Searching and analysing metal-organic structures

CCDC Virtual Workshop Summer 2021 – Session 3

Ilaria Gimondi, Suzanna Ward, Yinka Olatunji-Ojo, Seth Wiggin, Eva Myers

July 2021

advancing structural science

Learning outcomes

Advanced tips and tricks to search and analyse metal-organic structures by:

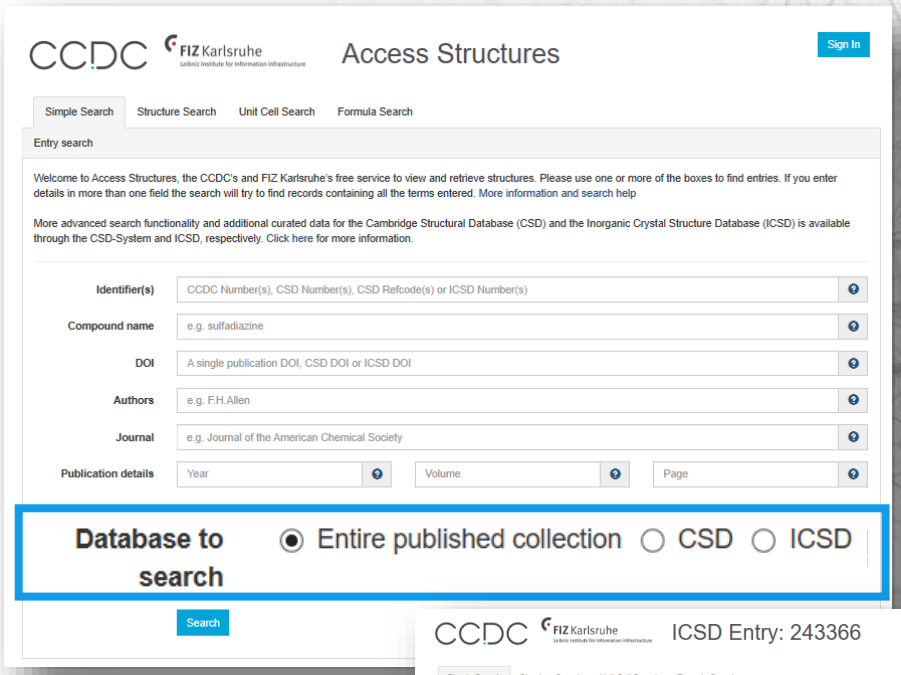
- Learning tactics to search metal-organic structures in ConQuest.
- Familiarity of metal-organic subsets in ConQuest.
- Learning how to export and analyse metal-organic hitlists in Mercury.

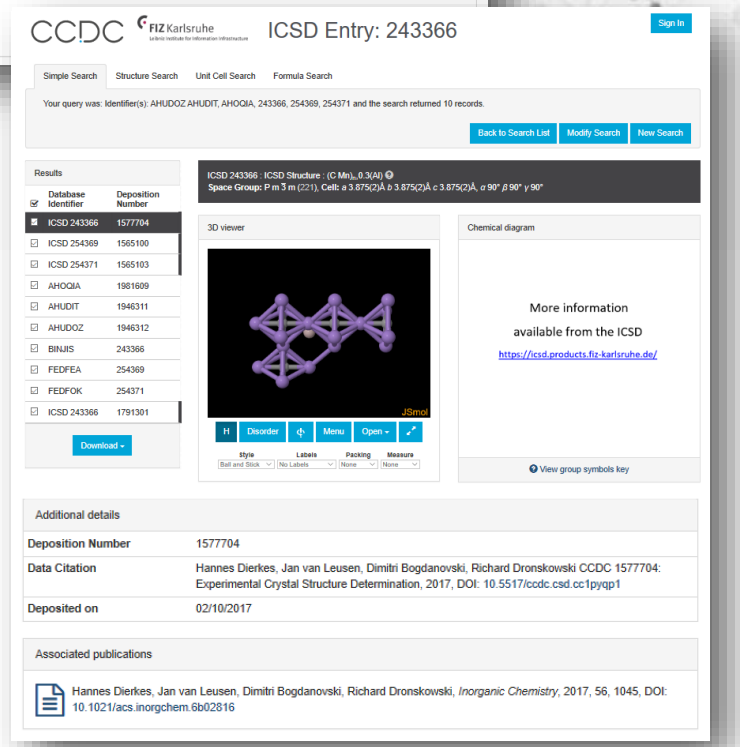
CCDC

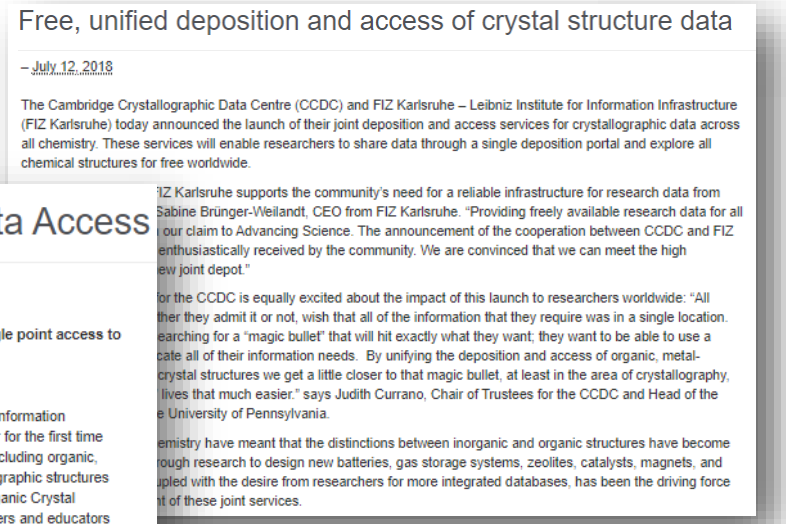
The CSD and the ICSD

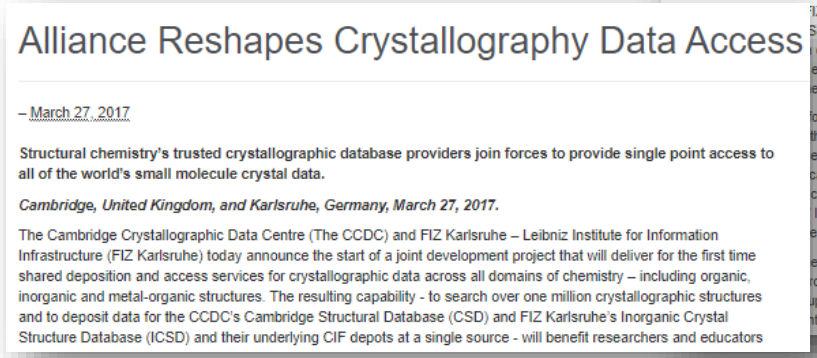
Connecting data resources, increase discoverability and simplify deposition of organic, metal-organic and inorganic structural data.

- Established, trusted databases
- Value quality and high levels of curation
- Comprehensive
- Shared communities









CSD criteria

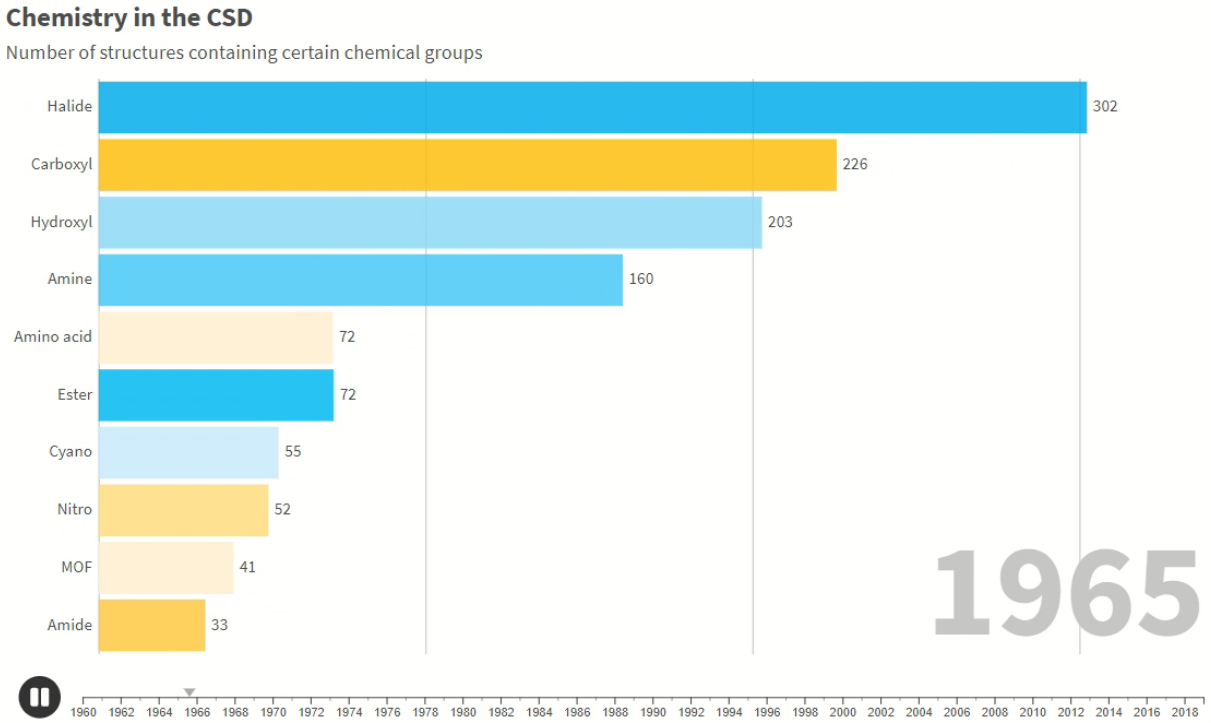
Organic and metal-organic experimental crystal structures

- Types of experimental determination include:
  - a single crystal study - where cell parameters are reported, or
  - a powder study, where cell parameters, atomic coordinates and constrained refinement (e.g. Rietveld) are reported
  - from x-ray, neutron or electron diffraction

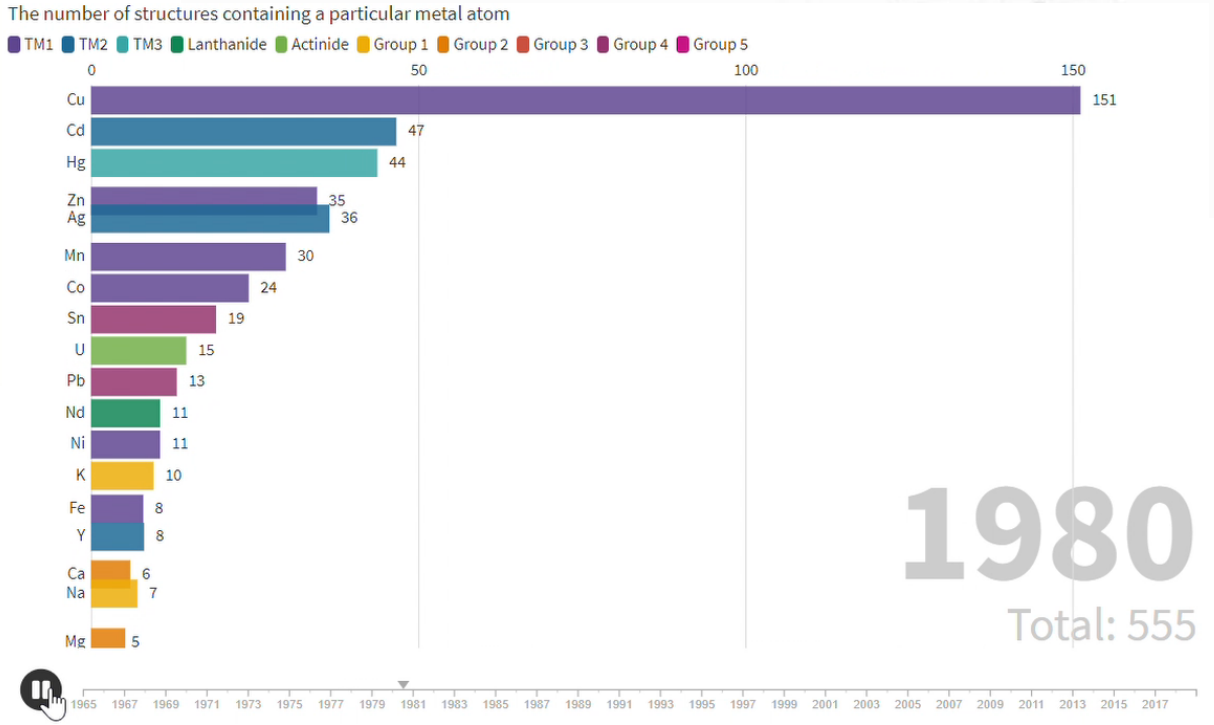
- Types of chemistry include:
  - Metal and metal-organic compounds
  - peptides and saccharides of up to 24 residues
  - mono-, di- and tri-nucleotides
  - metal carbonyls
  - boron compounds containing one or more B-H or B-OH bond and borazines
  - ring compounds containing any two of the following elements: N, P, S, Se and Te

CCDC

Exploring the CSD



Metals in the CSD MOF subset

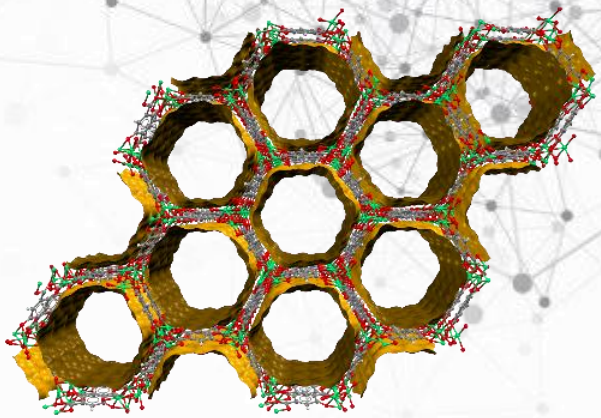


Images and graphics created using Flourish

What is a MOF?

"A metal–organic framework, abbreviated to MOF, is a coordination network with organic ligands containing potential voids."

IUPAC Recommendations 2013[1]

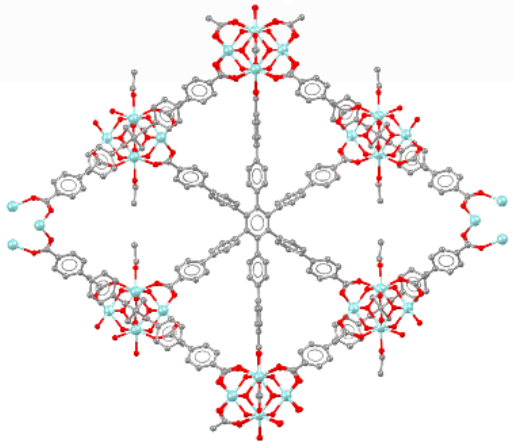


Interim 2012 report[2] noted:

- Questionnaire gave no clear consensus if MOFs should include 1D, 2D or 3D networks
- Almost 25% of respondents believed MOFs must have proven porosity
- No IUPAC definition of organic – include oxalates, cyanides etc.?

"*Finally, we should perhaps remind ourselves that what matters, or should matter, in the real world are the properties of our materials not what we chose to call them. As Juliet says: 'A rose by any other name would smell as sweet'*"

IZA Commission on Metal Organic Frameworks[3]

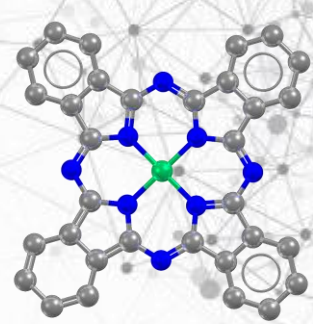


1. S. R. Batten *et al. Pure Appl. Chem.*, 2013, 85, 1715–1724 10.1351/PAC-REC-12-11-20; 2. S.R.Batten *et al.CrystEngComm*, 2012, 14, 3001-3004 10.1039/C2CE06488J; 3. http://www.iza-online.org/MOF/

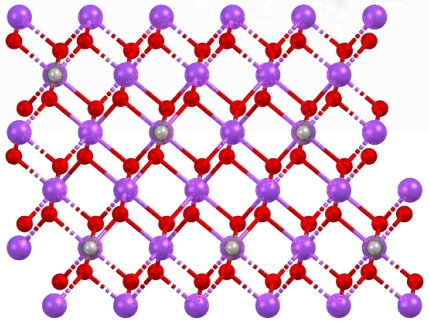
CCDC

MOF and metal-organic trivia

- 1st metal-organic structure with coords was published in 1938 by Monteath Robertson
- 1st MOF in our MOF subset was published in 1940
- >40,000 authors have contributed to CSD MOF structures
- Most prolific author of MOF structures is Pascal D.C.Dietzel
  - Time-resolved synchrotron X-ray diffraction measurements leads to many individual data collections, currently over 1,900!



NIPHTC
1st metal-organic structure with 3D coords

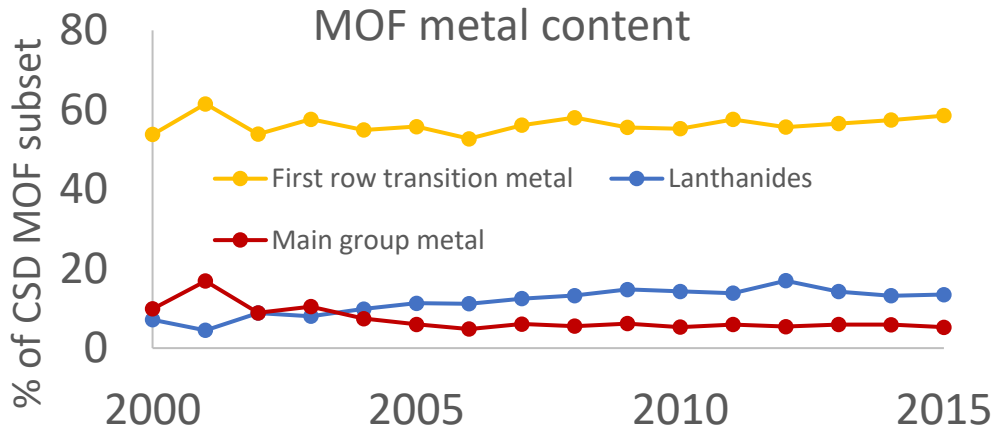


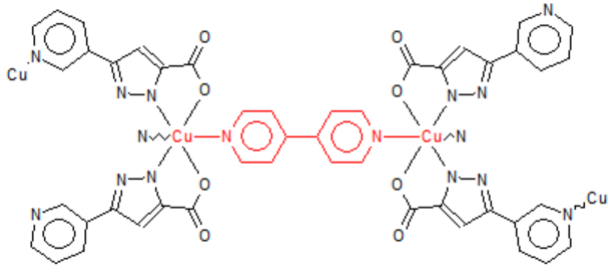
NAFORM
1st MOF in CSD subset

CCDC

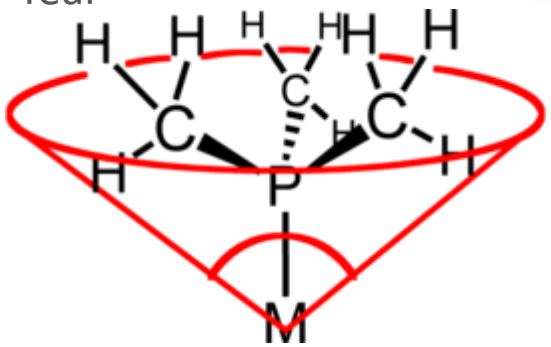
Using the CSD to explore metal-coordinations

- Ligand coordination modes
- Metal coordination numbers
- Bond lengths
- Symmetry and shape
- Spin states
- Ligand cone angles
- Polynuclear complexes
- Magnetism
- Software parametisation



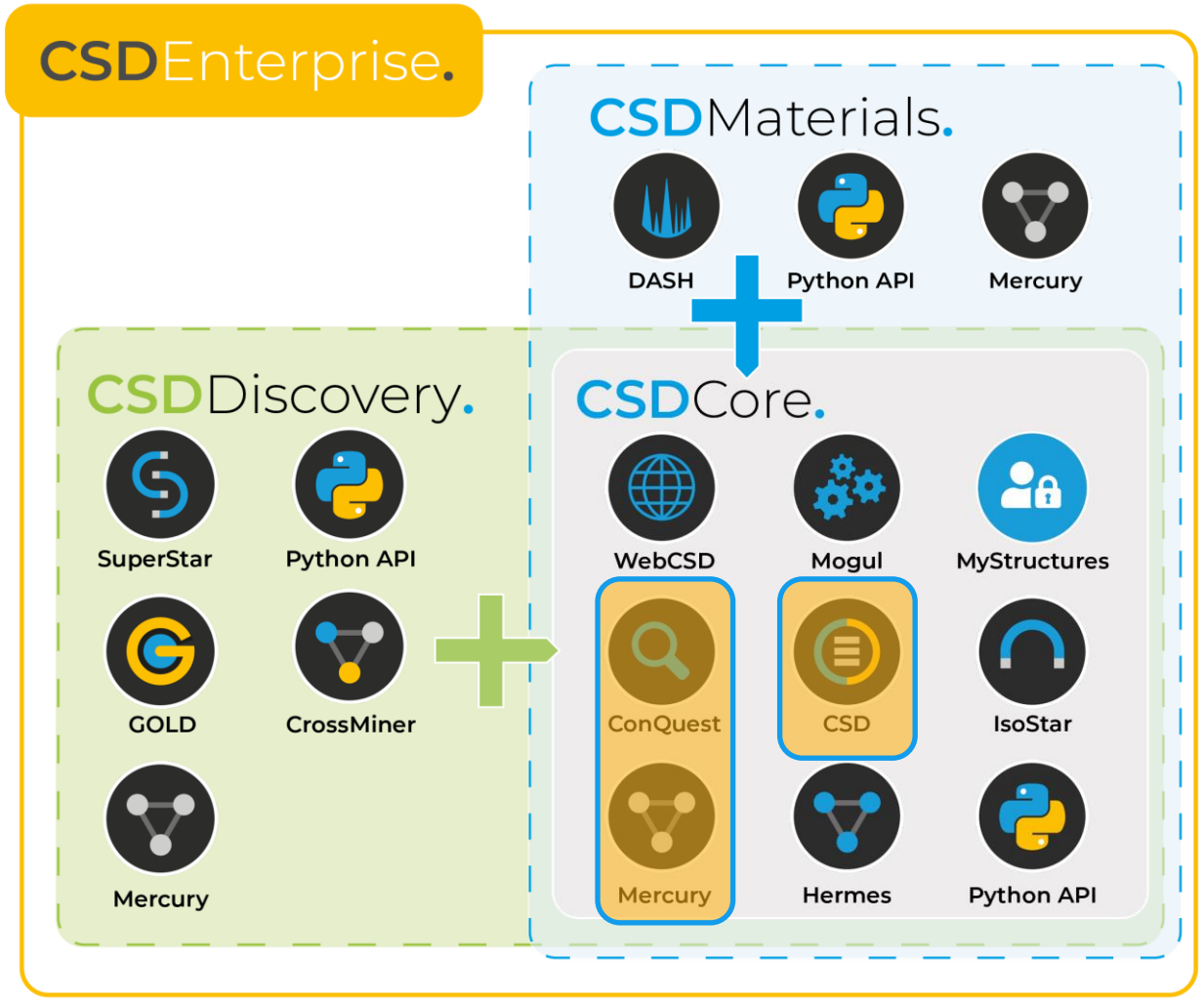


2,2-bipyridine one of the most frequent ligands in the CSD with >6,000 entries
~10% mono-coordinated, 90% di-coordinated



CCDC

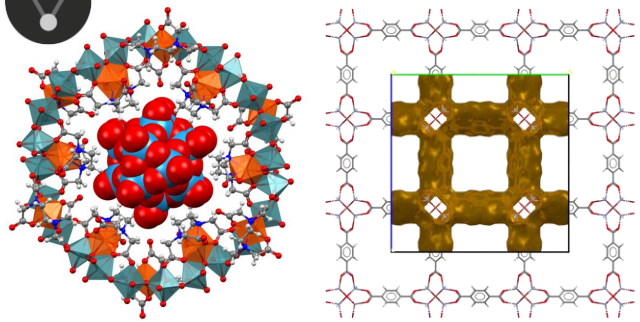
The CSD Portfolio



CCDC

Search, visualisation and analysis

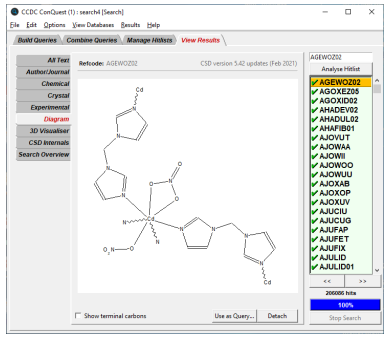
Visualisation - Mercury



Visualise data in the CSD or your own structures

- Explore how structures pack together including
  - 1D, 2D and 3D networks, inter-molecular and intramolecular bonds
- Visualise important features including
  - Voids, polyhedral display

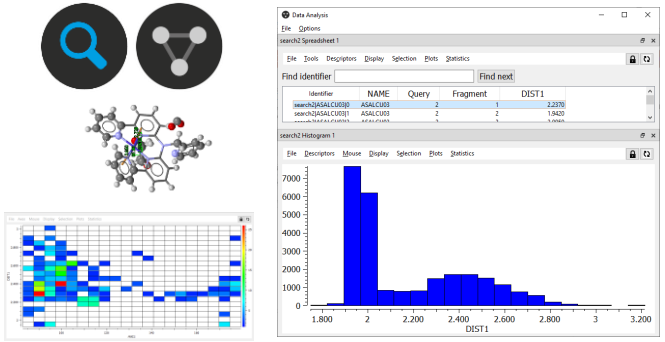
Search - ConQuest



Search and retrieval of data from the CSD

- Full range of text / numeric database search options
- More complex search functionality
  - Chemical structure searching
  - 3D Geometrical searching
  - Intermolecular non-bonded contact searching

