

Searching and analysing metal-organic structures

IUCr25 - CSD and ICSD Workshop – Session 2



CCDC and FIZ Karlsruhe

12th August 2021



CCDC
advancing structural science

Learning outcomes

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.

Agenda

- 10 mins: Introduction
- 15 mins: Show One: ConQuest overview
- 20 mins: *Show One*: Metal-organic demo
- 25 mins: *Try One*: Hands-on examples
- 10 mins: *Explore More*: Advanced functionality for metal-organics
- 5 mins: *Explore More*: Quiz
- 5 mins: Wrap up and summary

Inside the Cambridge Structural Database

The CSD is a database of all the published organic and metal-organic experimental crystal structures

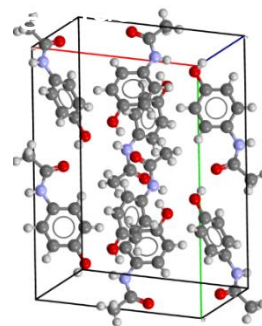
Organic
43%

Metal-Organic
57%

At least one transition metal,
lanthanide, actinide or any of Al,
Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po

Organic

- Drugs
- Agrochemicals
- Pigments
- Explosives
- Protein ligands



Additional data

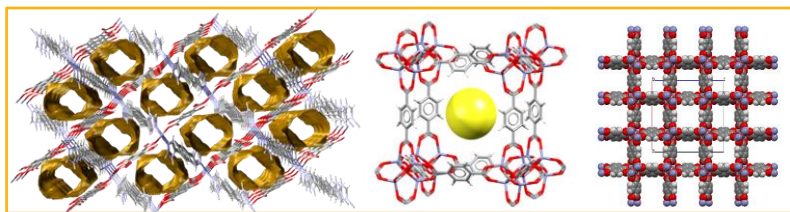
- 11,525 polymorph families
- 171,683 melting points
- 909,992 crystal colours
- 778,663 crystal shapes
- 24,916 bioactivity details
- 11,387 natural source data
- > 250,000 oxidation states

Not Polymeric
89%

Polymeric: 11%

Metal-Organic

- Metal Organic Frameworks
- Models for new catalysts
- Porous frameworks for gas storage
- Fundamental chemical bonding

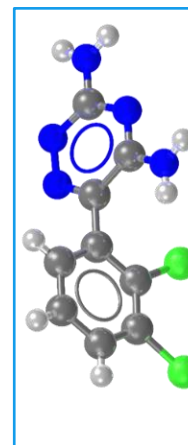


Single
Component
56%

Multi
Component
44%

Links and subsets

- DrugBank
- Druglike
- MOFs
- PDB ligands
- PubChem
- ChemSpider
- Pesticide PDB

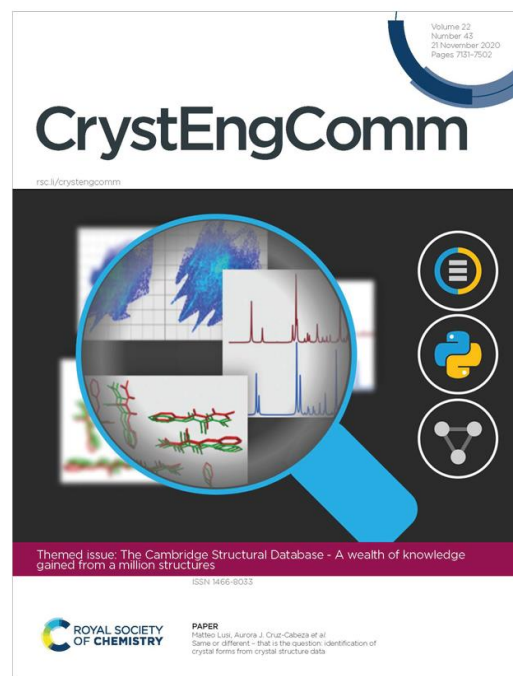


Supporting research globally

2020 CSD Special issue

- 33 articles from industry and academia

**1 million structures.
stories.
possibilities.**

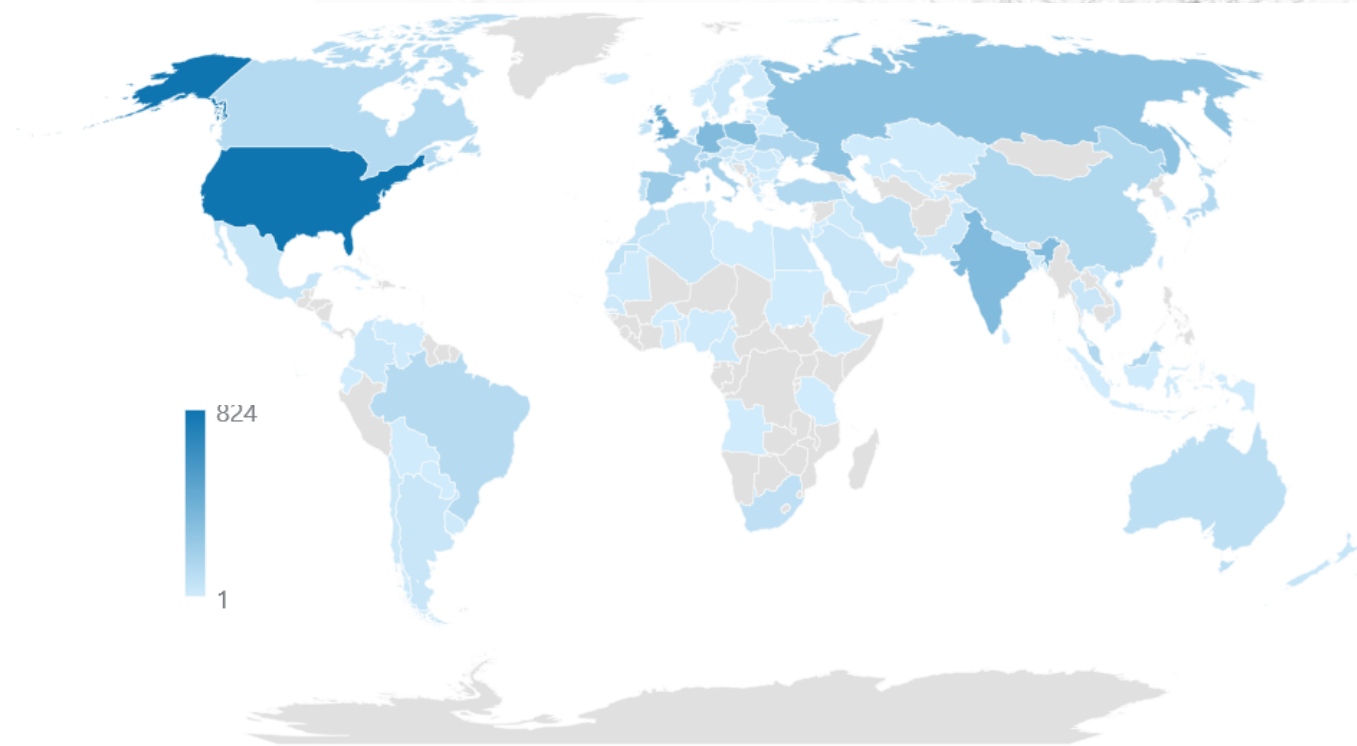


Cambridge Structural
Database, CSD



CCDC
advancing structural science

Publications on the CSD



Powered by Bing
© Australian Bureau of Statistics, GeoNames, Microsoft, Navinfo, OpenStreetMap, TomTom, Wikipedia

The CSD - A wealth of knowledge gained from a million structures. *CrystEngComm*, 2020,22, 7131-77502

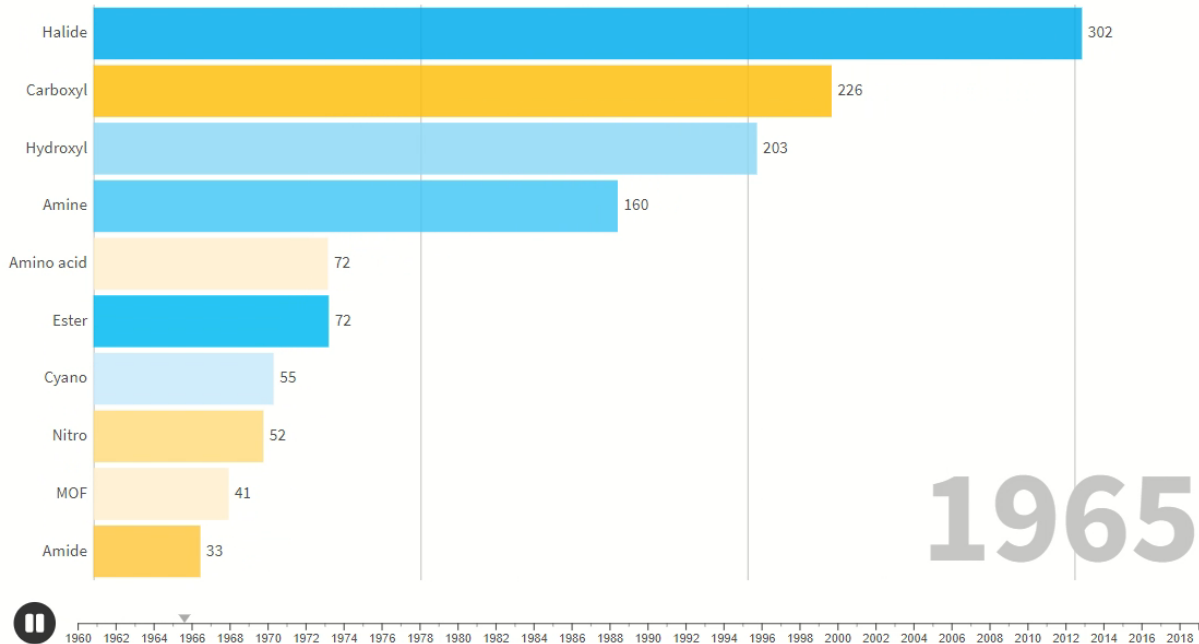
The CSD. C.R.Groom, I.J.Bruno, M.P.Lightfoot and S.C.Ward, *Acta Cryst.* (2016). B72, 171-179 DOI: 10.1107/S2052520616003954

CCDC

Exploring the CSD

Chemistry in the CSD

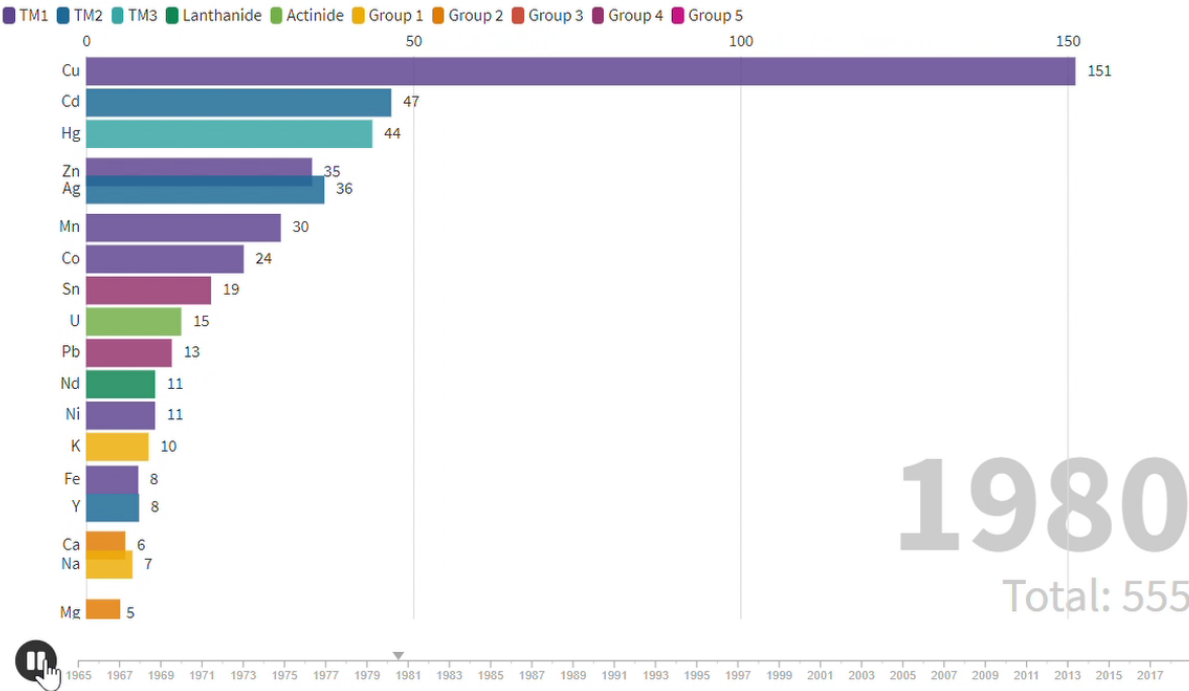
Number of structures containing certain chemical groups



1965

Metals in the CSD MOF subset

The number of structures containing a particular metal atom



1980

Total: 555

What is a MOF?

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

IUPAC Recommendations 2013¹

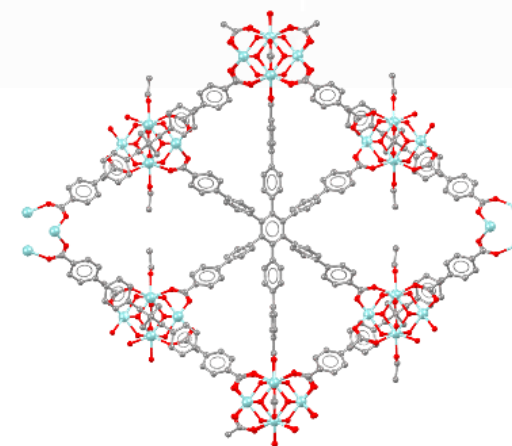
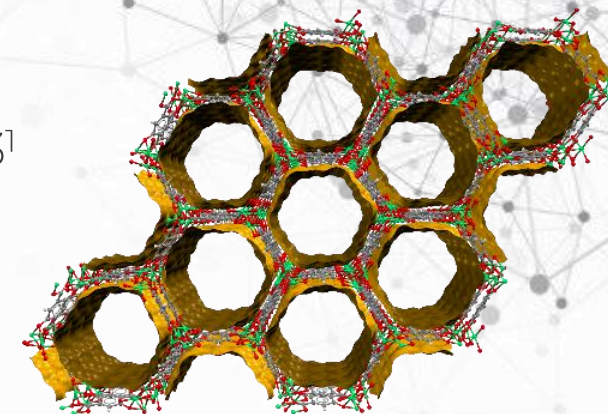
Interim 2012 report² noted:

- 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 Frameworks³

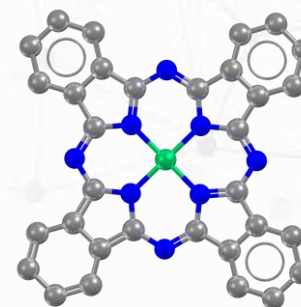
1. S. R. Batten et al. *Pure Appl. Chem.*, 2013, 85, 1715–1724 10.1351/PAC-REC-12-11-20; 2. S.R.Batten et al. *CrystEngComm*, 2012, 14, 3001-3004 10.1039/C2CE06488J; 3. <http://www.iza-online.org/MOF/>



CCDC

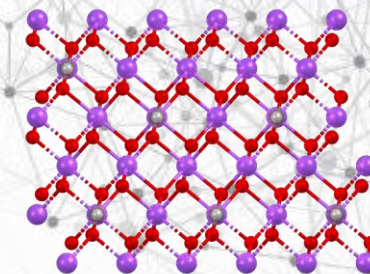
MOF and metal-organic trivia

- 1st metal-organic structure with coords was published in 1938 by Monteath Robertson
- 1st MOF in our MOF subset was published in 1940
- Metal-Organic entries now outweigh organic entries in the CSD
- >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!



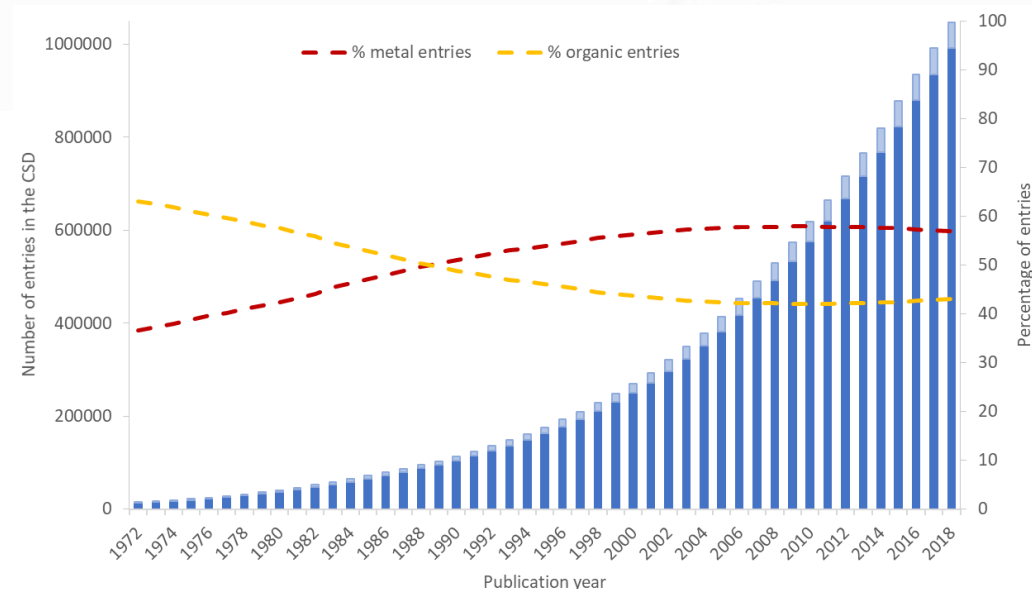
NIPHTC

1st metal-organic structure with 3D coords



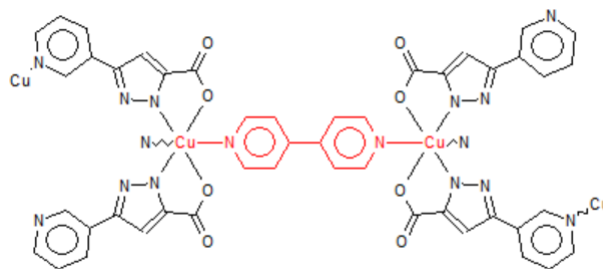
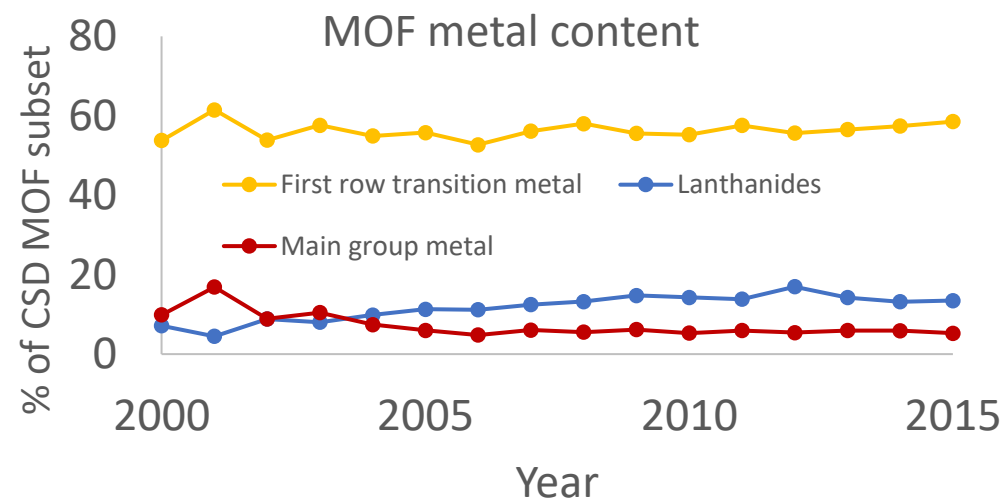
NAFORM

1st MOF in CSD subset

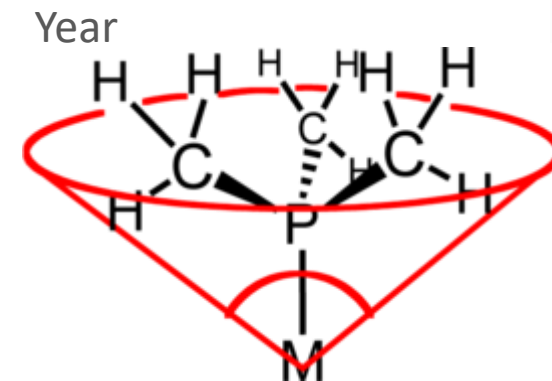


Using the CSD to explore metal-coordinations

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



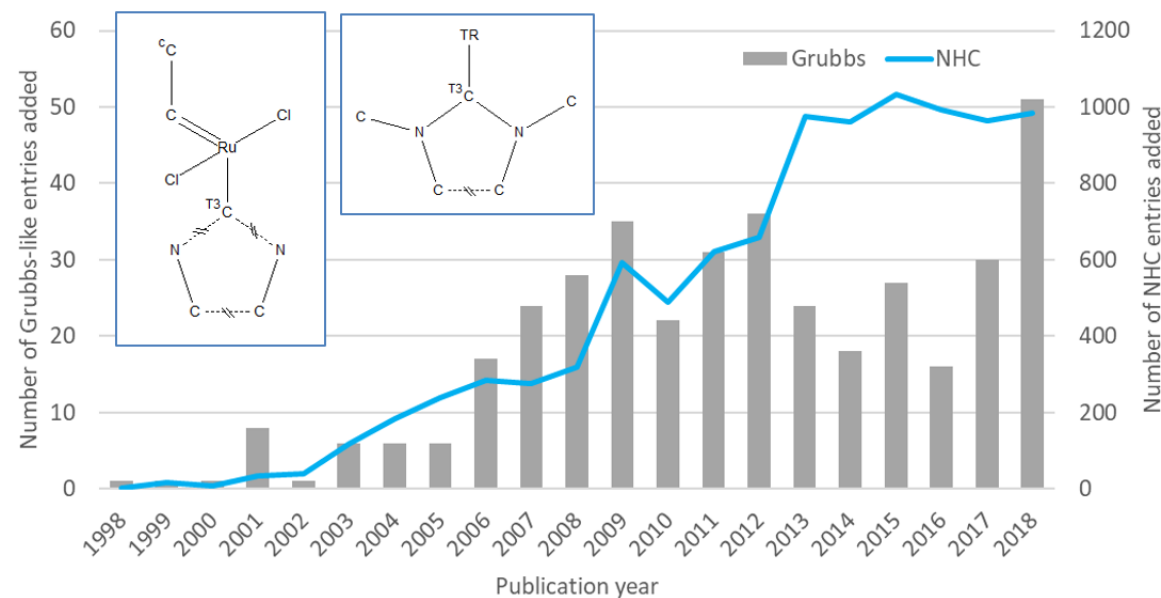
4,4-bipyridine one of the more frequent ligands in the CSD with >6,000 entries
~10% mono-coordinated, 90% di-coordinated



