SOME APPLICATION OF MATHEMATICS TO STRUCTURAL BIOLOGY

Garib N Murshudov
LMB-MRC, Francis Crick Avenue, Cambridge, UK, CB2 0QH
garib@mrc-lmb.cam.ac.uk

Macromolecules play role in all aspects of organisms. Knowledge of theie 3D structures is key for understanding how they work. Understanding their mechanisms of action is an important step in fixing them when they go wrong, and therefore in designing drugs to cure a range of illnesses. There are three main experimental techniques to study these structures: Macromolecular X-ray crystallography (MX), nuclear magnetic resonance (NMR) and single particle Electron Microscopy (EM). Usually the data produced by these techniques are very noisy and the number of parameters of the model to be estimated is very large, often exceeding 1000000. The problem is how to extract optimal information from such noisy data? Extracted information should also be consistent with prior knowledge about the macromolecules. Analysis of such large and noisy data and derivation of biologically useful models with large number of parameters require state-of-art statistical and computational tools. Solving such problems can naturally be formulated in one of several ways such as Bayesian statistics, regularization of ill-posed problems.

Our approach for solving this type of problems has several components[1]: 1) analysis and organisation of prior chemical and structural knowledge in machine readable from[2]; 2) designing of prior probability distribution encapsulating such information; 3) designing the likelihood function that links experimental data with the parameters of the atomic models to be derived[1]; 4) designing posterior probability distribution that combines prior knowledge and experimental data[1]; 5) optimisation of large system accounting for the fact often the problem is often ill-conditioned[1]; 6) producing the possible images that is used for manual and/or automatic critique and revision of the model. By its nature the problem is highly non-linear and therefore large number of iteration is needed for full convergence. To aid faster solution of the problem we also develop graphical tools[3] for manual inspection of model vs experimental data as well as prior knowledge.

Our approach to regularised de-blurring, that is essentially a solution of inverse problem with large number of parameters in the presence of noise and when blurring function is approximately known, has similarity to Tikhonov type regularisations and it is solved approximately using Fourier methods for one particular type of blurring function – position independent blurring[1]. Application of developed tools to EM model interpretation will also be presented[4].

References:

1) Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA "REFMAC5 for the Refinement of Macromolecular Crystal Structures" *Acta Cryst*: , 2011:D67;355-367

2) Long F, Vagin A, Young P, Murshudov GN. "BALBES: a molecular replacement pipeline" Acta Cryst 2008:D64;125-134

3) Emsley P, Lokhamp, B, Scott, WG, Cowtan K. "Features and developments of Coot", Acta Cryst. 2010:D66;486=501

4) Amunts, A, Brown, A, Bai, X-C, Llace, JL, Hussain, T, Emsley, P, Long, F, Murshudov G, Scheres, SHW, Ramakrishnan V, Structure of the Yeast Mitochondiral Large Ribosomal Subunit", Scince:343;1485-1489