

Crystal structure visualizations in three dimensions with database support

P. Moeck¹, O. Čertík^{1,2}, G. Upreti¹, W. Garrick³, P. Fraundorf⁴

¹ Department of Physics, Portland State University, P.O. Box 751, Portland, OR 97207-0751, pmoeck@pdx.edu

² Faculty of Mathematics and Physics, Charles University of Prague, Ke Karlovu 3, 121 16 Praha, The Czech Republic

³ Academic & Research Computing for Instruction and Research Services and the Office of Information Technology, Portland State University, Portland, OR 97201

⁴ Department of Physics and Astronomy and Center for Molecular Electronics, University of Missouri at St. Louis, MO 53121

ABSTRACT

Crystallographic databases for inorganic materials that are freely accessible over the internet are reviewed. The Nano-Crystallography Database project is described. Instructions are given on how to visualize in three dimensions the atomic arrangements of the several thousand entries of the Crystallography Open Database.

1. INTRODUCTION

Courses in materials science and engineering, crystal physics, crystal chemistry, and mineralogy typically employ two-dimensional sketches of the atomic arrangement in crystal structures and unit cells. The chapter on ideal crystal structures of the widely used introductory materials science and engineering text by Schaffer et al. [1] contains for example 14 such sketches. Textbooks on mineralogy contain a much larger number of such sketches (e.g. Hibbard's text [2] features a picture with more than ten ball-and-stick models on the dedication page and approximately one hundred sketches of atomic arrangements in crystal structures and unit cells in the body of the text). Physical three-dimensional models of crystal structures and unit cells are also popular. While academic departments in the developed part of the world may possess some tens of such three-dimensional models and allow their students to explore those hands on, their counterparts in the developing part of the world may consider such models too expensive and fragile for class room usages.

To help remedying this situation we made a survey on crystallographic databases that are freely accessible on the internet and also started our own nano-crystallography database project with on line three-dimensional (3D) visualizations of ideal and real structures. This paper reports on both the results of the survey and our ongoing project.

2. FREE INTERNET ACCESSIBLE CRYSTALLOGRAPHIC DATABASES

Fifteen years after the development of the World Wide Web, there is already a variety of internet based free access crystallographic databases. In this part of the paper we discuss only the subset of free databases that deals primarily with inorganic crystals.

The free database MINCRYST [3] is maintained and hosted by the Russian Academy of Sciences and contains more than 6,000 entries for minerals from which X-ray powder

diffraction patterns can be calculated on the fly. 3D visualization of the entries are provided by means of the java-based freeware program “Jmol” [4]. The Naval Research Laboratory’s Center for Computational Materials Science provides a free on-line database for education and research support under the name “Crystal Lattice Structures” [5]. Currently this database contains 254 entries in 90 space groups. Visualizations of the entries are also provided by Jmol applets. The “Reciprocal Net” is a distributed database of molecular structure information and its main web site is hosted at Indiana University at Bloomington [6]. Information from approximately 400 common molecules (and crystal structures) can be downloaded in the standard Crystallographic Information File (CIF) [7] format (file extension *.CIF). 3D visualizations of these common molecules (and crystal structures) are provided by a range of Java based applets.

Since its adoption in 1990 by the International Union of Crystallography as the standard file format and syntax for communicating crystallographic information between both human beings and computers and between different kinds of software programs and computers, CIF has become a crystallographic entity in its own right of such a significance that there are plans to dedicate a future volume of the International Tables for Crystallography to describing it comprehensively. A free on-line check of the syntax of CIFs is provided by the International Union of Crystallography at ref. [8].

The free on-line American Mineralogist Crystal Structure Database (AMCSD) was erected in 2003 and contains more than 8,000 critically evaluated entries [9],[10]. Files can be downloaded from this database in CIF and the American Mineralogist Crystal data format (file extension *.amc). A range of PC based freeware programs, for e.g. the visualization of crystal structures (XtalDraw) and for the calculation of powder X-ray and neutron diffraction patterns (XPow), is available for downloading from the AMCSD site at ref. [11]. The database is maintained under the care of both the Mineralogical Society of America and the Mineralogical Association of Canada. Ref. [9] states that approximately 75 % of the submitted manuscripts to the journals “American Mineralogist” and the “Canadian Mineralogist” possess errors of one sort or another. (For comparison, ref. [12] states that about 40 % of the CIFs receive for inclusion into the commercial Cambridge Structural Database contain errors that arise mainly from the manual editing of these files by their authors.) The critical evaluation of data to be included in a crystallographic database is, therefore, a major concern for both commercial and free databases. This might be one of the reasons that no provisions for uploading data to the AMCSD are provided.

