



IUCrData

ISSN 2414-3146

Received 27 May 2026

Accepted 18 August 2026

Edited by S. Coles, University of Southampton, United Kingdom

Keywords: raw diffraction data archiving; community data metrics; raw data reprocessing; electron-density model fitting.

Structural data: full structural data are available from iucrdata.iucr.org

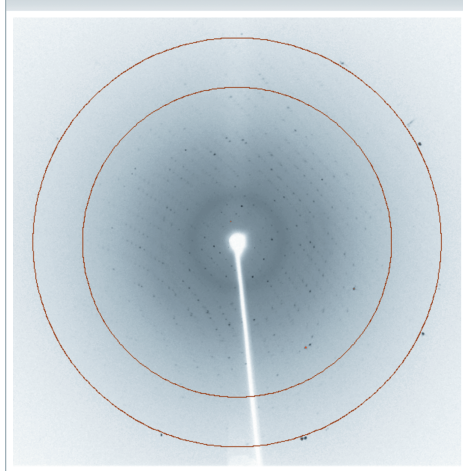
Developed community standards on setting the diffraction resolution yields improved density maps and a new monosaccharide model in *pdb_00004dda*

John R. Helliwell,^{a*} Robbie P. Joosten,^b Loes Kroon Batenburg^c and Simon Tanley^a

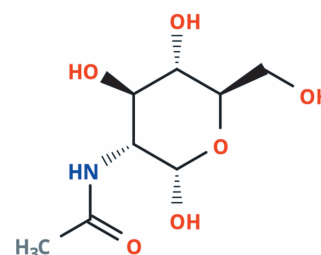
^aDepartment of Chemistry, University of Manchester, Manchester, M13 0PL, United Kingdom, ^bDivision of Biochemistry, Netherlands Cancer Institute, 1066CX Amsterdam, The Netherlands, and ^cUtrecht University, 3584 CG Utrecht, The Netherlands. *Correspondence e-mail: john.helliwell@manchester.ac.uk

A systematic survey of monosaccharides in the PDB archive threw up a possible improvement in model fit in *pdb_00004dda*. Reprocessing of the publicly open archive of the raw diffraction images was made using *EVAl*. The diffraction resolution was assessed by the $CC_{1/2} = 0.5$ criterion and confirmed by paired model refinement implemented in *PDB-REDO*. The resulting change in diffraction resolution, from 2.4 Å to 2.0 Å, improved the clarity of the electron density at the monosaccharide binding site. The NAG (*N*-acetyl- β -D-glucosamine) was refitted as NDG (*N*-acetyl- α -D-glucosamine). This example illustrates the advantages of raw diffraction data archiving, as aimed at by FAIR data polices, allowing reprocessing as agreement on community metrics.

Raw data



Chemical scheme



Raw diffraction data and reprocessed data: <https://zenodo.org/records/18258544>

Raw diffraction images converted to fullCBF format: <https://zenodo.org/records/20326720>

Metadata imgCIF file: <https://doi.org/10.1107/S2414314626008515/ii4005img.cif>

Introduction

A systematic survey of monosaccharides in the PDB archive threw up a possible improvement in model fit in *pdb_00004dda*. The structure was published in Tanley *et al.* (2012) with a diffraction resolution of $\langle I/\sigma I \rangle = 4.0$. The ways to establish the diffraction

Table 1
Experimental details.

Raw Data	
DOI	https://doi.org/10.5281/zenodo.18258544 https://doi.org/10.5281/zenodo.20326720
Data archive	Zenodo
Data Format	OSC and CBF
Data Collection	
Diffractometer	Rigaku Micromax-007 generator
Detector type	R-Axis IV image plate
Radiation type	Cu K α
Wavelength (Å)	1.5418
Beam centre (mm) (fast, slow)	150.00, −149.89
Detector axis	−Z
Detector distance (mm)	140.00†
Pixel size (mm)	0.100 × 0.100
No. of pixels	3000 × 3000
No. of scans	1
Scan axis	ω , χ
Start angle, increment per frame (°)	0.0, 1.0
Scan range (°)	180.0
No. of frames	180
Exposure time per frame (s)	4.5

The detector distance in the raw image files is given as 140.0 mm, while after indexing and refinement it is approximately 135 mm.

