Saulius Gražulis, Justas Butkus and Andrius Merkys. The deposition software performs rigorous checks of syntax and semantics.

The COD website allows searching on COD numerical identifier, unit-cell parameters, chemical composition and bibliographic data. Substructure searches using SMILES and SMARTS strings have been implemented. The free software package *OpenBabel* (O'Boyle *et al.*, 2011; Hutchison, 2007) is used for both the CIF-to-SMILES transformation and the actual search.

The retrieved records can be viewed online or downloaded for further processing. For massive data mining, COD permits downloads and updates of the whole database using Subversion, Rsync or http protocols. The ease of access to the COD data and its open nature has spurred the use of this resource for software testing (Grosse-Kunstleve & Gildea, 2011), teaching (Moeck, 2004) and research (First & Floudas, 2013). Multiple mirrors around the globe (Quirós-Olozábal, 2006; Gražulis, 2007; Moeck, 2007*a*; Chateigner, 2010) ensure data preservation, provide off-site backups, offer improved search interfaces (Moeck, 2007*b*) and increase reliability.

For the powder-diffraction community, the COD is interesting not only as an archive of structures solved by powder-diffraction methods, but also as a possibility for use in search/match procedures to identify crystalline compounds. Recently, the development of an open full-pattern search/match internet tool was launched by the COD developers. It allows phase quantifications from X-ray, neutron and electron powder patterns (with high- or medium-resolution instruments) provided that the structures are already in the COD. This tool is particularly suited to nanocrystalline powders, in which severe line broadening appears, precluding phase identification from only peak positions (Lutterotti *et al.*, 2012). COD-derived databases are also offered for software produced by several diffractometer vendors (Rigaku, 2011; PANalytical, 2012*a*,*b*; Bruker, 2013). In addition to the COD, searches and matches can be performed against its sister database, the PCOD, which contains structures predicted by the *GRINSP* program (Le Bail, 2005) and hypothetical zeolites (Pophale *et al.*, 2013). The power of such an approach is demonstrated by PCOD entry 3102887 (formulated as SiO_2). It was recently identified as corresponding structurally to a new phosphorus(V) oxonitride polymorph δ-PON (Baumann *et al.*, 2012).

3.7.9. Other internet databases

Other useful databases include the following:

(i) The American Mineralogist Crystal Structure Database (http://rruff.geo.arizona.edu/AMSamcsd.php).
(ii) The Mineralogy Database (http://webmineral.com).
(iii) MinCryst (http://database.iem.ac.ru/mincryst/index.php).
(iv) The International Zeolite Association Database of Zeolite Structures (http://www.iza-structure.org/databases).
(v) The Incommensurate Structures Database (http://webbdcrista1.ehu.es/incstrdb/).
(vi) The Nucleic Acid Database (http://ndbserver.rutgers.edu).

I thank Timothy G. Fawcett and Cyrus Crowder of the International Centre for Diffraction Data, Colin Groom of CCDC, Stephan Ruehl of FIZ Karlsruhe, Pierre Villars and Karen Cenzual of Material Phases Data Systems, and Saulius Gražulis of Vilnius University Institute of Biotechnology for their valuable comments and additions.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst.* D**66**, 213–221.

Allaire, M., Moiseeva, N., Botez, C. E., Engel, M. A. & Stephens, P. W. (2009). *On the possibility of using polycrystalline material in the development of structure-based generic assays. Acta Cryst.* D**65**, 379–382.

Allen, F. H., Cole, J. C. & Verdonk, M. L. (2011). *The relevance of the Cambridge Structural Database in protein crystallography. International Tables for Crystallography*, Vol. F, 2nd ed., edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 736–748. Chichester: Wiley.

Allmann, R. & Hinek, R. (2007). *The introduction of structure types into the Inorganic Crystal Structure Database ICSD. Acta Cryst.* A**63**, 412–417.

Barr, G., Dong, W. & Gilmore, C. J. (2009). *PolySNAP3: a computer program for analysing and visualizing high-throughput data from diffraction and spectroscopic sources. J. Appl. Cryst.* **42**, 965–974.

Barr, G., Gilmore, C. J. & Paisley, J. (2004). *SNAP-1D: a computer program for qualitative and quantitative powder diffraction pattern analysis using the full pattern profile. J. Appl. Cryst.* **37**, 665–668.

Basso, S., Besnard, C., Wright, J. P., Margiolaki, I., Fitch, A., Pattison, P. & Schiltz, M. (2010). *Features of the secondary structure of a protein molecule from powder diffraction data. Acta Cryst.* D**66**, 756–761.

Baumann, D., Sedlmaier, S. J. & Schnick, W. (2012). *An unprecedented AB_2 tetrahedra network structure type in a high-pressure phase of phosphorus oxonitride (PON). Angew. Chem. Int. Ed.* **51**, 4707–4709.

Behrens, H. & Luksch, P. (2006). *A bibliometric study in crystallography. Acta Cryst.* B**62**, 993–1001.

Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. (2002). *New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. Acta Cryst.* B**58**, 364–369.

Bergerhoff, G. & Brandenburg, K. (1999). *Typical interatomic distances: inorganic compounds. International Tables for Crystallography*, Vol. C, edited by E. Prince, pp. 770–781. Dordrecht: Kluwer Academic Publishers.

Bergerhoff, G. & Brown, I. D. (1987). *Crystallographic Databases*, edited by F. H. Allen, G. Bergerhoff & R. Sievers. Chester: International Union of Crystallography.

Berman, H. M., Henrick, K., Kleywegt, G., Nakamura, H. & Markley, J. (2011). *The Worldwide Protein Data Bank. International Tables for Crystallography*, Vol. F, 2nd ed., edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 827–832. Chichester: Wiley.

Berman, H., Henrick, K. & Nakamura, H. (2003). *Announcing the Worldwide Protein Data Bank. Nature Struct. Mol. Biol.* **10**, 980.

Berndt, M. (1994). Thesis, University of Bonn, Germany. Updates by O. Shcherban, SCC Structure-Properties Ltd, Lviv, Ukraine.

Berndt, M. (2003). *Open crystallographic database – a role for whom?* http://www.cristal.org/SDPD-list/2003/msg00025.html.

Bigelow, W. C. & Smith, J. V. (1964). *Two new indexes to the Powder Diffraction File. ASTM Spec. Tech. Publ.* STP372, 54. https://doi.org/10.1520/STP48334S.

Boldyrev, A. K., Mikheev, V. I., Dubinina, V. N. & Dovalev, G. A. (1938). *X-ray determination tables for minerals, Ft. I. Ann. Inst. Mines Leningrad*, **11**, 1–157.

Boles, M. O., Girven, R. J. & Gane, P. A. C. (1978). *The structure of amoxycillin trihydrate and a comparison with the structures of ampicillin. Acta Cryst.* B**34**, 461–466.

Bravais, A. (1866). *Etudes Cristallographiques*. Paris: Gathier Villars.

Bruker-AXS (2013). Crystallography Open Database for DIFFRAC.EVA. https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/x-ray-diffraction/xrd-software/eva/cod.html.

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., Pearson, J. & Taylor, R. (2002). *New software for searching the Cambridge Structural Database and visualizing crystal structures. Acta Cryst.* B**58**, 389–397.

Bruno, I. J., Cole, J. C., Kessler, M., Luo, J., Motherwell, W. D. S., Purkis, L. H., Smith, B. R., Taylor, R., Cooper, R. I., Harris, S. E. & Orpen, A. G. (2004). *J. Chem. Inf. Comput. Sci.* **44**, 2133–2144.