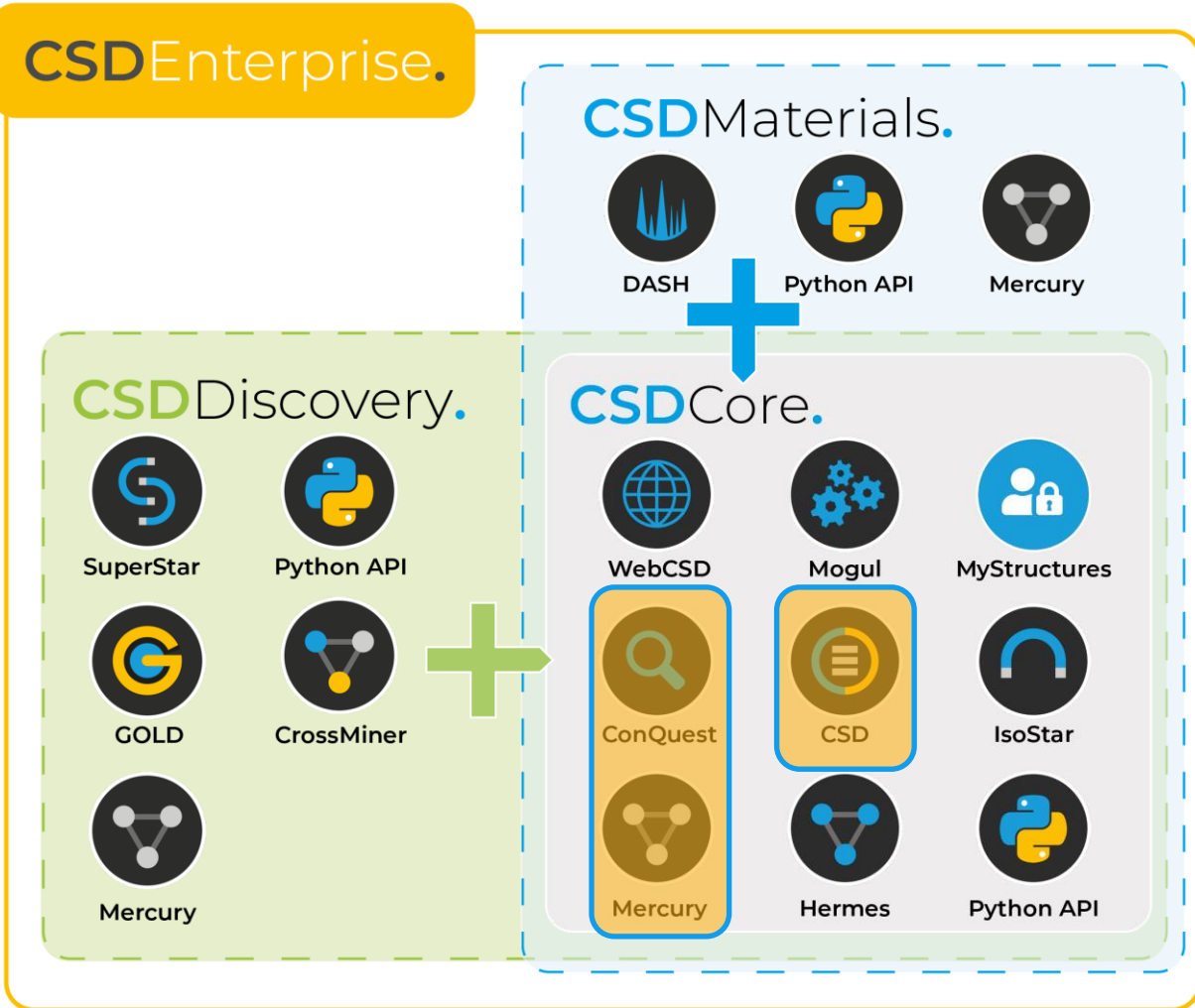
CCDC

Industrial applications

- Energetic materials
- Ferroelectricity
- Magnetic
- Catalysts
- Gas storage and separation
- Drugs

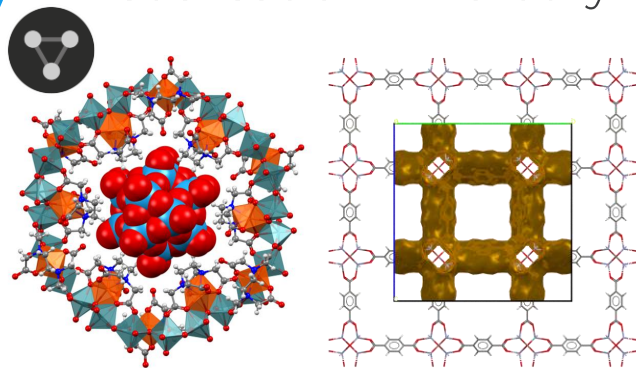


The CSD Portfolio



Search, visualisation and analysis

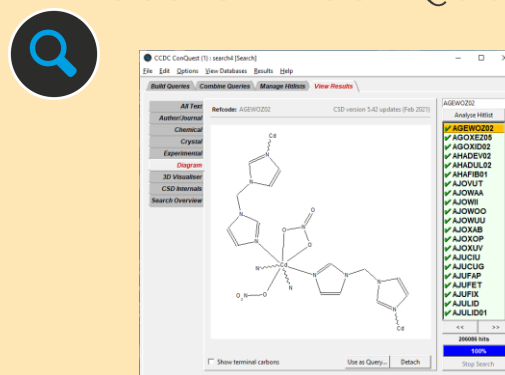
Visualisation - Mercury



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

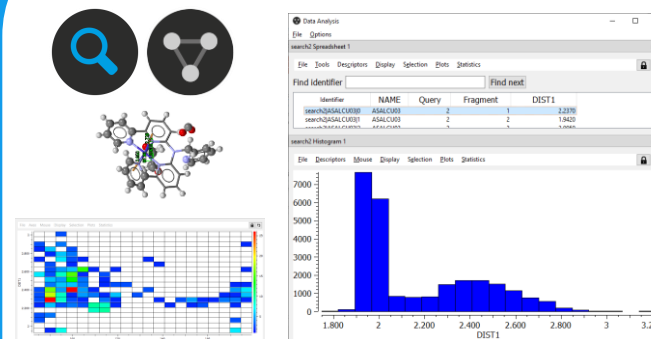
Search - ConQuest



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, porosity, 1D, 2D and 3D networks

CSD-Core: Search



Programmatically:
CSD Python API

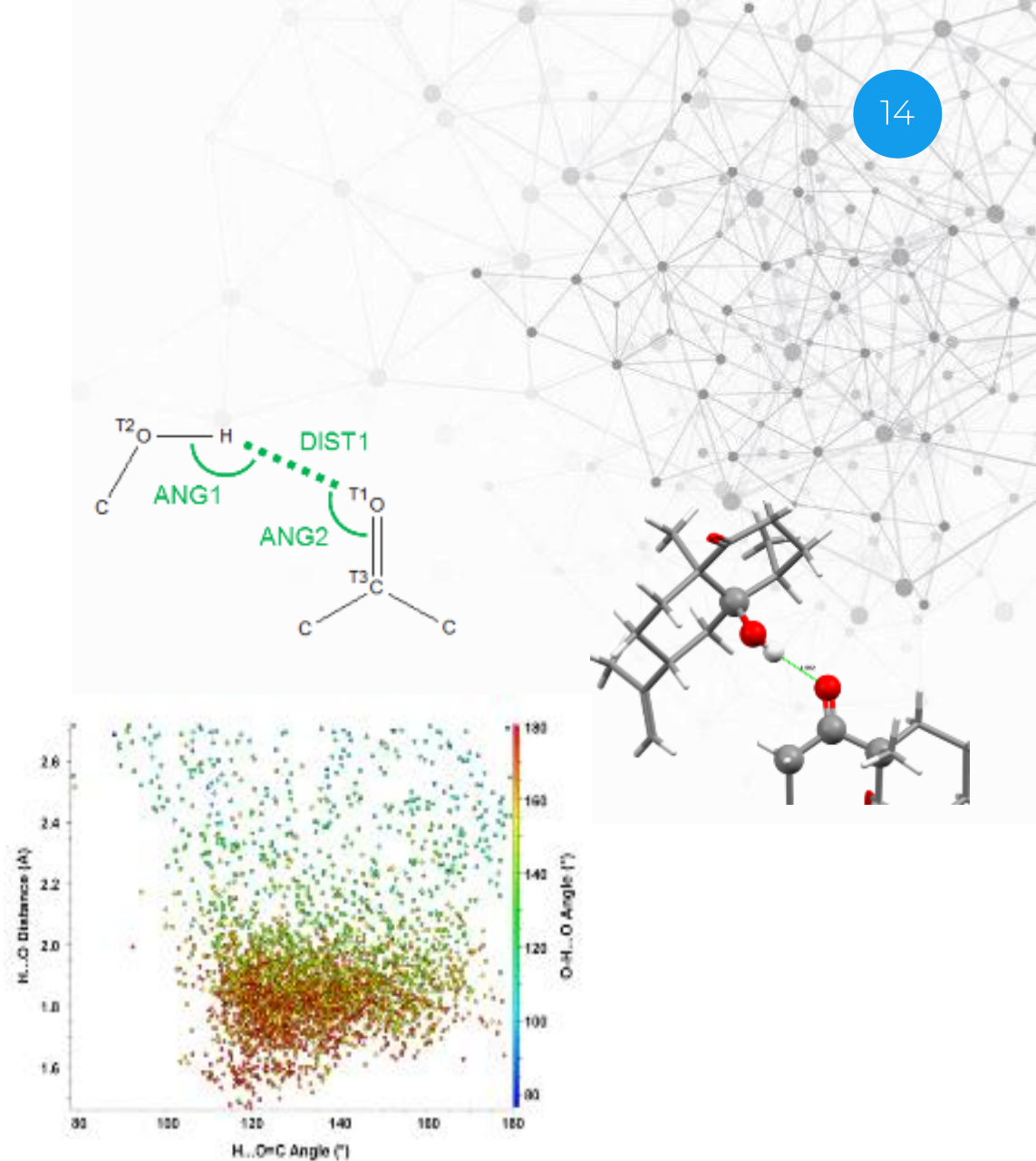


On your desktop: [ConQuest](#)

On a browser: [WebCSD](#)
Or Free [Access Structures](#)

What is ConQuest?

- Enables [search](#) and retrieval of information from the CSD
- Provides full range of [text / numeric](#) database search options
- More complex search functionality includes:
 - [Chemical structure](#) searching
 - [3D Geometrical](#) searching
 - Intermolecular non-bonded [contact searching](#)



Beware...

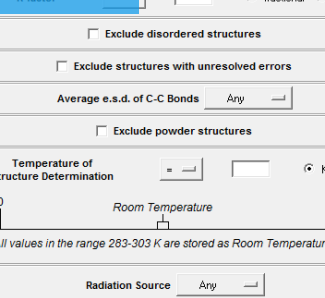
You can get answers to many important questions...



...but be careful when constructing your queries!!!

ConQuest features

Search for information relating to the structure determination



ation

R-factor Any fractional %

☐ Exclude disordered structures

☐ Exclude structures with unresolved errors

Average e.s.d. of C-C Bonds Any

☐ Exclude powder structures

Temperature of Structure Determination Room Temperature 0 610 K

All values in the range 283-303 K are stored as Room Temperature

Radiation Source Any

Search Store Cancel Reset

Structure sketcher, enabling set-up of basic substructure searches and complex 3D queries

The screenshot shows the ChemDraw 12.0x software interface. The 'Draw' menu is open, displaying options: EDIT, ERASE, ADD 3D, and CONTACT. A yellow tooltip box is visible over the main canvas area, containing the text: 'Click and drag to create a bond. Drag to an existing atom to make a connection.' The 'Bond' dialog box is open at the bottom, showing a list of elements (C, H, O, N, S, P, F, Cl, Any, More...) and the 'Bond' type set to 'Single'. The '3D Parameters' panel on the right is partially visible, showing 'Options...' and 'Delete' buttons.

Bibliographic search

Authors' Names **New Box**

☐ Exact surname

(Required format: F.H.Allen, O'Hara, Murray-Rust etc.
Brown will hit Browning unless 'Exact surname' is selected)

Journal Name

Type part of Journal name above to narrow list displayed
Select required journal in list below

- A.C.A. (Spring) [1974-1975]
- A.C.S. Mtg.172,Inorg. [1976]
- AAPS PharmSciTech [2004-2005]
- ACA Abstr.Papers(Winter) [1967-1986]
- ACA,Ser.2 [1977-1984]
- ADGC Chem.Res.Comm. [2005]
- ADGC Chem.Res.Comm. [2001-2008]
- ACH-Models Chem. [1994-2000]
- ACS Applied Materials and Interfaces [2009-2012]
- ACS Catalysis [2011-2012]

Volume (14, 1,2 etc.) **Page (212,6-A etc.)** **Year (1998, 2001 etc.)**

during

CCDC Number (Enter numeric part only, e.g. 123456 or 123/456)

Generic text search

Text Search Required Fields

New Box

either select from list or enter in box(es) below

absolute configuration
acouster
activity
agent
ar-sensitive
bai
black
blade
block
blue
brown
colorless
column

GenText search

Search Store Cancel Reset

Analysis of results

Elemental make-up

Elements Required to be Present

Type in elements, e.g. C H S

or

Elements must be in

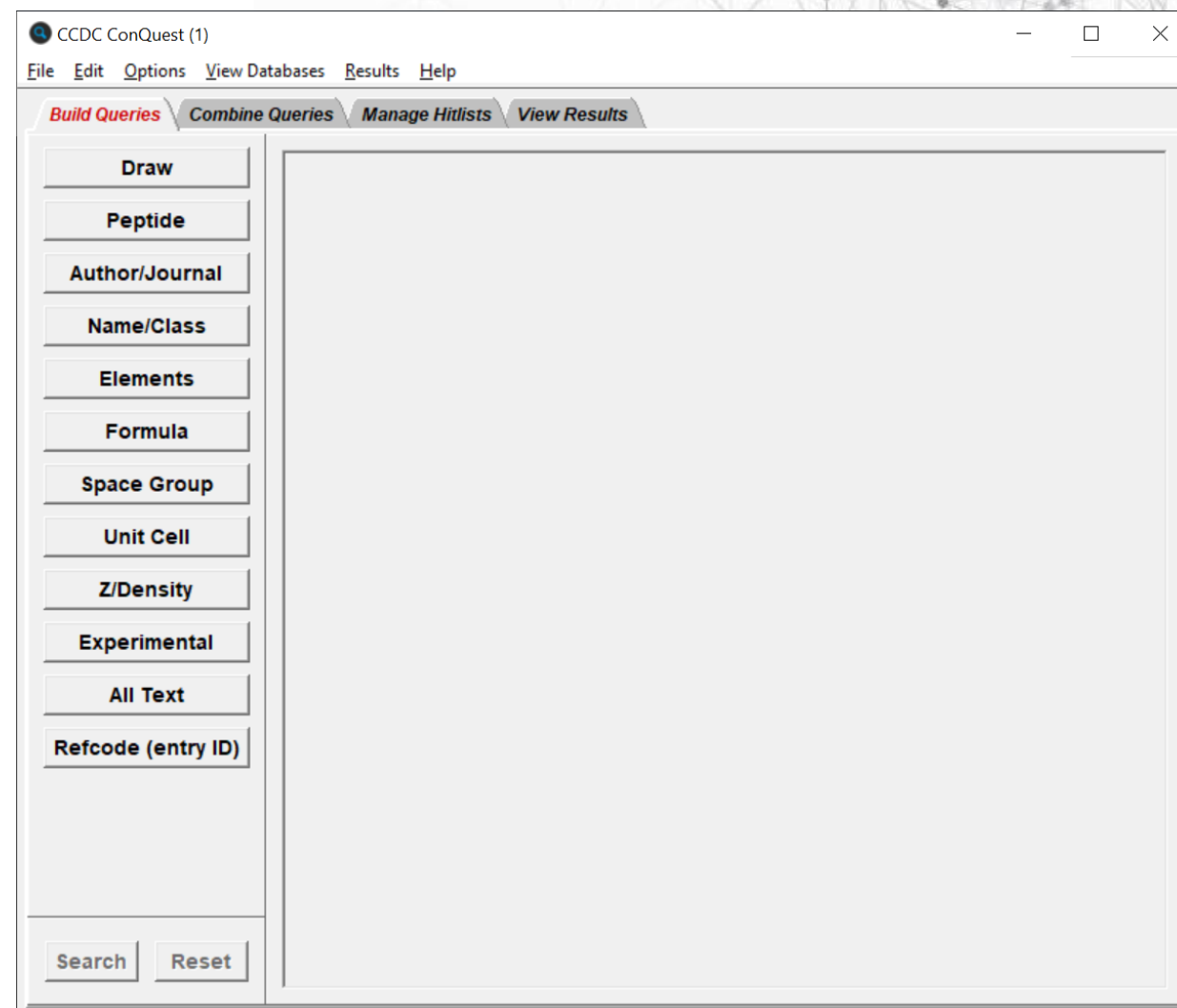
☐ same molecule

☒ same crystal structure

☒ Other elements allowed in molecule/structure

How many Permitted Element in Formula Unit

Show One: Searching metal-organics



ConQuest – Searching and menu options

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

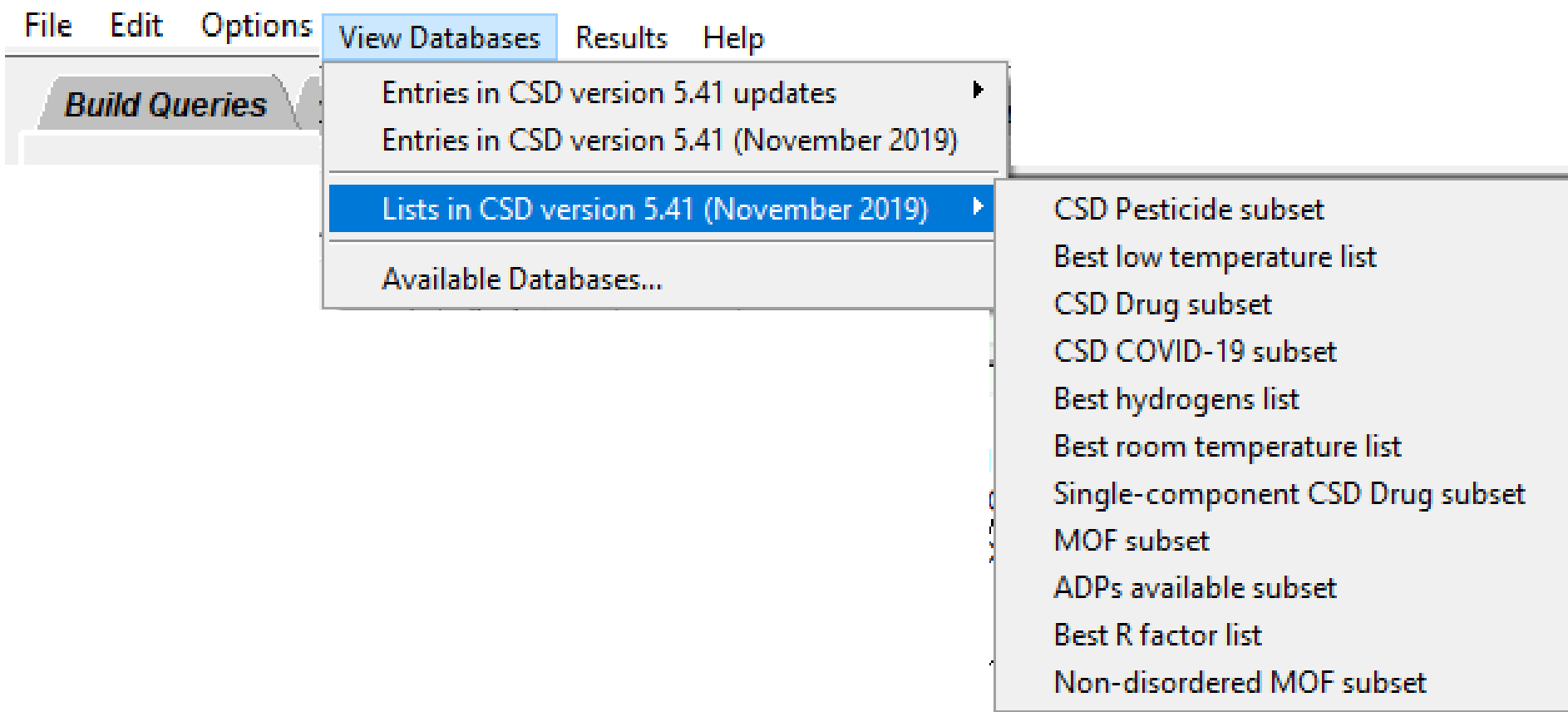
Z/Density

Experimental

All Text

Refcode (entry ID)

Search Reset



Filters

Search Setup

Search Name:

Available Databases: ☒ Show Updates separately

- ☒ CSD version 5.42 updates (Feb 2021)
- ☒ CSD version 5.42 (November 2020)
- ☒ CSD version 5.42 updates (May 2021)

You can search complete database(s) or a subset (e.g., hits found in a previous search)

Single query being used. Search will find structures:
where this query is true:

Filters **Advanced Options**

- ☐ 3D coordinates determined
- ☐ R factor ☒ ≤ 0.05
☐ ≤ 0.075
☐ ≤ 0.1
- ☐ Only ☒ Non-disordered
☐ Disordered
- ☐ No errors
- ☐ Not polymeric
- ☐ No ions
- ☐ Only ☒ Single crystal structures
☐ Powder structures
- ☒ Only ☐ Organics
☒ Organometallic

- 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

ConQuest – Author and journal search

Author/Journal (1) - Query 1

Authors' Names New Box

☐ Exact surname

(Required format: F.H.Allen, O'Hara, Murray-Rust etc.
Brown will hit Browning unless 'Exact surname' is selected)

Journal Name

Type part of Journal name above to narrow list displayed
Select required journal in list below

3D Mat. [2017]
J. Mater. Chem. [2015]
J. Mater. Chem. A (Spring) [1974-1975]
A.C.S. Mtg. 172, Inorg. [1976]
AAPS PharmSciTech [2004-2013]
ACA Abstr. Papers (Winter) [1967-1986]
ACA, Ser. 2 [1977-1984]
ACGC Chem. Res. Commun. [2001-2009]
ACH-Models Chem. [1994-2000]
ACS Appl. Bio Mater. [2019]

Volume (14, 1.2 etc.) Page (212, 6-A etc.) Year (1998, 2001 etc.)

during

CCDC Number (Enter numeric part only, e.g. 123456 or 123/456)

Search **Store** **Cancel** **Reset**

Author names should be initials followed by surname with no spaces e.g. S.J.Coles

Search Setup

Search Name:

Available Databases:

☒ CSD version 5.42 (November 2020)

You can search complete database(s) or a subset (e.g., hits found in a previous search)

Single query being used. Search will find structures:

where this query is true: **Query 1**

Filters **Advanced Options**

☐ 3D coordinates determined

☐ R factor ☐ ≤ 0.05 ☐ ≤ 0.075 ☐ ≤ 0.1

☐ Only ☒ Non-disordered ☐ Disordered

☐ No errors

☐ Not polymeric

☐ No ions

☐ Only ☒ Single crystal structures ☐ Powder structures

☐ Only ☒ Organics ☐ Organometallic

Query highlighted

Left clicking each tab will display different information

CCDC ConQuest (1) : search3 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists **View Results**

Refcode: AGACAP CSD version 5.42 (November 2020)

Author(s)	K.E.Thomas, K.J.Gagnon, L.J.McCormick , A.Ghosh
Reference	J.Porphyrins Phthalocyanines (2018), 22, 596
Publication DOI	10.1142/S1088424618500815
Deposition	CCDC 1534683
Formula	$C_{37}H_{12}AuBr_8F_{15}N_4S_3$
Compound	(2,3,7,8,12,13,17,18-octabromo-5,10,15-(4'-pentafluorosulfanylphenyl)corrolato)-gold(III)
CCDC Class	Porphyrins, corrins & complexes
Spacegroup	Name: C2/c Number: 15
Cell	a: 32.932(1) b: 10.418(<1) c: 27.887(1) alpha: 90.00 beta: 96.76(<1) gamma: 90.00 Volume: 9501.419
Reduced Cell	a: 10.418 b: 17.271 c: 27.887 alpha: 96.44 beta: 90.00 gamma: 107.56 Volume: 4750.708
Molecular Volume	1187.677
Chemical Units	1
Z, Z'	Z: 8.0 Z': 1.0

Detach

AGACAP

Analyse Hitlist

- ✓ AGACAP
- ✓ AMASEO
- ✓ AMASIS
- ✓ CIMYAA
- ✓ DIJQEV
- ✓ EWEPEC
- ✓ EWEPIG
- ✓ FADJAX
- ✓ FADJEB
- ✓ FADJIF
- ✓ FADJOL
- ✓ FADJUR
- ✓ FADKAY
- ✓ FADKEC
- ✓ FEYPEH
- ✓ FEYPIL
- ✓ FEYPOR
- ✓ FIFXEA
- ✓ FIFXIE
- ✓ FIHNIV
- ✓ FIHNOB
- ✓ FIHNI IH

<< >>

120 hits

100%

Stop Search

A hit list will be returned within the View Results tab of ConQuest. Left clicking on each REFCODE will toggle on and off that hit for future manipulation.

From CSD entry to data and back

CCDC ConQuest (1) : search3 [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: AGACAP CSD version 5.42 (November 2020)

Author(s)	K.E.Thomas, K.J.Gagnon, L.J.McCormick, A.Ghosh
Reference	J.Porphyrins Phthalocyanines (2018), 22, 596
Publication DOI	10.1142/S1088424618500815
Deposition	CCDC 1534683
Formula	C ₃₇ H ₁₂ AuBr ₈ F ₁₅ N ₄ S ₃
Compound	(2,3,7,8,12,13,17,18-octabromo-5,10,15-(4'-pentafluorosulfanylphenyl)corrolato)-gold(III)
CCDC Class	Porphyrins, corrins & complexes
Spacegroup	Name: C2/c Number: 15
Cell	a: 32.932(1) b: 10.418(<1) c: 27.887 alpha: 90.00 beta: 96.76(<1) gamma: 90.00 Volume: 9501.419
Reduced Cell	a: 10.418 b: 17.271 c: 27.887 alpha: 96.44 beta: 90.00 gamma: 107.56 Volume: 4750.708
Molecular Volume	1187.677
Chemical Units	1
Z, Z'	Z: 8.0 Z': 1.0

AGACAP

Analyse Hitlist

- ✓ AGACAP
- ✓ AMASEO
- ✓ AMASIS
- ✓ CIMYAA
- ✓ DIJQEV
- ✓ EWEPEC
- ✓ EWEPIG
- ✓ FADJAX
- ✓ FADJER

120 hits

100%

Detach

Stop Search

CSD entries

- Contain curated and enhanced information based on the deposited CIF

Deposited CIFs

- Can be downloaded from the CCDC# link

CCDC FIZ Karlsruhe Leibniz Institute for Information Infrastructure

CSD Entry: AGACAP

Sign In

Licensed to: CCDC Main Site

Simple Search Structure Search Unit Cell Search Formula Search

Your query was: Identifier(s): 1534683, DOI: doi:10.1142/S1088424618500815, Year: 2018, Volume: 22, Page: 596 and the search returned 1 record.

Modify Search New Search

Results

Database Identifier	Deposition Number
✓ AGACAP	1534683

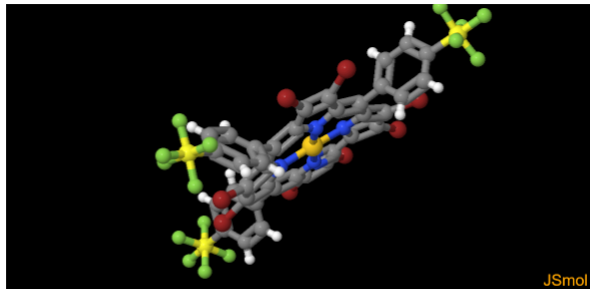
Download

- Download current entry
- Download all selected entries
- Download GCD File

AGACAP : (2,3,7,8,12,13,17,18-octabromo-5,10,15-(4'-pentafluorosulfanylphenyl)corrolato)-gold(III)

Space Group: C 2/c (15), Cell: a 32.9324(14)Å b 10.4183(5)Å c 27.8867(11)Å, α 90° β 96.759(3)° γ 90°

3D viewer



JSmol

H Disorder Menu Open

Style: Ball and Stick Labels: No Labels Packing: None Measure: None

Left clicking
each tab will
display
different
information

Link to
International
Tables

CCDC ConQuest (1) : search10 [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: FOPZUH

Spacegroup

Name: P43212 Number: 96

Cell Parameters

a: 21.500(3) b: 21.500(3) c: 14.525(3)

alpha: 90.00 beta: 90.00 gamma: 90.00

Volume: 6714.181

Reduced Cell Parameters

a: 14.525 b: 21.500 c: 21.500

alpha: 90.00 beta: 90.00 gamma: 90.00

Volume: 6714.181

Other Parameters

Molecular Volume: 839.273 Z: 8.0

Chemical Units: 1 Z': 1.0

Additional Information

From: DMF

Habit: needle

Pol

Recrystallisation solvents and crystal shape

FOPZUH

Analyse Hitlist

- ✓ FIHPUJ
- ✓ FOPZOB
- ✓ FOPZUH
- ✓ FULLII01
- ✓ GIWFUP01
- ✓ GUPBEA
- ✓ GUPBIE
- ✓ GUPBOK
- ✓ GUPBUQ
- ✓ HAJHUW01
- ✓ HAPVEC
- ✓ JIKJAR
- ✓ KERZEP
- ✓ KIKCOY
- ✓ KOCMEX
- ✓ KOCMIB
- ✓ LAZMAD
- ✓ LAZRIQ
- ✓ LAZROW
- ✓ LAZSAJ
- ✓ LEXQIR
- ✓ NESLOR

<< >>

120 hits

100%

Stop Search

CCDC ConQuest (1) : search10 [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: JIKJAR CSD version 5.41 (November 2019)

Formula: $C_{48}H_{18}F_{12}N_4 \cdot 2(C_4H_4O) \cdot 0.8(H_2O)$

Name: 2,3,7,8,12,13,17,18-octaodo-5,10,15,20-tetrakis[4-(trifluoromethyl)phenyl]porphyrin methanol solvate hydrate

Synonym:

CCDC Class: Porphyrins, corrins & complexes

Source:

Melting Point:

Colour: red

Extra Information: Heat-sensitive, Light-sensitive, Photo-sensitive

JIKJAR

Analyse Hitlist

- ✓ FIHPUJ
- ✓ FOPZOB
- ✓ FOPZUH
- ✓ FULLII01
- ✓ GIWFUP01
- ✓ GUPBEA
- ✓ GUPBIE
- ✓ GUPBOK
- ✓ GUPBUQ
- ✓ HAJHUW01
- ✓ HAPVEC
- ✓ JIKJAR
- ✓ KERZEP
- ✓ KIKCOY
- ✓ KOCMEX
- ✓ KOCMIB
- ✓ LAZMAD
- ✓ LAZRIQ
- ✓ LAZROW
- ✓ LAZSAJ
- ✓ LEXQIR
- ✓ NESLOR

<< >>

120 hits

100%

Stop Search

Detach

CCDC ConQuest (1) : search10 [Search]

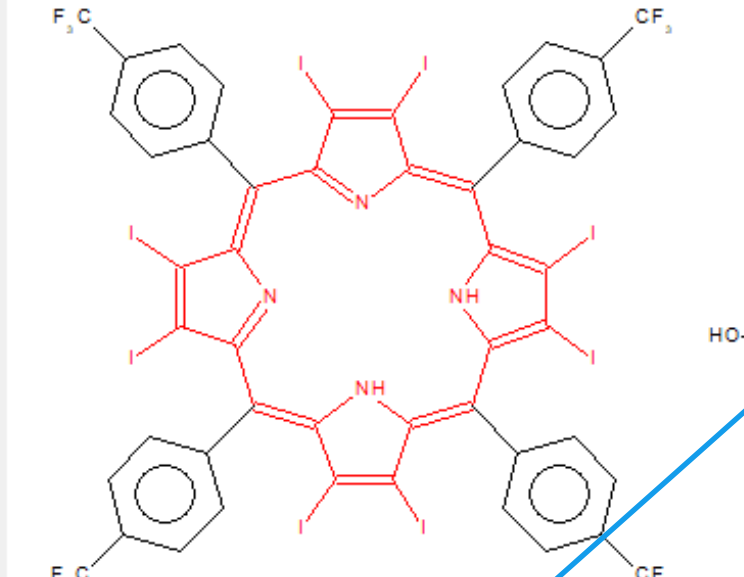
File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists **View Results**

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

Refcode: JIKJAR CSD version 5.41 (November 2019)

Search highlighted in 2D



HO-CH₃
H₂O

☒ Show terminal carbons

Use as Query... Detach

120 hits
100%
Stop Search

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

CCDC ConQuest (1) : search10 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

Refcode: JIKJAR CSD version 5.41 (November 2019)

3D Visualiser

Right-click in visualiser for options menu

Display Style

- Wireframe
- Capped Sticks
- Ball & Stick
- Spacefill

Packing

- Molecule or unit cell

Highlight Hit

Display Hydrogens

Display Disordered Atoms

Picking Mode

Measure

Clear Measurements

Label

Settings

Save as PNG...

Save as JPEG...

Save current setting

Reset View

120 hits

100%

Stop Search

ODFAQ

ODFEU

POTID01

WITAQ

WITEU

WITIY

WITOE

WITUK

WITUL

WITUR

WITUS

WITUT

WITUV

WITUX

WITUY

WITUZ

WITVQ

WITVR

WITVS

WITVT

WITVU

WITVV

WITVW

WITVX

WITVY

WITVZ

WITWA

WITWB

WITWC

WITWD

WITWE

WITWF

WITWG

WITWH

WITWI

WITWJ

WITWK

WITWL

WITWM

WITWN

WITWO

WITWP

WITWQ

WITWR

WITWS

WITWT

WITWU

WITWV

WITWW

WITWX

WITWY

WITWZ

WITXA

WITXB

WITXC

WITXD

WITXE

WITXF

WITXG

WITXH

WITXI

WITXJ

WITXK

WITXL

WITXM

WITXN

WITXO

WITXP

WITXQ

WITXR

WITXS

WITXT

WITXU

WITXV

WITXW

WITXZ

WITYA

WITYB

WITYC

WITYD

WITYE

WITYF

WITYG

WITYH

WITYI

WITYJ

WITYK

WITYL

WITYM

WITYN

WITYO

WITYP

WITYQ

WITYR

WITYS

WITYT

WITYU

WITYV

WITYW

WITYX

WITYY

WITYZ

WITZA

WITZB

WITZC

WITZD

WITZE

WITZF

WITZG

WITZH

WITZI

WITZJ

WITZK

WITZL

WITZM

WITZN

WITZO

WITZP

WITZQ

WITZR

WITZS

WITZT

WITZU

WITZV

WITZW

WITZX

WITZY

WITZZ



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

Packing > Molecule or unit cell

Display or hide disordered atoms

CCDC ConQuest (1) : search3 [Search]

File Edit Options View Databases Results Help

New Window Ctrl+N

Open

Save Search Ctrl+S

Save Search As...

Write PDF file to view/print... Ctrl+P

Read Queries... Ctrl+R

Save Queries

Export Entries as...

View Entries in Mercury Ctrl+M

Analyse Data in Mercury Alt+M

Export Parameters and Data...

View in Excel...

Close

Exit Ctrl+Q

Manage Hitlists View Results

CSD version 5.41 updates (Mar 2020)

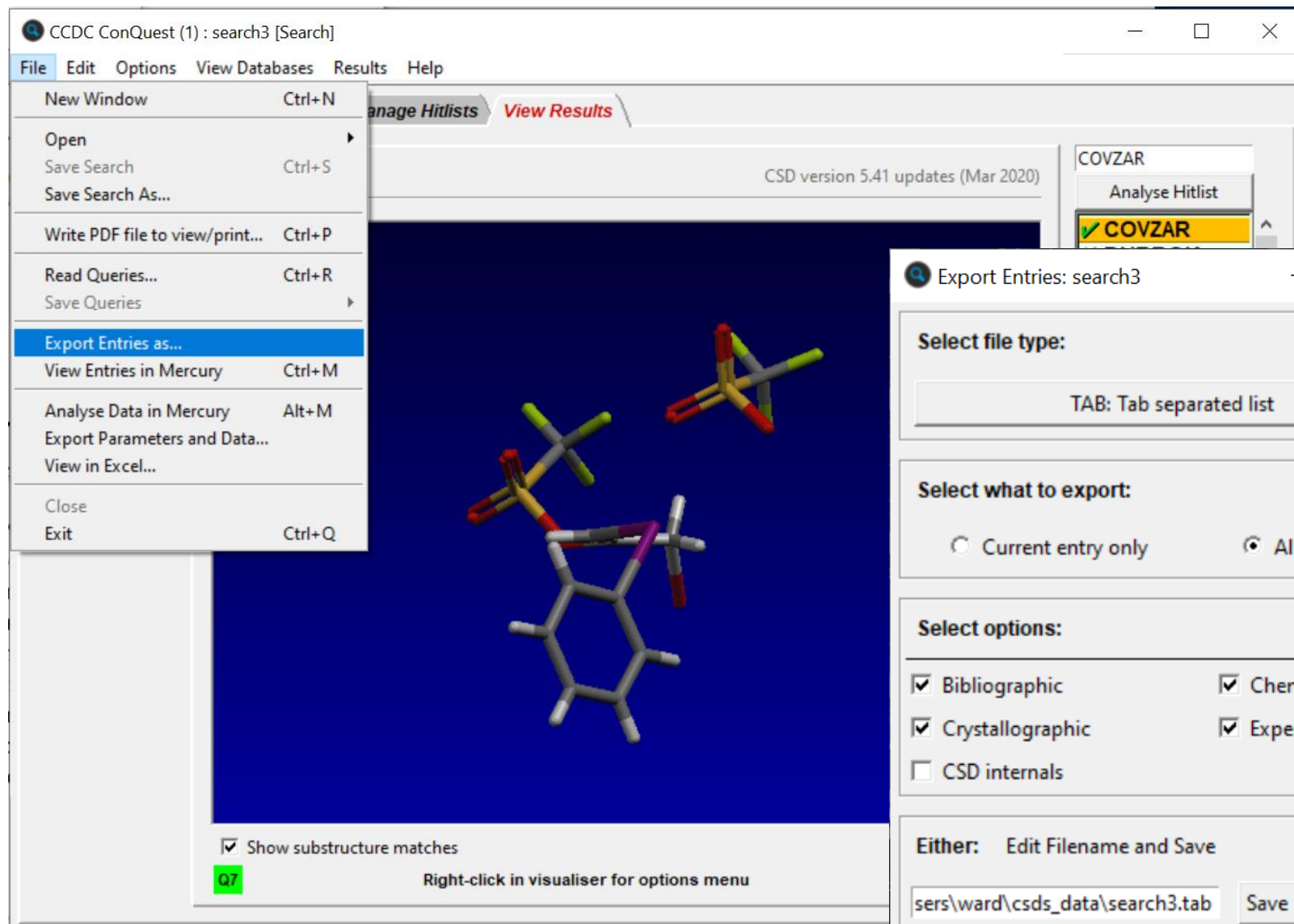
COVZAR

Analyse Hitlist

COVZAR

Q7

Right-click in visualiser for options menu



CIF: Crystallographic Information File

CIFMIF: Combined CIF and MIF file

COORD: CSD Coordinate file

TSV: Endnote (import as Tab Delimited)

FDAT: Legacy CSD crystallographic format

MOL2: Tripos exchange format

PDB: Original PDB format

Refcode: CSD entry identifier list

SD: SDfile format (MDL MOL)

SHELX: simulated SHELX .res file

SMILES: SMILES strings

TAB: Tab separated list

TEXT: Text representation

XML: Hit Fragments (XML)

You can export results in a range of different formats to save data for future manipulation.

ConQuest – Compound name search

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

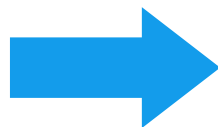
Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)



Name/Class (1) - Query 1

Compound Name

Contains: **Paracetamol**

☐ Ignore non-alphabetic characters, e.g. "butadiene" will match "buta-1,3-diene"

☐ Find exact word, e.g. "hydrazine" will not match "acetylhydrazine"

Add Replace Delete

Chemical Class

CCDC Chemical Class is assigned to some categories of entries in the CSD, particularly when it would be difficult to locate these categories using a substructure or compound name search.

Note that the results of a search on Class may not be comprehensive since it is not always evident from a publication that an entry belongs to one of the specified categories.

Find entries classified as: ----not defined----

and: ----not defined----

Search Store Cancel Reset

CSD contains IUPAC style names as well as more common names when known

Chemical classes can help you identify structures of interest

Compound	N-(4-hydroxyphenyl)acetamide
Synonym	acetaminophen; Paracetamol ; DrugBank: DB00316

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

Build Queries Combine C

Draw

Peptide

Author/Journal

Name/Class

Elements

Name/Class (1) - New

Compound Name

Contains: (ii)

☐ Ignore non-alphabetic characters, e.g. "butadiene" will match "buta-1,3-diene"

☐ Find exact word, e.g. "hydrazine" will not match "acetylhydrazine"

Add Replace Delete

Search Store Cancel Reset

Tip: use parenthesis to avoid

- Other oxidation states
- Polymorphic forms!

