STRUCTURAL DATABASES OF INORGANIC MATERIALS

R. Kužel, S. Daniš

Charles University in Prague, Faculty of Mathematics nad Physics
Ke Karlovu 5, 121 16 Praha 2

Keywords
structural databases, Inorganic Crystal Structure Database, Crystmet, Powder Diffraction File, Crystal Lattice Structures, mineralogical databases

Abstract
Structural databases are one of the basic tools of modern crystallographer. They are necesary for many tasks from just drawing the picture of the structure and study of symmetry to phase analysis and design of new materials. In this contribution, some databases for inorganic compounds are characterized together with their availability.The following databases and other sources of the database type information are mentioned - ICSD - inorganic crystal structure database, Crystmet, Powder Diffraction File, Crystal Structure Lattices, Pauling file, Pearson's Crystal Data, Mincryst and some others.

Links to the most important structural databases can be found for example at address [1].

Inorganic Crystal Structure Database - ICSD

Basic structural database of inorganic compounds is the database **ICSD** - Inorganic Crystal Structure Database [2, 3] developed and maintained in the Information center FIZ in Karlsruhe. It includes records of all inorganic structures published since 1913. The scope was gradually extended in the direction of intermetallic compounds with known atomic coordinates. Since 2003 FIZ Karlsruhe started to fill in the remaining gaps and will in future include all published intermetallic compounds. Until 2003 ICSD didn't contain any structures with C-H or C-C bonds ("organic compounds"). In 2003 this rule was modified: Now new entries should not contain both C-H AND C-C bonds (i.e. compounds with tetramethylammonium will be included now as well as oxalates).

The editorial team continuously extract and abstract the original data from over 80 leading scientific journals and additional 1,900 scientific journals. ICSD is updated twice a year, each time adding approximately 3,000 new records.

At present, the ICSD contains 600 crystal structures of the elements, 15,000 records for binary compounds, 33,000 records for ternary compounds, 35,000 records for quarternary and quintenary compounds. Last half-year increase of entries was about 4000. All the entries are checked several times and only data from good quality sources are accepted.

The database contains fully determined structures with atomic coordinates. Coordinates for hydrogen atoms or vagabonding atoms like Na in Zeolites may be missing.
Most of the structures contained in ICSD are published in journals, only a few entries were submitted as private communications.

The entries are tested for formal errors, plausibility and logical consistency. The data are stored as published, they are not being standardised. As a rule the setting of the space group chosen by the author is considered as a valuable information that should not be changed.

In particular, the database provides information on: structural data - atomic coordinates of pure elements, minerals, metals, and intermetallic compounds, structural descriptors (Pearson symbol, ANX formula, Wyckoff sequences), bibliographic data, synthesis conditions

A search in ICSD for a phase (defined by temperature, pressure and chemical composition) may yield in more than one entry. There are comments/remarks that give additional information on the quality of the data and refer to other related datasets of the same phase. Entries of "poor quality" are not removed from the database. The final evaluation must be done by carefully studying the results of a search by the user.

Basic search fields are chemical name, mineral name, chemical formula, title of publication, ciation, authors, cell, density, space group, atomic parameters, Wyckoff positions, atomic coordinates, isotropic displacement parameters, site occupation, bonded H-atoms, anisotropic displacement parameters, atomic distances.

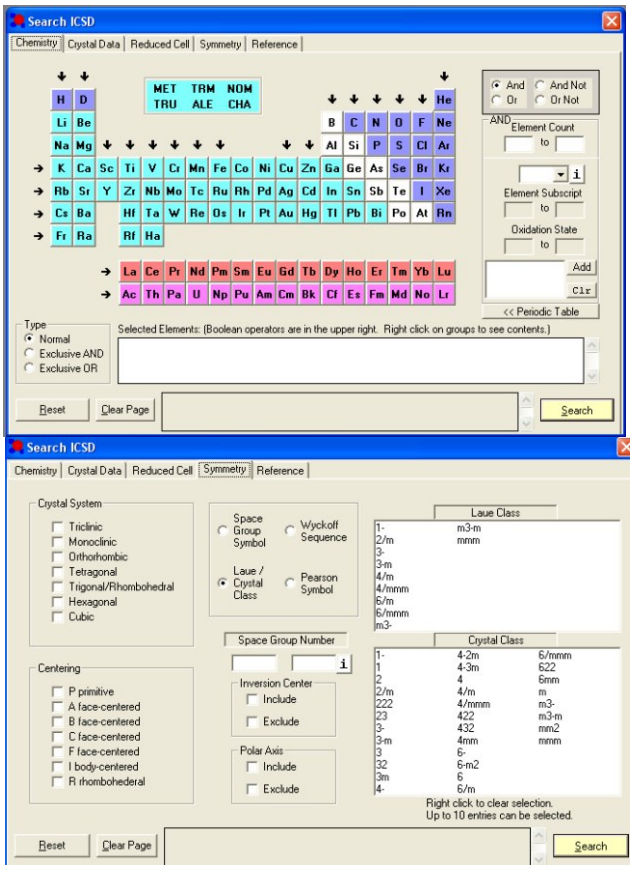
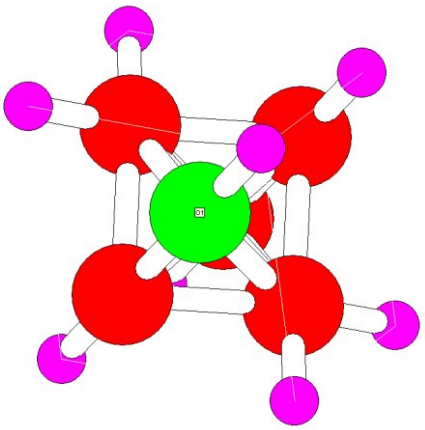
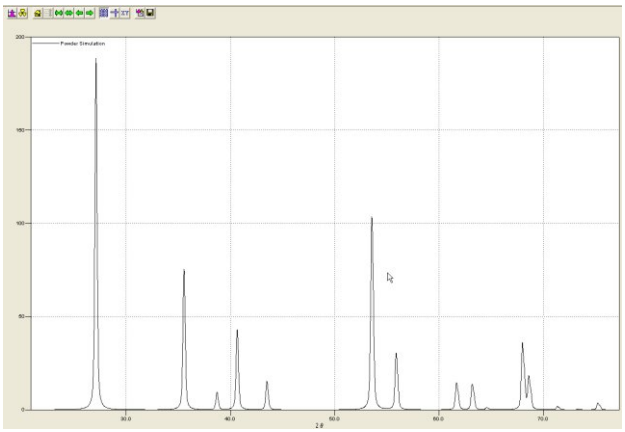


