

Fantastic structures and how to find them

Searching crystal structure data in the CSD with ConQuest and WebCSD

May 2022

Learning outcomes for today

- How to search over a million published crystal structures using our desktop software ConQuest.
- How to refine your results in ConQuest to structures in targeted subsets allowing you to explore drugs, pesticides, MOFs and more in more detail.
- Tips and tricks to make your ConQuest searches more effective.
- Basics of searching using WebCSD, our web browser search component.
- Structure similarity searches in WebCSD.

Exploring structural databases



PDB

>180,000

Mogul in dep,
Ligand linking
CSD-CrossMiner
BioChemGRAPH



CSD

>1.1 million
organic and
metal-organic
structures

ICSD

>250,000
Joint access
and
deposition

ICDD

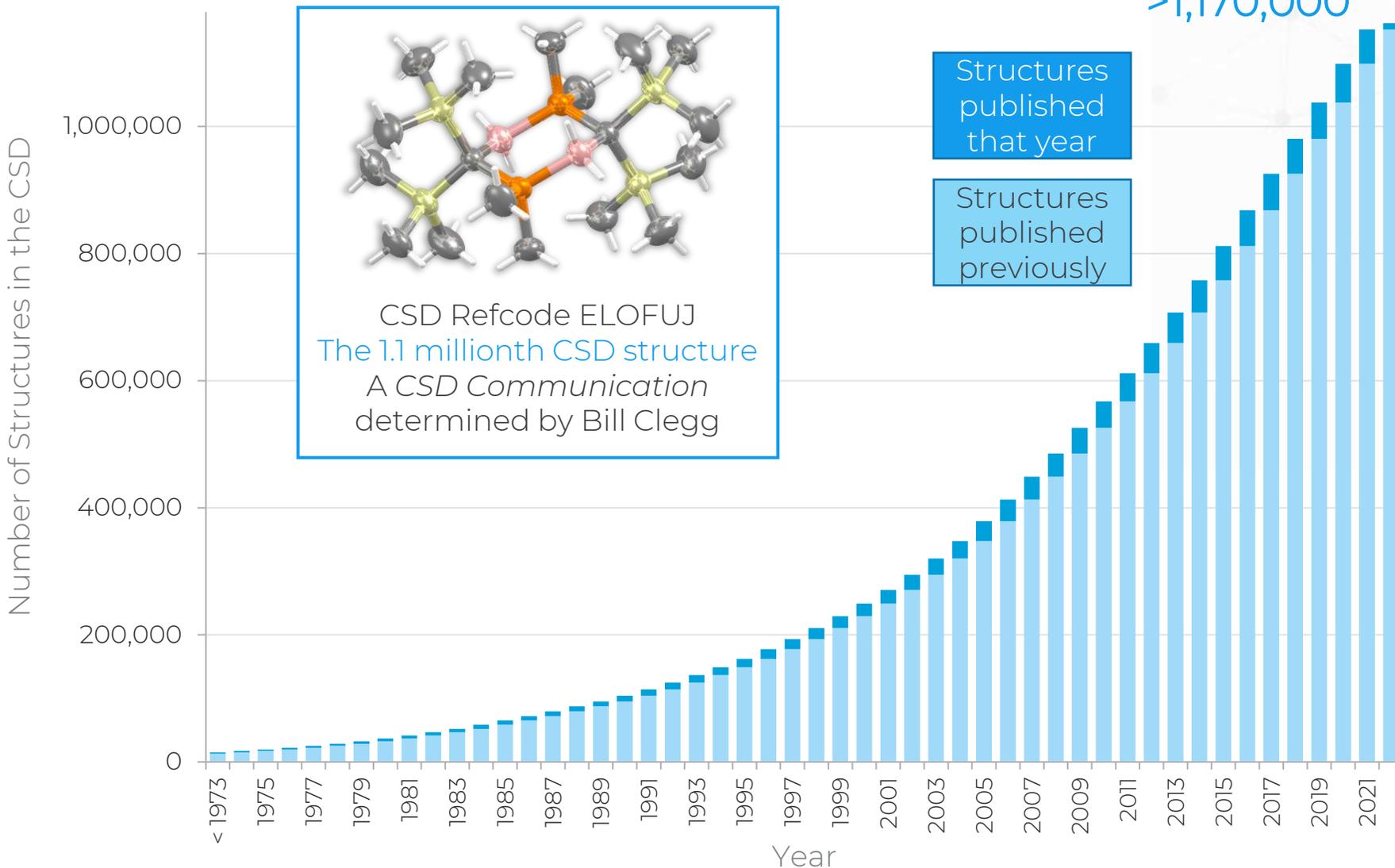
PDF-
4/Organics
>560,000
Includes data
derived from
CSD



CCDC



The Cambridge Structural Database



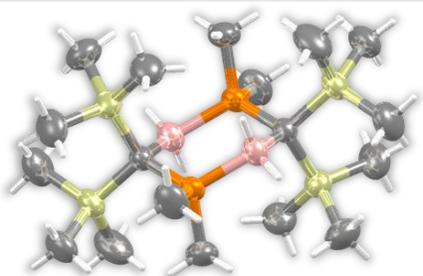
- Every published structure
 - ASAP & early view
 - CSD Communications
 - Patents
 - University repositories
 - Theses
- Every entry enriched and annotated by experts
- Discoverability of data and knowledge
- Sustainable for over 56 years
- A CoreTrustSeal repository



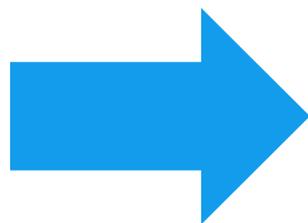
Certified as Trustworthy
by CoreTrustSeal

CCDC

CSD Refcodes



CSD Refcode -
ELOFUJ



What is ELOFUJ?

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

Refcode families

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

Inside the Cambridge Structural Database

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

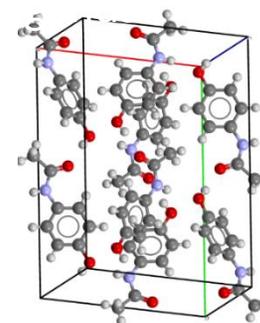
Organic
43%

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

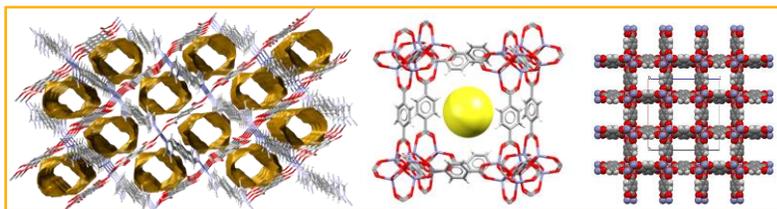
- 11,525 polymorph families
- 171,683 melting points
- 909,992 crystal colours
- 778,663 crystal shapes
- 24,916 bioactivity details
- 11,387 natural source data
- > 250,000 oxidation states

