Finding the optimal resolution cutoff with *PAIREF*

M. Malý[1,2], K. Diederichs[3], J. Stránský [2], K. Adámková [2,4], T. Skálová[2], J. Dohnálek[2], P. Kolenko[1,2]

[1]Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, [2]Institute of Biotechnology of the Czech Academy of Sciences, Biocev, [3]Univ. of Konstanz, [4]Univ. of Chemical and Technology Prague, Dept of Biochemistry and Microbiology
martin.maly@fjfi.cvut.cz

The decision on the high-resolution cutoff has an apparent impact on the quality of a structure model. To determine the optimal cutoff automatically, we developed a software tool *PAIREF* [1]. The program performs the *paired refinement* protocol that allows linking the data and structure model quality. This analysis goes beyond the conventional criteria based on the indicators of data quality only (*e.g.* $I/\sigma(I)$, R_{meas}).

PAIREF is freely available for multiple platforms and can be run from the command-line or graphical user interface. Two refinement engines are currently supported: *REFMAC*5 from the *CCP*4 software suite [2] and *PHENIX.REFINE* [3]. The program creates a compact comprehensive report. The final decision on the cutoff is based on several statistics that are calculated and monitored: *R*-values, correlation coefficients, optical resolution, merging statistics, *etc*. The consequent comparison between CC_{work} and CC* allows the assessment of overfitting. Moreover, a unique feature of the program is the complete cross-validation scenario: the protocol is run in parallel for each free-reflection set selection individually which leads to averaged, more general and meaningful results.

During the work on *PAIREF*, we confirmed previous findings and proved that useful signal can be often still present in the high-resolution data not fulfilling the obsolete conventional criteria. To give an example: In the particular case of interferon gamma from *Paralichthys olivaceus* (PDB entry 6f1e), the cutoff was originally applied at 2.3 Å, according to the criterion for $I/\sigma(I)$ higher than 2 in the highest resolution shell. Nevertheless, we ran paired refinement up to 1.9 Å and observed a systematic decrease in R_{free} while including data up to 2.0 Å [1]. Hence, the structure was improved, despite very poor statistics relating to the last resolution shell 2.1-2.0 Å ($I/\sigma(I)$ = 0.1, $CC_{1/2}$ = 0.03).

Furthermore, we similarly examined the high-resolution data from endothiapepsin (PDB entry 4y4g). This structure was originally solved at 1.44 Å resolution. However, we could observe a significant improvement in the quality of electron density of the partially occupied fragment after refinement up to 1.20 Å (Fig. 1). This observation was in harmony with corresponding drops in R_{free} [1].

Generally, the quality of a structure model can benefit from the involvement of even weak high-resolution data. Thus, the application of paired refinement could be recommended for any structural project in X-ray macromolecular crystallography. *PAIREF* provides automation of the routine and gives all the relevant statistics for users to make a precise decision on the cutoff.

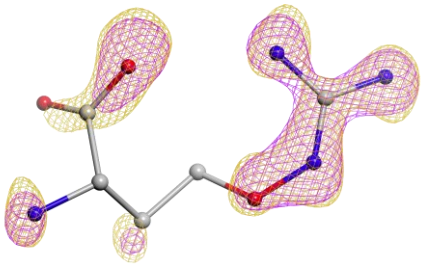

Figure 1. Improvement in omit maps of the partially occupied fragment B53. Electron density after refinement up to 1.44 (purple) and 1.20 Å (orange) is shown at a level of 0.56 eÅ^{-3}. Atomic coordinates were adapted from PDB entry 4y4g.

[1] Malý, M., Diederichs, K., Dohnálek, J. & Kolenko, P. (2020). *IUCrJ* **7**, pp. 681–692.

[2] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). *Acta Cryst. D* **67**, pp. 235–242.

[3] 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**, pp. 213–221.

Keywords: resolution; software; automation; refinement; macromolecular crystallography; IUCr2020; abstracts

This work was supported by the MEYS CR (projects CAAS – CZ.02.1.01/0.0/0.0/16_019/0000778 and BIOCEV – CZ.1.05/1.1.00/02.0109) from the ERDF fund, by the Czech science foundation (project 18-10687S), and by the GA CTU in Prague (SGS19/189/OHK4/3T/14).

Acta Cryst. (2021), A**77**, C88