

MS01 MX/Cryo-EM software development

MS1-03

Automatic decision on optimal resolution cut-off with PAIREF

M. Mař, K. Diederichs², J. Stránský³, K. Adámková³, J. Dohnálek³, P. Kolenko¹

¹Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering - Prague (Czech Republic), ²University of Konstanz - Konstanz (Germany), ³Institute of Biotechnology of the Czech Academy of Sciences, Biocev - Vestec near Prague (Czech Republic)

Abstract

High-resolution cut-off is an often carelessly estimated parameter during diffraction data processing despite its apparent impact on the quality of a solved structure model. Nevertheless, the paired refinement protocol has been shown to be a 'golden standard' for the determination of the cut-off. To provide this procedure to be run easily and effortlessly, we developed its automation – program *PAIREF* [1]. The resulting comprehensive analysis allows linking the data and structure model quality.

PAIREF has been included in the *CCP4 Suite* and can be run from its graphical user interface or the command line. For refinement, two engines are supported: *REFMAC5* and *Phenix.refine* [2].

Recently, we developed a new feature: an automatic interpretation of the calculated results that suggests an optimal resolution cut-off. The decision-making algorithm takes into account several statistics: overall R_{free} and their trends, R -values from high-resolution shells, CC_{work} and CC^* . This allows *PAIREF* to be involved in automated data-processing pipelines.

The automatic evaluation of results can be illustrated on a particular example of interferon gamma from *Paralichthys olivaceus* (PDB entry 6f1e). This structure was originally solved at 2.3 Å resolution [3]. We ran paired refinement with an increment of 0.1 Å up to the resolution of 1.9 Å using *PAIREF*; the interpretation of results is shown in Figure 1. The data in resolution shells 2.3-2.2 Å and 2.2-2.1 Å satisfy all the criteria as the overall R_{free} -values have a decreasing trend. However, in the next shell 2.1-2.0 Å, high-resolution R -values are close to exceeding an R -value of a perfect model against random data that is approximately 0.42 [4]. Thus, a warning sign is displayed for this shell. Finally, the last shell 2.0-1.9 Å does not comply with several criteria. To conclude, a strict cut-off of 2.1 Å and a permissive cut-off of 2.0 Å are suggested by the program. This interpretation is in good agreement with our previously published results [1].

References

- [1] Mař, M., Diederichs, K., Dohnálek, J., Kolenko, P. (2020). *IUCrJ* **7**, 681–692.
- [2] Mař, M., Diederichs, K., Dohnálek, J., Kolenko, P. (2021). *Acta Cryst. F77*, 226–229.
- [3] Zahradník, J., Kolářová, L., Pařzková, H., Kolenko, P., Schneider, B. (2018). *Fish Shellfish Immunol.* **79**, 140–152.
- [4] Evans, P. R., Murshudov, G. N. (2013). *Acta Cryst. D69*, 1204–1214.

The decision-making algorithm of *PAIREF* results in

Shell	Accepted?	Reason
2.30-2.20 Å	Yes	Overall R_{free} decreased while using data in the shell 2.30-2.20 Å
2.20-2.10 Å	Yes	Overall R_{free} decreased while using data in the shell 2.20-2.10 Å
2.10-2.00 Å	Warning	Overall R_{free} decreased while using data in the shell 2.10-2.00 Å R_{free} in high resolution is higher than 0.40 while using data in the shell 2.10-2.00 Å R_{work} in high resolution is higher than 0.40 while using data in the shell 2.10-2.00 Å
2.00-1.90 Å	No	Overall R_{work} increased and R_{free} remained constant while using data in the shell 2.00-1.90 Å R_{free} in high resolution is higher than 0.40 while using data in the shell 2.00-1.90 Å R_{work} in high resolution is higher than 0.40 while using data in the shell 2.00-1.90 Å $CC1/2$ in high resolution is negative or undefined while using data in the shell 2.00-1.90 Å But statistics deteriorate in a previous resolution shell.

Acknowledgements

This work was supported by the MEYS CR (projects CAAS - CZ.02.1.01/0.0/0.0/16_019/0000778 and ELIBIO - CZ.02.1.01/0.0/0.0/15_003/0000447) from the ERDF fund; by the AS CR (86652036); by the GA CTU in Prague (SGS22/114/OHK4/2T/14); and from the grant of Specific University Research (A1_FPBT_2021_003).