Not Polymeric
89%

Polymeric: 11%

Metal-Organic

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

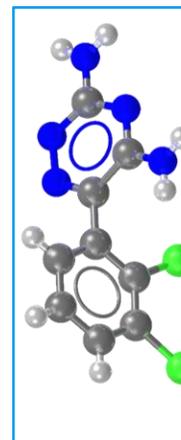


Single
Component
56%

Multi
Component
44%

Links and subsets

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

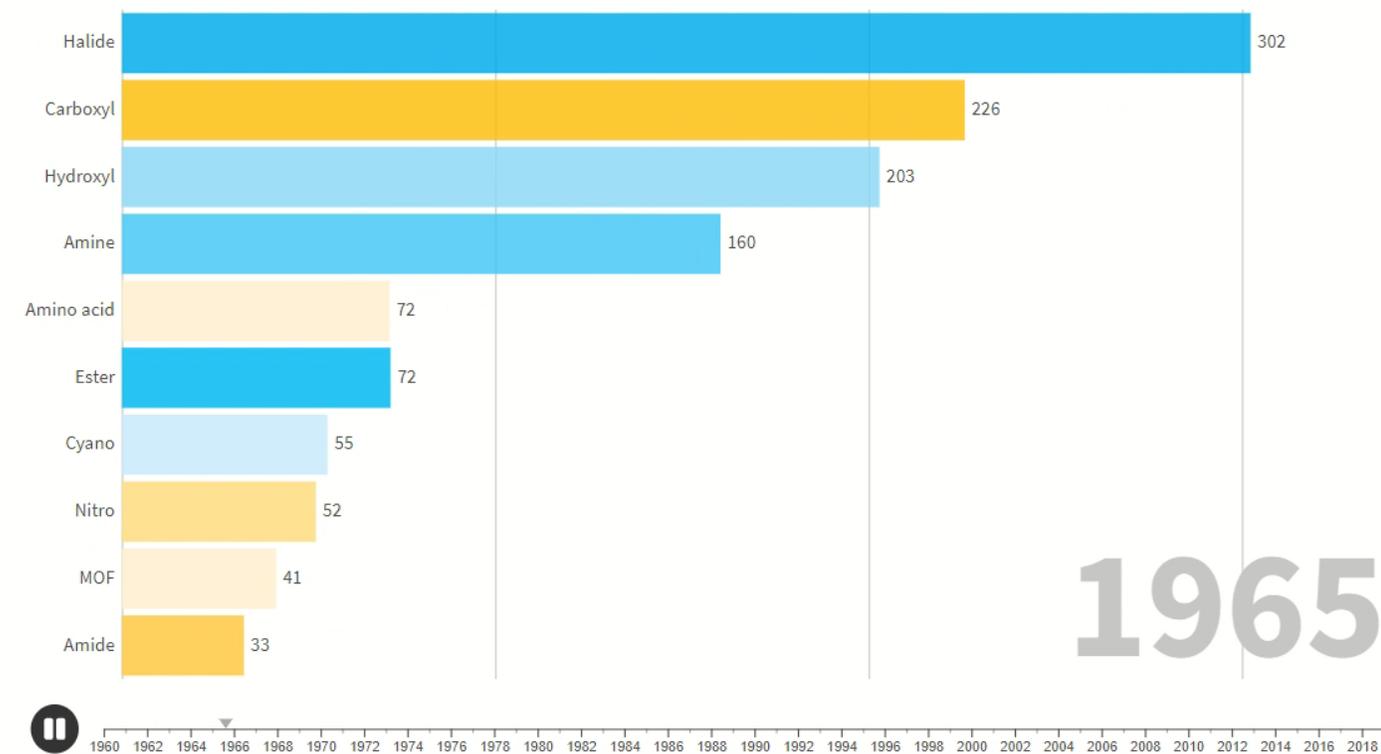


Exploring the CSD

- > 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

Chemistry in the CSD

Number of structures containing certain chemical groups

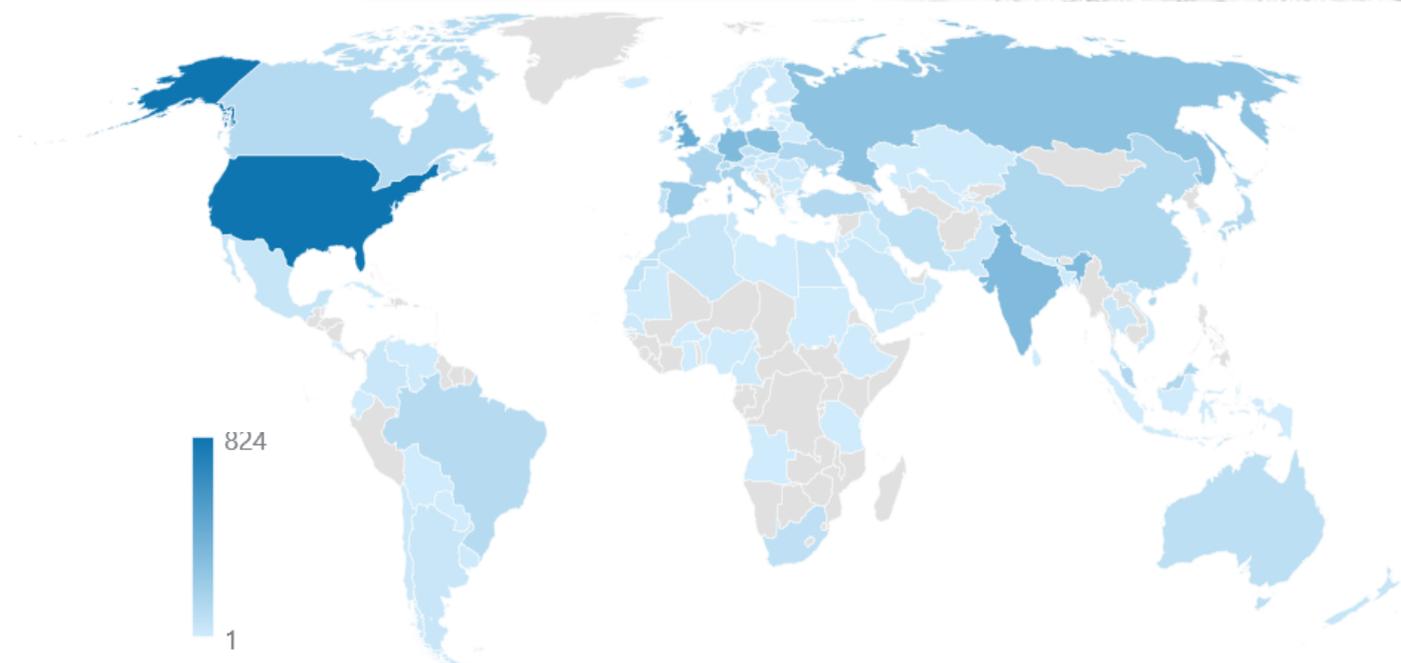
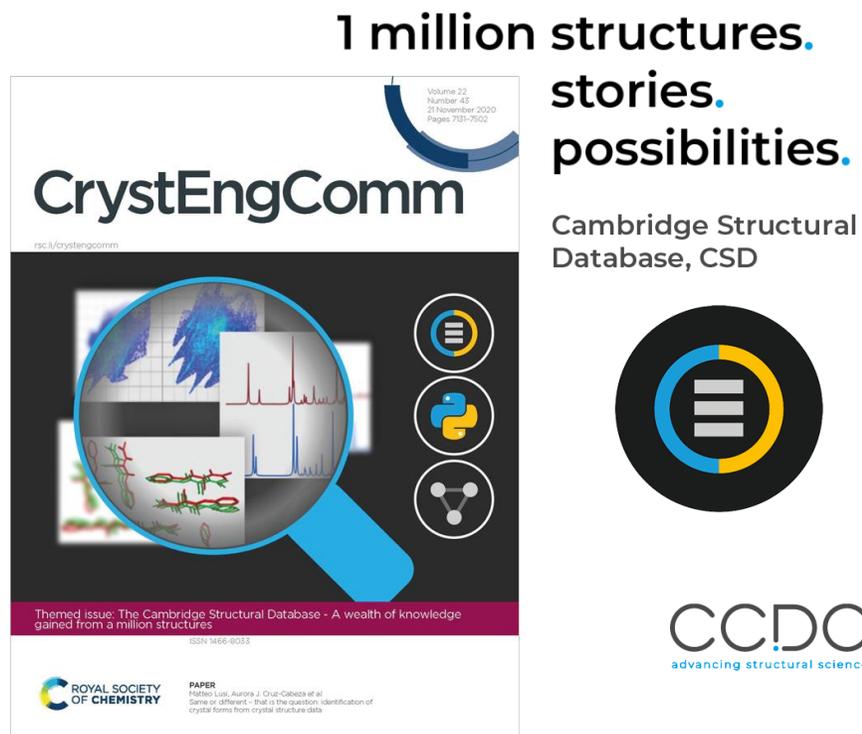


1965

A wealth of structures, stories and possibilities.

CSD one million special issue

- 33 articles from industry and academia



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

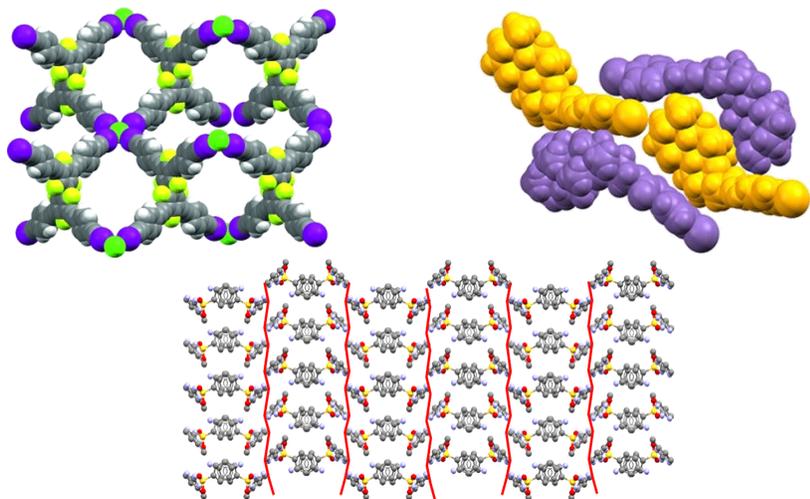
The CSD. C.R.Groom, I.J.Bruno, M.P.Lightfoot and S.C.Ward, *Acta Cryst.* (2016). B72, 171-179 DOI: 10.1107/S2052520616003954

CCDC

The whole is greater than the sum of its parts

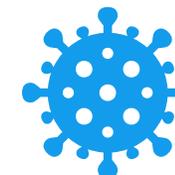
CSD Insights

- Molecular geometries
- Molecular interactions
- Molecular assemblies



CSD Applications

- Teaching
- Drug Design
- Materials Design
- Structure Solution
- Energetic materials
- Paints, Pigments and Dyes
- Organic Semiconductors
- Nonlinear Optical Materials
- Catalysts
- Gas Storage and Separation
- Crystal Engineering



The CSD Portfolio

CSDEnterprise.

CSDMaterials.



DASH



Python API



Mercury



CSDDiscovery.



SuperStar



Python API



GOLD



CrossMiner



Mercury

CSDCore.



WebCSD



Mogul



MyStructures



ConQuest



CSD



IsoStar



Mercury



Hermes



Python API



CSDCommunity.



Mercury



enCIFer



Symmetry



Deposit



CellCheck



Educational



Access



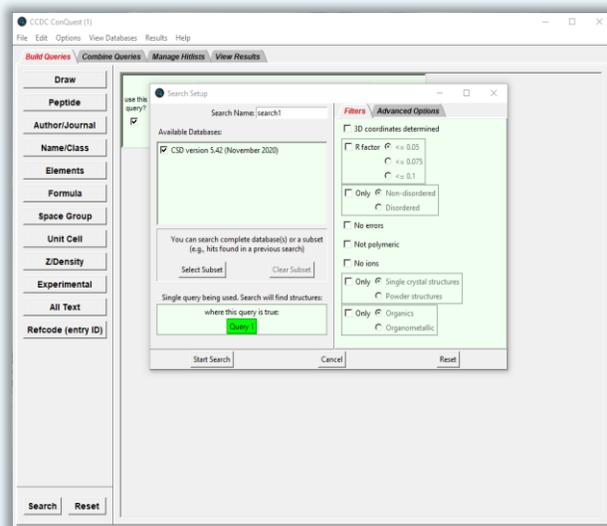
MyStructures

CSD-Core – What does it enable?

- It makes it easy to **answer very specific structure questions** on your desktop using highly flexible 3D searching of the CSD
- Quickly **identify relevant structures** across the CSD based on a wide range of properties including chemical constraints such as cyclicity
- Use **interaction and structure-property knowledge** gained to drive design decisions

