Validating Protein Structures in the AlphaFold 3 Era: A Contact Map-Based Approach with conkit-validate

A. J. Simpkin[1], R. M. Keegan[2], D. J. Rigden[1]

[1]Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 7ZB, United Kingdom
UKRI–STFC, Rutherford Appleton Laboratory, Research Complex at Harwell, Didcot OX11 0FA, United Kingdom

hlasimpk@liverpool.ac.uk

Reliable structural data in the Protein Data Bank (PDB) is essential for a wide range of biological and biomedical applications. Nonetheless, errors can still occur, particularly in models derived from low-resolution experimental data. To help address this, we previously developed a validation method that highlights inconsistencies between the residue-residue contacts and distances observed in a model and those predicted by state-of-the-art computational tools, specifically AlphaFold 2 [1].

In this work, we extend that approach by removing the requirement for distograms and instead derive contact maps directly from predicted atomic coordinates. This broadens compatibility with newer prediction methods such as AlphaFold 3 [2], Boltz-1 [3], and Chai-1 [4], and makes the method more accessible to users, as only a predicted model is required.

[1] Sánchez Rodríguez, F., Simpkin, A. J., Chojnowski, G., Keegan, R M. & Rigden, D. J. (2024). *IUCrJ*. **11**(6), 938-950.

[2] Abramson, J. et al. *(*2024*) Nature,* **630**, 493–500.

[3] Wohlwend, J. et al. *(*2024*) bioRxiv,*

[4] Discovery, C. et al. *(*2024*). bioRxiv.*