. All reconstructions are shown at the same contour level (5.4σ). The region shown in B and C is focussing on a β sheet in AAV. Separation between strands of the β sheet can be observed at this resolution.
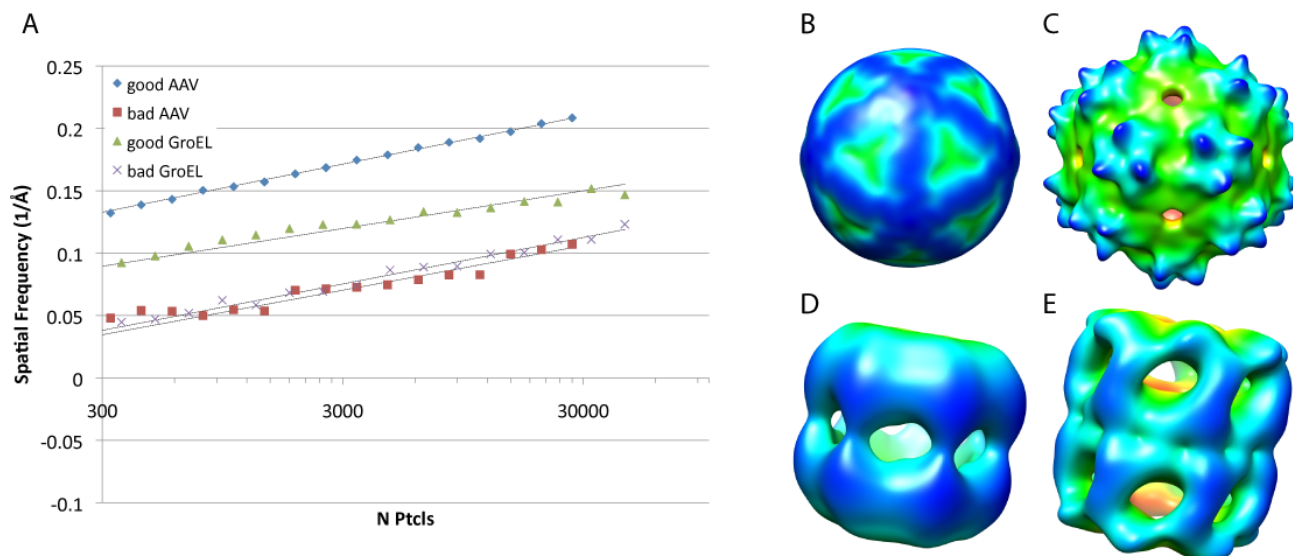
Figure 2



Fig. 2. Comparison of ResLog plots for correct and incorrect reconstructions. A) ResLog plots for correct and incorrect reconstructions of AAV-DJ and GroEL. The ResLog trend for the incorrect reconstructions was fit based on the linear portion of the curve. This shows that the y-intercept for ResLog plots of incorrect reconstructions is substantially lower than for good reconstructions. B) Incorrect AAV-DJ reconstruction at 26 Å resolution. C) Correct AAV-DJ reconstruction at 26 Å. D) Incorrect GroEL reconstruction at 26 Å E) Correct GroEL reconstruction at 26 Å.