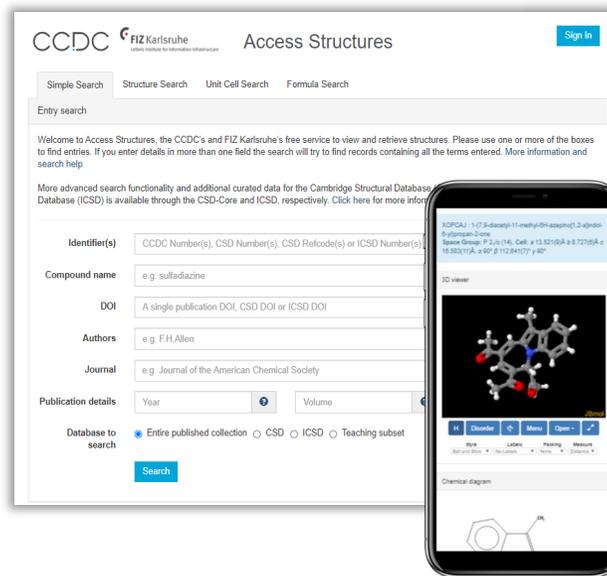
Finding fantastic structures

Desktop



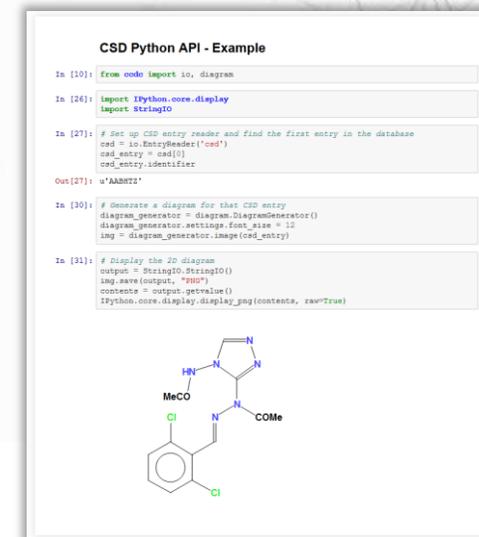
 ConQuest

Web Browser



 WebCSD /
Access Structures

Programmatic



```
CSD Python API - Example

In [10]: from ccsd import io, diagram
In [24]: 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'AASHTZ'

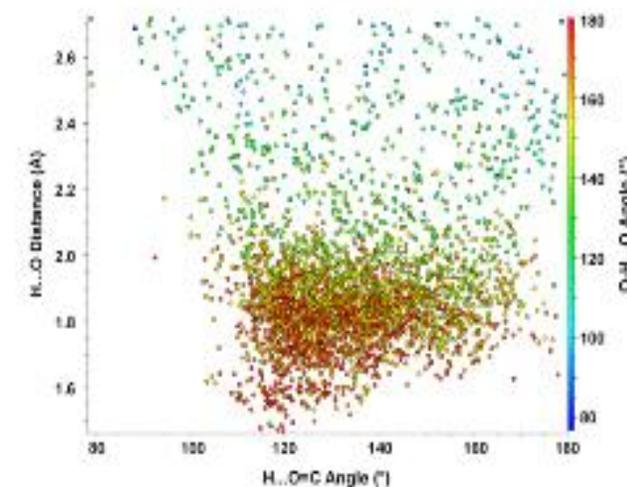
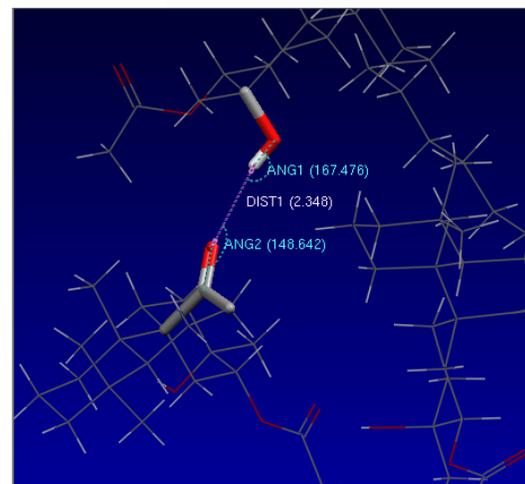
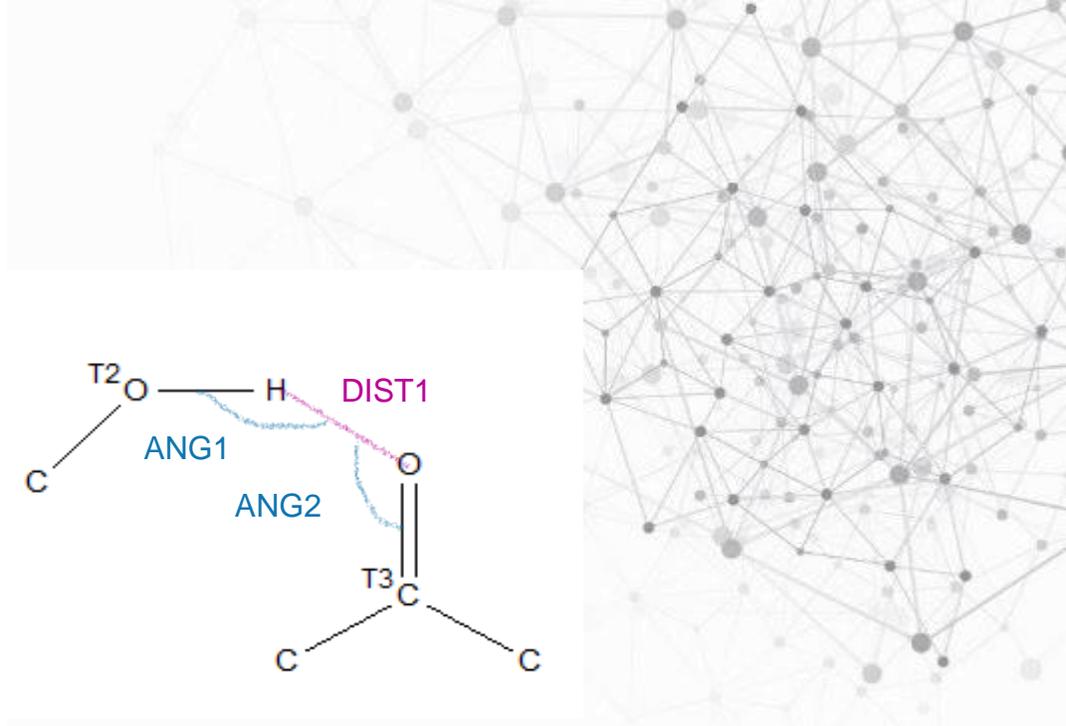
In [30]: # Generate a diagram for that CSD entry
diagram_generator = diagram.DiagramGenerator()
diagram_generator.settings.font_size = 22
img = diagram_generator.image(csd_entry)

In [31]: # Display the 3D diagram
output = StringIO.StringIO()
img.save(output, "png")
contents = output.getvalue()
IPython.core.display.display_png(contents, raw=True)
```

 CSD Python API

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](#)



ConQuest searches

- **Filtering options** based on experimental or chemical considerations (e.g., *R*-factor, temperature, only organics)
- **Combine search terms** with Boolean options and manage hitlists post-search
- **Export results** to Mercury for visualisation and advanced numerical analysis & plotting

The screenshot displays the CCDC ConQuest (1) software interface. The main window has a menu bar (File, Edit, Options, View Databases, Results, Help) and a toolbar with tabs for Build Queries, Combine Queries, Manage Hitlists, and View Results. The main area is divided into sections for defining search criteria:

- Find entries that:**
 - must have (boolean AND):** Contains Query 1 (with a chemical structure icon) and Query 2.
 - must not have (NOT):** Contains Query 3.
 - must have at least one of (OR):** An empty box.

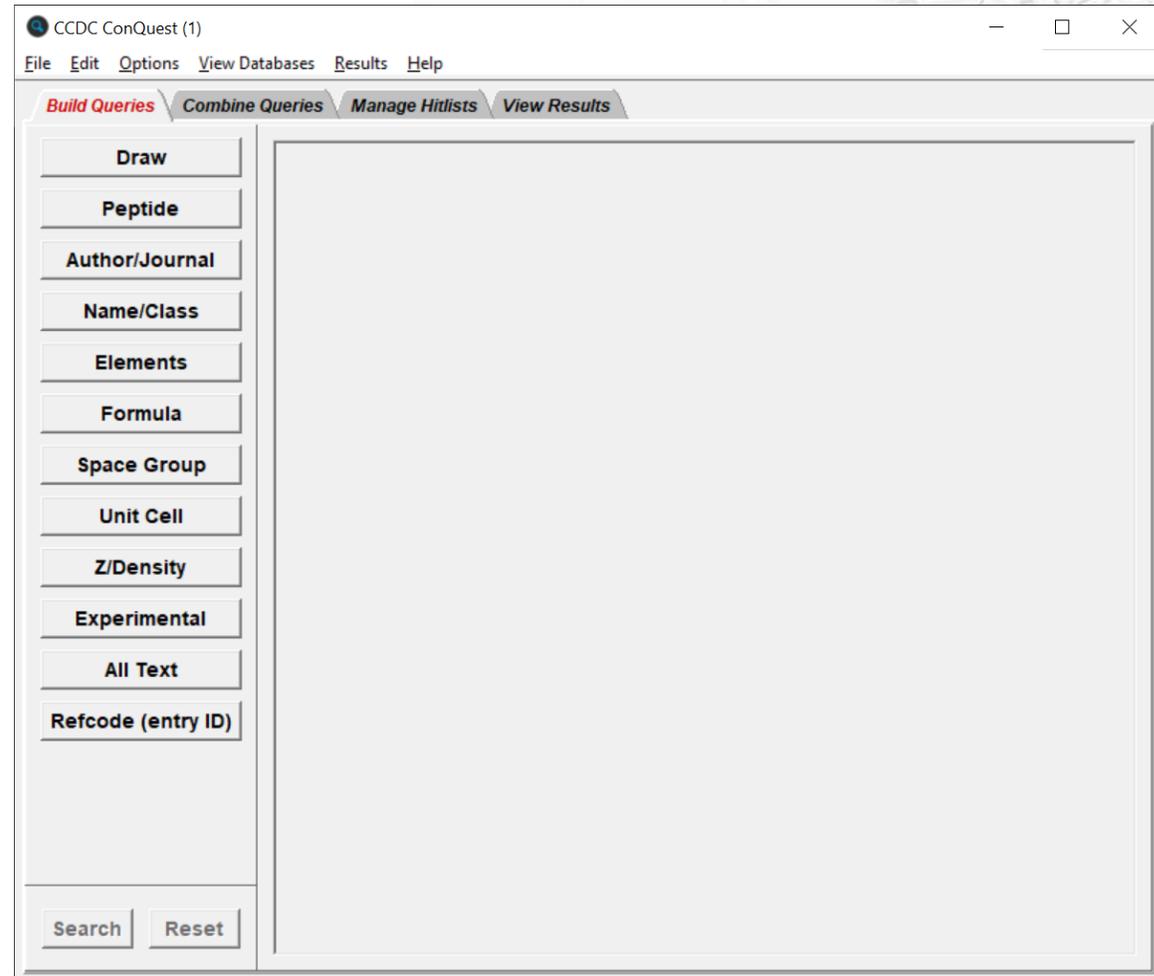
Buttons for Search and Reset are at the bottom of this section.

A **Search Setup** dialog box is open, showing:

- Search Name:** search1
- Available Databases:** CSD version 5.43 (November 2021) + 1 update
- 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

Buttons for Start Search, Cancel, and Reset are at the bottom of the dialog. A summary section at the bottom of the dialog shows: "Summary of queries to be used. Search will find structures: where these queries are true: Query 1, Query 2; where this query is not true: Query 3."

Show One: 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?