Figure 1. Two search screens in FindIt software.



a)



b)

Figure 2. Screens with crystal structure (a) and powder pattern (b).

There are two basic software tools - PC version with the software FindIt and Web access. Main search screens can be seen on Fig. 1a, b

The software can show all the selected structures in interactive 3D visualization (Fig. 2 a) with several but not too many (with respect to the dedicated visualization software) options. Export to VRML files is possible. Structural parameters can be exported into standard CIF file. For selected structures also powder diffraction patterns can be simulated (Fig 2 b). The following parameters can be given for the simulation - wavelength of radiation, background polynomial coefficients, theta limits, peak shape - pseudo-Voigt, Pearson with necessary shape parameters and U, V, W constants of Cagliotti polynomial describing theta dependence of the peak width. Demo version available on the Web page is fully working but on very limited subset of the database. The annual price of single academic licence is about 435 EUR (incl taxes), institutional license is for 920 EUR. The software and database usage is correspondingly disabled each half of year and a new licence key must be given. However the licence is paid annually.

Internet access has already been available for several years at ILL site in Grenoble (Fig. 3). The access can be based on password protection or by user's IP address. Single acess is for about 1 000 EUR. The Web page can be found at [4]. The features are similar to FindIt software. Structures can be visualized in VRML and powder patterns

Figure 3. Screen with Web interface of ICSD

are generated in PDF files. Both interfaces can be used for free generation of structure pictures and powder patterns for any structure, if the user just fill in correctly the Web forms with the own data.The users without password can run the interface in demo mode. Intranet versions of demo can be installed in Linux, HP, SGI, SUN.

CRYSTMET

CRYSTMET is a database of critically evaluated crystallographic data for metals, including alloys, intermetallics and minerals. Started in 1960 by Cromer and Larson at Los Alamos, its development was continued by the National Research Council of Canada. In 1996 the production and dissemination of CRYSTMET was transferred to Toth Information Systems [5].

Crystmet database contains 109,877 entries now. They include crystallographic data, atomic coordinates and calculated powder diffraction patterns for intermetallics and inorganic materials.

The CRYSTMET database is embedded in a generalized crystallographic database software environment (Fig. 4). This environment provides all the necessary tools for database searching, data analysis and manipulation, crystallographic visualization and associated statistical graphics. All databases within this environment are of a relational type and use the Microsoft Jet relational engine.

Entry into CRYSTMET is via a number of search screens (Fig. 4) which provide intuitive and easy access to searching, analysis and visualization. Major entry points are chemistry, bibliography, cell and powder spectra (d-spacing can be added). Results of queries reside in sets which can be further manipulated using logical operations.

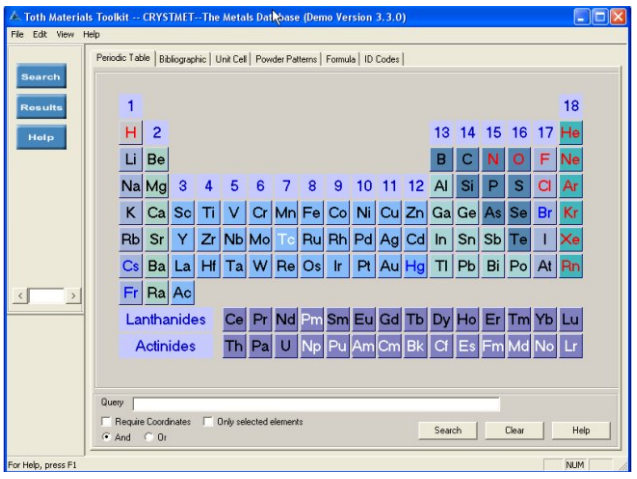


Figure 4. Main screen of the Crystmet database software.

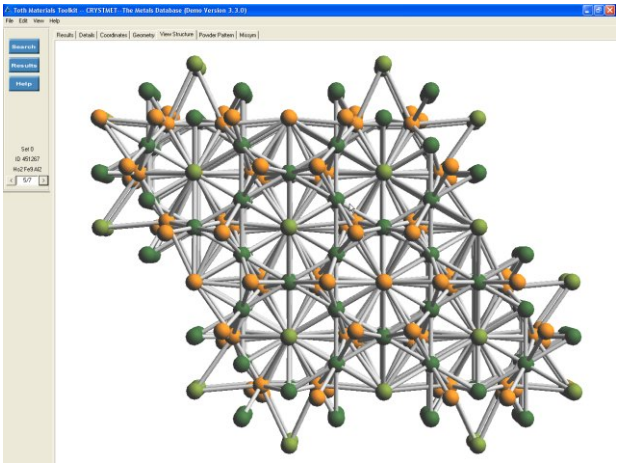


Figure 5. Output screen of Materials Toolkit showing atomic structure.

