

By

Prof. Vladislav A. Blatov
Department of Chemistry
Samara State University
Samara, Russia
Email: blatov@ssu.samara.ru

and

Prof. Davide M. Proserpio
DCSSI - University of Milano
Milano, Italy
Email: davide.proserpio@unimi.it

This manual is devoted to solving the problem of topological classification of extended atomic arrays in crystal structures of any composition and complexity by means of TOPOS4.0 Pro.

General concepts

To classify atomic nets in crystals TOPOS uses the following conventional topological descriptors (see for more details: Delgado-Friedrichs O., O’Keeffe M. Crystal nets as graphs: Terminology and definitions. J. Solid State Chem. 178 (2005) 2480–2485, and the chapter 1.4 Networks, topologies and entanglements by Lucia Carlucci, Gianfranco Ciani and Davide M. Proserpio in Making Crystals by Design - from molecules to molecular materials, methods, techniques, applications. 2007 Wiley):

Coordination sequence (CS) $\{N_k\}$ is a sequence of numbers N_1, N_2, N_3, \dots counting the atoms in the 1st, 2nd, 3rd etc. coordination spheres of any given atom in the net.

Point Symbol (also Schläfli Symbol) lists the numbers and sizes of *circuits* (closed chains of connected atoms) starting from any non-equivalent atom in the net. A **Total Point Symbol for net (TS)** summarize all the Point Symbols for the non-equivalent atoms with their relation/weight. The terminology **Schläfli Symbol** should be abandoned being used in mathematics with a different definition.

Extended Point symbol (ES) lists all shortest circuits for each angle for any non-equivalent atom.

Vertex symbol (VS) gives similar information as ES, but only for rings (circuits without shortcuts).

These descriptors for all non-equivalent atoms are collected in the binary *.ttf* (TOPOS Topological Database) files. The following structure corresponds to the *.ttf* file, and is realized in the textual *.nnt* (New Net Topology) file:

An NNT and TTD entry example: TiO₂, Rutile, rtl topological type

```
'$TiO2',  
'{4;6^2}2{4^2;6^10;8^3}',  
'3 14 19 62 51 144 99 254 163 400',  
'[4.6(2).6(2)]',  
'[4.6(2).6(2)]',  
'6 10 38 34 102 74 198 130 326 202',  
'[4.4.6.6.6.6.6.6.6.6(2).6(2).8(2).8(4).8(4)]',  
'[4.4.6.6.6.6.6.6.6.6(2).6(2).*.*.]*',
```

Detailed description:

'\$TiO2',
Name of the record with the '\$' prefix
'{4;6^2}2{4^2;6^10;8^3}',
TS, Total point symbol for the whole net: $\{4.6^2\}_2\{4^2.6^{10}.8^3\}$. In this case the numbers of the two non-equivalent nodes (O and Ti) relate as 2:1
'3 14 19 62 51 144 99 254 163 400',
CS, Coordination sequence $\{N_k\}$ for the first non-equivalent node (oxygen atom); $k=1-10$
'[4.6(2).6(2)]',
ES, Extended point symbol: $[4.6_2.6_2]$ for the first non-equivalent node
'[4.6(2).6(2)]',
VS, Vertex symbol for the first non-equivalent node (here coincides with the Extended point symbol)
'6 10 38 34 102 74 198 130 326 202',
'[4.4.6.6.6.6.6.6.6.6(2).6(2).8(2).8(4).8(4)]',
'[4.4.6.6.6.6.6.6.6.6(2).6(2).*.*.*)',
Similar triples for other non-equivalent nodes (Ti atom) * means that there are no rings for this angle: $[4.4.6.6.6.6.6.6.6.6_2.6_2. *. *. *]$ (alternatively to * the ∞ symbol could be used in the Vertex symbol notation)

Normally TOPOS assumes that CS, ES, and VS descriptors determine the net topology up to isomorphism. Thus, if several atoms in asymmetric unit have the same CS, ES, and VS they are assumed to be topologically equivalent (a topological supersymmetry exists), and a unique triple set CS,ES,VS will correspond to them in .ttd or .nnt files. Also, two nets are assumed to possess the same topology if an isomorphic relation can be established for their sets of CS, ES, and VS. However, in some very rare cases topologically different nets may have the same CS, ES, and VS sets. Such nets can be distinguished extending the Vertex symbol including *all-ring*, where *all* (not only *shortest*) rings are collected. The *all-ring* VS is computed if requested (see below) and reported only in TOPOS output.