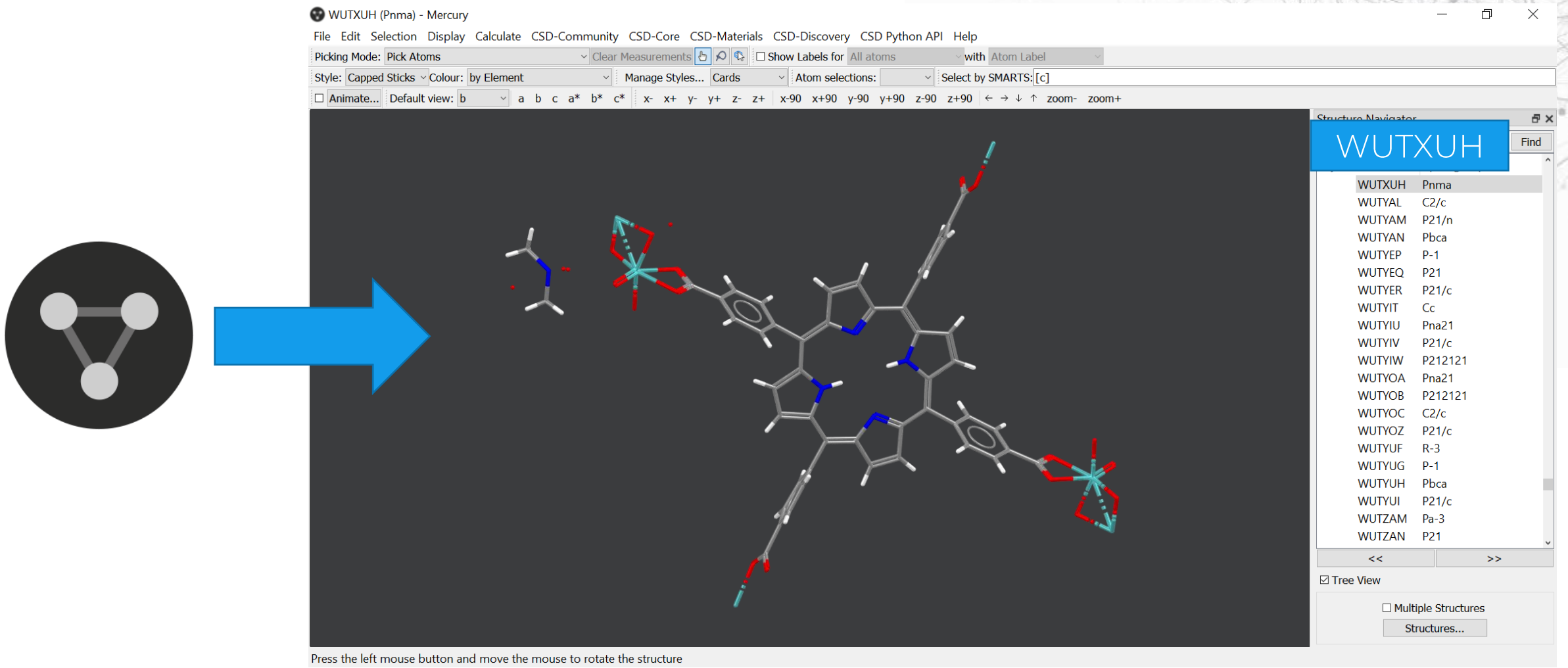
Analysis – CQ + Mercury



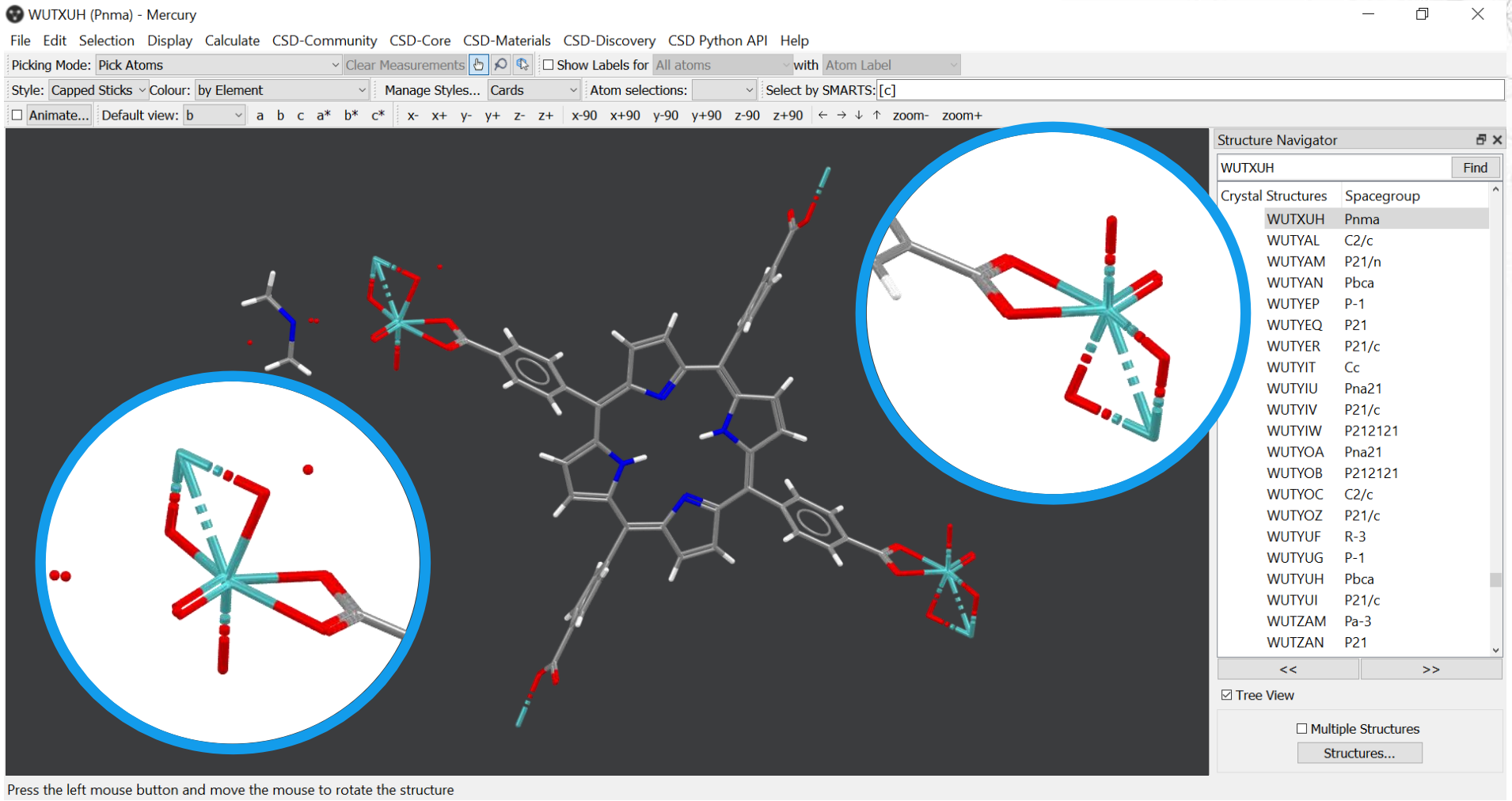
Analyse data in the CSD and compare your own structures

- Learn about metal-coordinations
  - Ligand coordination modes, metal coordination numbers, bond lengths, symmetry and shape, spin states, ligand cone angles, porousity, 1D, 2D and 3D networks

Show One: Visualisation tips and tricks for metal-organic structures



Polymeric bonds representations

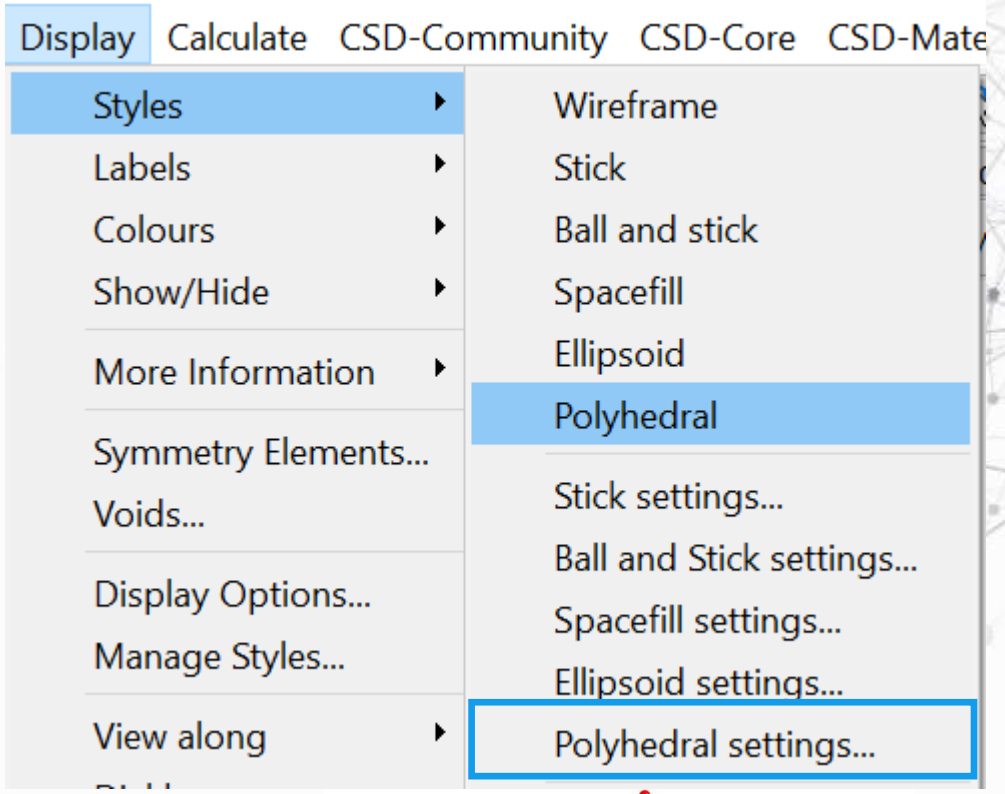


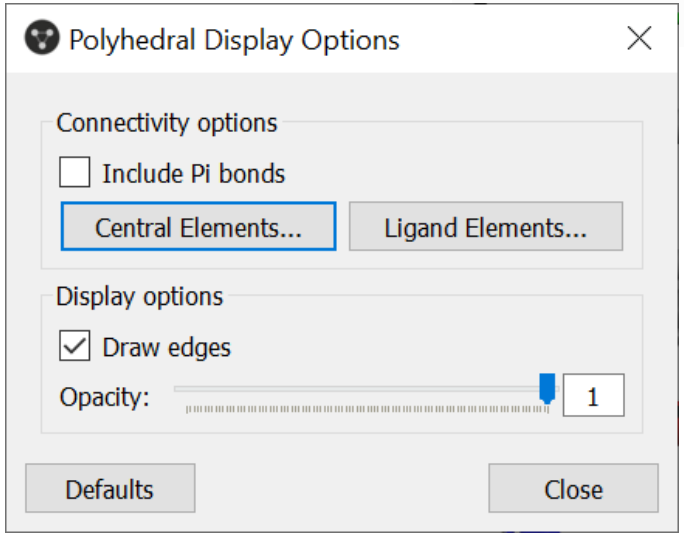
In Mercury, polymeric bonds are drawn alternating long and short lines.

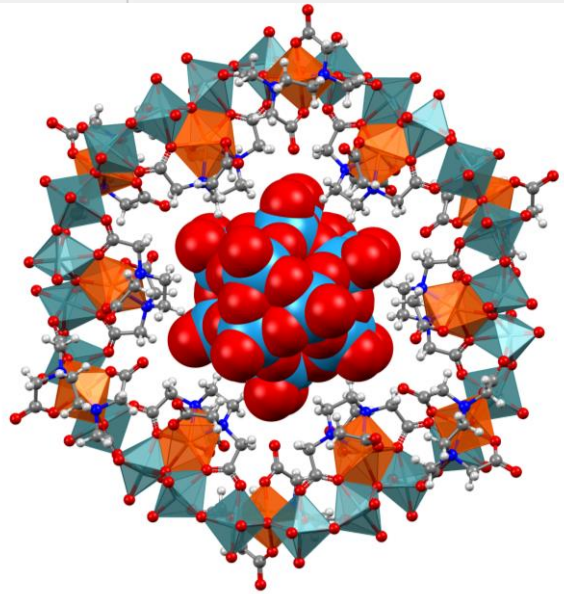
CCDC

Polyhedral display

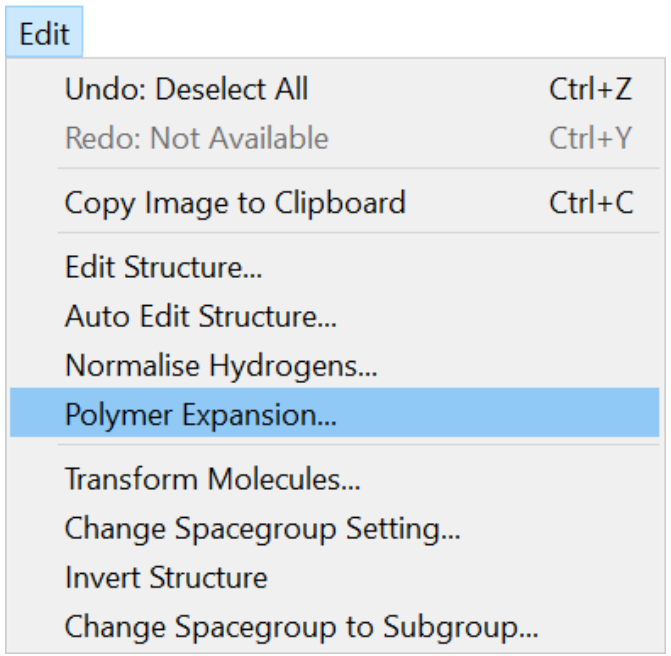
- Visualise metal-organic structures using polyhedral representations for the metal-organic coordination centres.
  - Central Elements - atoms around which polyhedra are centred
  - Ligand Elements - neighbouring atoms which create the corners of the polyhedron



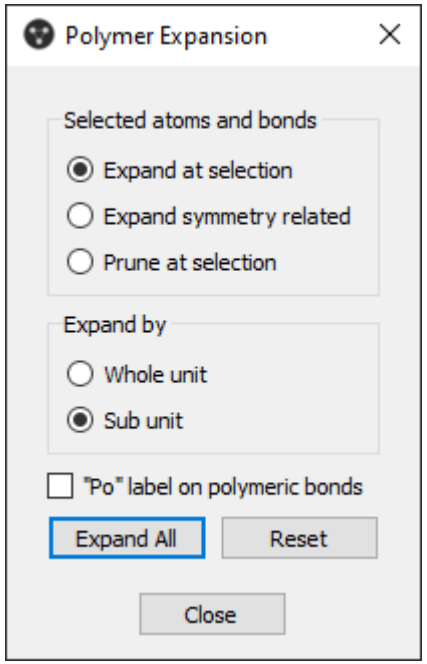


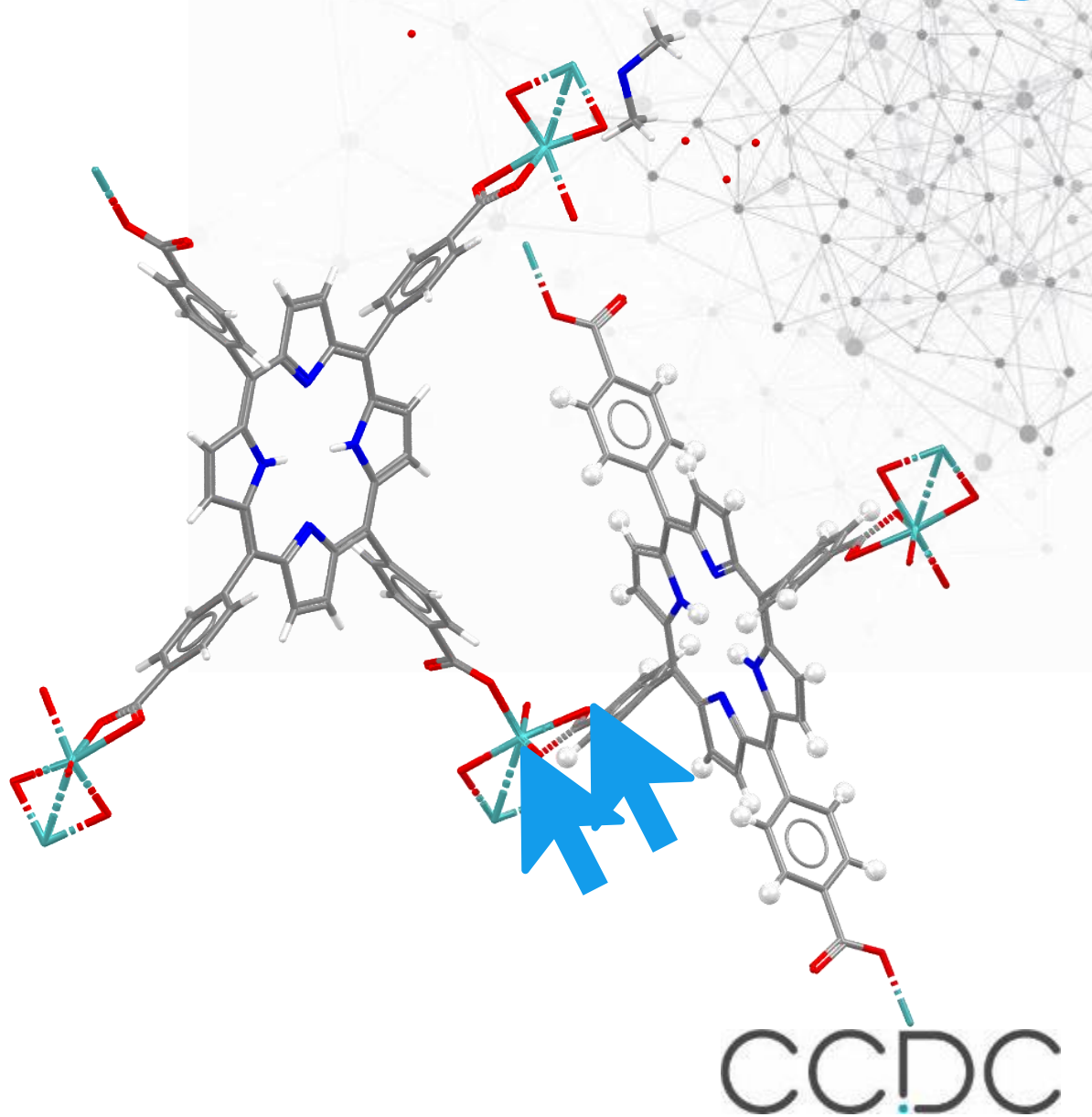


Polymer Expansion

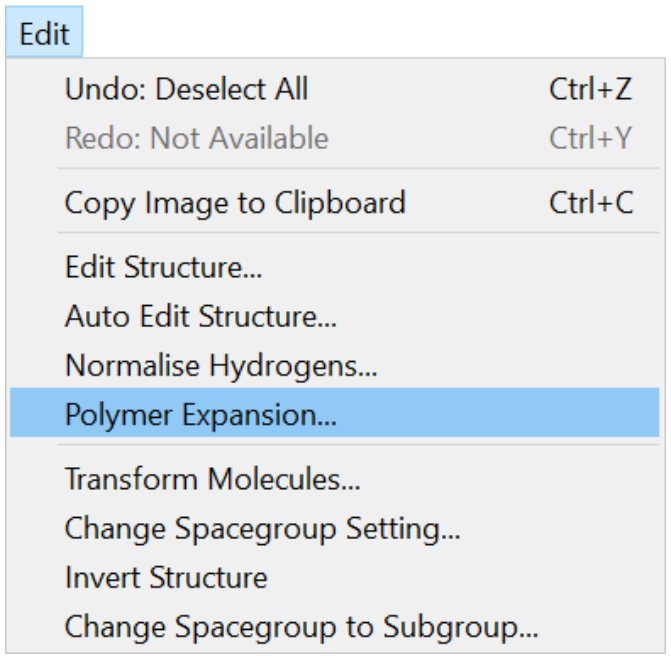


By **Sub unit**
it will add atoms and bonds up to the next polymeric atom or bond in the structure.

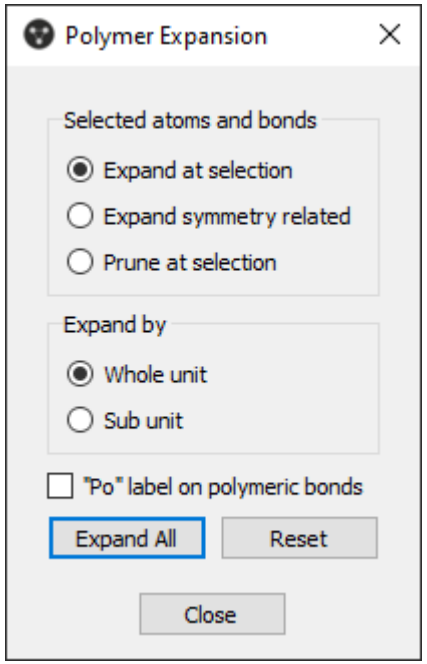


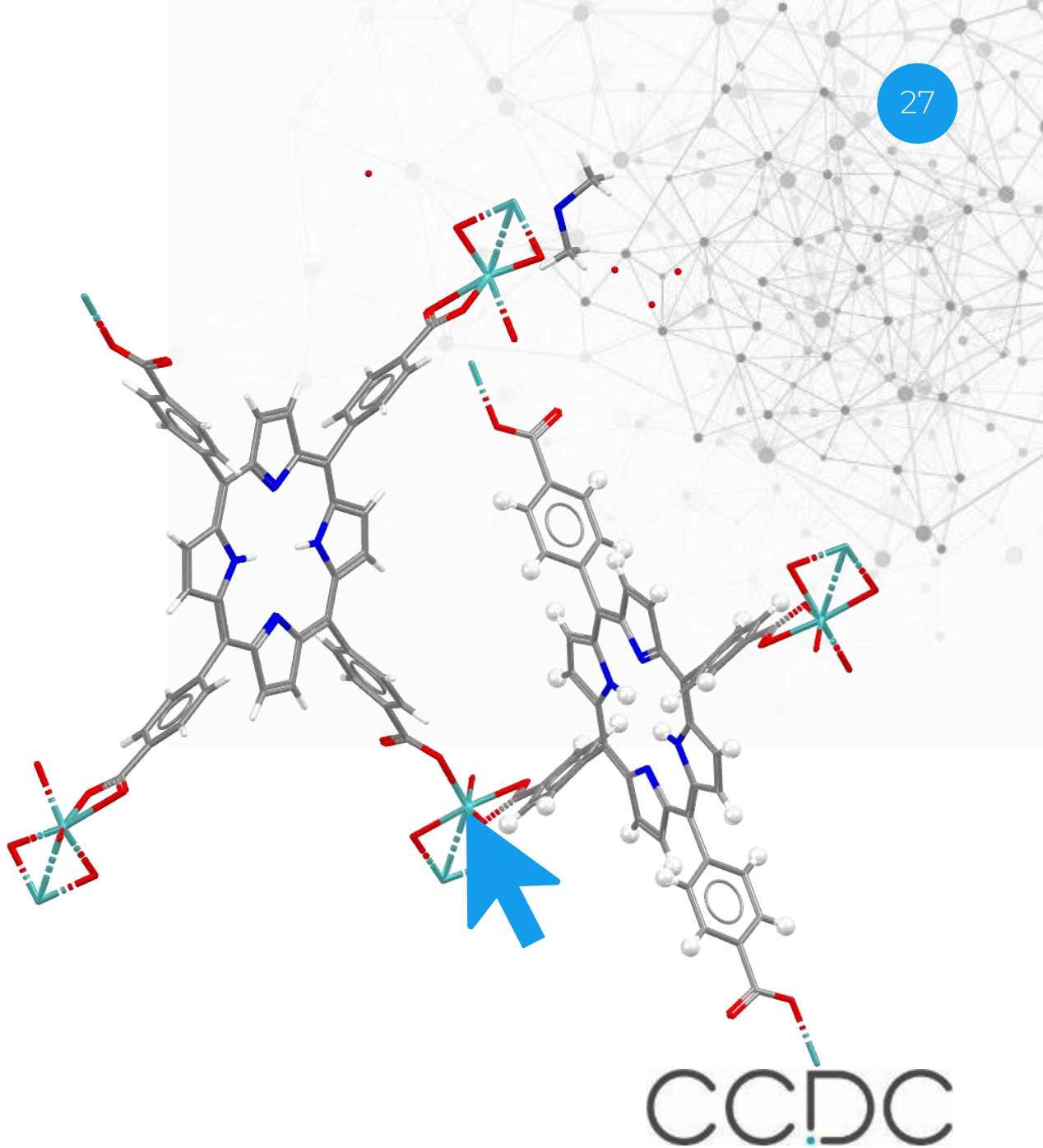


Polymer Expansion

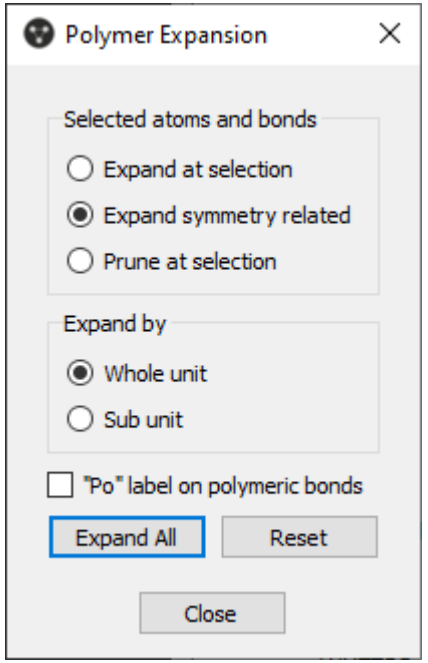


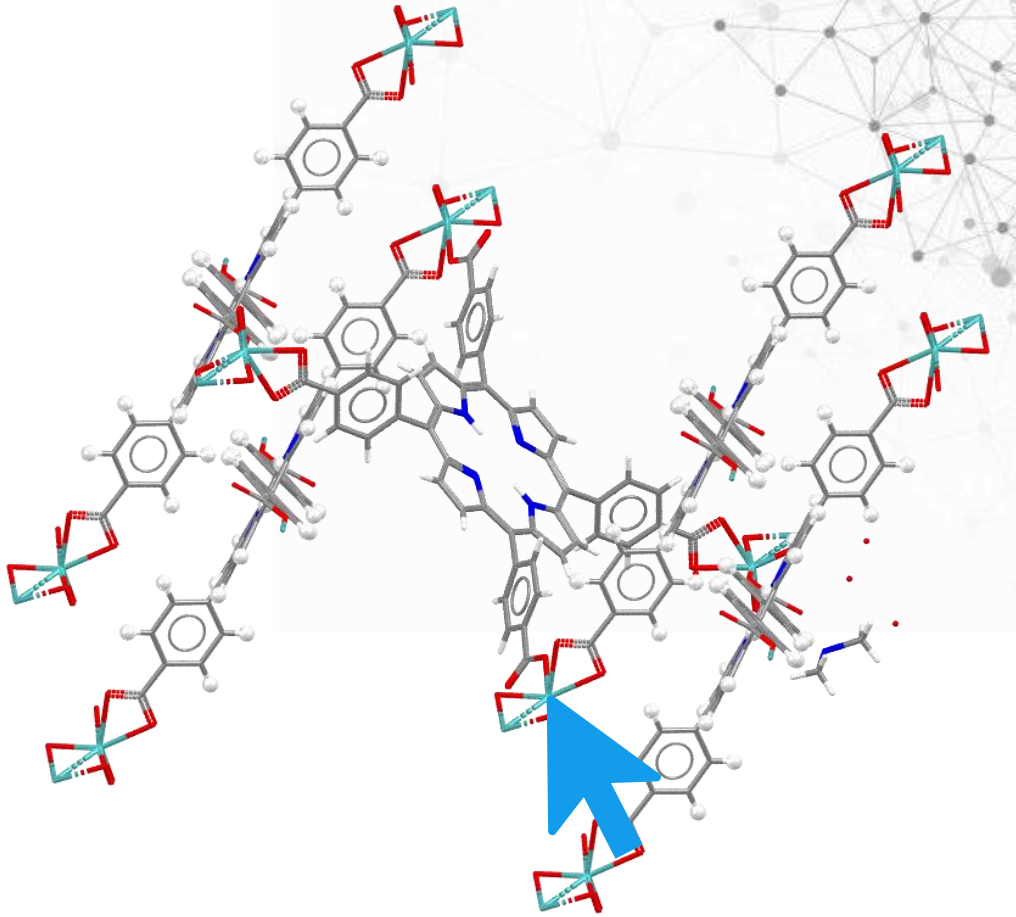
By **Whole unit**
It will add an additional repeat of the initial crystal chemical unit to the structure





