ConQuest to Mercury: From Searching to Data Analysis

CCDC Virtual Workshop

October 2024

Guest Speaker:

Catherine Esterhuysen

Professor of Physical Chemistry

Stellenbosch University

Stellenbosch University
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

CCDC
advancing structural science

Learning outcomes for today

- How to construct a CSD search using ConQuest.
- How to define 3D parameters in ConQuest searching.
- How to export ConQuest results to Mercury
- How to analyse results using the Data Analysis module.
- Tips and tricks for effective searching in ConQuest .
- How this functionality has been used by global researchers through the exploration of recent case studies.

CCDC

Agenda

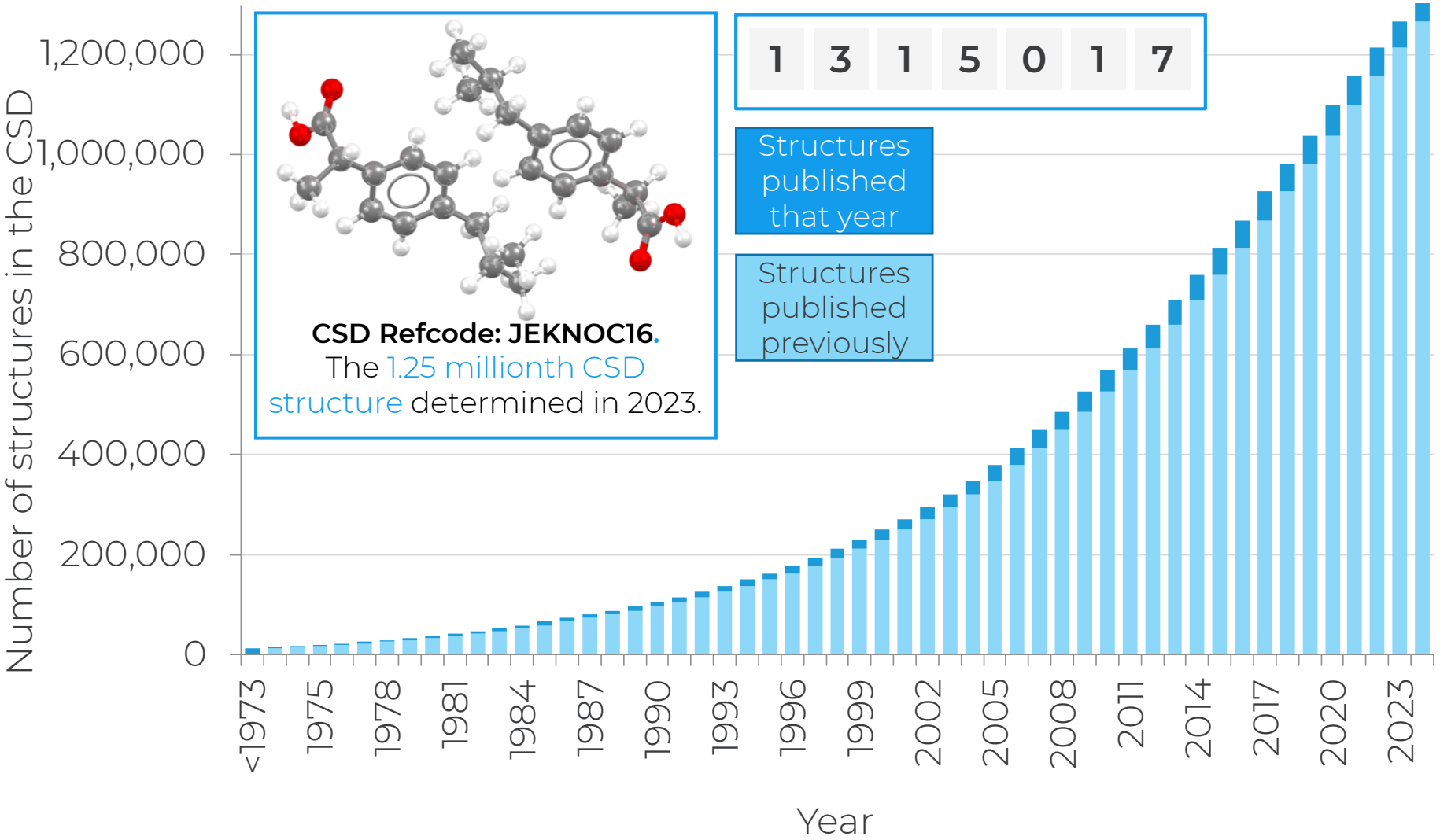
We will make the recording available to you in the next few days.

- Introduction to the CSD
- *Show One:* Research exploring Hydrogen and Halogen bonding
- *Show One:* Introduction and demonstration of ConQuest and Mercury
- *Try One:* Hands-on examples
- *Explore More:* More advanced tips and tricks and case-studies
- *Explore More:* Quiz and Summary
- *Extra time:* More time for hands on and Q&A

After the session you can earn a completion certificate for today by taking the test.

CCDC

The Cambridge Structural Database



- Every published structure
  - Inc. ASAP & early view
  - *CSD Communications*
  - Patents
  - University repositories
  - Thesis
- Every entry enriched and annotated by experts
- Discoverability of data and knowledge
- Sustainable for over 59 years
- A trusted CoreTrustSeal repository

Certified as Trustworthy by CoreTrustSeal

Inside the Cambridge Structural Database

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

Organic
45%

Metal-Organic
55%

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

Not Polymeric
89%

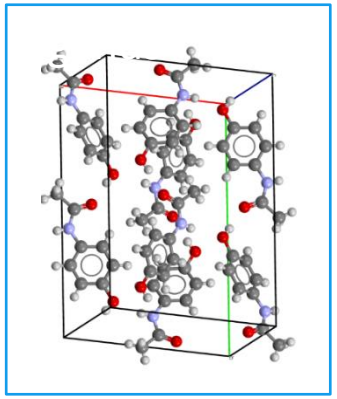
Polymeric: 11%

Single Component
58%

Multi Component
42%

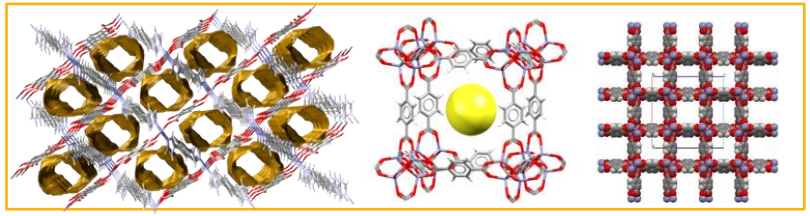
Organic

- Drugs
- Agrochemicals
- Pigments
- Explosives
- Protein ligands



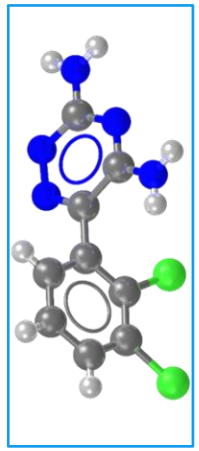
Metal-Organic

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



Additional data

- 13,478 polymorph families
- 174,987 melting points
- 1,075,904 crystal colours
- 951,746 crystal shapes
- 30,275 bioactivity details
- 13,641 natural source data
- > 350,000 oxidation states



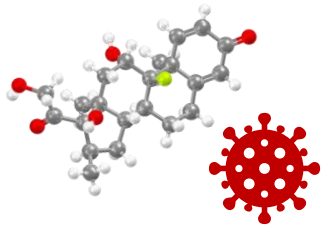
Links and subsets

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

CSD Subsets

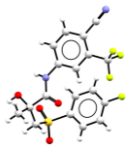
Groups of structures that may be more difficult to find in CSD from searching alone

COVID-19

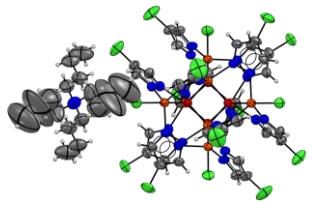


Pesticides

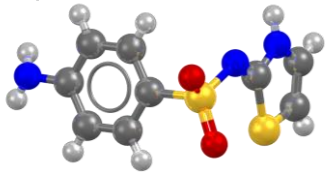
Drugs



ADPs



Best representative

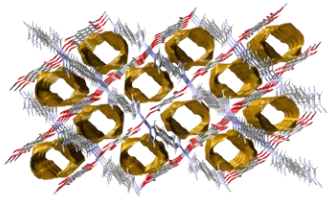


Teaching

MOF Dimensionality

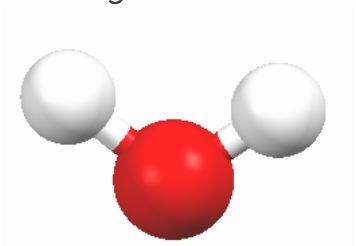
- 1D
- 2D
- 3D

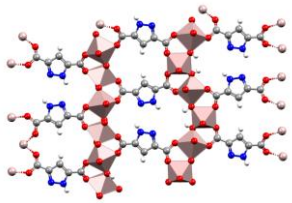
MOFs



- Easy access to the most relevant structures.
- Using our in-house and external expertise.
- Convenient starting point for research and analysis.

Hydrates

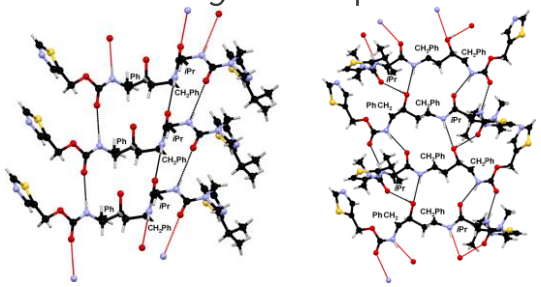




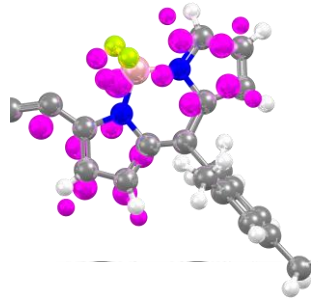
Electron Diffraction

High Pressure

Polymorphs

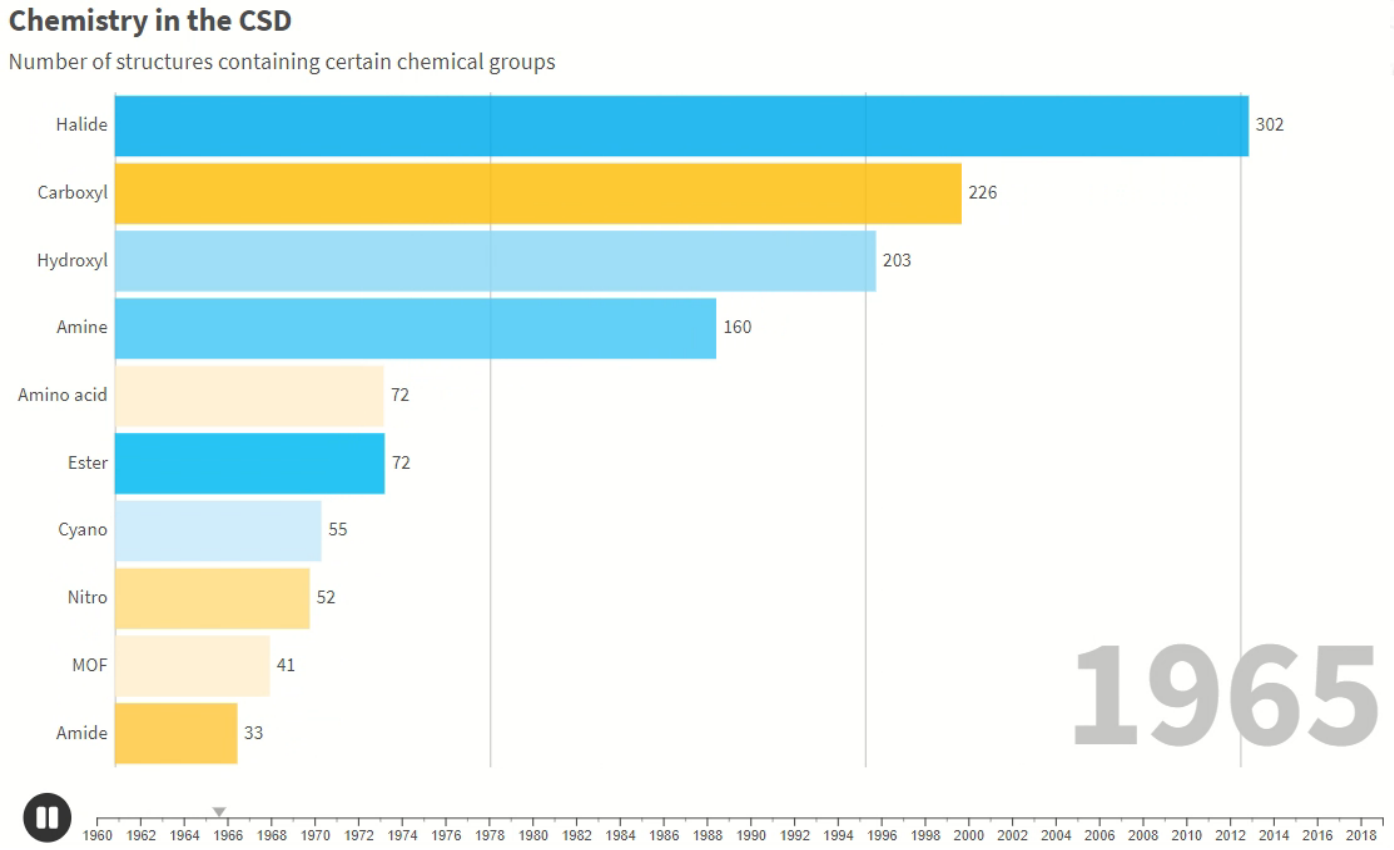


Disorder



Exploring the CSD

- > 1 million structures
  - > 94M 3D coordinates
- > 28 million bond lengths
  - > 2M unique distributions
- > 40 million valence angles
  - > 3M unique distributions
- > 14 million torsion angles
  - > 800K unique distributions
- > 2 million rings
  - > 400K unique distributions
- > 2 million hydrogen bonds
  - >30 million Isostar contacts



Images and graphics created using the CSD Python API and Flourish

CCDC

The CSD Portfolio

CSDCore.

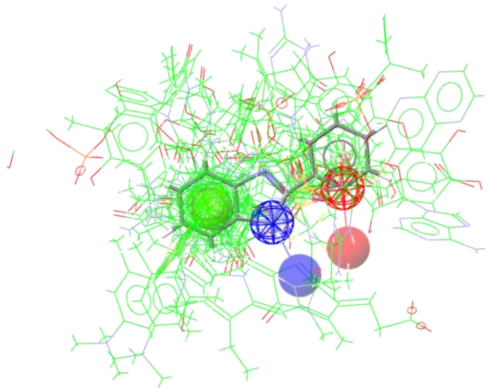
ConQuest

Mercury

CSDEnterprise.

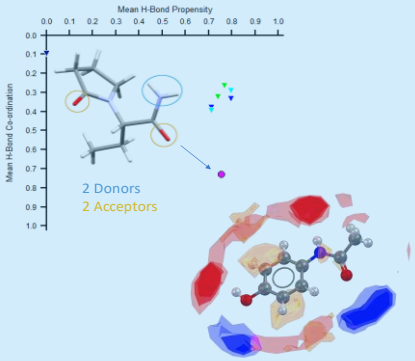
CSDDiscovery.

