eCrystallographyDataReports:
An Open Archive Route for the Reporting and Dissemination of Crystal Structures

Simon J. Coles*[1], Jeremy G. Frey[1], Michael B. Hursthouse[1], Leslie A. Carr[2] and Christopher J. Gutteridge[2]
[1]School of Chemistry and [2]School of Electronics and Computer Science,
University of Southampton, Southampton, Hampshire SO17 1BJ, UK
(s.j.coles@soton.ac.uk)

Advances in crystallographic instrumentation and computational resources have caused an explosion of crystallographic data, as shown by the recent exponential growth of the Cambridge Structural Database (CSD) [1]. However, even this is considered to be lower than expected, following the introduction of area detection. The reason for this is clearly identified as a publication bottleneck [2], which can only become even more severe with developments in high throughput crystallography [3]. As a result of this situation, the user community is deprived of valuable information, and the funding bodies are getting a poor return for their investments!

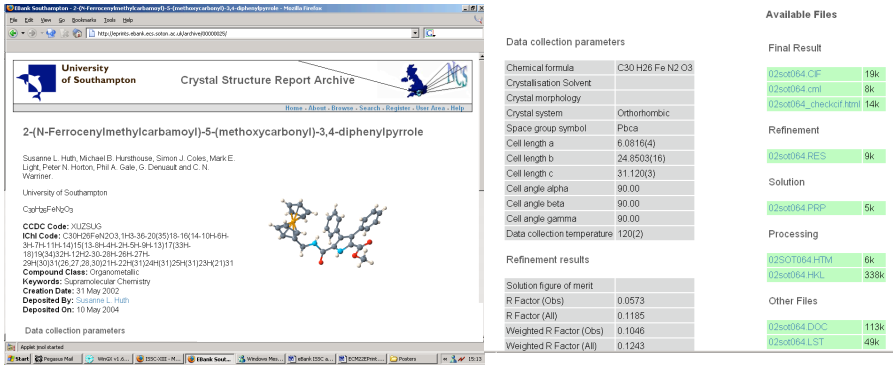
Unlike the mathematical and electronic sciences, the chemical sciences have been reluctant to embrace the 'preprint concept' [4]: the one exception has been the efforts of rapid electronic communications journals. This poster outlines a publication at source procedure for the rapid and effective dissemination of structural information to the scientific community which removes the lengthy peer review process that hampers traditional publication routes, but provides an alternative mechanism. eCrystallographyDataReports are built on a concept developed in the Computer Science community, the Open Archive Initiative (OAI) [5], whereby an author may reveal to the public archives of information. OAI software, known as ePrints [6], has been modified to implement this new type of open archive. An eCrystallographyDataReport makes available all raw, derived and results data from a crystallographic experiment via a searchable and hierarchical system. At the top searchable level this metadata includes bibliographic and chemical identifier items which allow access to a secondary level of crystallographic items which are directly linked to the associated archived data.

The 'core bibliographic data' is made available to the public domain by the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) [7] so that it may be 'harvested' by data aggregator services. This enables linking of the archive record with other, perhaps not structurally related, published literature. Perhaps more importantly it notifies those monitoring the archive that a new entry is available, hence the CSD may automatically download the structural data and assimilate it into the database. This will return a deposition number which, along with any future literature citations of this entry, can be retrospectively added to the record to link the data or information sources.

eCrystallographyDataReports are intended for the publication of data only. Hence the results of a crystal structure determination may be disseminated in a manner that anyone wishing to utilise the information may access the entire archive of data related to it and assess its validity and worth. The archive entry can then be referenced, which allows learned society journals to streamline articles to contain more intellectual discussion and less experimental reports. This archive software is to be freely available for public download and installation and is simple and inexpensive to set up and maintain.

Figure 1.

An eCrystallographyDataReport archive entry



References

1. Allen, F. H. *Crystallogr.Revs*. **2004**, *10 (1)*, 3-15. High Throughput Crystallography: The Challenge of Publishing, Storing and Using the Results.
2. Bond, A. D.; Davies, J.E. *Chemistry in Britain*. **2003**, 39, 44. Data Overload.
3. Hursthouse, M. B. *Crystallogr.Revs*. **2004**, *10(1)*, 85-96. High Throughput Chemical Crystallography (HTCC): Meeting and Greeting the Combichem Challenge.
4. Garson, L. R. *Acc.Chem.Res*. **2004**, *37(3)*, 141-148. Communicating Original Research in Chemistry and Related Sciences.
5. http://www.openarchives.org/
6. Gutteridge, C. *Proc. 11th Panhellenic Academic Libraries Conf*. **2002** (http://eprints.ecs.soton.ac.uk/archive/00006840/).
7. http://www.openarchives.org/OAI/openarchivesprotocol.html