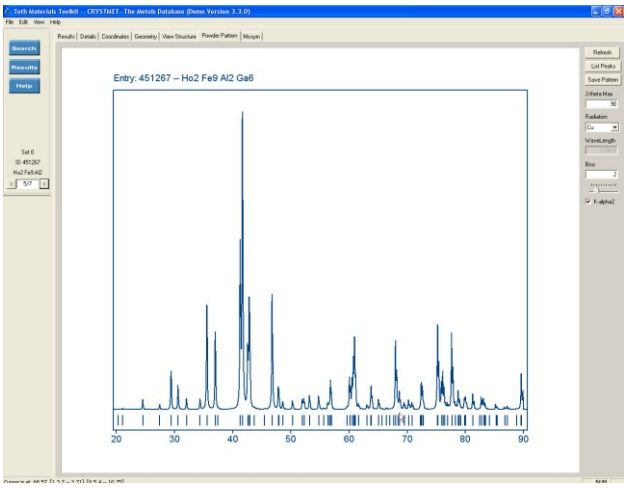


Figure 6. Output screen of Materials Toolkit showing powder diffraction pattern.

The output for each entry can be manipulated using tools provided to calculate geometry, powder diffractometer trace with peak positions highlighted, 3-D structure plots with extensive plot options for compound manipulation. All software analysis, calculation of diffraction patterns and structure visualization are instantaneous, being calculated directly from the database entries. Changing plot options and the calculation of diffraction spectra using a number of X-ray sources can be performed interactively. Plotting tools enable the user to construct a variety of solids such as surfaces, spheres, cylinders and disc from any structure in the database.

Tools of Materials Toolkit software are provided to output the results in a variety of file formats. The collection of fully general and highly graphical tools operates on the structure description stored in an editable ASCII screen. A single command screen puts at the user's fingertips simple utilities like distances and angles or powder-pattern calculations, but also less common interactive modular utilities like editable surface models parallel to any lattice plane; editable content for fully general supercells; adsorption modeling for catalysis studies; epitaxy modeling; push-button derivation of the conventional description within a distance tolerance for any structure model. Additional tools comprise 3D plots of predicted morphology, of Voronoi polyhedra, of voids and channels, of radial plots for selected properties calculable from their tensor expression, like uniaxial compression or shear, or the three speeds of sound, cluster modeling and their calculated diffraction pattern, listing of possible twin laws, an aid for indexing of convergent-beam electron diffraction patterns, easy storage and retrieval of multi-model files etc.

Price of annual single academic licence is about 1800 USD. The software can be run only on single particular machine which is checked by the software.

With respect to FindIt software, there are several more features (see above), the output data are shown in suitable and clear forms. The visualization software (Fig. 5) is of better quality and with more options. By contrast for diffraction patterns (Fig. 6) only wavelength and constant width can be given. One interesting tool is the possibility of continous change of temperature factor with online pattern visualization.

Powder Diffraction File - PDF

This database is well-know by all the powder diffractionists. It was founded in a form of card sets - ASTM, later on as JCPDS and now PDF. Nowadays, it is developed, edited and maintained by the ICDD - the International Center on Diffraction Data [6].

Release 2006 of the Powder Diffraction File contains over 500,000 unique material data sets. Each data set contains diffraction, crystallographic and bibliographic data, as well as experimental, instrument and sampling conditions and select physical properties in a common standardized format. The PDF contains the collective works of > 2,000 journals and hundreds of thousands of authors. Each year, ~50 leading material laboratories are awarded grants to analyze and characterize new materials for the PDF.

The ICDD and the collaborating database organizations each have editorial mechanisms for updating historic data. Updates are frequently made by authors. Corrections in the data are made by users of the database or editorial reviewers. Typically, tens of thousands of historical entries are updated each year.

Traditional form of the database is called **PDF-2** and it is a collaborative product between ICDD, FIZ (Karlsruhe) and NIST (National Institute of Standards). Last issue