Design of new molecules



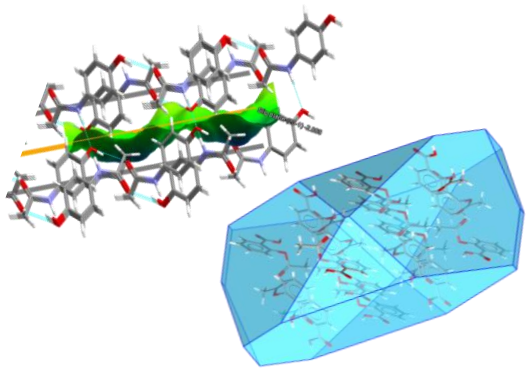
CSDMaterials.

Assessment of solid form stability and properties



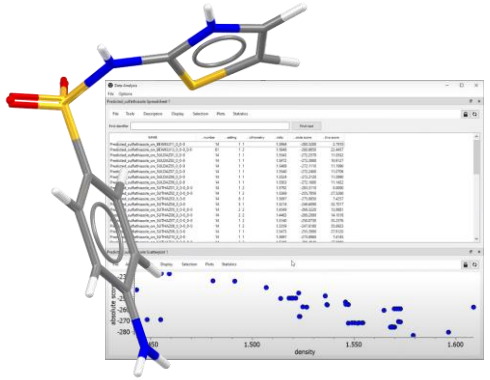
CSDParticle.

Anticipate particle properties and behaviour



CSDTheory.

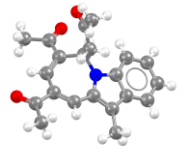
Generate solid form landscapes

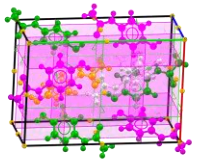


CSDCommunity.

Deposit, publish, access and visualise structural data

Free functionality to share and learn from structures





CSDU

Medicinal & Computational Chemists ♦ Crystallographers & Structural Biologists ♦ Solid Form & Crystallisation Scientists ♦ Functional Materials Scientists ♦ Educators ♦ Industry and Academia

CCDC

Agenda

We will make the recording available to you in the next few days.

- Introduction to the CSD
- *Show One:* Research exploring Hydrogen and Halogen bonding
- *Show One:* Introduction and demonstration of ConQuest and Mercury
- *Try One:* Hands-on examples
- *Explore More:* More advanced tips and tricks and case-studies
- *Explore More:* Quiz and Summary
- *Extra time:* More time for hands on and Q&A

CCDC

Agenda

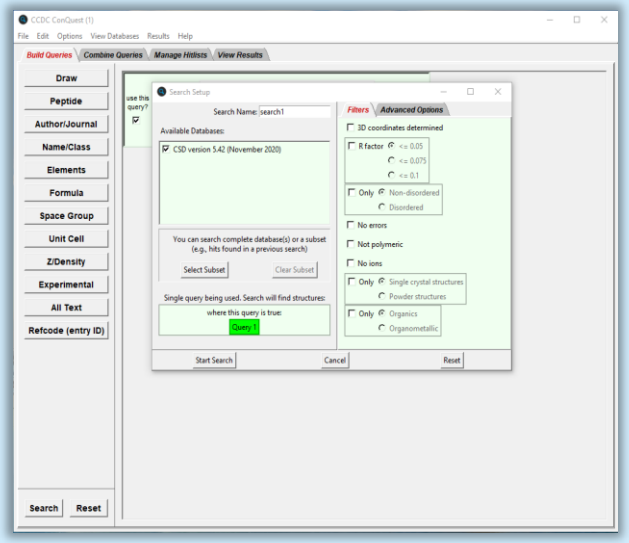
We will make the recording available to you in the next few days.

- Introduction to the CSD
- *Show One:* Research exploring Hydrogen and Halogen bonding
- *Show One:* Introduction and demonstration of ConQuest and Mercury
- *Try One:* Hands-on examples
- *Explore More:* More advanced tips and tricks and case-studies
- *Explore More:* Quiz and Summary
- *Extra time:* More time for hands on and Q&A

CCDC

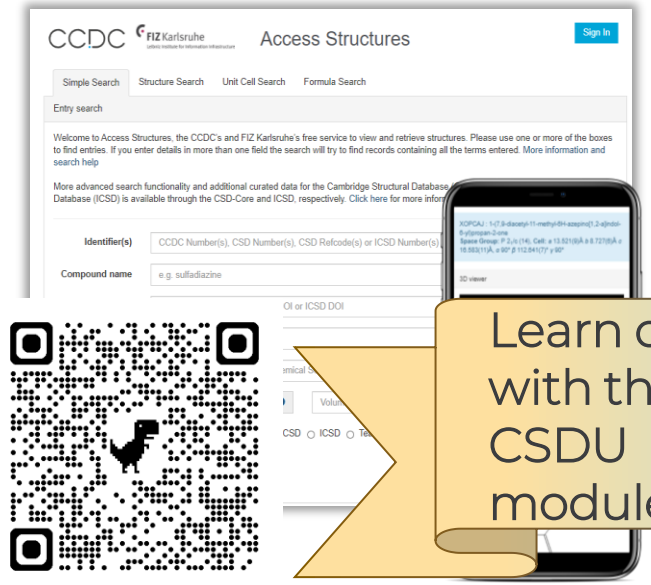
Structure searching

Desktop



ConQuest

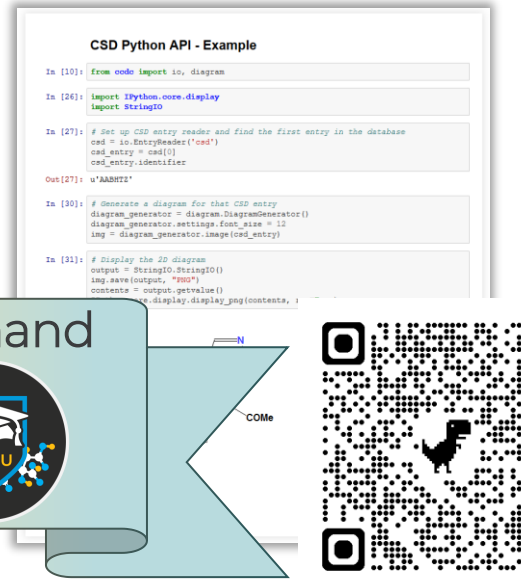
Web Browser



Learn on demand with the CSDU modules!

WebCSD / Access Structures

Programmatic

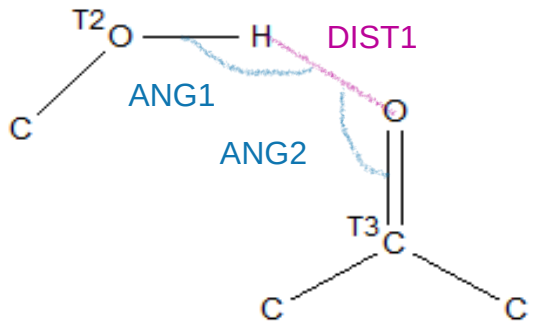


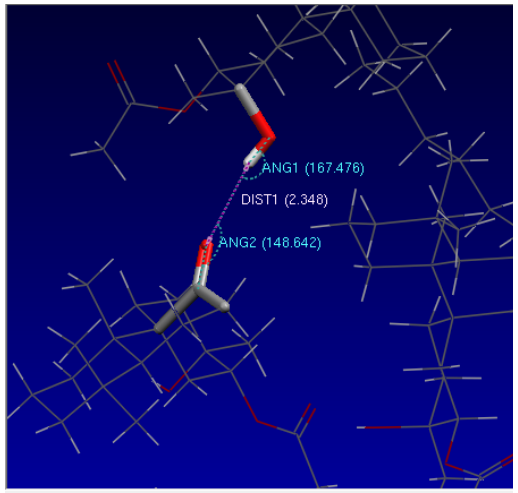
CSD Python API

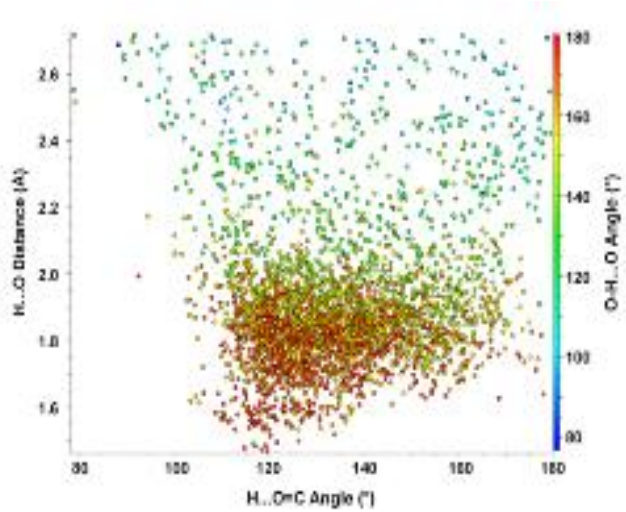
CCDC

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

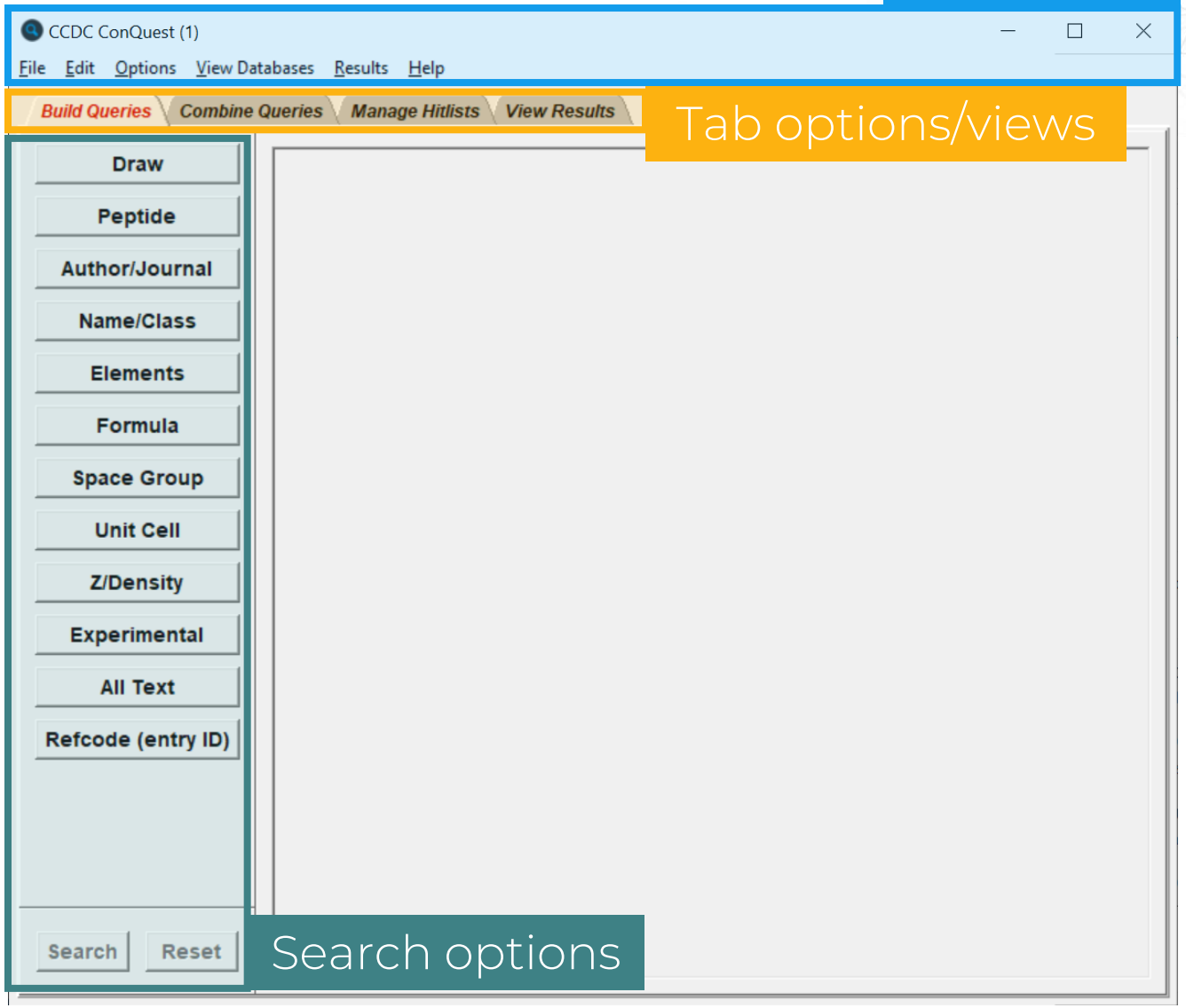






CCDC

ConQuest: Opening and search options

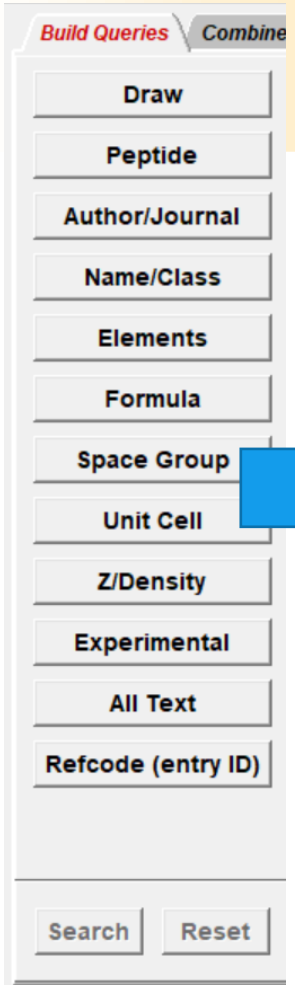


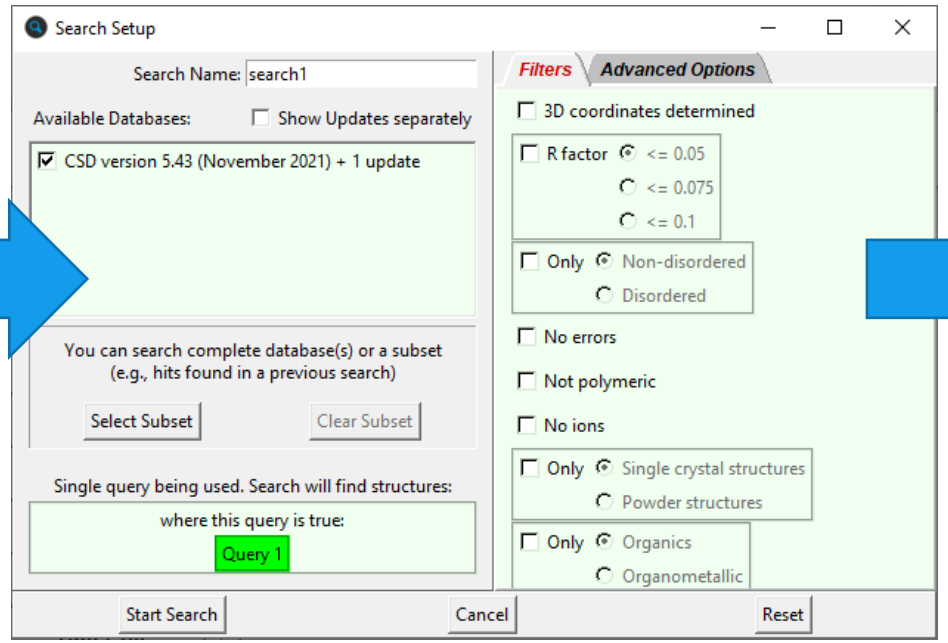
How to search in ConQuest

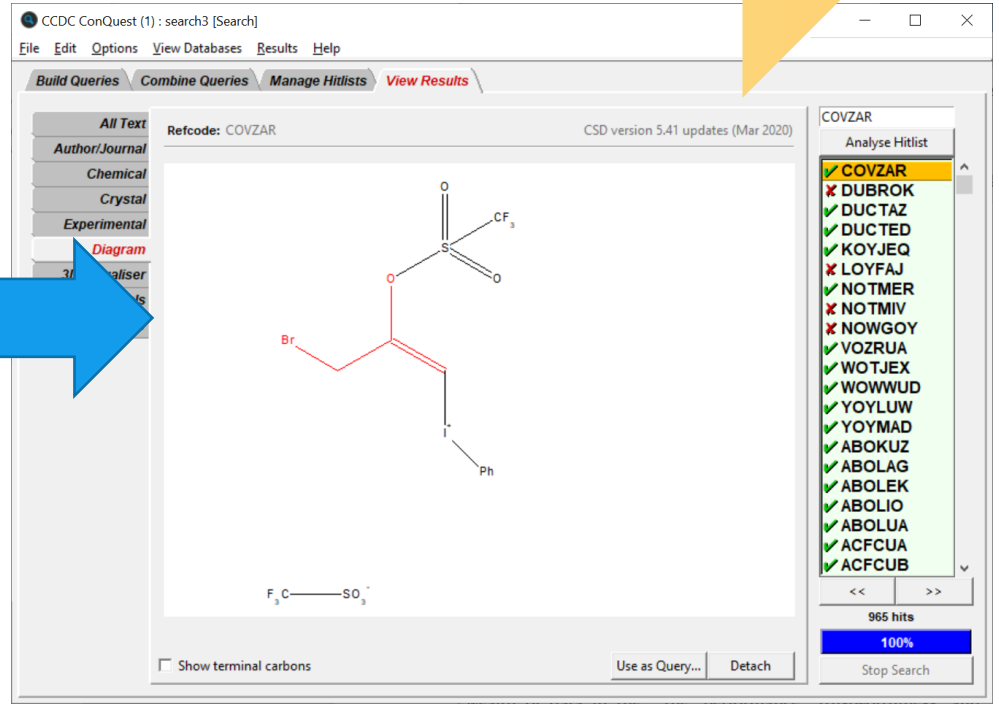
1. Build a Query: what do you want to find?

