

## From fake news to new insights: showing what the anomalous signal and PIXE does to enhance metalloprotein biochemistry.

Elsbeth F. Garman<sup>a</sup>, Geoff W. Grime<sup>b</sup>, Edward H. Snell<sup>c</sup>

<sup>a</sup>Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK.

elsbeth.garman@bioch.ox.ac.uk

<sup>b</sup>Ion Beam Centre, Advanced Technology Institute, University of Surrey, Guildford, Surrey, GU2 7XH, UK. g.grime@surrey.ac.uk

<sup>c</sup>Hauptman-Woodward Medical Research Institute, 700 Ellicott St., Buffalo, NY 14203, USA.  
esnell@hwi.buffalo.edu

The Protein Data Bank (PDB) contains information on over 150,000 proteins. Of these, approximately 30,000 have been labelled by their depositors as metalloproteins, which are either molecules in which a small number of metal atoms define the folded three-dimensional shape (and hence function) or those which are involved in metal transport or metabolism. In virtually all metalloprotein entries, the identity of the metal atoms in the PDB entries has been inferred indirectly from the X-ray diffraction data or by other methods, including the presence and magnitude of the anomalous signal.

In the 1990s, two of the authors developed a method for unambiguously identifying and quantifying the metal atoms (to  $\pm 10\%$  on stoichiometric ratios) using microbeam Proton Induced X-ray Emission (PIXE) in combination with proton Rutherford Backscattering (RBS) [1, 2] which has since been used in many important studies of metalloproteins. Although PIXE allows the different metals in a protein to be identified, it does not give positional information on which metal site contains which metal atom.

We have recently automated the PIXE pipeline to permit the analysis of up to 100 liquid protein deposited in an array of small ( $\sim 60 \mu\text{m}$  diameter  $200 \mu\text{m}$  apart) spots using non-contact inkjet printing methods compared with 10 samples per day using our previous low throughput method. As a demonstration of this method we analysed 32 previously structurally characterized metalloproteins with the surprising result that over a third of the models deposited in the PDB had misidentified metals and another third contained additional metals [3]. In one of the former cases we used the experimental and theoretical anomalous signal ratios to position the 3 different metals detected by PIXE and rerefined the PDB deposited structure and this revealed new biological information.

This contribution will review the current state of our method and will discuss possible explanations for the misattributed metals. Crucially, we have demonstrated that using the correct metal in the metalloprotein structure allows the model fitted to the X-ray diffraction data to be refined and improved, so revealing previously unknown insights into biological function of the proteins studied.

[1] Garman, E. (1999) *Structure* **7**, R291-299.

[2] Garman, E.F. & Grime, G.W. (2005) *Progress in Biophysics and Molecular Biology* **89/2**, 173-205.

[3] Grime, G.W. *et al.* (submitted)