MedeA InfoMaticA: Rapid Access to Comprehensive Experimental Structure Data

Contents

1 Overview

MedeA InfoMaticA provides a graphical user interface to the experimental databases COD, ICSD, NCD, and Pearson's. Together these databases contain over 1.2 million structure entries, thus representing one of the largest integrated sources of solid-state structure data accessible today.

InfoMaticA

Structure data from powder diffraction and scattering experiments

Lattice parameter and atomic positions

Bibliographic references

Experimental setup

Symmetry, stoichiometry, partial disorder

Phase diagrams

InfoMaticA provides an easy-to-use keyword search language to formulate complex queries on the full set of experimental and computed data available.

MedeA offers a simple interface to store computed structure data in *InfoMaticA*. You can build up your database of computed structures along with comments and references and search this database in combination with the experimental ones.

DOCUMENTATION

2 *InfoMaticA*'s Structure Databases

2.1 COD (Crystallography Open Database, University of Cambridge, UK)

COD is an open-access collection of crystal structures of organic, inorganic, metal-organic compounds and minerals, excluding biopolymers. People registered to COD can add published and unpublished structures of small molecules and small to medium-sized unit cell crystals. As of November 2022, the database has more than 495,000 entries. COD is an academic project, and its development has been published in several peer-reviewed papers. More information: MedeA COD [1]

2.2 ICSD (Inorganic Crystal Structure Database, FIZ Karlsruhe, Germany)

ICSD contains crystal structure information for non-organic compounds mainly including ceramics and minerals and since recently also metallic systems. ICSD entries go back to 1915; currently, the database has around 185,000 structure entries. More information: MedeA ICSD [2]

2.3 Pearson's File (ASM International, USA)

The Pearson's data file is a crystallographic database published by ASM International and originates from the well-known Pearson Crystal Structure Database [3]. The database contains crystal structures of the full range of inorganic compounds. The present *MedeA* release contains close to 303,000 structural data sets (including atom coordinates and displacement parameters, when determined) for more than 95,000 different phases. More information: MedeA Pearson [4]

2.4 NCD (NIST Crystal Data, National Institute of Standards and Technology, USA)

NCD contains chemical, physical, and crystallographic information on approximately 237,000 inorganic and organic crystalline materials. NCD includes reliable data across the entire range of solid-state materials including minerals, inter-metallic phases, metals, alloys, drugs, antibiotics, and pesticides. Data included in NIST comprise standard cell parameters, cell volume, space group number and symbol, calculated density, chemical formula, chemical name, and classification by chemical type. The National Institute of Standards and Technology (NIST), USA, maintains NCD. More information: Medea NCD [5]

3 *Materials Design Database* for Computed Data

Computed structures can be saved to the "computational database" aka *Materials Design Database*. To save a structure, bring up the structure window in *MedeA* (e.g. by loading a minimized structure from a previous calculation) and select Save to database from the *MedeA* File menu. When saving a structure, you should provide a name and comments to save along with the structure. Structures saved this way can be retrieved through a standard search request in the *InfoMaticA* dialog along with experimental database entries. Structures saved to the *Materials Design Database* are tagged with *MD*.

Hint: The *Materials Design Database* file is *MaterialsDesign.db* which is located in the *MedeA* installation directory, namely *MD/Databases*. This file can be easily overwritten upon updating *MedeA* or installing a newer version over an older *MedeA* version. The MedeA installer is prepared for such situations and asks to overwrite and replace existing *Materials Design Database* files.

---

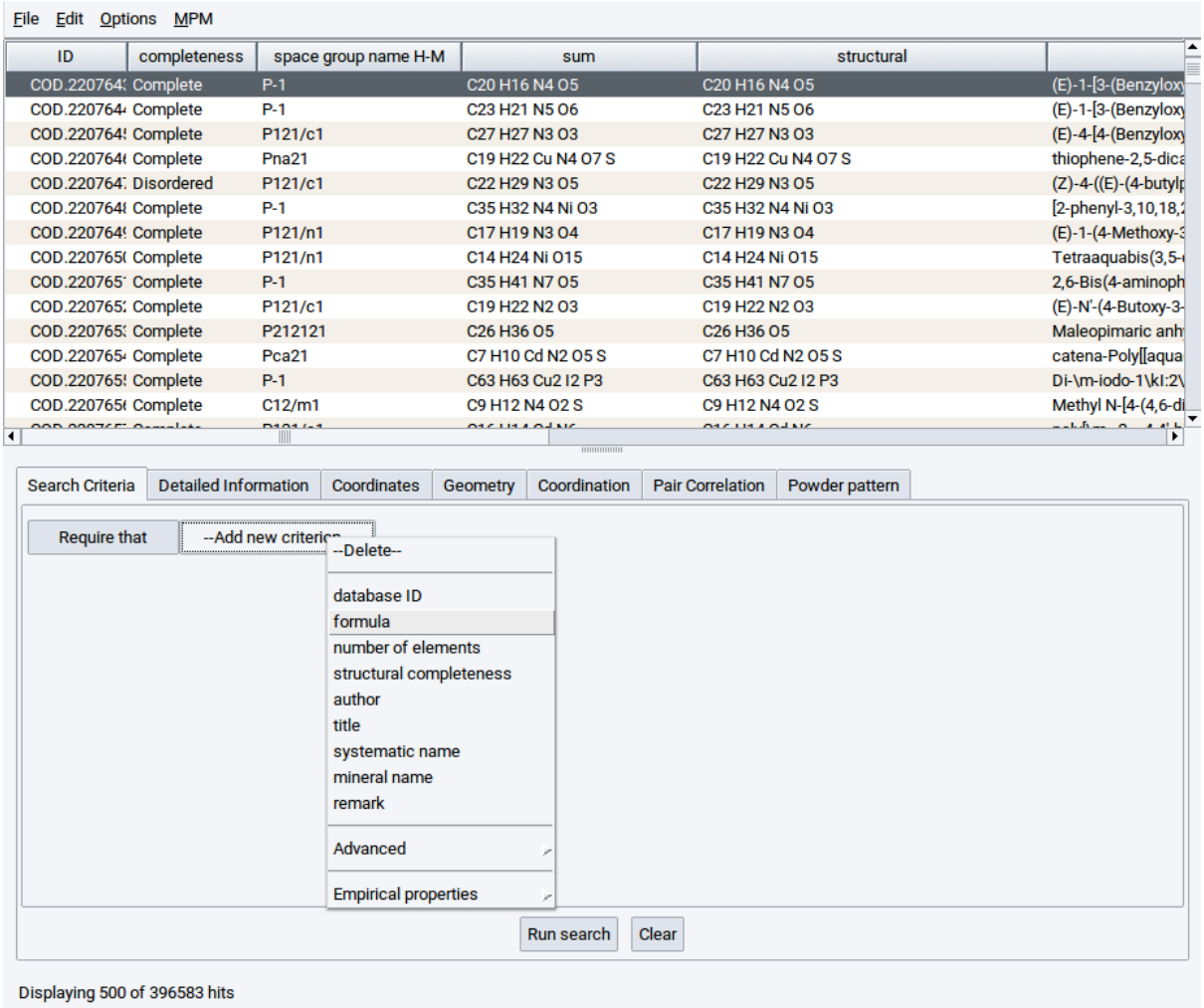
[1] http://www.materialsdesign.com/datasheet-cod
[2] http://www.materialsdesign.com/datasheet-icsd
[3] https://www.asminternational.org/materials-resources/online-databases/pearson
[4] http://www.materialsdesign.com/datasheet-pearson
[5] https://www.materialsdesign.com/datasheet-ncd

