A "post-mortem" analysis of radiation damage in the Protein Data Bank with the B_{net} metric

Kathryn Shelley[1,2], Elspeth Garman[2]

[1]*School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, United Kingdom;*
[2]*Dept of Biochemistry, University of Oxford, Oxford, OX1 3QU, United Kingdom*

elspeth.garman@bioch.ox.ac.uk

During macromolecular crystallography (MX) data collection one or more crystals are exposed to a high flux of ionising radiation, which even at cryo-temperatures can damage the crystal(s), causing structural and chemical changes. If sufficiently severe, this damage can prevent solution of the crystal structure. However, even when structure solution remains possible, specific damage artefacts can confuse the biological conclusions drawn from a structure: hence their identification is important. Traditionally however the detection of specific radiation damage artefacts within crystal structures has proven difficult.

To address this problem, previously the Garman group developed the B_{Damage} metric[1], calculated by the CCP4[2] program RABDAM[3]. B_{Damage} is a per-atom metric that highlights potential sites of specific radiation damage as atoms with high B-factor values as compared to other atoms in a similar local environment in the parent crystalline structure. Whilst this metric is useful at identifying damage artefacts in individual structures, unfortunately B_{Damage} values can not be compared between different structures. To address this limitation, here we present a derivative of the B_{Damage} metric, B_{net}, a per-structure metric that can be used to compare the relative damage suffered by different protein crystal structures. After validating that B_{net} is an appropriate metric on a dataset of structures known to contain specific damage artefacts, we use B_{net} to analyse the specific radiation damage present in a dataset of 94,145 protein crystal structures in the Protein Data Bank (PDB). Notably, many of the structures identified as damaged by B_{net}, and which on closer inspection contain obvious damage artefacts, have reasonable or excellent values for the metrics typically reported for PDB structures.

[1] Gerstel, M., Deane, C. M. & Garman, E. F. (2015). *J. Synchrotron Rad.* **22**, 201–212.

[2] Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

[3] Shelley, K. L., Dixon, T. P. E., Brooks-Bartlett, J. C. & Garman, E. F. (2018). *J. Appl. Cryst.* **51**, 552–559.

Keywords: Protein Data Bank, Radiation Damage, B factors, Bdamage