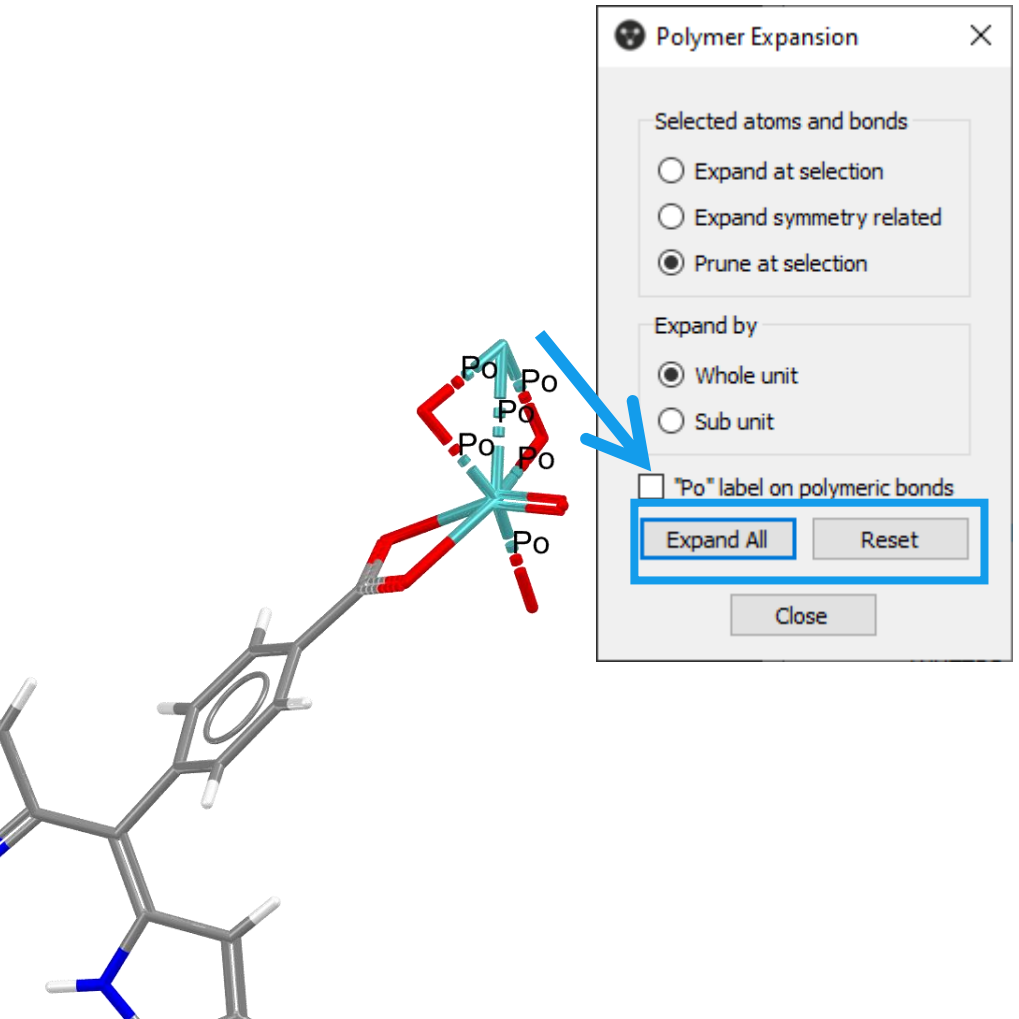
Polymer Expansion

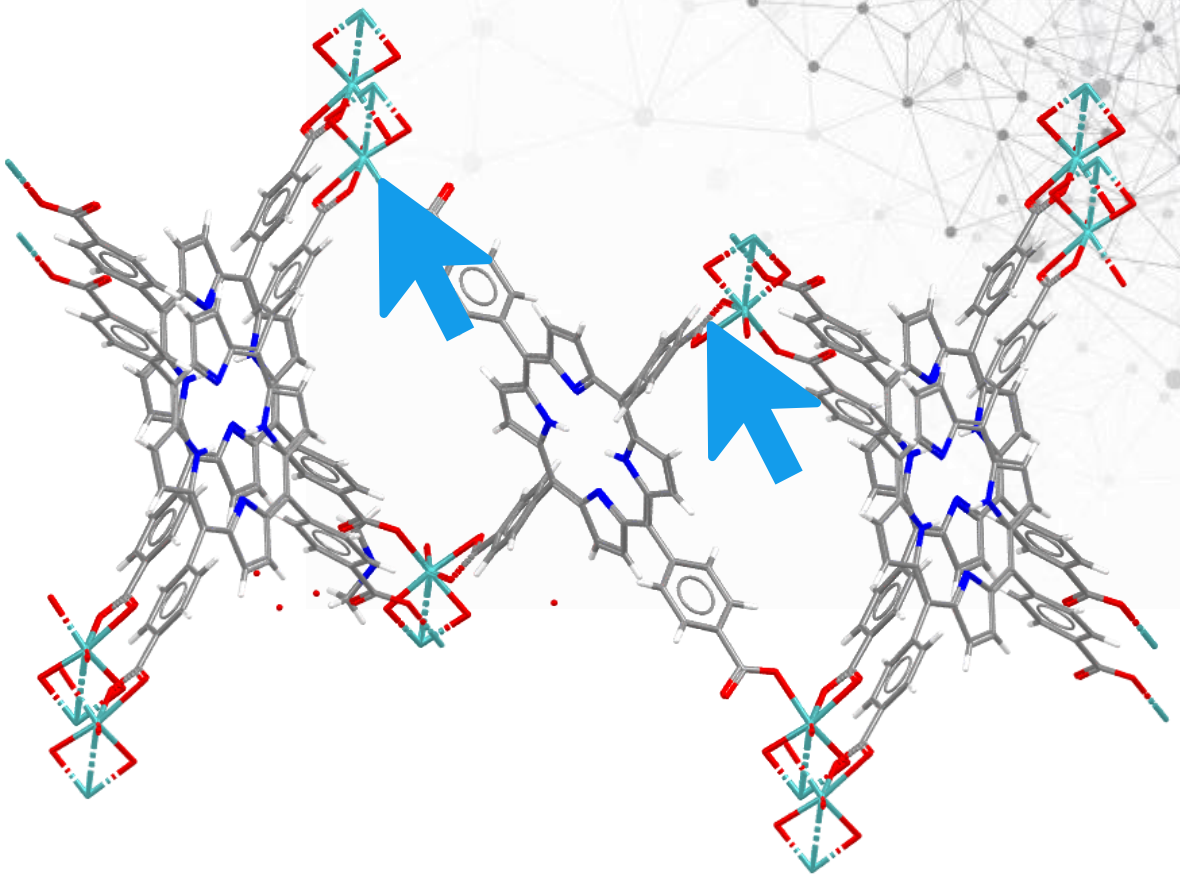




CCDC

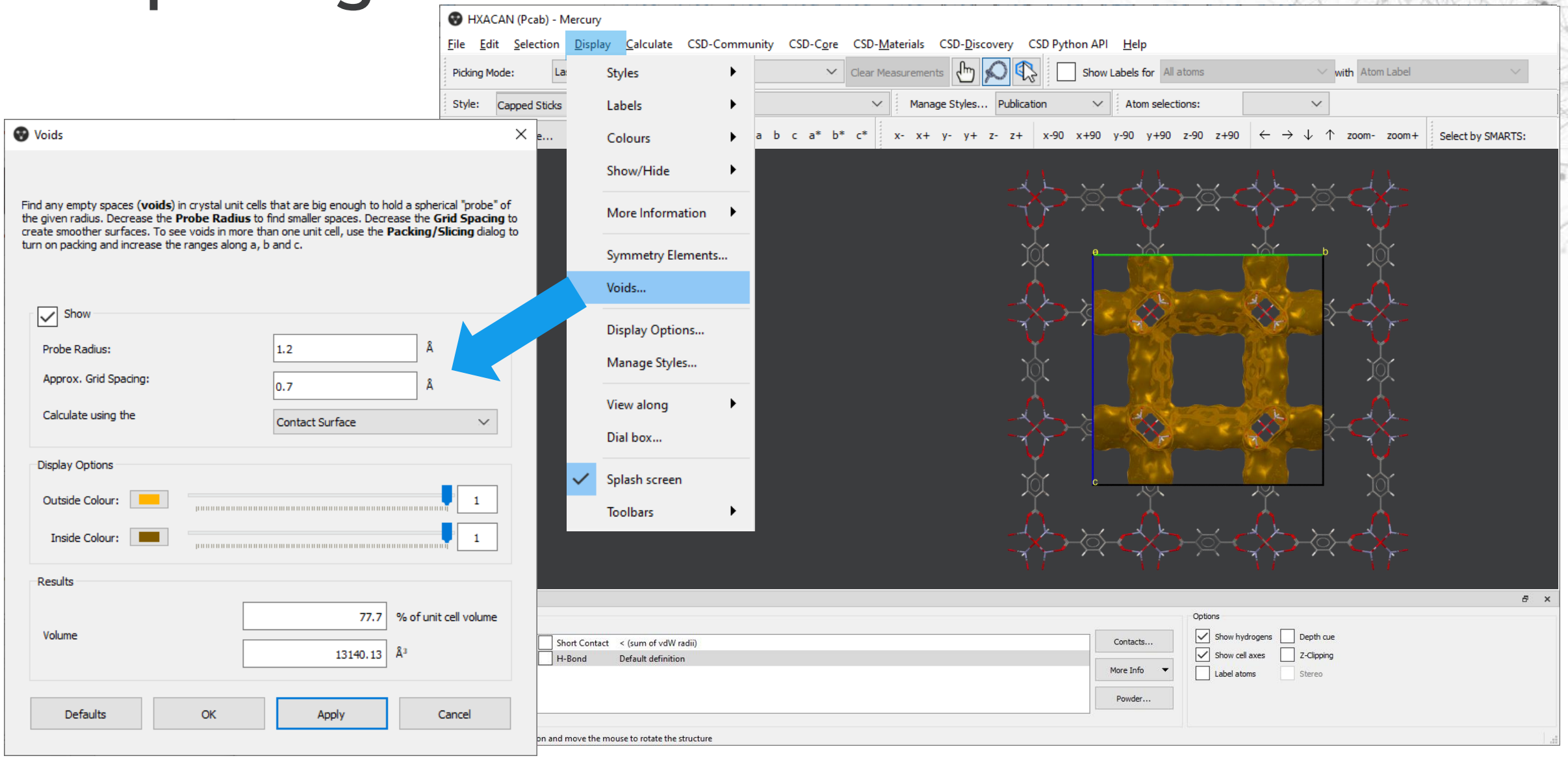
Polymer Expansion



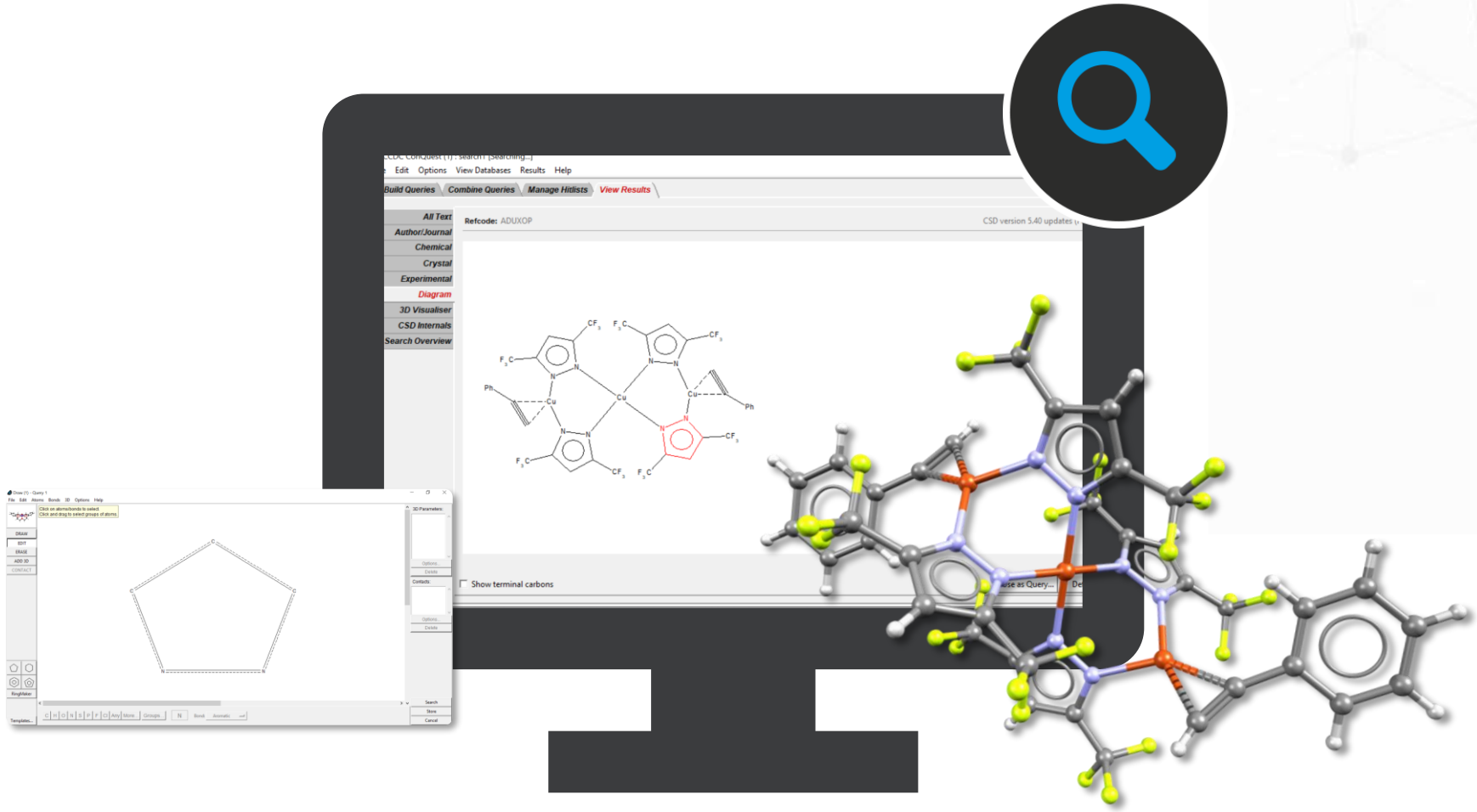


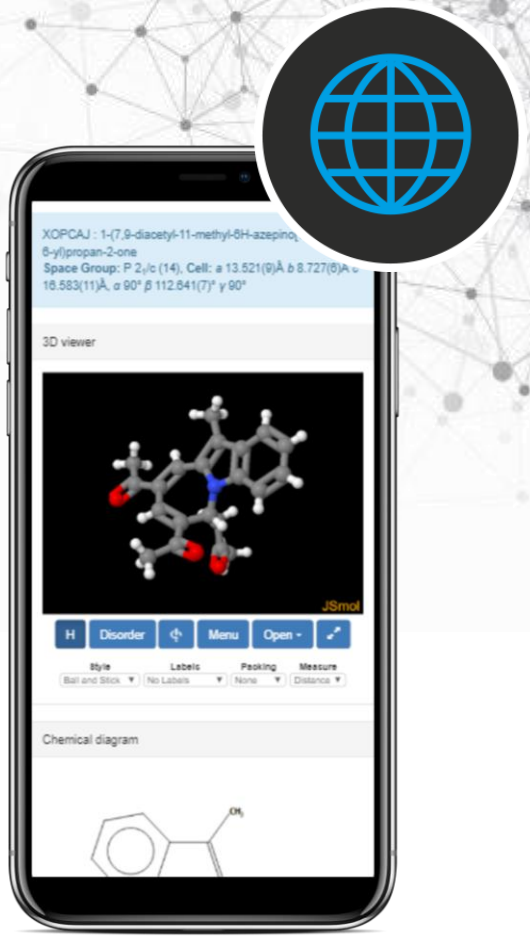
CCDC

Exploring voids



CSD-Core: Search

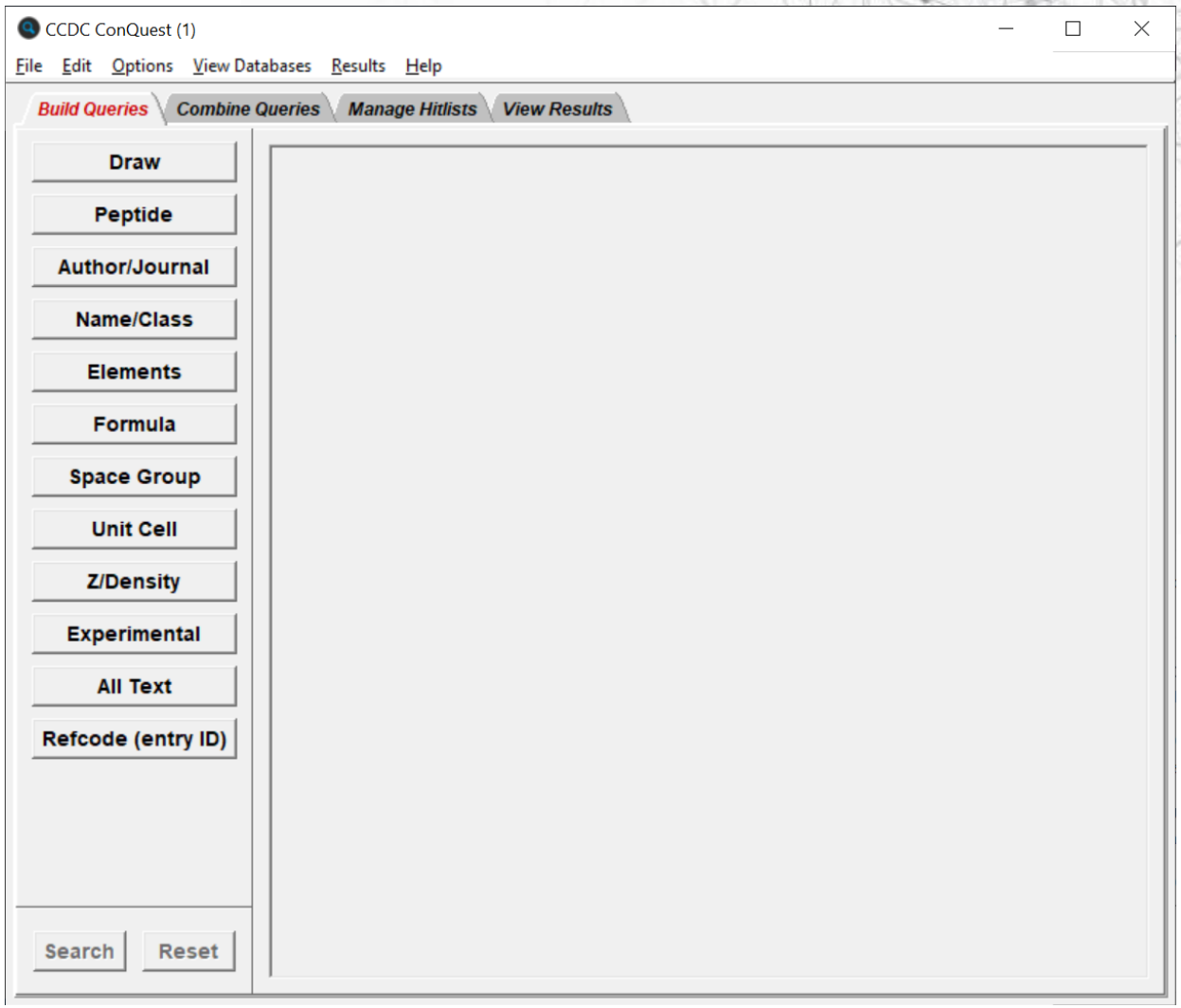




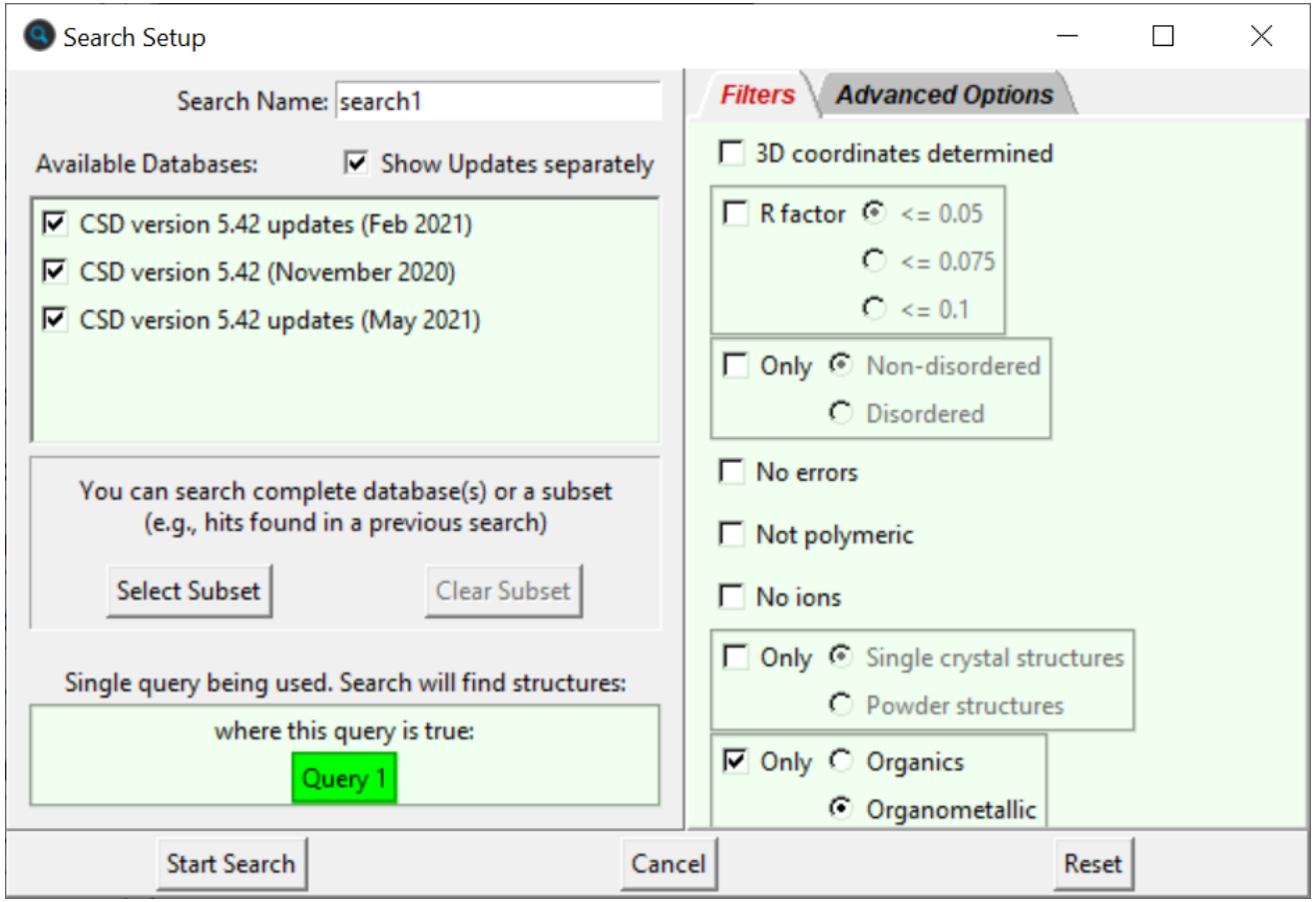
On your desktop: ConQuest

On a browser: WebCSD
Or Free Access Structures

Show One: Searching metal-organics



Organometallic filter

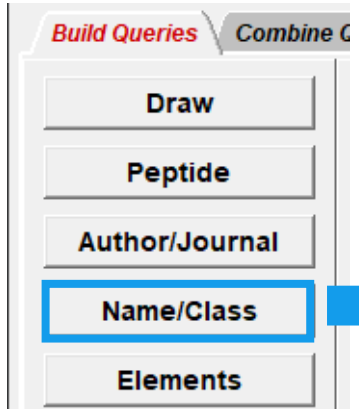


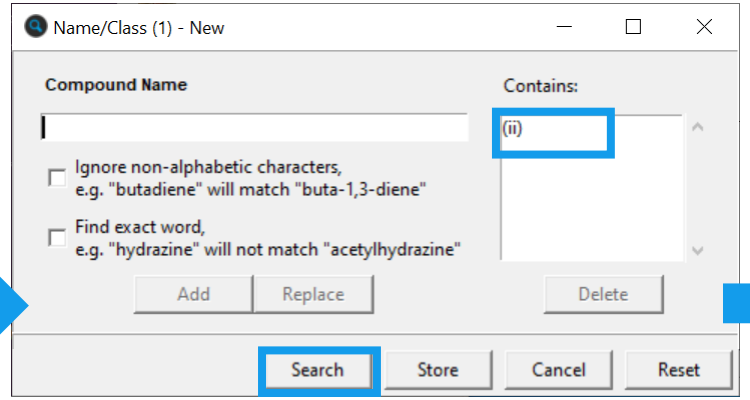
- Restricts search to structures containing at least one:
  - Transition metal
  - Lanthanide
  - Actinide
  - Or any of Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po

