Let’s Treat Multi-Dataset Crystallography Data like a Brain

Nicholas D. DeBouver, Ylva Strid Holmertz, Nicholas M. Pearce

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

Crystallographic Fragment Screening (CFS) identifies small 'fragment' molecules that serve as starting points in structure-based drug design. The PanDDA method demonstrates that a real-space approach can be employed to routinely identify weakly binding molecules through statistical analysis of electron densities. However, the quality of this analysis heavily depends on the homogeneity of the data. Despite being collected from crystals of the same protein grown under identical conditions; heterogeneity is frequently observed within CFS data sets. Currently, global clustering of data sets is necessary to obtain homogeneous data suitable for PanDDA analysis.

Here, we introduce developments in the PanDDA software that incorporate MRI image-alignment algorithms to enhance real-space analysis by reducing local inter-dataset heterogeneity. This approach allows for continuous distortions of the electron density, enabling significant local changes to be distinguished from uninteresting large-scale effects. Improved alignment of electron densities from CFS experiments will further enhance the sensitivity of the analysis.