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

Search Setup

Search Name: search1

Available Databases: Show Updates separately

CSD version 5.43 (November 2021) + 1 update

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

Select Subset Clear Subset

Single query being used. Search will find structures: where this query is true: Query 1

Start Search Cancel Reset

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

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

Refcode: COVZAR CSD version 5.41 updates (Mar 2020)

COVZAR Analyse Hitlist

COVZAR

DUBROK

DUCTAZ

DUCTED

KOYJEQ

LOYFAJ

NOTMER

NOTMIV

NOWGOY

VOZRUA

WOTJEX

WOWWUD

YOYLUW

YOYMAD

ABOKUZ

ABOLAG

ABOLEK

ABOLIO

ABOLUA

ACFCUA

ACFCUB

965 hits 100%

Show terminal carbons Use as Query... Detach Stop Search

Author and journal search

Build Queries / **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

Z/Density

Experimental

All Text

Recode (entry ID)

Author/Journal (1) - Query 5

Authors' Names New Box

M.Zema 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]
- 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.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**

Search Setup

Search Name: search1

Available Databases: Show Updates separately

- CSD version 5.43 (November 2021) + 1 update

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

Select Subset **Clear Subset**

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

Start Search **Cancel** **Reset**

Results

Query highlighted

Left clicking each tab will display different information

CCDC ConQuest (1) : search1 [Search]

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

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	$C_{10}H_{24}N_4Ni^{2+} \cdot 2(ClO_4^-)$
Compound	(3-(4-(3-Aminopropyl)piperazin-1-yl)propylamine)-nickel(ii) diperchlorate
Spacegroup	Name: P6ca 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 c: 16.171 alpha: 90.00 beta: 90.00 gamma: 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

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 data to publication

CCDC ConQuest (1) : search1 [Search]

McKervey, A. R. Maguire, S. M. Tuladhar and M. Fiona Twohig, *J. Chem. Soc.* 1047–1054 DOI: [10.1039/P19900001047](https://doi.org/10.1039/P19900001047); (b) H. Duddeck, *J. Chem. Soc.*, 1055–1063 DOI: [10.1039/P19900001055](https://doi.org/10.1039/P19900001055); (c) P. Parne and J. M. Fox, *J. Appl. Cryst.* 1990, 23, 1055–1063 DOI: [10.1039/P19900001055](https://doi.org/10.1039/P19900001055); External Links.

ROYAL SOCIETY OF CHEMISTRY

Footnote

Electronic supplementary information (ESI) available: Experimental procedure for the synthesis of 15 and 16. For ESI and crystallographic data in CIF or other electronic format see DOI: [10.1039/C9PY00061F](https://doi.org/10.1039/C9PY00061F).

This journal is © The Royal Society of Chemistry 2009

CCDC ConQuest (1) : search1 [Search]

Databases Results Help

Refcode: AVETAW CSD version 5.43 (November 2021)

Author(s)	M. Bolocchi, M. Bonizzoni, L. Fabbrizzi, F. Foti, M. Licchelli, A. Taglietti, M. Zema
Reference	10.1039/C9PY00061F , 653
Publication DOI	10.1039/b312980b
Chemical Formula	$C_{10}H_{24}N_4Ni^{2+} \cdot 2(ClO_4^-)$
Compound	(3-(4-(3-Aminopropyl)piperazin-1-yl)propylamine)-nickel(II) diperchlorate
Spacegroup	Name: P6ca Number: 61
Reduced 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
Molecular Volume	448.788
Chemical Units	2

ACS Publications
Most Trusted. Most Cited. Most Read.

Polymorph α

Polymorph β

Polymorph β , is exceptionally resistant to forming to the β phase. We have studied the transition between polymorph α and β in resorcinol polymorphs in a range of solvents and pressures. The transition between polymorph α and β is reversible and occurs at pressures above 0.5 GPa, consistent with the calorimetric and NMR results. The high-temperature, high-pressure polymorph β achieves the more dense packing through the changed conformation of one of the hydroxyl groups and the considerable twisting of the hydrogen bonds necessary for the formation of additional C–H... π bonds. The large temperature and pressure hysteresis of the polymorphs α and β are consistent with the different topologies of their O–H...O networks.

Synthesis

The pressure- and temperature-dependent transition between resorcinol polymorphs α and β has been rationalized in terms of transforming H-bonds and their networks.

Introduction

Resorcinol, an intermediate often used in chemical practice and a pharmaceutical agent, was one of the first organic compounds for which the phenomenon of polymorphism was described and the first organic compound for which the structures of both polymorphs were determined in 1938 by Robertson and Ubbelohde.^(1,2) Until today, the resorcinol crystals belong to the best known examples of polymorphs.⁽³⁾

Keywords: celecoxib; active pharmaceutical ingredient; API; solvate; crystal structure; isostructurality; disorder; PIXEL; anti-inflammatory.

CCDC references: 2011633; 2011634; 2011635; 2011636; 2011637; 2011638

Similar articles PowerPoint slides

IUCr Journals
CRYSTALLOGRAPHY JOURNALS ONLINE

Wiley Online Library

Two successive benzyne cycloadditions which is slightly higher than that obtained in the case of 15 (1.2). Bis-cycloadduct 15 was subjected to aromatization (TiCl₄, Zn, THF, RT, 1 h)[16] followed by hydrolysis of the silyl acetal (16) to give 17. The aromatization of 16 was carried out in a solvent mixture of CH₂Cl₂ and MeOH at 0 °C, 1 h. The aromatization proceeded well, and the yield of 17 was 73%. The E/Z content of 17 was determined to be 1:1. Nitrile oxide 18 could be stored at –18 °C for at least one month.

16 M. A. Meador, H. Hart, *J. Org. Chem.* 1989, 54, 2336–2341.

CrossRef | CAS | Web of Science® Times Cited: 16 | ejournals@cambridge.org - find full text'

17 CCDC 1543805 (16) contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre.

Outline

Abstract

Graphical abstract

1. Introduction

2. Results and discussion

3. Conclusions

4. Experimental

Acknowledgements

Supplementary data

Research Data

References and notes

An efficient phosphate sensor: tripodal quinoline excimer transduction

ELSEVIER SCHOLIX

Research data for this article

Cambridge Crystallographic Data Center
Crystallographic data
Data associated with the article:
CCDC 689113: Experimental Crystal Structure Determination

References

Accelrys Software Inc., San Diego, CA, USA. Google Scholar

Cryst. Growth Des. **12**, 2147–2152. Web of Science CrossRef CAS Google Scholar

gia, A. (2014). *CrystEngComm*, **16**, 24–27. Web of Science CSD CrossRef CAS

IUCrJ, **6**, 751–760. CSD CrossRef CAS PubMed IUCr Journals Google Scholar



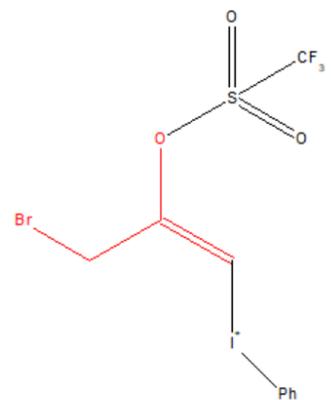
CCDC ConQuest (1) : search3 [Search]

File Edit Options View Databases Results Help

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

Refcode: COVZAR CSD version 5.41 updates (Mar 2020)

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



Show terminal carbons
 Use as Query... Detach

COVZAR

Analyse Hitlist

- COVZAR
- DUBROK
- DUCTAZ
- DUCTED
- KOYJEQ
- LOYFAJ
- NOTMER
- NOTMIV
- NOWGOY

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



Left clicking each tab will display different information



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

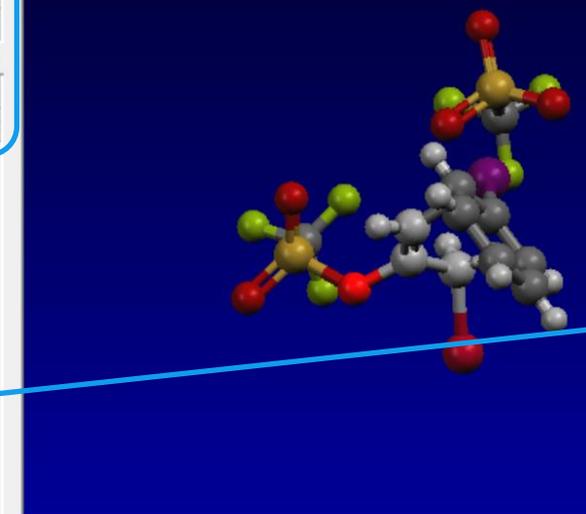
CCDC ConQuest (1) : search3 [Search]

File Edit Options View Databases Results Help

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

Refcode: COVZAR CSD version 5.41 updates (Mar 2020)

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



Show substructure matches
 Right-click in visualiser for options menu

COVZAR

Analyse Hitlist

- COVZAR
- DUBROK
- DUCTAZ
- DUCTED
- KOYJEQ
- LOYFAJ
- NOTMER

Display Style

- Wireframe
- Capped Sticks
- Ball & Stick**
- Spacefill

YOYLW
 YOYMD
 ABOKUZ
 ABOLAG
 ABOLEK
 ABOLIO
 ABOLUA
 ACFCUA
 ACFCUB

965 hits 100%

Stop Search

The screenshot displays the CCDC ConQuest (1) : search3 [Search] application window. The main interface shows a search results view with a 3D molecular model of a complex organic structure. The 'File' menu is open, highlighting the 'Export Entries as...' option. An 'Export Entries: search3' dialog box is overlaid on the right, allowing the user to select a file type (TAB: Tab separated list), choose what to export (All selected entries), and select options (Bibliographic, Chemical, Crystallographic, Experimental, CSD internals). The dialog also includes fields for filename and save location, a progress bar at 0%, and a 'Keep window open when finished' checkbox.

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

ConQuest search features

Experimental (1) - New

R-factor: [] fractional [x] %

Exclude disordered structures

Exclude structures with unresolved errors

Average e.s.d. of C-C Bonds: Any []

Exclude powder structures

Room Temperature: 610K

Radiation Source: Any []

Search Store Cancel Reset

Search for information relating to the structure determination

Name/Class (1) - New

Compound Name: [] 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

Chemical Class

Find entries classified as: and: []

- not defined----
- Carbohydrates
- Nucleosides & nucleotides
- Amino-acids, peptides & complexes
- Porphyryns, corrins & complexes
- Steroids
- Terpenes
- Alkaloids
- Organic polymers

Search Store Cancel Reset

Chemical name searching and categories

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
- blade
- block
- blue
- brown
- colorless
- column
- conductor
- cream
- cube

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

Generic text search

Author/Journal (1) - New

Authors' Names: [] 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]
- A.K.U.Int'l.J.Engg.Tech.App.Sci. [2004-2020]
- AAPS PharmSciTech [2004-2020]
- ACA Abstr.Papers(Winter) [1967-1977]
- ACA.Ser.2 [1977-1984]
- ACGC Chem.Res.Comm. [2001-2002]
- ACH-Models Chem. [1994-2000]

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

Bibliographic search

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

Include alternative settings of selected space group

Spacegroup Symmetry: ----not defined----

Crystal System: ----not defined----

Search Store Cancel Reset

Search by space group, crystal system or space group symmetry

Formula (1) - New

Formula: []

Type in formula, e.g. C6 H12 O6. You may specify an inexact formula to find entries with 1,2 or 3 Calciums.