CCDC

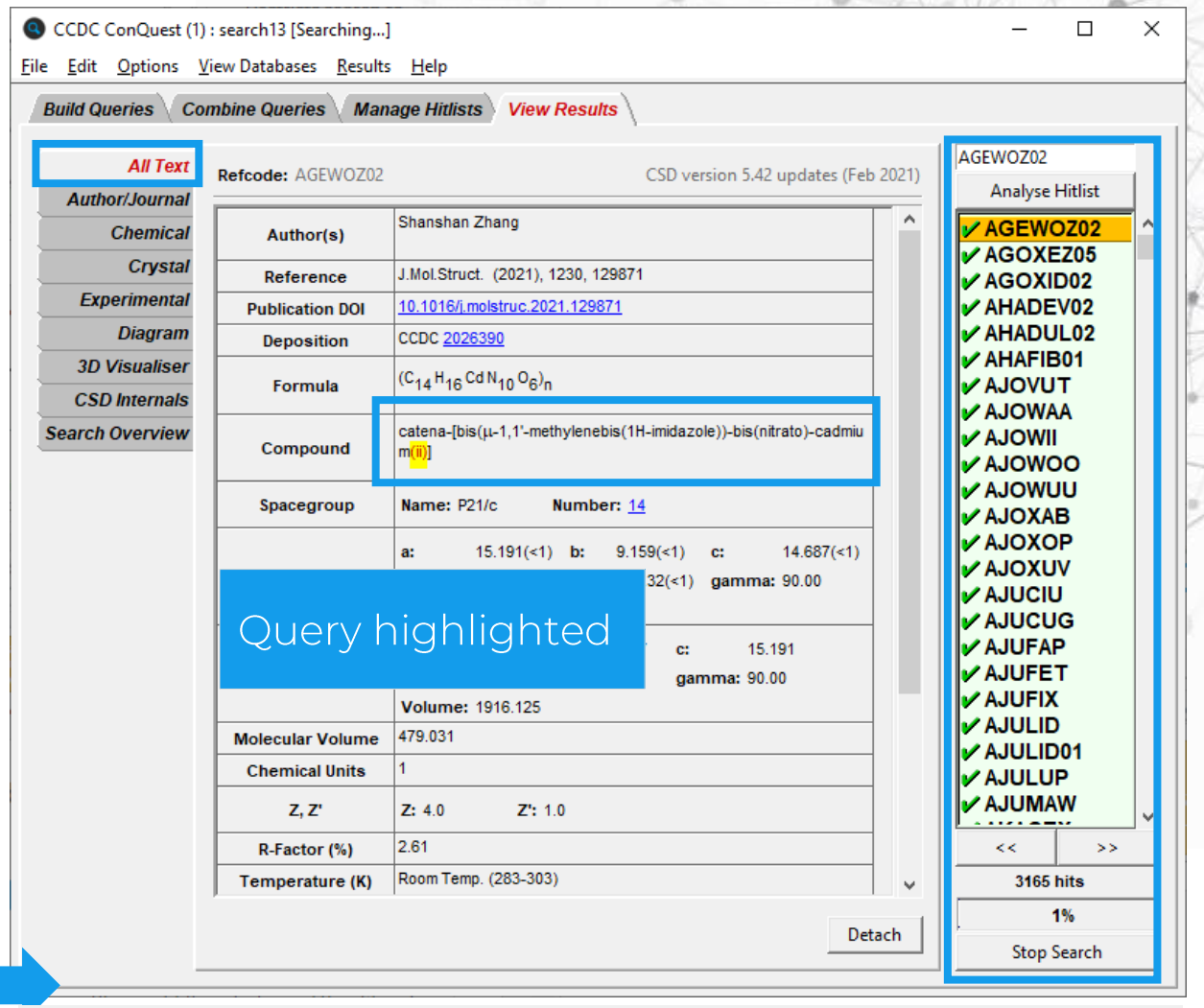
Oxidation states

- Charges only added to 2D/3D for charge balancing purposes
- Oxidation states are provided in compound name
  - Convention is (i-viii)
- >340,000 entries with oxidation states





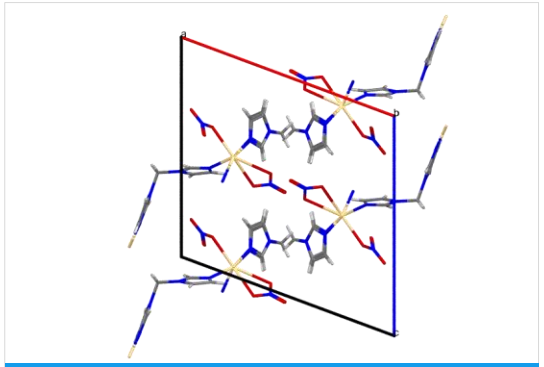
Tip: use parenthesis to avoid
- Other oxidation states
- Polymorphic forms!



catena-[bis(μ-1,1'-methylenebis(1H-imidazole))-bis(nitrato)-cadmium(ii)]

Compound names include information about polymeric nature, oxidation states, ligand names, clathrates and solvates and hydrate

CSD Refcodes



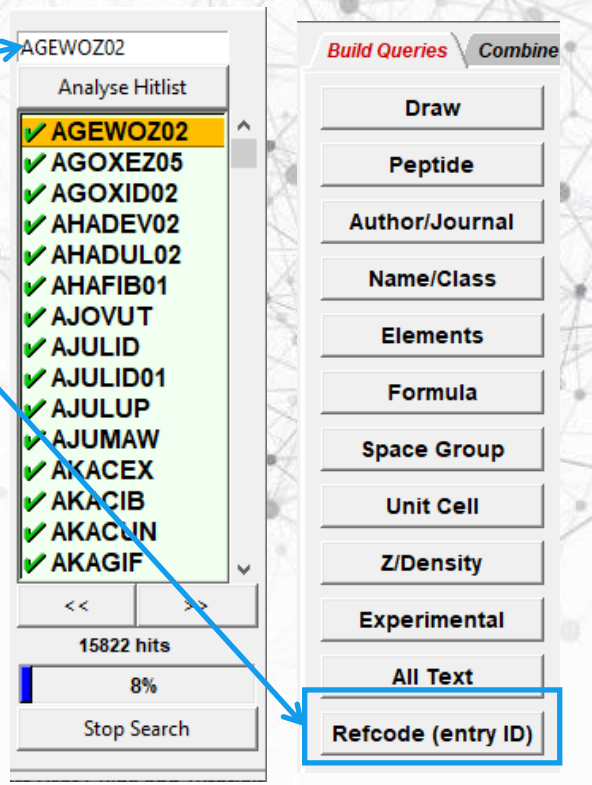
CSD Refcode
AGEWOZ02

You can use the box to jump to a recode in your search results

You can search the entire CSD for a Recode too

What is AGEWOZ02?

- A CSD Refcode
- A database reference code
- Containing 6-8 characters
- Used to identify entries in the CSD

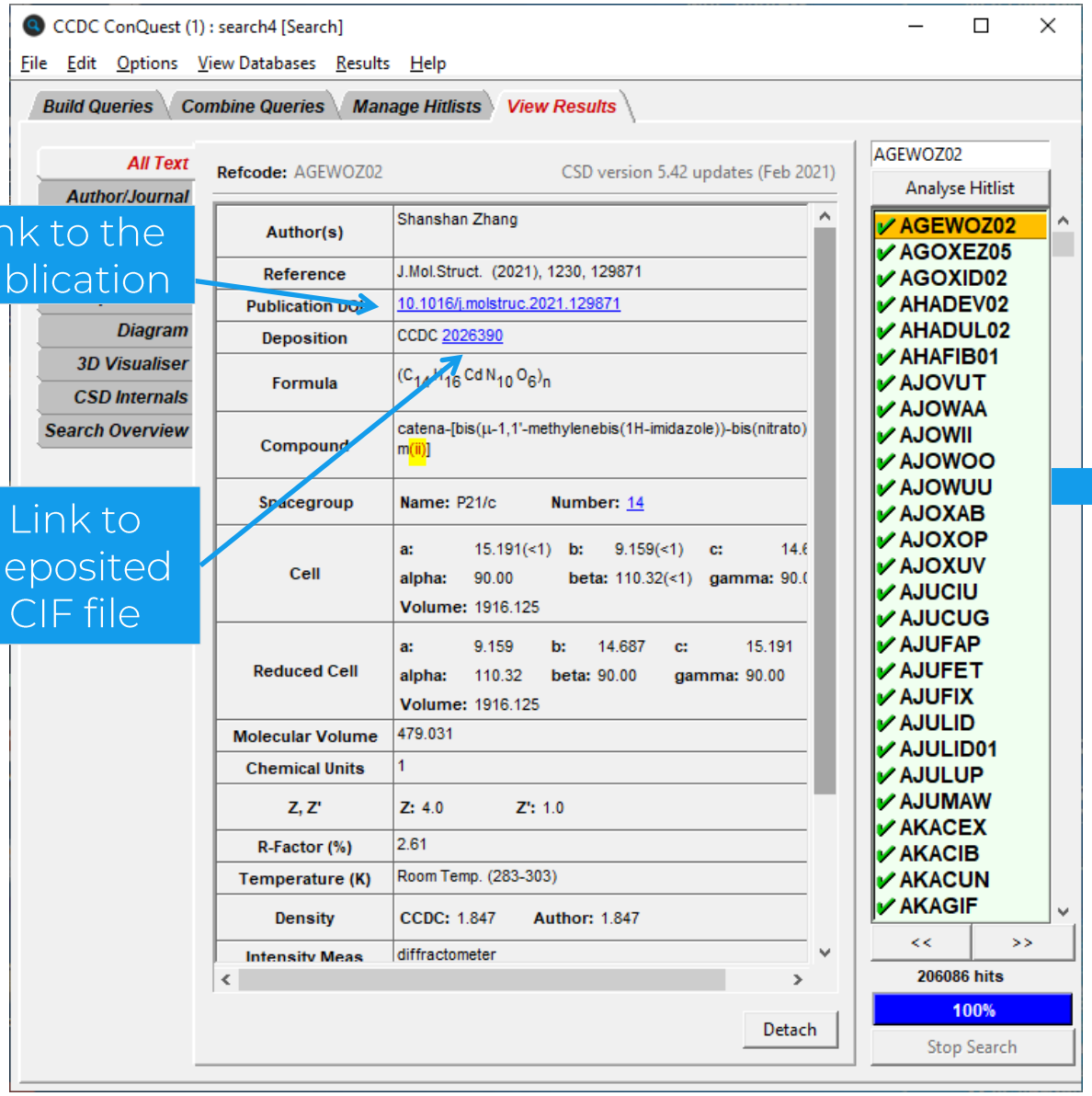


Refcode families

- The same substances are assigned the *same* 6 letter code plus an additional 2 numbers
  - Polymorphs
  - New determinations or re-refinements of the same substance
  - Determinations at different temperatures/pressures
- Stereoisomers or different solvates, co-crystals, etc are assigned *different* refcode families

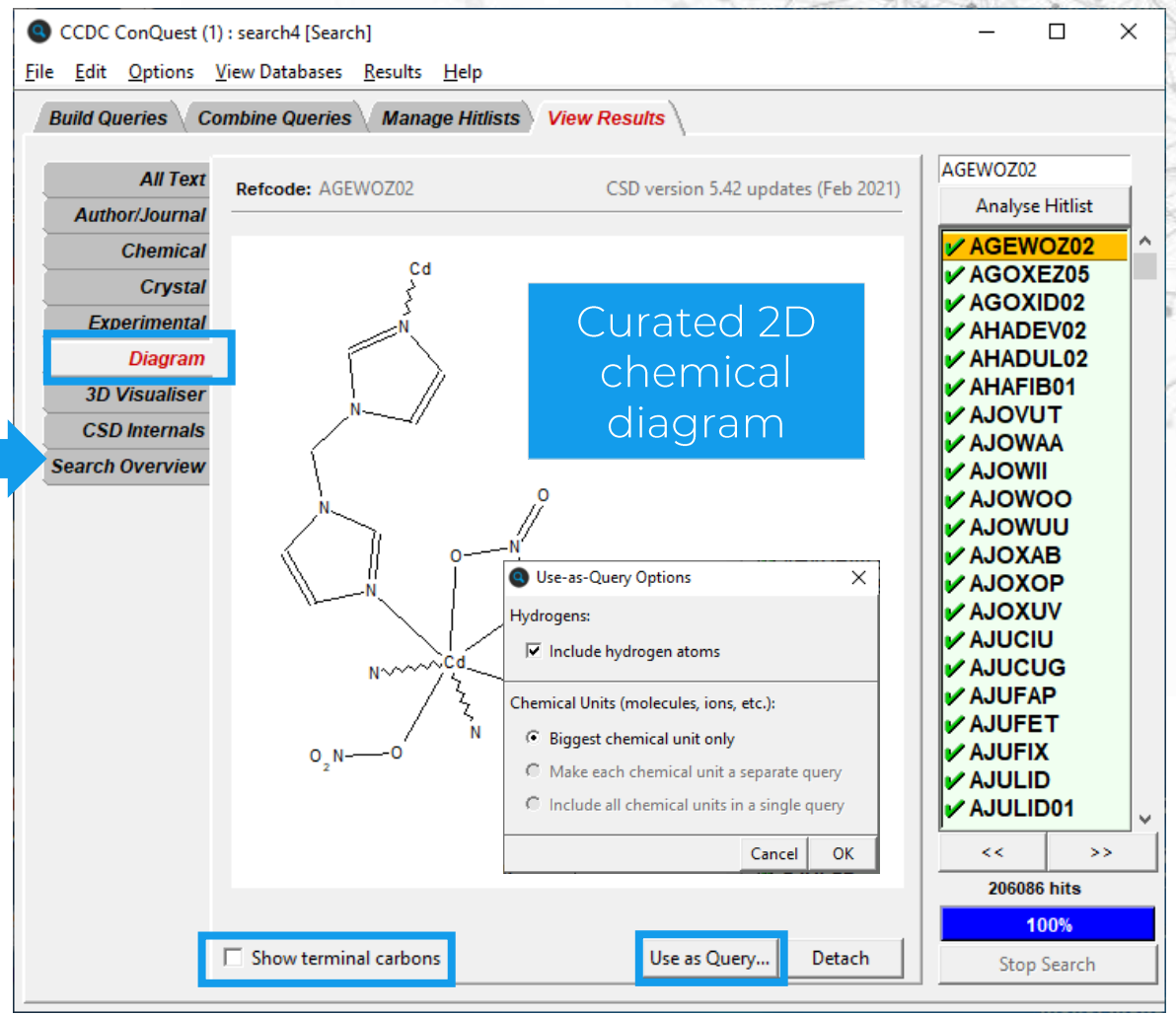
- Some other recognisable refcodes are: KITTEN, DISNEY, GAUTAM, NELSON, GLYCIN
- Do you have a favourite refcode? You can use the chat box to tell us yours!

CCDC

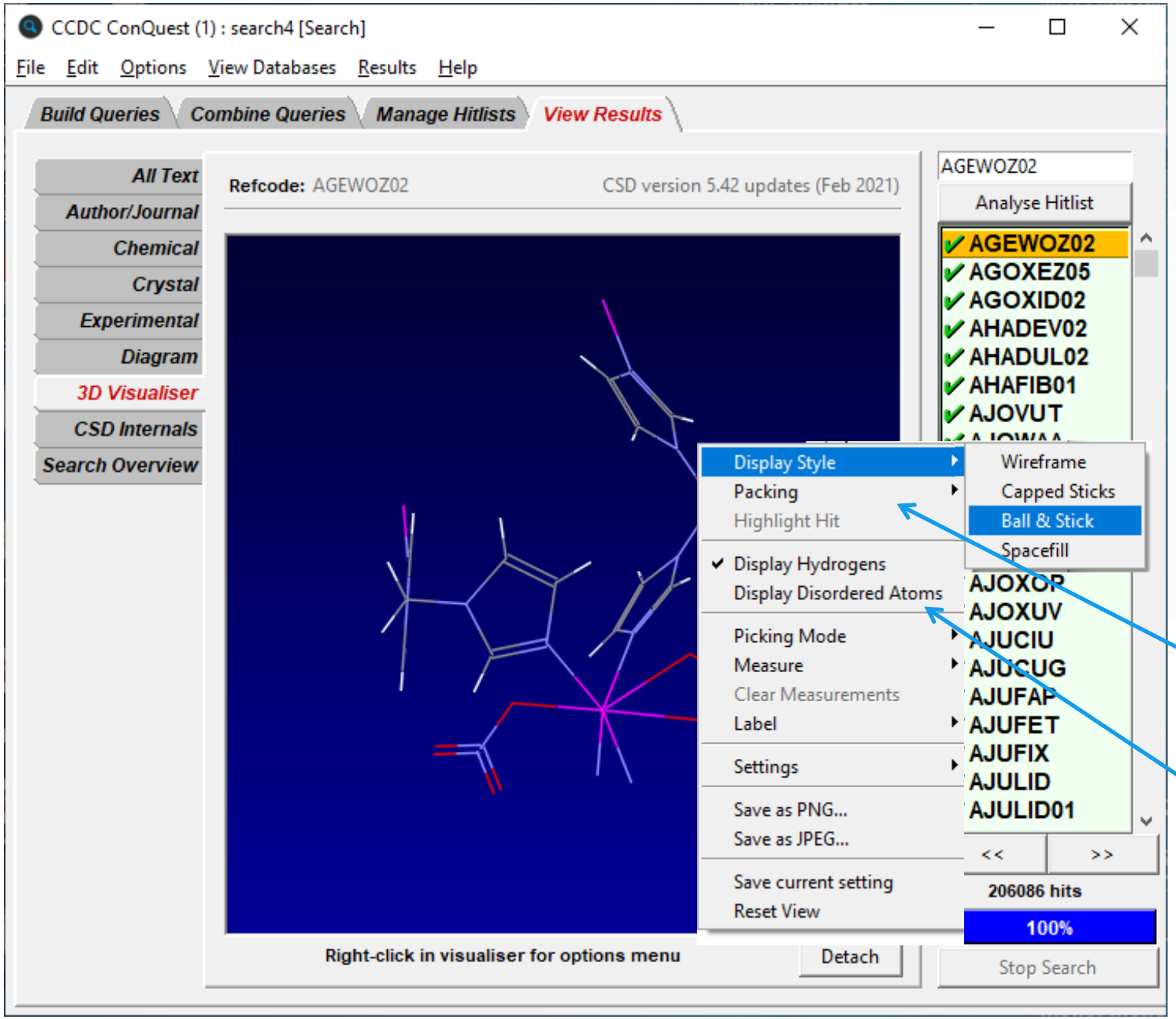


Link to the publication

Link to deposited CIF file



CCDC



Left clicking each tab will display different information

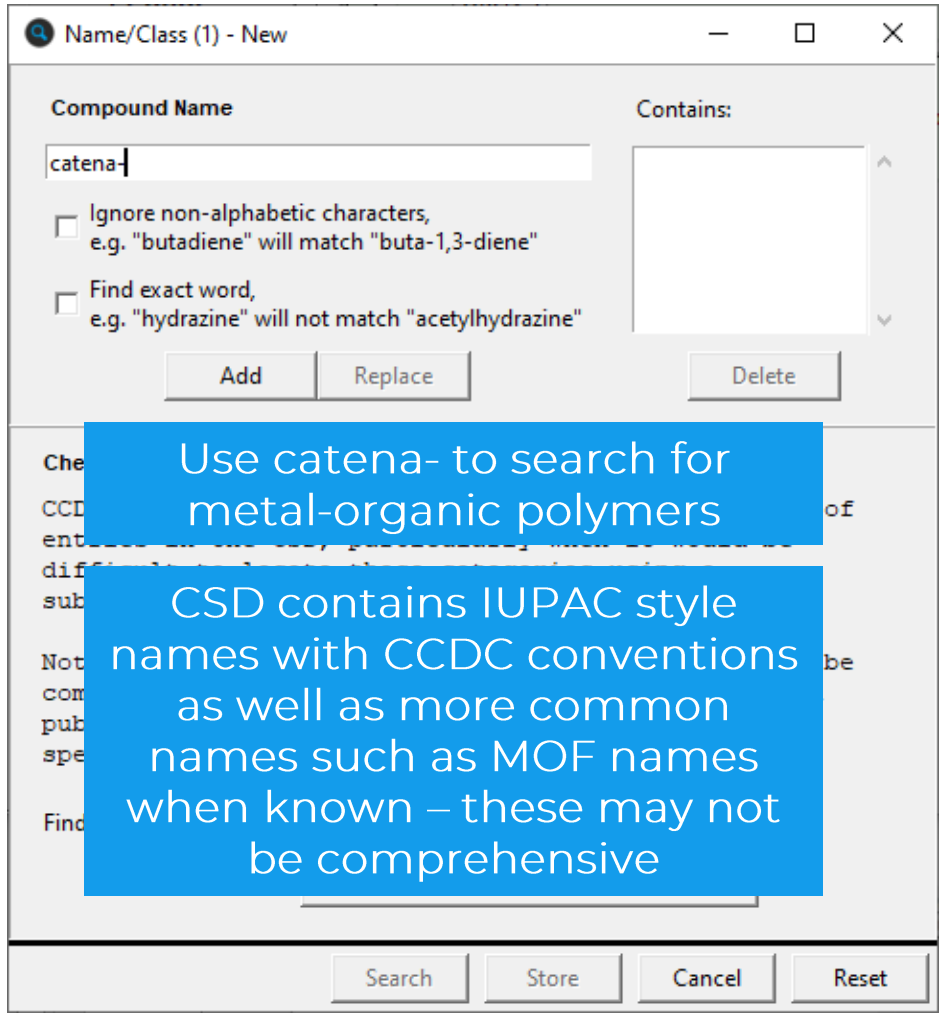
Right clicking on 3D visualiser enables you to change the display style

Packing > Molecule or unit cell

Display or hide disordered atoms

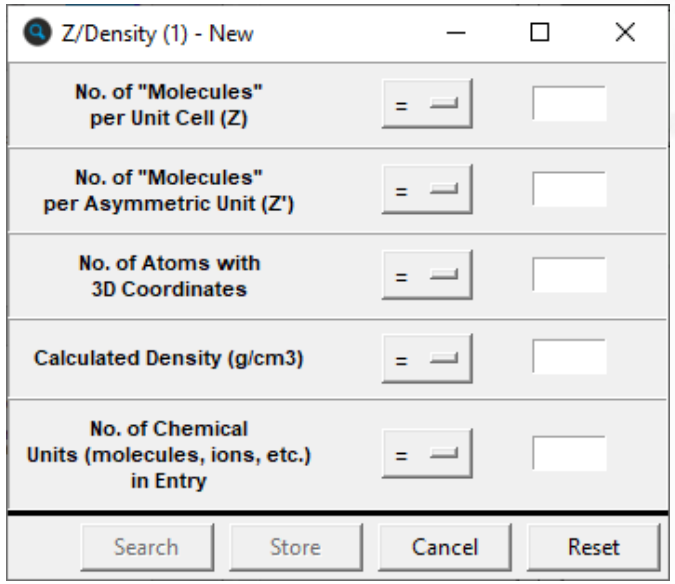
CCDC

ConQuest – more text/numeric searches

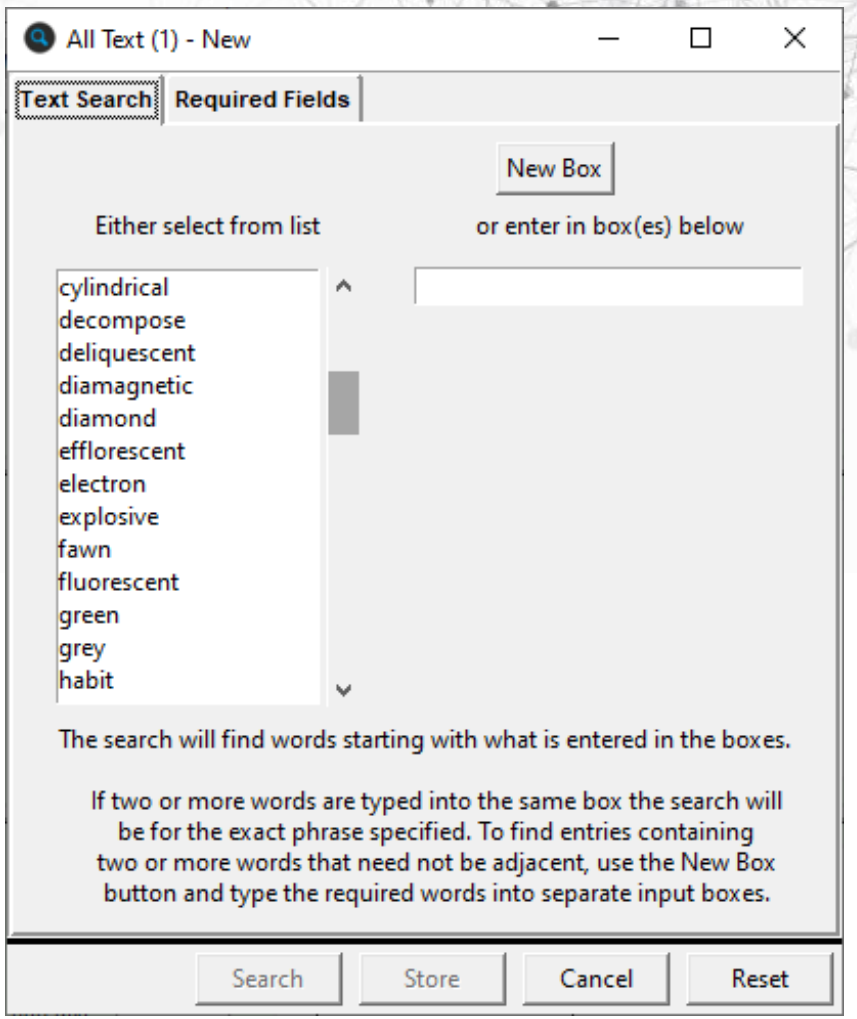


Use catena- to search for metal-organic polymers

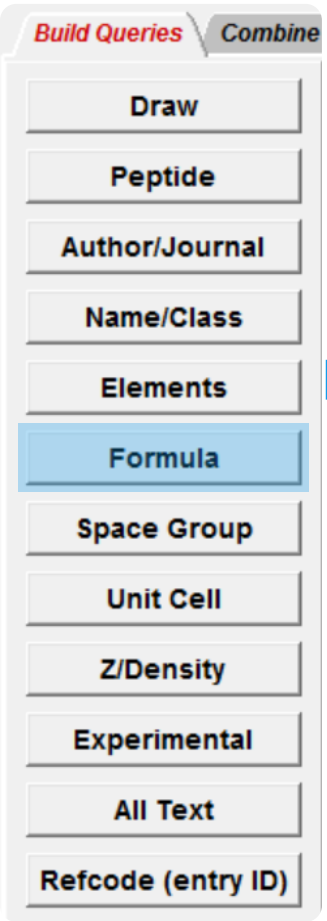
CSD contains IUPAC style names with CCDC conventions as well as more common names such as MOF names when known – these may not be comprehensive