4 Starting *InfoMaticA*

To bring up the *InfoMaticA* entry in the *MedeA* main menu bar, select the InfoMaticA item from the *MedeA* Tools menu. Next click InfoMaticA >> Search to bring up the *InfoMaticA* search window/dialog.

The main elements of the *InfoMaticA* interface are the following (top to bottom):

- A menu bar containing entries Files, Edit, and Options.
- A configurable table for displaying search results in the upper region
- An adjustable separator
- Several panels in the lower region to formulate search phrases/criteria and to visualize structure specific results and properties



5 Searching and Retrieving Data

Start your search in the tab labeled *Search Criteria* by formulating a search phrase. To do so, click on – Add new criterion – and select a search option. The following criteria are available:

- – Delete – : Remove search criterion
- Database ID : Provide identification numbers (IDs) under which structures are registered in the *MedeA* databases

- formula : Provide an entire chemical formula or parts of the formula
- number of elements : Allow a certain number or range of elements in the formula
- structural completeness : Display only structures with with particular structure type (select one option of the options complete , missing atoms , disordered , none available)
- author : Search for names of people who published relevant structure data in the literature
- title : Search in the titles of publications with structure data
- systematic name : Search by the systematic name of compounds (IUPAC nomenclature)
- mineral name : Search by mineral name of compounds (trivial/non-systematic name)
- remark : Search for keywords in the content of *Remark section* of structure records
- Advanced : Define additional search criteria (can be database specific)

Note that, as you select a search criterion from the selection, additional *sub-options* may appear on the right. Click these options to create a complete search phrase.

For example, the formula criterion has the following sub-options:

- is [ ]: Exact chemical formula, e.g. Cr_2O_3 , Be, TiC, etc.
- contains exactly [ ] atoms of [ ]: Number and type of atoms, e.g. contains exactly 3 atoms of Fe
- contains any number of atoms of [ ]: elements of a particular group or period of the periodic system of the elements, e.g. contains any number of atoms from group 3 (Sc..La Ac)
- contains less than [ ] atoms of [ ]: upper boundary for the number of particular elements
- contains more than [ ] atoms of : [ ]: lower boundary for number of particular elements
- contains from [ ] to [ ] atoms of [ ]: define a range of number of particular elements
- does not contain atoms of : [ ]: Exclude elements
- has a ratio of [ ] atoms of [ ] to [ ] atoms of [ ]: The exact ratio of two elements

---

Hint: Recognition of element symbols is case sensitive! For instance, the search algorithm has to decide whether structures with cobalt are requested (element symbol Co) or the structures that contain CO (carbon monoxide) molecules. Another example: Should the compound contain three units of bismuth (Bi_3) or BI_3. The latter is a binary compound that is formed by boron and iodine.

Some criteria like author , remarks , and title let you search for text strings of the data files. These options have the following sub-options:

- Is [ ]: Search for an exact text string
- Is Not [ ]: Search for anything but this exact string
- Is like [ ]: Search for a fraction of a text string defined in the field that is part of another text strings as names, terms, etc.. Use wildcards *?* for a single character and *%* or *** for multiple characters.

- Is Not Like [ ]: Search for text strings that do not have the fraction of the text string defined in the field
- Contains [ ]: Search for a substring

Furthermore, a given search statement can be made active, non-active or conditional by clicking on Require that on the left-hand side of each search phrase and selecting either of the options given below:

- Require that : Require that the search criterion in the present line is fulfilled
- If present : Require criterion in present line to be fulfilled if present or possible
- Ignore : Criterion is ignored

Once you have built a search phrase, click Run search to perform the search and display results. If needed, add additional criteria using a second line and so forth, until you are satisfied with the results. With Clear all search criteria can be erased.

