We've sorted out the data, now what was the sample again?

C. Neylon[1,2], J.G. Frey[2], A.J. Milsted[2] and J. Hale[2]

S.T.F.C. Rutherford Appleton Laboratory, Didcot, UK OX11 0QX,

School of Chemistry, University of Southampton, Highfield Southampton, UK SO17 1BJ,

C.Neylon@rl.ac.uk

Sharing and making available the raw data from experiments is becoming an increasingly important requirement for the scientific community. Government funding agencies are moving towards requiring open access to the results of publicly funded research and there is an increasing expectation that much more of the detail of experimental data will made available as supplementary information in the published literature. Standardised data formats can make experimental data from SAS experiments fully accessible to the interested community. However without access to minimal information on the sample this information can be of limited use. Details of sample concentration, monodispersity, and oligomerisation state are crucial in determining the reliability of conclusions made on the basis of the data. However, even where the scattering data is made publicly available or reported in the peer reviewed literature there is often limited information on sample preparation and characterisation made available.

We propose the meeting considers whether a minimal standard of reporting on samples should be adopted. If such a standard were adopted it could be applied either at the level of publication or at the level of facility access through a requirement for minimal reporting on samples. Applying a standard for publication could be achieved by adopting a community standard based approach similar to that used for the reporting of DNA microarray experiments. However this would only apply to published data. The advantages of mandating minimal sample description at the level of facility access is that all data would then be collected in one place and available for the future in a useful form allowing for re-analysis and re-use.

Recording data on sample preparation requires tools and standards and providing the data requires systems that can capture data and provide access on request. We are developing and using an electronic laboratory notebook based on a Blog format to capture experimental data in a biochemistry laboratory. Within the system each sample is recorded in a single post. Analysis and manipulations of the sample are recorded in separate posts with links back to the sample. All the information is made immediately available on the Web as it is recorded. Our objective is to capture all the relevant detail of our sample preparation and characterisation and to make this freely available, along with all the relevant experimental data, and any analysis of the data. While this approach will not be suitable for all disciplines, nor all scientists, it provides a model of how sample preparation data can be shared and made accessible for the future.