CCDC ConQuest (1) : search13 [Searching...]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists **View Results**

Refcode: AGEWOZ02 CSD version 5.42 updates (Feb 2021)

Author(s) Shanshan Zhang

Reference J.Mol.Struct. (2021), 1230, 129871

Publication DOI [10.1016/j.molstruc.2021.129871](https://doi.org/10.1016/j.molstruc.2021.129871)

Deposition CCDC [2026390](https://www.ccdc.ac.uk/2026390)

Formula $(C_{14}H_{16}CdN_{10}O_6)_n$

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

Spacegroup Name: P21/c Number: 14

a: 15.191(<1) b: 9.159(<1) c: 14.687(<1)

32(<1) gamma: 90.00

c: 15.191 gamma: 90.00

Volume: 1916.125

Molecular Volume 479.031

Chemical Units 1

Z, Z' Z: 4.0 Z': 1.0

R-Factor (%) 2.61

Temperature (K) Room Temp. (283-303)

Detach

Analyze Hitlist

AGEWOZ02

✓ AGOXEZ05

✓ AGOXID02

✓ AHADDEV02

✓ AHADUL02

✓ AHAFIB01

✓ AJOVUT

✓ AJOWAA

✓ AJOWII

✓ AJOWOO

✓ AJOWUU

✓ AJOXAB

✓ AJOXOP

✓ AJOXUV

✓ AJUCIU

✓ AJUCUG

✓ AJUFAP

✓ AJUFET

✓ AJUFIX

✓ AJULID

✓ AJULID01

✓ AJULUP

✓ AJUMAW

<< >>

3165 hits

1%

Stop Search

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

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

ConQuest – Elements search

Build Queries

Combine

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

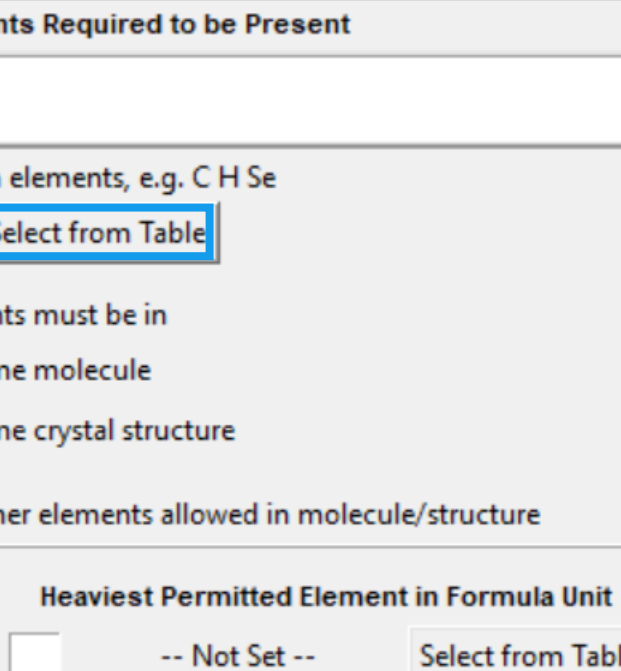
Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)



Elements (1) - New

Elements Required to be Present

Type in elements, e.g. C H Se
or **Select from Table**

Elements must be in

☐ same molecule
☒ same crystal structure

☒ Other elements allowed in molecule/structure

Heaviest Permitted Element in Formula Unit

-- Not Set -- **Select from Table**

Search Store Cancel Reset

[illegible]

ConQuest – Formula search

Build Queries **Combine**

- Draw
- Peptide
- Author/Journal
- Name/Class
- Elements
- Formula**
- Space Group
- Unit Cell
- Z/Density
- Experimental
- All Text
- Refcode (entry ID)

Formula (1) - New

Formula

C₁₀H₁₀Fe

Type in formula, e.g. C₆ H₁₂ O₆
You may specify an inexact formula, e.g. Ca₁₋₃ O_{>4}
finds entries with 1,2 or 3 Calciums and more than 4 Oxygens.
or **Select from Table**

Formula applies to

- ☒ an individual molecule
- ☐ all molecules in structure added together
- ☐ Other atoms allowed in molecule/structure

Search Store Cancel Reset

Select Formula Elements

Select required elements

1A 1M 2M 3M 4M 8A

HD D 2A 3A 4A 5A 6A 7A He

Li Be 8B B C N O F Ne

Na Mg 3B 4B 5B 6B 7B 8X 8Y 8Z 1B 2B Al Si P S Cl Ar

1R K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr

TR 2R Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe

3R Cs Ba La Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn

Fr Ra Ac

LN Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu

AN Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr

Select atom counts

C	=	10
H	=	10
Fe	=	1
	Range	
	Range	
	Range	
	Range	
	Range	
	Range	
	Range	
	Range	

Cancel Done

Other searches

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)

Space Group (1) - New

Space Group

Enter full or partial space group symbol or number ...

... or pick from list. Click on symbol to select

A1
B1
C1
F1
I1
P1
A-1

Current Space Group (with alternative settings)

☒ Include alternative settings of selected space group

Spacegroup Symmetry

Crystal System

Unit Cell (1) - New

Do you want to search on the reduced cell?

You should search on reduced cell if you want to find structures which match a particular set of cell dimensions (a,b,c,alpha,beta,gamma)

☐ Yes, do a reduced cell search ☐ No, do not do a reduced cell search

Tolerance 1.5 % of longest cell dimension

Lattice Type ----not defined----

Cell Parameters

a (Å)	=		alpha (°)	=	
b (Å)	=		beta (°)	=	
c (Å)	=		gamma (°)	=	

Search Store Cancel Reset

Z/Density (1) - New

No. of "Molecules" per Unit Cell (Z)

No. of "Molecules" per Asymmetric Unit (Z')

No. of Atoms with 3D Coordinates

Calculated Density (g/cm³)

No. of Chemical Units (molecules, ions, etc.) in Entry

Search Store Cancel Reset

Experimental (1) - New

R-factor

fractional %

Exclude d

Exclude structure

Average e.s.d. of C

Exclude powder structures

Temperature of Structure Determination

0 Room Temperature 610K

All values in the range 283-303 K are stored as Room Temperature

Radiation Source Any

Search Store Cancel Reset

All Text (1) - New

Text Search Required Fields

New Box

Either select from list or enter in box(es) below

acicular
air-sensitive
bar
black
cream
cube

The search will find words starting

If two or more words are typed be for the exact phrase specify two or more words that need not be adjacent, use the New Box button and type the required words into separate input boxes.

Search Store Cancel Reset

View Refcode

Enter Refcode (CSD entry code) below

Go to entry in Database

Find all Refcodes that begin with query text

Alternatively view full database(s) using the 'View Databases' options on the top menu

Refcode search

CCDC

ConQuest – more text/numeric searches

Name/Class (1) - New

Compound Name: catena-

Contains:

☐ Ignore non-alphabetic characters, e.g. "butadiene" will match "buta-1,3-diene"

☐ Find exact word, e.g. "hydrazine" will not match "acetylhydrazine"

Add Replace Delete

Search Store Cancel Reset

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

Z/Density (1) - New

No. of "Molecules" per Unit Cell (Z)	=	
No. of "Molecules" per Asymmetric Unit (Z')	=	
No. of Atoms with 3D Coordinates	=	
Calculated Density (g/cm3)	=	
No. of Chemical Units (molecules, ions, etc.) in Entry	=	

Search Store Cancel Reset

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

All Text (1) - New

Text Search Required Fields

New Box

Either select from list or enter in box(es) below

cylindrical
decompose
deliquescent
diamagnetic
diamond
efflorescent
electron
explosive
fawn
fluorescent
green
grey
habit

The search will find words starting with what is entered in the boxes.

If two or more words are typed into the same box the search will be for the exact phrase specified. To find entries containing two or more words that need not be adjacent, use the New Box button and type the required words into separate input boxes.

Search Store Cancel Reset

Storing and searching

Author/Journal (1) - Query 1

Authors' Names New Box

☐ Exact surname

(Required format: F.H.Allen, O'Hara, Murray-Rust etc.
Brown will hit Browning unless 'Exact surname' is selected)

Journal Name

Type part of Journal name above to narrow list displayed
Select required journal in list below

- 2D Mat. [2017]
- 3 Biotech [2015]
- A.C.A.(Spring) [1974-1975]
- A.C.S.Mtg.172.Inorg. [1976]
- AAPS PharmSciTech [2004-2013]
- ACA Abstr.Papers(Winter) [1967-1986]
- ACA,Ser.2 [1977-1984]
- ACGC Chem.Res.Comm. [2001-2009]
- ACH-Models Chem. [1994-2000]
- ACS Appl. Bio Mater. [2019]

Volume (14, 1.2 etc.) Page (212,6-A etc.) Year (1998, 2001 etc.) during

CCDC Number (Enter numeric part only, e.g. 123456 or 123/456)



CCDC ConQuest (1) : search1 [Search]

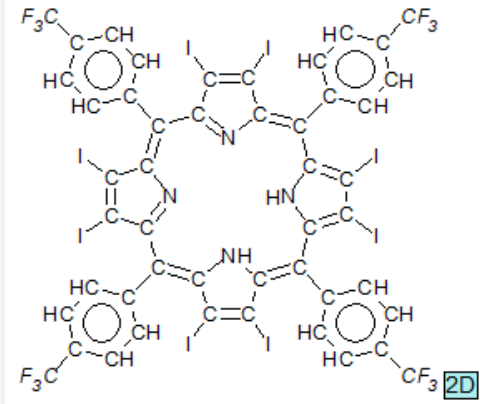
File Edit Options View Databases Results Help

Build Queries **Combine Queries** **Manage Hitlists** **View Results**

Draw
Peptide
Author/Journal
Name/Class
Elements
Formula
Space Group
Unit Cell
Z/Density
Experimental
All Text
Refcode (entry ID)

use this query? ☒ **Query 1** Author **L.J.McCormick**

use this query? ☒ **Query 2**



From JIKJAR

Combining queries and managing hitlists

CCDC ConQuest (1) : search1 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

Drag Query Icons into Boxes

Find entries that:

must have (boolean AND)

Query 1

must not have (NOT)

Query 2

must have at least one of (OR)

Search Reset

Query 1 Author L.J.McCormick Edit... Delete

Query 2

Chemical structure diagram of a complex molecule, labeled "From JIKJAR".

CCDC ConQuest (1) : search2 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

Combine Hitlists

Combination Name: combination1

List A List B

search1 search2

Include deselected entries in:

☐ List A ☐ List B

Generate a List of Entries:

☒ common to List A and List B

☐ in either List A or List B

☐ in List A but not in List B

OK

Hitlist Overview

search1 (120 Entries)

Author L.J.McCormick

Name	Hits	Type
search1	120	Search
search2	1	Search

Delete Rename... Notes... View

Tip - You can rename a search

Draw

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)

Draw (1) - New

File Edit Atoms Bonds 3D Options Help

Element
Add Group
Expand Chemical Groups
Hydrogens
Charge
Number of Bonded Atoms
Cyclicity
All Atoms in Same Molecule

Drawing Options...
Snap to Grid
Save Window Size
Limit Window Size
Auto-Generate H
Balloon Help

Next Atom: C
Next Bond: Single

3D Parameters:

Options...
Delete

Contacts:

Options...
Delete

Search
Store
Cancel

DRAW
EDIT
ERASE
ADD 3D
CONTACT

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

RingMaker

Templates...

C H O N S P F Cl Any More... Groups... C Bond: Single

Select specific or general atom types/functional groups

Select bond type

ConQuest – Draw/Structure search

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

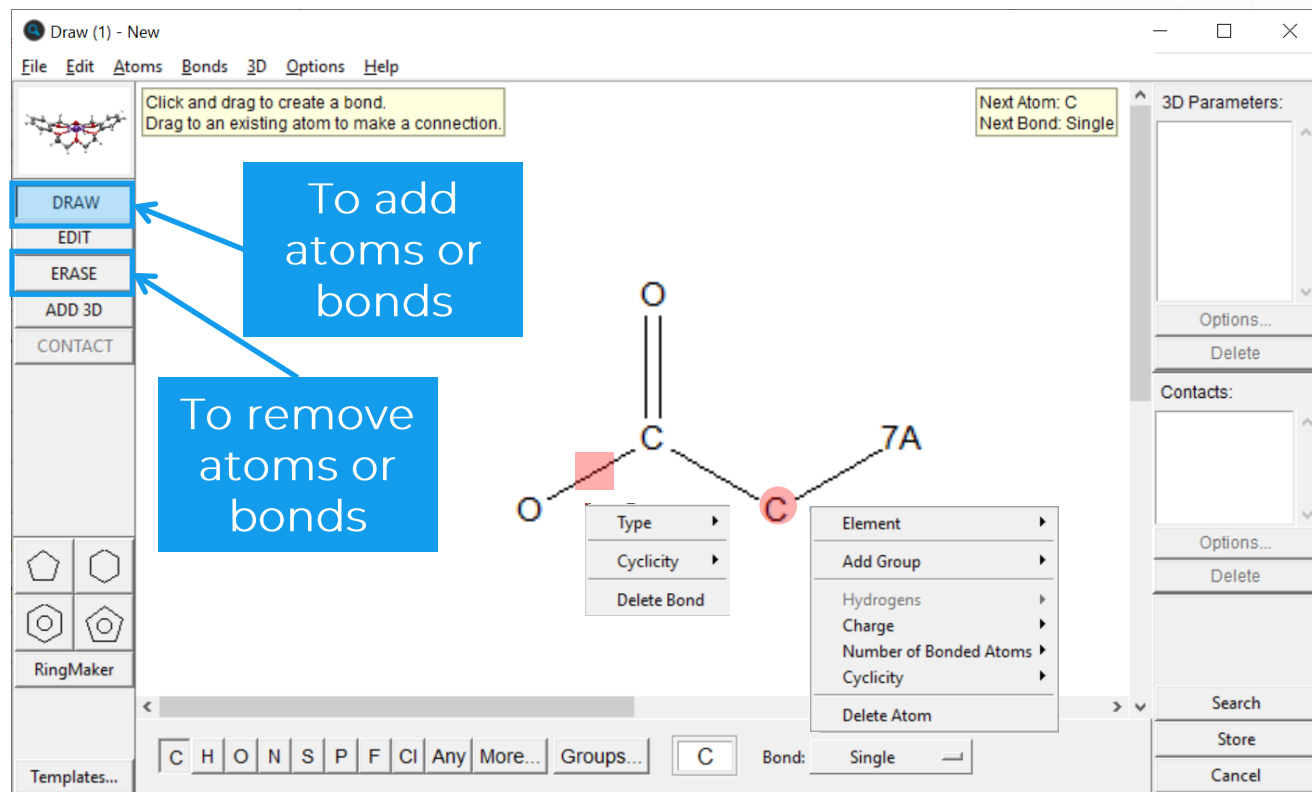
Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)



Left click in space you can add an atom



Left click and drag from an atom you can attach an atom



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

Example of a search, where 7A represents any halogen

ConQuest – Draw/Structure search

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

Z/Density

Experimental

All Text

Refcode (entry ID)

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

DRAW

EDIT

ERASE

ADD 3D

CONTACT

RingMaker

Templates...

CH₁ CH₁ CH₁ CH₁ CH₁ CH₁ N

Element

Add Group

Hydrogens

Charge

Number of Bonded Atoms

Cyclicity

Delete Atom

1

2

3

4

5

6

7

8

Unspecified

Other...

Next Atom: C

Next Bond: Single

3D Parameters:

Other Atom Type

Other Atom Type

Previous Selections

Picking Mode

Single Pick

Multi Pick

Current Selection: None

Reset Cancel Apply OK

Search

Store

Cancel

C H O N S P F Cl Any More... Groups... C

Bond: Single

Not Hydrogen
C or H

Any Non-metal

Any Metal

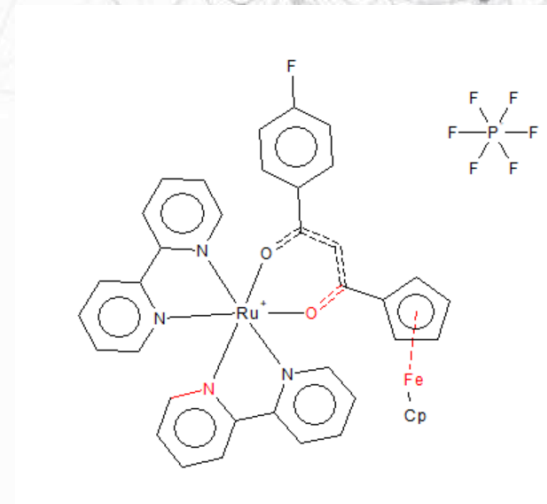
Any Transition Metal

Any Halogen

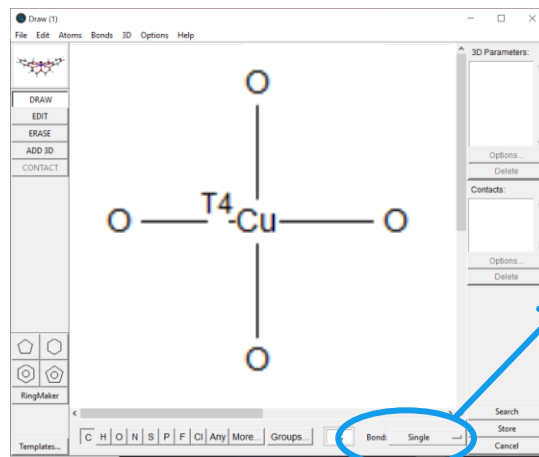
Other Elements...

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

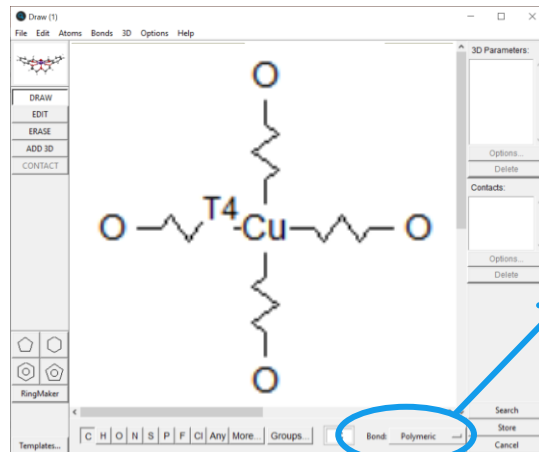


Substructure found



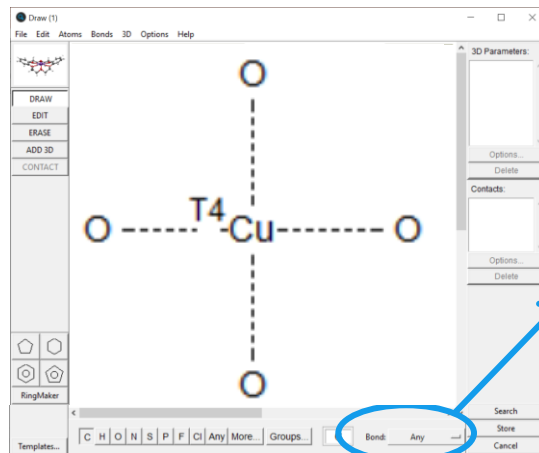
Bond:

Single
Double
Triple
Quadruple
Aromatic
Polymeric
Delocalised
Pi
Any
Variable...



Bond:

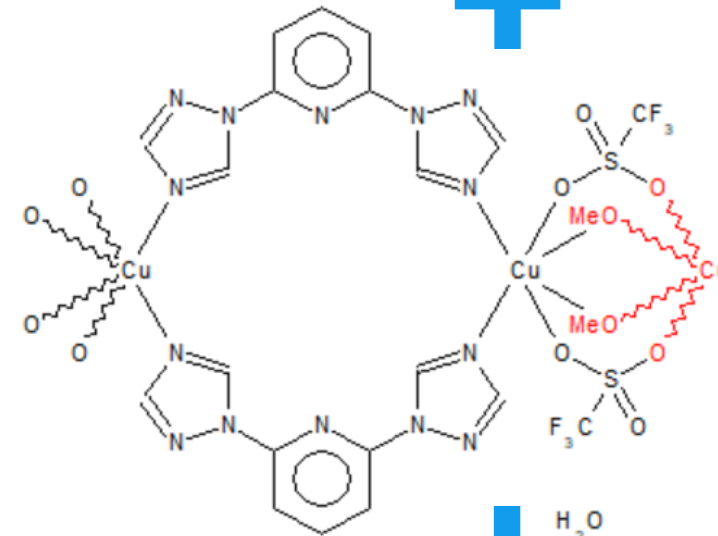
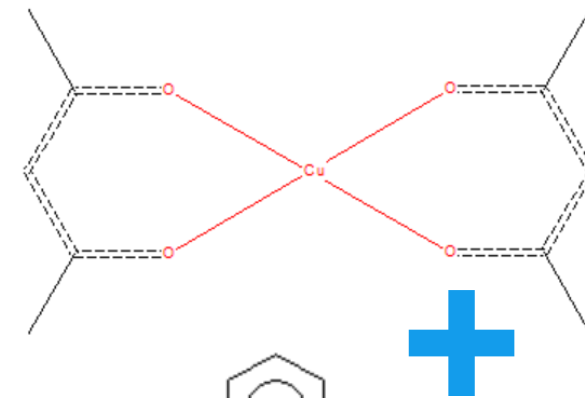
Single
Double
Triple
Quadruple
Aromatic
Polymeric
Delocalised
Pi
Any
Variable...



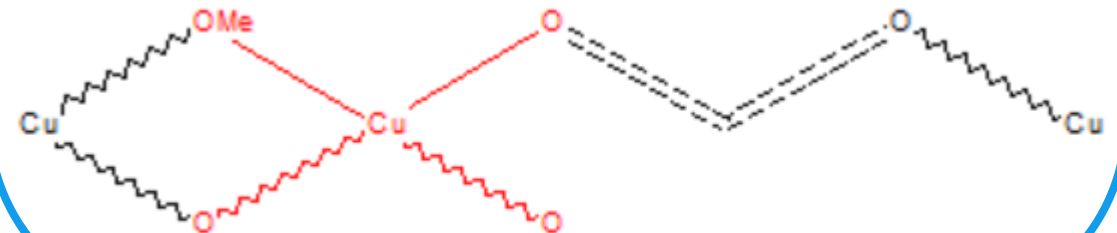
Bond:

Single
Double
Triple
Quadruple
Aromatic
Polymeric
Delocalised
Pi
Any
Variable...

