

**P.24.01.1**

*Acta Cryst.* (2005). **A61**, C481

**COD (Crystallography Open Database) and PCOD (Predicted)**

Armel Le Bail, Daniel Chateigner, Xiaolong Chen, Marco Ciriotti, Lachlan M.D. Cranswick, Robert T. Downs, Luca Lutterotti, Alexandre F.T. Yokochi, *COD Advisory Board*. E-mail: cod@crystallography.net

The COD was created in March 2003 and was built on the PDB model of open access on the Internet. It is intended that this database [1] consist of any small or medium crystal structure (inorganic, organic, organometallic). Currently the total entry number is close to 15,000, including 6,600 entries from the American Mineralogist Crystal Structure Database (AMCSD) [2], and CIF files donations from a few laboratories in Europe or from individuals. The distribution is made through an Apache/MYSQL/PHP system that takes queries on chemistry, ranges of cell parameters, volumes, etc, as well as combination of fields, and can download or upload CIF files. A large donation of CIFs is anticipated from the IUCr.

The PCOD, created in December 2003, is a COD subset of crystal structures predicted by the GRINSP computer program [3]. It is growing fast and already contains > 1000 CIF files corresponding to  $M_2X_3$ ,  $MX_2$ ,  $MX_3$  or  $M_aM'_bX_c$  formulations ( $X = O, F$ ;  $M/M' = B, Na, Si, Al, P, Ca, V, Fe, Ga, Re, Zr$ , etc), including hypothetical zeolites and other binary compounds with N-connected 3D frameworks of M atoms ( $N = 3, 4, 5, 6$ ) as well as ternary compounds with mixed  $M/M'$  frameworks. The PCOD is open for search, download and upload of predicted crystal structures (coming from any prediction computer program, inorganic or small and medium organic molecules).

Crystallographers are invited to deposit their data as CIF files.

[1] <http://www.crystallography.net/> [2] <http://www.geo.arizona.edu/AMS/amcsd.php> [3] <http://www.cristal.org/grinsp/>

**Keywords:** crystal structure database, open access, COD

**P.24.01.2**

*Acta Cryst.* (2005). **A61**, C481

**Hydrogen and Hydration DataBase for Bio-Macromolecules (HHDB)**

Ryota Kuroki<sup>1</sup>, Toshiyuki Chatake<sup>1</sup>, Kazuo Kurihara<sup>1</sup>, Takashi Ohhara<sup>1</sup>, Hisao Umino<sup>1</sup>, Raul E. Cachau<sup>3</sup>, Michael Blaber<sup>4</sup>, Nobuo Niimura<sup>1,2</sup>, <sup>1</sup>Neutron Science Research Center, Japan Atomic Energy Research Institute (JAERI). <sup>2</sup>Institute of Applied Beam Science, Graduate School of Science and Engineering, Ibaraki University. <sup>3</sup>Advanced Biomedical Computing Center, National Cancer Institute at Frederick. <sup>4</sup>Department of Chemistry and Biochemistry, Florida State University. E-mail: r-kuroki@neutrons.tokai.jaeri.go.jp

Solvation, desolvation and hydrogen bonding are key energetic contributors to biopolymer folding, dynamics and molecular recognition; however, no systematic, high-resolution stereochemical database dedicated to the characterization and analysis of hydrogen bonding exists. We have created a hydrogen and hydration database for bio-macromolecules (HHDB; <http://hhdb.tokai.jaeri.go.jp>) that categorizes all hydrogen atoms and hydration water molecules including hydrogen atoms in proteins. The HHDB includes H-bond data only from direct determination of hydrogen atoms by neutron diffraction, and certain extremely high resolution x-ray diffraction data. The HHDB provides a graphical user interface for the visualization of all hydrogen atom positions in proteins and solvent, and all hydrogen bonding (H-bond) interactions. For example, one type of plot featured in the database places the hydrogen atom positions at the origin and plots the direction of the H-bond donor along the horizontal axis, allowing the user to visualize the distribution of the acceptor atoms, distance and angle. The HHDB provides researchers with a much-needed resource for understanding and analyzing hydrogen-bonding.

