Oral presentation

Error estimates in atom coordinates and B factors in macromolecular crystallography

J.R.Helliwell[1]

[1]*Department of Chemistry, University of Manchester, Manchester, M13 9PL,*

john.helliwell@manchester.ac.uk

The overall diffraction precision index (DPI) of a biological macromolecule crystal structure was first described by Durward Cruickshank in 1999 [1]. Additional developments were introduced by the availability of a webserver http://cluster.physics.iisc.ernet.in/dpi/ , which provides a transformed PDB entry with individual atom coordinate errors derived from applying the DPI method, using the parameters provided by the authors, and then subsequently added to the PDB file. This webserver has been extensively used and harnessed in describing non-covalent distance error estimates as well as assessing the significance, or otherwise, of atom movements in a variety of studies. The standard uncertainties on a biological macromolecule's atomic displacement parameters (the 'B factors') has been an entirely different challenge but their absence is obviously important since assuming a false precision can convey a false certainty in the dynamics of a structure. A method involving parallelisation of workflows for raw diffraction image data processing does however offer estimates of the precision of B factors, provided the raw diffraction images are archived. These workflows are routinely provided at the synchrotron radiation macromolecular crystallography beamlines these days. So, it would be straightforward to implement an artificial intelligence/machine learning approach to explore the variance in protein molecular models from the different possible stages of a parallelised workflow of multiple softwares. It would be necessary for the researcher to obtain their preferred molecular model with their preferred workflow before submitting it into the AI/ML workflows procedure. Before estimating the uncertainties, checks must firstly be made that any gross errors of atom misidentification, or inclusion or exclusion when unwarranted by the evidence, are not present. Likewise, an incorrect choice of diffraction resolution limit should have been avoided. With the investment in X-ray laser and synchrotron radiation facilities in general, and in the current era especially for dynamic crystallography in the study of smaller movements on ever faster timescales, it is a major gap if uncertainties on distances and movements derived from individual atom coordinate uncertainties are not provided. In terms of basics is the core point that, as the IUCr Working Party on Expression of Uncertainty in Measurement reported (Schwarzenbach et al (1995)), *"a measurement result is considered complete only when accompanied by a quantitative statement of its uncertainty"*. I describe these matters in detail in my recent article [3].

Keywords: Coordinates; Atomic displacement parameters; B factors; Precision estimates

[1] Cruickshank, D. W. J. (1999) Acta Cryst. D55, 583-601.

[2] Schwarzenbach, D., Abrahams, S. C., Flack, H. D., Prince, E. & Wilson, A. J. C. (1995) Acta Cryst. A51, 565–569.

[3] Helliwell, J.R. (2023) Current Research in Structural Biology, 6, 100111, ISSN 2665-928X,

https://doi.org/10.1016/j.crstbi.2023.100111.

(https://www.sciencedirect.com/science/article/pii/S2665928X2300017X) .