An example of TOPOS output with all-ring Vertex symbols for rutile

Vertex symbols for selected sublattice

O1 Point (Schlafli) symbol: $\{4;6^2\}$

Extended point symbol: $[4.6(2).6(2)]$

Rings coincide with circuits

All rings (up to 10): $[(4,6(2)).(6(2),8(6)).(6(2),8(6))]$

Ti1 Point (Schlafli) symbol: $\{4^2;6^{10};8^3\}$

Extended point symbol: $[4.4.6.6.6.6.6.6.6.6(2).6(2).8(2).8(4).8(4)]$

Vertex symbol: $[4.4.6.6.6.6.6.6.6.6(2).6(2).*.*.*)$

All rings (up to 10):

$[4.4.(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).6(2).6(2).*.*.*)$

ATTENTION! Some rings * are bigger than 10, so likely no rings are contained in that angle

Point (Schlafli) symbol for net: $\{4;6^2\}_2\{4^2;6^{10};8^3\}$

In this case all rings were computed up to 10-ring (a parameter selected by the user, see below). So possibly larger rings may exist - TOPOS does not know this!

All rings (up to 10): $[(4,6(2)).(6(2),8(6)).(6(2),8(6))]$

This means that not only 4- (or 6-) rings, but also larger 8-rings meet at the same angle of the first non-equivalent node (oxygen atom, cf. ES or VS). The all ring VS could be reported as: $[(4,6_2).(6_2,8_6).(6_2,8_6)]$ and $[4.4.(6,8_3).(6,8_3).(6,8_3).(6,8_3).(6,8_3).(6,8_3).(6,8_3).(6,8_3).6_2.6_2. *. *. *]$

At last even a more detailed description of rings is possible. Rings of the same size are not always (symmetry) equivalent, TOPOS can distinguish them grouping them by equivalent kind. The groups are designated by one or more letters: a-z, aa-az, ba-bz, etc, e.g. 4a, 12ab, 20xaz. As a result a *all-ring labelled* Vertex symbol is calculated.

The *all-ring labelled* VS is computed if requested (see below) and reported only in TOPOS output.

An example of TOPOS output with *all-ring labelled* Vertex symbols for rutile

Vertex symbols for selected sublattice

```
-----
O1 Point (Schlafli) symbol:{4;6^2}
Extended point symbol:[4.6(2).6(2)]
Rings coincide with circuits
All rings (up to 10): [(4,6(2)).(6(2),8(6)).(6(2),8(6))]
All rings with types: [(4,6(2)).(6(2),8a(4),8b(2)).(6(2),8a(4),8b(2))]
-----
Ti1 Point (Schlafli) symbol:{4^2;6^10;8^3}
Extended point symbol:[4.4.6.6.6.6.6.6.6(2).6(2).8(2).8(4).8(4)]
Vertex symbol: [4.4.6.6.6.6.6.6.6(2).6(2).*.*.*]
All rings (up to 10):
[4.4.(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).(6,8(3)).6(2).6(2).*.*.*]
All rings with types:
[4.4.(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).(6,8a(2),8b).6(2).6(2).*.*.*]
ATTENTION! Some rings * are bigger than 10, so likely no rings are contained in that angle
-----
Point (Schlafli) symbol for net: {4;6^2}2{4^2;6^10;8^3}
```

For example, the first angle for the first node (oxygen atom) contains two non-equivalent 8-rings. The all ring labelled VS could be reported as: [(4,6₂).(6₂,8a₄,8b₂).(6₂,8a₄,8b₂)]

We believe that using CS+ES+VS+ when needed *all-ring* and/or *all-ring labelled* VS combination one will be able to distinguish any topologically non-equivalent nets (no counterexamples are known).

Beside the classification of a net via the TTD collection, starting from December 2007 a new collection is being created – TTO (Topological Types Observed). **In short, TTO collection matches topological types of abstract nets with examples of real crystal structures.**

TTO files contain the records of the following format:

RefCode – the Reference Code of a particular structure;
Dimen – dimensionality of the underlying net;
Z – number of nets;
TopType – topology of the net (its code in TTD collection)
ReprType – code of representation type of the structure.