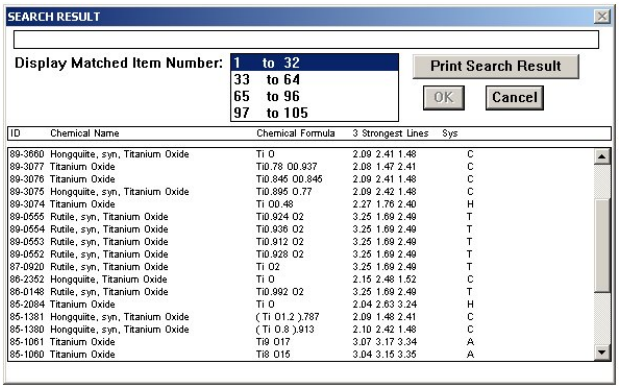


Figure 7. Example of search window of PDF-2

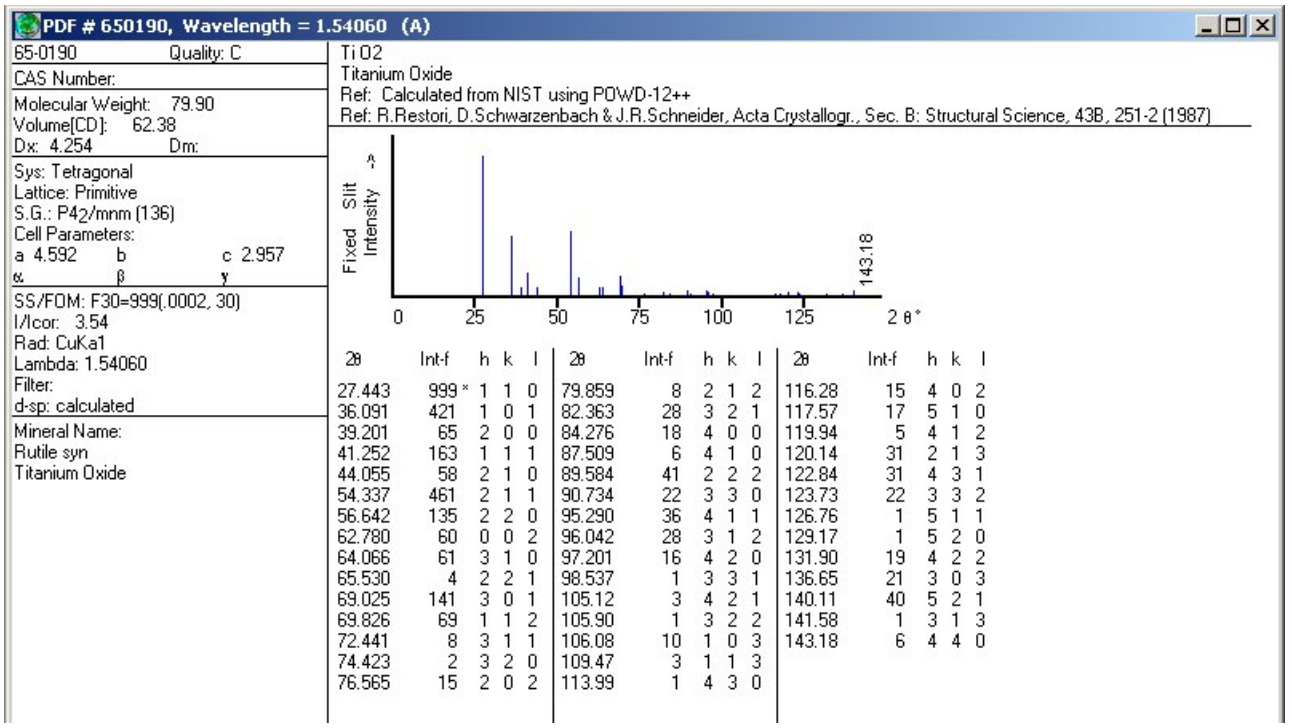


Figure 8. Classical card picture of PDF-2, software PCPDFWIN with typical data - 2 angles, intensities, hkl, reference, chemical name, chemical formula, diffraction method and wavelength, space group, crystal system, lattice parameters, density, molecular weight

(2006) contains about 160 000 inorganic and 29 000 organic entries (98 500 experimental entries, 77 500 entries calculated from ICSD database, 10 000 from NIST database). It must be take into account that quite often more records correspond to single phase. Each record contains table of interplanar spacings (*d*), relative intensities (*I*) and often also diffraction indices (*hkl*). In addition, chemical formula, chemical name, mineral name for minerals, crystal system, some physical characteristics, experimental parameters, bibliographic information and mark of data quality. Each entry is indexed for search in subfiles as minerals, organic compounds, polymere, metals and alloys, pharmaceuticals, ceramics etc.). The price of single academic license is about 5800 USD, annual renewal is for 1150 USD but this is not required in order to run the software. Search results window (Fig. 7) shows list of hits - entry number, chemical name, chemical formula, interplanar spacing of the three strongests lines and symbol of crystal system. Classical card picture is shown on Fig. 8. Older search software **PCPDFWIN** has been replaced by **DDView.** It allows to use different search criteria and their boolean combinations (chemical elements, chemical and physical characteristics, reduced cell size, centering, names, references and mainly interplanar spacing of strong or long lines).

Electronic search manual **PCSIWIN**, has been replaced by **SIEVE**. The search is based on Hanawalt, Fink and Long-8 lists. Input for the program consists of pairs *d/I* or 2 */I* and wavelength. Different filters can be combined. A flexible text importer can read a wide variety of input files. SIeve is linked with Global searches and Search Preferences from DDView. The results of SIeve can also be integrated into search criteria in DDView.

Preferred form of database distribution is, however, **PDF-4+**. It is complete change of conception. This database has comprehensive material coverage for inorganic materials and contains numerous features such as digitized patterns, molecular graphics and atomic parameters. It is working more with total powder patterns rather than pure *d-I* pairs. These patterns are calculated on-the-fly either