Formula applies to:

an individual molecule

all molecules in structure added to the search

Heaviest Permitted Element in Formula Unit: []

Search Store Cancel Reset

Elemental and formula search

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

Distance: 1.5 % of longest cell dimension

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

Cell Parameters

a (A) = [] alpha (*) = []

b (A) = [] beta (*) = []

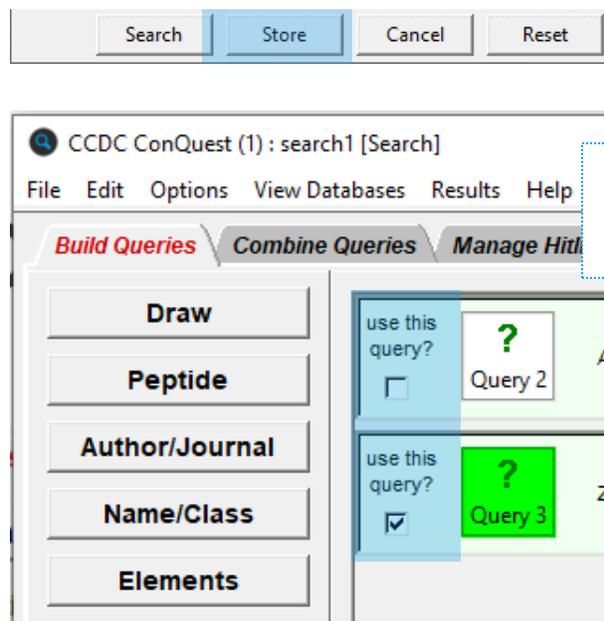
c (A) = [] gamma (*) = []

Search Store Cancel Reset

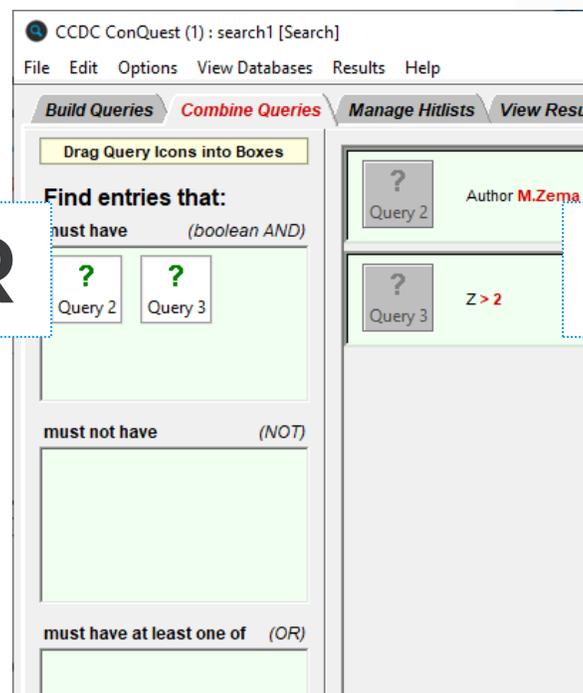
Unit cell search

Combining searches and results

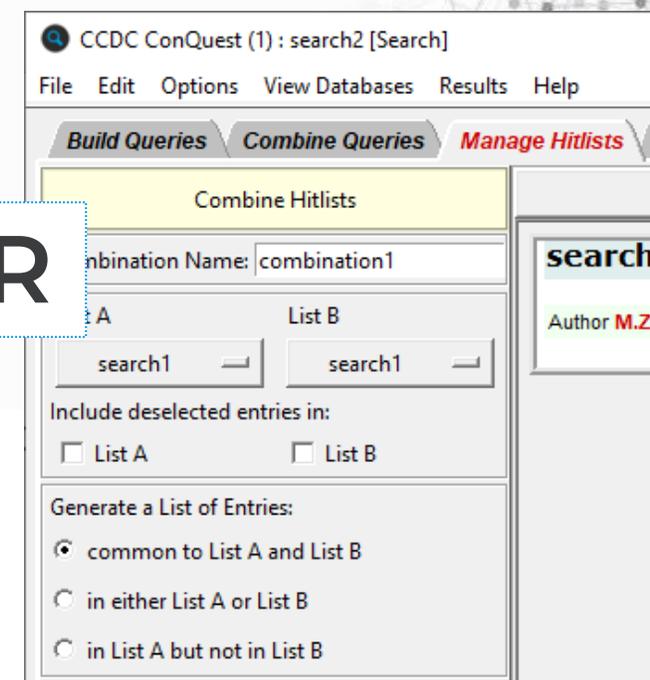
Store and Search



Combine Queries



Manage Hitlists



OR

OR

Store a query shown on Z/Density search

Build Queries **Combine**

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	>	4

Search **Store** **Reset**

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

use this query? **Query 2** Author **M.Zema** Edit... Delete

use this query? **Query 3** Number of Residues > 2 Edit... Delete

CCDC

Combine Queries

Build Queries Combine Queries Manage Hitlists View Results

Allows you to combine different queries using boolean AND, NOT, OR

Query 1: Authors include M.Zema
Query 2: Number of Residues > 2

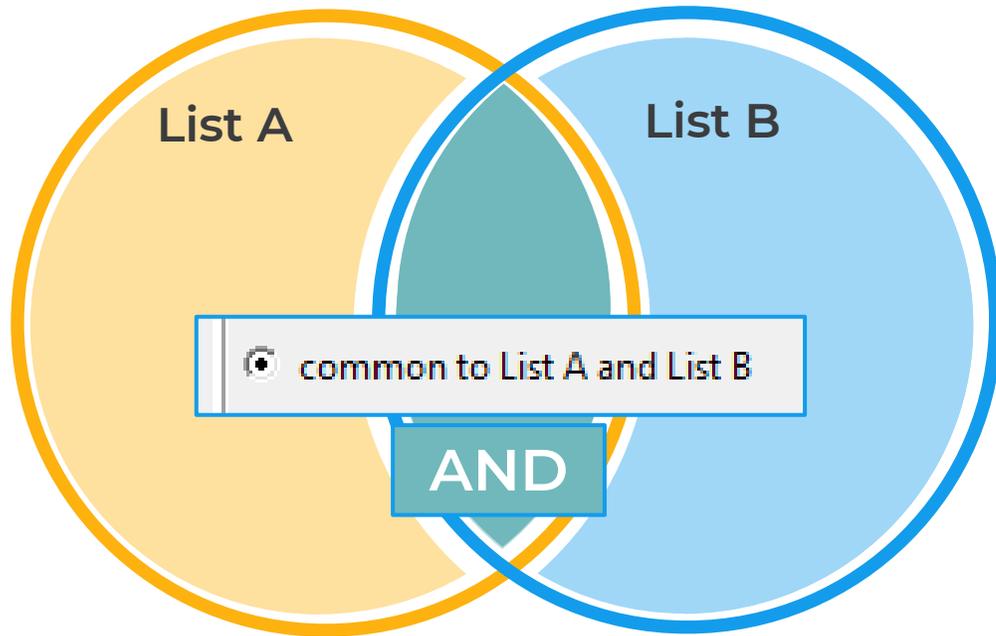
Must have: Query 1 AND Query 2

The screenshot displays the CCDC ConQuest (2) software interface. The main window is titled "CCDC ConQuest (2)" and has a menu bar with "File", "Edit", "Options", "View Databases", "Results", and "Help". The interface is divided into several sections:

- Build Queries**: A section with a tab "Combine Queries" selected. It contains a "Find entries that:" section with "must have (boolean AND)" and "must not have" options. Two query boxes, "Query 1" and "Query 2", are visible. Below them is a "must have at least one of" section.
- Combine Queries**: A section showing two query boxes. The first box is labeled "Query 1" and contains "Author M.Zema". The second box is labeled "Query 2" and contains "Number of Residues > 2". Each box has "Edit..." and "Delete" buttons.
- Search Results**: A window titled "CCDC ConQuest (2) : search1 [Search]" is open, showing search results for "BAQFOO". The results are displayed in a table with columns for "Refcode", "CSD version", and "Author(s)". The "Author(s)" column shows "M.Biagioli, L.Strinna-Erre, G.Micera, A.Panzanelli, M.Zema". The "Number" column shows "61". The "Chemical Units" column shows "4". The "Molecular Volume" is "893.106" and "Z" is "4.0". The "Z'" is "0.5". A chemical structure diagram is shown below the table.
- Hitlist**: A section on the right side of the search results window, titled "BAQFOO", with an "Analyse Hitlist" button. The hitlist shows "BAQFOO", "IBIZAU", and "KEMER" with checkmarks.

Manage Hitlists

- **Hitlists** = results from searches that you have already run



The screenshot shows the CDC ConQuest software interface. The main window is titled "CCDC ConQuest (1) : search2 [Search]". The "Manage Hitlists" tab is active. The "Combine Hitlists" dialog box is open, showing the following options:

- Combination Name: combination1
- List A: search1
- List B: 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

The "Hitlist Overview" window shows a table with the following data:

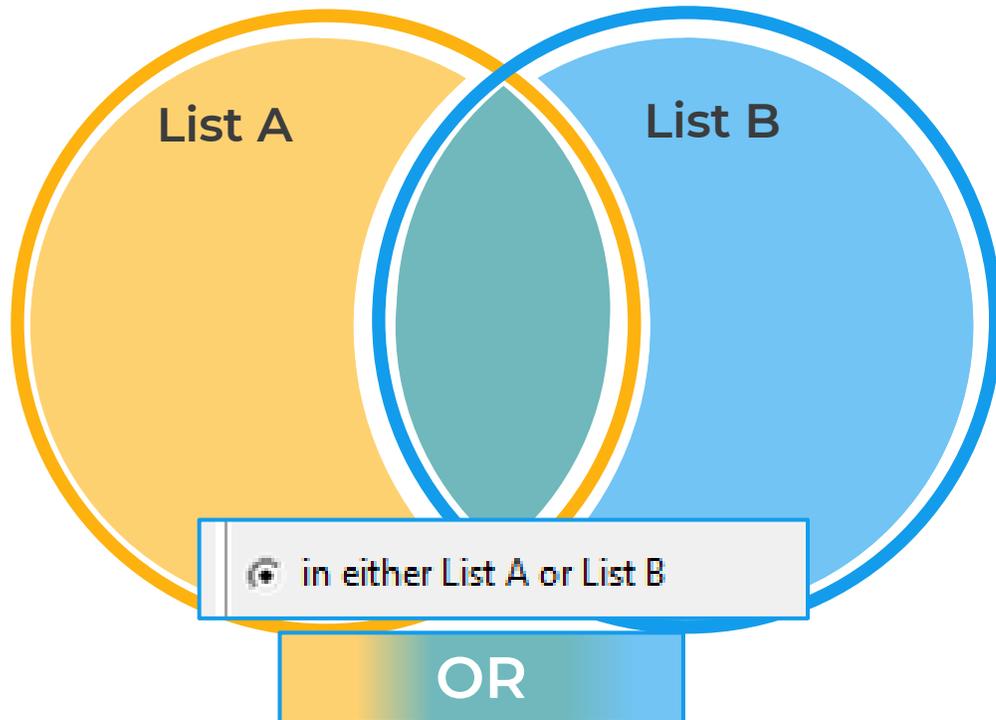
Name	Hits	Type
search1	50	Search
search2	32	Search

Callouts indicate "Combine results" pointing to the "Combine Hitlists" dialog and "Rename searches" pointing to the "Rename..." button in the Hitlist Overview window.