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



H₂O



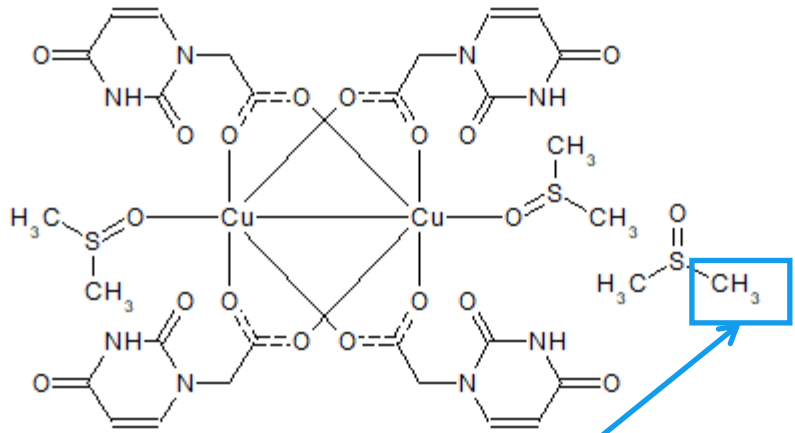
CCDC ConQuest (1) : search11 [Search]

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

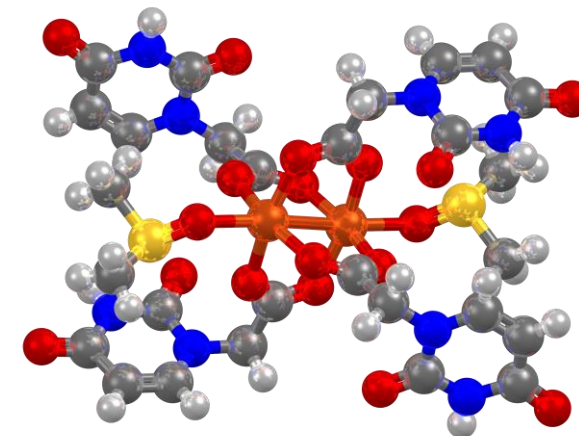
TUWMOP

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

<< >>

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.):

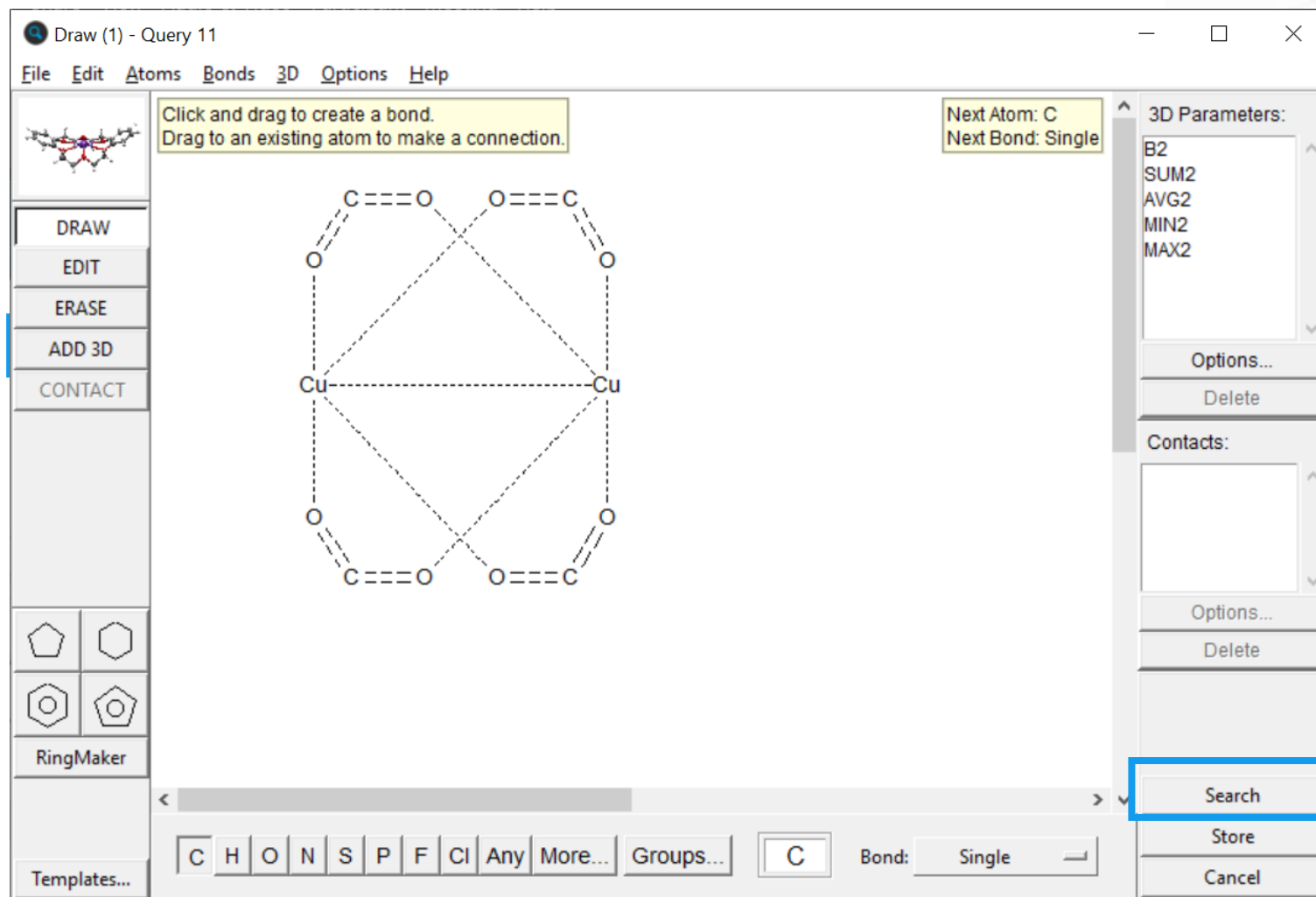
☒ Biggest chemical unit only

☐ 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

All Text

Author/Journal

Chemical

Crystal

Experimental

Diagram

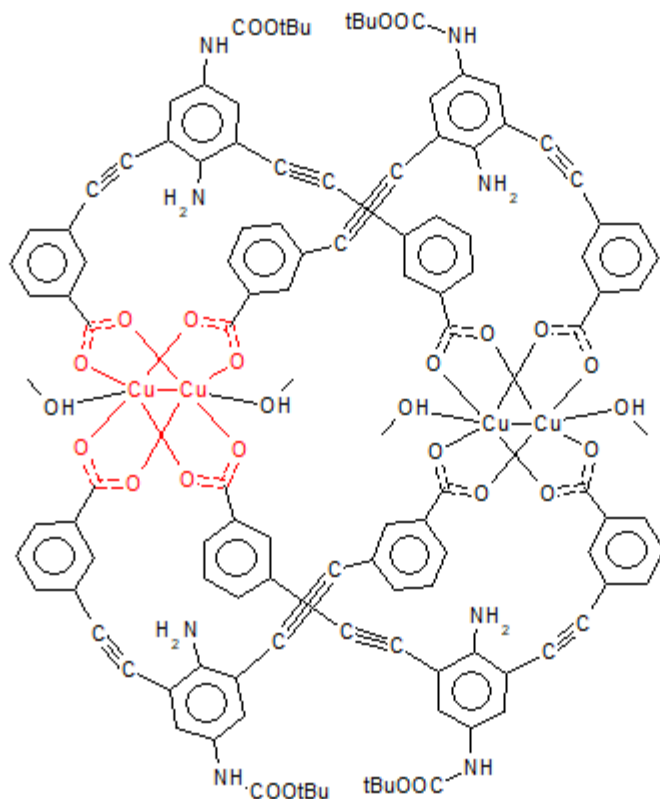
3D Visualiser

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

✓ CUYUUT

✓ JUZTOQ

✓ JUZTUW

✓ LOYYED

✓ LUYVEJ

✓ OKUSAR

✓ PYCUAC

✓ UJOKUC

✓ UJUNOF

✓ ABUXU

✓ ACACC

✓ ACACC

✓ ACACC

✓ ACAPO

✓ ACGLO

✓ ACNICU

✓ ACURC

✓ ACURC

✓ ACURC

✓ ACURC

✓ AGUTA

✓ AHEDO

✓ AMACU

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

829 h

0%

Reset

Cancel

Analyse in Mercury

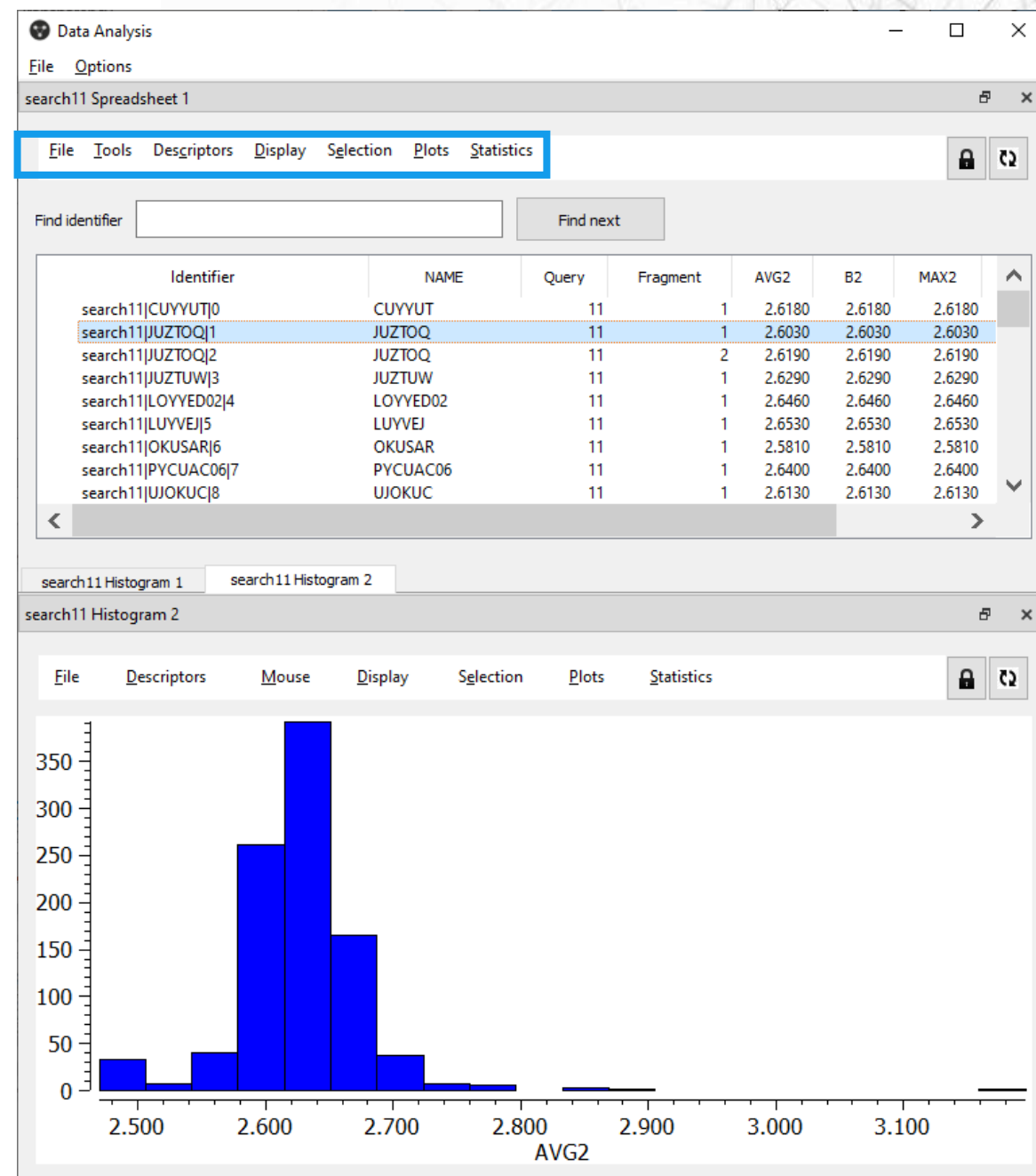
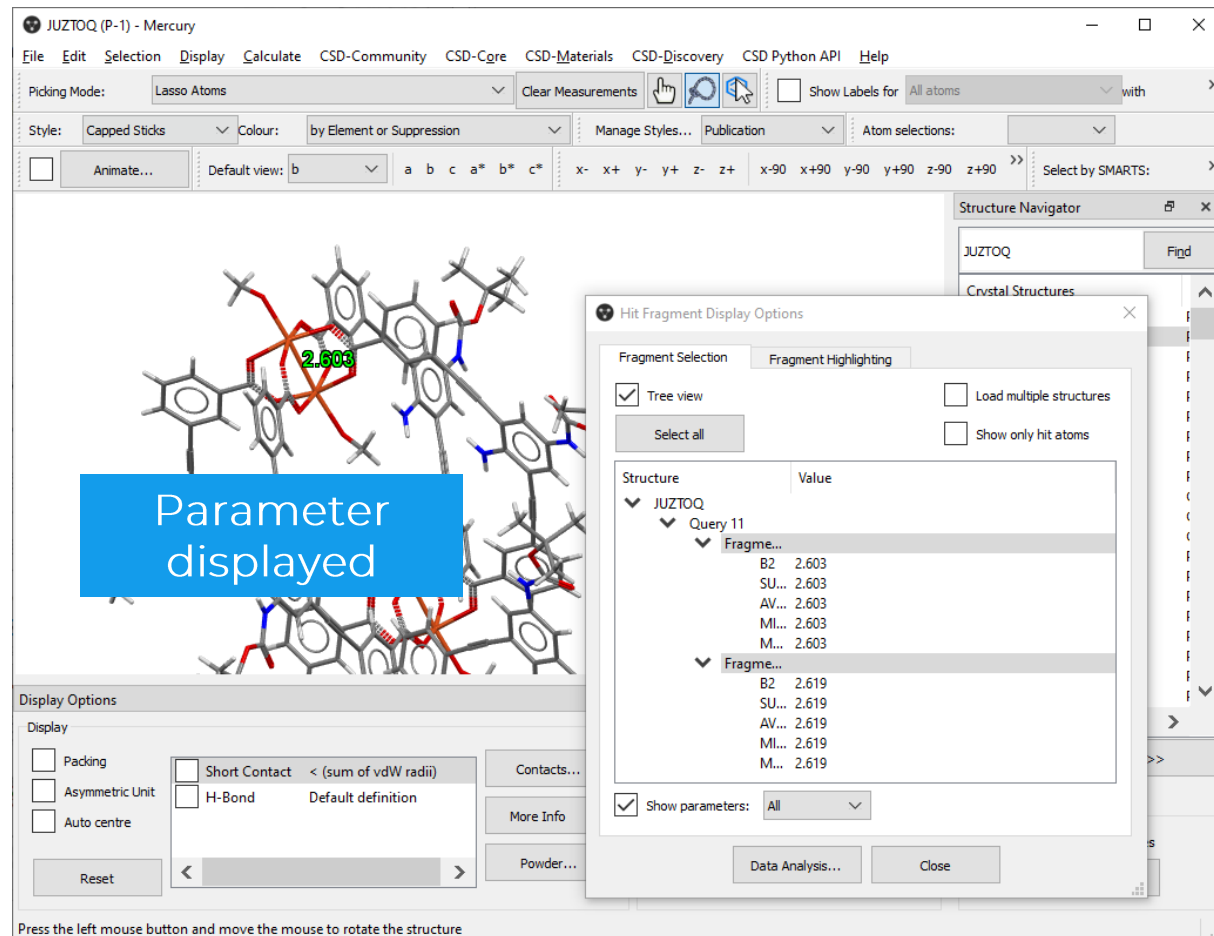
100%

Stop Search

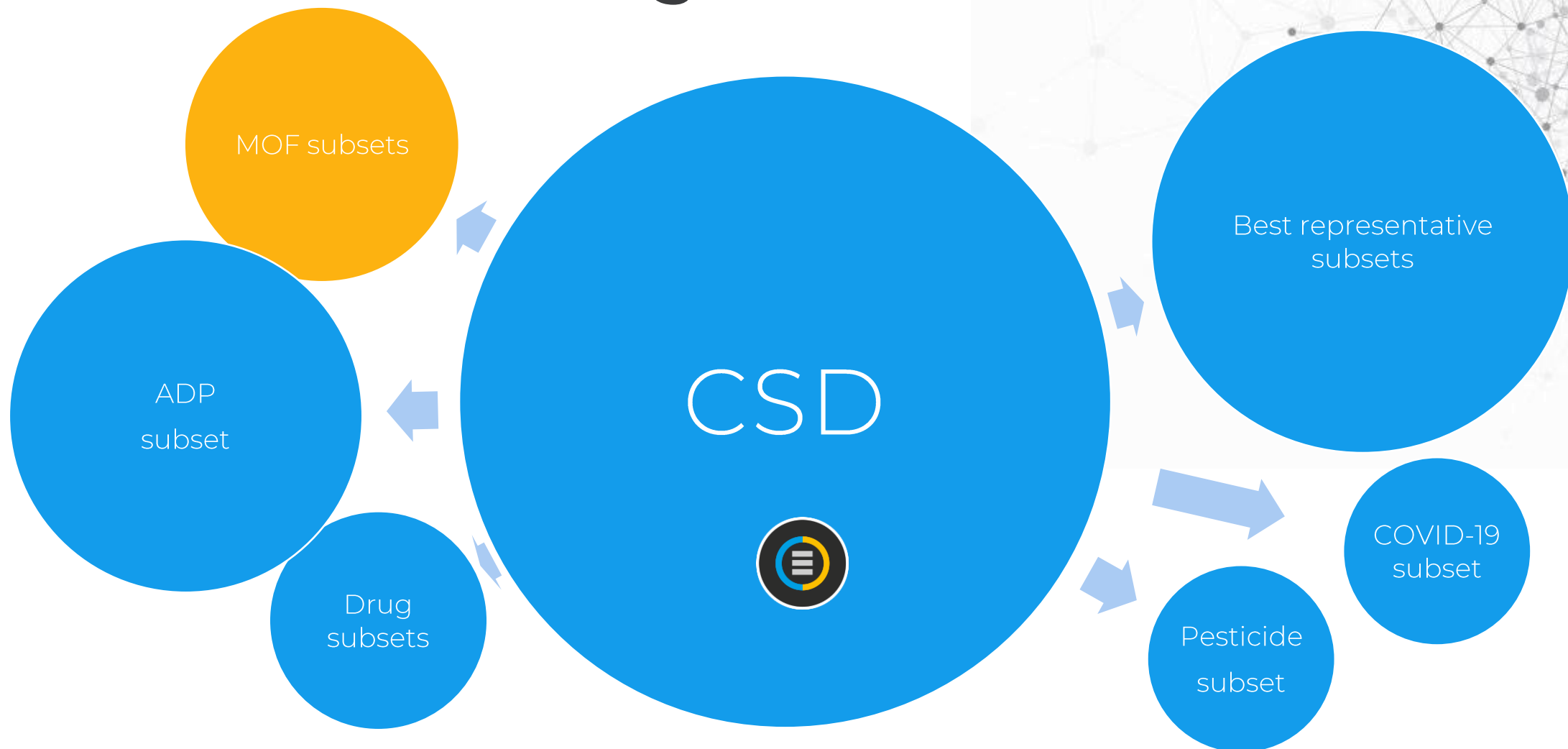
Search highlighted in 2D

Parameters displayed

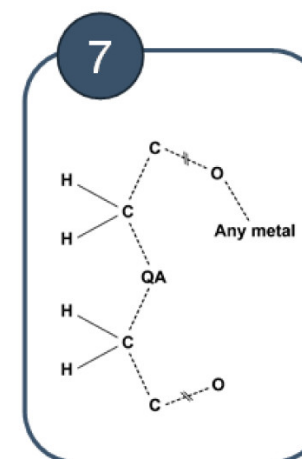
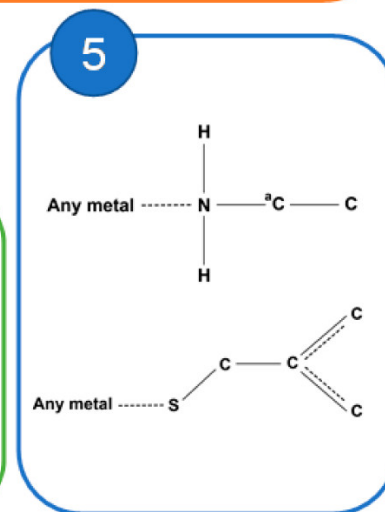
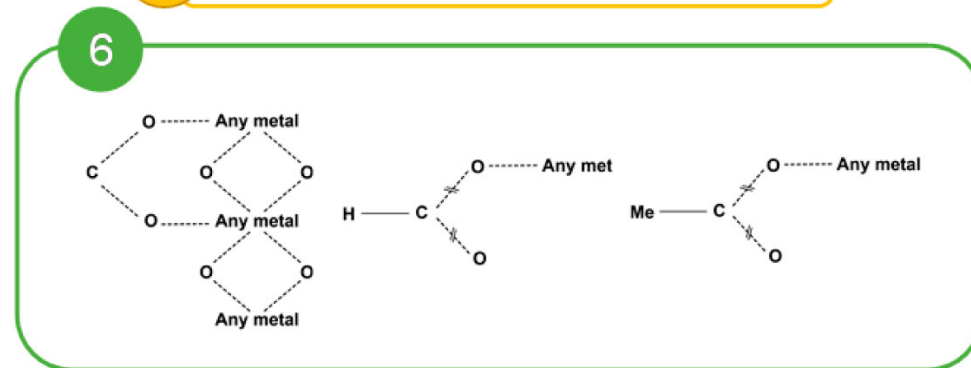
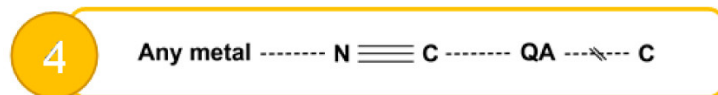
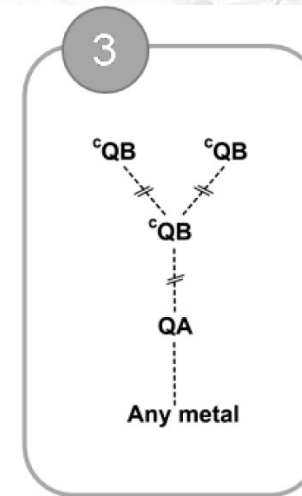
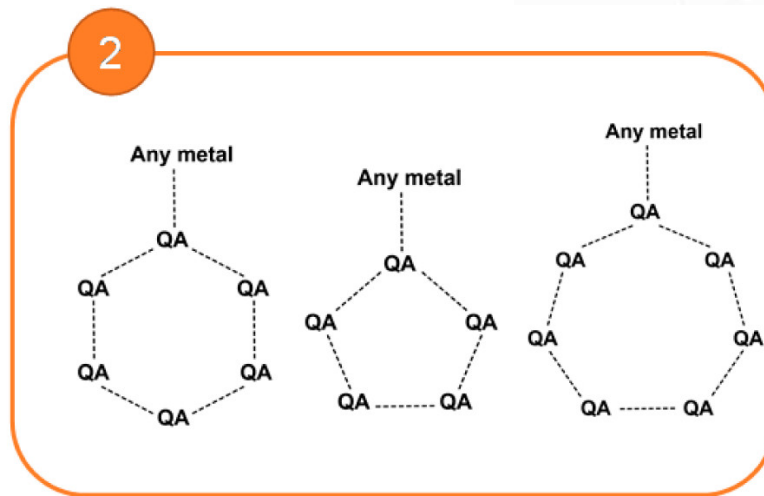
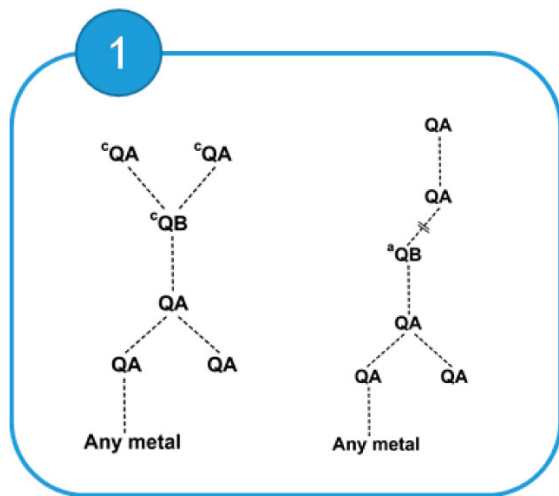
Data analysis



CSD subsets – targeted datasets



CSD MOF subset creation – Draw Queries



Using the subsets

CCDC ConQuest (1) : CSD_COVID-19_subset [Refcode List] - CSD_COVID-19_subset.gcd

File Edit Options View Databases Results Help

Build Queries

Available Databases...