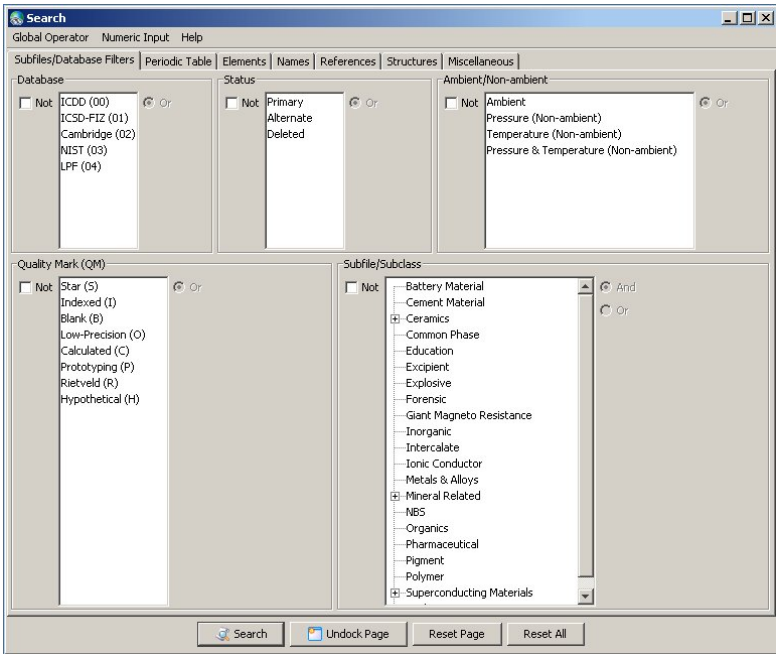


Figure 9. Search screens of PDF4+ with different subfiles (top), elemental choice (bottom). Additional options are for example names, references, structural limits (symmetry, cell)

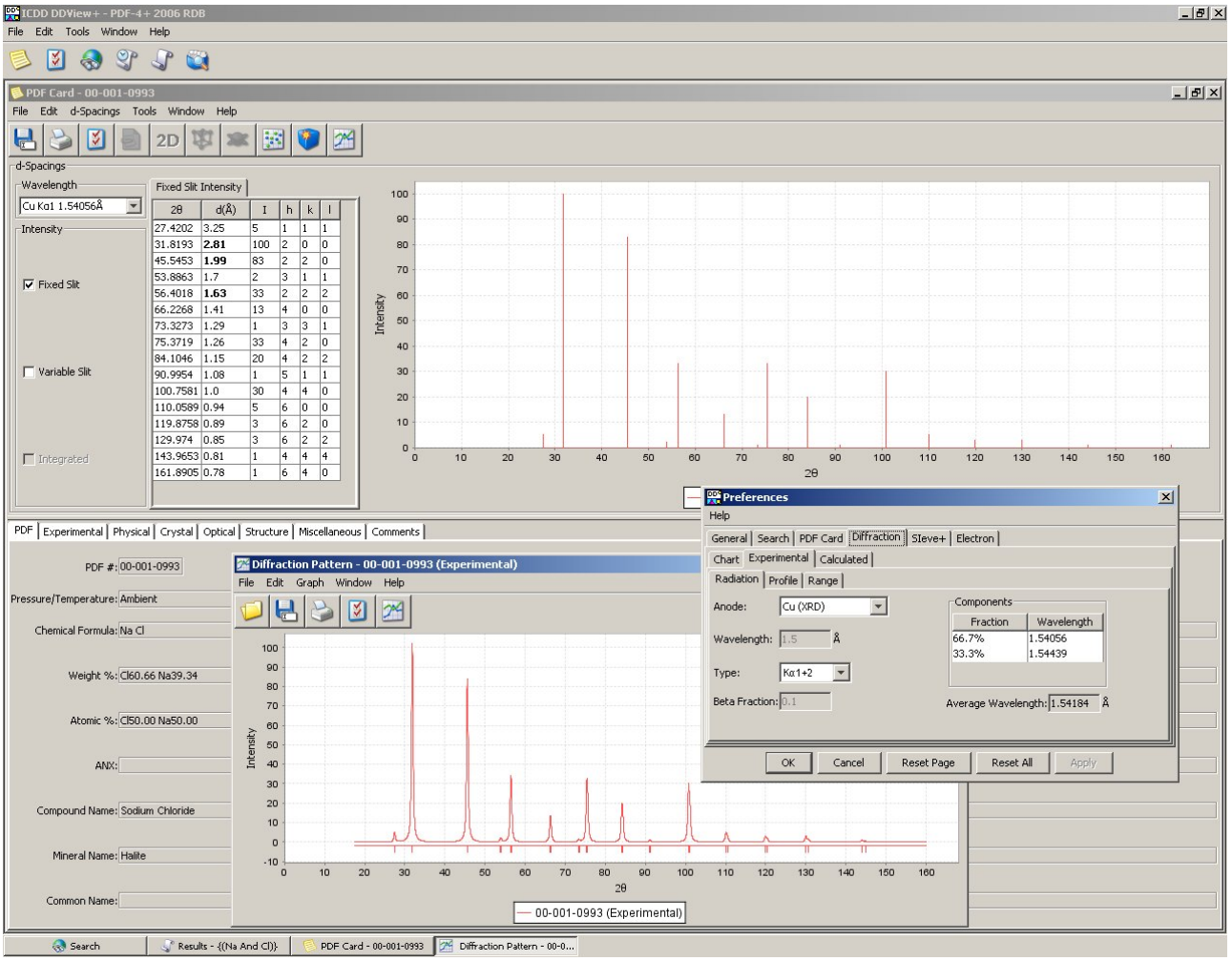


Figure 10. New card view in PDF-4+ DDVIEW software. In addition to line diffraction pattern, full apttern can be calculated with different options (radiation, profiles functions - pseudo-Voigt, Thomson-Cox-Hastings with full control of line broadening behaviour (U, V, W, A, B parameters well-known in the Rietveld analysis).

from structural data or experimental diffraction patterns, when certain parameters describing instrumental and physical effects. Database is created as a set of tables and it is so a classical relational database. Software is integrated in the systems. The database is distributed in several versions - PDF4+, Minerals and Organics. PDF4+ contains 255 000 entries (229 000 inorganic) and the price of single academic licence is 5 775 USD with renewal 1 150 USD each year. However, unlike PDF2, the software cannot run unless the licence is not regularly renewed. Morover, it is limited strictly to one particular machine. Organics 2007 version costs 6 500 USD with annual renewal of 2 940 USD. This database was created in collaboration with the CCSD and it contains over 300 000 entries. In addition to features of PDF4+ it includes also 2D chemical structure and additional search criteria specific for organic materials can be applied.

