

Towards archiving raw diffraction images for validating crystal structures.

Loes M.J. Kroon-Batenburg¹

¹*Crystal And Structural Chemistry, Utrecht University, Utrecht, Netherlands*

E-mail: l.m.j.kroon-batenburg@uu.nl

The importance of retaining and making available the experimental data to support scientific results is becoming increasingly evident to policy makers, funders and scientists themselves. The IUCr established a Diffraction Data Deposition Working Group (DDDWG) with the aim of developing standards for the representation of raw diffraction data in crystallography. Whilst discussions are actively going on about what data to archive (i.e. either only those related to published papers, or those associated with incomplete or unsuccessful research that could be particularly interesting for the development of new science), the field should prepare itself for depositing fully self-contained data. Two key issues play a role: the importance of persistent identifiers (such as the standard registered digital object identifier, DOI) and the full recording of metadata. A recent review [1] summarizes the ongoing developments.

The retention of raw data can be seen as complementing the extensive archives of derived data and processed data (structure factors) in the crystallographic databases. Until the advent of CIF and automated structure validation checks with e.g. the checkCIF suite [2] many structures were published which required subsequent correction. The availability of the raw data (i.e. original diffraction images) can enhance structure validation: is the indexing correct? is the lattice symmetry correct? was a second fragment missed (pseudo merohedral twinning)? were satellite reflections missed (modulated structure)? is data reduction done correctly (polarization, absorption, anisotropic scaling)? was diffuse scattering apparent but ignored? The usefulness of having raw data available for these various aspects will be discussed. Referees of a submitted article need access to the underlying data if they are to endorse these data as 'versions of record', and this access can include raw data if a referee deems it necessary.

[1] Kroon-Batenburg, L.M.J., Helliwell, J.R., McMahon, B. & Terwilliger, T.C. (2017). IUCrJ 4, 1–13.

[2] Strickland, P. R., Hoyland, M. A. & McMahon, B. (2005). International Tables for Crystallography, Vol. G, pp. 557–569.

Keywords: [Raw data archiving, validation, metadata](#)