MD-assisted refinement of x-ray coordinates

Oleg Mikhailovskii [a, b], Yi Xue [a, c], Nikolai R. Skrynnikov [a, b]

[a] Purdue University, West Lafayette, Indiana, USA
[b] St. Petersburg State University, St. Petersburg, Russia
[c] Tsinghua University, Beijing, China
E-mails: omikhail@purdue.edu, yixue@mail.tsinghua.edu.cn, nikolai@purdue.edu

We have developed a software solution to refine crystallographic protein structures using experimental diffraction data in conjunction with state-of-the-art MD modeling setup. For this purpose, a special module to calculate structure factors has been implemented within AMBER16 [1] biomolecular simulations package. The refinement is conducted in a form of a short MD simulation that models the entire unit crystal cell complete with interstitial solvent. The structure-factor-based restraints are imposed in a form of maximum likelihood potential [2], which is balanced against the regular force-field potential.

To evaluate the performance of the new refinement protocol, we tested it on a set of 84 protein structures. As a first step, we conducted the refinement of relatively poor initial structural models (average deviation from the target 0.75 Å). It has been found that in 83% of the cases our protocol outperformed the refinement procedures available in Phenix [3] as judged by the lower resultant R_{free} factors. Furthermore, our protocol consistently led to better geometries, as indicated by superior MolProbity scores [4]. As a second step, we refined the coordinates that are deposited in the PDB databank. We found that in 38% of the cases our protocol achieved improvements over both the original PDB depositions and the re-refined variants thereof obtained by Phenix. Of note, the results from Phenix equipped with the advanced Amber ff14SB force field [5] have also been included in this comparison.

The approach to structure refinement demonstrated in our work has a number of broad advantages. It explicitly models protein-protein interactions in the crystal, as well as protein-water interactions. It also offers a natural way to represent the protein conformational dynamics: multiple protein molecules in the simulated crystal cell (or block of cells) sample different local conformations similar to the actual crystal. Furthermore, our method is well suited to recover those portions of the protein structure that diffract poorly due to high mobility, such as flexible loops and termini. In this case, the algorithm is mainly guided by the force-field potential, although at the same time it automatically utilizes the (limited) information contained in the diffraction data. Finally, it should be emphasized that our protocol has no tunable parameters and the calculations can be conducted in a matter of several hours on the desktop computers equipped with graphical processing units (GPU). A simple web interface has been implemented to support the use of the program by remote users.

References

[1] DA Case et al., “AMBER 16”, University of California, San Francisco, 2016.

[2] VY Lunin, PV Afonine, AG Urzhumtsev, “Likelihood-based refinement. I. Irremovable model errors,” *Acta Crystallogr. Sect. A*, vol. 58, pp. 270–282, 2002.

[3] PV Afonine *et al*., “Towards automated crystallographic structure refinement with phenix.refine,” *Acta Crystallogr. Sect. D*, vol. 68, pp. 352–367, 2012.

[4] VB Chen *et al*., “MolProbity: All-atom structure validation for macromolecular crystallography,” *Acta Crystallogr. Sect. D*, vol. 66, pp. 12–21, 2010.

[5] DA Case, PA Janowski, NW Moriarty, JM Swails, PD Adams, “Improved chemistry restraints for crystallographic refinement by integrating Amber molecular mechanics in Phenix,” *Acta Crystallogr. Sect. A*, vol. 74, p. a145, 2018.