

Molecular Therapy for 2.5-4Å Models: Anecdotes and Progress from the CryoEM Model Challenge

Jane S Richardson, Christopher J Williams, Lizbeth L Videau, David C Richardson

Many of the most exciting macromolecular structures are currently in the 2.5-4Å resolution range -- considered "low" for X-ray and "high" for cryoEM. For those techniques there are distinct data characteristics, but rather similar opportunities and difficulties for model building. Over most of the 2.5 to 4Å resolution range, secondary structure and chain connectivity can be determined with reasonable reliability. However, as resolution drops from 2.5 to 4Å, detailed local backbone conformation, peptide orientation, sidechain position, and even sequence alignment progressively become unclear and error-prone, and these problems are not well diagnosed by traditional validation such as Ramachandran outliers. We have been developing new tools in the MolProbity system that are tuned for the harder case of these resolutions, both for protein and for RNA.

Acting as assessors for the EM Database's CryoEM Model Challenge was an excellent opportunity to test the effectiveness of these new tools, and to determine the most prevalent systematic errors made by currently used methods. The most serious errors we found were sequence misalignments, sidechains pushed to the wrong side of a strand by misfit backbone, or incorrect ribose puckers, because these can destroy the validity of a binding or active site. The most prevalent errors seen were peptide orientations off by up to 180°, which move Ramachandran-plot positions for the two adjacent residues into the wrong local minima. They are not correctable by refinement, and can falsely seem OK. However, they are especially well diagnosed by our CaBLAM tool, which looks at C α and O atoms across 5 residues. We are currently implementing protocols in the Phenix software suite that should be able to automatically correct most of those bad peptide orientations, and also to detect and correct several types of less common errors.