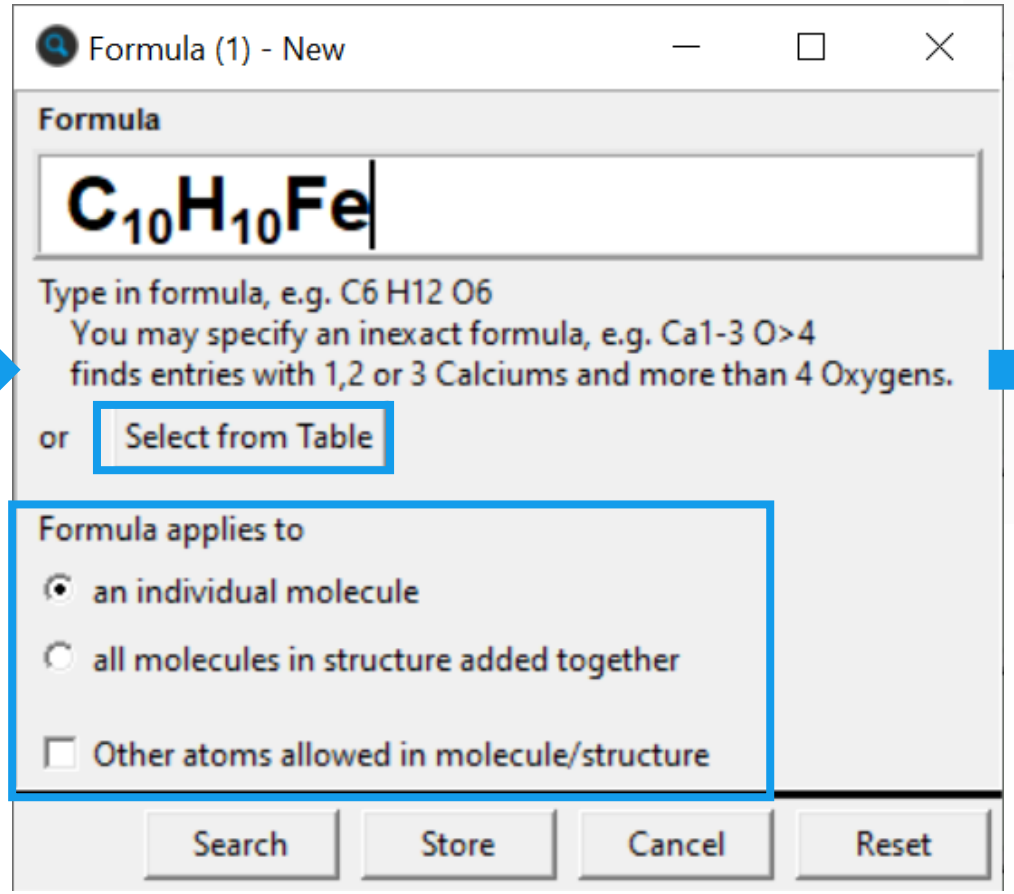


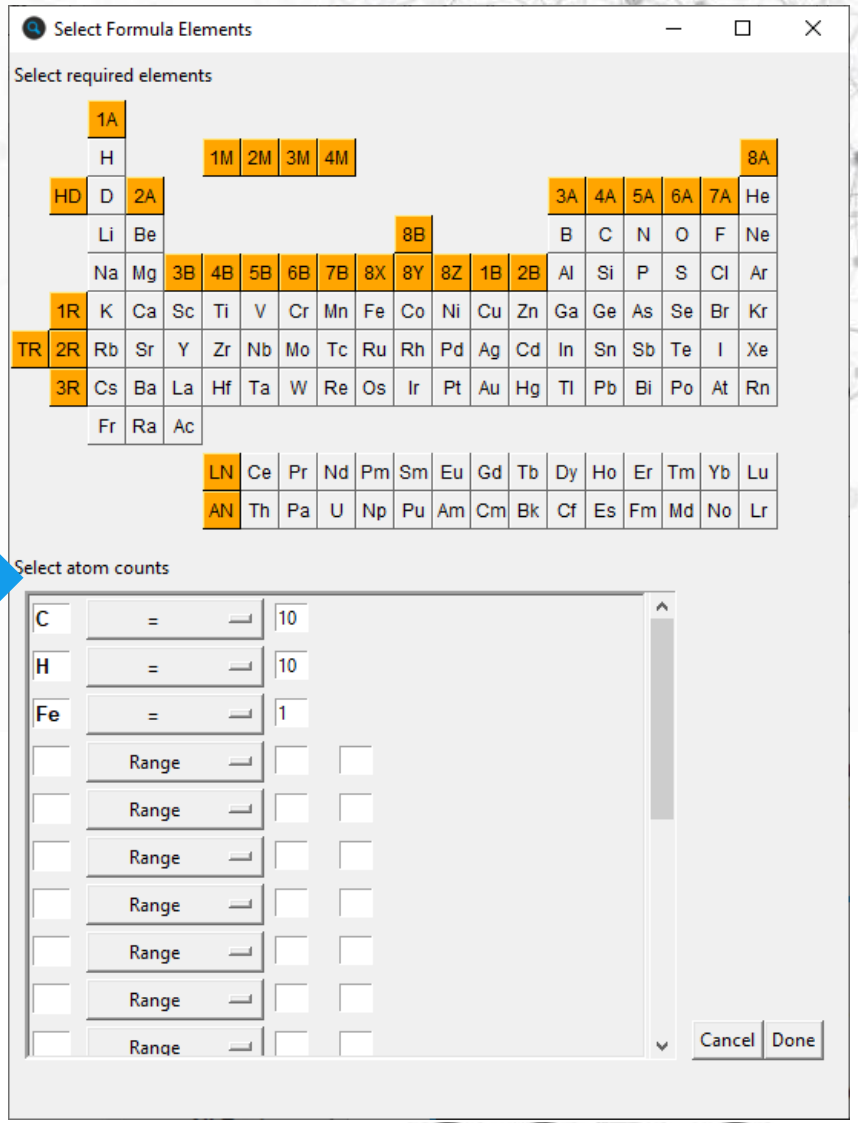
Tip: use a combined search to find structures with catena- or catena(and this will avoid finding catenanes



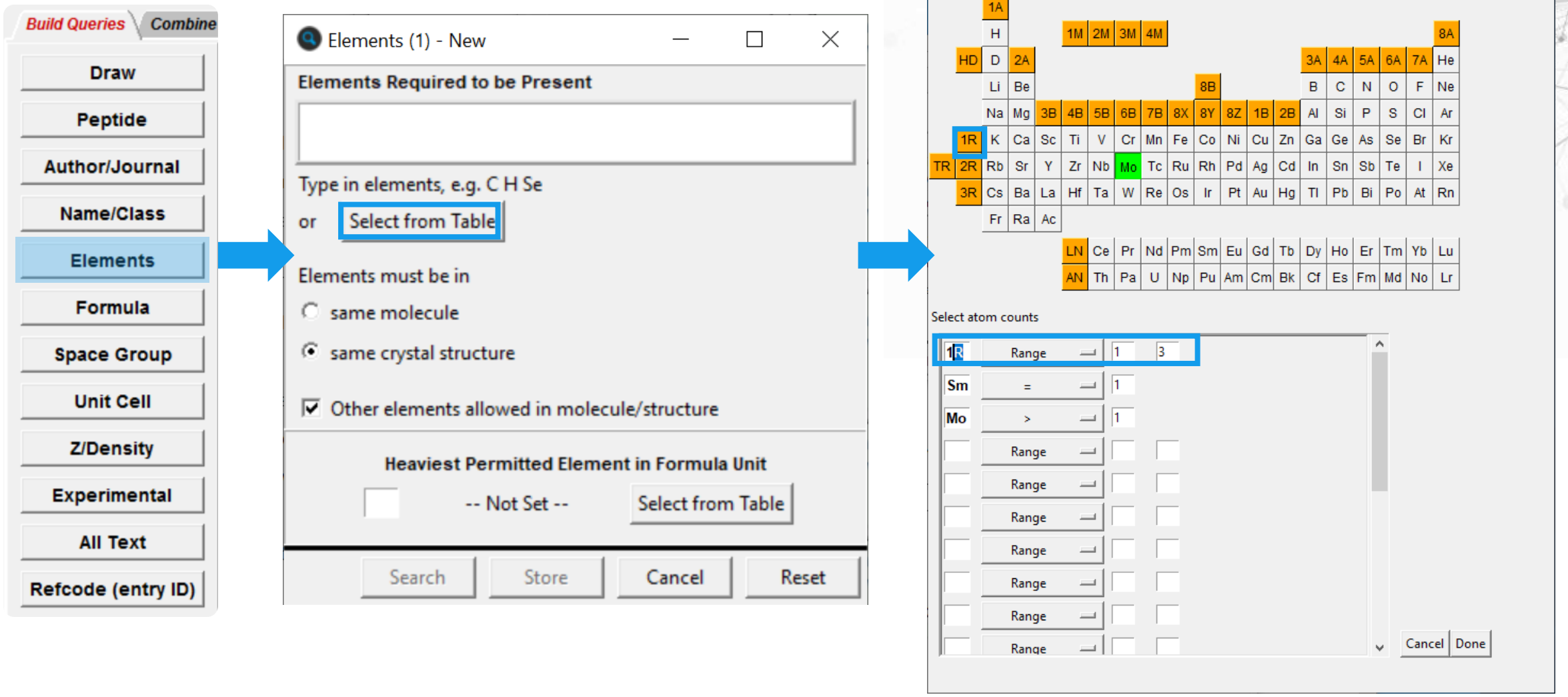
ConQuest – Formula search



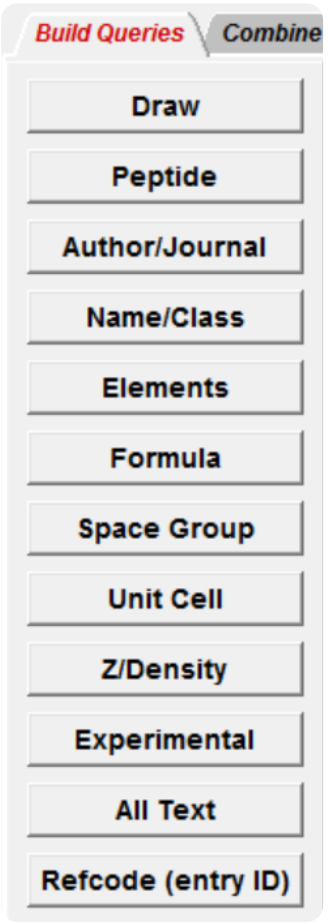


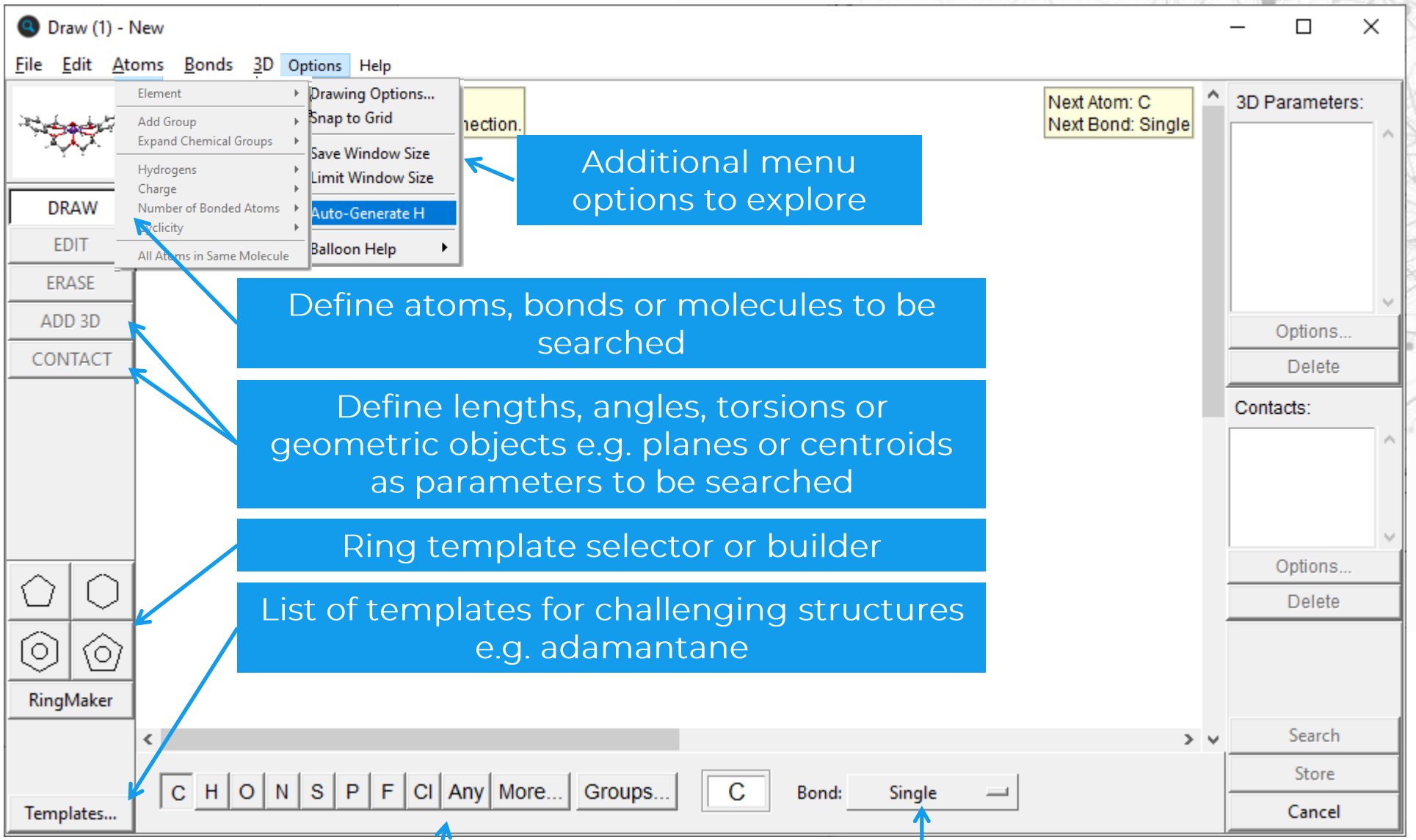


ConQuest – Elements search



Draw





Additional menu options to explore

Define atoms, bonds or molecules to be searched

Define lengths, angles, torsions or geometric objects e.g. planes or centroids as parameters to be searched

Ring template selector or builder

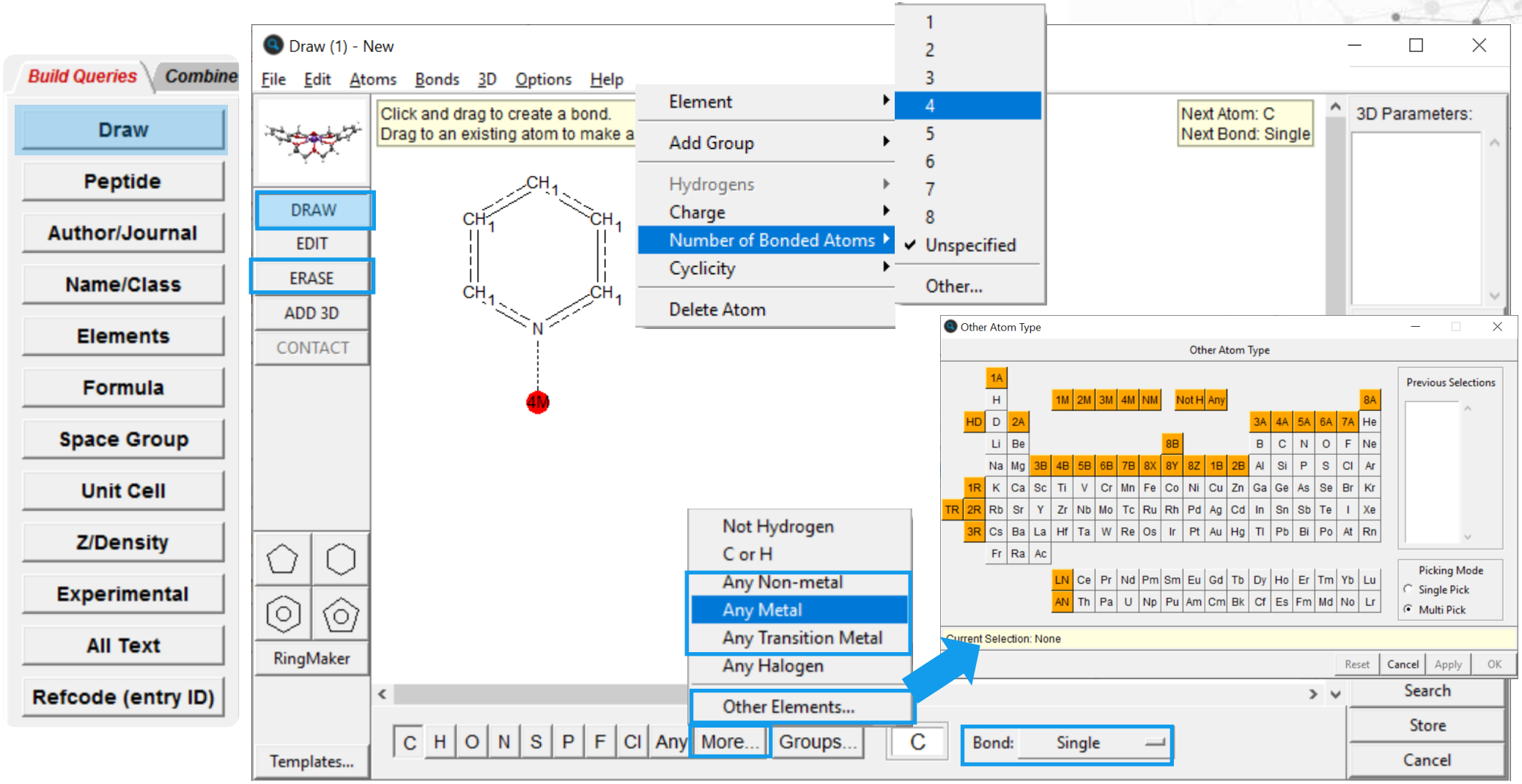
List of templates for challenging structures e.g. adamantane

Select specific or general atom types/functional groups

Select bond type

CCDC

ConQuest – Draw/Structure search



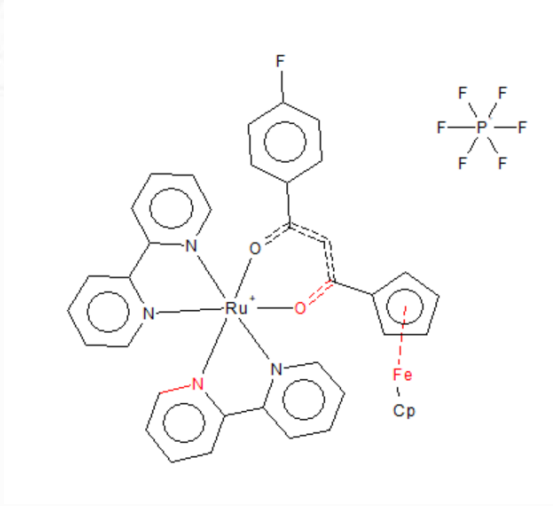
Left click in space to add an atom

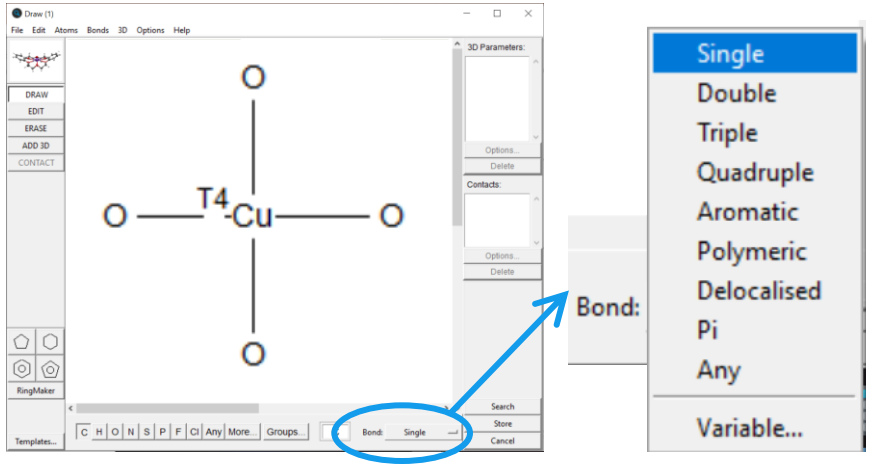
Left click and drag from an atom to attach an atom

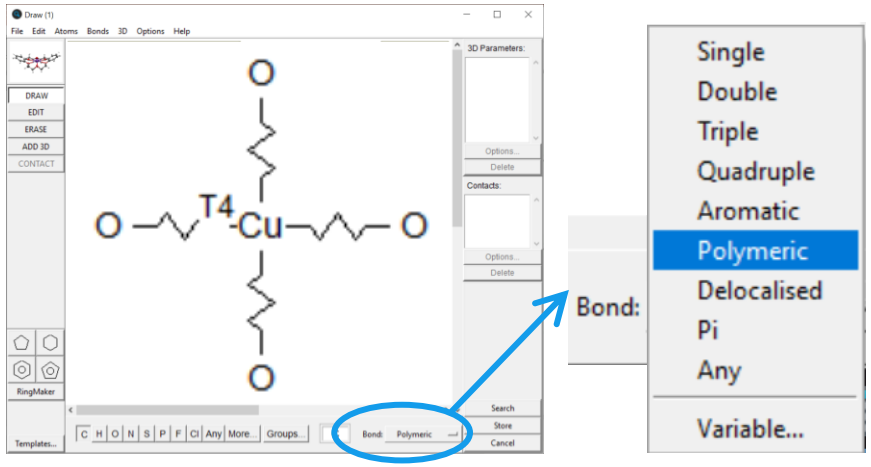
Right click on an atom or bond to edit or add extra properties to it

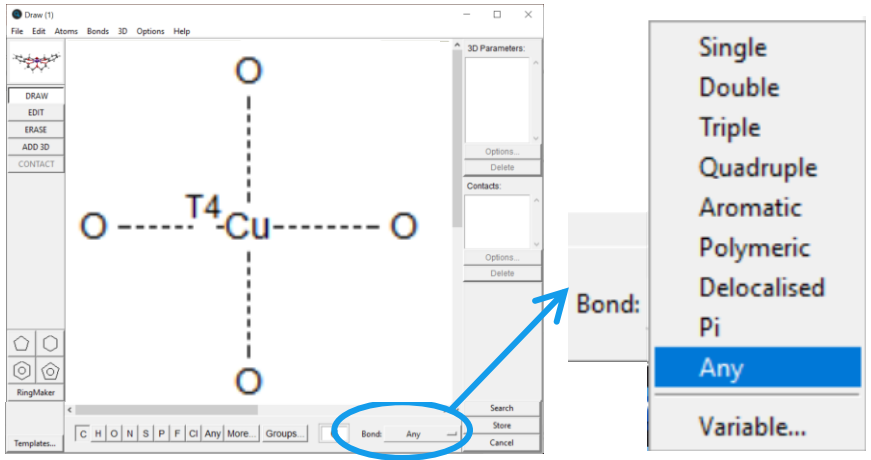
Some useful CCDC drawing conventions

- When appropriate the CSD uses:
  - Aromatic bonds
  - Delocalised bonds
  - Pi-bonds
  - Polymeric bonds – depicted polymeric unit may vary
- CSD 2D chemical representations usually include:
  - The CCDC representation of the structure
  - Missing H atoms/other atoms
  - Squeezed or undetermined parts of the structure
  - Charge – this is the property of an atom rather than an entity but is only used for charge balance purposes



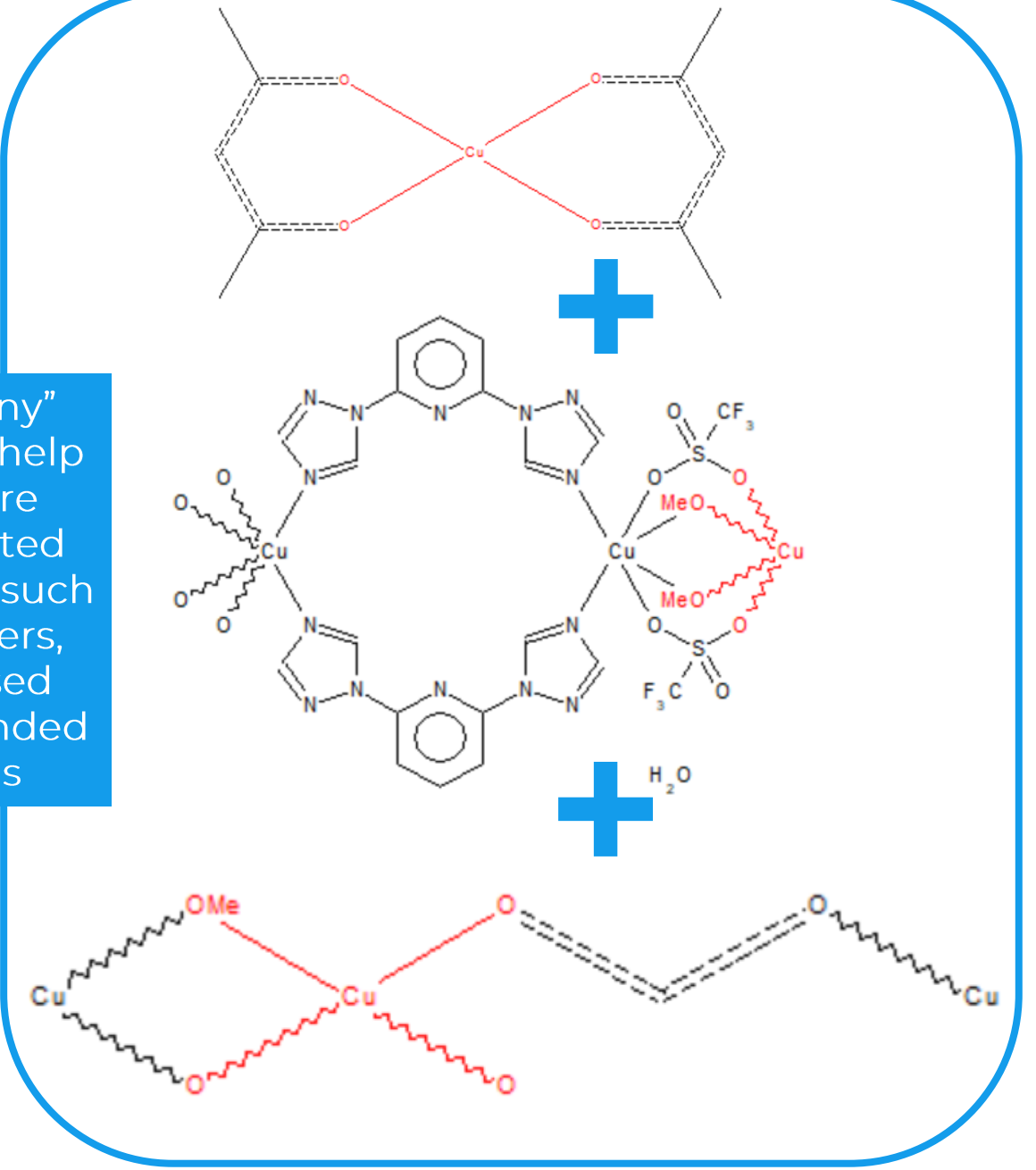






Using “Any” bond can help find more complicated structures such as polymers, delocalised and pi-bonded systems

Substructure found



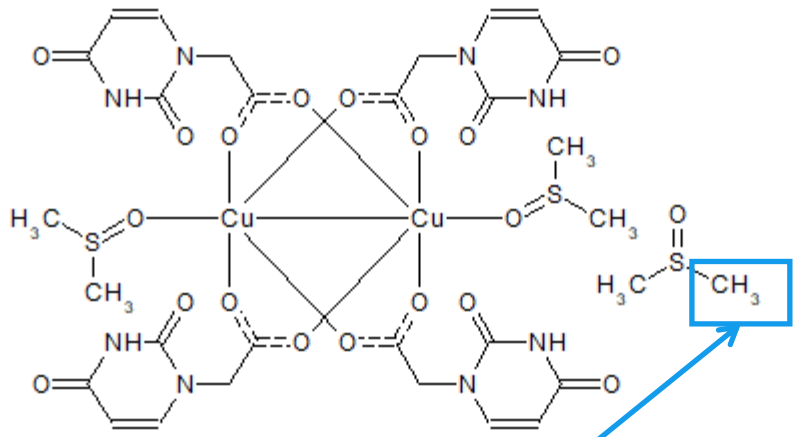
CCDC ConQuest (1) : search11 [Search]

File Edit Options View Databases Results Help

Build Queries | Combine Queries | Manage Hitlists | View Results

All Text
Author/Journal
Chemical
Crystal
Experimental
Diagram
3D Visualiser
CSD Internals
Search Overview

Refcode: TUWMOP CSD version 5.42 (November 2020)



Show terminal carbons

Use as Query... Detach

TUWMOP

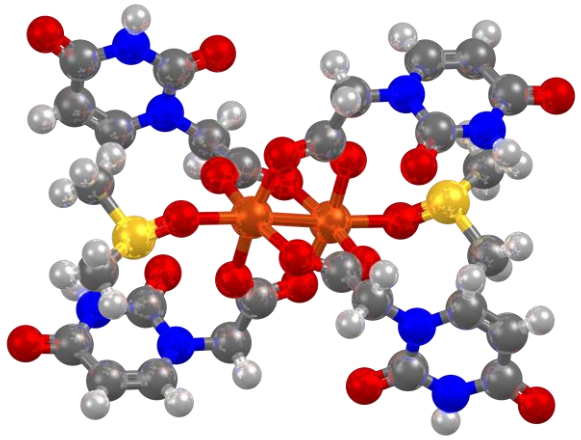
Analyse Hitlist

TUWMOP
TUWMOQ
TUWMUT
TUWMUU
TUWMUV
TUWMUW
TUWNAA
TUWNAB
TUWNAC
TUWNAD
TUWNEE
TUWNEF
TUWNEG
TUWNEH
TUWNII
TUWNIJ
TUWNIK
TUWNIL
TUWNOO
TUWNOQ
TUWNUU
TUWNUV
TUWNUW

<< >>

1097619 hits

100%



Could we use the CSD to explore the nature of copper carboxylate paddle structures and then go on to explore other metals?

