Statistical Modelling of Heterogeneity in Crystallographic Fragment Screening Data

Ylva Strid Holmertz, Nicholas DeBouver, Nicholas M Pearce

Physics, Chemistry and Biology (IFM), SciLifeLab, Linköping University.

Elucidation of macromolecular dynamics remains a core focus in structural biology. However, the direct identification of partially-occupied high-energy microstates remains difficult in macromolecular crystallography (MX) and cryo-electron microscopy (CryoEM) due to the inherent averaging over different microstates, whereby low-occupancy states are obscured in the resulting experimentally determined average electron density. Direct construction of multi-state models from this average is difficult given the number of degrees of freedom in macromolecular models, making the problem degenerate.

Fragment screening experiments aim to identify small molecules that bind to a target macromolecule. However, crystallographic fragment screening (CFS) experimental data have also displayed variations in the occupancies of alternate conformers unrelated to fragment binding. To model these variations, we have constructed a statistical model where the occupancies of different local states are sampled by a Dirichlet distribution; we propose that maximising the likelihood of the model will directly reveal the densities for the underlying states. Here, we present the preliminary implementation of the statistical model and show some preliminary results.