The correspondences between *ReprType* codes and the description of representations are given in the textual file *TTOReprTypes.txt*.

Having TTO collection the user may now perform the following operations:

1. Find all topologies for all representations of a particular crystal structure.
2. Find all structures with a particular topology of underlying net.
3. Find all structures with a particular topology in a given database.

All *.ttd* and *.tto* files have to be copied into separate subdirectories (/TTD and /TTO by default) within the TOPOS folder. *TTOReprTypes.txt* file has to be copied into /TTO directory. In alternative, the paths to TTD and TTO collections can be changed in *System/TOPOS Parameters/Paths*.

Interface with *Systre*

Another way to classify nets is to use *Gavrog Systre* program (<http://gavrog.org/>). TOPOS has an easy interface with *Gavrog Systre* format. To export TOPOS database entries into a *Gavrog Systre* input file (*.cgd) choose *Database/Export* and *Systre (*.cgd)* as file type. To import *Gavrog Systre* output file(s) into a TOPOS database choose *Database/Import* and *Systre files (*.out, *.cgd, *.log)* as file type.

Sources of the TTD collection

The TTD collection has been built using the following sources:

- RCSR (Reticular Chemistry Structure Resource), the original data are available for free at <http://rcsr.anu.edu.au/>.
- EPINET (Euclidean Patterns in Non-Euclidean Tilings), the original data are available for free at <http://epinet.anu.edu.au/>.
- Data on the sphere packings derived by Fischer, Koch & Sowa; the data were published in a number of papers (see, e.g. Sowa, H. & Koch, E. (2005). *Acta Cryst.* **A61**, 331-342).
- Crystallographic data from CSD (Cambridge Structural Database) and ICSD (Inorganic Crystal Structure Database).
- Atlas of Zeolite Framework Types, the original data are available for free at <http://www.iza-structure.org/databases/>.
- Data on supernet-subnet relations (see Blatov, V. A. (2007). *Acta Cryst.* **A63**, 329–343).

Nomenclatures for topologies

In the TTD collection the net topologies are designated according to the following nomenclatures:

- RCSR lower-case three-letter symbols, see <http://rcsr.anu.edu.au/> for details.
- EPINET *sqcXXXXX* symbols, see <http://epinet.anu.edu.au/> for details.
- Fischer's symbols *k/m/fn* for sphere packings (see Koch, E., Fischer, W. & Sowa, H. (2006). *Acta Cryst.* **A62**, 152-167).
- CSD Reference Codes or ICSD Collection Codes.
- Zeolite capital three-letter symbols, see <http://www.iza-structure.org/databases/> for details.
- Subnet *s-d-G-n* symbols (Blatov, V. A. (2007). *Acta Cryst.* **A63**, 329–343), where *s* is a conventional name of the initial net, *d* is a set of ascending integers equal to degrees of all inequivalent nodes in the subnet, *G* is the space group for the most symmetrical embedding of the subnet, *n* is optional and enumerates non-isomorphic subnets with a given *s-d-G* sequence.
- Subnet transformation symbols *s/G→S1→...→Sn;BS* where *s* is a conventional name of the initial net, *G* is the space group of the initial net, *S1, ..., Sn* is the sequence of group-subgroup transformations to obtain the symmetry of the resulting subnet, *Sn, BS* is the set of numbers of non-equivalent edges to be retained in the subnet. For instance, the notation $f_{ny/P} 63/m c m \rightarrow P 63 2 2$

$(0,0,1/4)$; Bond sets: 2,3,4,5 means that the subnet is derived from the RCSR net fny by decreasing its space-group symmetry from $P6_3/mcm$ to $P6_322$ with shifting the origin by $(0,0,1/4)$ vector and breaking all non-equivalent edges in the resulting net except the edges No 2, 3, 4 and 5. The subnet transformation symbols are being gradually replaced in the TTD collection with s-d-G-n symbols; **if you have obtained in TOPOS output a subnet transformation symbol, this means that the topology has not yet been found in crystal structures. Let the authors know about this case!**

Some nets fall into different nomenclatures, in this case TOPOS outputs all possible symbols separated by semicolons.

What to do with new topologies?