Manage Hitlists

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

- **Hitlists** = results from searches that you have already run



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 (50 Entries)

Author M.Zema

Combine results

Name	Hits	Type
search1	50	Search
search2	32	Search

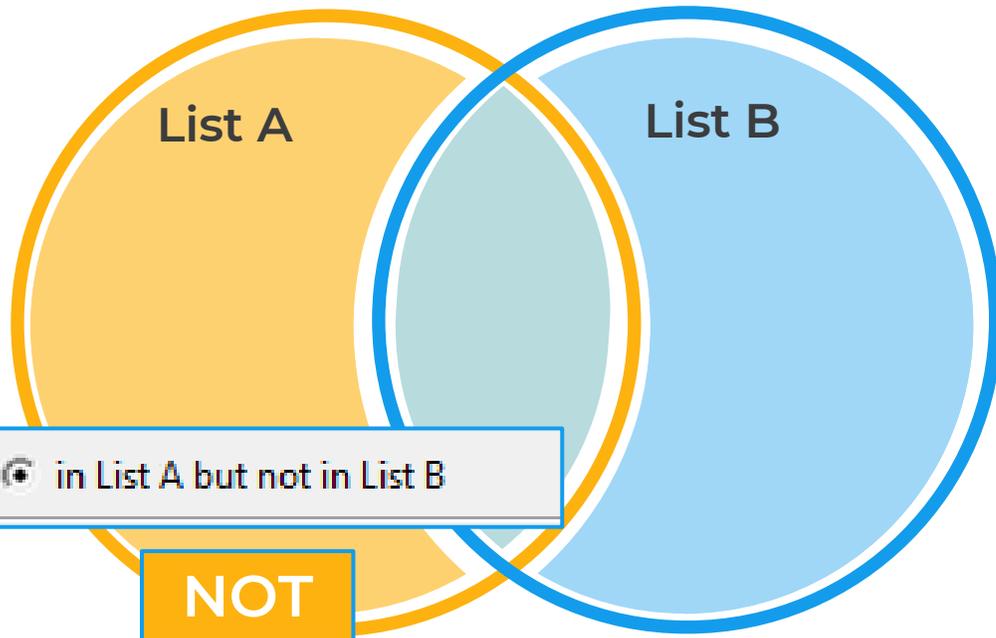
Rename searches

Delete Rename... Notes... View

Manage Hitlists

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

- **Hitlists** = results from searches that you have already run



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 (50 Entries)

Author M.Zema

Name	Hits	Type
search1	50	Search
search2	32	Search

Delete Rename... Notes... View

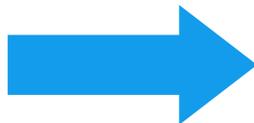
Combine results

Rename searches

Compound name search

Build Queries / **Combine**

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



Name/Class (1) - New

Compound Name

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

Chemical Class

CCDC Chemical C...
entries in the ...
difficult to lo...
substructure or...

Note that the r...
comprehensive s...
publication tha...
specified categ...

ories of
d be

y not be
from a
he

----not defined----
Carbohydrates
Nucleosides & nucleotides
Amino-acids, peptides & complexes
Porphyrins, corrins & complexes
Steroids
Terpenes
Alkaloids
Organic polymers

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

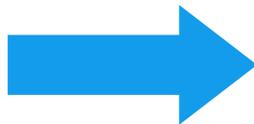
and: ----not defined----

Search Store Cancel Reset

Space Group and cell search

Build Queries Combine

- Draw
- Peptide
- Author/Journal
- Name/Class
- Elements
- Formula
- Space Group**
- Unit Cell**
- Z/Density
- Experimental
- All Text
- Recode (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 -----not defined-----

Crystal System -----not defined-----

Search Store Cancel Reset

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 % of longest cell dimension

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

Cell Parameters

a (A)	=	<input type="text"/>	alpha (°)	=	<input type="text"/>
b (A)	=	<input type="text"/>	beta (°)	=	<input type="text"/>
c (A)	=	<input type="text"/>	gamma (°)	=	<input type="text"/>

Search Store Cancel Reset

Elements and formula search

Build Queries **Combine**

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

Elements (1) - New

Elements Required to be Present

Type in elements, e.g. C H Se
or

Elements must be in

same molecule

same crystal structure

Other elements allowed in molecule/structure

Heaviest Permitted Element in Formula Unit

Formula (1) - New

Formula

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

Formula applies to

an individual molecule

all molecules in structure added together

Other atoms allowed in molecule/structure

Z/Density and experimental info search

Build Queries / **Combine**

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

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/cm ³)	=	
No. of Chemical Units (molecules, ions, etc.) in Entry	=	

Search Store Cancel Reset

Experimental (1) - New

R-factor = 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 = K °C

0 Room Temperature 610K

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

Radiation Source Any

Search Store Cancel Reset

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 connection.

Next Atom: C
Next Bond: Single

DRAW

EDIT

ERASE

ADD 3D

CONTACT

RingMaker

Templates...

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

3D Parameters:

Options...

Delete

Contacts:

Options...

Delete

Search

Store

Cancel

Example of a search, where 7A represents any halogen

Draw

Build Queries **Combine**

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

The screenshot shows the 'Draw (1) - New' window with a menu bar (File, Edit, Atoms, Bonds, 3D, Options, Help) and a toolbar. The 'Atoms' menu is open, showing options like 'Element', 'Add Group', 'Expand Chemical Groups', 'Hydrogens', 'Charge', 'Number of Bonded Atoms', 'Cyclic', and 'All Atoms in Same Molecule'. The 'Options' menu is also open, showing 'Drawing Options...', 'Snap to Grid', 'Save Window Size', 'Limit Window Size', 'Auto-Generate H', and 'Balloon Help'. The toolbar includes buttons for 'DRAW', 'EDIT', 'ERASE', 'ADD 3D', 'CONTACT', 'RingMaker', and 'Templates...'. The 'Templates...' button is highlighted with a blue box and an arrow pointing to a blue callout box. The 'RingMaker' button is also highlighted with a blue box and an arrow pointing to a blue callout box. The 'Number of Bonded Atoms' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Auto-Generate H' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Drawing Options...' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Snap to Grid' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Save Window Size' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Limit Window Size' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Auto-Generate H' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Balloon Help' menu item is highlighted with a blue box and an arrow pointing to a blue callout box. The 'Next Atom: C' and 'Next Bond: Single' status bar is highlighted with a blue box and an arrow pointing to a blue callout box. The '3D Parameters' panel is visible on the right side of the window. The 'Contacts' panel is also visible on the right side of the window. The 'Search', 'Store', and 'Cancel' buttons are visible at the bottom right of the window.

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

ConQuest – Draw/Structure search

The screenshot displays the ConQuest software interface. On the left is a sidebar with buttons for 'Draw', 'Peptide', 'Author/Journal', 'Name/Class', 'Elements', 'Formula', 'Space Group', 'Unit Cell', 'Z/Density', 'Experimental', 'All Text', and 'Refcode (entry ID)'. The main window is titled 'Draw (1) - New' and contains a menu bar (File, Edit, Atoms, Bonds, 3D, Options, Help) and a toolbar with buttons for DRAW, EDIT, ERASE, ADD 3D, and CONTACT. A chemical structure of a six-membered ring with a nitrogen atom and a red atom is shown. A context menu is open over the structure, listing options like Element, Add Group, Hydrogens, Charge, Number of Bonded Atoms (selected), Cyclicity, and Delete Atom. A sub-menu for 'Number of Bonded Atoms' is also open, showing options 1 through 8 and 'Unspecified'. A '3D Parameters' panel is visible on the right. Below the main window, an 'Other Atom Type' dialog box shows a periodic table with a search bar and 'Picking Mode' options. A context menu is also open over the periodic table, listing options like 'Not Hydrogen C or H', 'Any Non-metal', 'Any Metal', 'Any Transition Metal', 'Any Halogen', and 'Other Elements...'. At the bottom, there is a search bar with 'C H O N S P F Cl Any More... Groups...' and a 'Bond: Single' dropdown menu.



Left click in space to add an atom



Left click and drag from an atom to attach an atom



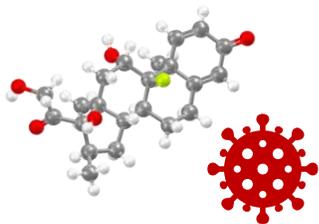
Right click on an atom or bond to edit or add extra properties to it

Making Drawing Easier: RingMaker

The screenshot displays the RingMaker software interface. A central dialog box titled "RingMaker" is open, showing a "Ring Size" input field with the value "8". Below this, a list of bond types is shown with checkboxes: Single, Double, Triple, Quadruple, Aromatic, Polymeric, Delocalised, Pi, and Any. The "Aromatic" checkbox is selected. The "OK" button is highlighted with a yellow box. At the bottom of the dialog, a yellow bar displays "Size: 8 Bonds: Aromatic".

In the background, the main drawing area shows a skeletal structure of an 8-membered ring with alternating single and double bonds, with each carbon atom labeled "C". A yellow box highlights the "Next Atom: C" and "Next Bond: Single" status. The interface includes a menu bar (File, Edit, Atoms, Bonds, 3D, Options, Help), a toolbar on the left with buttons for DRAW, EDIT, ERASE, ADD 3D, and CONTACT, and a bottom toolbar with element buttons (C, H, O, N, S, P, F, Cl, Any, More...), Groups..., and Bond: Single. A yellow box highlights the RingMaker icon in the bottom-left toolbar.