Author(s)	Reference	Publication DOI	Deposition	Formula	Compound	Synonym	CCDC Class	Spacegroup	Cell	Reduced Cell
	Acta Crystallogr., Sect. E: Suppl.	10.1107/S1600536807020	CCDC 651371 ; IUCr FL212	$C_{22}H_{29}F_5O_5$	9 α -Fluoro-16 α -methyl-11 β ,20-dione	(11 β ,16 α)-9-Fluoro-11,17,21-trihydroxy-16-methylpregna-1,4-diene-3,20-dione; Decadron; Dexamethasone; DrugBank: DB00443; PDB Chemical Component code: DEX	Steroids	Name: P212121 Number: 19	a: 10.364(1) b: 16.157(1) c: 23.206(1) alpha: 90.00 beta: 90.00 gamma: 90.00 Volume: 3885.871	a: 10.364 b: 16.157 c: 23.206 alpha: 90.00 beta: 90.00 gamma: 90.00 Volume: 3885.871

Available Databases...

- Entries in CSD version 5.41 updates
- Entries in CSD version 5.41 (November 2019)
- Lists in CSD version 5.41 (November 2019)
- CSD Pesticide subset
- Best low temperature list
- CSD Drug subset
- CSD COVID-19 subset
- Best hydrogens list
- Best room temperature list
- Single-component CSD Drug subset
- MOF subset
- ADPs available subset
- Best R factor list
- Non-disordered MOF subset

DEXMET11

Analyse Hitlist

- ✓ CLQUON
- ✓ CLQUON01
- ✓ CORTSL
- ✓ COYRUD
- ✓ CPROMZ
- ✓ CPROMZ01
- ✓ CURHOM
- ✓ DEKSAN
- ✓ DEXMET11
- ✓ DIBBUL
- ✓ DIHSAO
- ✓ DIJWOJ
- ✓ DOHVED
- ✓ DOJZAD
- ✓ DOJZAD01
- ✓ DONRAB
- ✓ DUTRUF
- ✓ EJEQAL
- ✓ ESOURCE12
- ✓ FAHFAV
- ✓ FINWEE
- ✓ FIXGAU

121 hits

100%

Detach

Search Setup

Search Name: search4

Available Databases: ☐ Show Updates separately

☒ CSD version 5.42 (November 2020) + 1 update
☐ CSD version 5.42 (November 2020)

You can search complete database(s) or a subset (e.g., hits found in a previous search)

Single query being used. Search will find structures: where this query is true:

Filters **Advanced Options**

☐ 3D coordinates determined

☐ R factor ☒ <= 0.05
☐ <= 0.075
☐ <= 0.1

☐ Only ☒ Non-disordered
☐ Disordered

☐ No errors

☐ Not polymeric

☐ No ions

☐ Only ☒ Single crystal structures
☐ Powder structures

☐ Only ☒ Organics
☐ Organometallic

Restrict Search

Restrict Search by Refcode

Current Restrictions

Type: Refcode List Title: unknown Percent: Unknown

Time: Thu Feb 18 11:16:17 2021

File: ...subsets\Non-disordered_MOF_subset.gcd

Database CSD version 5.42 updates (Feb 2021)
1421 Refcodes

Database CSD version 5.42 (November 2020)
80303 Refcodes

Database CSD version 5.42 (November 2020)
0 Refcodes

Entries not found
0 Refcodes

☐ Entries in a hitlist loaded this session

☒ Entries in a pre-defined hitlist

☐ Entries in saved search file

☐ Entries in saved hitlist combination

☐ Entries in saved refcode list

Restrict based on pre-defined hitlists

Non-disordered MOF subset

100%

CSD Version... Select to restrict based on CSD version

Agenda

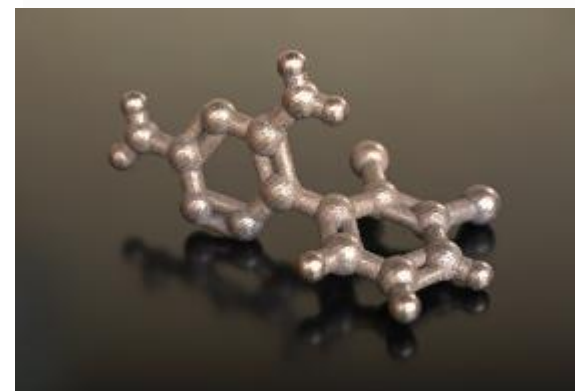
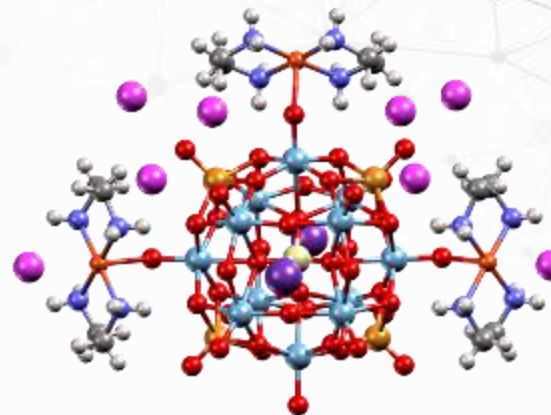
- 10 mins: Introduction
- 15 mins: Show One: ConQuest overview
- 20 mins: *Show One*: Metal-organic demo
- 25 mins: *Try One*: Hands-on examples
- 10 mins: *Explore More*: Advanced functionality for metal-organics
- 5 mins: *Explore More*: Quiz
- 5 mins: Wrap up and summary



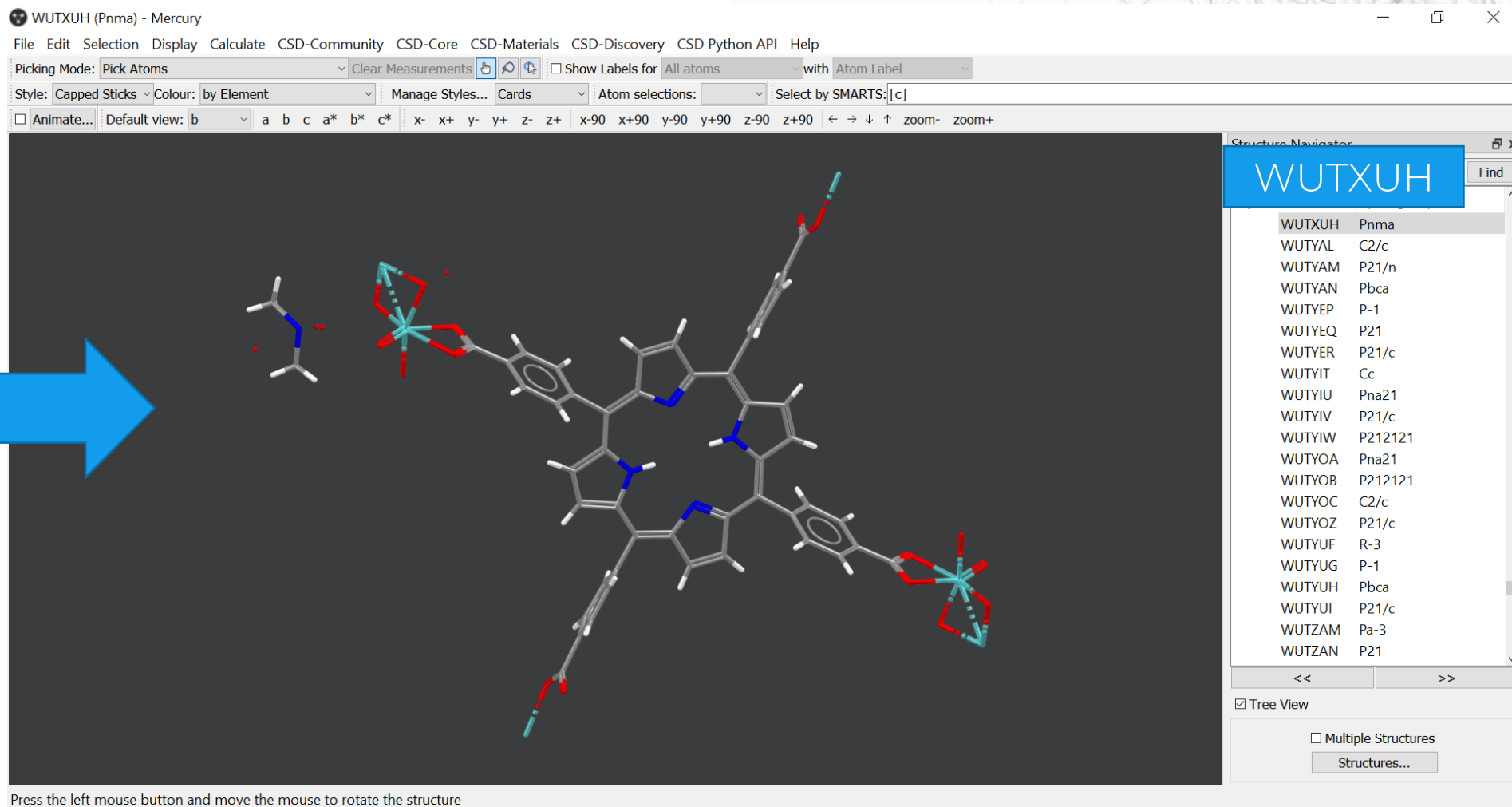
Mercury – Structure visualisation

With Mercury you can:

- Explore crystal structures, molecular conformations, crystallographic planes and simulated morphologies
- Generate high quality structural images for effective scientific communication
- Output model files for 3D printing
- Analyse geometries, interactions and the packing inside structures



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



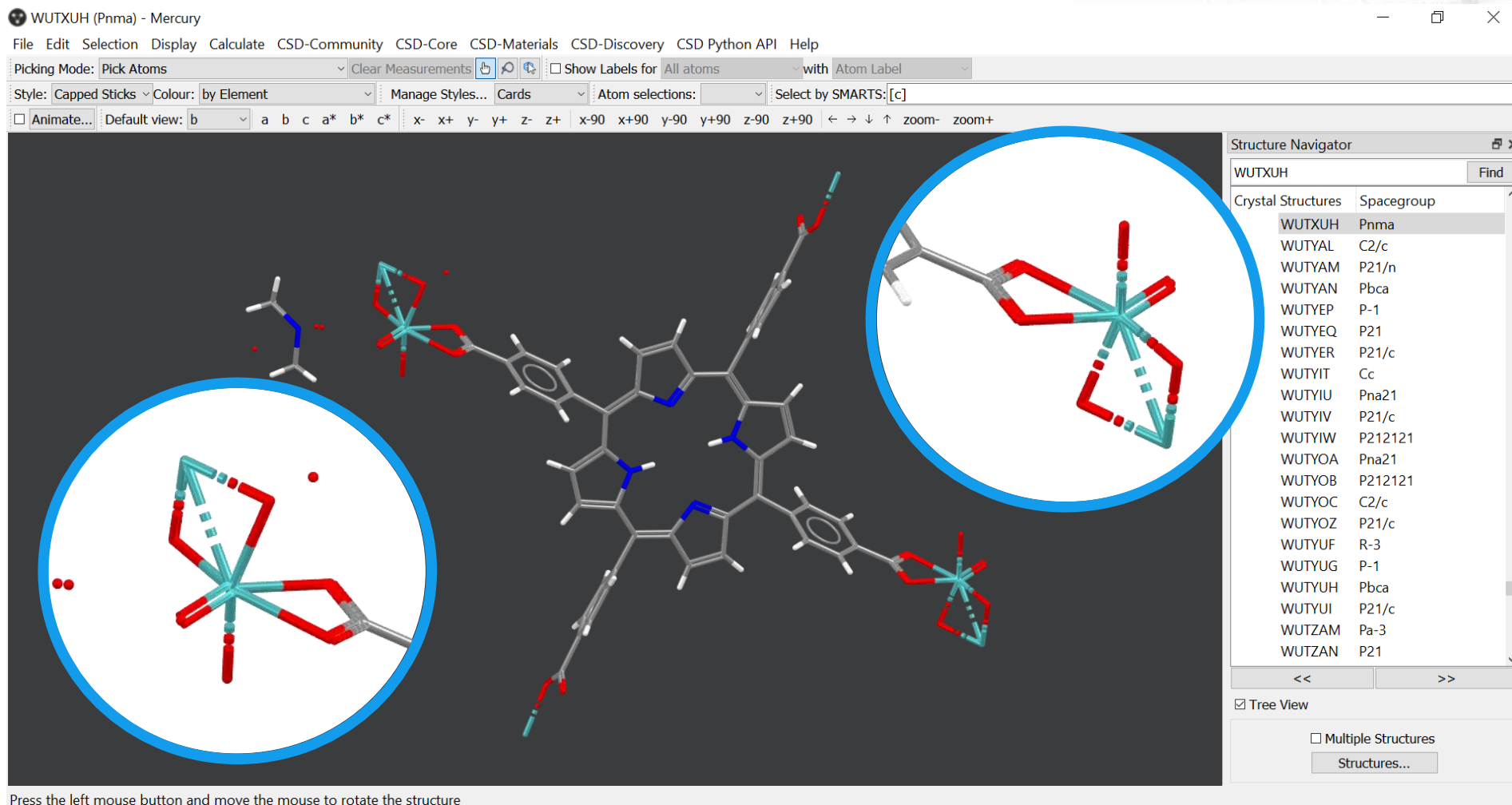
The basics

- **Left mouse** button and move – allows you to rotate structure
- **Middle Mouse** button and move – allows you to move structure
- **Right mouse** button and move – allows you to zoom in and out of structure
- **Structure Navigator** – Allows you to view any structure in the CSD



Structure Navigator	
Type in a refcode	
Find	
Crystal Structures	Spacegroup
AABHTZ	P-1
AACANI10	P21/c
AACANI11	P21/c
AACFAZ	Pbcn
AACFAZ10	Pbcn
AACMAL	P21/c
AACMHX10	Pbca
AACRHA	Pncm
AACRHC	P-1
AACRUB	Cc
AACRUB01	C2/c
AADAMC	P21/c
AADMPY	P-1
AADMPY10	P-1
AADRIB	P21
AAGAGG10	P212121
AAGGAG10	P21

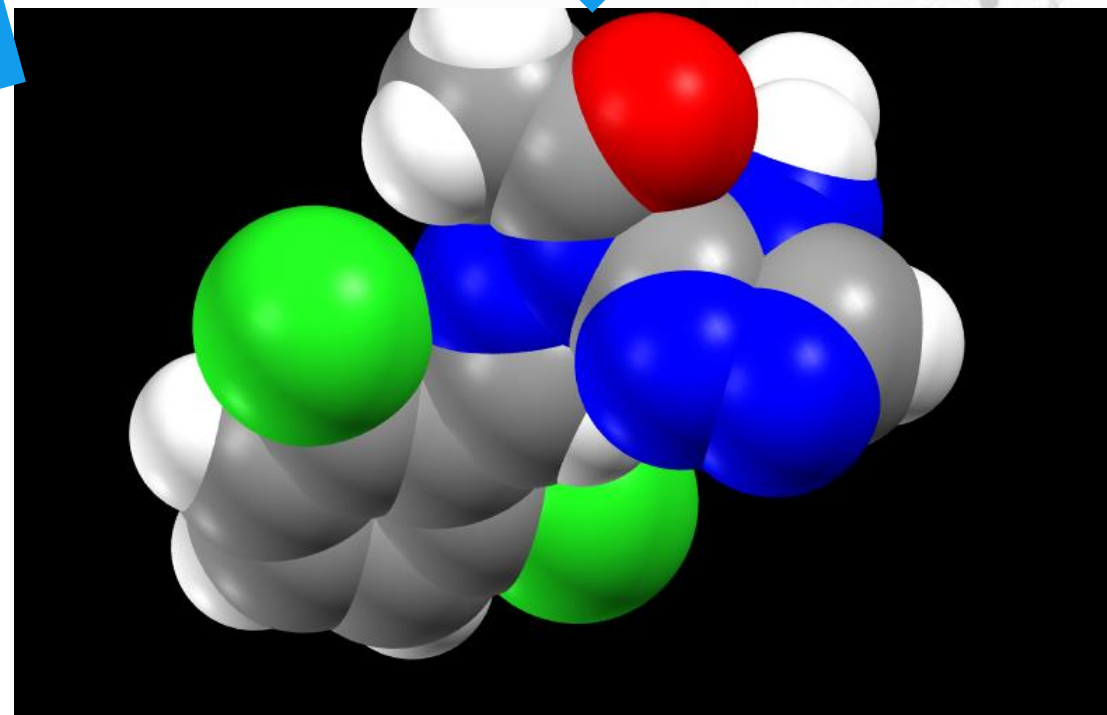
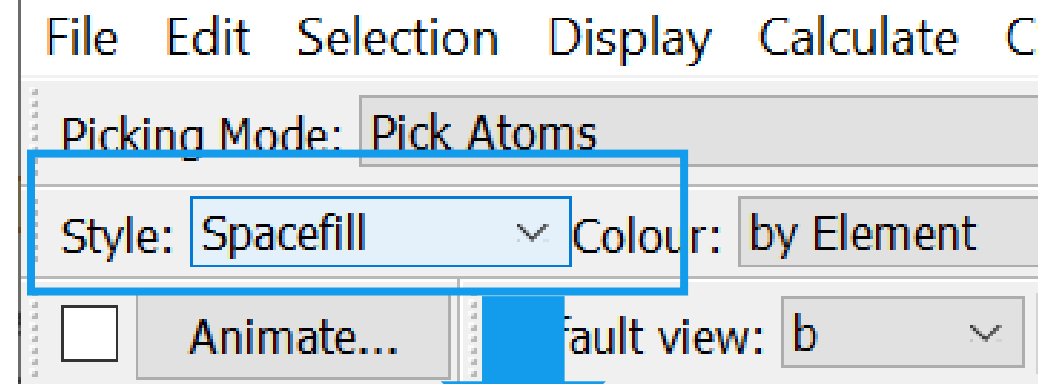
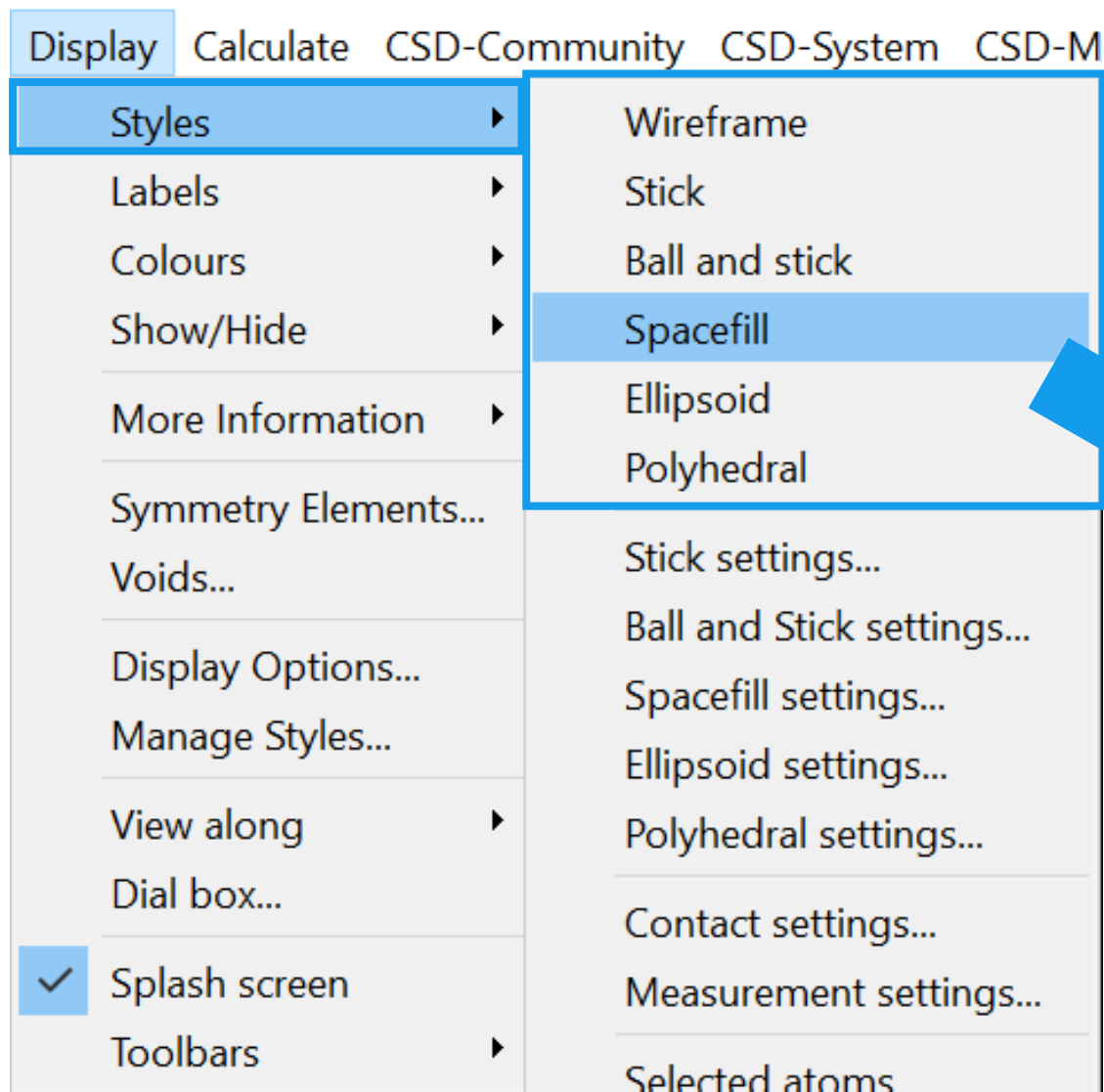
Polymeric bonds representations



