

Use of ultra-high resolution structures in validation

Mariusz Jaskolski

*Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University
and Center for Biocrystallographic Research, Institute of Bioorganic Chemistry,
Polish Academy of Sciences, Poznan, Poland*

In variance with small-molecule crystallography, where structure refinement is highly overdetermined by a large surplus of diffraction data, the observation/parameter ratio in macromolecular crystallography is usually low and only at about 2.7-2.5 Å resolution reaches 1.0 for reasonably constructed models. It is, therefore, not only useful but mathematically necessary to use stereochemical restraints in macromolecular refinement at low resolution. At higher resolution the restraints are still used, for a variety of reasons, but at some point, usually when atomic resolution (1.2 Å) has been reached, they may be relaxed as the diffraction terms are "taking the refinement over". It is a valid question, however, if the restraints could be dropped altogether, and if yes, under what conditions. A separate question concerns the restraint targets themselves and the strictness with which they should be obeyed. The most popular library of stereochemical standards was compiled in 1991 by Engh and Huber from careful analysis of small-molecule structures available at that time in the Cambridge Structural Database (CSD). A survey of the entries deposited in the Protein Data Bank (PDB) indicates that often the models are forced to imitate the standards more closely than justified by the errors with which those standards were originally estimated. Additionally, with a nearly six-fold expansion of the CSD from 80,000 entries in 1991, it might be interesting to see if the "old" stereochemical standards are still valid. An even more interesting possibility is opened up by the explosive growth of the PDB (100-fold since 1990), now holding more than 52,000 entries, and especially by the rapid accumulation of ultra-high resolution protein structures. For example, 0.8 Å resolution macromolecular structures were unknown in 1997, while today there are about two dozen of them. Such structures are usually refined with utmost care and are only minimally "contaminated" by the prior knowledge enforced by stereochemical restraints. They offer, therefore, a unique possibility to review (and if necessary to adjust) the stereochemical standards of protein structure. A preliminary analysis indicates that while some of the "old" standards have withstood the test of time, some others might need small but clear adjustments. This is especially true of the peptide group, which can show higher deviations (up to 20°) from strict planarity than allowed by the restraints. It is also important conceptually that we are now able to obtain "protein parameters from proteins". With the currently attainable level of accuracy, one can investigate if there are any detectable idiosyncrasies of protein structure, related for instance to the specific nature of protein conformation or to specific interactions with the environment. Finally, the parameters derived from ultra-high resolution protein structures can serve not only as more appropriate stereochemical targets for model refinement but may also be used as validation criteria for lower resolution models.