News in Open structure and property Databases

Daniel Chateigner[1,*], Saulius Grazulis[2], Justas Butkus[2], Andrius Merkys[2], Adriana Daškevič[2], Olivier Pérez[1], Giancarlo Pepponi[3], Luca Lutterotti[1,4], Miguel Quirós Olozábal[5], Armel Le Bail[6], Robert Downs[7], Peter Moeck[8], Alexandre F.T. Yokochi[9]

[1] CRISMAT-ENSICAEN-CNRS, IUT Caen, Normandie Universités, Caen, France, [2] Inst. of Biotechnology, Vilnius, Lithuania, [3]Fundazione Bruno Kessler, Trento, Italy, [4]Dept. of Industrial Engineering, Univ. of Trento, Italy, [5]Dept de Química Inorgánica, Univ. de Granada, Spain, [6]Lab des Oxydes et Fluorures, CNRS and Univ. du Maine, Le Mans, France, [7]Dept of Geosciences, Univ. of Arizona, Tucson, Arizona, USA, [8]Portland State Univ., Dept of Physics, Portland, Oregon, USA, [9]School of Chemical, Biological and Environmental Engineering, Oregon State Univ., Corvallis, Oregon, USA

The COD project (abbreviated from the "Crystallography Open Database", http://www.crystallography.net/) aims at collecting in a single open access database all organic, inorganic and metal organic structures [1]. It grew to more than 235000 cif files in the past 10 years. Since December 2007 the main database server is maintained and new software is developed in the Vilnius University Institute of Biotechnology. Structure file harvesting is now routinely operated and automated for a growing number of publishers' journals [2]. Registered users can deposit new data into the database, either form the previous publications or as personal communications. The deposition software performs rigorous checks of syntax and semantics, thus ensuring high quality depositions. The search engine acquired new capabilities, and COD records can be viewed on-line or downloaded. For massive data mining, COD permits downloads and updates of the whole database using Subversion, Rsync or HTTP protocols. The ease of access to COD data has spurred the use of this resource for software testing [3], teaching [4], and research [5].

For the powder diffraction community, COD is recently used for open Full-Pattern Search-Match (cod.iutcaen.unicaen.fr). It allows phase quantification from x-rays, neutron and electron powder patterns, with high or medium resolution instruments, provided the structures are already in COD. This tool is particularly suited for nanocrystalline powders in which severe line broadening appears precluding phase identification from only peak maxima. COD-derived databases are offered for software produced by many diffractometer vendors. In addition to COD, search match can be done against its sister database, PCOD, that contains structures predicted by the GRINSP program [6]. Recently launched the new TCOD database collects structures optimized from COD using first-principles calculations. The open nature of the COD permitted numerous mirrors around the globe [7] and specifically tailored COD database variants. Meanwhile, the Material Properties Open Database [8] collects tensor properties. This latter sister is on a demanding growth.

References

1. Gražulis, S.; Chateigner, D.; Downs, R. T.; Yokochi, A. F. T.; Quirós, M.; Lutterotti, L.; Manakova, E.; Butkus, J.; Moeck, P. & Le Bail, A. (**2009**). *Crystallography Open Database -- an open-access collection of crystal structures*, Journal of Applied Crystallography 42 : 726-729.

2. Gražulis, S.; Daškevič, A.; Merkys, A.; Chateigner, D.; Lutterotti, L.; Quirós, M.; Serebryanaya, N. R.; Moeck, P.; Downs, R. T. & Le Bail, A. (**2012**). *Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration*, Nucleic Acids Research 40 : D420-D427.

3. Grosse-Kunstleve, R. & Gildea, R. (**2011**). *Computational Crystallography Initiative: COD stats*, http://cci.lbl.gov/cod_stats/ (retrieved 2013.01.31).

4. Moeck, P. (**2004**). *EDU-COD: Educational Subset of COD*, http://nanocrystallography.research.pdx.edu/search/edu/ (retrieved 2013.01.31).

5. First, E. L. & Floudas, C. A. (**2013**). *MOFomics: Computational pore characterization of metal–organic frameworks*, Microporous and Mesoporous Materials 165 : 32-39.

9 L. Lutterotti, H. Pilliere, C. Fontugne, P. Boullay, D. Chateigner (**2013**), in preparation

6. Le Bail, A. (**2005**). *Inorganic structure prediction with it GRINSP*, Journal of Applied Crystallography 38 : 389-395.

7. Quirós-Olozábal, M. (**2006**). *COD Mirror of Granada University*, http://qiserver.ugr.es/cod/. Moeck, P. (**2007**). *Crystallography Open Database Mirror*, http://nanocrystallography.research.pdx.edu/search/codmirror/. Gražulis, S. (**2007**). *COD Mirror in Vilnius*, http://cod.ibt.lt/. Chateigner, D. (**2010**). *Crystallography Open Database Mirror at ENSICAEN*, http://cod.ensicaen.fr/.

8 G. Pepponi, S. Grazulis, D. Chateigner (**2012**). MPOD: a Material Property Open Database linked to structural information. *Nuclear Instruments and Methods in Physics Research B* **284**, 2012, 10-14. http://www.materialproperties.org/