Use-as-Query Options

Hydrogens:

Include hydrogen atoms

Chemical Units (molecules, ions, etc.):

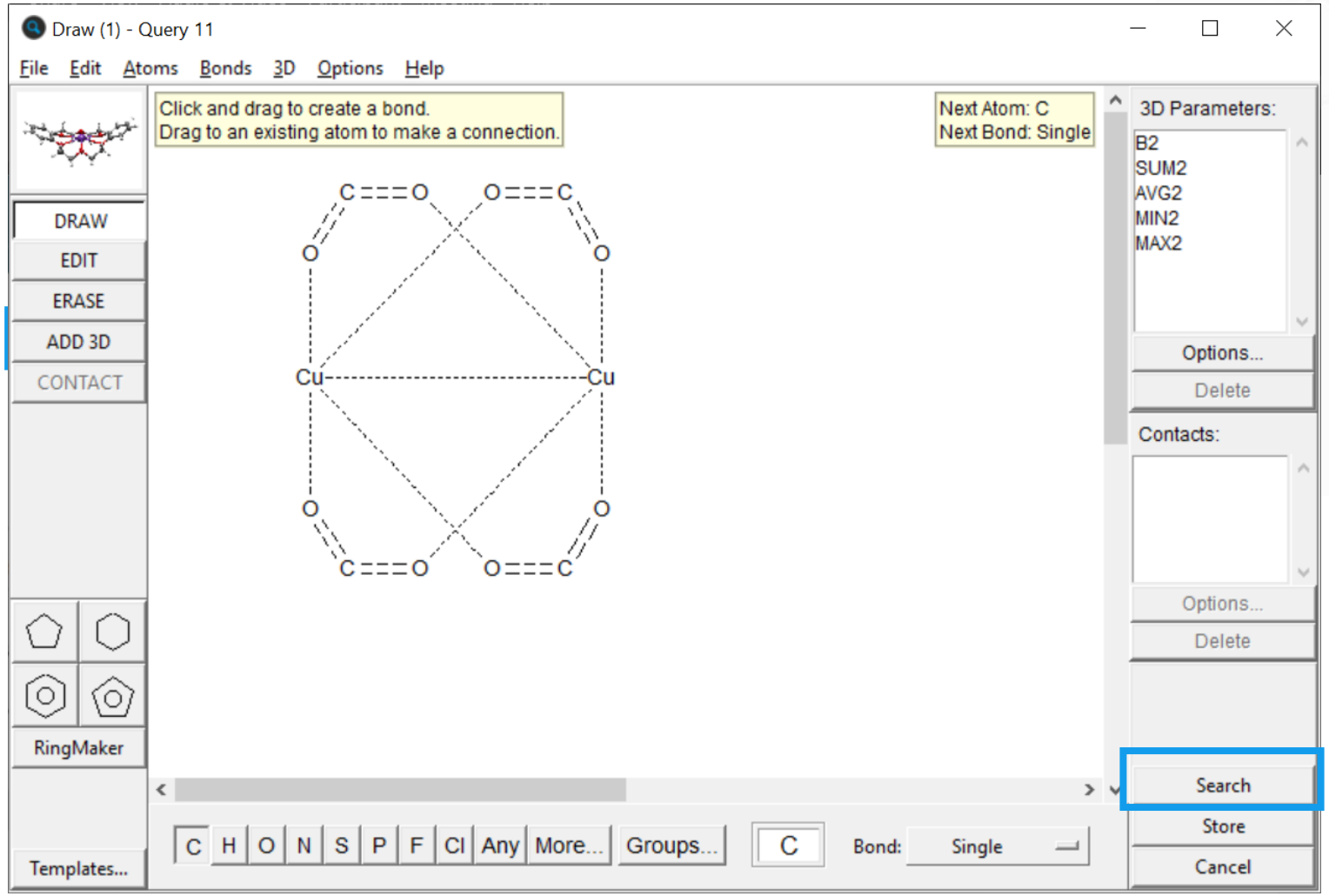
Biggest chemical unit only

Make each chemical unit a separate query

Include all chemical units in a single query

Cancel OK

3D searching



1. Draw your substructure
2. Click Add 3D
3. Select parts of the structure you are interested in
4. Define the parameters you want to analyse
5. Search or store your query
6. View the results
7. Analyse the hit list

Search highlighted in 2D

Parameters displayed

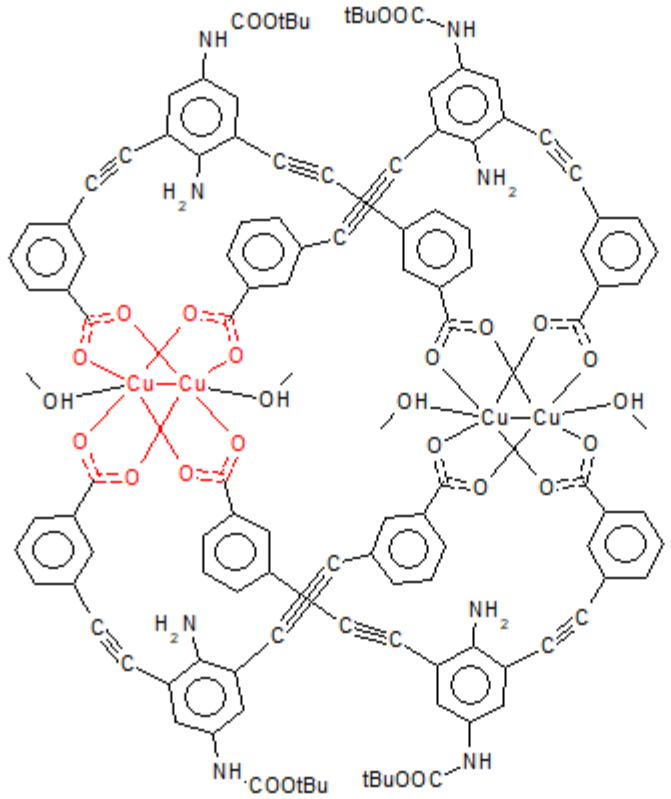
CCDC ConQuest (1) : search11 [Search]

File Edit Options View Databases Results Help

Build Queries | Combine Queries | Manage Hitlists | View Results

All Text
Author/Journal
Chemical
Crystal
Experimental
Diagram
3D Visualiser
CSD Internals
Search Overview

Refcode: JUZTOQ CSD version 5.42 updates (Feb 2021)



Parameters
B2
2.603
SUM2
2.603
AVG2
2.603
MIN2
2.603
MAX2
2.603

Multiple Hits: Show 1 of 2 | Show Parameters

Show terminal carbons | Use as Query... | Detach

JUZTOQ

Analyse Hitlist

✔ CUYYUT
✔ JUZTOQ
✔ JUZTUW
✔ LOYYED
✔ LUYVEJ
✔ OKUSAR
✔ PYCUAC
✔ UJOKUC
✔ UJUNOF
✔ ABUXU
✔ ACACC
✔ ACACC
✔ ACACC
✔ ACAPC
✔ ACGLC
✔ ACNICU
✔ ACURC
✔ ACURC
✔ ACURC
✔ ACURC
✔ AGUTA
✔ AHEDO
✔ AMACU

<< | 829 h

100%

Stop Search

Visualise Structures
Analyse Data
Select All
Deselect All
Invert Selection

Analyse in Mercury

Select the items you wish to include from the choices below

File type: Mercury data file (.c2m)

☑ Include Defined Parameters

Crystal data

☐ R-factor ☐ Space Gp. Symbol ☐ Space Gp. Number ☐ No. of Coordinates
☐ Z Value ☐ Z Prime ☐ Study Temp. ☐ Calc. Density

Cell data

☐ a ☐ b ☐ c
☐ Alpha ☐ Beta ☐ Gamma ☐ Cell Volume
☐ Reduced Cell a ☐ Reduced Cell b ☐ Reduced Cell c ☐ Reduced Cell Volume
☐ Reduced Cell Alpha ☐ Reduced Cell Beta ☐ Reduced Cell Gamma

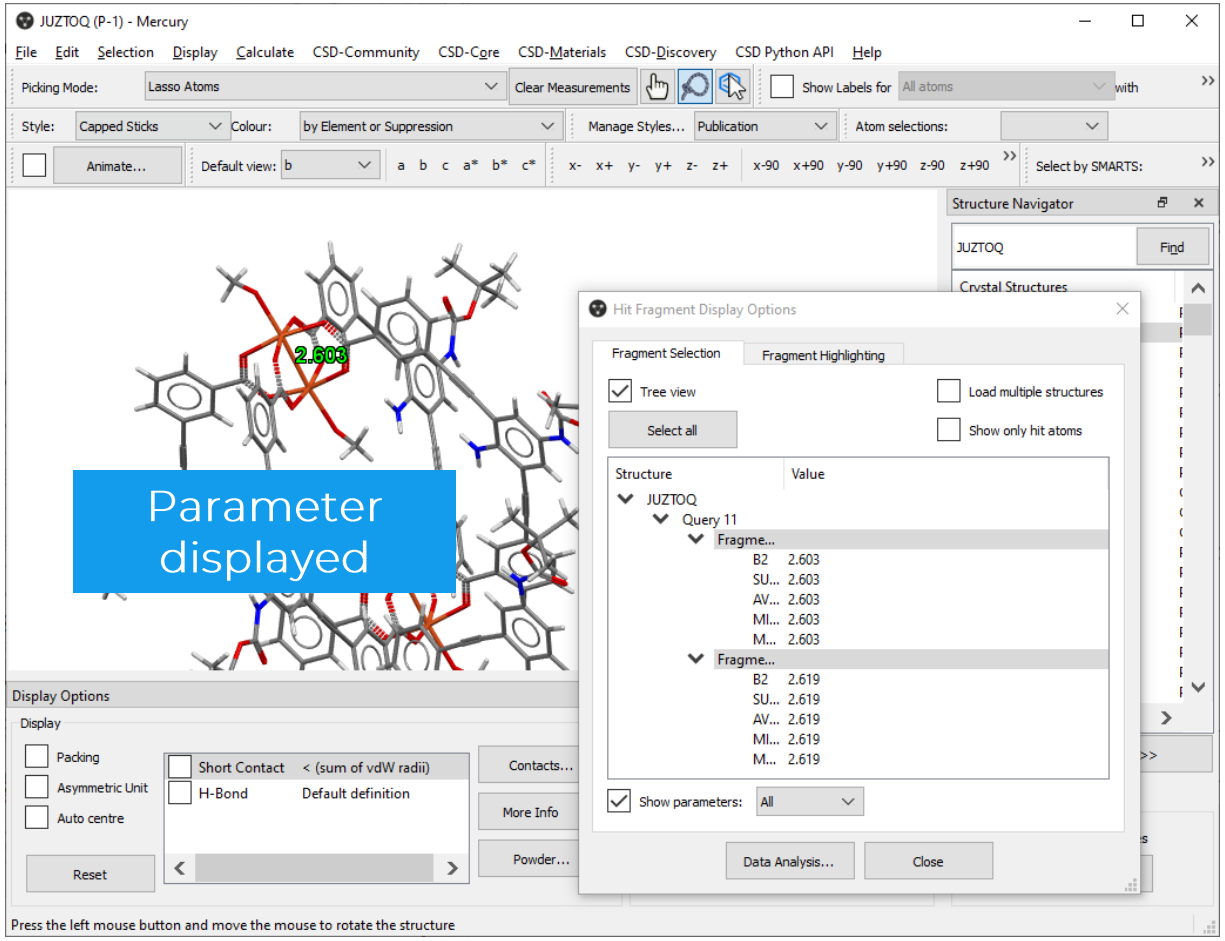
Other

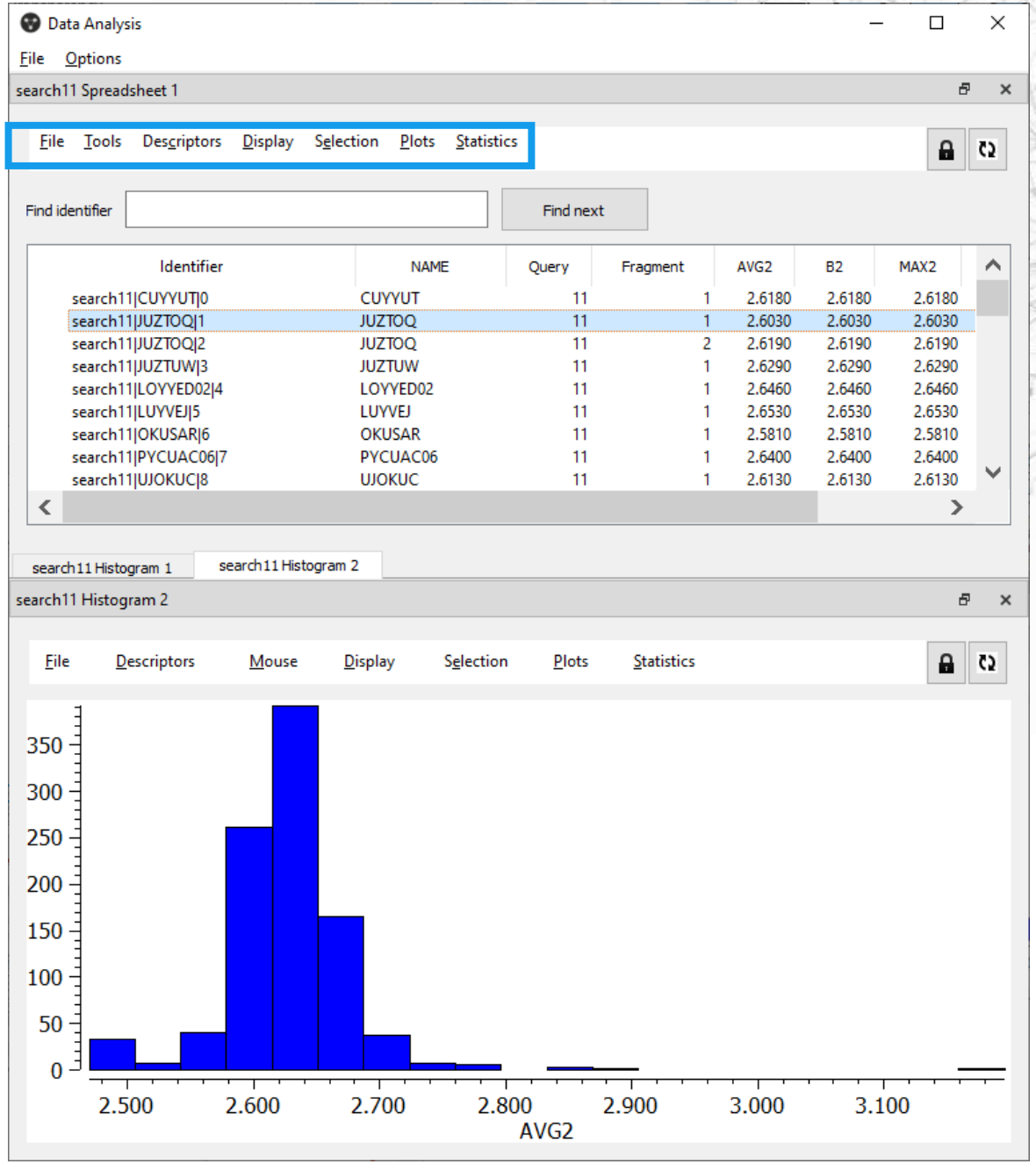
☐ Publication Year ☐ Unique Chemical Units ☐ Multiplier Sum ☐ Compound Name

0% | Reset | Cancel | Analyse in Mercury

CCDC

Data analysis

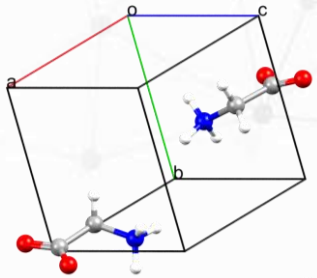


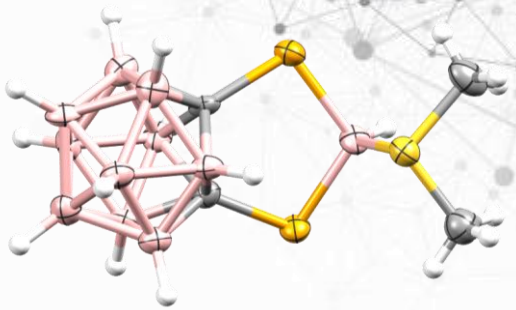


CSD Subsets: Helping find one in a million+

Why provide subsets?

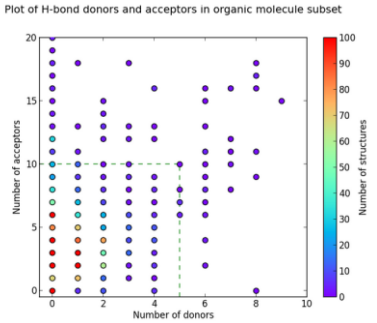
- The CSD contains a huge range of compounds; subsets allow easy access to the most relevant structures of interest
- Metal-organics in particular can be hard to name and draw consistently
- Gives CSD users the benefit of analysis from in-house and external experts
- Convenient starting point for analysis using CSD or 3rd-party tools