Press the left mouse button and move the mouse to rotate the structure

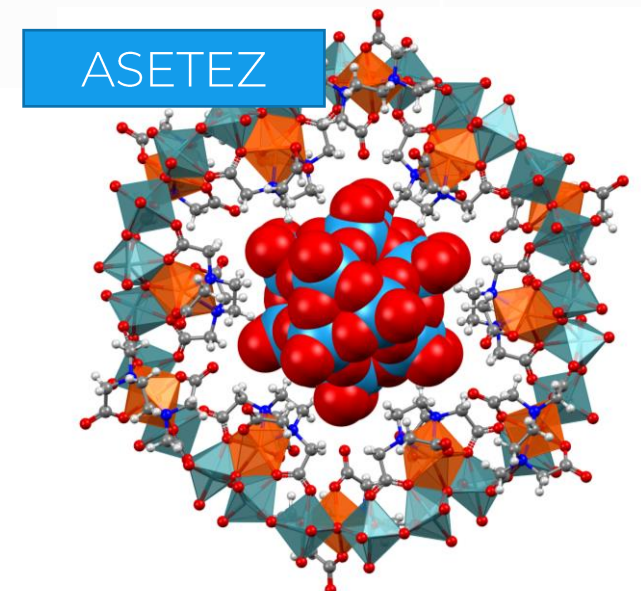
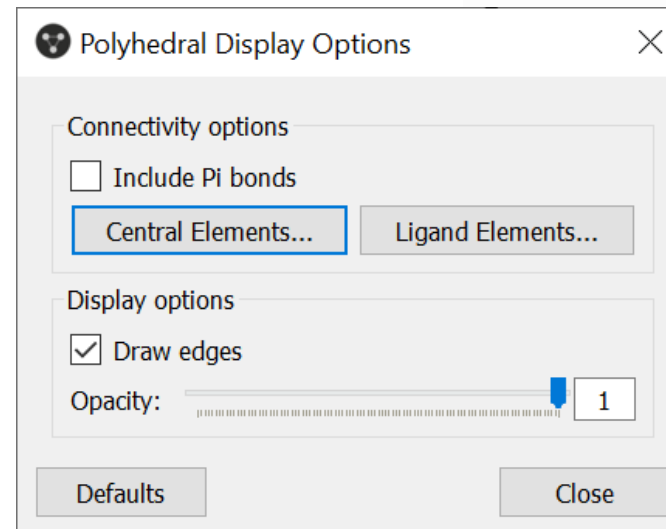
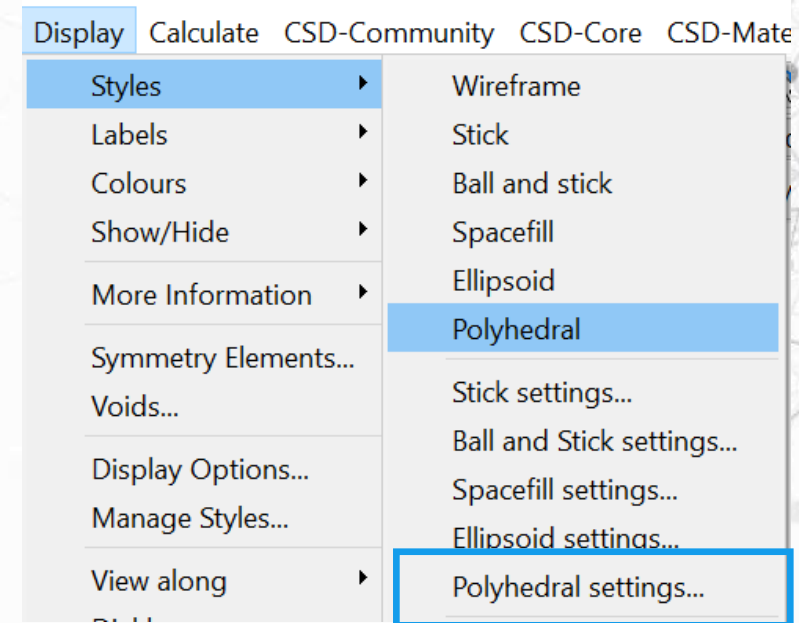
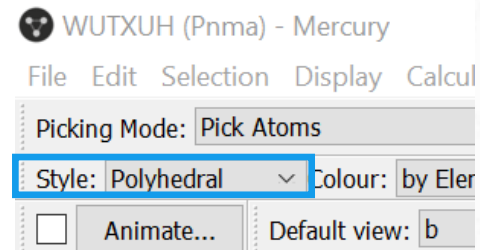
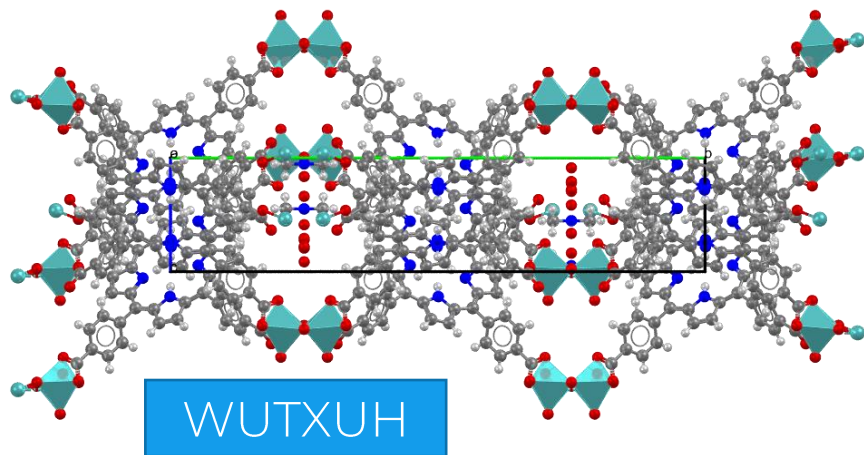
In Mercury,
polymeric
bonds are
drawn
alternating
long and
short lines.

Changing display - Style



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

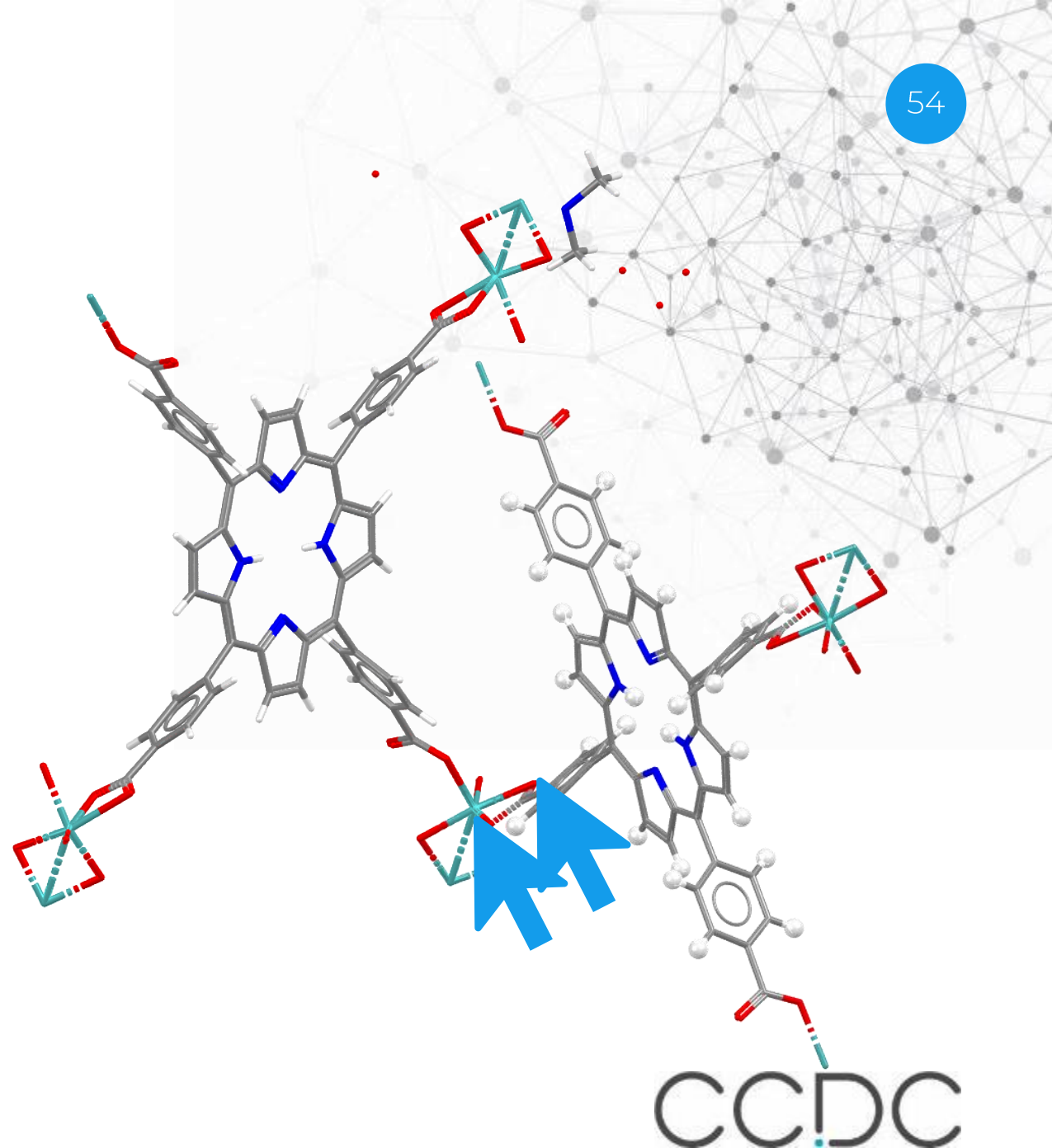
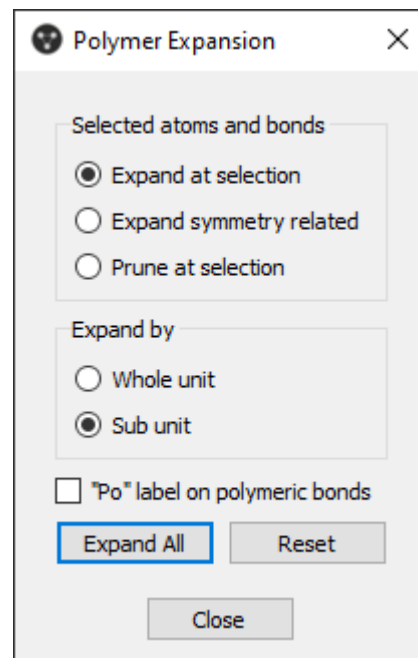


Polymer Expansion

Edit

Undo: Deselect All	Ctrl+Z
Redo: Not Available	Ctrl+Y
Copy Image to Clipboard	Ctrl+C
Edit Structure...	
Auto Edit Structure...	
Normalise Hydrogens...	
Polymer Expansion...	
Transform Molecules...	
Change Spacegroup Setting...	
Invert Structure	
Change Spacegroup to Subgroup...	

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



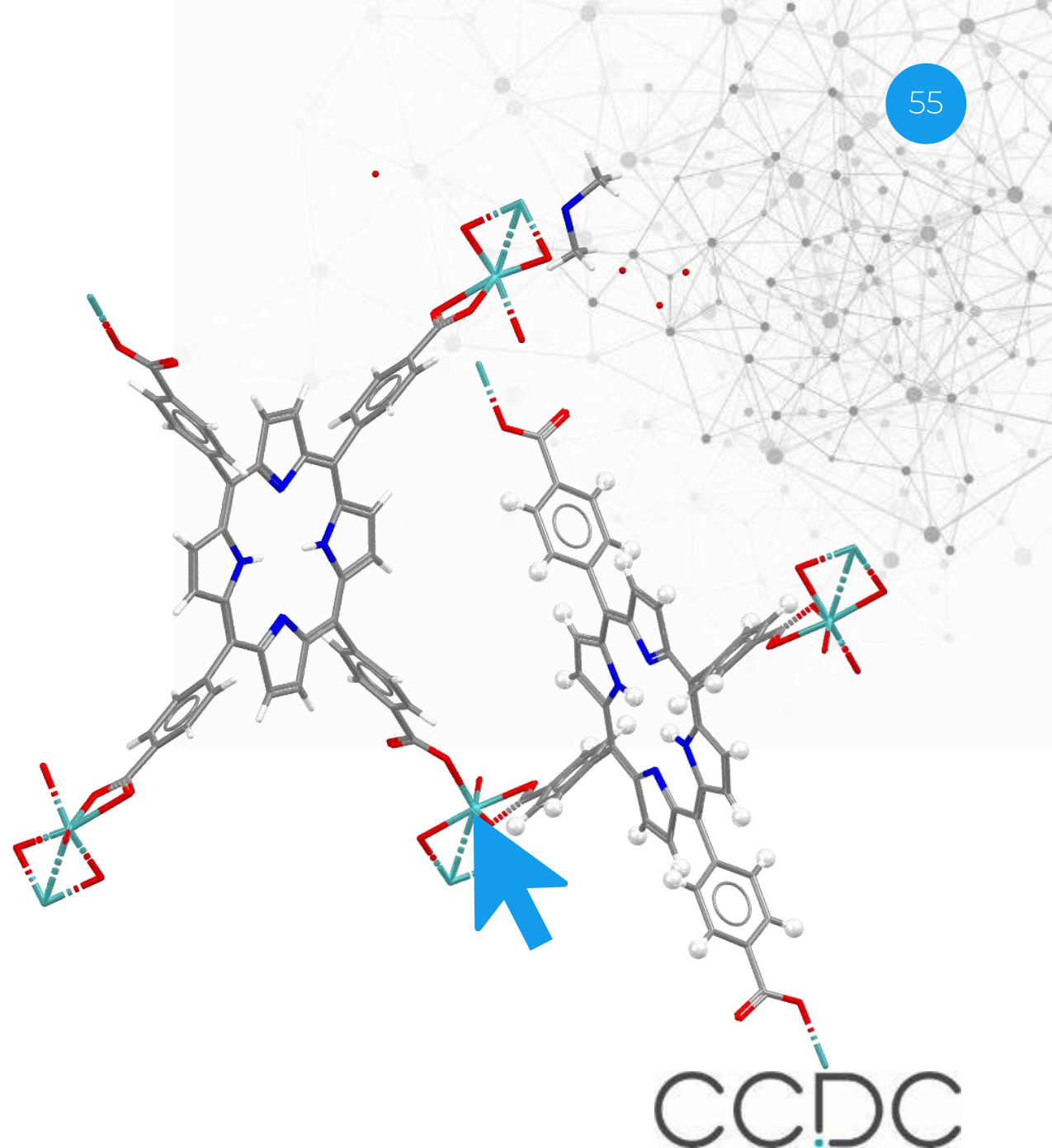
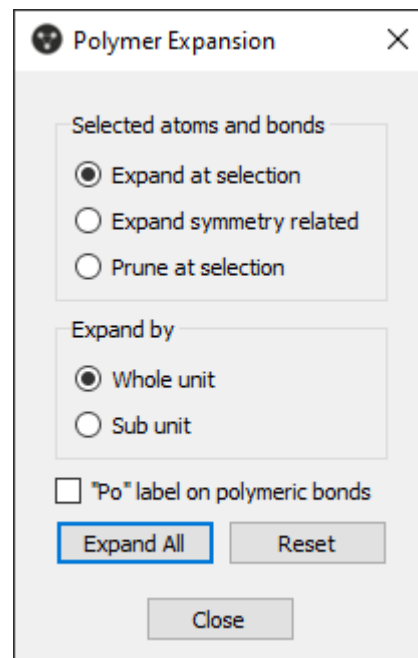
Polymer Expansion

Edit

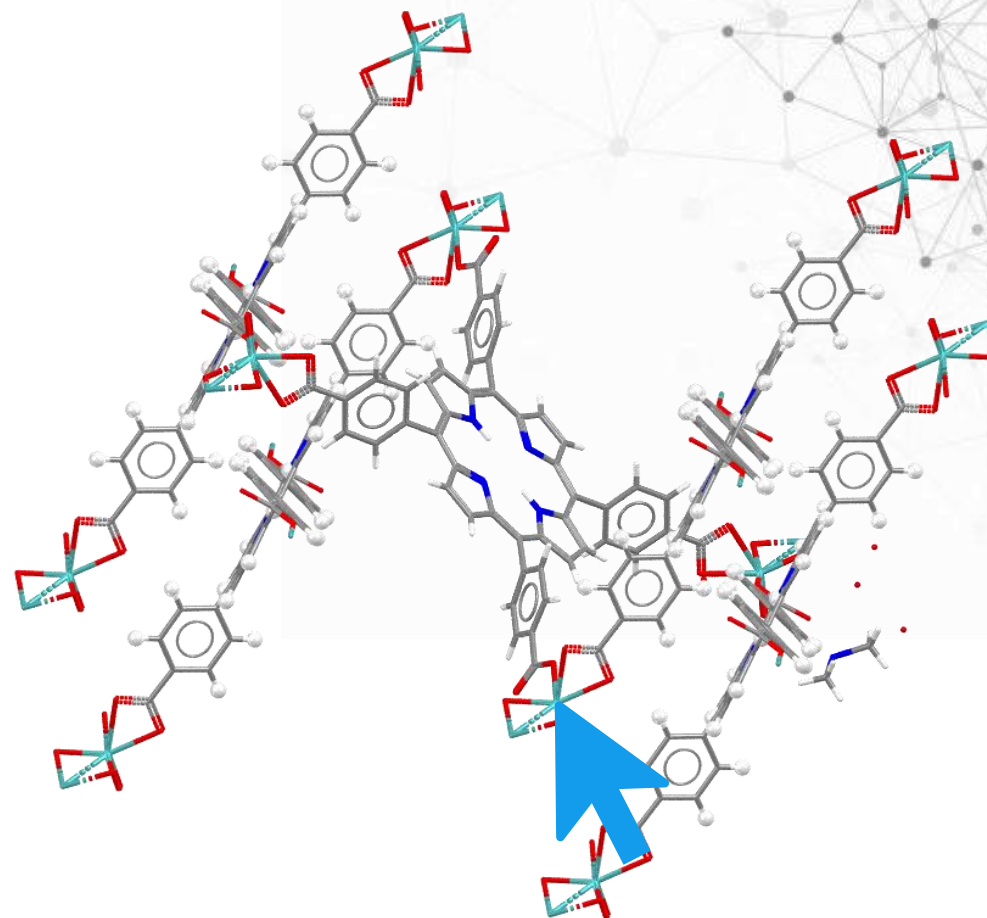
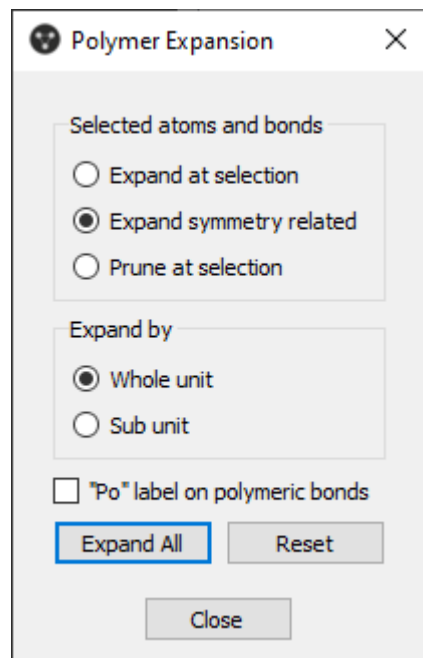
Undo: Deselect All	Ctrl+Z
Redo: Not Available	Ctrl+Y
Copy Image to Clipboard	Ctrl+C
Edit Structure...	
Auto Edit Structure...	
Normalise Hydrogens...	
Polymer Expansion...	
Transform Molecules...	
Change Spacegroup Setting...	
Invert Structure	
Change Spacegroup to Subgroup...	