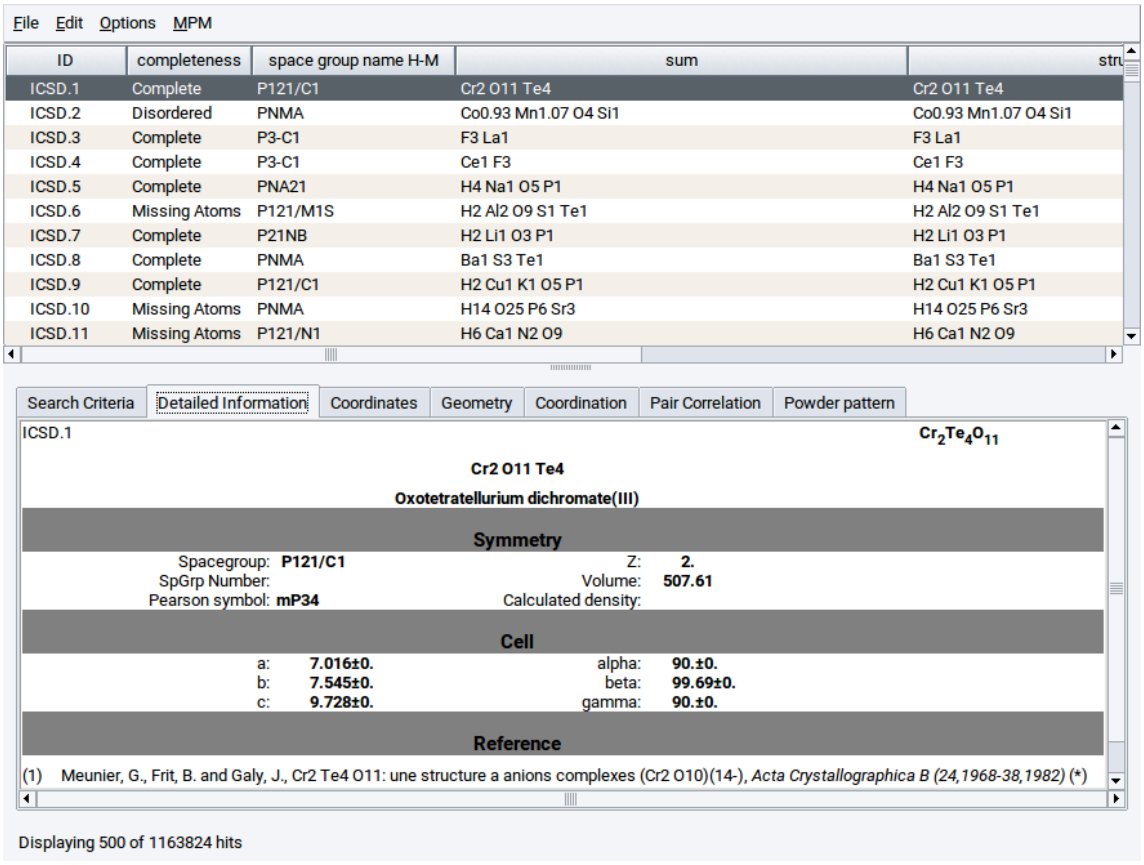
Note: There is **no Undo** function to bring back search criteria that were cleared!

---

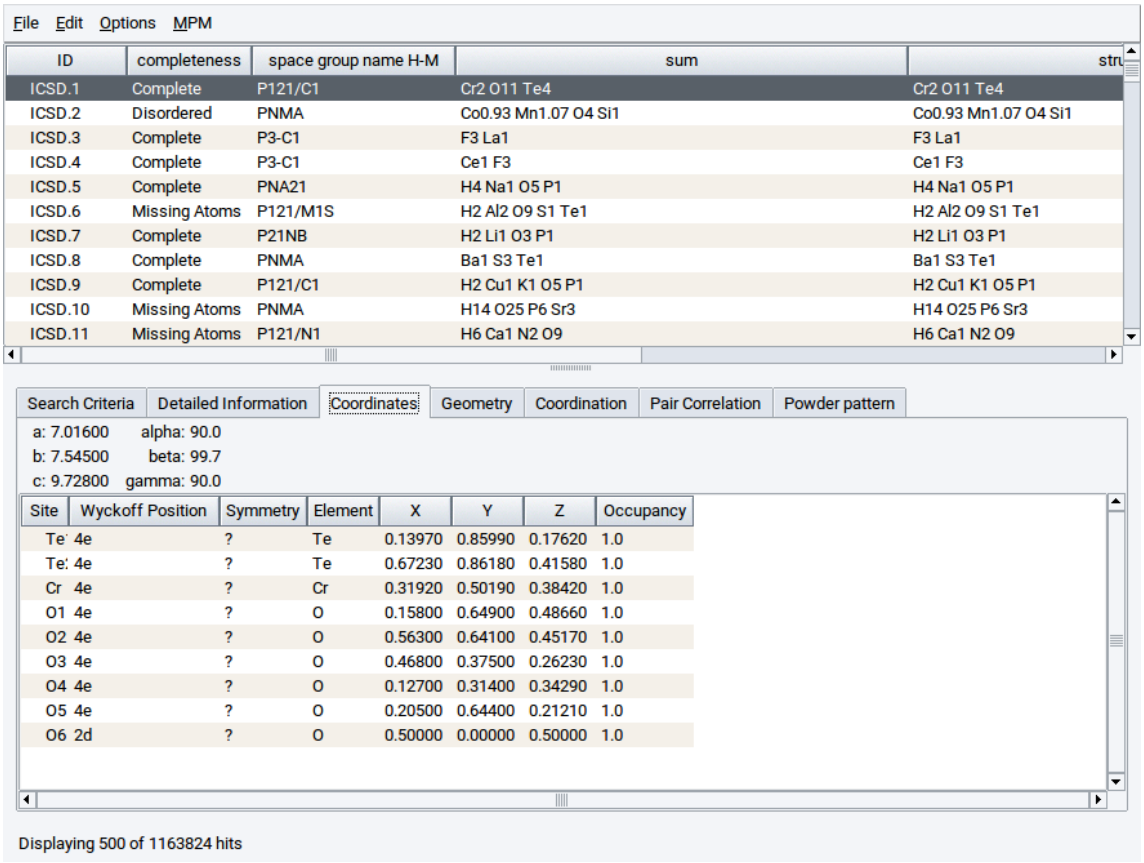
6 Displaying Detailed Information

To get more information on a structure entry, highlight the corresponding row in the table of results with the pointer (simple click onto the record). Afterward switch to one of the following tabs:

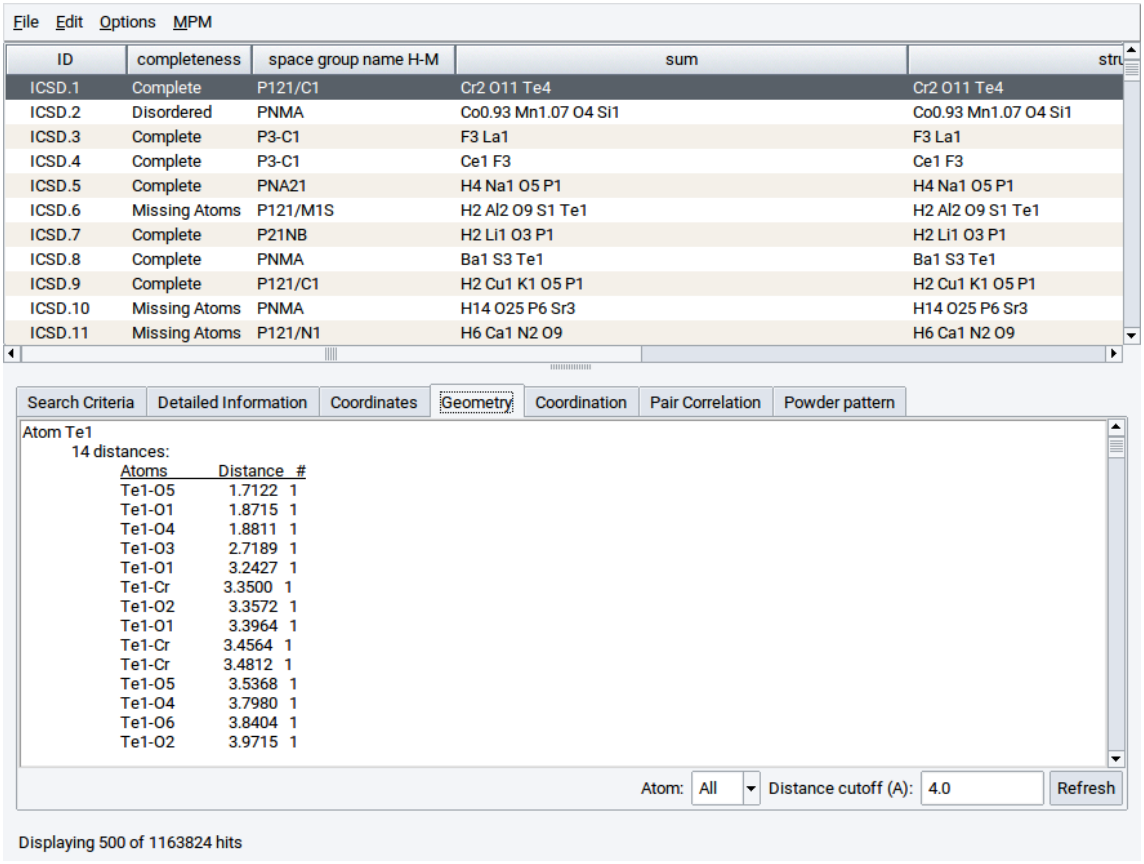
- Search criteria tab: Formulate search phrase and run searches on all databases
- Detailed information tab: Display main structural data, symmetry data and bibliographic references, and remarks (if present)



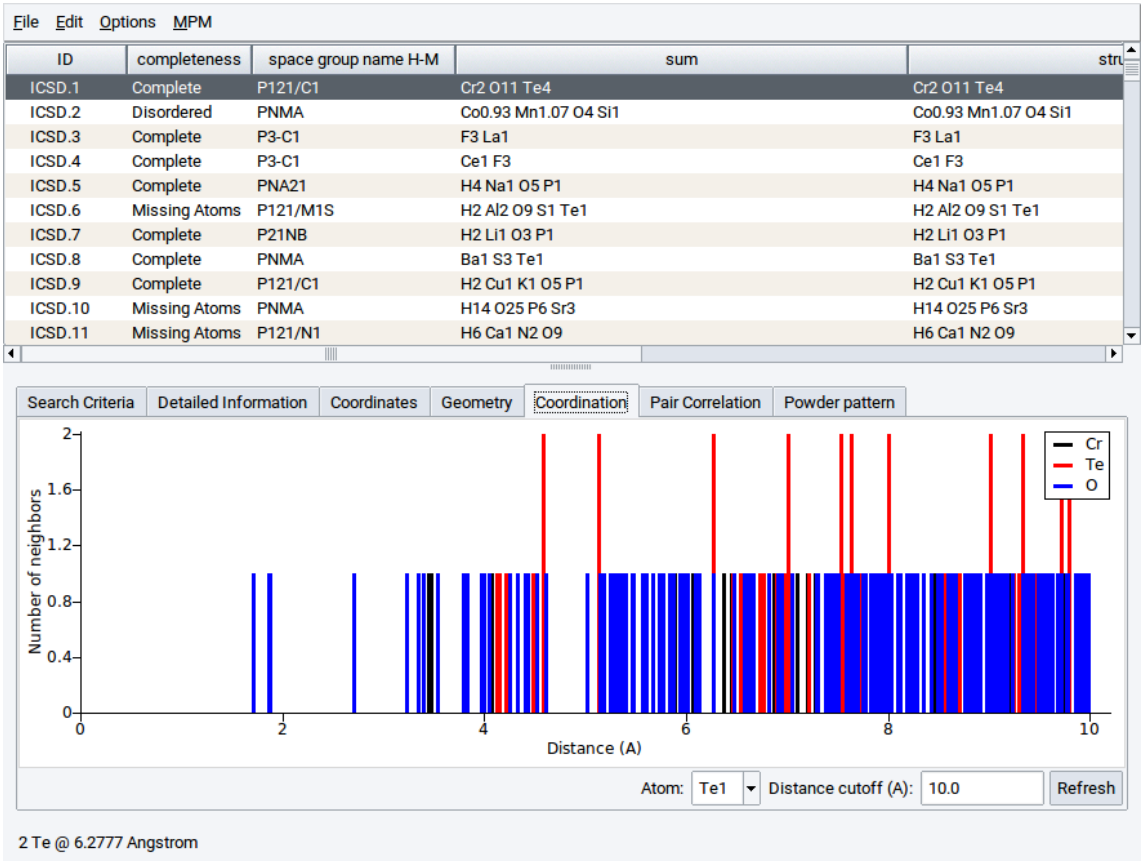
- Coordinates tab: Display positions of atoms, Wyckoff sites and their occupancies of structure records that have atomic coordinates