Exploration of the crystal structure topology becomes more and more popular and researchers often encounter topologies not described in the TTD collection. How to designate them, how to refer to them and how to be sure that your colleagues will be advised about them? Here we announce a strategy to resolve these problems.

1. If you have found the message

New topology, please, contact the authors

in the ADS output please email a TOPOS database (*.adm, *.cd and *.cmp files) with this structure to one of the authors of this manual.

2. After checking the data we assign a code **abcX** to the topology, where **abc** is your three-letter personal identifier and X is the ordinal number of the topology you have discovered (e.g. **smi12** that means '12th topology discovered by Smith'). If you have no identifier we will consult with you about its name.

3. The list of authors' personal identifiers together with contact information will be available at the TOPOS website.

4. The novel topology will be included into the TTD collection and published in a two-week update of the *personal.ttd* database. Anyone who will download the database will be aware of the new topology.

5. If the topology will further be included into other databases it will get other names according to the nomenclatures described above, but the initial **abcX** name will be retained as a prove of priority.

Practical Guide

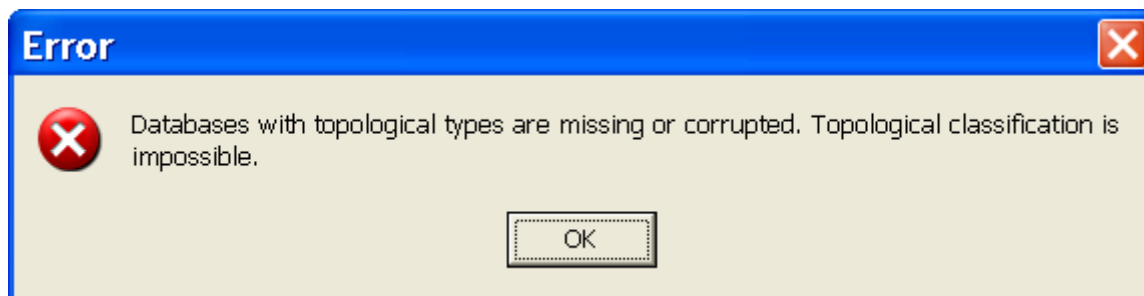
This part includes the strict algorithm how to classify the topology of a crystal structure in TOPOS.

1. Normally, the topology of the net should be simplified before the classification. *E.g.* in metal-organic frameworks (MOFs) the net nodes are usually metal centers (atoms or clusters) and centroids of organic ligands; in H bonded arrays of molecules the simplified net consists of connected molecular centroids, *etc.* To simplify the net and to save the new topology you need to check the *Save Centroid ADS* option (*Common* tab), see also the *Interpenetration with TOPOS4.0 Professional* manual (the part **Studying topological properties of the structure**). After running ADS choose central atoms (*e.g.* metal atoms for MOFs) or press the *No Atoms* button if the simplified net should describe a molecular packing. As a result a database *<old name>_c* will be created containing a record with the simplified net. The nodes for complex structural groups will be assigned to virtual atoms (starting from 'Sc'), and the legend of the assignment will be stored in *Comment* tab of the *Crystal Data* for the simplified net. The next steps should be made with the simplified net. You may also create no simplified net, but perform the simplification during the classification procedure. In this case go to the next step without checking the *Save Centroid ADS* option.

2. Be sure that there is at least one *.ttd* file in TOPOS directory. You may find a frequently updating collection of *.ttd* files (TTD collection) on the TOPOS official site. At the beginning of the topological analysis TOPOS automatically loads all the *.ttd* files, so you may easily expand your library of topological types by copying new *.ttd* files into TOPOS directory. The information about loaded databases is output after performing the analysis, for example:

(23793 types in 8 databases)

If there are no *.ttd* files the error message will be output:



IMPORTANT: if you have copied a new *.ttd* file you must rerun TOPOS to accept this file.