The program **DDVIEW+2006** contains 44 different search methods to filter the database contents with custom display of a selection of 65 separate data fields. Using a Java™ point and click interface, various search methods and field selections can be combined to produce a nearly limitless choice of data mining options (Fig. 9).

PDF-4+ is designed to support automated quantitative analyses by providing key reference data required for these analyses. Every pattern, independent of type or source, can be displayed as a digitized pattern, enabling an ability to perform total pattern analysis. Three different algorithms are used to calculate the display, depending on the source of the data and information required. PDF-4+ automatically chooses the appropriate algorithm for display purposes and then allows the user to change the display based on their local conditions. The user has the ability to apply various common instrumental parameters, optical configurations, and experimental parameters, including crystallite size, into the digital display (Fig. 10).

The atomic coordinate parameters provided in 98,291 entries, from the LPF and NIST sources, permit several calculations based on scattering and diffraction physics. Rietveld techniques can use the atomic parameters for quantitative analysis and structural determination. The PDF-4+ database contains several integrated scattering factor tables so that electron diffraction, variable wavelength X-ray diffraction, and neutron diffraction patterns can be calculated from first principles.

Crystal Lattice Structures

In demonstration database of crystal lattice structures [7] all individual structure types can bee found classified according space groups, Pearson symbols or *Strukturbericht* symbols (Fig. 11). It includes 273 structures in 94 of the 230 space groups. Web interface offers static views of all the structures from basic directions (Fig. 12), posibility of rotation (Java applet), atomic coordinates, primitve and basal vectors. The database is suitable for looking for any structure type and can be used also for generating of useful pictures without commercial databases.

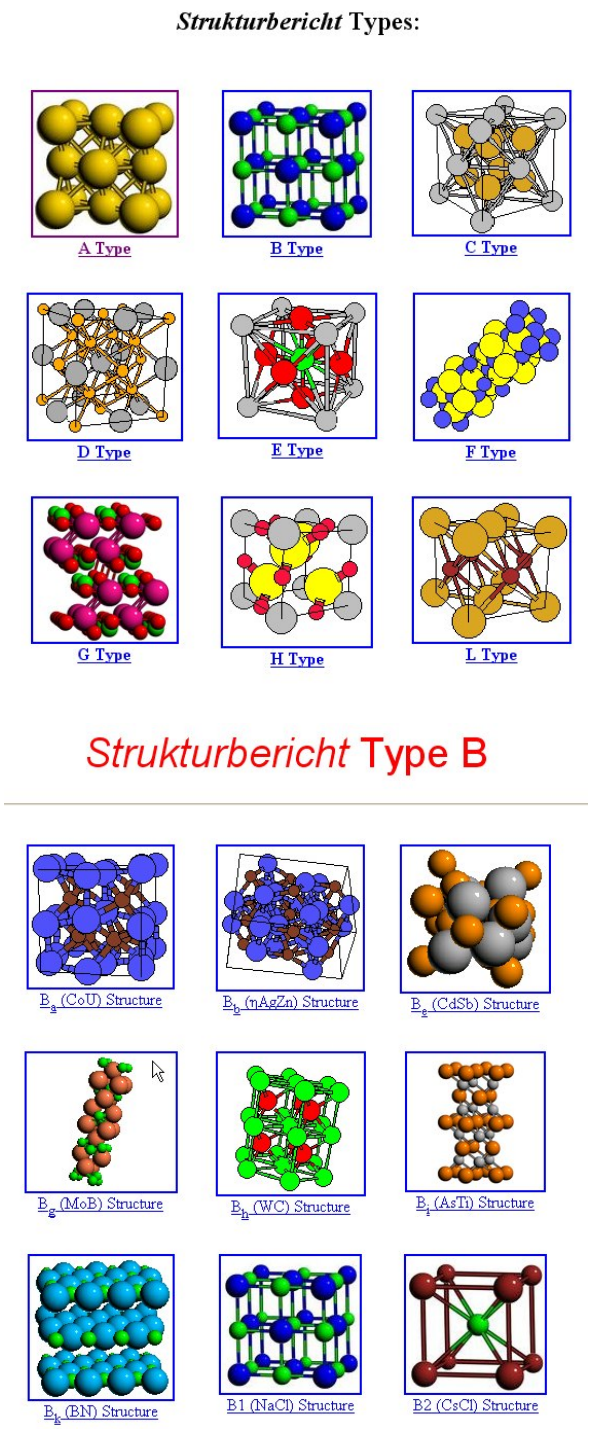


Figure 11. Parts of basic menus of Crystal Lattice Structures site

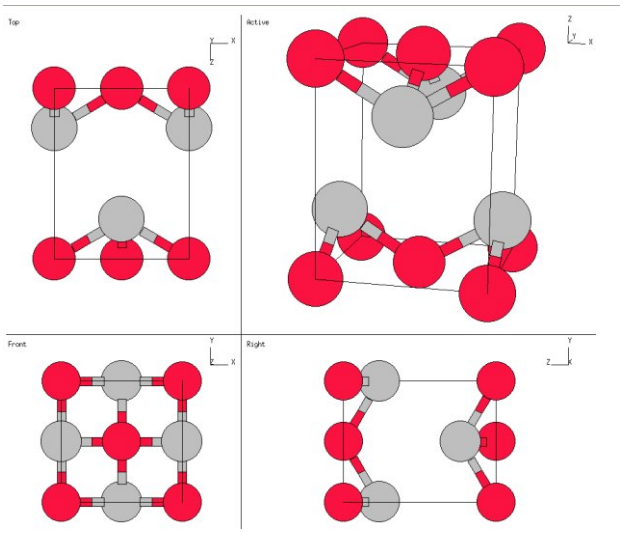


Figure 12. Static views on structure in Crystal Lattice Structures.