2. Click search and select filters: do you need to restrict your search?

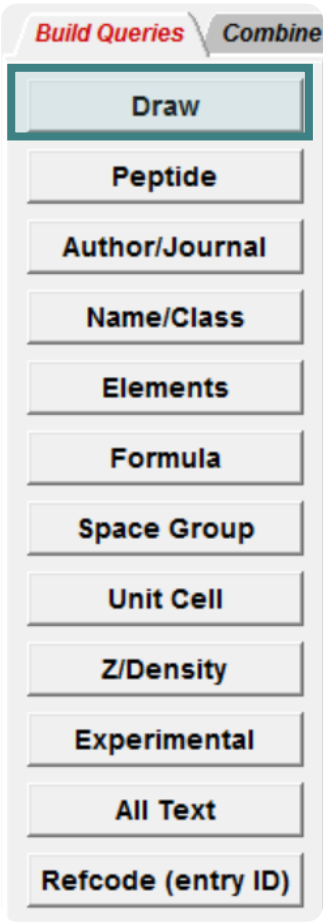
3. Visualise and analyse results: what can you learn from this data?

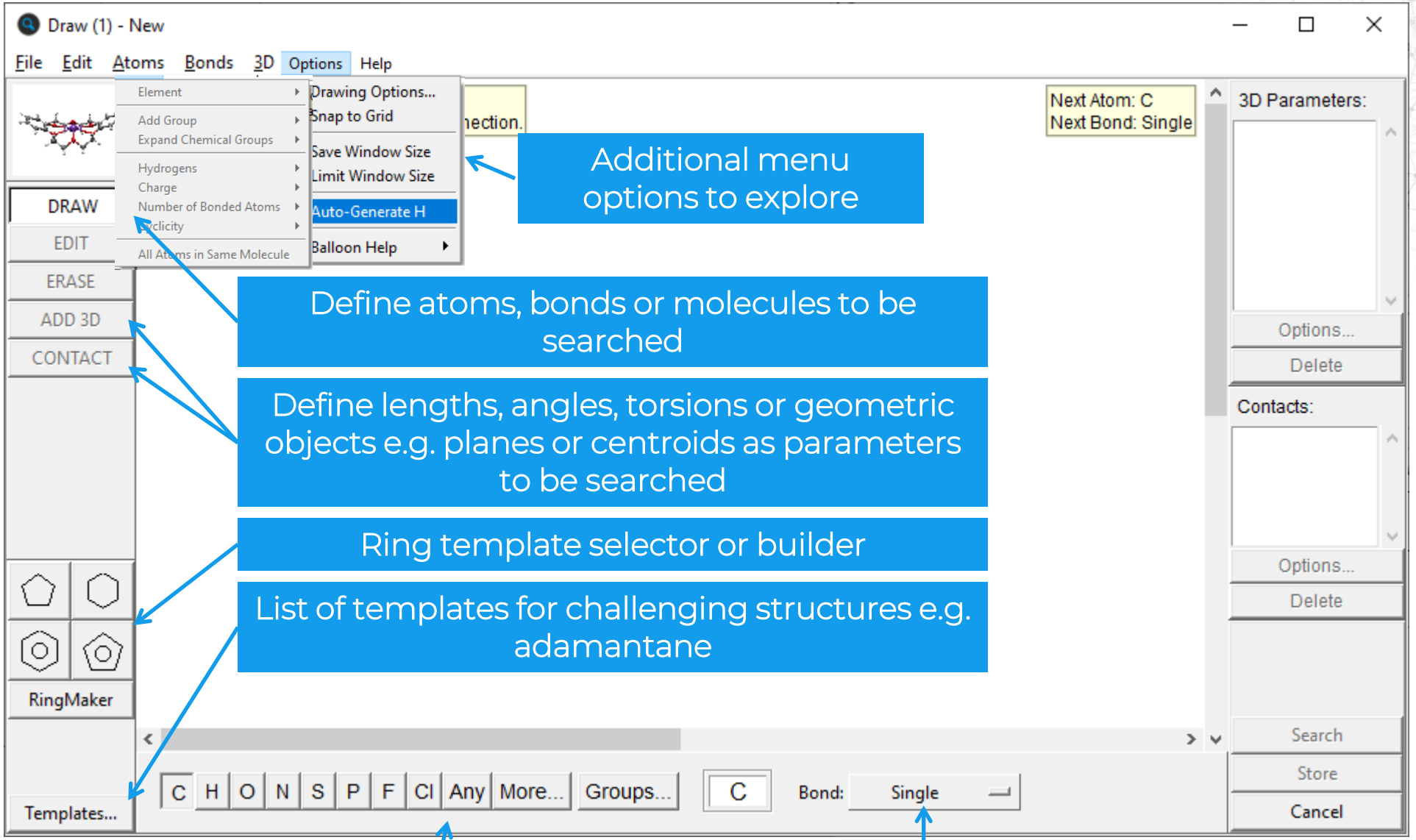






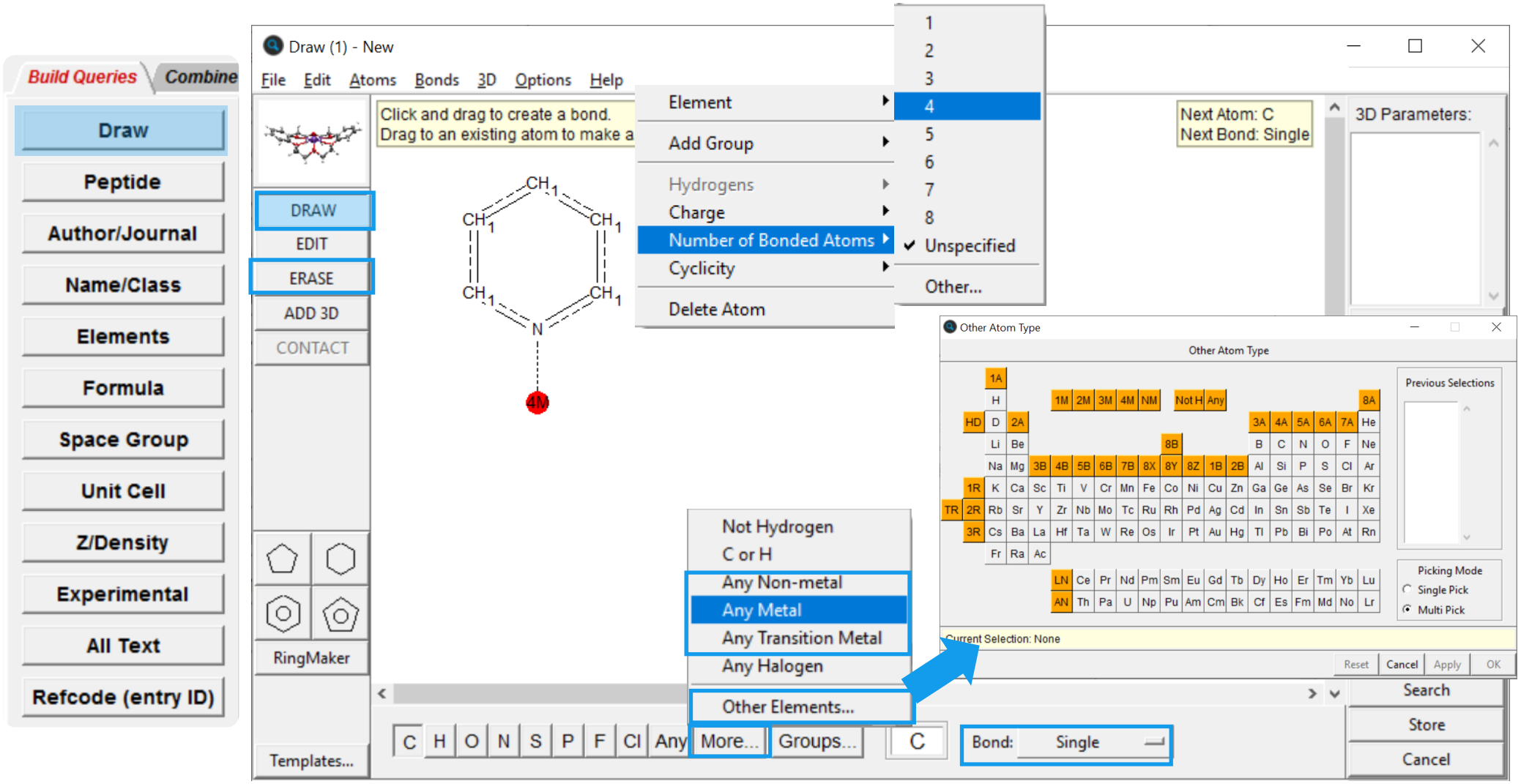
Draw


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
DRAW
EDIT
ERASE
ADD 3D
CONTACT
RingMaker
Templates...
Next Atom: C
Next Bond: Single
3D Parameters:
Options...
Delete
Contacts:
Options...
Delete
Search
Store
Cancel
C H O N S P F Cl Any More... Groups...
C
Bond: Single
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