Records of approximately 20,000 full crystal structure determinations in the form of CIFs are freely accessible on-line at the Crystallography Open Database (COD) [13,14]. This database was also started in 2003 and is rapidly growing as more and more crystallographers and scientific institutes/societies/academies donate their collections of CIFs and upload them over the internet so that anybody with access to the World Wide Web can access them for free. The COD also contains all the data sets that are in the AMCSD. The long term objectives of the COD initiators are summarized on the COD web page [14] as: (i) providing free access to comprehensive crystallographic data (including the atomic coordinates) on all known inorganic, metallic, organometallic and organic crystalline compounds and (ii) complementing the existing commercial databases (which typically specialize on only one or two classes of crystals and contain in addition to crystallographic information a range of physical properties). Another major concern of the COD initiators is supporting crystallographers in emerging countries. By providing the opportunity to check rapidly and early on in an investigation (i.e. already after the

cell parameters have been determined) if the structure of the particular crystal under investigation has already been solved, the COD also helps increasing the productivity of structure determinations world wide. To insure a high quality of the data entries, the policy of the COD is that only those CIFs that result from structure refinement software can be uploaded.

3. THE NANO-CRYSTALLOGRAPHY DATABASE PROJECT AND VISUALIZATION OF STRUCTURES IN THREE DIMENSIONS

As a complementary project with a similar philosophy to both the COD and the AMCSD, the nanocrystallography group at Portland State University started in 2005 the Nano-Crystallography Database (NCD) project [15]. Since structure* and morphology of nanocrystals are crucial to their physical properties, the NCD is collecting entries on both the full structure (including atomic coordinates) and the typical morphology (tracht and habit) of inorganic nanocrystals in the form of CIFs. The experimental morphological information that is collected in the “Bestimmungstabellen für Kristalle” [16] will be the first to be included in the NCD.

As many electron microscopist work on rather simple inorganic structures, we will write and upload CIFs to support their work. A good starting point for this are the descriptions of the 161 full structure determinations in the second edition of Wyckoff’s classical reference book “Structure of Crystals” [17]. To insure a high quality of the data entries, we provide a comprehensive on-line syntax check for CIFs [18]. Only those CIFs that pass both this test and the test that is available at the IUCr web site [8] will be uploaded to the NCD.

Since the NCD is being developed to support not only materials science education but also image-based nanocrystallography in general (i.e. methods that determine both structure and morphology of nanocrystals from images [20-22] in transmission electron microscopes), visualizations in 3D of the entries in this database are crucial. For a start, we provide 3D visualizations [18] by means of Jmol applets for the structural information that is contained in the entries of the COD, Figs. 1 to 3. Next we will work on providing visualizations of the morphological information that will be contained in the entries of the NCD by means of the program “XMorph” [23]. Fig. 1a shows the opening screen of the website that hosts our local version of the COD. The menu entries are self explanatory and read from top to bottom: *Search the Database, Upload Data, Check syntax errors and Insert symmetry equivalent positions.*

