

Better bulk-solvent models can improve model-to-data fit

Pavel V. Afonine

Lawrence Berkeley National Lab, Berkeley, CA, USA

International Center for Quantum and Molecular Structures, Shanghai University, Shanghai, PRC

PAfonine@lbl.gov

On average, crystals of bio-macromolecules contain about 50% of solvent. Some of this solvent is structured and therefore can be visualized in crystallographic Fourier maps and in turn modeled using atomic representation. The rest of the solvent is disordered and cannot be modeled using individual atoms. Since this solvent represents a substantial amount of scattering matter, its contribution to the total scattering needs to be accounted for. There have been several models proposed in the past to model disordered solvent. To date, the flat (mask-based) model is almost universally used in modern crystallographic software. This is because of its simplicity and yet relatively good modeling power. The flat model assumes electron density is constant anywhere in the unit cell where there is no atomic model placed. The advantage of this model is that it is easy to implement and very fast to calculate. The latter is particularly important given that a typical refinement job performs several dozens of such calculations. Major deviations from the assumption of a flat model include: 1) local concentration of solvent components of specific type, 2) unmodeled ligands, 3) partial occupancy and 4) possible lack of solvent in certain regions, such as hydrophobic cores. These deviations contribute to the mismatch between experimental and model-calculated structure factors across all resolution range. Also, this is a contributing factor to systematically larger R factors for macromolecules (compared to small molecule cases) as well as hard-to-explain residual features in difference maps.

Attempts to improve the existing flat bulk solvent model by using non-uniform solvent mask will be presented.