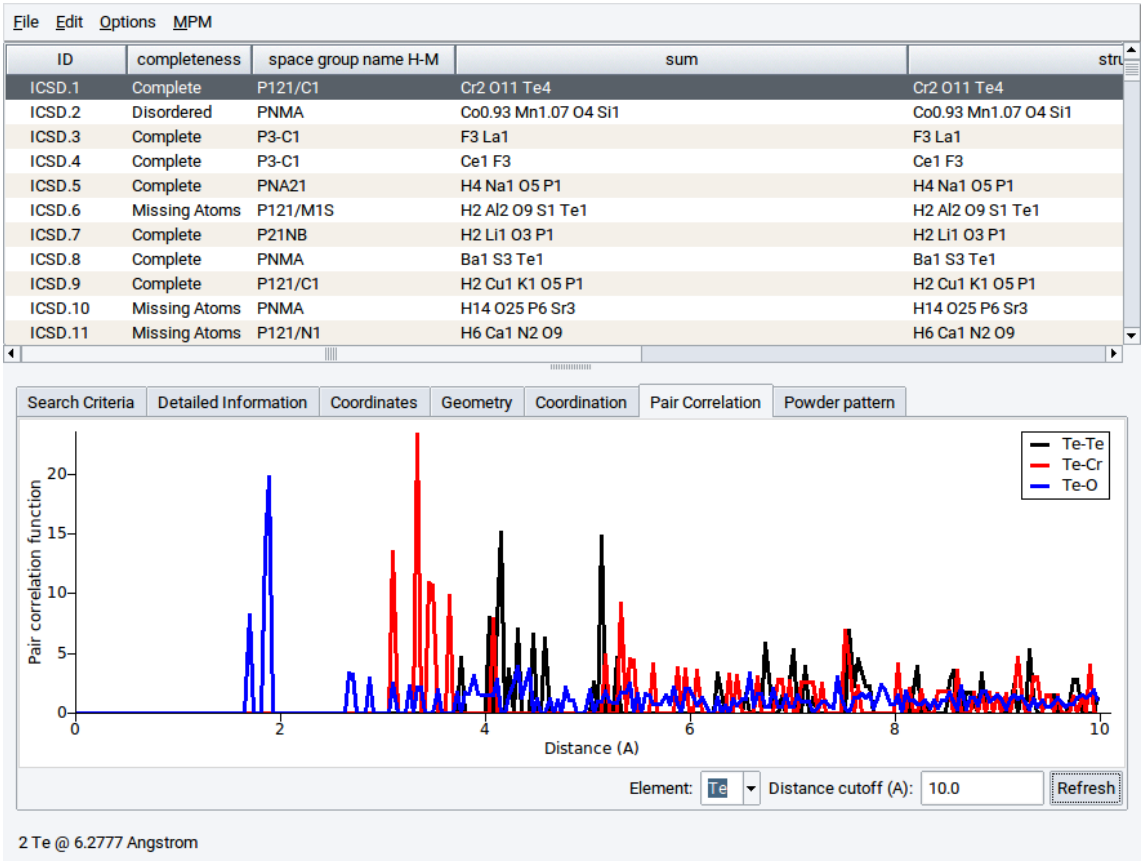
- Geometry tab: Display bond lengths and angles of structure records with atomic coordinates; define for which constituting Atom: and the corresponding Distance cutoff (A): in Å and confirm with Refresh



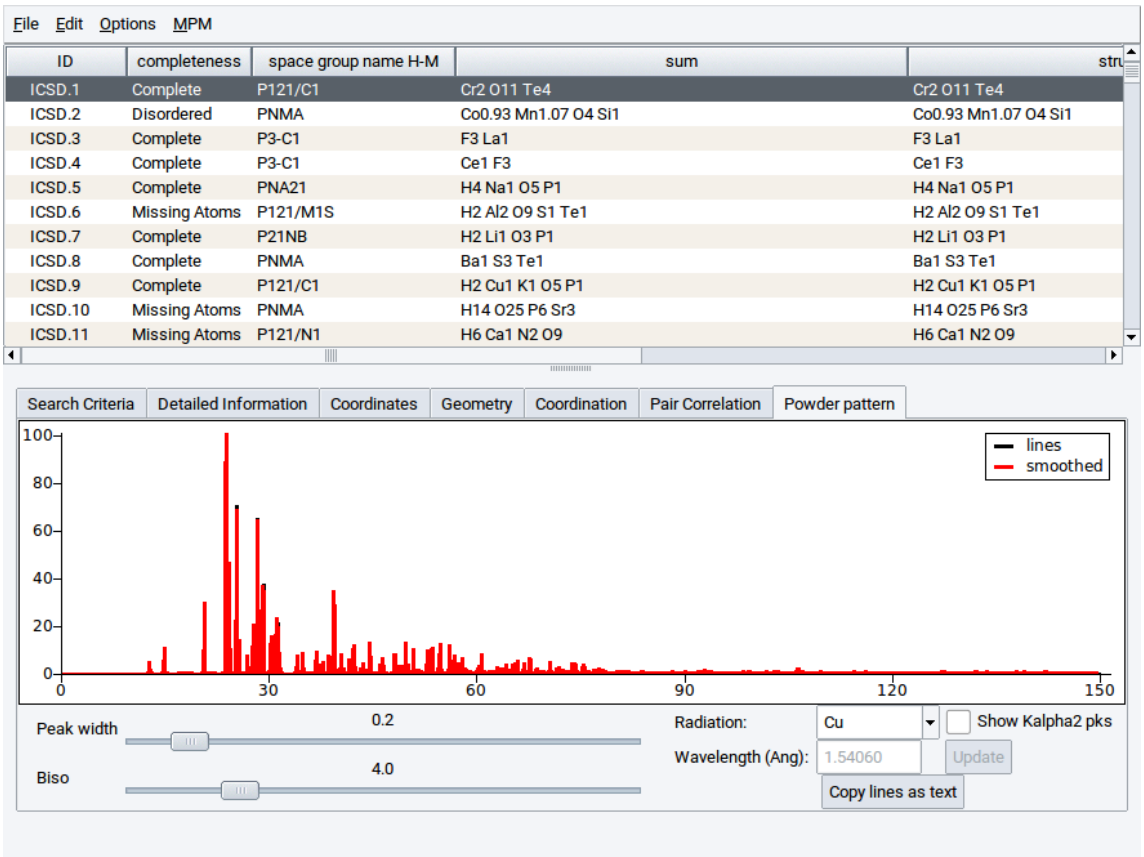
- Coordination tab: Select a lattice site (atom position) and display a chart with the nearest neighbors, next-nearest neighbors, etc. (requires that structure records have atomic coordinates) Define for which constituting Atom: the coordination shell should be calculated, the corresponding Distance cutoff (A): in Å and confirm with Refresh