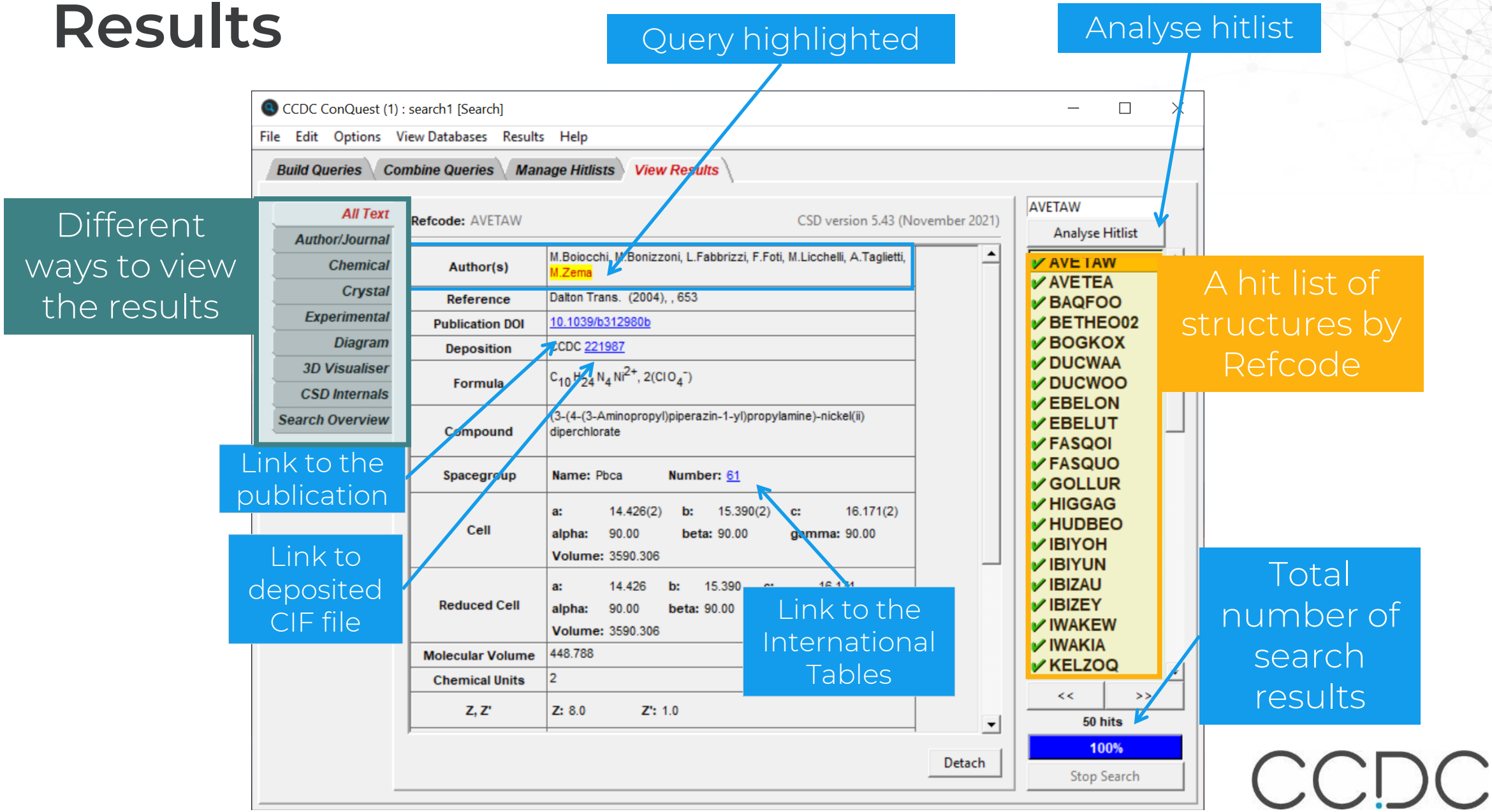


Left click in space to add an atom

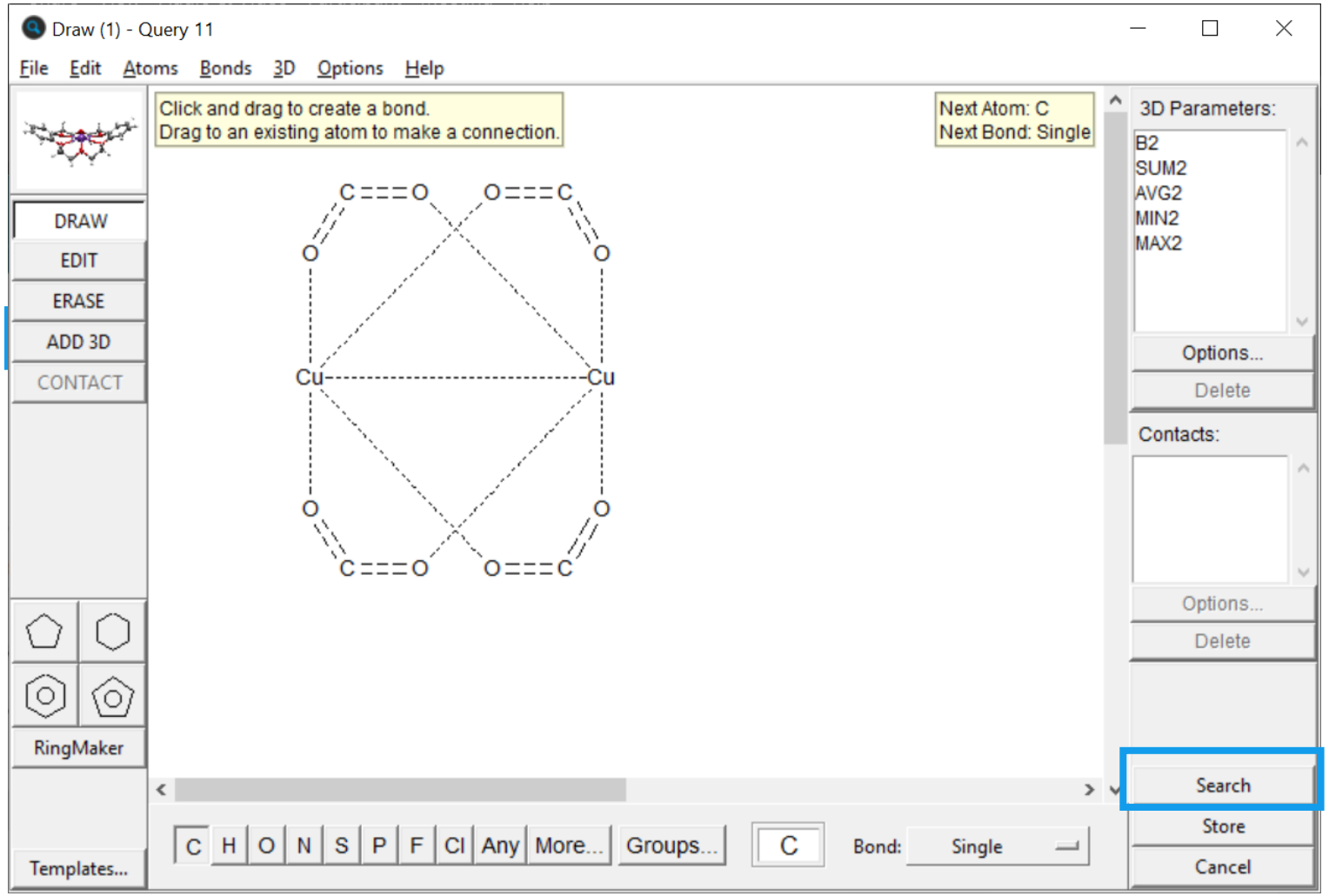
Left click and drag from an atom to attach an atom

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

Results


Query highlighted
Analyse hitlist
Different ways to view the results
A hit list of structures by Refcode
Link to the publication
Link to deposited CIF file
Link to the International Tables
Total number of search results
CCDC ConQuest (1) : search1 [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: AVETAW
CSD version 5.43 (November 2021)
Author(s) M.Boiocchi, M.Bonizzoni, L.Fabbrizzi, F.Foti, M.Licchelli, A.Taglietti, M.Zema
Reference Dalton Trans. (2004), , 653
Publication DOI 10.1039/b312980b
Deposition CCDC 221987
Formula $C_{10}H_{24}N_4Ni^{2+}, 2(ClO_4^-)$
Compound (3-(4-(3-Aminopropyl)piperazin-1-yl)propylamine)-nickel(ii) diperchlorate
Spacegroup Name: Pbca Number: 61
Cell a: 14.426(2) b: 15.390(2) c: 16.171(2) alpha: 90.00 beta: 90.00 gamma: 90.00 Volume: 3590.306
Reduced Cell a: 14.426 b: 15.390 alpha: 90.00 beta: 90.00 Volume: 3590.306
Molecular Volume 448.788
Chemical Units 2
Z, Z' Z: 8.0 Z': 1.0
Detach
AVETAW
Analyse Hitlist
AVETAW
AVETEA
BAQFOO
BETHEO02
BOGKOX
DUCWAA
DUCWOO
EBELON
EBELUT
FASQOI
FASQUO
GOLLUR
HIGGAG
HUDBEO
IBIYOH
IBIYUN
IBIZAU
IBIZEY
IWAKEW
IWAKIA
KELZOQ
<< >>
50 hits
100%
Stop Search

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

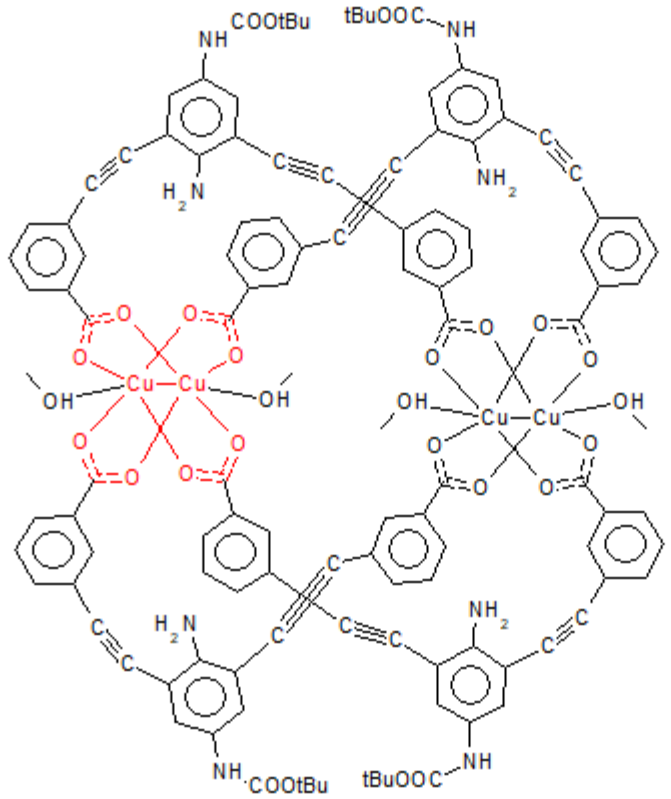
CCDC ConQuest (1) : search11 [Search]

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

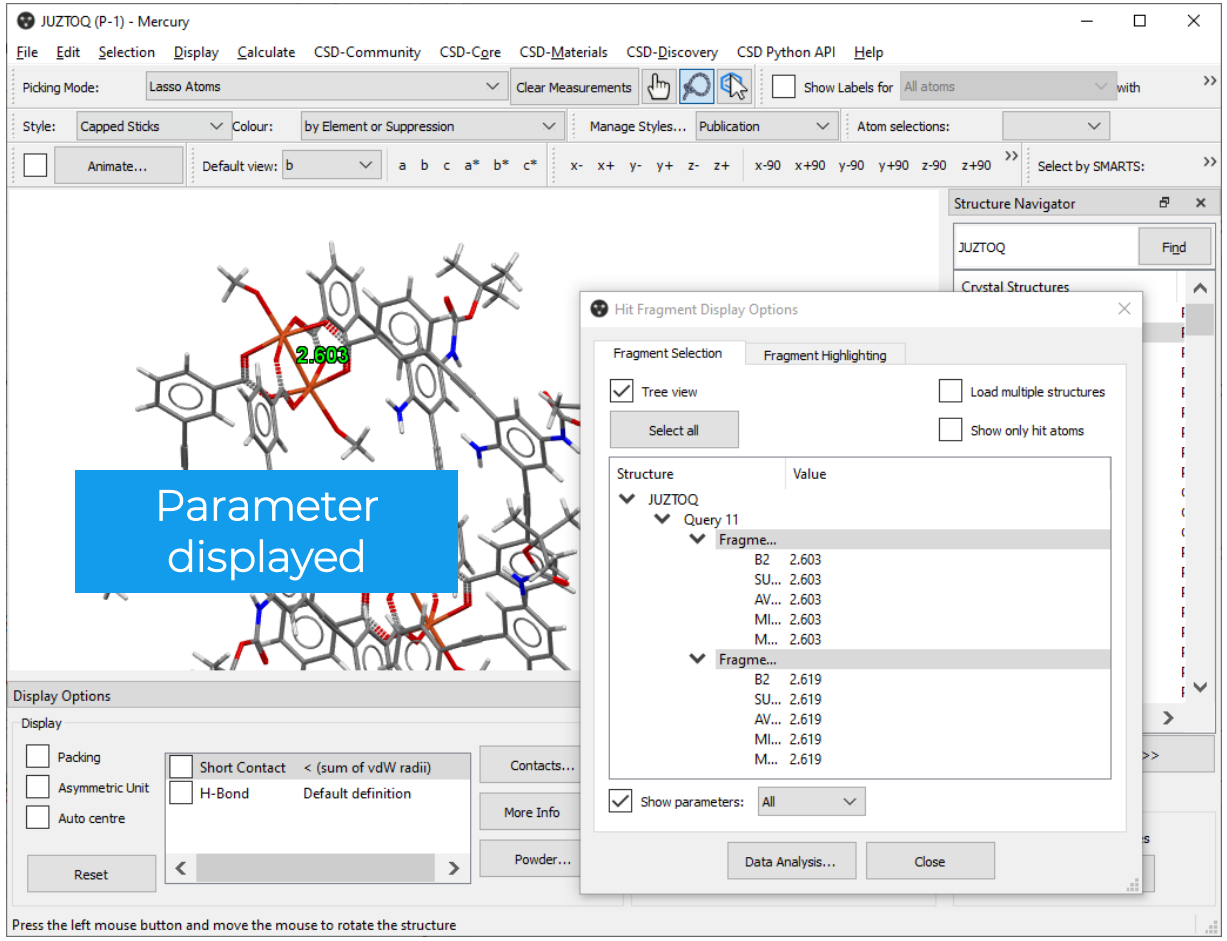
0% | Reset | Cancel | Analyse in Mercury