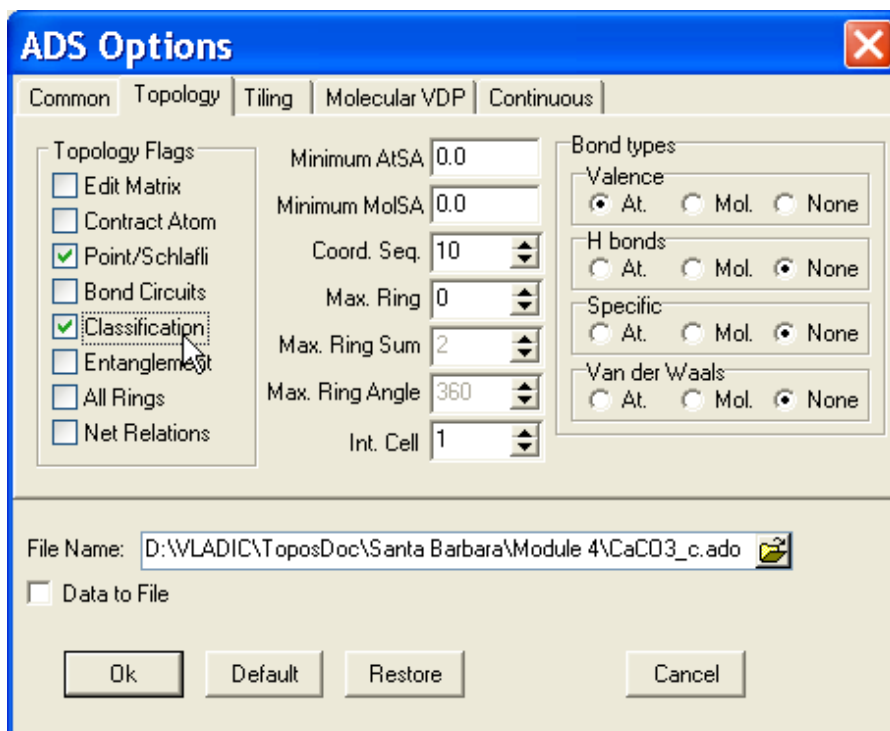
3. Open ADS program and check the flag *Classification* on the *Topology* tab of the *ADS Options* window. The following parameters will also be specified:

Dimen. Calc. (*Common* tab) – to recognize all non-bonded structural components (*e.g.* interpenetrating 3D nets). This flag may be unchecked to speed the analysis if you are sure that the crystal structure consists of the single net.

Point/Schlafl (*Topology* tab) – to calculate extended Schläfli symbols.

Coord. Seq. (*Topology* tab) – to calculate coordination sequences (default number of coordination spheres is 10).

Other flags will be unchecked.



You may not uncheck the *Point/Schlafl*i flag or specify *Coord. Seq.* less than 10 until the *Classification* flag is unchecked. As a result a minimum combination CS+ES will be used to classify nets.

4. If you want to make the analysis more precise you may extend the combination of topological parameters with rings (VS); for this purpose specify *Max. Ring* >0 to determine maximum size of rings to be analyzed. If *Max. Ring* >0 you may check the *Data on New Nets* flag (*Common* tab) to save the topological parameters on the nets with new topologies in a <database name>.nnt file. The .nnt file can be sent to the authors of this manual to include the new topologies into a new edition of the TTD collection.

5. To calculate *all-ring* Vertex symbols you need to check the *All Rings* flag, and to get *all-ring labelled* Vertex symbols specify *Max. Ring Sum* >2. In most cases to compute all rings (with or without label) are not required to recognize nets, but can highly slow the calculation. Do not use these options unless necessary.

6. Close the *Options* window and run ADS. You should follow the recommendations of the *Interpenetration with TOPOS4.0 Professional* manual (the part **Test on interpenetration**) to analyze both a single crystal structure and a set of compounds. As a result of the analysis TOPOS outputs a message with the information on the topological type:

```
Topological type: dia Diamond; 4/6/c1; sqc6 {6^6} - VS [6(2).6(2).6(2).6(2).6(2).6(2)] (23793 types in 8 databases)
```