- Pair Correlation tab: Chart to show which pairs of atoms of which distance; by default pair distances of all atoms are shown; select an element of the relevant system to display specific distributions of pair distances (requires that structure records have atomic coordinates) Define for which constituting Atom: the pair correlation should be calculated, the corresponding Distance cutoff (A): in Å and confirm with Refresh



- Powder pattern tab: Show the powder diffraction pattern of the selected structure; can be calculated for different radiation sources and manually defined radiation wavelengths (requires that structure records have atomic coordinates)

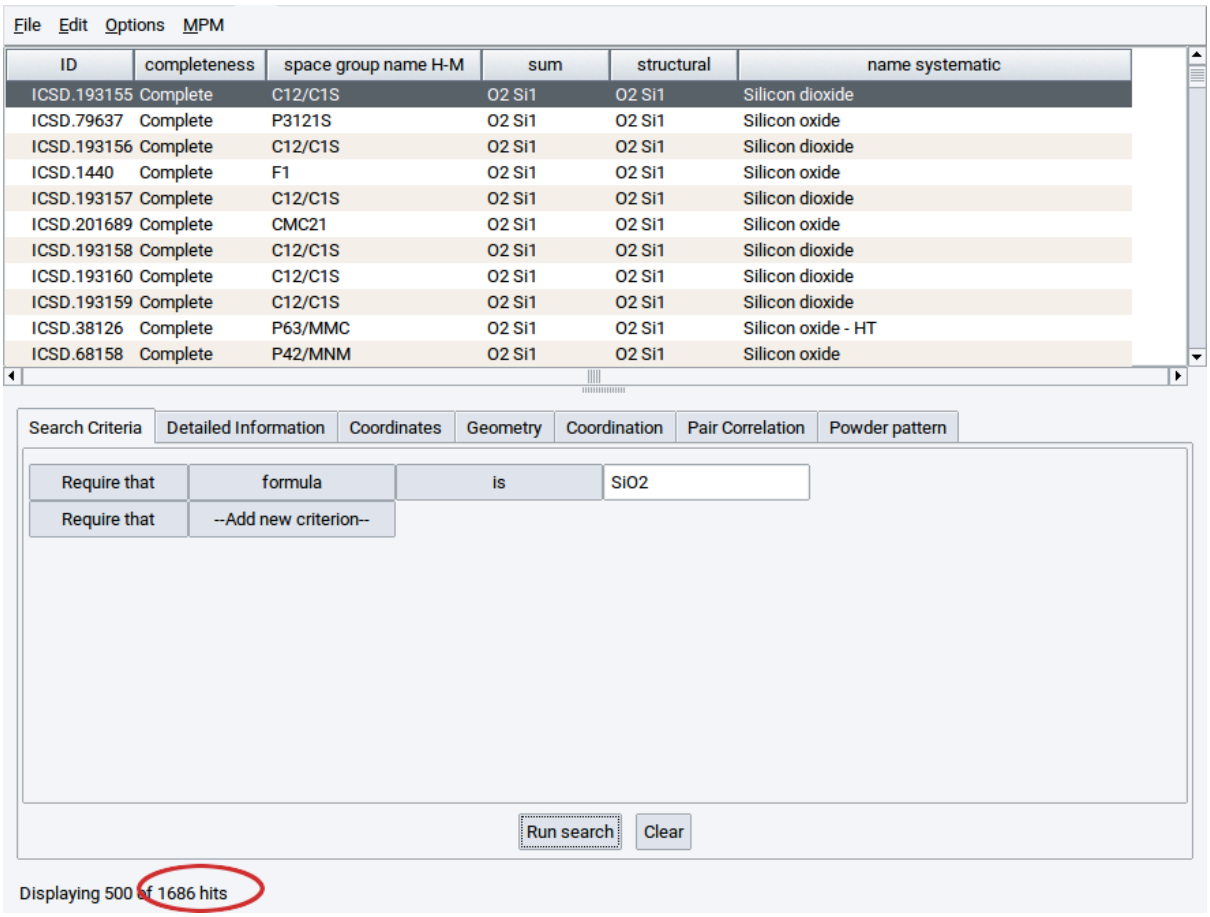


7 Assessing Structures and Overview

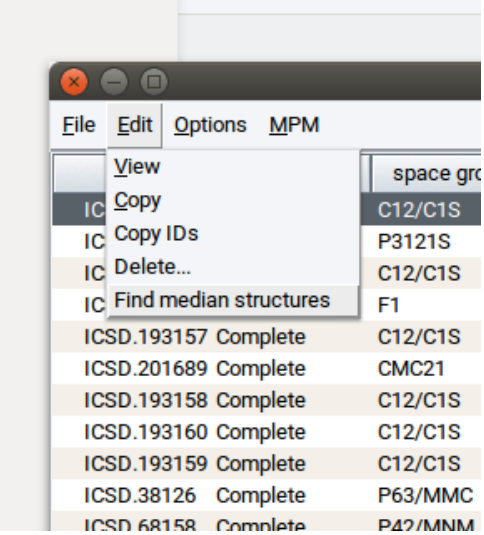
Often the number of structures matching your search criteria is rather large. For single compounds, there are multiple entries (measured at different temperatures and pressures, hence structures represent different phases), the assignment of space groups changed with higher resolution or improved sample preparation techniques in the measurements, etc. *InfoMaticA* has features to efficiently select an appropriate set of structures from the search results.