By **Whole unit**

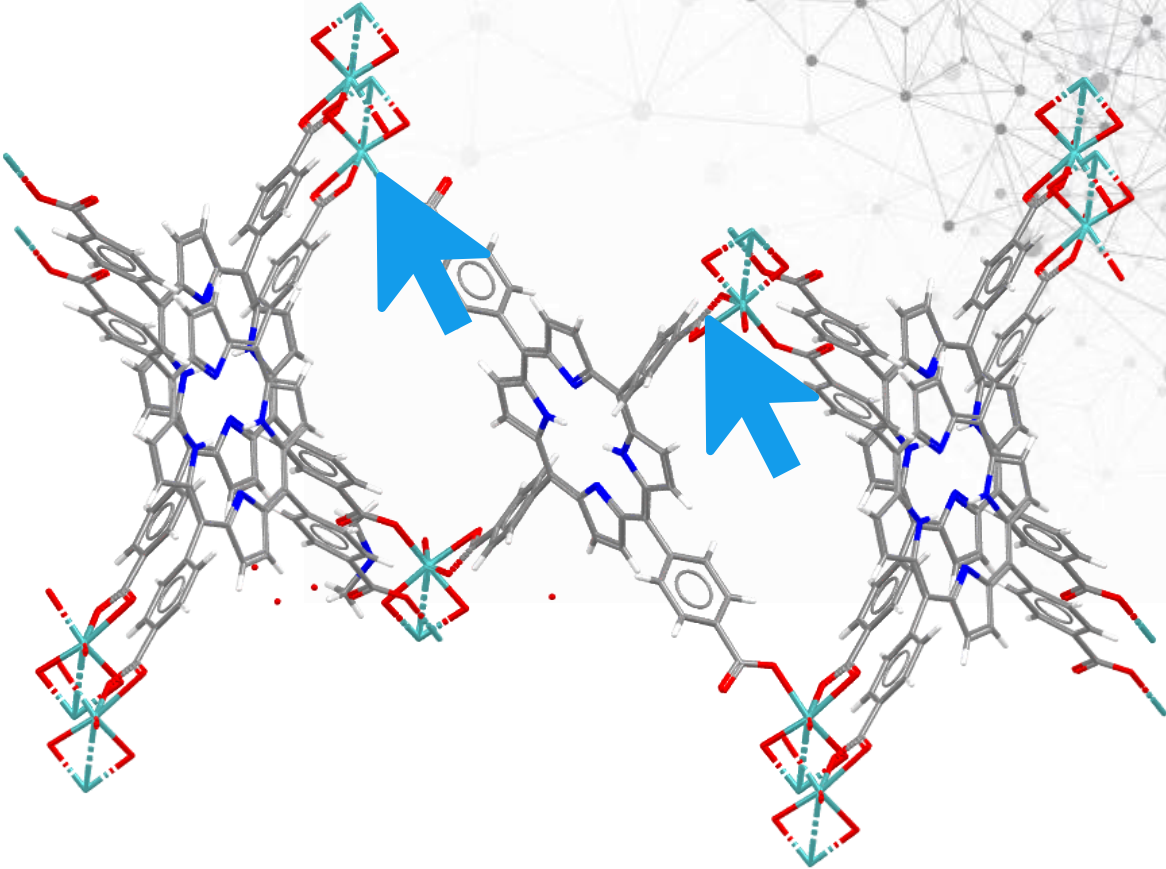
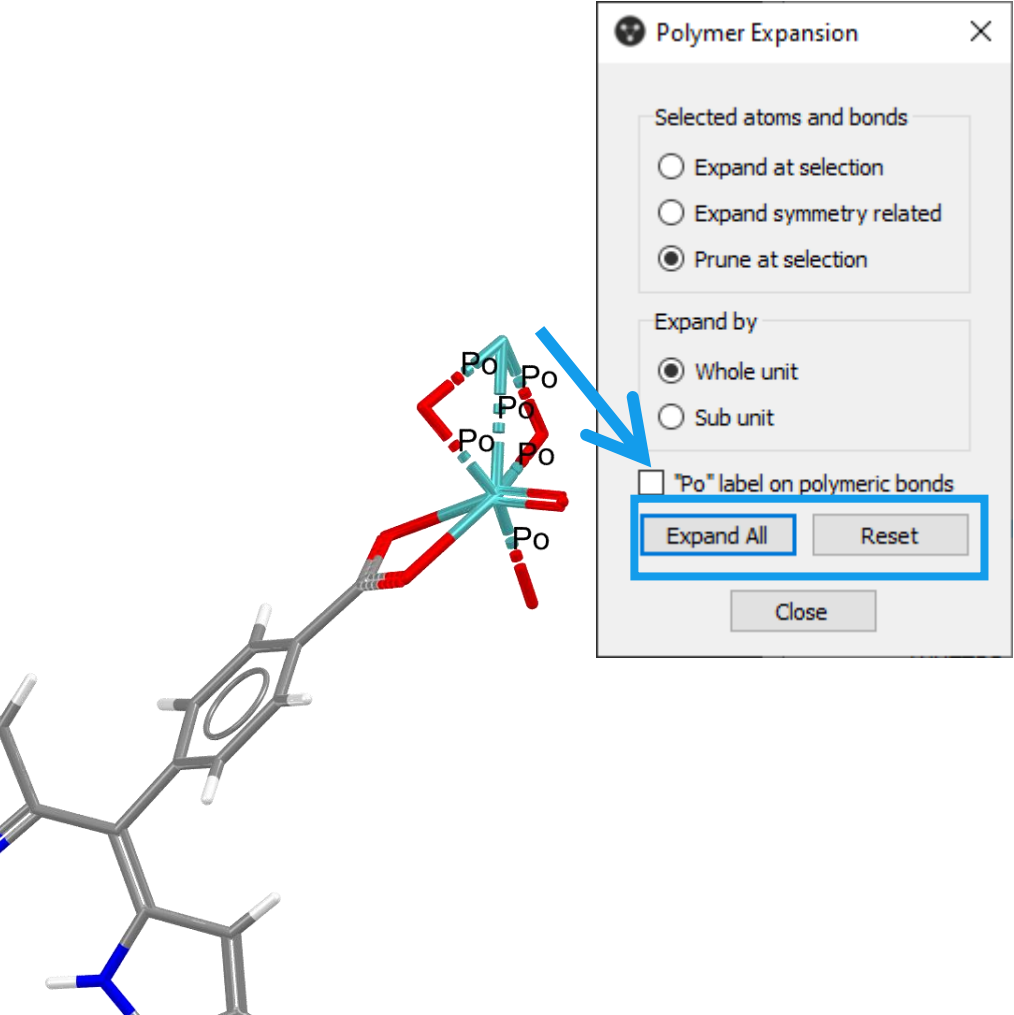
It will add an additional repeat of the initial crystal chemical unit to the structure



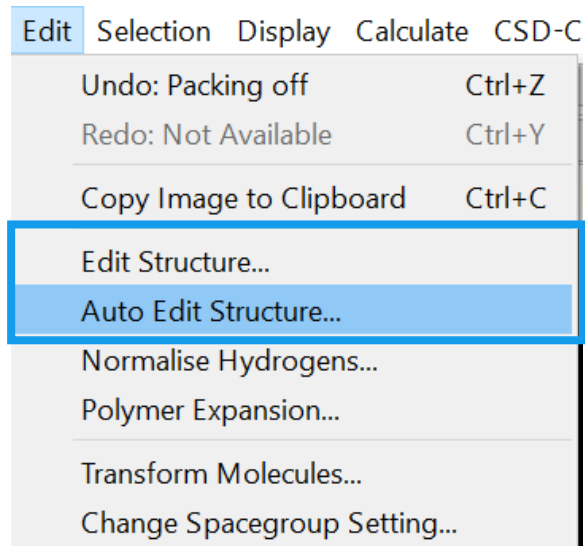
Polymer Expansion



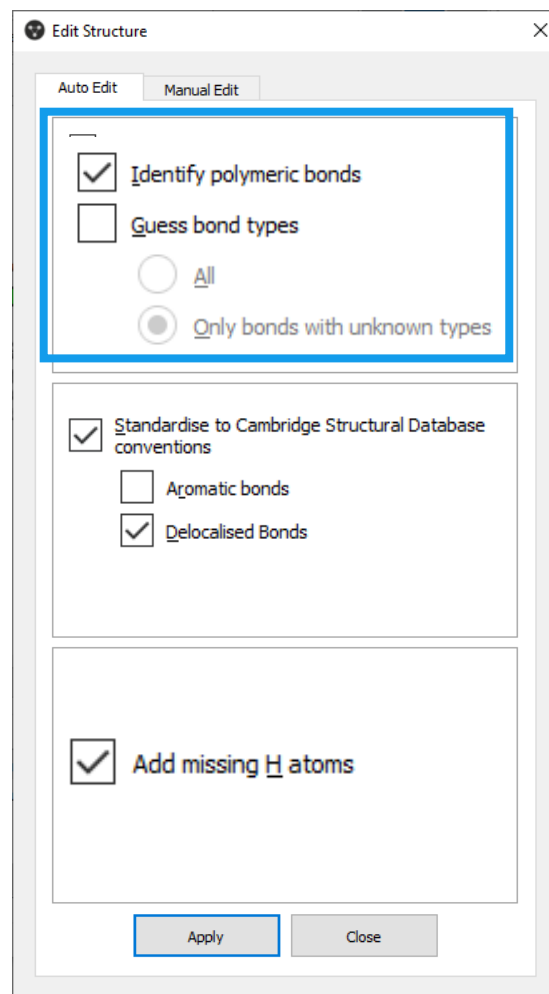
Polymer Expansion



What about editing polymers?



CSD structures have curated chemical representations including 3D bond types, 2D chemical diagrams and compound names



Reminder: you can save your Mercury display by going to File > Save As > File Type: Mercury compressed XML crystal files

Reminder: you can create high resolution images for presentations and publications by going to File > POV-Ray image... > Render

Exploring voids

Voids

Find any empty spaces (**voids**) in crystal unit cells that are big enough to hold a spherical "probe" of the given radius. Decrease the **Probe Radius** to find smaller spaces. Decrease the **Grid Spacing** to create smoother surfaces. To see voids in more than one unit cell, use the **Packing/Slicing** dialog to turn on packing and increase the ranges along a, b and c.

☒ Show

Probe Radius: Å

Approx. Grid Spacing: Å

Calculate using the

Display Options

Outside Colour:

Inside Colour:

Results

Volume % of unit cell volume

Volume Å³

HXACAN (Pcab) - Mercury

File Edit Selection **Display** Calculate CSD-Community CSD-Cgre CSD-Materials CSD-Discovery CSD Python API Help

Picking Mode:

Style:

☒ Splash screen

Display Options

Display

☒ Packing ☐ Short Contact < (sum of vdW radii)

☐ Asymmetric Unit ☐ H-Bond Default definition

☐ Auto centre

Options

☒ Show hydrogens ☐ Depth cue

☒ Show cell axes ☐ Z-Clipping

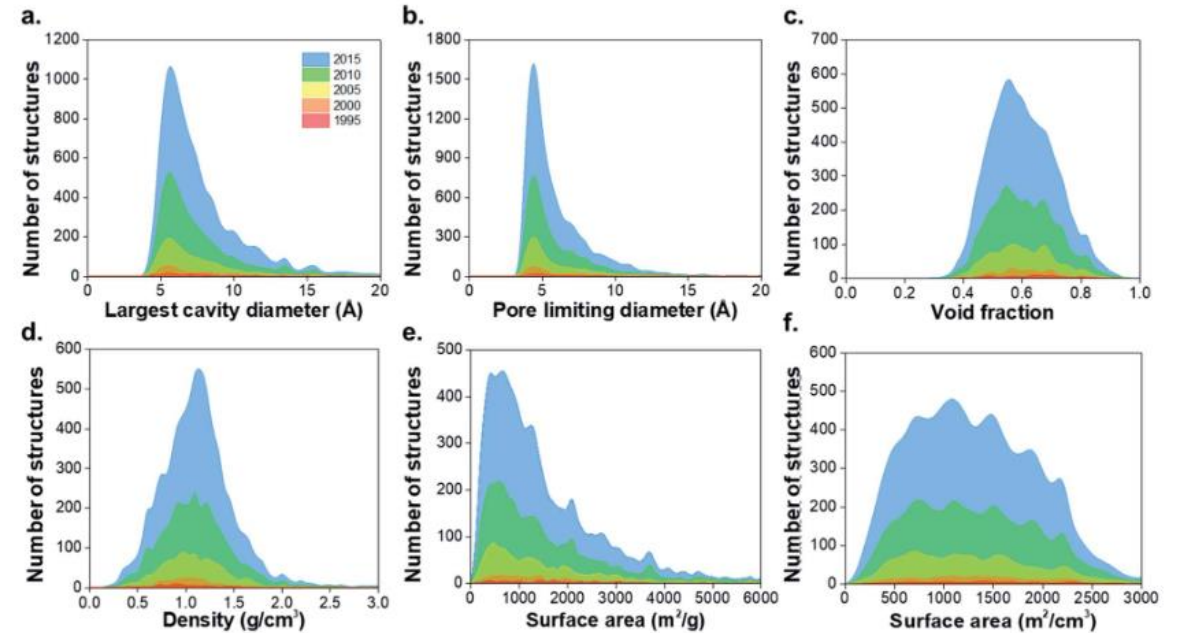
☐ Label atoms ☐ Stereo

Press the left mouse button and move the mouse to rotate the structure

Analysing MOFs

Aspects of MOF research are well-suited to high-throughput analysis

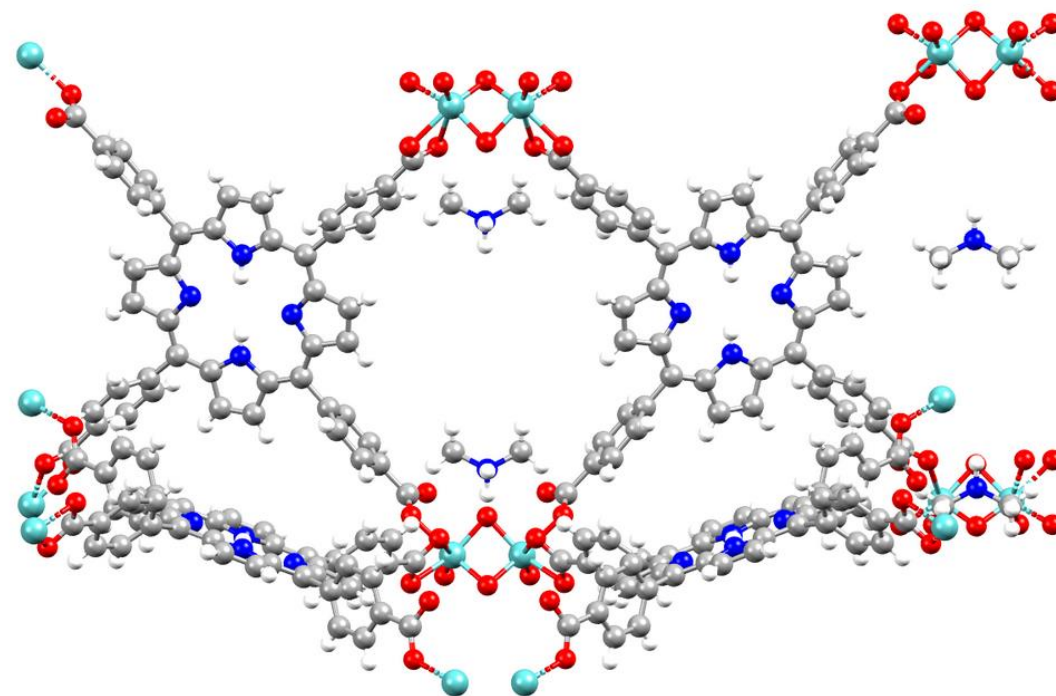
- Topology
- Pore size
- Void space
- Surface area



CSD MOF Collection

- Complements our CSD MOF subsets embedded in CSD Portfolio
- A good starting point for high-throughput analysis
- 10,636 3D porous MOFs
 - All 3D MOFs containing over 10 % void space
 - Converted to space group P1
 - Non-bonded solvent removed
 - Missing hydrogen atoms added

MOF Collection Entry



CSD MOF Collection

- The CSD MOF Collection contains reference spreadsheets
 - Allows further categorisation within the collection

CSD refcode	CIF filename	original crystal system	Sohncke space group	Percentage void space	Charged framework	Hydrogen added	Unreliable chemistry
ABAVIJ	abavij_P1	monoclinic	-	15.538	-	-	-
ABAVOP	abavop_P1	monoclinic	-	13.946	-	-	-
ABAYIO	abayio_P1_charged	cubic	-	62.844	Yes	-	-
ABAZAF	abazaf_P1	monoclinic	-	21.272	-	-	-
ABAZAF01	abazaf01_P1	monoclinic	-	22.353	-	-	-

- The CSD MOF Collection has a Creative Commons licence: CC BY-NC-SA 4.0
- <https://www.ccdc.cam.ac.uk/Community/csd-community/csd-mof-collection/>

CSD Python API

The **CSD Python API**
(Application Programming Interface)

- Enables you to use many capabilities of the CSD-Core without being bound by graphical interfaces
- You can readily create CSD-driven analyses and workflows tailored to your needs
- Our Mercury/Hermes interface allows you to easily access your scripts and run analysis
- You can also access CSD Subsets and alter the MOF subset script to match your requirements



CSD Python API - Example

```
In [10]: from ccdc import io, diagram

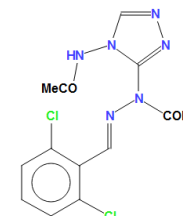
In [26]: import IPython.core.display
import StringIO

In [27]: # Set up CSD entry reader and find the first entry in the database
csd = io.EntryReader('csd')
csd_entry = csd[0]
csd_entry.identifier

Out[27]: u'AABHTZ'

In [30]: # Generate a diagram for that CSD entry
diagram_generator = diagram.DiagramGenerator()
diagram_generator.settings.font_size = 12
img = diagram_generator.image(csd_entry)

In [31]: # Display the 2D diagram
output = StringIO.StringIO()
img.save(output, "PNG")
contents = output.getvalue()
IPython.core.display.display_png(contents, raw=True)
```



Entry examples

Create indexes of useful information for subsets of CSD entries

Note that this script makes use of functionality from the `cookbook` utility module.

```
#!/usr/bin/env python
#
# This script can be used for any purpose without limitation subject to the
# conditions at http://www.ccdc.cam.ac.uk/Community/Pages/Licence/v2.aspx
#
# This permission notice and the following statement of attribution must be
# included in all copies or substantial portions of this script.
#
# 2015-06-17: created by the Cambridge Crystallographic Data Centre
#
'''
Provide information on a set of structures in the CSD.

This script takes as input a gsd file (a text file with CSD refcodes) and
writes out the identifier, author(s), literature reference, formula, compound
name and compound synonym(s). The output can be formatted as csv or html.
'''
from future__ import division, absolute_import, print_function
import six
import sys
import os
import csv
import html
import argparse
import codecs

from ccdc.io import EntryReader

class Writer(object):
    def __init__(self, infile, out, format='csv'):
        try:
            self.rdr = EntryReader(infile, format='identifiers')
        except RuntimeError:
            print('Failed to read input file %s!' % infile)
            exit(1)

        self.out = out
        getattr(self, format + '_header')()
        for e in self.rdr:
            getattr(self, format + '_line')(e)
            getattr(self, format + '_footer')()

    def csv_header(self):
        data = ','.join([
            'Identifier',
            'Author',
            'Literature Ref',
            'Formula',
            'Compound Name',
            'Synonyms'
        ])
        self.out.write(data + '\n')
```

Agenda

- 10 mins: Introduction
- 15 mins: Show One: ConQuest and Mercury overview
- 20 mins: *Show One*: Metal-organic demo
- 25 mins: *Try One*: Hands-on examples
- 10 mins: *Explore More*: Advanced functionality for metal-organics
- 5 mins: *Explore More*: Quiz
- 5 mins: Wrap up and summary

What have we learnt?

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.

Want to explore more?

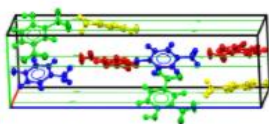
Training and Educational Resources

The wealth of information contained within the Cambridge Structural Database (CSD) extends far beyond a collection of crystal structures. Knowledge derived from these materials informs much of chemistry, biochemistry, and biology. Chemical and structural concepts are often difficult to grasp without real world, interactive examples for students to explore.

The CCDC and our colleagues continually produce educational materials for use in classroom and computer lab settings, or as independent study modules. Many of these materials make use of the Teaching Subset - a freely available set of over 750 structures that can be investigated with the free version of our Mercury visualisation and analysis program. Of course, our database of over one million entries are available for free through our Access Structures portal.

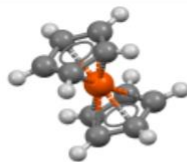
If you are an educator looking for supplementary teaching materials, find out more about the Teaching Database [here](#). If you have developed your own modules using the CSD and would like to share them with the broader community, please contact us at education@ccdc.cam.ac.uk.

To keep up to date with the latest news from education and outreach at the CCDC, sign up for the Education and Outreach Newsletter [here](#).



CSD-Core

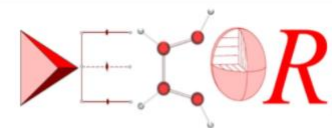
Essential crystallographic and structural chemistry capabilities.



Information on the Teaching Subset



Access a series of teaching modules for use in the classroom



DECOR: Educational Resources for Teaching Crystallography

Register for E&O newsletter

Self-guided workshops



Download a series of self-guided workshop materials for CCDC tools and features



Watch software training and support videos

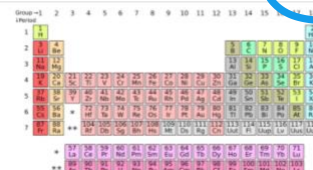


CSD University modules

On-demand modules with completion certificate



Access fun science activities for kids through the CCDC Home learning page



Explore the Periodic Table through Crystal Structures

YouTube and LabTube channels

CCDC

More opportunity to learn during the IUCr

- 14th August 19:00-19:50 CEST – [Ewald Prize lecture honouring Olga Kennard](#)
- 15th August 10:20-12:45 CEST - Social media and new frontiers for spreading crystallographic
 - [Talk: Using social media to share the wonders of crystallography - perspectives from a structural database](#)
- 15th August 14:45-17:10 CEST - Pre and post publication peer review of crystallographic data
 - [Talk: Crystallographic review - perspectives from a chemical structural database](#)
- 15th August 18:10-19:00 CEST – Keynote lecture: Crystal forms in pharmaceutical applications
- 16th August 14:45-17:10 CEST - Exemplary practice in chemical, biological and materials database archiving
 - [Talks from the CSD, ICSD, ICDD and PDB](#)
- 16th August 14:45-17:10 CEST - New trends in pharmaceutical materials
 - [Talk: “Particle Informatics”: Evolving methods for understanding particle properties](#)
- 21st August 10:20-12:45 CEST - Crystallography schools to promote interdisciplinarity in science
 - [Talk: Crystallography for all – Using the CSD to help promote interdisciplinarity in science](#)

More chance to interact with us

As well as talks we have:

- Two posters
 - 16th August 17:10-18:10 CEST: Metadata for better data - Growing and improving the Cambridge Structural Database
 - 15th August 17:10-18:10 CEST: Come for the drug, stay for the solvent!
- A virtual Exhibition
 - CCDC has a virtual stand!
 - Drop in to say hi
 - Get your name on our virtual leaderboard