if the TTD collection contains the topology, or

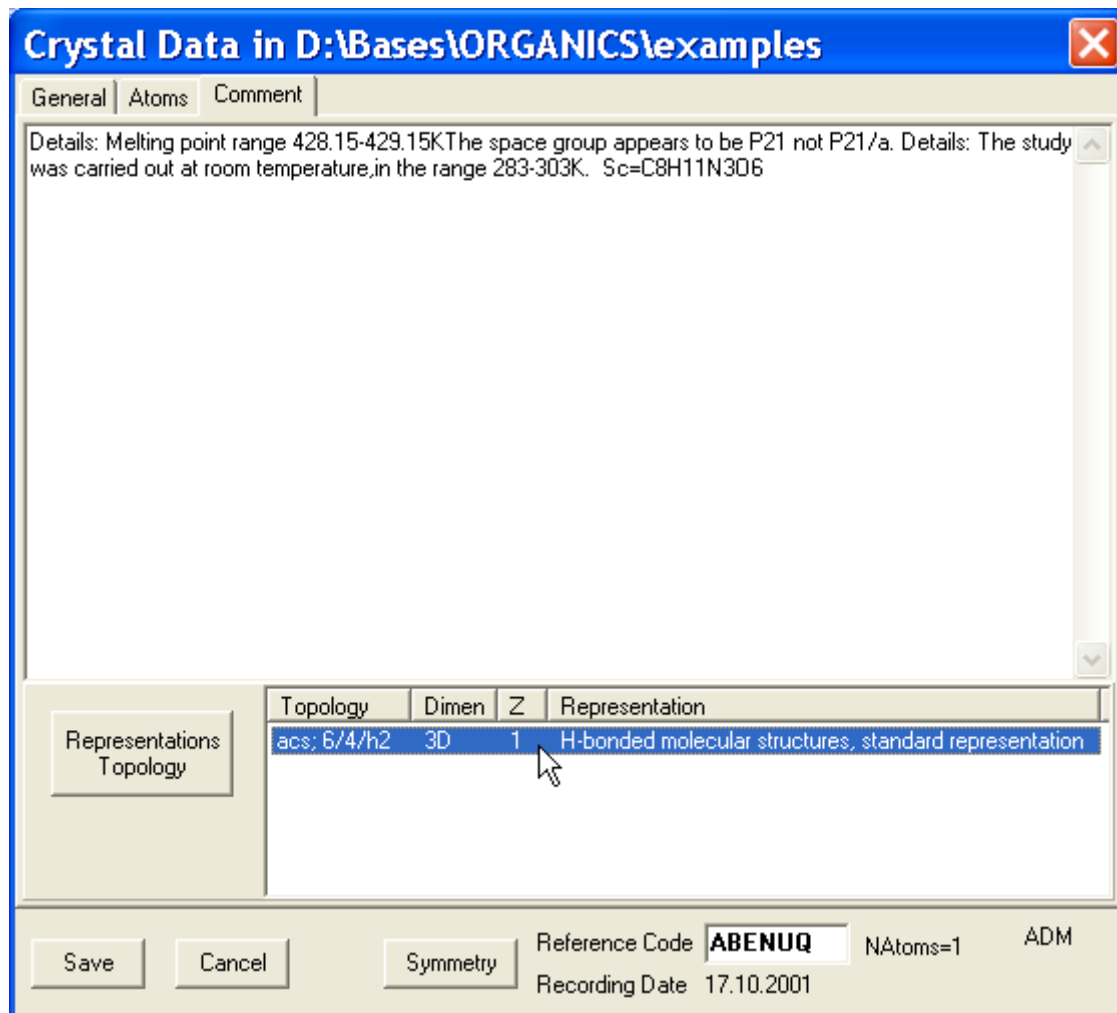
New topology, please, contact the authors

if the topology is unknown. For details on the topological types you may ask the authors of this manual, or the authors of original databases from the TTD collection (see TOPOS site).

Examples: see demo database and exercises in the *Interpenetration with TOPOS4.0 Professional* manual.