One possibility to create an essence of relevant structures is to reduce the list of found structures to a list of the median structures. The latter are representatives for each class of structures found in the search.

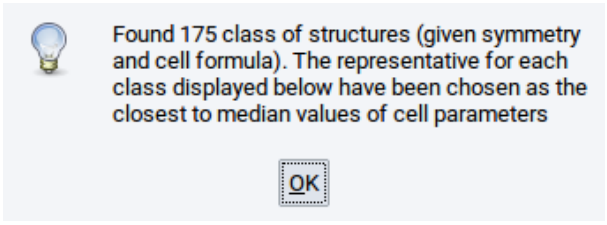
For instance: A search for structures with the search criterion formula is *SiO2* yields more than 1000 hits if all *MedeA* databases are considered.



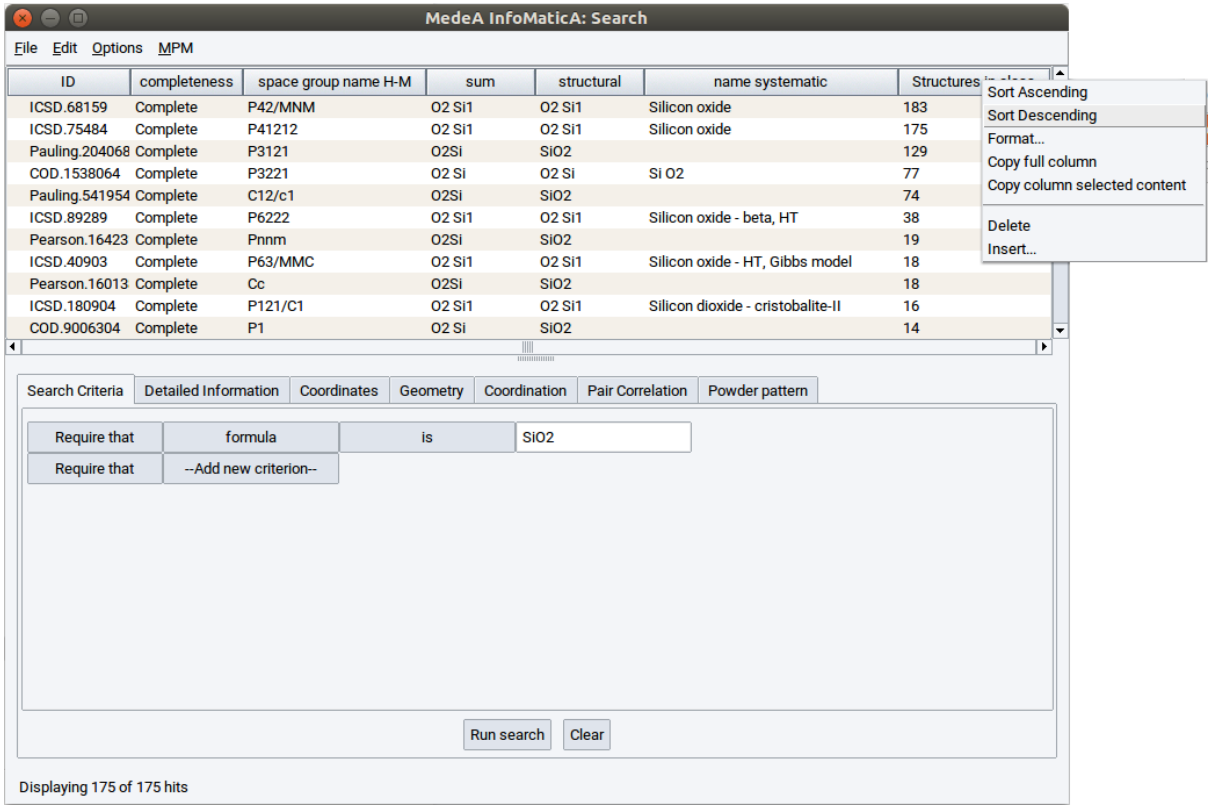
With the menu sequence Edit >> Find median structures *InfoMaticA* examines the structural properties of the complete structures that have been found (those with atomic coordinates).



InfoMaticA takes into account the space groups, lattice parameters, atomic positions, compositions, etc. In this example, *InfoMaticA* finds approximately 175 classes of structures whereby each structure found is a representative of each class.



Upon confirming with OK the 175 structures are summarized in the main table of the *InfoMaticA* dialog. The largest class of structures found in the current example is that for alpha quartz which is thermodynamically the most stable SiO_2phase.