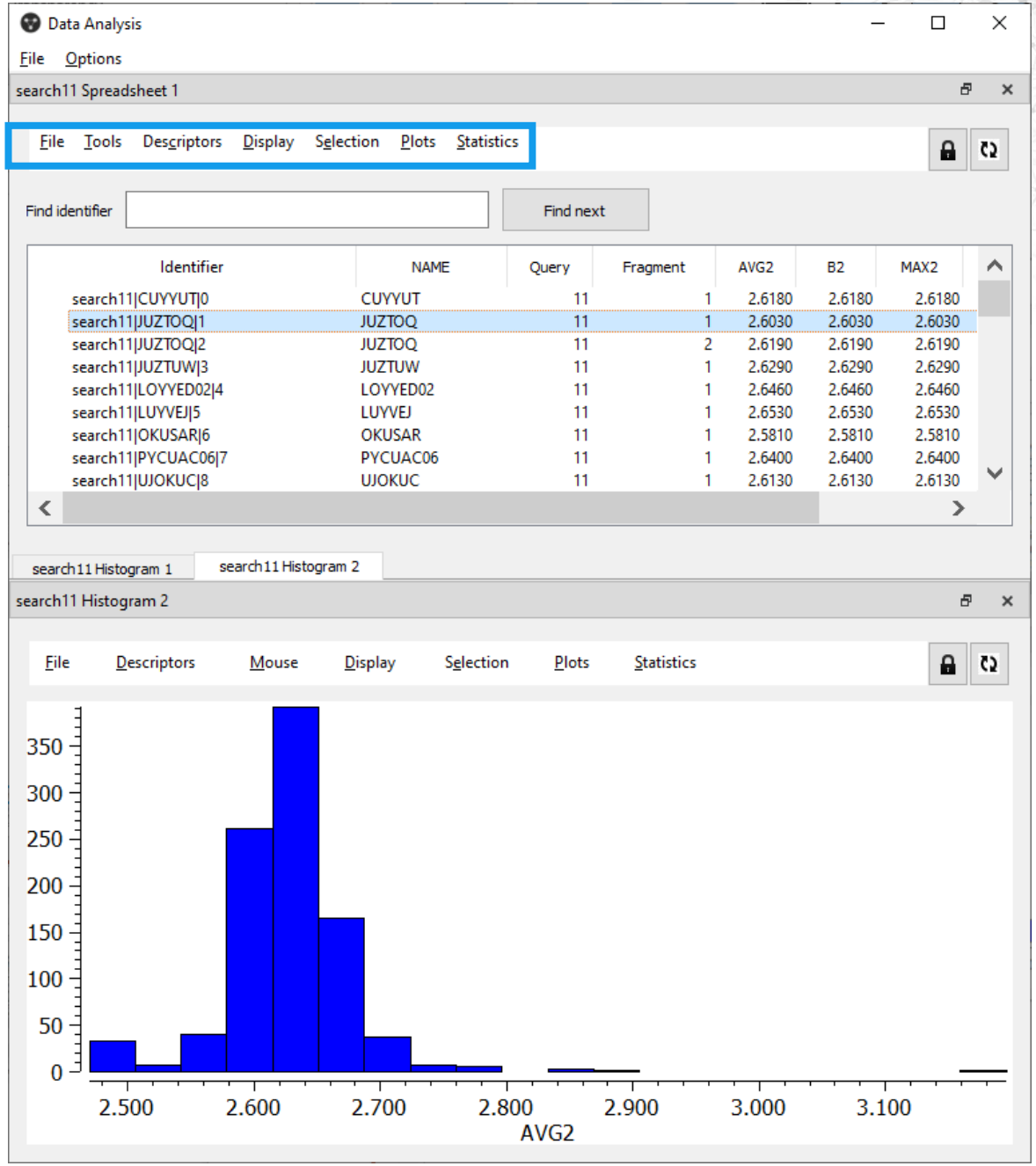
Search highlighted in 2D

Parameters displayed

CCDC

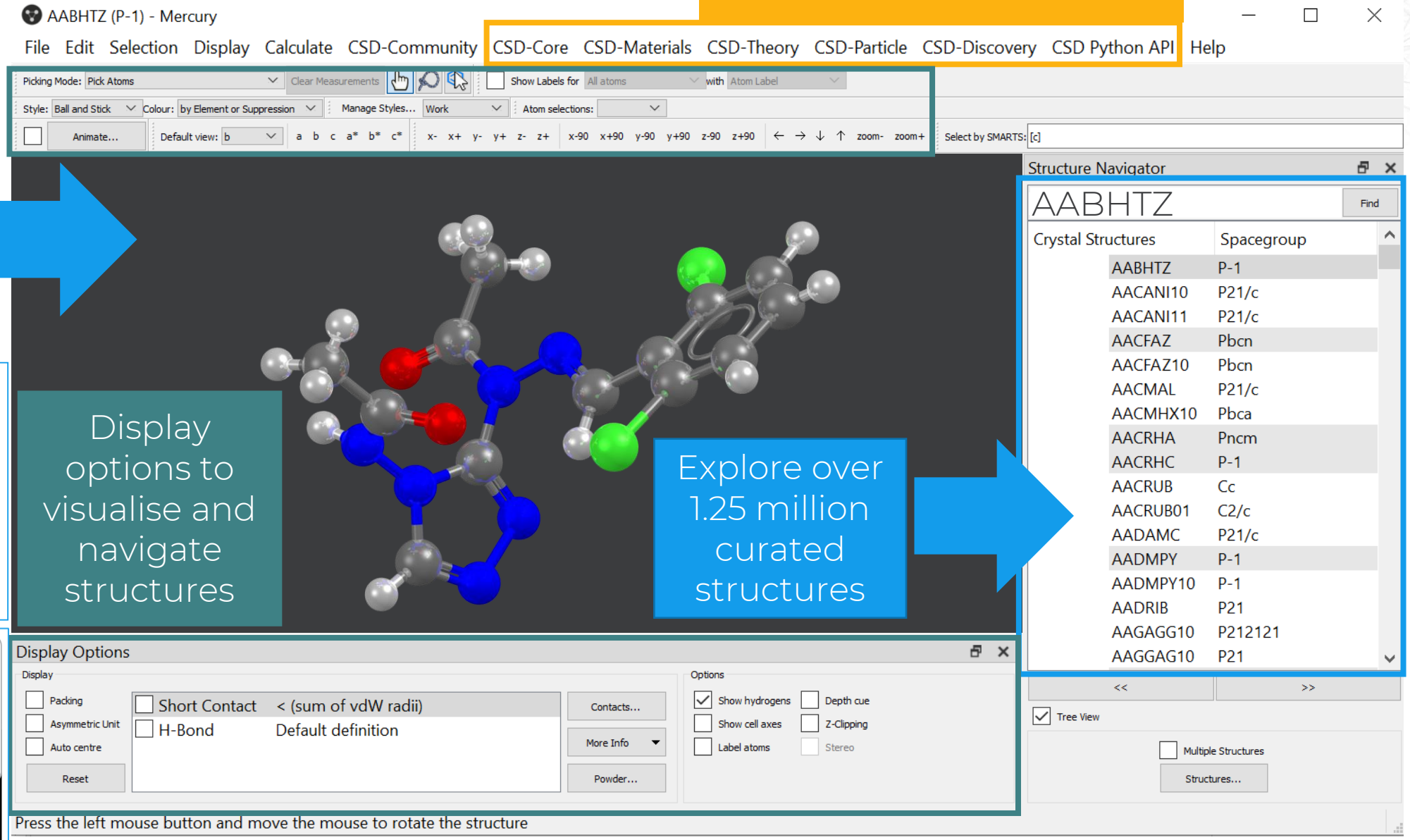
Data analysis

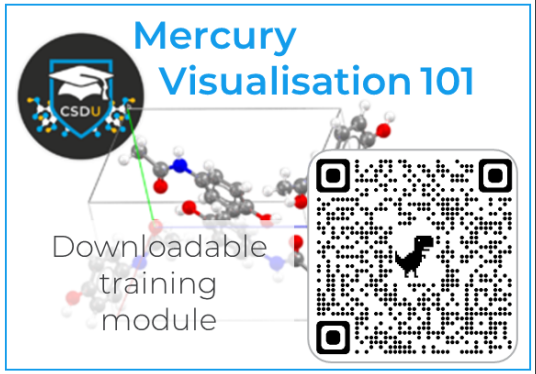




Mercury Overview

More advanced functionality to analyse and learn from structures





Show One: Demo of ConQuest and Mercury

- In this demo we will use ConQuest and Mercury to follow one of Catherine's investigations in *Cryst. Growth Des.* 2024, 24, 2, 859–870.
  - Construct a 3D search on ConQuest.
  - Export results to Mercury.
  - Use the Data Analysis tool to explore Hydrogen Bonded $R^2_2(8)$ Rings.

We will make the recording available to you in the next few days.

Catharine Esterhuysen, *Cryst. Growth Des. 2024, 24, 2, 859–870.* DOI: 10.1021/acs.cgd.3c01343

CCDC

Try One: hands-on exercise

We will make the recording available to you in the next few days.

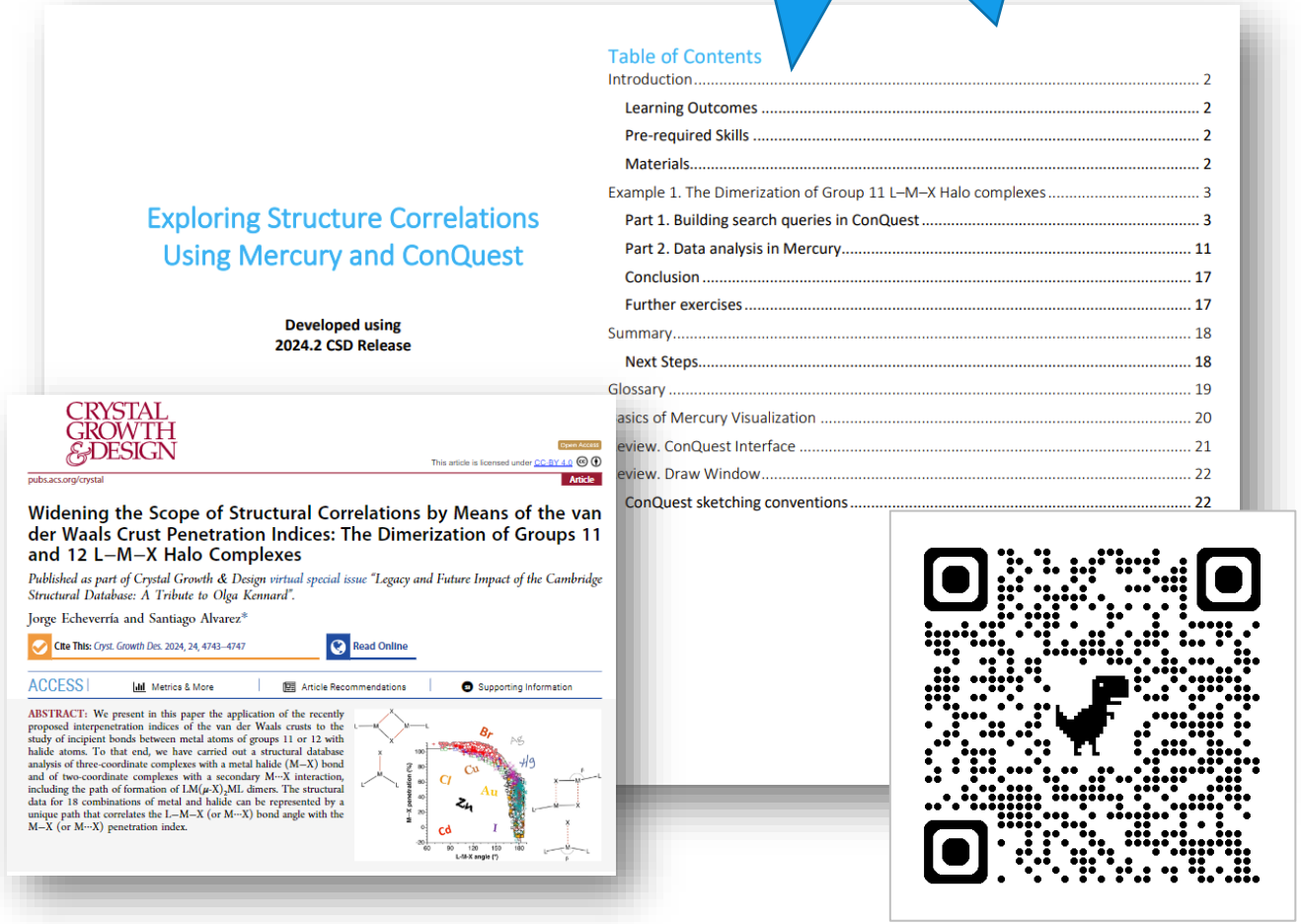
It's your turn!

- Try the example from the handout.
- Your tutors are on hand to help you!
- To ask questions during this time type a message in the chat box.

These exercises follow research in another recent article:

- J. Echeverría and S. Alvarez, *Cryst. Growth Des.*, 2024, 24, 4743–4747. DOI: 10.1021/acs.cgd.4c00335

https://info.ccdc.cam.ac.uk/2024-autumn-virtual-workshop



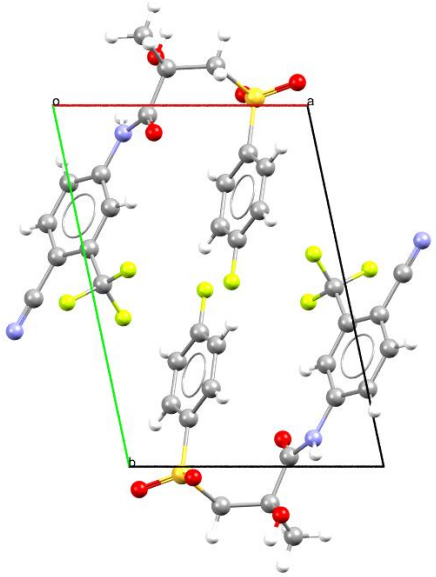
Agenda

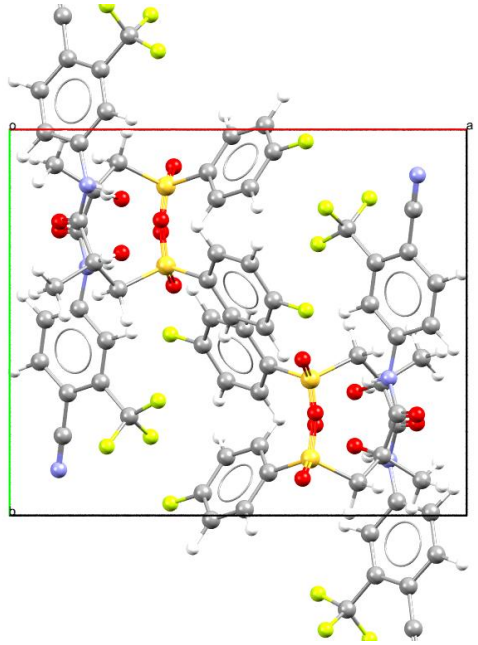
- Introduction to the CSD
- *Show One:* Research exploring Hydrogen and Halogen bonding
- *Show One:* Introduction and demonstration of ConQuest and Mercury
- *Try One:* Hands-on examples
- *Explore More:* More advanced tips and tricks and case-studies
- *Explore More:* Quiz and Summary
- *Extra time:* More time for hands on and Q&A

CCDC

Intermolecular interactions and packing

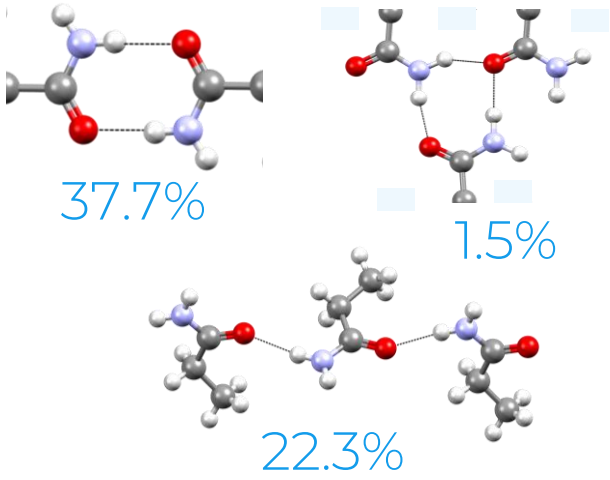
- To go further you can look at the packing environment of the molecules using a range of tools.
- Using the CSD for context (unusual hydrogen bond parameters for example) as well as enabling comparison of structures.





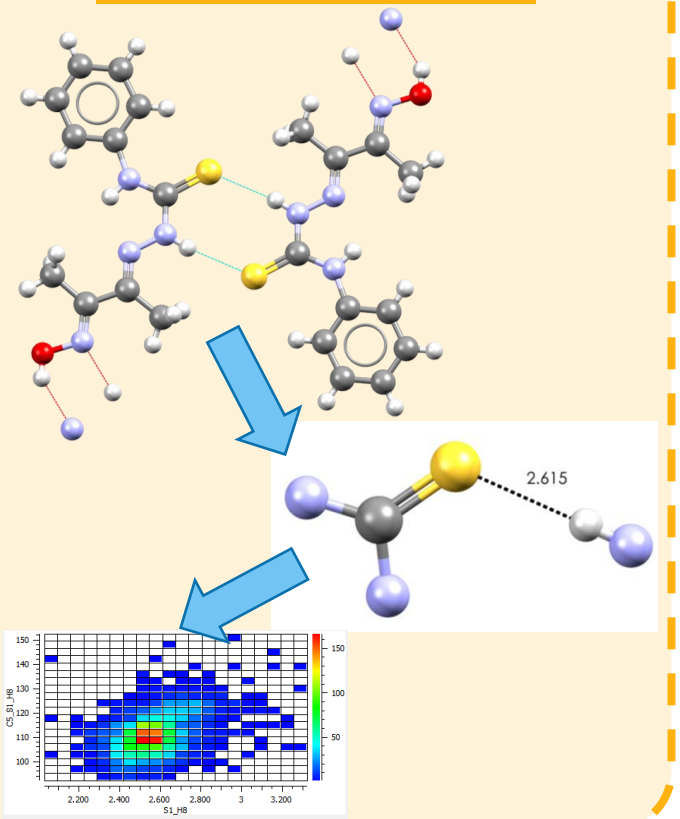
Searching within Mercury

Motifs

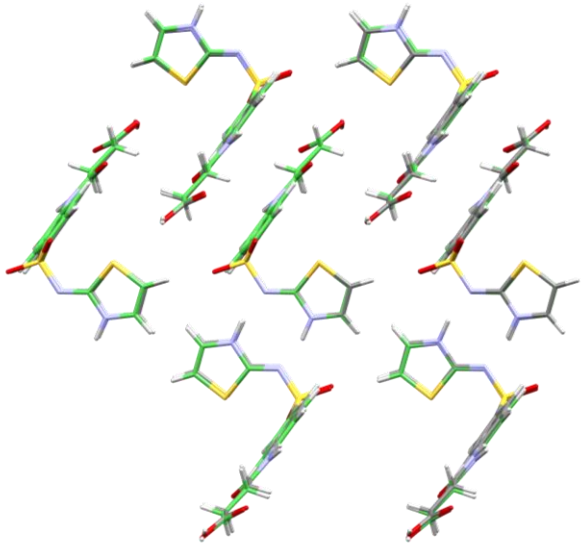


- H-bond motifs

Crystal Packing Feature



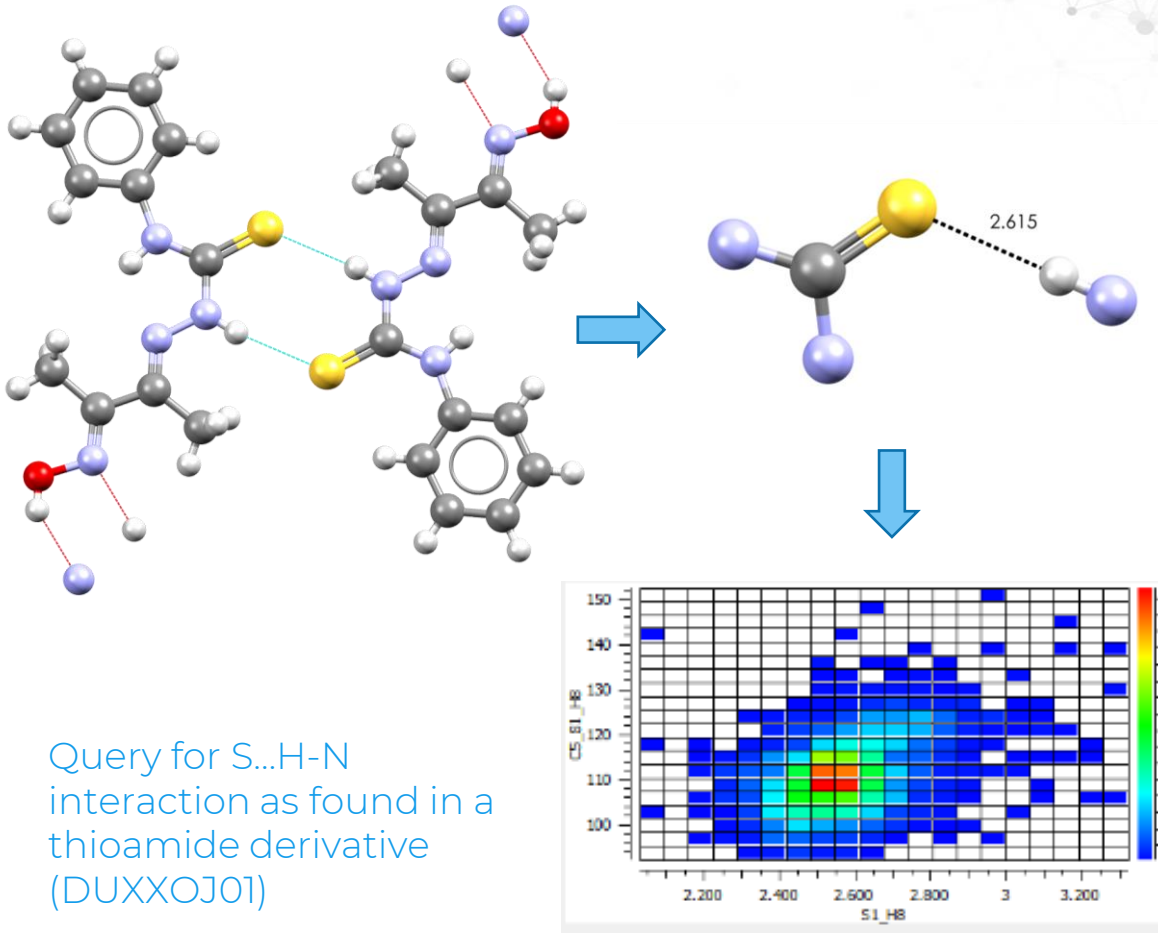
Crystal Packing Similarity



CCDC

Crystal Packing Feature Search

- Perform a substructure search
- Investigate conformations of molecules or bonded fragments.
- Search for:
  - non-covalent interactions such as π-π or hydrogen bond interactions.
  - particular spatial arrangements of functional groups.
  - particular spatial arrangements of molecules.