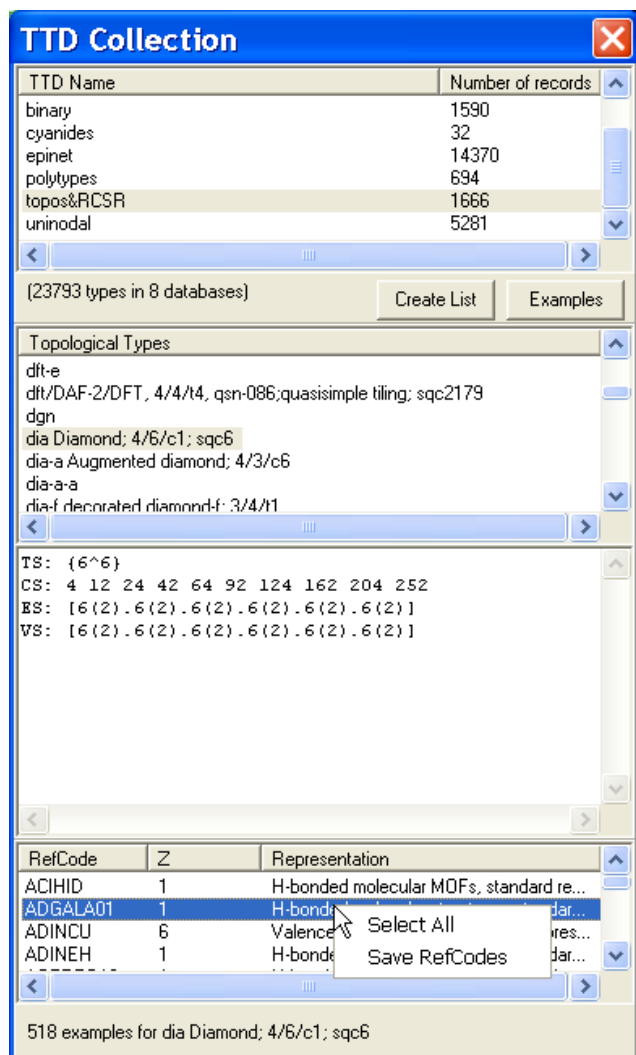
To work with TTD and TTO collections the following operations are intended.

1. Find all topologies for all representations of a particular crystal structure. For this purpose open *Crystal Data* window for the structure, go to *Comment* tab and push *Representations Topology* button. In the example below, for the ABENUQ crystal structure there is found one representation, in which H-bonded molecules are considered as whole; the topology of the resulting 3D single (Z=1) net is **acs** (6/4/h2 sphere packing).