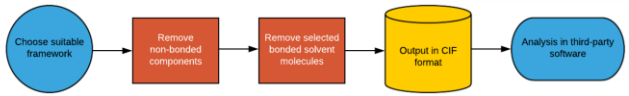




DRUGBANK

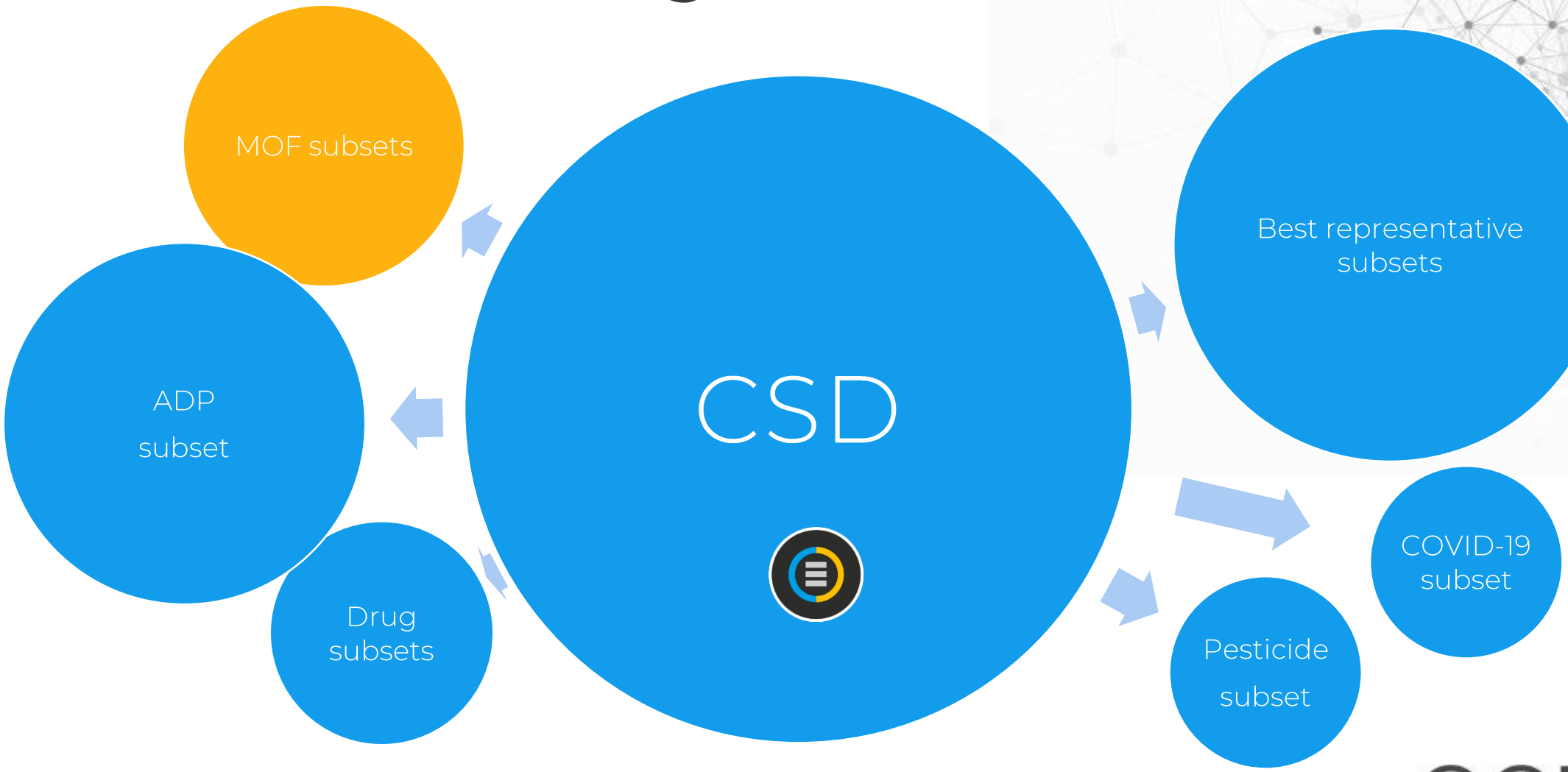
PPDB





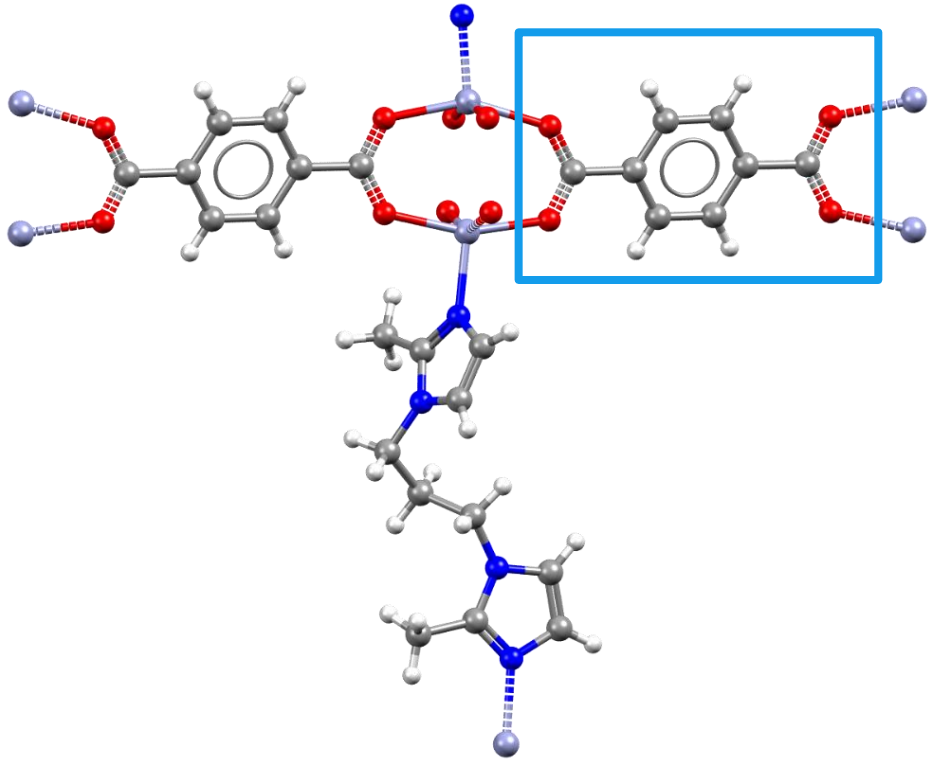
CCDC

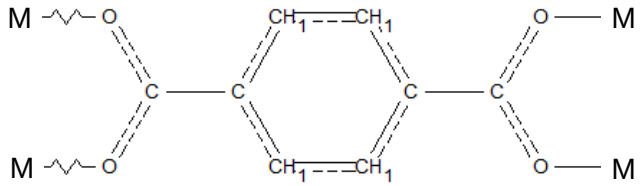
CSD subsets – targeted datasets



CCDC

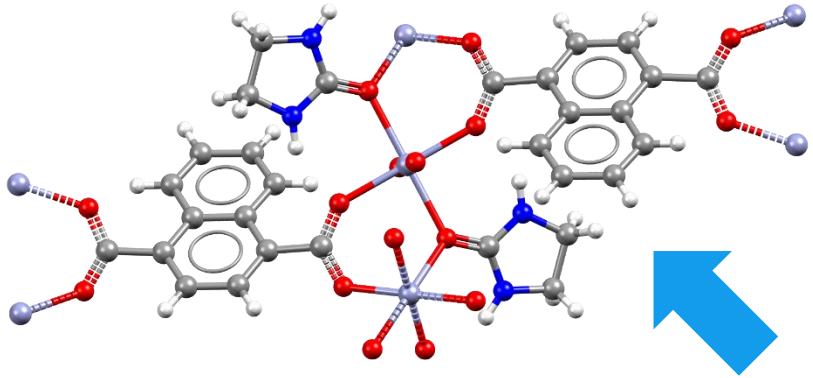
CSD MOF subset creation

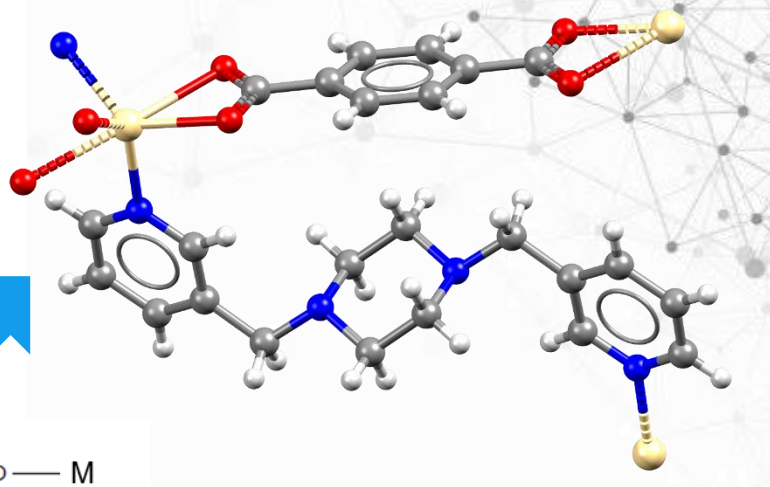


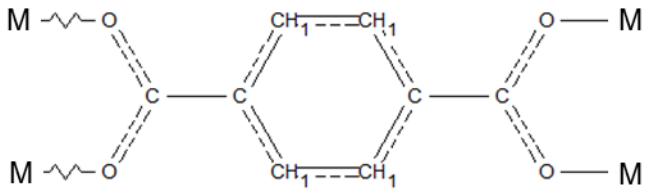


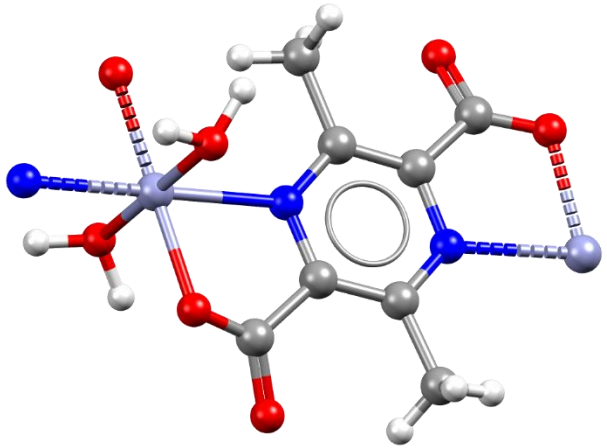
CCDC

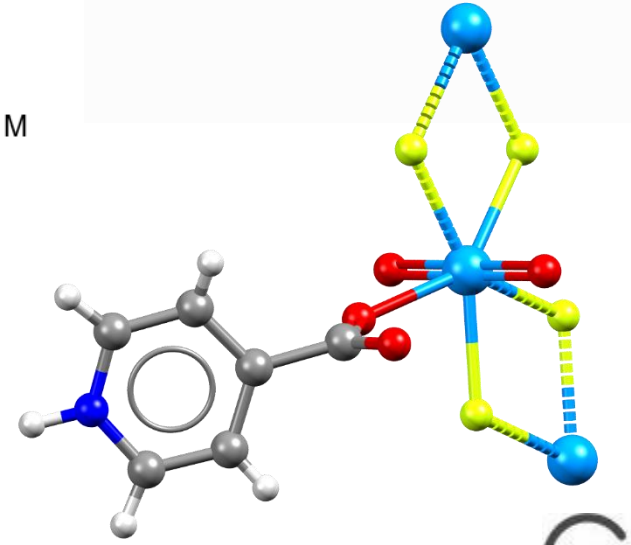
CSD MOF subset creation





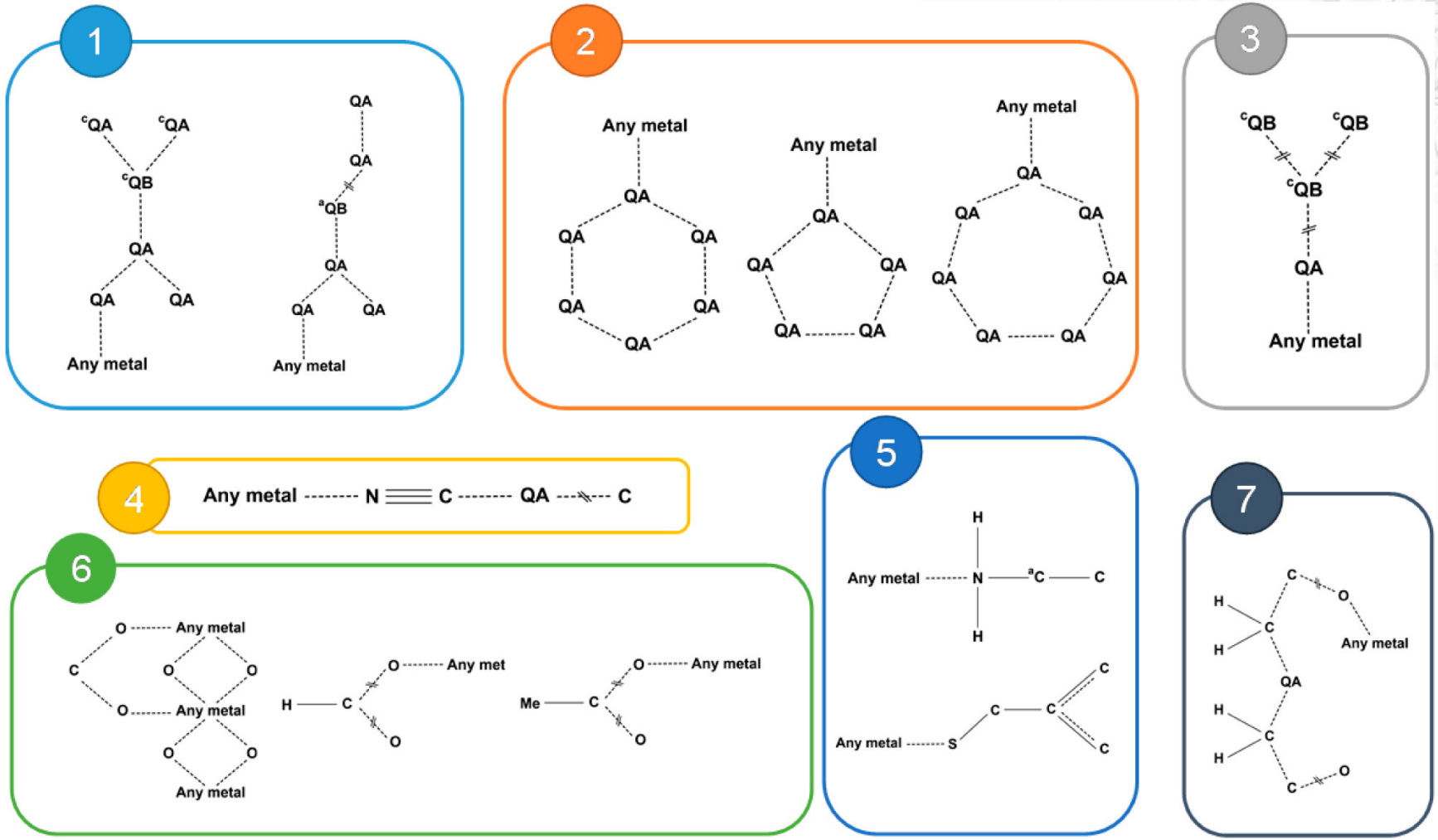




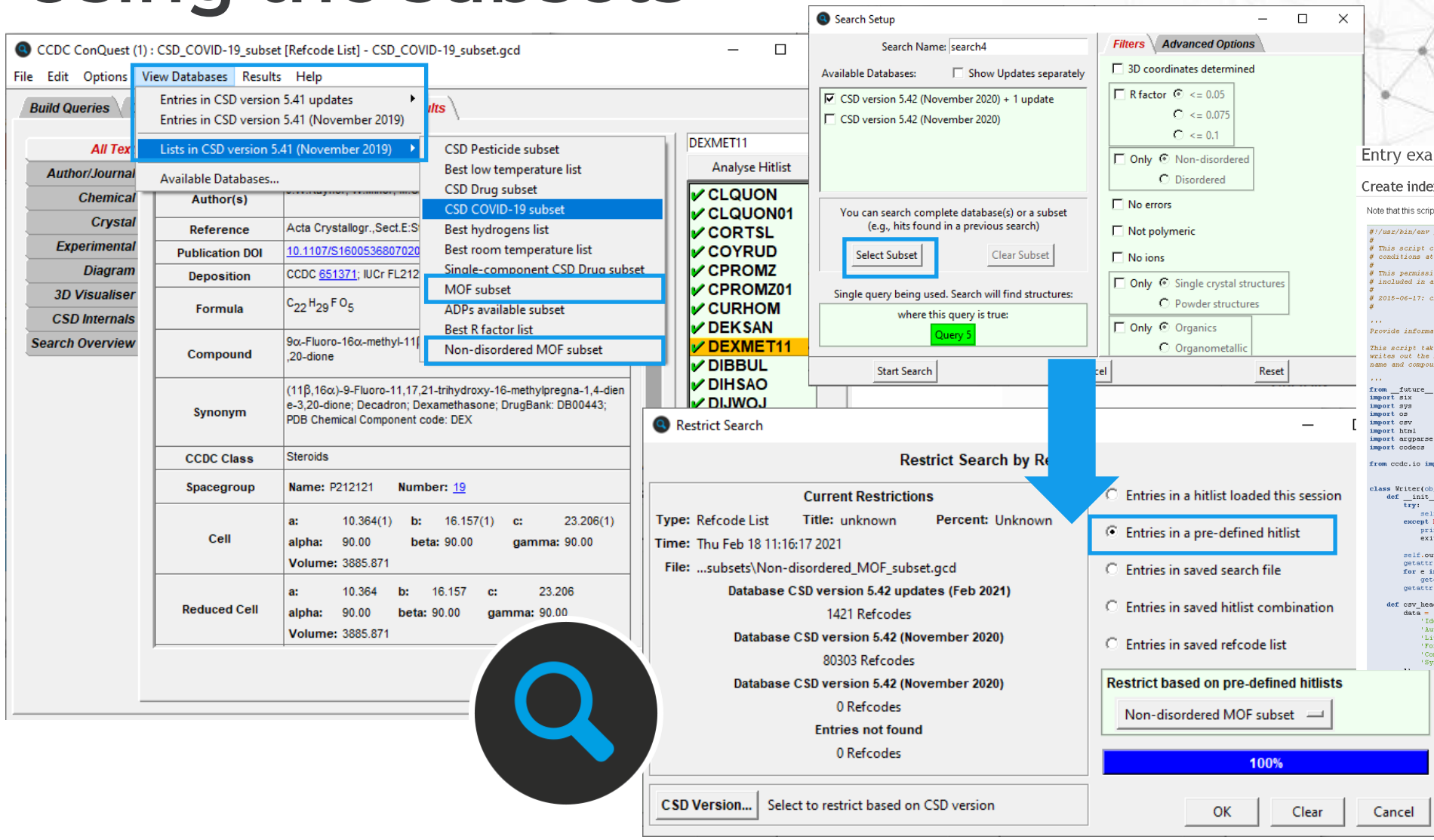


CCDC

CSD MOF subset creation – Draw Queries



Using the subsets

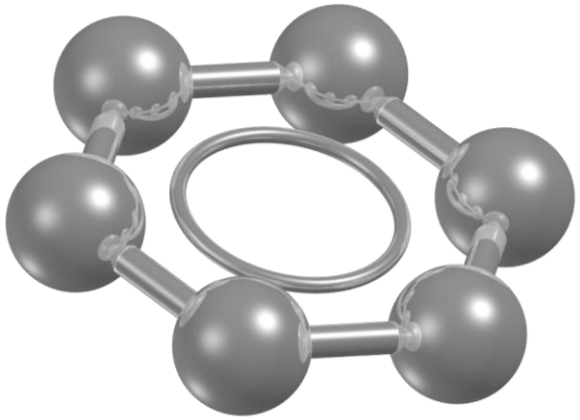


CCDC

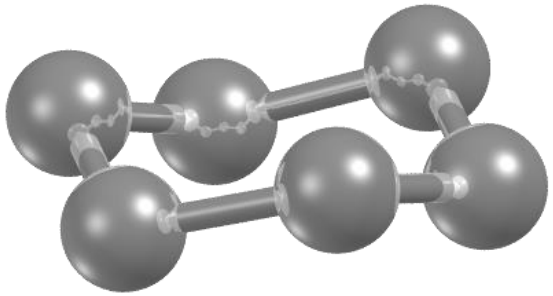
Explore More: More advanced uses

- Not enough time to explore all the insights that you can get from metal-organic structures in the CSD
- But here are a selection of more advanced tips and tricks and case studies

Exploring chemical concepts

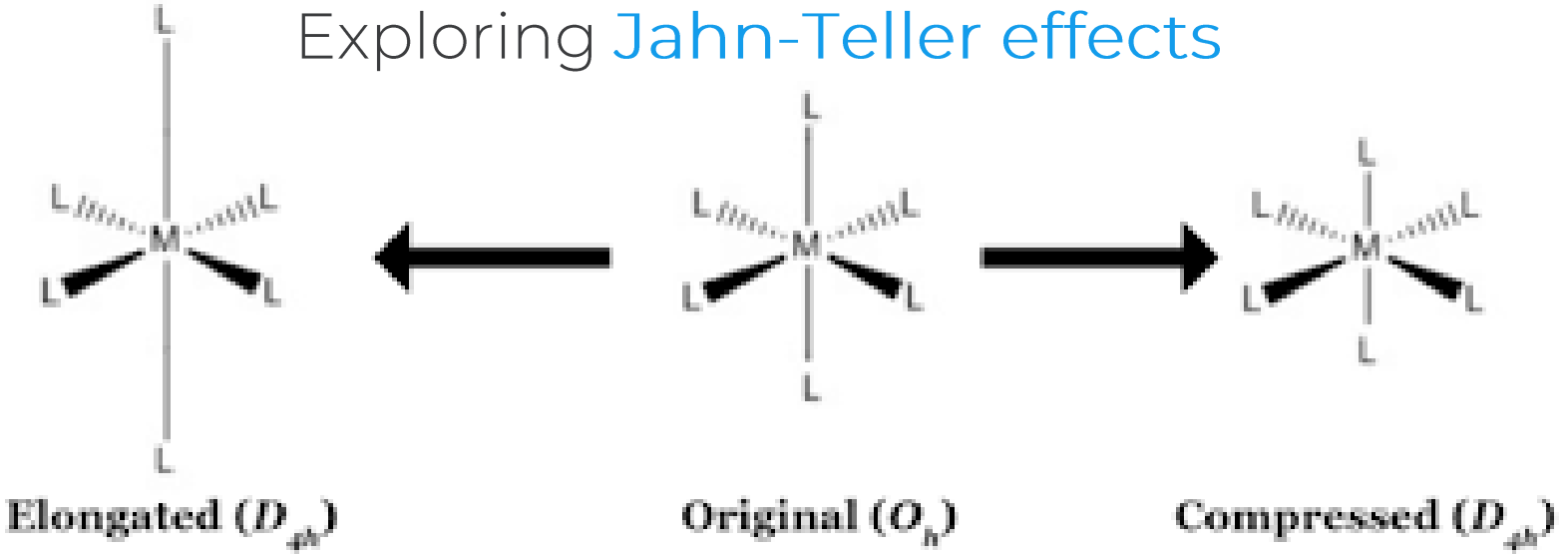


BENZEN07



FUHVEM

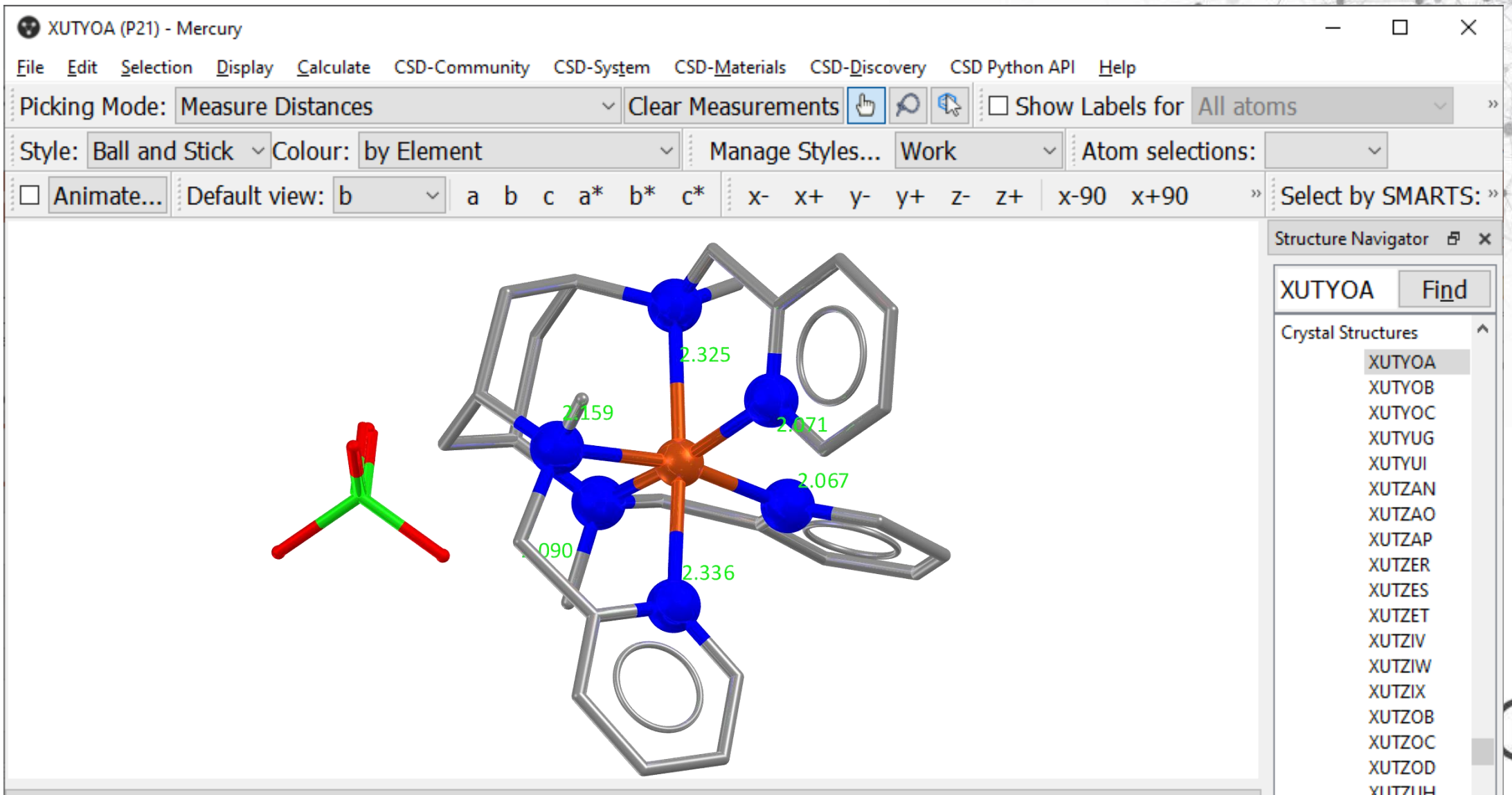
Exploring Jahn-Teller effects

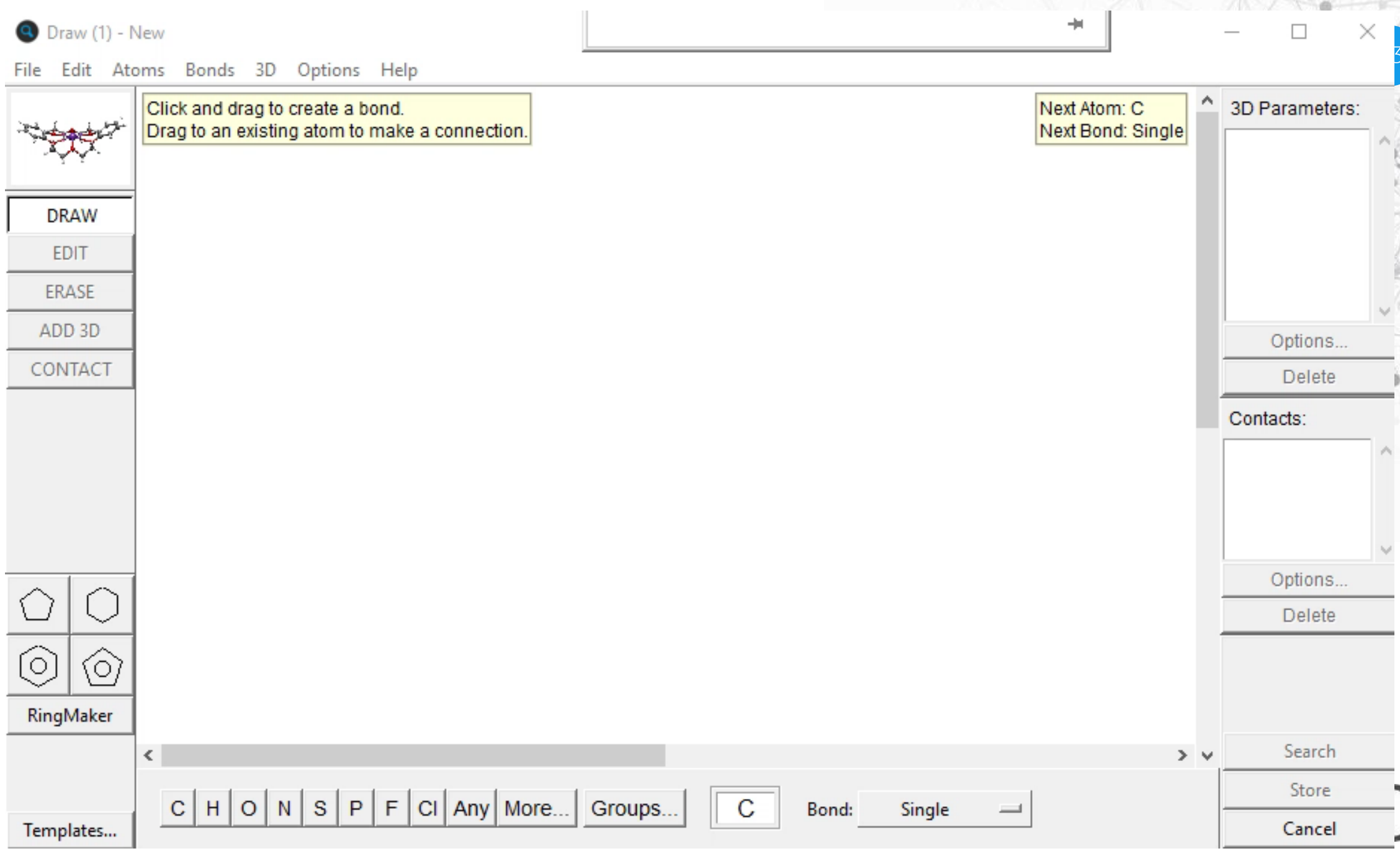


The Jahn–Teller effect is most often encountered in octahedral complexes of the transition metals.[3] The phenomenon is very common in six-coordinate copper(II) complexes.