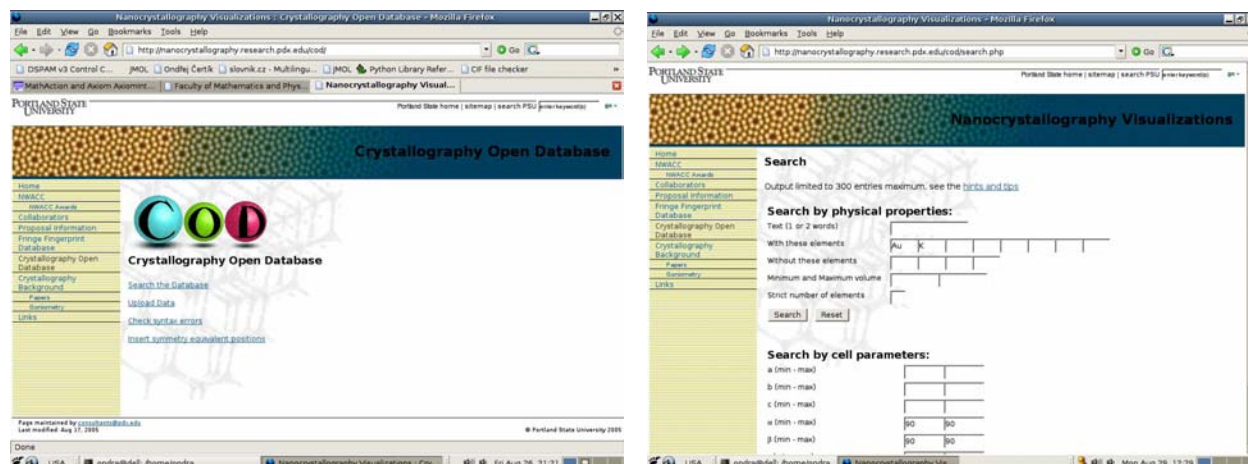


Fig. 1: Screenshots of (a, right) opening screen of the website that hosts our local version of the COD and (b, left) search screen with entries for the elements Au and K.

Figure 2 consists of two screenshots from a web browser. The left screenshot (a) shows the search results page for Au and K. The right screenshot (b) shows the CIF file content for the second entry in the list.

(a) Search Results:

Result: There are 6 entries in the selection Searching elements including Au K

Entry 1:
 Formula: $P 1 00 Au 1 00 K 2 00$
 Comments: Eisenmann, B Klein, J Somer, M Crystal structure of dipotassium catena-phosphonate Zefschroft fuer entalographie (149,1979) 197 (1991) 277
 Space group: $C 2/m$
 Cell volume: 445.40
 Cell parameters: 9.7870, 7.9950, 6.1480, 90.000, 90.000, 90.000

Entry 2:
 Formula: $As 2 00 Au 1 00 K 5 00$
 Comments: Eisenmann, B Klein, J Somer, M Linear ansions [CuAu225-, (Au)225- and [AuAu225- in potassium compounds Journal of Alloys Compd. 178 (1992) 431, 439
 Space group: $P 63/m$
 Cell volume: 546.40
 Cell parameters: 5.7300, 5.7300, 19.2160, 90.000, 90.000, 120.000

Entry 3:
 Formula: $P 2 00 Au 1 00 K 5 00$
 Comments: Eisenmann, B Klein, J Somer, M Linear ansions [CuAu225-, (Au)225- and [AuAu225- in potassium compounds Journal of Alloys Compd. 178 (1992) 431, 439
 Space group: $P 63/m$
 Cell volume: 514.90
 Cell parameters: 5.6300, 5.6300, 18.7160, 90.000, 90.000, 120.000

Entry 4:
 Formula: $As 2 00 Au 1 00 K 5 00$
 Comments: Zachewia, W, Wlozdzka, J, KAu5Pb: Layers of [20Infnb](AuAu32) gold tetrahedra and [12Infnb](Pb-2Z) lead chains - ZEITSCHRIFT FUR ANORGANISCHE UND ALLEGEMEINE CHEMIE 624 (1998) 1369-1372
 Space group: $Im\bar{3}m$
 Cell parameters: 6.5100, 11.9300, 4.2300, 90.000, 90.000, 90.000

Entry 5:
 No CIF
 Formula: $Au 5 00 K 3 00 Pb 1 00$
 Comments: Zachewia, W, Wlozdzka, J, KAu5Pb: Layers of [20Infnb](AuAu32) gold tetrahedra and [12Infnb](Pb-2Z) lead chains - ZEITSCHRIFT FUR ANORGANISCHE UND ALLEGEMEINE CHEMIE 624 (1998) 1369-1372
 Space group: $Im\bar{3}m$
 Cell parameters: 6.5100, 11.9300, 4.2300, 90.000, 90.000, 90.000