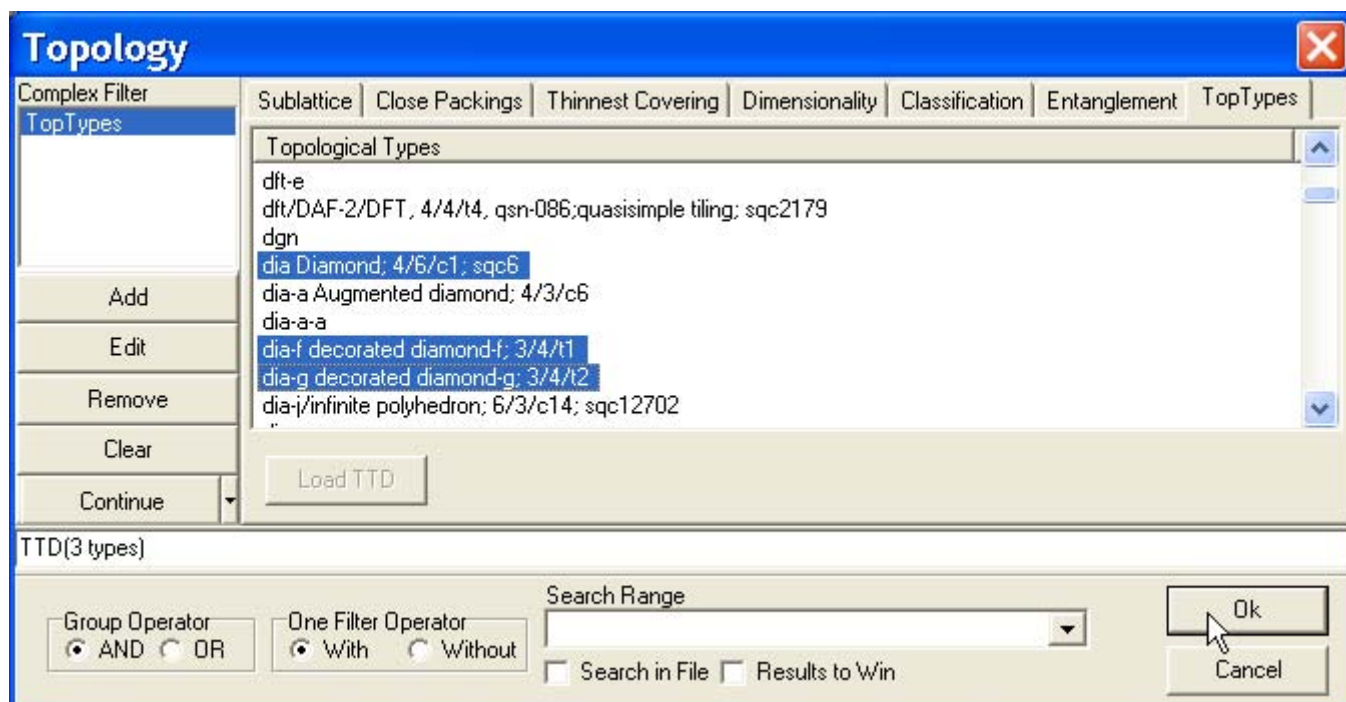


2. Find all structures with a particular topology of underlying net. Go to *Database/TTD Collection*, load TTD collection and perform *Database/TTD Collection/Content* command. In the upper list of *TTD Collection* window you have all *.ttf* files loaded. Select one or several *.ttf* names and press *Create List* button; you will get all the topological types of this *.ttf* in the next list. Select any topological type; you will have all the topological indices in the next window. You may easily find the type by name, typing the name on the keyboard (use *BackSpace* to go backward the name). And, finally, pressing *Examples* button you will get all structures that have this topology in the bottom list. you may output Reference Codes from the list of examples by selecting them, right-clicking and choosing *Save RefCodes*. The list of Reference Codes will be stored in the TTO directory as a textual file *<top.type name>.gcd*. Then you may use *TOPOS Acquisition Data/Ref. Code* filter (or *CSD ConQuest*) to find the records.

In the example below a TTD collection of 8 databases with 23793 topological types is loaded. The list of topological types composing *topos&RCSR* database is built, the topological indices for **dia** topological type are output, and 518 examples of underlying nets in real crystal structures are found in TTO collection.



3. Find all structures with a particular topology in a given database. Open *Filter/Topology/TopTypes* tab and press Load TTD button. Then select one or several topological types in the list and run the filter.



Remember that TOPOS can find only those correspondences that are collected in the TTD collection.

4. Browse TTO collection. Go to *Database/TTO Collection*, load TTO collection, select TTO databases and press *Create List* button.

The screenshot shows a window titled "TTO Collection" with a close button in the top right corner. It contains two tables. The first table is a summary table with three columns: "TTO Name", "Number of records", and "Representation Type". The second table is a detailed list of records with five columns: "RefCode", "Topological Type", "Dimen", "Z", and "Representation". Below the summary table, it indicates "4606 records" and a "Create List" button. The detailed list table has a scrollbar on the right side.

TTO Name	Number of records	Representation Type
MOF_HB_stand	757	H-bonded molecular MOFs, standard representation
MOF_VB_cluster_stand	146	Valence-bonded MOFs, standard cluster representation
MOF_VB_stand	1985	Valence-bonded MOFs, standard representation
org_HB_stand	1718	H-bonded molecular structures, standard representation

4606 records Create List

RefCode	Topological Type	Dimen	Z	Representation
ABALUL	dia Diamond; 4/...	3	1	Valence-bonded MOFs, standard cluste...
ABAVIJ	rtl rutile 3,6-conn	3	1	Valence-bonded MOFs, standard repres...
ABAVOP	rtl rutile 3,6-conn	3	1	Valence-bonded MOFs, standard repres...
ABIWOX	crb/BCT; 4/4/t5...	3	1	Valence-bonded MOFs, standard repres...
ACAJIY	pcu alpha-Po pri...	3	1	Valence-bonded MOFs, standard repres...
ACAZEK	sqc502-3,5-P43...	3	1	Valence-bonded MOFs, standard repres...
ACDMAL	etb; 3/8/h1; sqc...	3	1	Valence-bonded MOFs, standard repres...
ACENIF	dia Diamond; 4/...	3	1	Valence-bonded MOFs, standard cluste...
ACMPCD	zfd-4-1422	3	1	Valence-bonded MOFs, standard repres...
ACOGAB	roa-4-Fddd	3	1	Valence-bonded MOFs, standard repres...

You may order the list with any column by clicking the title of the column and then find required record by typing the value of the parameter corresponding to the column. Right click in the list provides the operations like in the *Examples* list of TTD collection.