CCDC

Using Mercury to see chemical concepts




Draw (1) - New
File
Edit
Atoms
Bonds
3D
Options
Help
Click and drag to create a bond.
Drag to an existing atom to make a connection.
Next Atom: C
Next Bond: Single
DRAW
EDIT
ERASE
ADD 3D
CONTACT
RingMaker
Templates...
3D Parameters:
Options...
Delete
Contacts:
Options...
Delete
Search
Store
Cancel
C
H
O
N
S
P
F
Cl
Any
More...
Groups...
C
Bond:
Single

CCDC ConQuest (1) : search2 [Search]

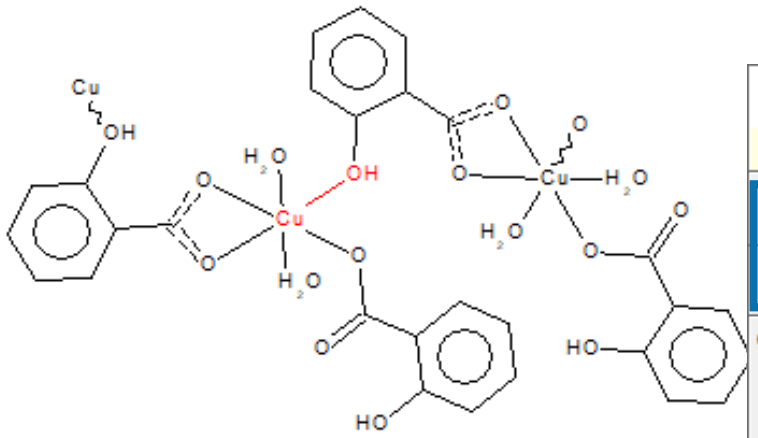
File Edit Options View Databases Results Help

Build Queries | Combine Queries | Manage Hitlists | View Results

All Text
Author/Journal
Chemical
Crystal
Experimental
Diagram
3D Visualiser
CSD Internals
Search Overview

Refcode: ASALCU03 CSD version 5.41 updates (Mar 2020)

Parameters
DIST1
2.237



Multiple Hits: Show 1 of 11

Show terminal carbons

Use as Quer

ASALCU03

Analyse Hitlist

- ✔ ASALCU03
- ✔ ASALCU05
- ✔ BEYRAY07
- ✔ BOVNEH01
- ✔ BOWNOT
- ✔ BOZXEW
- ✔ BUBMOD
- ✔ BUCBIN
- ✔ BUFZUA

Visualise Structures
Analyse Data
Select All
Deselect All
Invert Selection

Analyse in Mercury

Select the items you wish to include from the choices below

File type: Mercury data file (.c2m)

☑ **Include Defined Parameters**

Crystal data

☐ R-factor ☐ Space Gp. Symbol ☐ Space Gp. Number ☐ No. of Coordinates
☐ Z Value ☐ Z Prime ☐ Study Temp. ☐ Calc. Density

Cell data

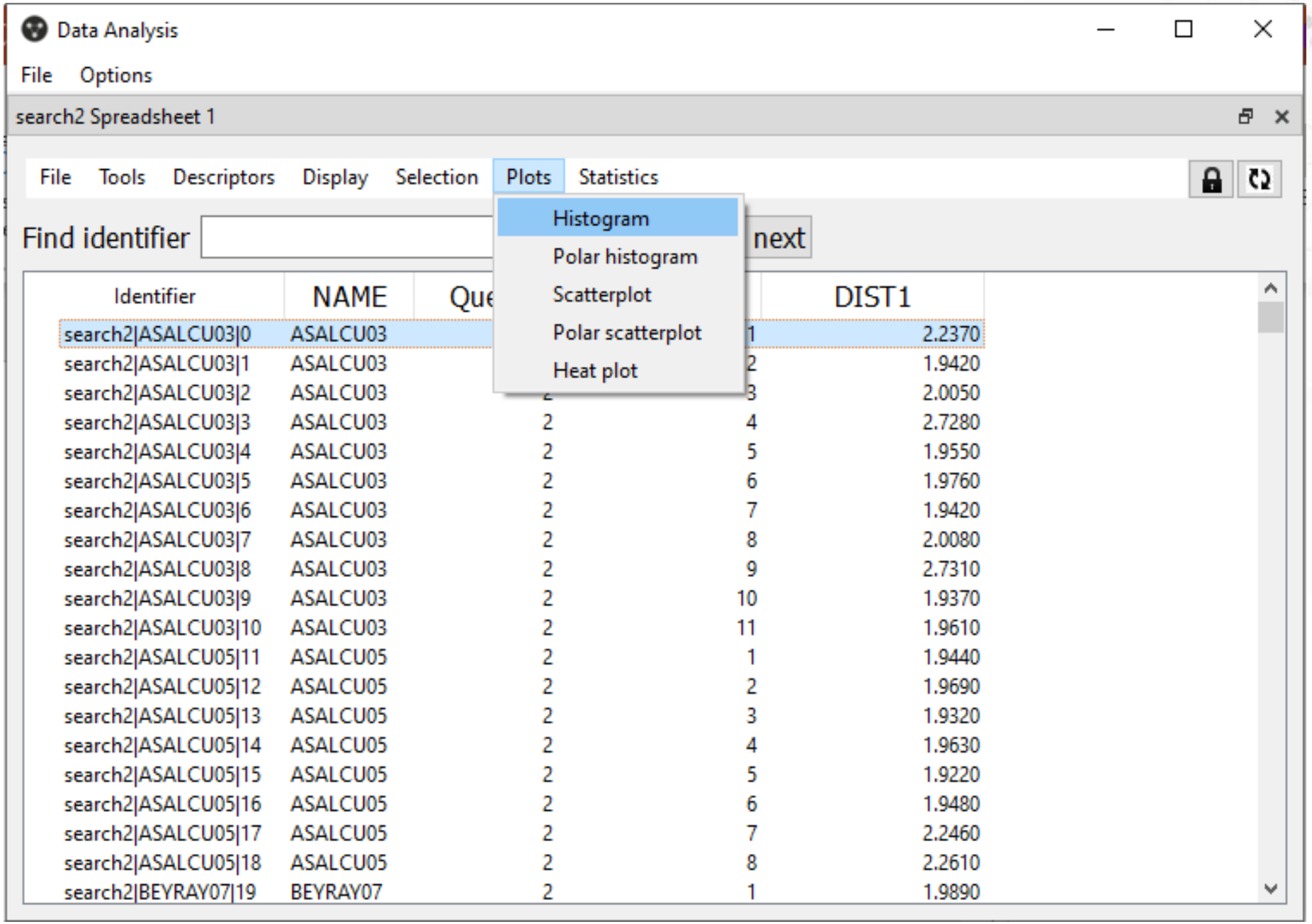
☐ a ☐ b ☐ c
☐ Alpha ☐ Beta ☐ Gamma ☐ Cell Volume
☐ Reduced Cell a ☐ Reduced Cell b ☐ Reduced Cell c ☐ Reduced Cell Volume
☐ Reduced Cell Alpha ☐ Reduced Cell Beta ☐ Reduced Cell Gamma

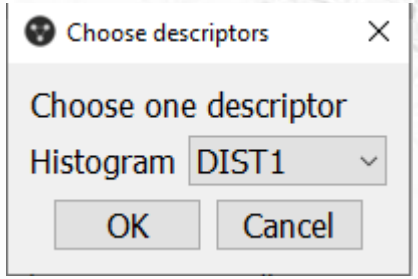
Other

☐ Publication Year ☐ Unique Chemical Units ☐ Multiplier Sum ☐ Compound Name

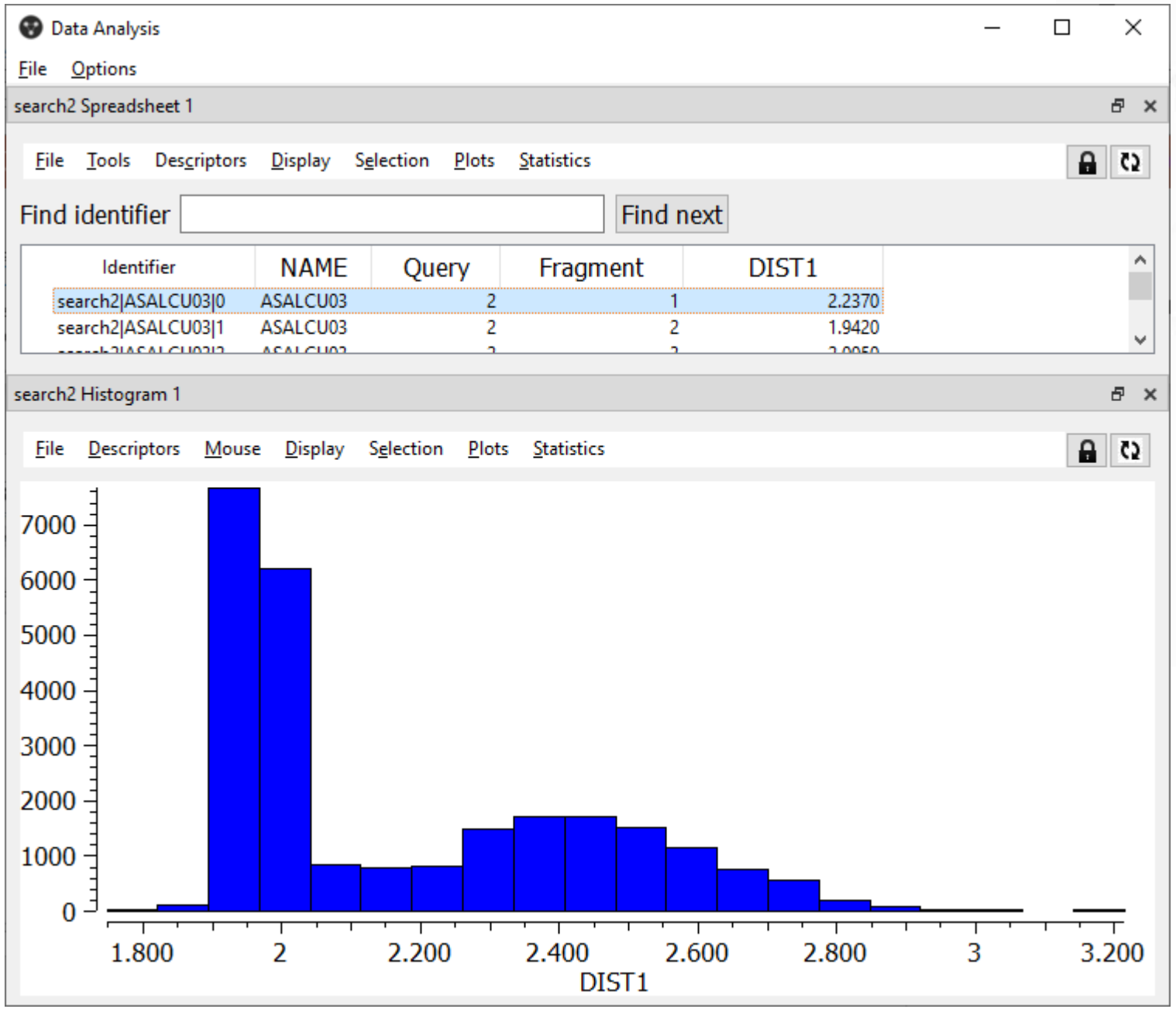
0%

Reset Cancel Analyse in Mercury


Data Analysis
File Options
search2 Spreadsheet 1
File Tools Descriptors Display Selection Plots Statistics
Histogram
Polar histogram
Scatterplot
Polar scatterplot
Heat plot
Find identifier
next
Identifier NAME Que DIST1
search2|ASALCU03|0 ASALCU03 1 2.2370
search2|ASALCU03|1 ASALCU03 2 1.9420
search2|ASALCU03|2 ASALCU03 3 2.0050
search2|ASALCU03|3 ASALCU03 2 4 2.7280
search2|ASALCU03|4 ASALCU03 2 5 1.9550
search2|ASALCU03|5 ASALCU03 2 6 1.9760
search2|ASALCU03|6 ASALCU03 2 7 1.9420
search2|ASALCU03|7 ASALCU03 2 8 2.0080
search2|ASALCU03|8 ASALCU03 2 9 2.7310
search2|ASALCU03|9 ASALCU03 2 10 1.9370
search2|ASALCU03|10 ASALCU03 2 11 1.9610
search2|ASALCU05|11 ASALCU05 2 1 1.9440
search2|ASALCU05|12 ASALCU05 2 2 1.9690
search2|ASALCU05|13 ASALCU05 2 3 1.9320
search2|ASALCU05|14 ASALCU05 2 4 1.9630
search2|ASALCU05|15 ASALCU05 2 5 1.9220
search2|ASALCU05|16 ASALCU05 2 6 1.9480
search2|ASALCU05|17 ASALCU05 2 7 2.2460
search2|ASALCU05|18 ASALCU05 2 8 2.2610
search2|BEYRAY07|19 BEYRAY07 2 1 1.9890


Choose descriptors
Choose one descriptor
Histogram DIST1
OK
Cancel

CCDC

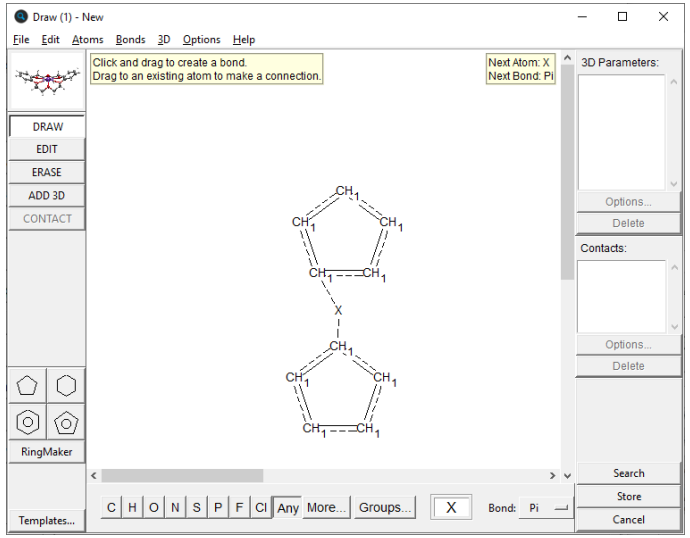


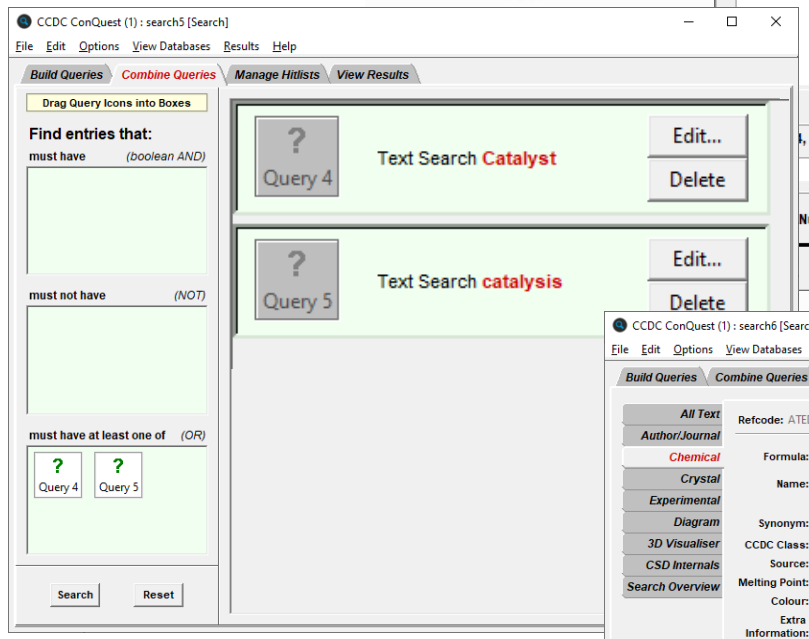
You could use this method to see what other metal atoms exhibit Jahn-Teller distortions

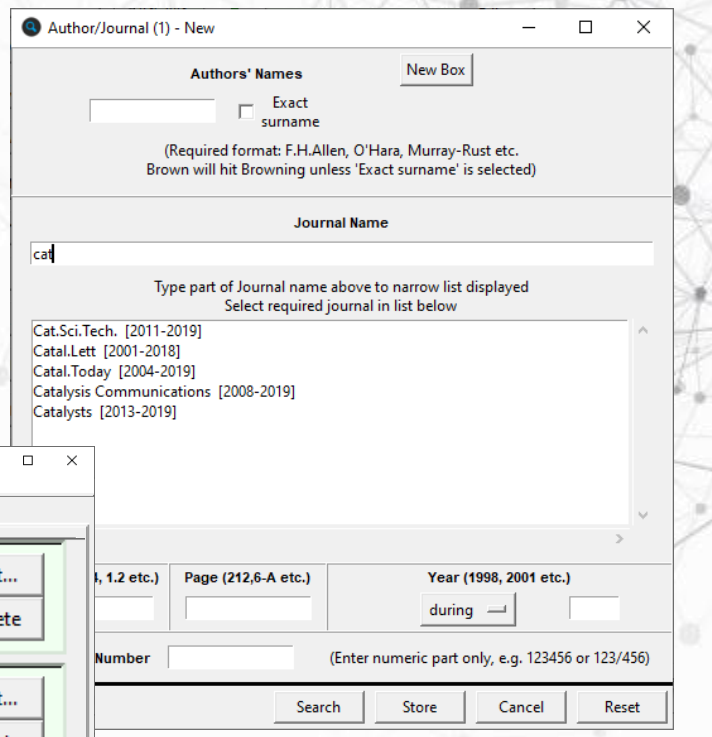
CCDC

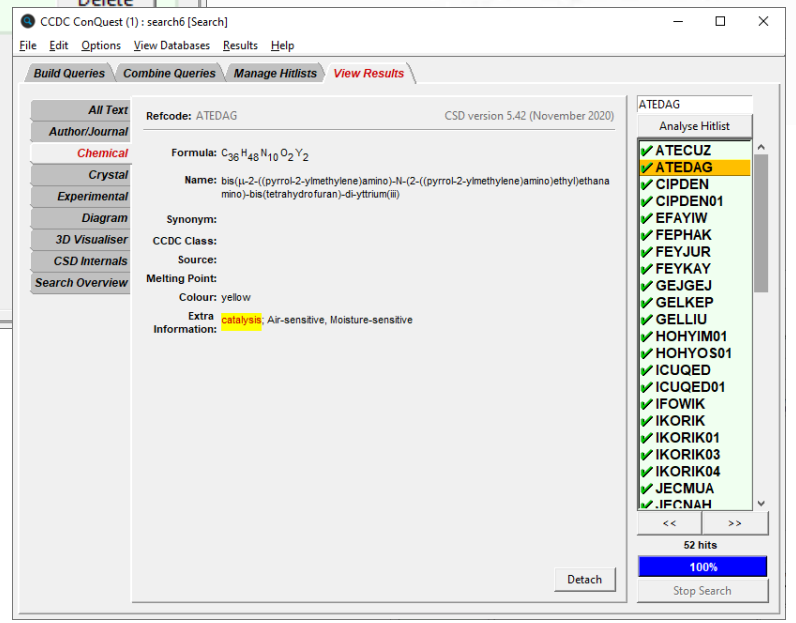
Catalysts

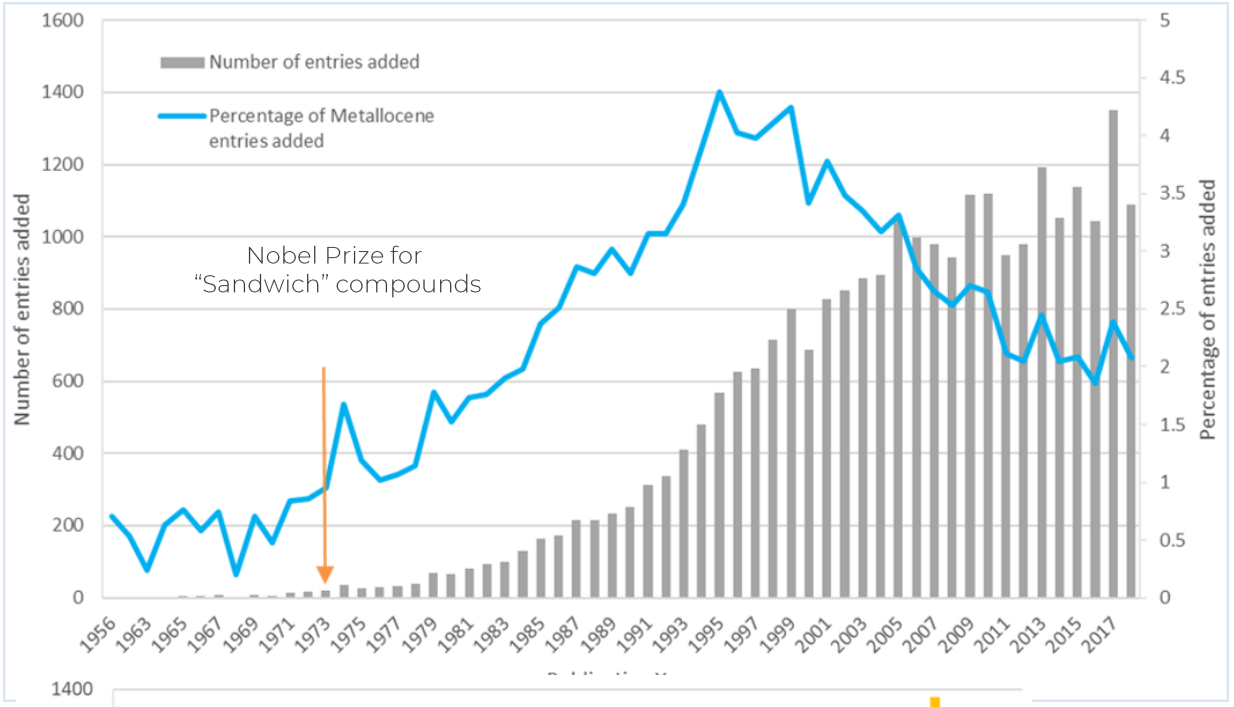
- Searching for Journal
  - Organometallics
  - Other catalysis based journals
- All-text search
- Similarity search on results
- Structure/Draw searches

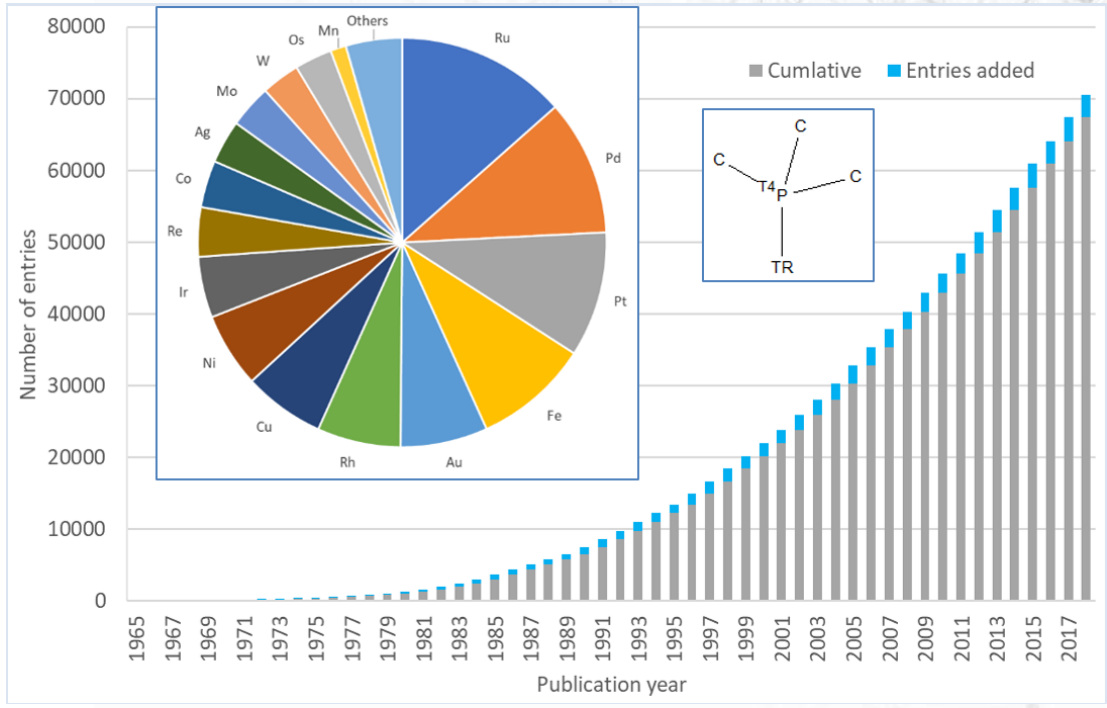


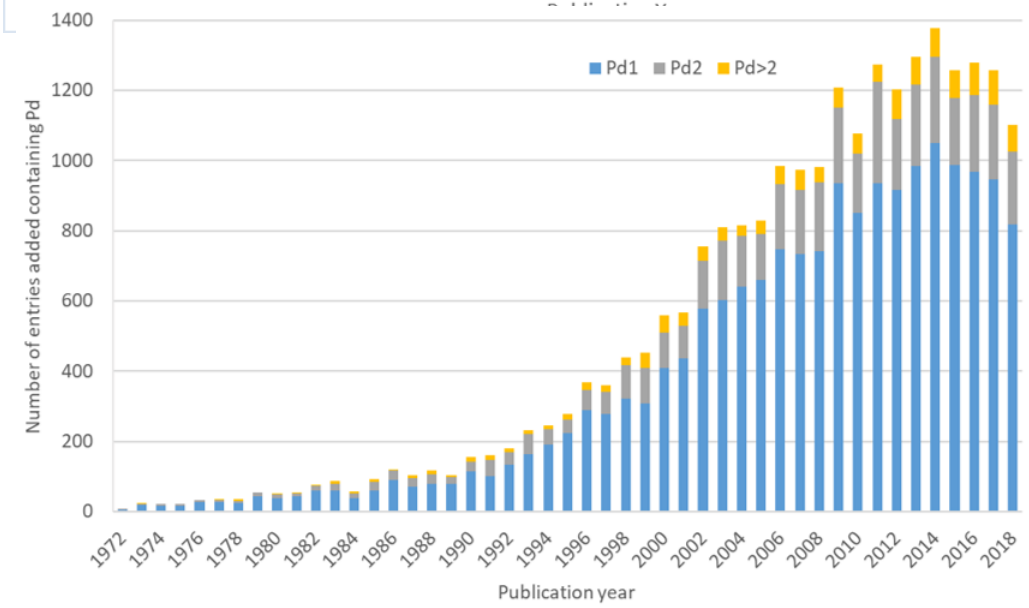


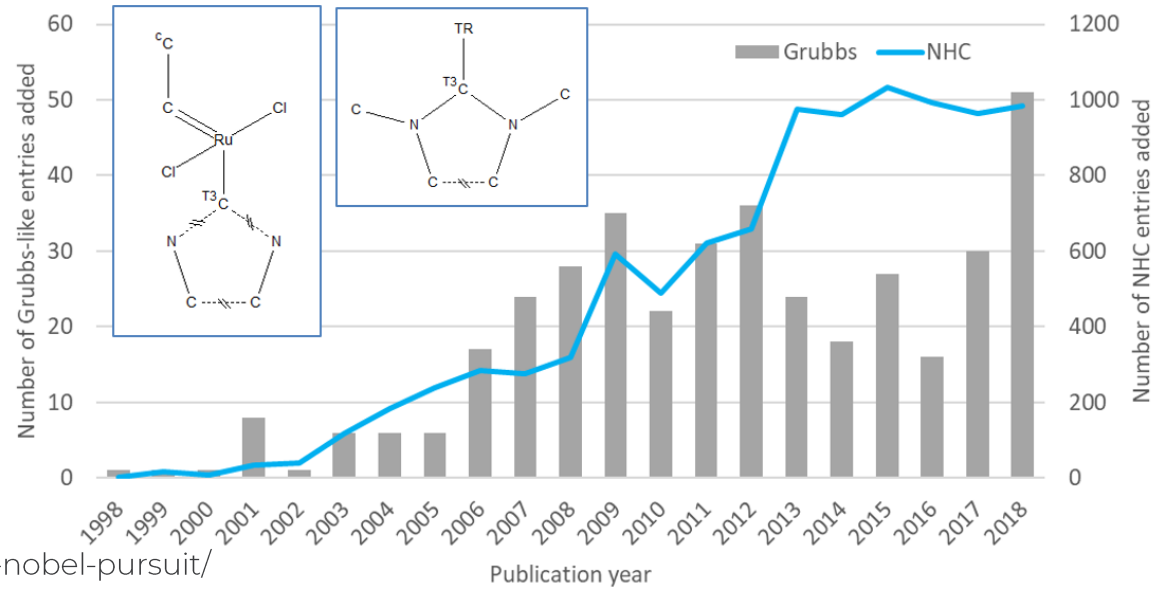












https://www.ccdc.cam.ac.uk/Community/blog/2020-02-20-better-catalysts-a-nobel-pursuit/