(b) CIF File Content:

```
data_1008878
_chemical_name_systematic 'Pentapotassium arsenoborate'
_chemical_formula_structural 'K5 Au As2'
_chemical_formula_sum 'As2 Au K5'
_title 'Linear ansions [CuAu225-, (Au)225- and [AuAu225- in potassium compounds'
loop
_jmol_author_name
11 'Eisenmann, B'
12 'Sommer, M'
13 'Sommer, M'
14 _jmol_name_half 'Journal of Alloys Compd.'
15 _jmol_volume_half 'JALCEU'
16 _jmol_volume 178
17 _jmol_year 1992
18 _jmol_page_first 431
19 _jmol_page_last 439
20 _cell_length_a 5.73002
21 _cell_length_b 5.73002
22 _cell_length_c 19.21599(40)
23 _cell_angle_alpha 90
24 _cell_angle_beta 90
25 _cell_angle_gamma 120
26 _cell_volume 546.4
27 _cell_formula_units_z 2
endloop
```

Fig. 2 Screenshots of (a, right) results of the search for the simultaneous presence of the two elements Au and K and (b, left) CIF of the second entry in the result list of (a) displayed by the free LINUX version of the program “enCIFer” [24].

Figure 3 is a screenshot of the Jmol visualization screen for one unit cell of As_2AuK_5 . The screen shows a 3D ball-and-stick model of the unit cell with axes labeled X, Y, and Z. The model is displayed in a Jmol applet window.

Showing file '1008878.cif' in a Jmol applet.

Unit cell parameters:
 $a = 5.730 \text{ \AA}$
 $b = 5.730 \text{ \AA}$
 $c = 19.216 \text{ \AA}$
 $\alpha = 90, 0^\circ$
 $\beta = 90, 0^\circ$
 $\gamma = 120, 0^\circ$

Buttons: 45 | 45 | show unit cell
 Rot X | Rot Y | Rot Z
 Reset to original orientation | mode 1 - asymmetric unit | mode 2 - primitive cell | mode 3 - conventional cell

Fig. 3: Screenshots of the visualization screen for one unit cell of As_2AuK_5 displayed by Jmol.

Fig. 1b shows the search screen that is accessible from the opening screen, Fig. 1a. The elements Au and K were entered into this screen to conduct a search for all the entries that contain these

two elements together in one crystal phase. Fig. 2a shows the result screen for the search for the simultaneous presence of Au and K. By clicking on *CIF file* for the second entry in this list, the respective CIF is displayed on the user's computer with the program he or she may select for this purpose. The text editor "Notepad" can be employed to display and edit CIFs, but the free program "enCIFer" was specifically designed for these purposes and can be downloaded for all common platforms at ref. [24]. Using enCIFer, Fig. 2b shows the CIF for As_2AuK_5 .

Fig. 3 shows the 3D visualization of the As_2AuK_5 structure in Jmol. Alternatively, the *asymmetric unit* (which is typically just one molecule in molecule crystals) can be displayed. By clicking on *3 x 3 x 3 primitive unit cells*, the content of nine unit cells can be displayed at once. There are options to display only the atoms or the atoms together with a sketch of the unit cell superimposed on the atoms. The cell parameters are listed in addition when a structure is displayed with the latter option. As with any Jmol applet, one can zoom into the display of the atomic arrangement by using the Shift key together with the mouse. A range of tools that allow, for example, for the measurement of distances between atoms or the bond angles between three atoms are accessible by a left mouse click directly onto the Jmol structure display. (To display the 3D visualizations with Jmol, the internet browser must support Java 1.4. If not already installed within the browser, the JAVA™ 2 platform is available for free at ref. [19].)

The actual visualization of a COD database entry in three dimensions is obtained by turning the displayed atomic arrangement, e.g. Fig. 3, around three mutually perpendicular axes by clicking on *Rot X*, *Rot Y*, and *Rot Z*. Rotations by 45 degrees are default, but arbitrary rotations (with a negative or positive sign) can be realized by entering the desired rotations in the respective boxes of the rotation menu. The option to *Reset to (the) original orientation* also exist.

In the future, we will add more crystallographic functionality to the display of structures in order to support image-based nanocrystallography [20-22] so that the NCD can become a free internet based research support tool in addition to providing world wide support for materials science and engineering, crystal physics, crystal chemistry, and mineralogy education.

SUMMARY

Crystallographic database for inorganic materials that are freely accessible over the internet were reviewed. The objectives, current state of the affair, and future directions of the Nano-Crystallography Database project were described. Instructions were provided on how to visualize the atomic arrangements in three dimensions for several thousand crystals for which entries exist in the Crystallography Open Database.