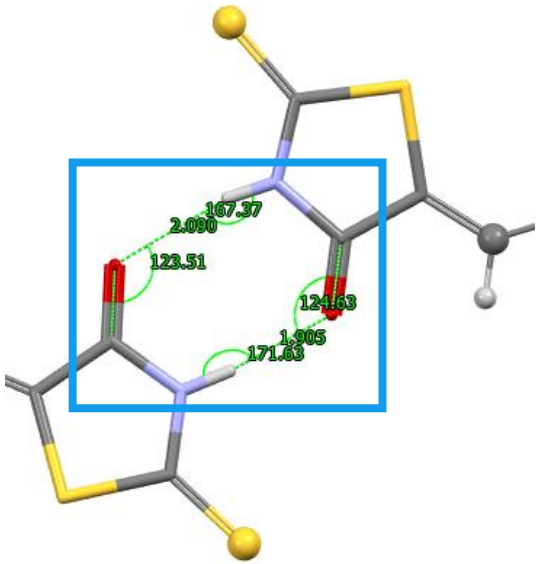


Query for S...H-N interaction as found in a thioamide derivative (DUXXOJ01)

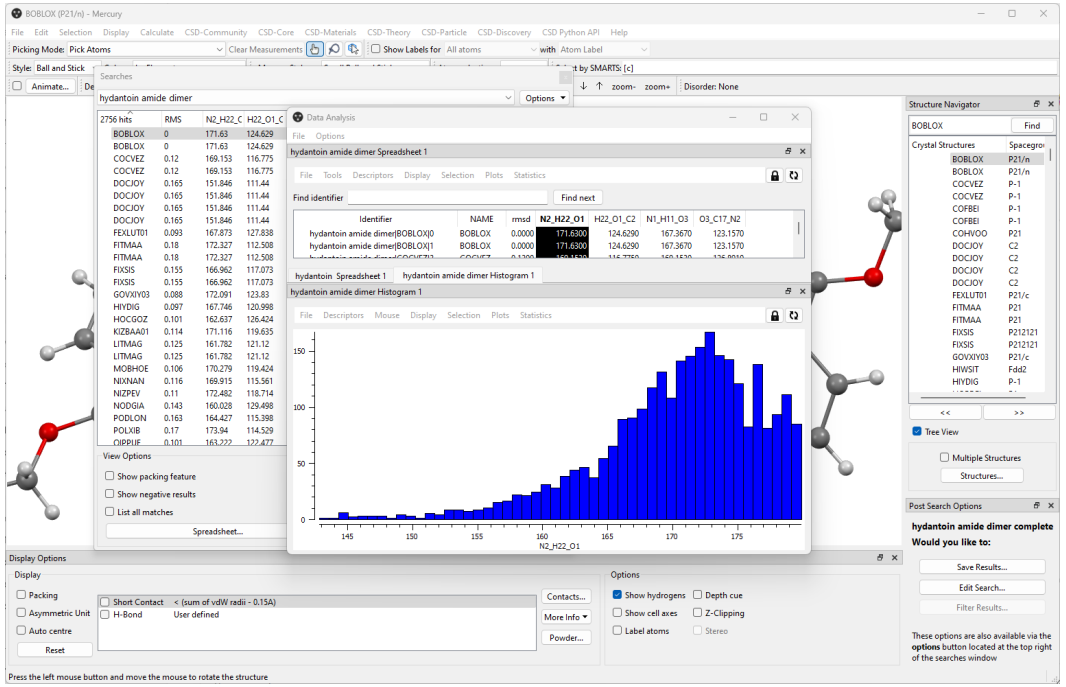
CCDC

Searching within Mercury: Crystal Packing Feature

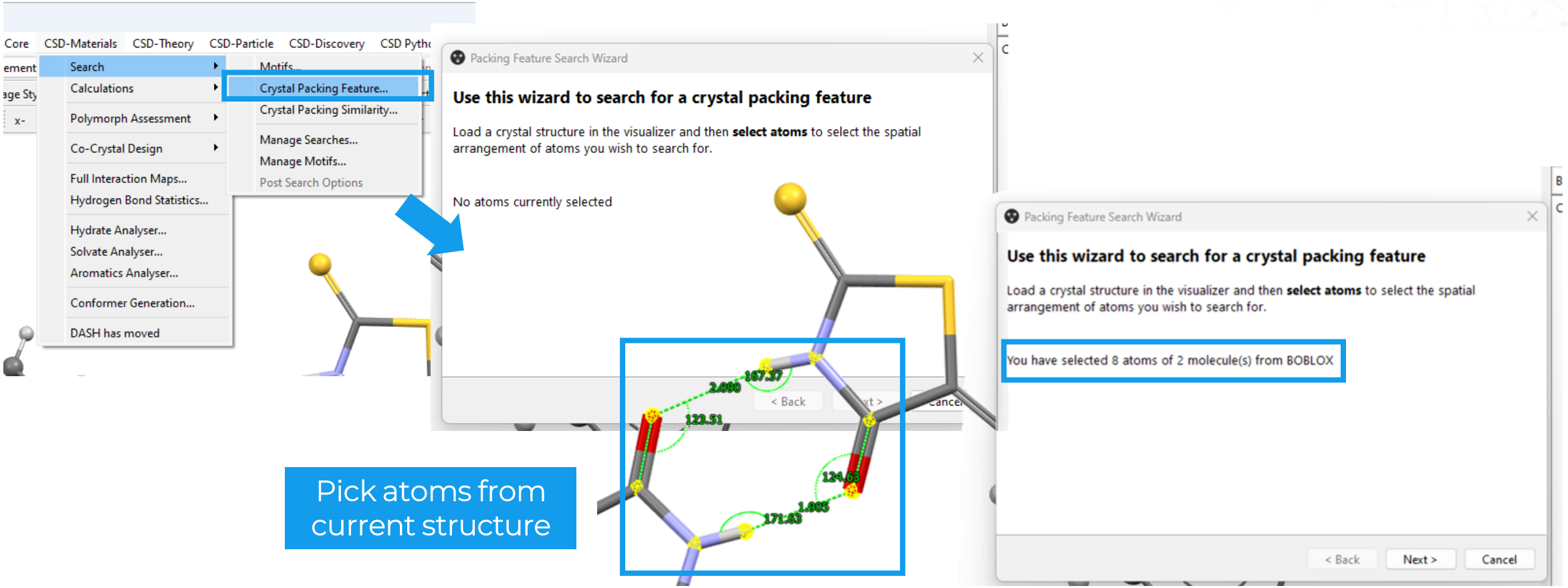
- Pick a feature from the current structure in Mercury
- Search the whole CSD or within ConQuest hitlist



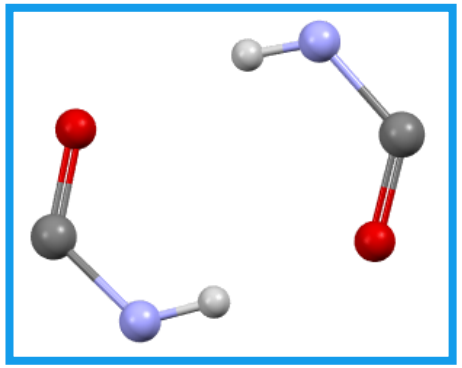
Packing Feature Search



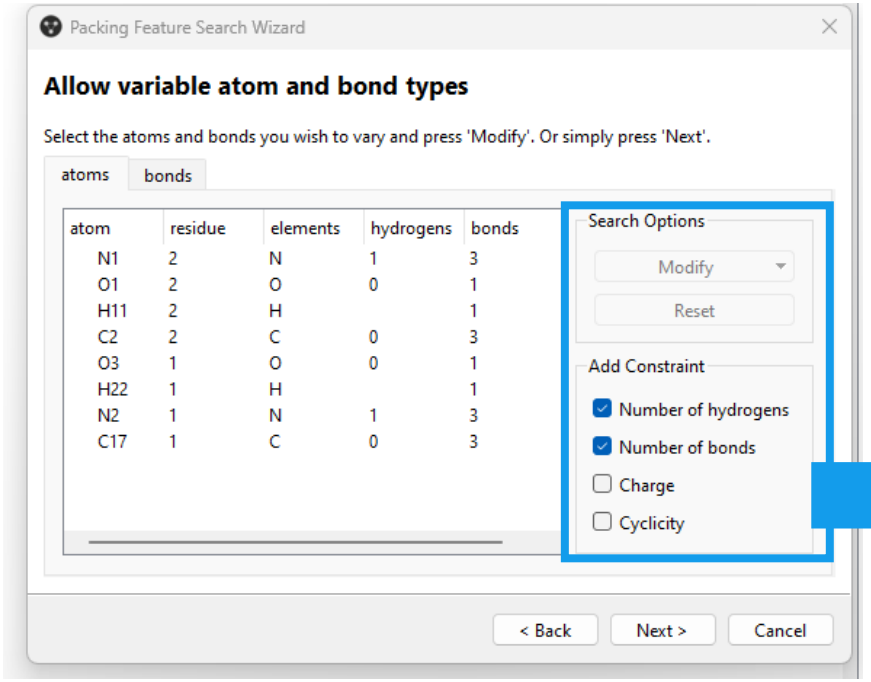
Crystal Packing Feature : Selecting a Feature



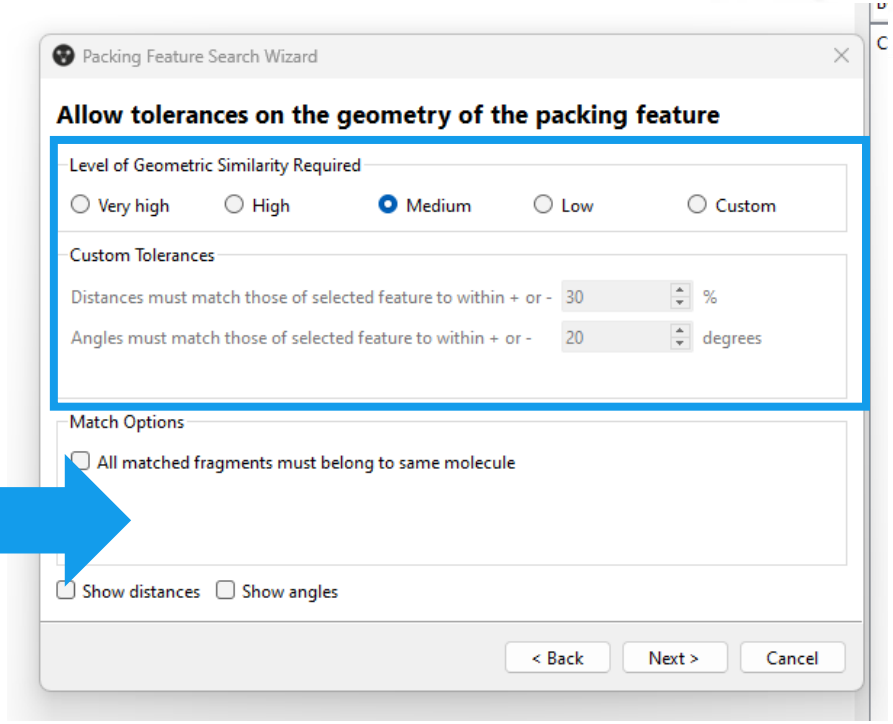
Crystal Packing Feature: Selecting Options



Picked fragment of structure



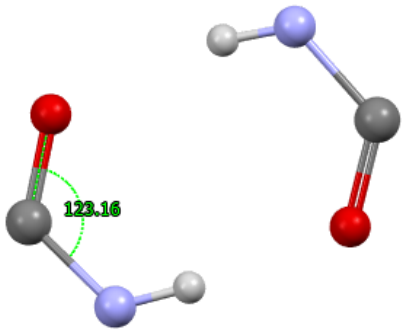
Add constraints if required

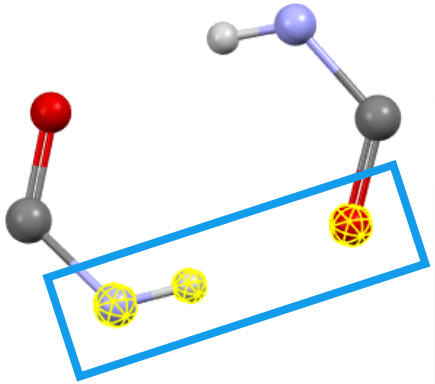


Select search tolerance for RMSD calculations

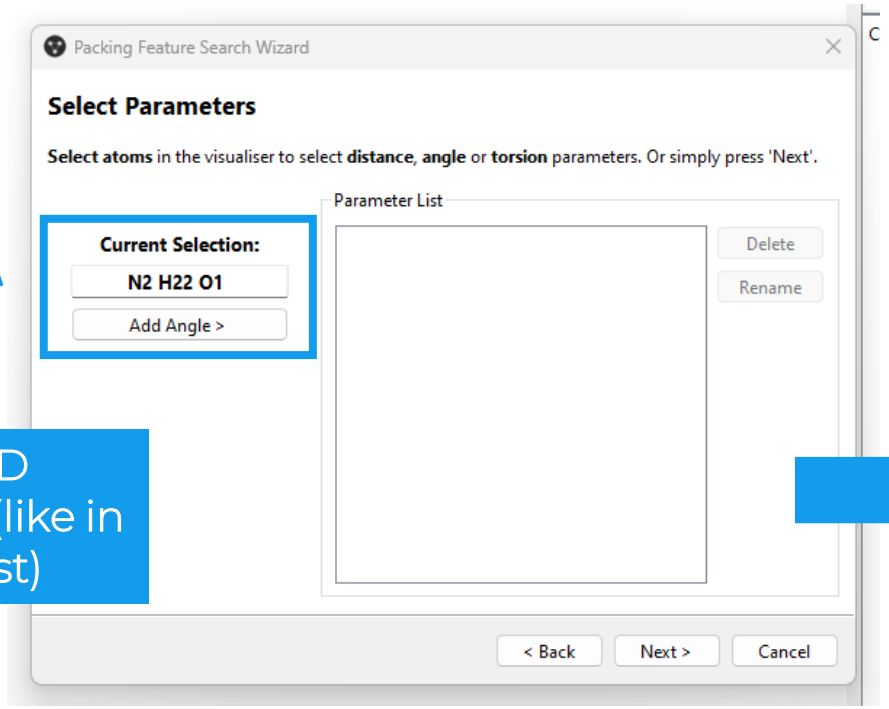
CCDC

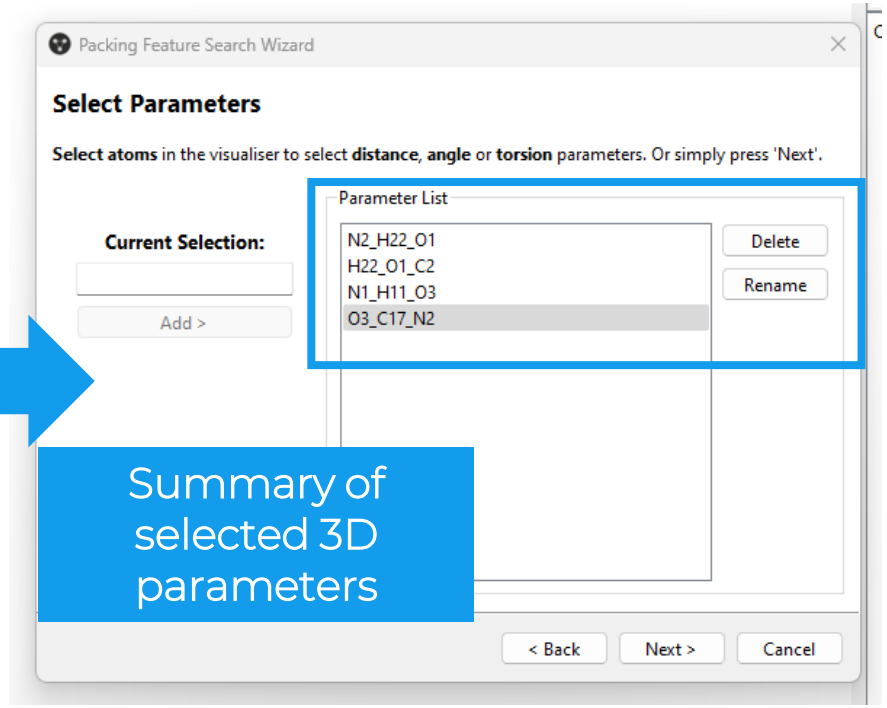
Crystal Packing Feature: Selecting Parameters