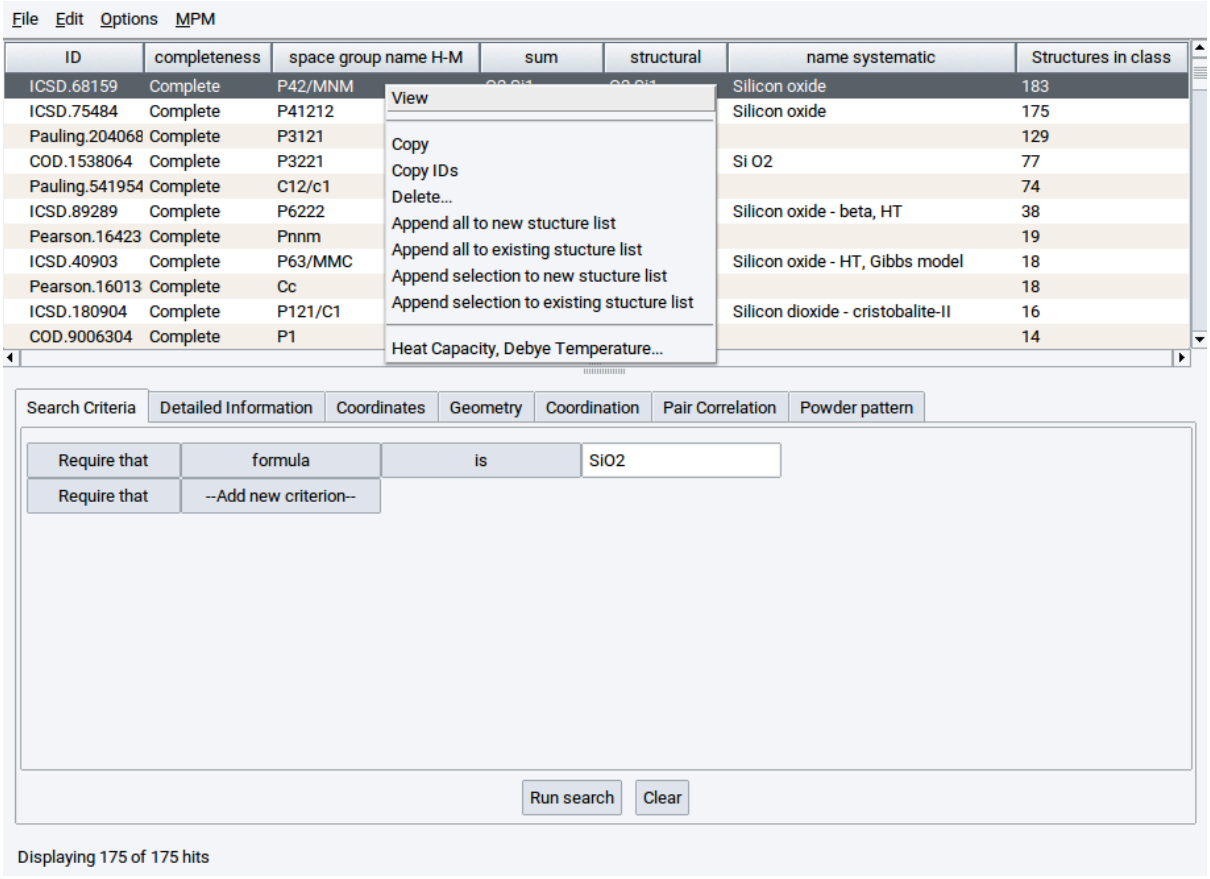


A right-click in the header cells of the table (the black row on top) opens a context menu to change what and how information is displayed:

- Sort Ascending and Sort Descending: Allow to sort data in columns. *MedeA* keeps a memory of previous sorts, so you can sort first by volume to find the most likely candidates for the low temperature structure with smallest volume, then sort by space-group: The results are grouped by space group, then within each space group by volume.
- Insert...: Add any available property to the results table such as cell lengths and volume. Some other properties might not be reported for all structures from search results.
- Delete: Remove the selected column
- Format: Change how results are displayed: numbers are displayed with all digits as reported in the database. With the format option restrict the display of numbers to a required precision, cutting off digits and filling missing numbers as 0.

8 Viewing and Editing Structures

Structures indicated by *MedeA* to be complete have all information on unit cell and atomic positions needed to display them in a 3D structure model. To view such a structure, left click on the corresponding row in the table of results to select it, then right-click >> select View from the pop-up menu (see below).



Alternatively, you can use the menu items Copy and Copy ID to copy structure information to *MedeA*'s internal clipboard for later use.

The Delete... option can be used to delete structure entries from the *Materials Design Database* but cannot be applied to the experimental databases such as e.g. *ICSD*.

Hint: As known from spreadsheet programs as *Microsoft Excel* or *Libreoffice* it is possible to select (high-

light) several table rows, i.e. structures at the same time. Multiple rows can be selected by holding the Shift key (keyboard) or the Ctrl key (keyboard) while selecting rows with the pointer. With multiple table rows selected it is possible to view, copy, etc. several structures at the same time.

With a valid license for *MedeA HT-Launchpad* it possible to add structures found in *InfoMaticA* to existing structures list and to create new structure lists. The relevant menu items are

- Append all to new structure list : Create a new structure list with all structures collected in the table
- Append all to existing structure list : Add all structures in the table to an existing structure list
- Append selection to new structure list : Create a new structure list with a selection of structures, i.e. records that are highlighted in the table - either select one structure or a multiple of structures
- Append selection to existing structure list : Add one highlighted or several selected structures in the table to an existing structure list

9 Configuration Options

Use the Option menu in *InfoMaticA*'s top menu bar to change *InfoMaticA*'s default settings:

- Options : Set the number of maximum hits to display in the results table
- Properties : Add/Remove properties to be displayed in the results table
- Databases : Select the databases to include in a search

10 Missing Databases

If you are missing a database in your *InfoMaticA* results, please check whether your *C:/MD/Databases* and *~/MD/Databases* directories include the following files: COD.db, ICSD.db, and Person.db. If not, you can download any of them from update.materialsdesign.com/db [6]. Please use your my.materialsdesign.com [7] login to access this link. Please make sure the downloaded file has an extension name *.dbz*.

Once downloaded, please save them to *MD/Databases/* and restart MedeA. Upon restarting MedeA, the database is automatically extracted and installed. You only need these database files on the computer with the MedeA GUI installed - you do not need these databases on computers that serve as JS/TS only.

After MedeA restarted, please start the InfoMaticA tool:

1. In the *MedeA GUI* click on Tools >> InfoMaticA , followed by InfoMaticA >> Search
2. In *InfoMaticA* Options >> Databases
3. Make sure all available databases are tagged with a green mark:

[6] http://update.materialsdesign.com/db/
[7] http://my.materialsdesign.com/

DOCUMENTATION