In some cases more complicated approaches to refinement of macromolecular structures are not necessary

Zbigniew Dauter and Alexander Wlodawer*

Center for Structural Biology, Center for Cancer Research, National Cancer Institute, Frederick, Maryland, USA. *Correspondence e-mail: wlodawer@nih.gov

In the manuscript published in *IUCrJ* under the title 'Modeling a unit cell: crystallographic refinement procedure using the biomolecular MD simulation platform *Amber*' (Mikhailovskii *et al.*, 2022), the authors described the results of the refinement of 84 protein structures using a novel approach that utilizes molecular dynamics and the maximum-likelihood potential that encodes the structure-factor based restraints. They claimed the procedure was superior to traditional approaches used for refinement of macromolecular structures, exemplified by programs such as *Refmac5* (Murshudov *et al.*, 2011) or *Phenix* (Adams *et al.*, 2010). In our opinion, the only example for which the re-refined structure was deposited in the Protein Data Bank (PDB) does not provide convincing proof for the superiority of this refinement method.

The structure chosen as an example represented Type III antifreeze protein isoform HPLC12 (PDB entry 2msi; DeLuca *et al.*, 1998). The structure was originally refined at 1.9 Å resolution, but it appears that the refinement was not finalized. In particular, the deposited coordinates lacked any solvent that should have been visible at this resolution, thus despite good validation statistics (Table 1), that structure has to be considered of uncertain quality.

In the procedure used by Mikhailovskii *et al.* (2022), the original unit cell that contained a single protein molecule in the asymmetric unit in the space group $P2_12_12_1$ was expanded to $P1$ with identical unit-cell parameters, but containing four independent molecules. Thus, the number of refinable parameters was significantly increased, though the amount of available diffraction data remained the same. Their re-refined structure consisted of 1952 protein atoms and 155 water molecules (Table 1). The resulting model (PDB entry 7q3v) corrected obvious mistracing at the N-terminus and added missing solvent. As described by the authors, this was accomplished after very extensive calculations that could only be done on a computer equipped with GPUs.

To put this effort into perspective, we re-refined the structure 2msi with *Refmac5*, after manually correcting the erroneous tracing of the N-terminus and several improper side-chain rotamers with the program *Coot* (Emsley *et al.*, 2010). Waters were added automatically in *Coot* on the basis of 3.5σ peaks in the $F_o - F_c$ electron density map. Statistics of the resulting structure (PDB entry 9cbe), some obtained with *MolProbity* (Chen *et al.*, 2010), are also shown in Table 1. The process took less than an hour of human time and negligible computing time on a standard PC.

In principle, the more adjusted parameters, the better the agreement between the observed and calculated functions. However, if more parameters do not lead to better agreement, the strict Ockham razor should be mercilessly applied to avoid unnecessary complication of the problem. It is clear that, at least in this case, the complicated
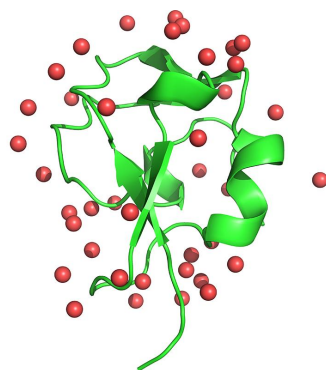
2msi => 7q3v => 9cbe

Table 1
Selected statistics of the three models of Type III antifreeze protein isoform HPLC12.

	Original structure (PDB entry 2msi)	*Amber*-refined (PDB entry 7q3v)	*Refmac5*-refined (PDB entry 9cbe)
Protein atoms (non-hydrogen)	485	1952	495
Solvent atoms	0	155	39
R/R_{free}	0.193/0.261	0.160/0.194	0.152/0.189
Clashscore	3.99 (98th percentile)	4.45 (97th percentile)	1.94 (100th percentile)
MolProbity score	1.61 (92nd percentile)	1.64 (90th percentile)	1.37 (98th percentile)
Bad bonds	0	0	0
Bad angles	8	16	1

procedure involving molecular dynamics did not work any better than manual refitting and standard crystallographic refinement. Although the procedure utilized by Mikhailovskii *et al.* (2022) might be potentially very useful in some specific cases, the sole example out of the 84 that were deposited in the PDB was maybe not an optimal selection. A more difficult or complicated case might have provided a clearer indication of the usefulness of this novel structure refinement procedure.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

DeLuca, C. I., Davies, P. L., Ye, Q. & Jia, Z. (1998). *J. Mol. Biol.* **275**, 515–525.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Mikhailovskii, O., Xue, Y. & Skrynnikov, N. R. (2022). *IUCrJ*, **9**, 114–133.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.