Select 3D parameters (like in ConQuest)

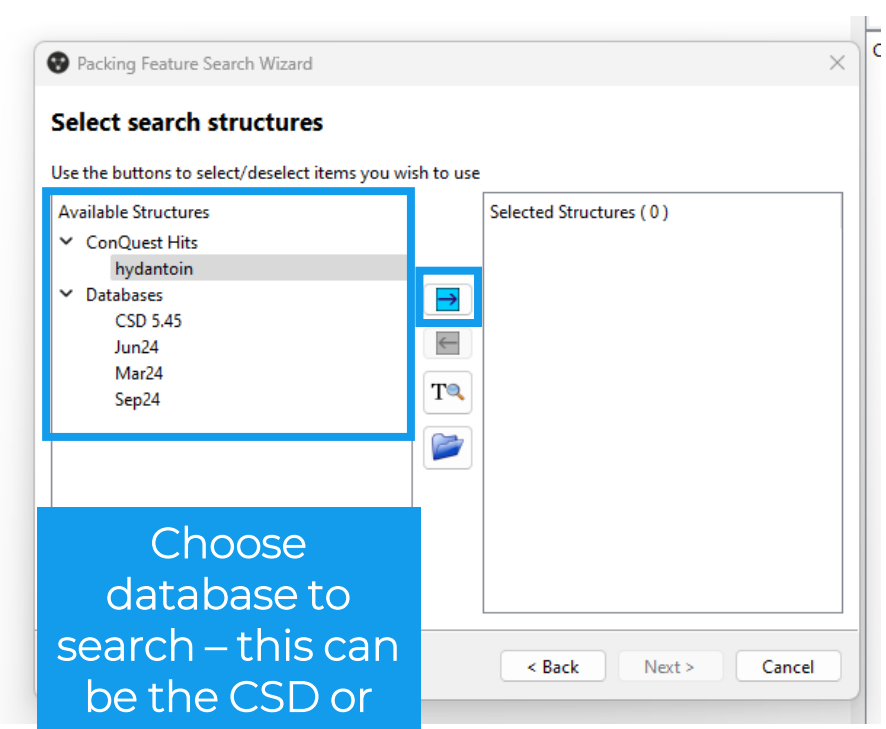




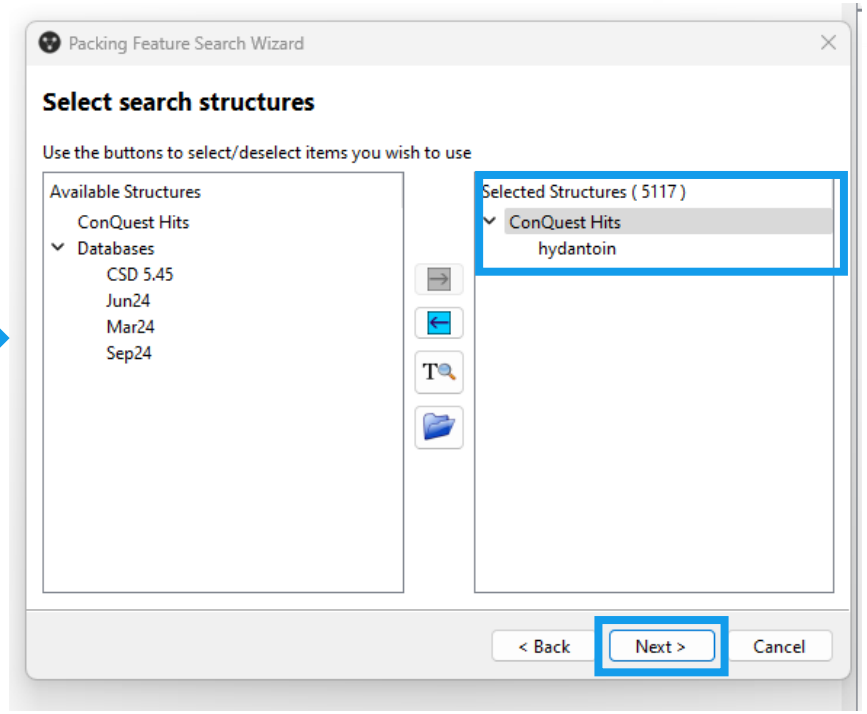
Summary of selected 3D parameters

CCDC

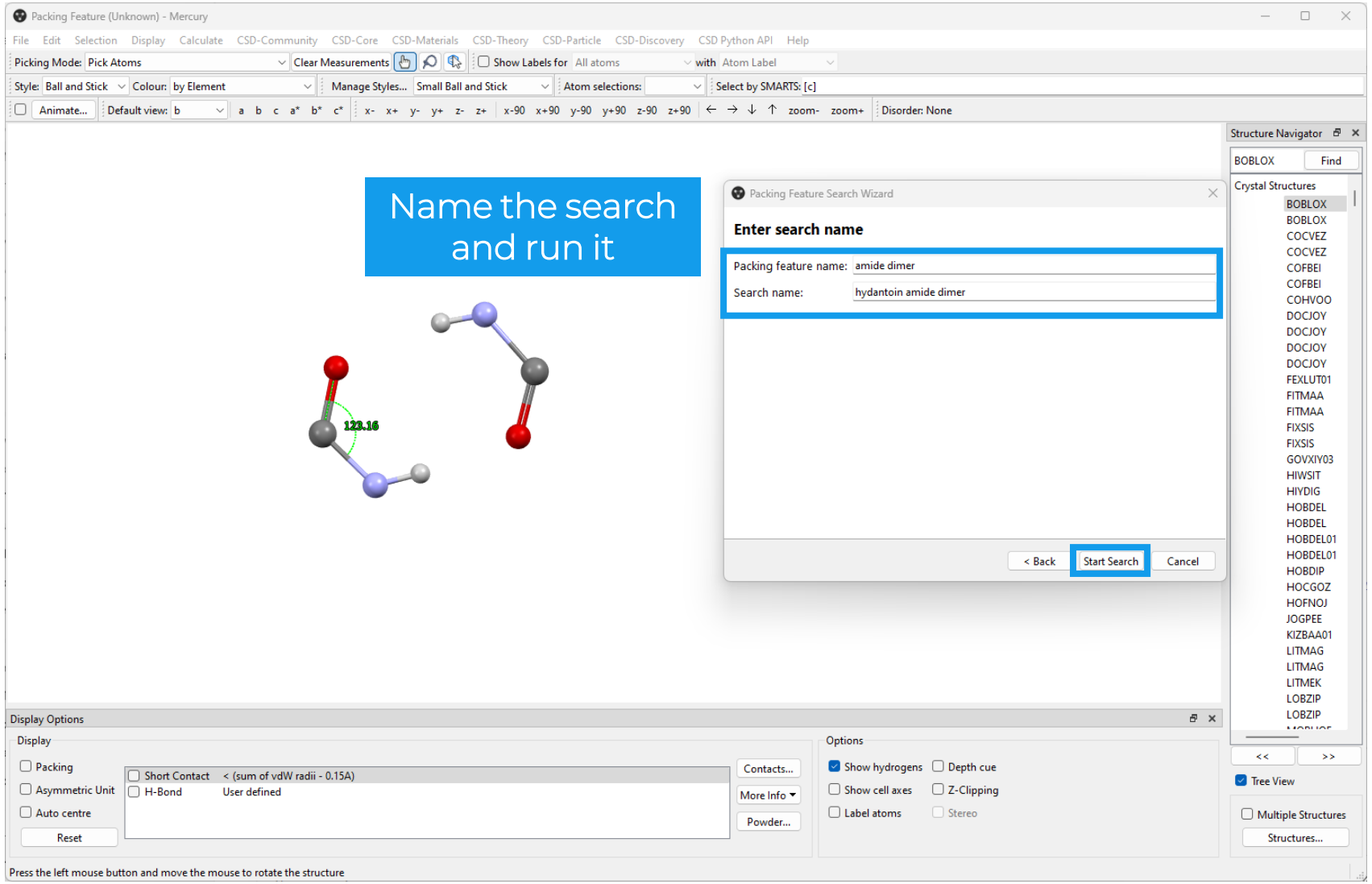
Crystal Packing Feature: Selecting Search Structures



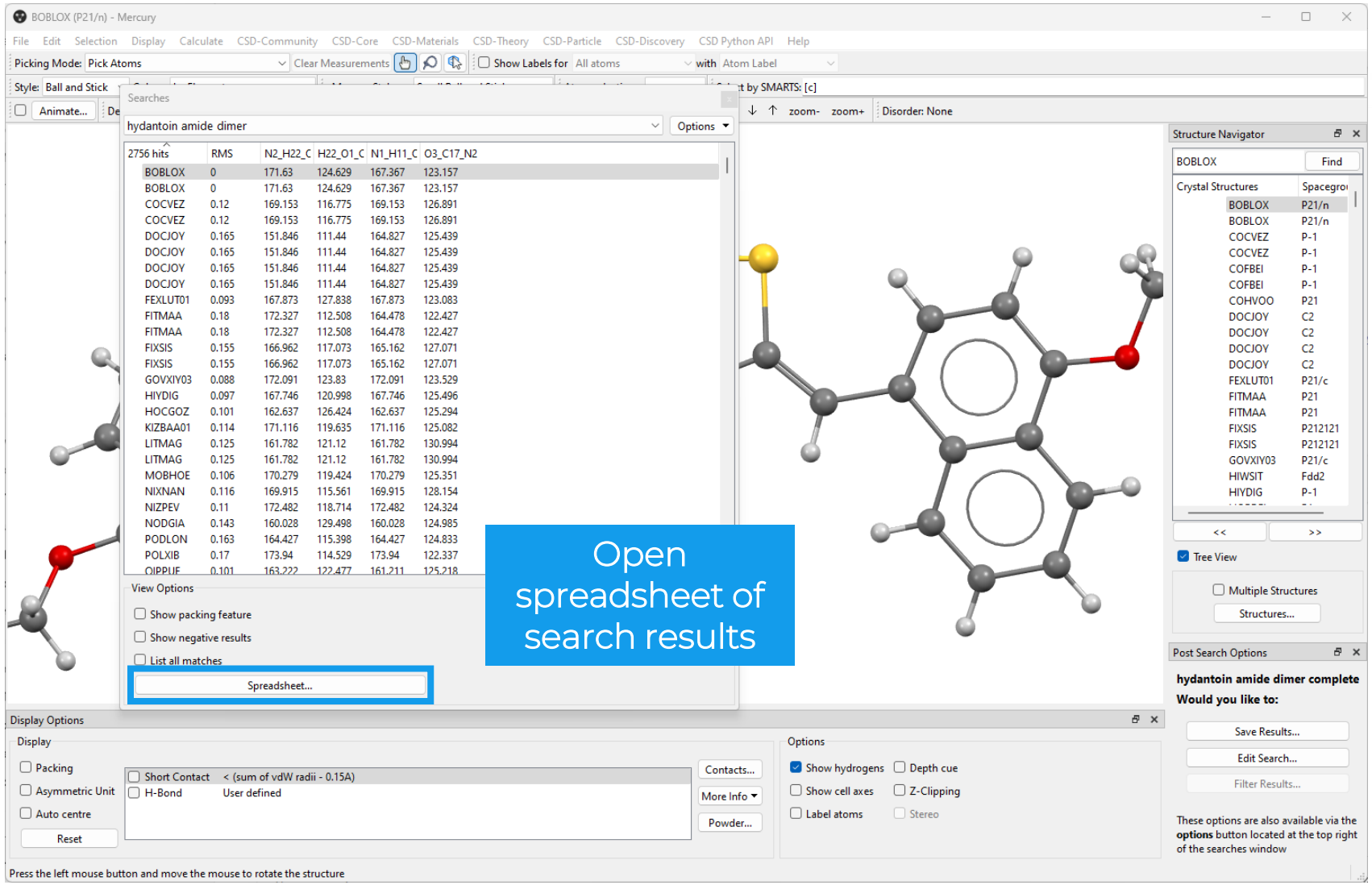
Choose database to search – this can be the CSD or ConQuest hits



