

MS46 Reproducibility in crystallography

MS46-04

Reproducibility and the CSD

N.T. Johnson¹, **S.C. Wiggin**¹, **I.J. Bruno**¹, **S.C. Ward**¹

¹*Cambridge Crystallographic Data Centre - Cambridge (United Kingdom)*

Abstract

The Cambridge Structural Database (CSD) is a powerful resource for scientists around the world, containing over 1.1 million small molecule organic and metal-organic crystal structures.¹ As a repository for experimental structural data, it has an important role to play in helping the community set and maintain good data practises for sharing crystallographic data to help ensure that the data stored in the CSD is FAIR.²

Although FAIR data is important, we recognise that this does not completely capture our responsibility as a data repository. Alongside other factors, like securing the sustainability of our resource, we need to play our role to help ensure data is not only reusable but also reproducible. The CCDC's web deposition service encourages users to provide structure factor information, share a DOI to any raw experimental data, as well as supply additional experimental information (such as the crystallisation details or any properties of the material) to help capture as much metadata as possible. Through our education and outreach work, we aim to help scientists understand the importance of providing rich metadata and how it benefits them and the wider academic community.

We will highlight some of the steps we have taken to promote the reproducibility of experimental crystal structure data and beyond. We will also reflect on what we can learn from the CSD to help us on this journey and conclude by exploring some of the challenges and opportunities in striving towards the reproducibility of data.

References

1. Groom, C., Bruno, I., Lightfoot, M., & Ward, S. (2016). *Acta Crystallographica Section B Structural Science, Crystal Engineering And Materials*. 72, 171-179.
2. Helliwell, J. (2019). *Structural Dynamics*. 6, 054306.