Pauling file. Pearson's Crystal Data

The PAULING FILE project is a collaboration between Japan Science and Technology Corporation (JST) and Material Phases Data System (MPDS). The project started 1995. National Institute for Materials Science (NIMS) obtains the right to provide the online service of this system. The Pauling File aims at a comprehensive materials database which covers all non-organic solid state materials and consists of structure, diffraction, constitution, and physical property data.

The source of Pauling File data are around 150,000 original publications taken from more than 1,000 scientific journals since 1900. The data are processed by an international, highly-experienced group of scientists with going through a sophisticated data evaluation, standardization and derivation procedure.

The present Pauling File includes about 80,000 structure entries, 34,000 diffraction entries, about 52,000 property data counts, about 6,000 constitution entries and 6,000 images of phase diagram. The data of binary is accomplished, while of multinay is at an underdeveloped stage.

The main idea came from P. Villars and the database is distributed on CD ROM by ASM International (price 1950 USD, single user, [8]) - as Pauling File: Inorganic Materials Database and Design System-Binaries. It consists of binary phase diagrams, crystallographic structure, including lattice parameters, structure types, viewable atom positions, etc., diffraction data and patterns, including 'inferred' patterns from crystal structure data, physical properties of binary crystalline materials.

Search can be done in more than 200 fields for: journal, author, country, organization, chemical composition, atomic environments, physical properties (mechanical, thermal, electrical, magnetic, optical), structure determination details. The software also allows to search and plot different statistical dependencies, find interesting rules and correlations. Therefore it is said to be suitable for materials design.

The file can be used on-line at [9] but free registration is required. Original software - database browser is distributed by ASM, demo can be found at [8].

Pearson's Crystal Data is the is the software front-end for the database on the Microsoft Windows platform and it is distributed by Crystal Impact from this year [10]. The first version will contain 150,000 structural data sets (including atom coordinates and displacement parameters, when determined) for more than 95,000 different phases, roughly 12,000 experimental powder diffraction patterns and about 130,000 calculated patterns (interplanar spacings, intensities, Miller indices).

Selected features of the program are for example following.

Visualization (3D pictures) of crystal structures, radii/volumes diagram for comparison of similar crystal structures , chemical system (system matrix), phases list, link to phase diagrams online, link to original publication (if available online), visualization of atomic environments/coordination polyhedra, manual measurement of selected distances and angles, calculated powder pattern for user-defined wave length, published pattern (if available), table of distances and angles: 3D picture of selected atomic environment, distance statistics histograms

Search for: interatomic distances, phase information (e.g. phase formula, phase prototype, mineral name), chemical composition, atomic environment (coordination number, atom coordinations), crystallographic data and classifications, structure determination details, processing information (e.g. PDF-number), bibliographic data

Complete flexibility to edit an entry report by adding a column with another database field, or moving or deleting columns, as well as sorting

Comparison of nearest neighbor histogram of any specific atom with its statistical plot containing all distances in the database

Dynamic changes of the selected atomic environments by clicking on its nearest neighbor histogram

Searching for pre-defined atomic environments (coordination polyhedra) including selection criteria to the central atom and to the coordination atoms, as well as its interatomic distances

A session together with its answer sets, selection criteria, and visited views can be stored and used as a starting point for next session.

Typical program screen is on Fig. 12.