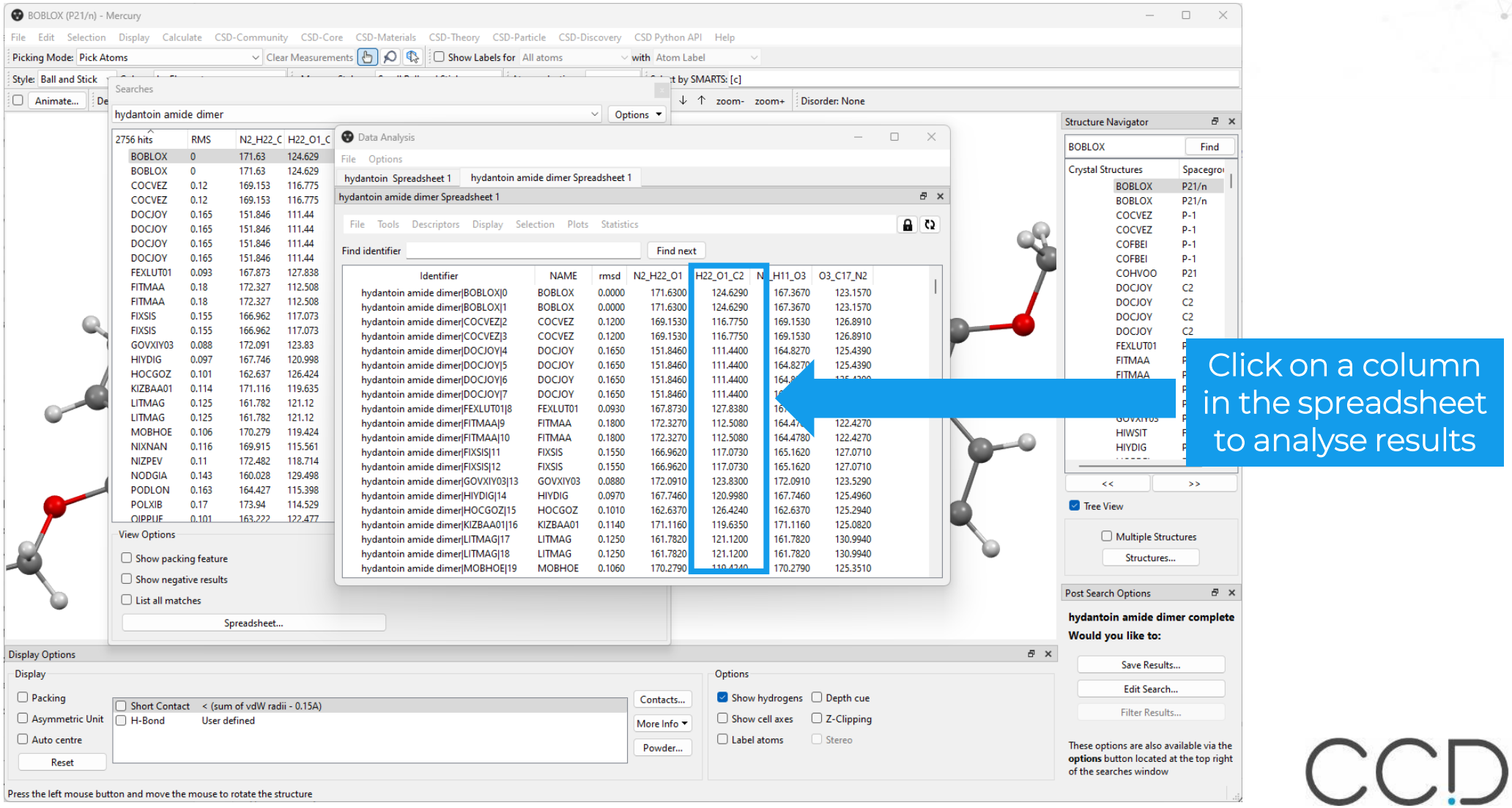
Crystal Packing Feature: Launching Search



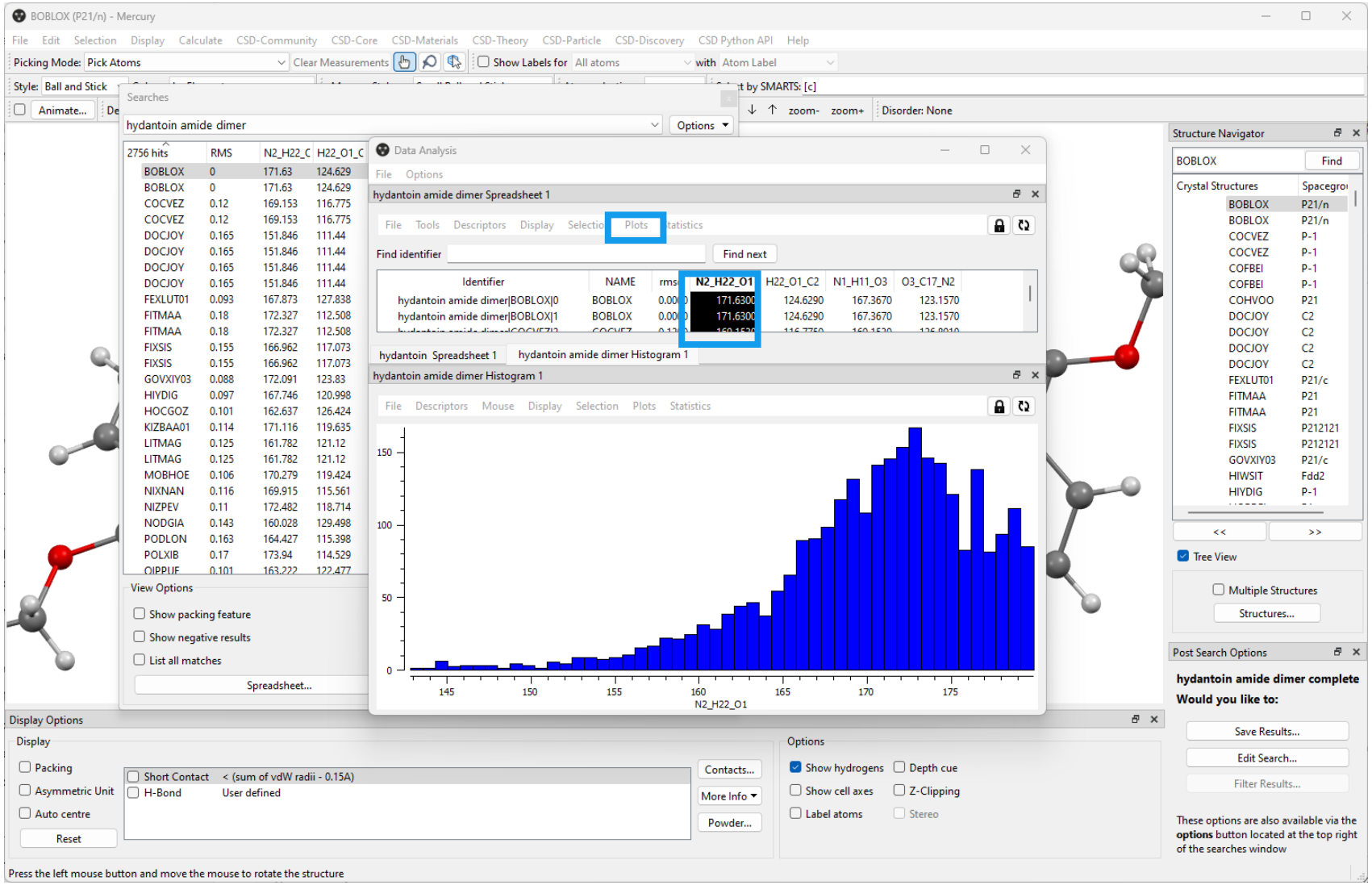
Crystal Packing Feature: Viewing Results



Crystal Packing Feature: Viewing Spreadsheet



Crystal Packing Feature: Analysing Results



Plot graphs and conduct other analysis

CCDC

Try the feature!

- CSD-Materials self-guided workshops: https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-materials-workshops/
- “Motifs, Crystal Packing Feature and Crystal Packing Similarity Search”

Search
(Motifs, Crystal Packing Feature, Crystal Packing Similarity MAT-006)

Developed using 2020.0 CSD Release

CCDC
advancing structural science

Table of Contents

CCDC

How are others using ConQuest and Mercury?

Mercury in Action: A Geometry Analysis on Complexes With 2,2'-Dipyridylamine

Dr Tetteh investigated the geometry of first-row transition metal complexes with 2,2'-dipyridylamine and halogen ligands by searching the Cambridge Structural Database (CSD) with ConQuest and analysing the data with Mercury. Density Functional Theory (DFT) calculations were also performed to study the stability of the complexes.

The research is part of the Frank H. Allen International Research and Education (FAIRE) Programme, which supports academic researchers and educators from underrepresented communities with access to the wealth of information contained within the CSD.

Why?

The ligand 2,2'-dipyridylamine (dpa, **Figure 1 a**) is widely used in coordination and organometallic chemistry for its ability to bind to a variety of metal centres and adopt different coordination modes, acting as a monodentate, bidentate, or bridging ligand. Among these, the bidentate coordination mode (**Figure 1 b**) is particularly stable thermodynamically, and hence it is frequently found in metal complexes.

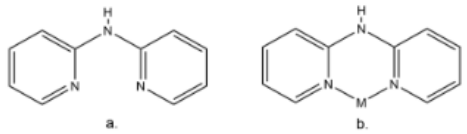
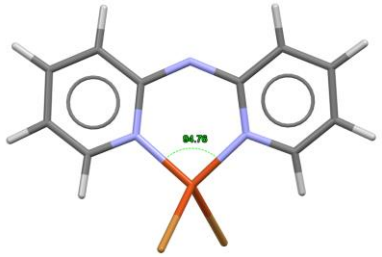


Figure 1. a) Structure of the ligand dpa and **b)** bidentate coordination mode for dpa. M = metal centre.

Metal complexes that involve dpa are used as catalysts and other photophysical materials. The ligand is also used for MOFs design and as coligand for magnetic complexes. Despite its popularity, research still needs to be done to understand the chemistry of dpa and its behaviour when in crystal forms.

This work aimed to investigate first-row transition metal complexes with dpa and halogen ligands (F, Cl, Br, and I), and looked at rationalizing their geometry, electronic stability, and reactivity.



Samuel Tetteh, *Cryst. Growth Des.* 2024, 24, 1, 506–513.

ConQuest in Action: Introducing Intramolecular π-Interactions into Heteroleptic Complexes

Here we highlight how the use of ConQuest to search the Cambridge Structural Database (CSD) helped guiding the design and synthesis of new heteroleptic compounds – coordination complexes with more than one type of ligands.

Read the full article and find out more about the CCDC data, software, and services to advance structural science that were used in this work.

Why?

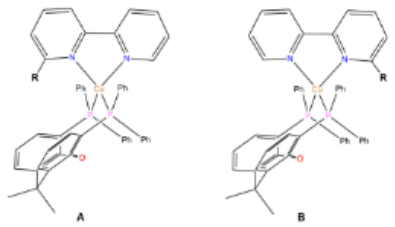
Heteroleptic copper(I) compounds can exhibit enhanced photoluminescence quantum yields (PLQY). This property is particularly interesting for applications in organic light emitting diodes (OLEDs) and light-emitting electrochemical cells (LECs).

In this work published in *CrystEngComm*, the scientists from the University of Basel reported the synthesis and characterization of six new heteroleptic copper(I) compounds. Beside performing solid-state photoluminescence studies, the team also analysed the crystal structures of the compounds and investigated the intramolecular interactions responsible for the enhanced emission.

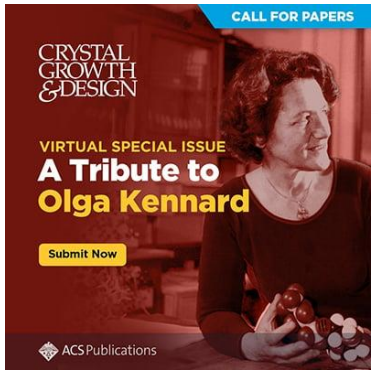
How?

Species with formula $[Cu(N^N)(P^P)]^+$, where N^N = aromatic diimine and P^P = chelating bis(phosphane) are known to be efficient LECs.

When the bis(phosphane) involved presents a wide bite angle (angle on a central atom between two bonds to a bidentate ligand), intramolecular π-stacking interactions can arise between the ligands, enhancing the emission and PLQY in these species. For this reason, the team chose to use **xantphos** and **POP** as P^P ligands.



Catherine Housecroft, et al., *Cryst. Eng. Comm* 2023, 25, 3000–3012.



CSD in Action: How Metal Ammine Complexes Interact With Aromatic Rings

Here we highlight a paper by Snežana D. Zarić and co-workers from the University of Belgrade.

In this work, the team analysed the Cambridge Structural Database (CSD) and used quantum chemical calculations to perform an in-depth study of the NH/π interactions between coordinated ammonia (NH_3) and C_6-aromatic rings.

Why?

NH/π interactions are common in nature. They are found in proteins and related structures, and are involved in important mechanisms such as the transport of ammonia through the cell membrane.

Snežana D. Zarić *et al. Cryst. Growth Des.* 2024, 24, 4, 1705–1714.

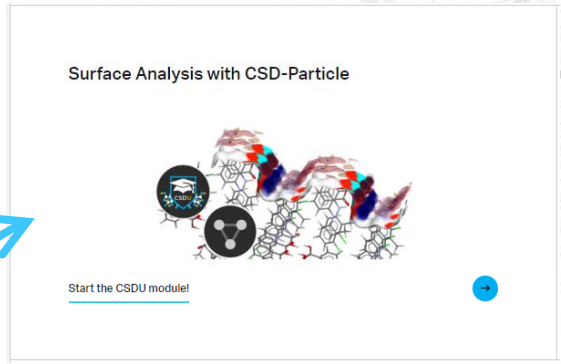
Want to explore more?

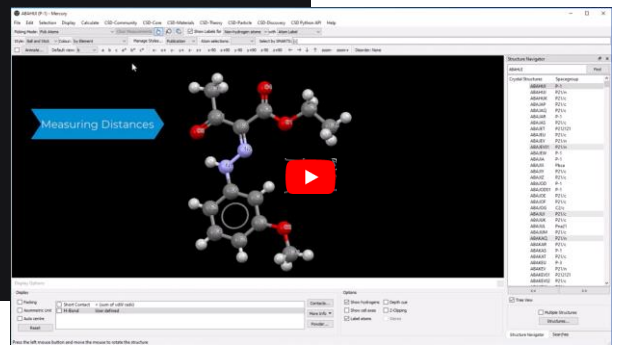
On-demand training resources

Solutions | Community | Discover | Consultancy

- Access/Deposit Structures
- CCDC for the Community
- Education and Outreach
- Events
- Free Products
- Training and Learning

- Free Online Courses
- Self-Guided Workshops
- Training and Support Videos





CCDC

Free online training courses

With completion certificates!

CSDU

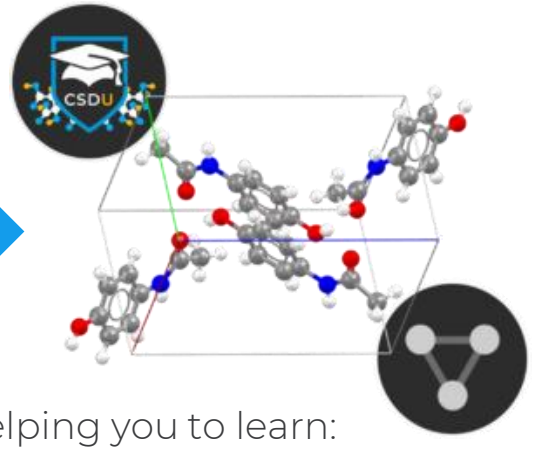
On-demand modules to learn how to use the CSD software at your own pace.

UWatch · UTry · UTest

Mercury · Python API · WebCSD · GOLD Docking

Mogul · FIMs · Particle

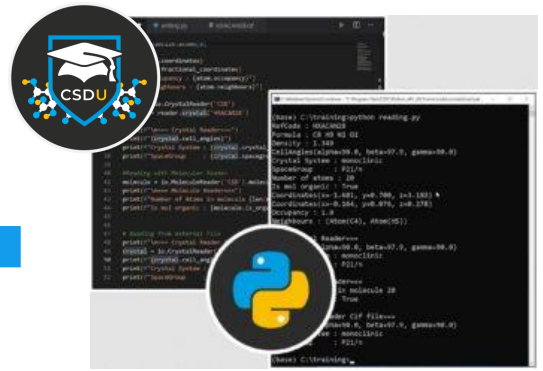
Visualization 101



Helping you to learn:

- The basics of Mercury software.
- How to explore and pack structures.
- How to create high resolution images.

CSD Python API



Helping you to learn:

- How CSD entries are represented in the API.
- How to access CSD entries programmatically.
- How to read different file formats.
- How to run a search and output results.

https://www.ccdc.cam.ac.uk/community/training-and-learning/csdu-modules/

More learning events

CCDC Virtual Workshops

- **5th Nov CSD-CrossMiner**: Introducing interactive pharmacophore searching across the CSD and the PDB.

CHEMAI Virtual Satellite Event

- **27th Nov** Unlocking **CSD data** for Functional Materials innovation.

CCDC Webinars

- **Coming soon!** The CCDC also regularly host webinars alongside these workshops. Check out our website, social media or sign-up to our newsletter to stay up to date.

https://www.ccdc.cam.ac.uk/community/events/

CCDC