The Q|R Project: Quantum-Based Refinement of Biomacromolecules

Malgorzata Biczysko[1], Nigel W. Moriarty[2], Holger Kruse[3], Mark P. Waller[4], Pavel V. Afonine[2]

[1] Shanghai University, Shanghai, 200444, China
[2]MBIB Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8235, USA
[3]Institute of Biophysics of the Czech Academy of Sciences, Brno, 61265, Czech Republic
[4]Pending AI Pty Ltd., Eveleigh, NSW 2015, Australia
biczysko@i.shu.educ.cn

Protein structure determination is largely reliant on crystallography (X-ray, neutron or electron), electron cryo-microscopy (Cryo-EM) or NMR experiments. Refinement is the final step in obtaining accurate three-dimensional atomic model based on experimental data. Since the quality of the data (e.g., resolution) is rarely sufficient to utilize these data alone, this step has traditionally relied on parameterized libraries that describe stereochemistry of the molecules in question. The libraries used in major refinement packages do not describe unusual local arrangements of protein residues in the Ramachandran space, novel ligands, or non-covalent interactions such as π stacking, halogen, hydrogen or salt bridges. In particular, structures obtained using low-to-medium resolution data are biased by simple harmonic geometry restraints derived from these libraries. Quantum chemical computations can yield accurate geometries for standard protein or RNA/DNA molecules as well as novel ligands.

The methods we are developing in the Q|R project [1-4], an open-source software package (http://github.com/qrefine), combine experimental data with chemical restraints derived from quantum-chemical computations [1-5]. These procedures allow at present rather straightforward quantum refinement of proteins containing amino acids and organic ligands, in conjunction to both X-ray crystallography and Cryo-EM experiments. Replacing standard restraints by the ones derived from QM has shown to significantly improve model geometry, considering both the overall aspects of model and model-to-data fit statistics, as well as specific detailed structural features, in particular the hydrogen bonding. However, extension of the Q|R procedures to very large proteins, and its routine use by the community still poses challenges on the employed algorithms and QM-based computations, which will be discussed in this talk.

Figure 1. Q|R: blend of expertise and software tools

References

[1] M. Zheng, J. R. Reimers, M. P. Waller, P. V. Afonine, *Acta Cryst. D* 73. (2017). 45.

[2] M. Zheng, N. W. Moriarty, Y. Xu, J. R. Reimers, P. V. Afonine, M. P. Waller, *Acta Cryst. D* 73. (2017). 1020.

[3] M. Zheng, M. Biczysko, Y. Xu, N. W. Moriarty, H. Kruse, A. Urzhumtsev, M. P. Waller, P. V. Afonine, *Acta Cryst. D* 76. (2020). 41.

[4] L. Wang, H. Kruse, O. V. Sobolev, N. W. Moriarty, M. P. Waller, P. V. Afonine, M. Biczysko, *Acta Cryst. D* 76 (2020). 1184.

[5] Y. Liu, M. Biczysko, N. W. Moriarty, *Acta Cryst. D* 78 (2022). 43.