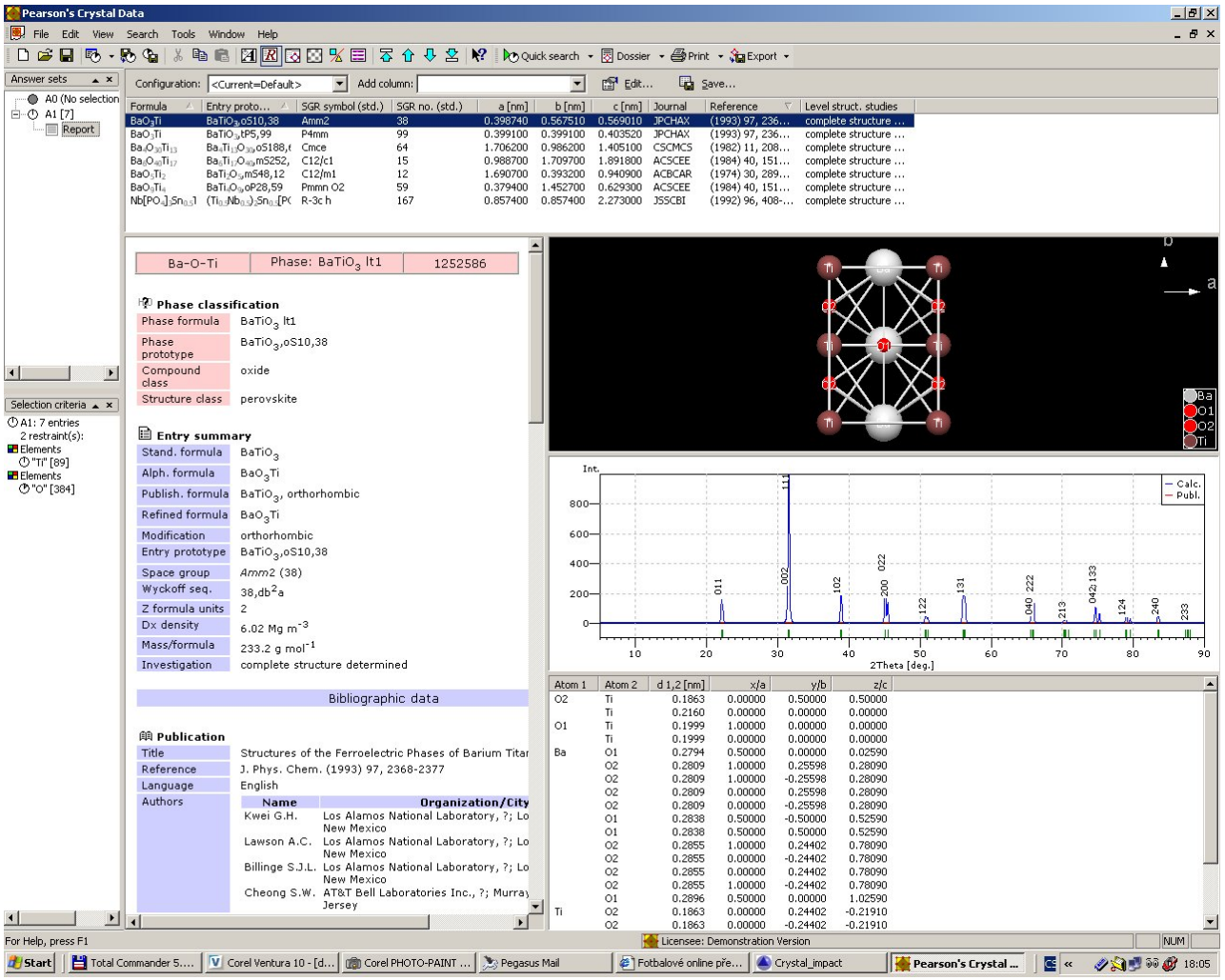


Figure 12. Typical screen of Pearson's Crystal Data (Crystal Impact). Separate parts of the window display - structural data, interactive structure view, powder pattern, atomic positions.

NIST databases

Two structural database are distributed by NIST.

NIST Surface Structure Database - SSD [11]. The database is the only complete critical compilation of reliable crystallographic information now available on surfaces and interfaces. SSD brings instant access to detailed text and graphical displays of over 1250 experimentally-determined atomic-scale structural analyses.

Crystal Data [12] database include the standard cell parameters, cell volume, space group number and symbol, calculated density, chemical formula, chemical name, and classification by chemical type.

Other useful databases can be found at the site.

Mineralogical databases

It is interesting to note that some mineralogical databases are distributed free. Two of them should be mentioned here - **Mineralogy Database** [13] and **Mincryst** [14]. The formed contains 4442 minerals. The information included is for example - chemical composition of minerals, crystallographic information (powder pattern, unit cell), photos of minerals (Fig. 13), physical characteristics like hardness, density, optical properties) and mineralogical classification - Dana's New classification, Strunz classification, alphabetical listings of mineral species. There also are extensive links to other external sources of mineral data and information. As a new feature interactive structure pictures can be shown by using applet (similar to Crystal Lattice Structures applet). The latter, Russian database (Institute of Experimental Mineralogy, Moscow), is called "Crystallographic and Crystallochemical Database for Mineral and their Structural Analogues" and it is a combination of structural database, theoretical powder patterns and software. Two versions are available: *stand-alone* and *WWW*. Over 5000 entries for around 2500 unique mineral names can be found there [14]. Each entry has a mineral name, chemical formula, space group, lattice parameters, atomic coordinates, temperature factors, occupation and bibliographic information. Software package includes calculation of powder pattern or mixture of phases, standardless phase analysis, structure visualization etc.

Other databases

Specialized databases can be found for incommensurate structures [15] and zeolites [16].

Crystallographic server in Bilbao [14] provides information on space groups (e.g. generators of positions in individual space groups, subgroups etc.) similar to the information in International Tables for Crystallography vol. 1.

Data of general interest like periodic element tables [18-20]. The last one [20] is focused on X-ray characteristics.

Acknowledgement

This contribution was parially supported within the framework of the program MSM 0021620834 financed by the Ministry of Education of the Czech Republic.

Figure 13. Mineral picture from Mineralogy Database.

References

1. http://www.iucr.org/iucr-top/data/index.html.
2. http://www.fiz-karlsruhe.de/icsd.html.
3. Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. *Acta Cryst.* B58 (2002), 364–369.
4. http://icsdweb.fiz-karlsruhe.de.
5. http://www.tothcanada.com.
6. http://www.icdd.com.
7. http://cst-www.nrl.navy.mil/lattice/.
8. http://www.asminternational.org/.
9. http://crystdb.nims.go.jp/
10. http://www.crystalimpact.com.
11. http://www.nist.gov/srd/nist42.htm.
12. http://www.nist.gov/srd/nist3.htm.
13. http://webmineral.com/.
14. http://database.iem.ac.ru/mincryst/.
15. http://www.cryst.ehu.es/icsdb/index.html.
16. http://www.iza-structure.org/databases/.
17. http://www.cryst.ehu.es/.
18. http://www.webelements.com/.
19. http://www.physics.nist.gov/PhysRefData/Elements/cover.html
20. http://www.csrri.iit.edu/periodic-table.html.