**Keywords:** neutron diffraction, protein, hydrogen bond

**P.24.01.3**

*Acta Cryst.* (2005). **A61**, C481

**Four Years of the EPSRC SRS Service**

Ross W. Harrington, W. Clegg, *School of Natural Sciences*

(Chemistry) University of Newcastle. E-mail: r.w.harrington@ncl.ac.uk

Since October 2001, the EPSRC national crystallography service in the UK has had a SRS component, run in conjunction with the national crystallographic service based at Southampton University, through the use of station 9.8 at Daresbury laboratory.

Over the course of the past 3.5 years regular monthly visits have taken place, and a total of 385 data collections made, averaging 6 data collections per day of beam time.

This poster reports the many successes of the service, and includes statistics on users and publications benefiting from the service.

The reader will hopefully gain an insight into a service utilizing the most successful small molecule synchrotron beam line in the world.

**Keywords:** EPSRC, national crystallographic service, statistics

**P.24.02.1**

*Acta Cryst.* (2005). **A61**, C481

**The Crystallographic Semantic Web**

Peter Murray-Rust, Simon Tyrrell, Yong Zhang, *Unilever Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, CB2 1EW, UK*. E-mail: pm286@cam.ac.uk

The semantic Web [Berners-Lee, <http://www.w3.org/2001/sw/>] is a vision of a global knowledge network where machines understand and reason from web-based resources. We provide working demonstrations that this technology is ideally suited to create an Open crystallographic knowledge base. The complete information chain - experiment, publication, storage, dissemination, searching and re-use can be completely managed by machines. CIFs can be converted to XML, annotated, and redistributed as WebServices and indexed by conventional search engines. New structures can be announced through crystallographically enhanced RSS feeds.

Our World Wide Molecular Matrix (WWMM, <http://wwmm.ch.cam.ac.uk>) uses free-text XML indexing in a repository (eXist) to collect hundreds of thousands of compounds. IUPAC InChI and chemical substructure searching (e.g. <http://openbabel.sf.net>) to provide an instant, Open and freely redistributable crystallographic knowledge base. Data are abstracted from publishers who allow Open access to CIFs and authors can publish CIFs directly into XML (<http://www.xml-cml.org>). Institutional and national collections (<http://eprints.soton.ac.uk/1633/>) with appropriate metadata allow conventional search engines to index the data and effectively create a complete Open database of crystal structures.

**Keywords:** XML, CIF, semantic web

**P.24.02.2**

*Acta Cryst.* (2005). **A61**, C481-C482

**ECRYSTALS(CHEM.SOTON.AC.UK): Open Archive Publication of Crystal Structure Data**

Mike B. Hursthouse<sup>a</sup>, S. J. Coles<sup>a</sup>, J. G. Frey<sup>a</sup>, L. Carr<sup>b</sup>, C. Gutteridge<sup>b</sup>, L. Lyon<sup>c</sup>, R. Heery<sup>c</sup>, M. Duke<sup>c</sup>, M. Day<sup>c</sup>, <sup>a</sup>School of Chemistry, <sup>b</sup>School of Electronics and Computer Science, University of Southampton, <sup>c</sup>UKOLN, University of Bath., UK. E-mail: mbh@soton.ac.uk

The eCrystals Repository has been developed as part of the eBank-UK project, to serve as an exemplar and testbed for open archive publication, linking and aggregation of digital scientific data.

This poster outlines a "pre-print"[1] procedure for the rapid and effective dissemination of structural information "@source", based on the "e-print" concept. An eCrystals record makes available, for assessment and/or re-use, all raw, derived, results and validation data generated during the course of a crystallographic experiment. During the deposition process metadata, comprising bibliographic and chemical identifiers, are associated with a dataset, and key items of data are automatically extracted and displayed on a HTML jumpoff page, which is then offered over the internet by standard Open Archive Initiative (OAI) protocols. The bibliographic, crystallographic and chemical metadata relating to an eCrystals record may be