Zero Ramachandran outliers does not guarantee a “good” model

O Sobolev[1], P Afonine[1], N Moriarty[1], M Hekkelman[2, 3], R Joosten[2, 3], A Perrakis[2, 3], P Adams[1, 4]
[1]Molecular Biosciences and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [2]Division of Biochemistry, The Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam 1066CX, The Netherlands, [3]Oncode Institute, The Netherlands [4]Department of Bioengineering, University of California, Berkeley, CA 94720, USA
osobolev@lbl.gov

Ramachandran plots report the distribution of the (φ, ψ) torsion angles of the protein backbone and are one of the best quality metrics of experimental structure models. Typically, validation software reports the number of residues belonging to "outlier", "allowed" and "favored" regions. While "zero unexplained outliers" can be considered the current "gold standard", this can be misleading if deviations from expected distributions, even within the favored region, are not considered. We therefore revisited the Ramachandran Z-score (Rama-Z), a quality metric introduced more than two decades ago, but underutilized. We describe a re-implementation of the Rama-Z score in the Computational Crystallography Toolbox along with a new algorithm to estimate its uncertainty for individual models; final implementations are available both in Phenix and in PDB-REDO. We discuss the interpretation of the Rama-Z score and advocate including it in the validation reports provided by the Protein Data Bank. We also advocate reporting it alongside the outlier/allowed/favored counts in structural publications.

Figure shows examples of Ramachandran plots: Left: a good-looking Ramachandran plot for (1ix9, 0.9 Å), Middle: an obviously bad Ramachandran plot (5a9z, 4.7 Å) and Right: a suspicious Ramachandran plot (6dzv, 4.2 Å). PDB ID code of the models in top right corner. Triplets of numbers on the bottom right on each plot indicate, from top to bottom: percentage of residues in favored and outlier regions, Rama-Z.

This research was supported by the NIH (grant GM063210), the Phenix Industrial Consortium, and by the Netherlands Organization for Scientific Research (NWO; Vidi grant 723.013.003). This work was partially supported by the US Department of Energy under Contract DE-AC02-05CH11231.

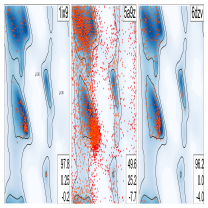

Figure 1