ACKNOWLEDGMENTS

This project is supported by the NorthWest Academic Computing Consortium. Additional support was provided by both an award from Research Corporation and a Scholarship of Teaching Resources Team Incentive Grant from Portland State University.

REFERENCES

- [1] J.P. Schaffer, et al., *The Science and Design of Engineering Materials*, WBC/McGraw-Hill, Boston, 1999
- [2] M.J. Hibbard, *Mineralogy: A Geologist's Point of View*, McGraw-Hill, Boston, 2002
- [3] <http://database.iem.ac.ru/mincryst/>

- [4] <http://Jmol.sourceforge.net>
- [5] <http://cst-www.nrl.navy.mil/lattice/index.html>
- [6] <http://www.reciprocalnet.org/>
- [7] I.D. Brown and B. McMahon, CIF: the computer language of crystallography, *Acta Cryst. B* **58** (2002) 317-324
- [8] <http://checkcif.iucr.org/>
- [9] R.T. Downs and M. Hall-Wallace, The American Mineralogist crystal structure database, *American Mineralogist* **88** (2003) 247–250
- [10] <http://rruff.geo.arizona.edu/AMS/amcsd.php>
- [11] <http://rruff.geo.arizona.edu/AMS/extra.php>
- [12] F.H. Allen, The Cambridge Structural Database: a quarter of a million crystal structures and rising, *Acta Cryst. B* **58** (2002) 380-388
- [13] A. Le Bail, COD (Crystallography Open Database), Newsletter No. **29**, June 2003, 39-40, The International Union for Crystallography's Commission on Powder Diffraction, ISSN 1591-9552, see also D. Chateigner, X. Chen, M. Ciriotti, L.M.D. Cranswick, R.T. Downs, A. Le Bail, L. Lutterotti, A.F.T. Yokochi, COD (Crystallography Open Database) and PCOD (Predicted), Book of Abstracts, XX Congress of the International Union of Crystallography, Florence, (Italy), August 23-31, 2005, <http://www.iucr2005.it/pdf/17.pdf>
- [14] <http://www.crystallography.net>
- [15] <http://www.nanocrystallography.research.pdx.edu>
- [16] A.K. Boldyrew and W.W. Doliwo-Dobrowolsky, Bestimmungstabellen für Kristalle (Определитель Кристаллов), Vol. I, Part 1, Einleitung, Tetragyrische Syngonie; W.W. Doliwo-Dobrowolsky and G.P. Preobraschensky, Vol. I, Part 2, Trigyrische and Hexagyrische Syngonien allgemeine Ergänzungen zu den mittleren Syngonien, Zentrales Wissenschaftlichers Institute der Geologie und Schürfung, Leningrad and Moscow, 1937 and 1939 (in *Russian* and *German*)
- [17] R.W.G. Wyckoff, The Structure of Crystals, 2nd edition & Supplement for 1930-1934 to the 2nd edition, American Chemical Society Monograph Series 19 & 19a, Reinhold Publishing, 1931 & 1935
- [18] <http://www.nanocrystal.research.pdx.edu/cod>
- [19] <http://www.java.com/en/download/index.jsp>
- [20] P. Fraundorf, W. Qin, P. Moeck, and E. Mandell, Making sense of nanocrystal fringes, arXiv:cond-mat/0212281 v2 (2005), *J. Appl. Phys.*, *accepted*
- [21] P. Moeck, W. Qin, and P. Fraundorf, Towards 3D image-based nanocrystallography by means of transmission electron goniometry, *Mat. Res. Soc. Symp. Proc. Vol. 839* (2005) P4.3.1-P4.3.6
- [22] P. Moeck, B. Seipel, W. Qin, E. Mandell, and P. Fraundorf, Fringe fingerprinting and transmission electron goniometry, supporting image-based nanocrystallography in two and three dimensions, *Proc. 9th World Multi-Congress on Systemics, Cybernetics and Informatics*, Vol. **IX**, 249-254, July 10-13, 2005, Orlando, Florida
- [23] Courtesy of Dr. Dedong Wu, AstraZeneca Pharmaceuticals LP, 1800 Concord Pike, Wilmington, DE 19850
- [24] http://www.ccdc.cam.ac.uk/free_services/encifer/

* We consider here the surface of nanocrystals and the possible presence of single or multiple twins as the most important real structure. In the future, the NCD project will collect (and later on display) such information as well (together with visualizations of other crystal defects).