resolution have since then developed into community-agreed improved criteria. Reprocessing of the publicly open archive of the raw diffraction images was made using *EVAL* (Schreurs *et al.* 2010). The diffraction resolution was assessed by the $CC_{1/2} = 0.5$ criterion (Karplus & Diederichs 2012) and confirmed by paired-model refinement using *Refmac2* (Yamashita *et al.* 2023) implemented in *PDB-REDO* (Joosten *et al.* 2014). The resulting change in diffraction resolution, from 2.4 Å to 2.0 Å, improved the clarity of the electron density at the monosaccharide binding site. The NAG (*N*-acetyl- β -D-glucosamine) was refitted as NDG (*N*-acetyl- α -D-glucosamine). The raw diffraction images for *pdb_00004dda* in the original RAXIS.osc format, along with the reprocessing files, can be found at <https://zenodo.org/records/18258544>. The raw data were converted to fullCBF with the *make-cbf* tool as part of *imgCIF_Creator* (Kluyver *et al.*, 2026) and deposited to Zenodo (<https://zenodo.org/records/20326720>). This example illustrates the advantages of raw diffraction data archiving, as aimed at by FAIR data policies (Wilkinson *et al.*, 2016), allowing reprocessing as agreement on community metrics improve.

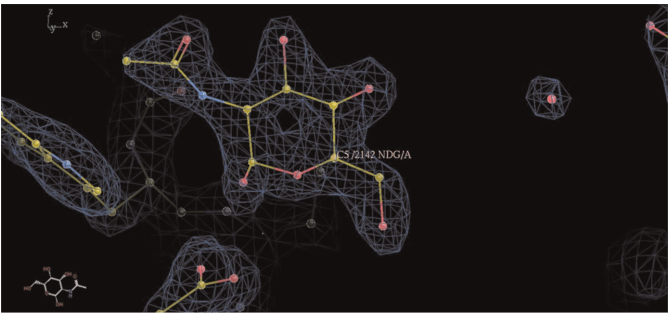


Figure 1
Electron density $2mF_o - DF_c$ (1.58 r.m.s.) with the monosaccharide NDG fitted (this study at 2.0 Å). The real-space correlation coefficient of NDG is 0.93. This figure was produced using *Coot* (Emsley *et al.*, 2010).

Table 2
4dda X-ray crystallographic data processed via *EVAL15* and refinement statistics for HEWL co-crystallized with *N*-acetyl-D-glucosamine (Tanley *et al.* 2012) and new data processing and refinement statistics undertaken by *PDB-REDO* with TLS refinements and hydrogens in riding positions.

Space group	<i>P4</i> ₃ <i>2</i> ₁ <i>2</i>	<i>P4</i> ₃ <i>2</i> ₁ <i>2</i>
Unit-cell parameters (Å)	$a = b = 78.37$, $c = 36.58$	$a = b = 78.37$, $c = 36.58$
Molecular mass (Da)	14700	14700
Molecules per asymmetric unit	1	1
Observed reflections	49543	93073 (10160)
Unique reflections	4120	7468 (796)
Resolution (Å)	19.59–2.40 (2.48–2.40)	19.59–2.0 (2.07–2.0)
Completeness (%)	85.3 (100)	91.5 (100.0)
R_{merge}	0.147 (0.607)	0.243 (1.960)
$\langle I/\sigma(I) \rangle$	13.6 (4.02)	8.15 (1.14)
$CC_{1/2}$	†	0.996 (0.508)
Multiplicity	12.1 (12.8)	12.5 (12.8)
Cruickshank DPI (Å)	0.806	0.287
Average Bfactor (Å ²)	29.5	21.9
Refinement $R_{\text{factor}}/R_{\text{free}}$ (%)	20.0/28.5	20.4/26.0
R.m.s. Z bonds	1.201	0.204
R.m.s. Z angles	0.967	0.439
Molprobability Clashscore	5.53	2.00
Ramachandran Z-score	−2.12	−0.06
Rotamer normality Z-score	−1.88	−1.11

Note 1. $CC_{1/2}$ was not a metric in use in 2012.

The benefits of raw diffraction data archiving linked to deposited PDB files were described with an example by the IUCr abstract (Helliwell *et al.*, 2023) as a case study within the Global Open Science Cloud (Chen *et al.* 2023). Their example documented a case of a good choice of diffraction limit by the authors (Sato *et al.* 2021) and PDB entry *pdb_00007ccy*. In the case of *pdb_00004dda* we can now illustrate below the corrective power and an improvement with one of our own studies.

Methods and results

Data were processed in the same way as in Tanley *et al.* (2012) and detailed in Tanley *et al.* (2013) by *EVAL* (Schreurs *et al.* 2010), which uses a diffraction data integration method based on *ab initio* predicted profiles, but only extending the resolution and with the summary statistics now including $CC_{1/2}$. The new diffraction resolution was chosen to be at a $CC_{1/2}$ criterion of 0.5 and confirmed by paired model refinement implemented in *PDB-REDO* (Joosten *et al.* 2014). The electron density at the monosaccharide binding site after the reprocessing and re-refinement is shown in Fig. 1. The raw data collection details are provided in Table 1. The data statistics of *pdb_00004dda* and of the reprocessed and re-refined models are given in Table 2. The metadata for the raw diffraction images have been automatically checked by the IUCrData checkcif-for-raw-data software (<https://publbio.iucr.org/imgcif/>).

Discussion and conclusions

As community metrics in macromolecular crystallography evolve, the processing of archived raw diffraction images can be successfully re-evaluated as shown in this case with an

improved diffraction resolution limit and model. The full range of situations suitable for a *Raw Data Letter* are described in Kroon-Batenburg *et al.* (2022).

Acknowledgements

We thank Jasmine Young of the RCSB for assistance. Simon Tanley was supported by an EPSRC PhD studentship in the Department of Chemistry, University of Manchester. We thank James Hester for writing a tool to convert RAXIS.osc images to fullCBF.

Data availability

The raw diffraction data and the reprocessed data to 2.0 Å using *EVAl* as well as the molecular model are available at <https://zenodo.org/records/18258544>. The raw diffraction images converted to fullCBF format are available at <https://zenodo.org/records/20326720>. The improved Protein Data Bank entry can be found at pdb_000030IG. The pdb_00004dda entry is now obsolete.

References

- Chen, Y., Zhang, L., Li, J., Hodson, S., Uhler, P. F., Li, X., Pergl, H., Haagstrom, I., Yue, X., Wang, J., Zhang, X., Helliwell, J. R., Kurisu, G., Miller, Yingchao Piao, Monthip Sriratana, Gensuo Jia, Bapon Fakhruddin, Zhang Ying, J., Bouwman, J., Liu, L., Maxwell, L., Oladipo, F., Sithole, H., Luo, Z. & Zhang, H. (2023). *The Global Open Science Cloud: Vision and Initial Successes*. GOSC IPO, 25 Aug 2023. <https://doi.org/10.5281/zenodo.8296517>.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Helliwell, J. R., Kurisu, G. & Kroon-Batenburg, L. (2023). *Acta Cryst. A* **79**, C822.
- Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ* **1**, 213–220.
- Karplus, P. A. & Diederichs, K. (2012). *Science* **336**, 1030–1033.
- Kluyver, T., Dall'Antonia, F., Hester, J. R. & Kroon-Batenburg, L. M. J. (2026). *imgCIF_Creator*. https://github.com/COMCIFS/imgCIF_Creator.
- Kroon-Batenburg, L. M. J., Helliwell, J. R. & Hester, J. R. (2022). *IUCrData* **7**, x220821.
- Sato, H., Sugishima, M., Tsukaguchi, M., Masuko, T., Iijima, M., Takano, M., Omata, Y., Hirabayashi, K., Wada, K., Hisaeda, Y. & Yamamoto, K. (2021). *Biochem. J.* **478**, 1023–1042.
- Schreurs, A. M. M., Xian, X. & Kroon-Batenburg, L. M. J. (2010). *J. Appl. Cryst.* **43**, 70–82.
- Tanley, S. W. M., Schreurs, A. M. M., Kroon-Batenburg, L. M. J., Meredith, J., Prendergast, R., Walsh, D., Bryant, P., Levy, C. & Helliwell, J. R. (2012). *Acta Cryst. D* **68**, 601–612.
- Tanley, S. W. M., Schreurs, A. M. M., Helliwell, J. R. & Kroon-Batenburg, L. M. J. (2013). *J. Appl. Cryst.* **46**, 108–119.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). *Sci. Data* **3**, 160018.
- Yamashita, K., Wojdyr, M., Long, F., Nicholls, R. A. & Murshudov, G. N. (2023). *Acta Cryst. D* **79**, 368–373.