COVID-19



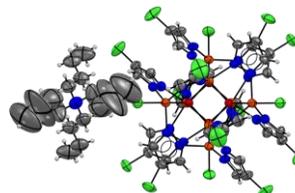
Pesticides



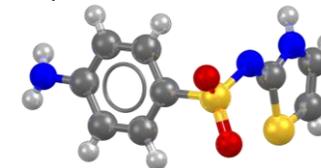
Drugs



ADPs



Best
representative



MOF
Dimensionality

- 1D
- 2D
- 3D

MOFs



Teaching

CSD Subsets

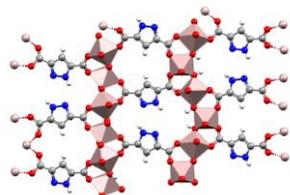
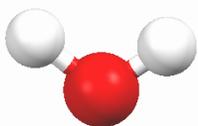
Groups of structures that may be more difficult to find in CSD from searching alone

Best representative lists provide a single example of every structure in the CSD (including polymorphs) under a particular condition;

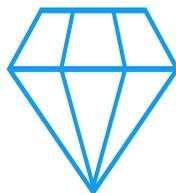
- low temperature,
- room temperature,
- with 3D determined hydrogen atoms,
- or the lowest R factor measurement

New in CSD v5.43

Hydrates

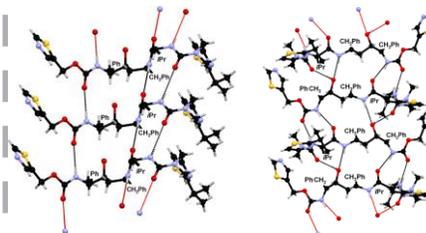


Electron
Diffraction



High
Pressure

Polymorphs



Included structures must have an $R < 10\%$, no disorder and 3-D coordinates present (amongst other criteria).

van de Streek, J. (2006). *Acta Cryst. B62*, 567-579. DOI: [10.1107/S0108768106019677](https://doi.org/10.1107/S0108768106019677)

Using the subsets

1

1. You can Load a subset

2

2. You can restrict a search to a subset

3

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.

3. Bonus tip – they are also available in our API!

```
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
        getatt(self, format + '_header')()
        for e in self.rdr:
            getatt(self, format + '_line')(e)
            getatt(self, format + '_footer')()

    def csv_header(self):
        data = ', '.join([
            'Identifier',
            'Author',
            'Literature Ref',
            'Formula',
            'Compound Name',
            'Synonym'
        ])
        print(data)
```



Making Drawing Easier: Other Elements

You can select other elements – or even multiple elements!

Not Hydrogen
C or H
Any Non-metal
Any Metal
Any Transition Metal
Any Halogen

Other Elements...

On the periodic table
Single Element: click on the element you need.

Pre-defined selection (e.g. groups, periods, metals, ...): click on the corresponding orange tile.

Personalised selection: tick Multi Pick, then select the elements you need.

single element

pre-defined selections

personalised selection

Other Atom Type

Current Selection: Cu

This screenshot shows the 'Other Atom Type' dialog with a periodic table. The element Copper (Cu) is highlighted in green. The 'Picking Mode' is set to 'Single Pick'.

Other Atom Type

Current Selection: 1M

This screenshot shows the 'Other Atom Type' dialog with a periodic table. The '1M' group tile is highlighted in orange. The 'Picking Mode' is set to 'Single Pick'.

Other Atom Type

Current Selection: 2B

This screenshot shows the 'Other Atom Type' dialog with a periodic table. The '2B' group tile is highlighted in orange. The 'Picking Mode' is set to 'Single Pick'.

Other Atom Type

Current Selection: C N

Picking Mode
 Single Pick
 Multi Pick

Reset Cancel Apply OK

This screenshot shows the 'Other Atom Type' dialog with a periodic table. The 'C' and 'N' tiles are highlighted in green. The 'Picking Mode' is set to 'Multi Pick'.

How can I be more specific in my draw searches?

The image displays a ChemDraw interface with several windows. On the left, there are search result windows for 'VIPZAV' and 'ABALIZ'. The 'VIPZAV' window shows a list of 20 compounds, with 'VIPZAV' highlighted. The 'ABALIZ' window shows a list of 20 compounds, with 'ABALIZ' highlighted. In the center, there is a 'Draw (1) - New' window showing a pyrene structure (a four-ring fused benzene system) with dashed lines indicating the drawing process. The window title is 'Draw (1) - New' and it has a menu bar with 'File', 'Edit', 'Atoms', 'Bonds', '3D', 'Options', and 'Help'. Below the menu bar are buttons for 'DRAW', 'EDIT', 'ERASE', 'ADD 3D', and 'CONTACT'. A yellow tooltip says 'Click and drag to create a bond. Drag to an existing atom to make a connection.' On the right side of the window, there are '3D Parameters' and 'Contacts' sections. At the bottom of the window, there is a 'Search' bar with the text 'C H O N S P F Cl Any More...' and a 'Bond' dropdown menu set to 'Single'. A text box at the bottom of the image asks: 'How can I improve my search to get only pyrene structures?'

Refcode: VIPZAV CSD version 5.43 (November 2021)

Refcode: ABALIZ CSD version 5.43 (November 2021)

Refcode: ABUREW CSD version 5.43 (November 2021)

Example: Pyrene

How can I improve my search to get only pyrene structures?

Draw (1) - New

File Edit Atoms Bonds

Bonds 3D Options Help

Type

Cyclicity

Variable Attachment...

Exclude...

Click and drag
Drag to an element

DRAW

EDIT

ERASE

ADD 3D

CONTACT

13 C 11 C 12 C 15 C 16 C

1 C 10 C 14 C

2 C 6 C 8 C

3 C 5 C 9 C

4 C 7 C

Templates...

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

Cancel

Exclude Bonds

Ring Closure Direct Links

Disallow additional cyclic bonds from selected atoms

Note: this will not apply to direct links to other atoms in the query.

Current Atom Selection:

Select All Deselect All

Cancel OK

Bonds > Exclude...

- Ring Closure:
Exclude molecules where further rings are bonded to specified atoms from search results.
- Direct Links:
Select atoms that cannot be bonded in the search results.

No additional rings

Can click on one atom at a time or use **Select All**

Bonds > Exclude...

- Ring Closure:
Exclude molecules where further rings are bonded to specified atoms from search results.
- Direct Links:
Select atoms that cannot be bonded in the search results.

Draw (1) - New
File Edit Atoms Bonds 3D Options Help

Click and drag to create a bond.
Drag to an existing atom to make a connection.

DRAW
EDIT
ERASE
ADD 3D
CONTACT

Options...
Delete

Contacts:

Options...
Delete

Search
Store
Cancel

RingMaker

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

Set number of hydrogen atoms

Can set number of hydrogens bonded to an atom.

Draw (1) - New

File Edit Atoms Bonds 3D Options Help

Element

Add Group

Expand Chemical Groups

DRAW

EDIT

ERASE

ADD 3D

CONTACT

Hydrogens

Charge

Number of Bonded Atoms

Cyclicity

All Atoms in Same Molecule

Generate

Clear

0

1

2

3

4

Unspecified

Other...

Options...

Delete

Contacts:

Options...

Delete

Search

Store

Cancel

RingMaker

Templates...

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

Set number of hydrogen atoms

To assign H=1 to multiple atoms, go to the 'Atoms' menu, and select 1 from the 'Hydrogens' option.

Draw (1) - New

File Edit Atoms Bonds 3D Options Help

Click and drag to create a bond.
Drag to an existing atom to make a connection.

DRAW
EDIT
ERASE
ADD 3D
CONTACT

Options...
Delete

Contacts:

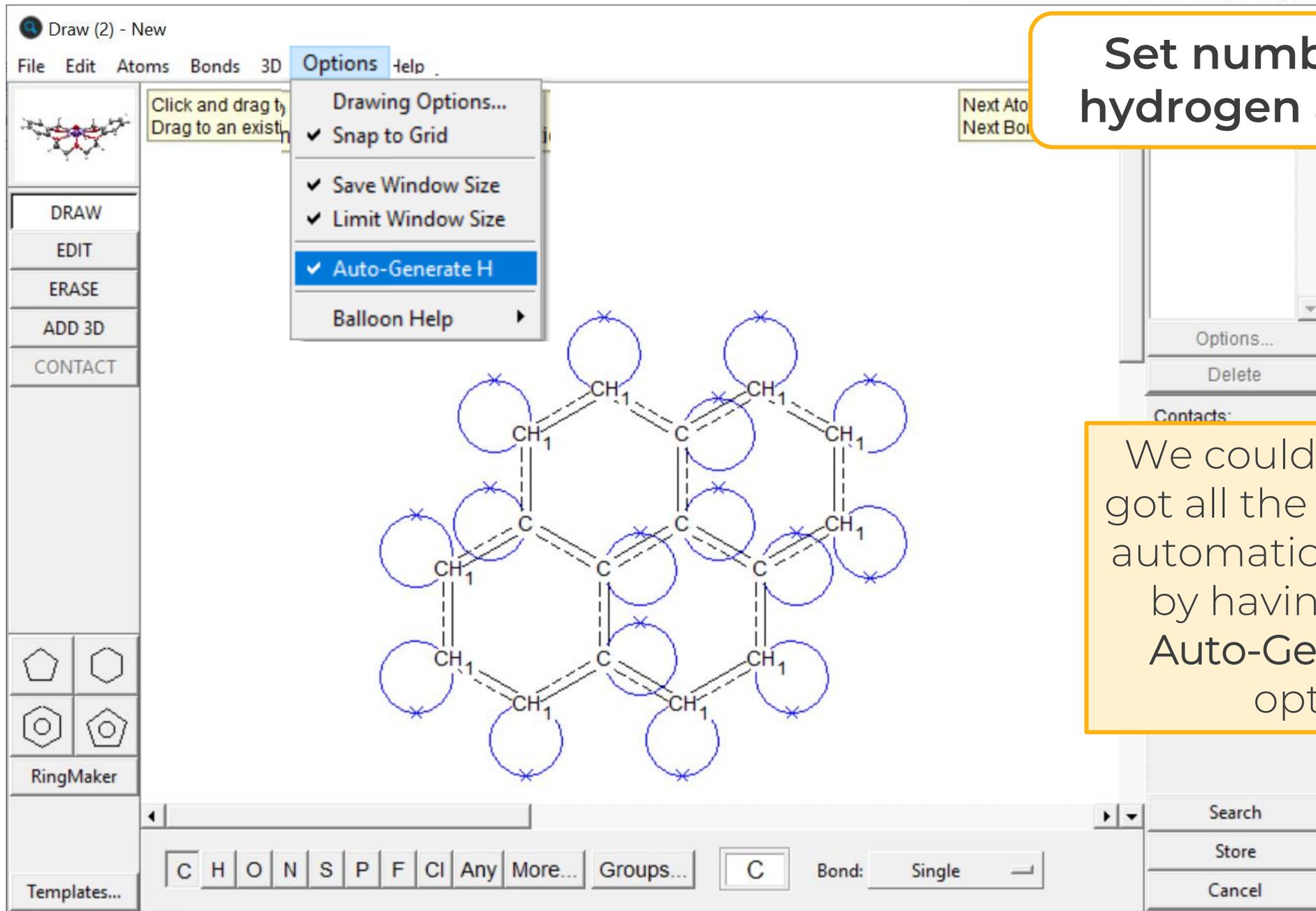
Options...
Delete

Search
Store
Cancel

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

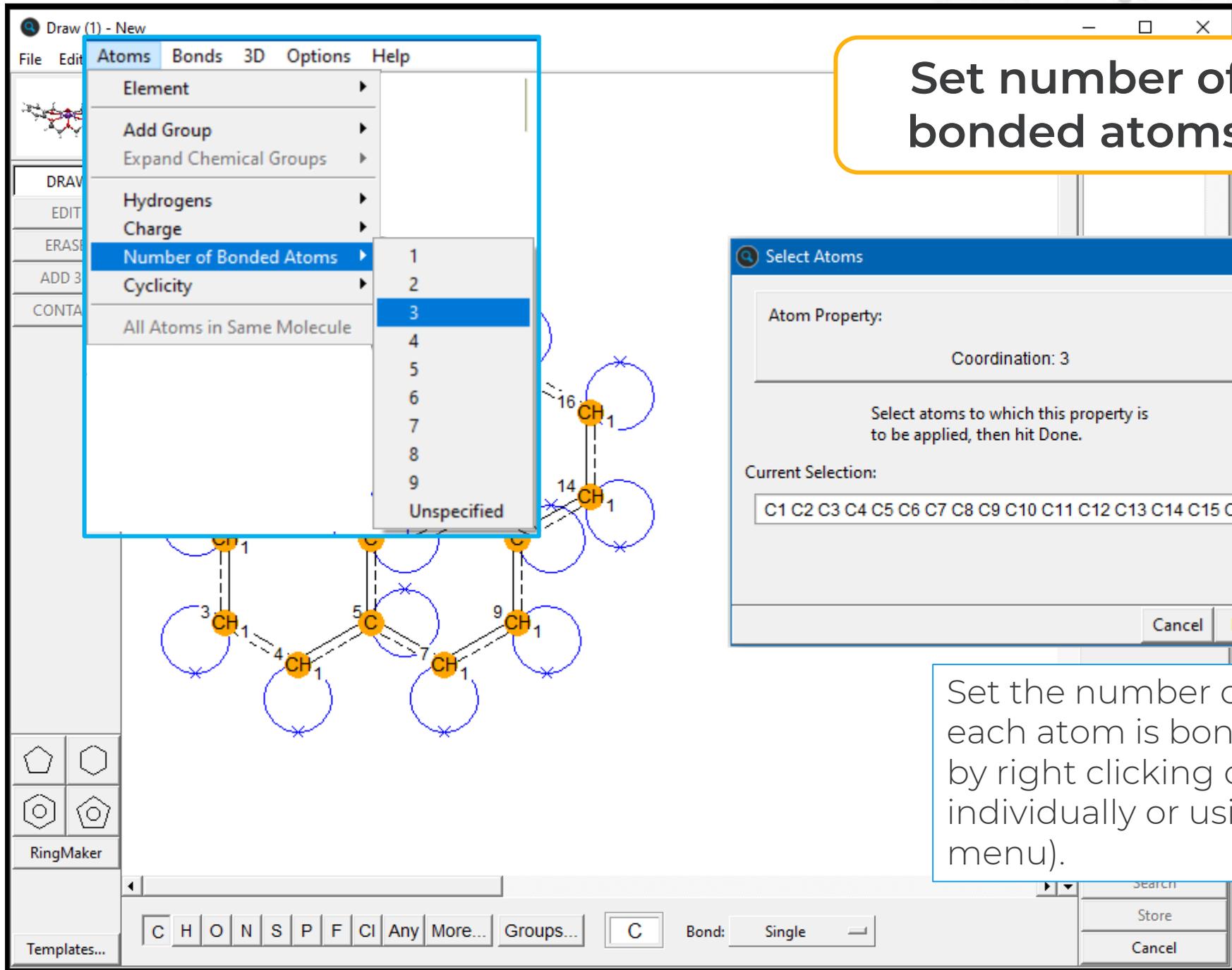
The screenshot shows a chemical drawing software window titled "Draw (1) - New". The interface includes a menu bar (File, Edit, Atoms, Bonds, 3D, Options, Help), a toolbar with buttons for DRAW, EDIT, ERASE, ADD 3D, and CONTACT, and a "RingMaker" section with icons for pentagons and hexagons. The main workspace displays a skeletal structure of a bicyclic compound, specifically a decalin derivative, with several carbon atoms labeled "C" and hydrogen atoms labeled "CH1". A yellow callout box with a black border and orange outline is positioned in the upper right, containing the text "Set number of hydrogen atoms". The right sidebar contains "Options..." and "Delete" buttons, and a "Contacts:" section. The bottom status bar shows a template bar with elements C, H, O, N, S, P, F, Cl, and "Any More...", a "Groups..." button, a "C" atom button, and a "Bond: Single" dropdown menu.

Set number of hydrogen atoms



Set number of
hydrogen atoms

We could also have
got all the hydrogens
automatically added
by having on the
Auto-Generate H
option



Set number of bonded atoms

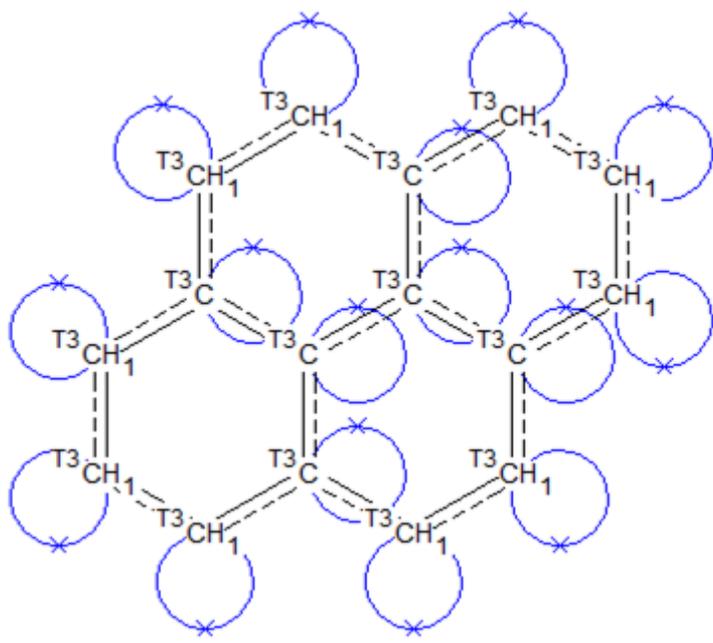
Set the number of other atoms each atom is bonded to (either by right clicking on atoms individually or using the atoms menu).

Draw (1) - New

File Edit Atoms Bonds 3D Options Help

Click and drag to create a bond.
Drag to an existing atom to make a connection.

DRAW
EDIT
ERASE
ADD 3D
CONTACT



Options...
Delete

Contacts:

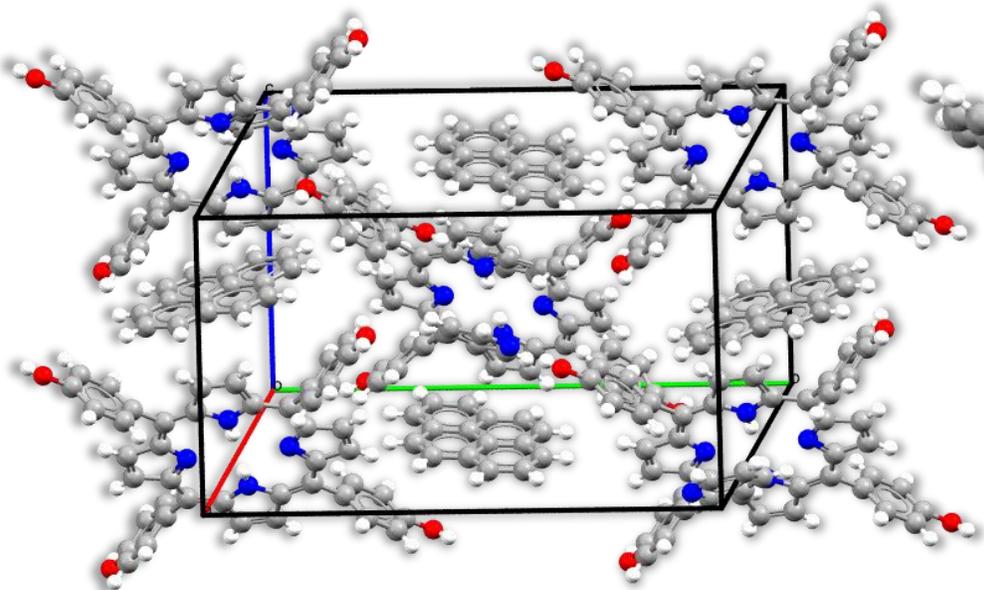
Options...
Delete

Search
Store
Cancel

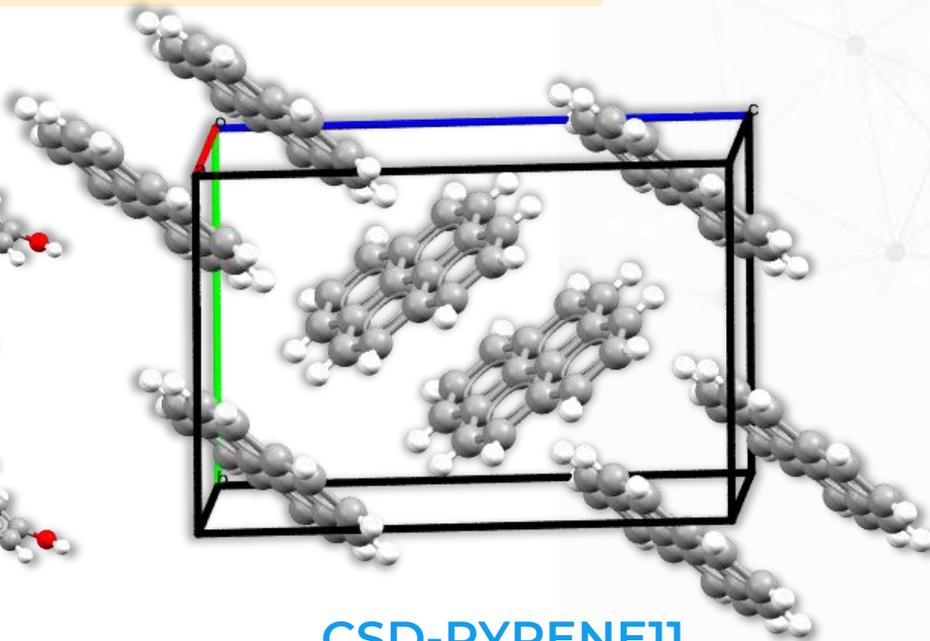
Templates... C H O N S P F Cl Any More... Groups... C Bond: Single

Set number of bonded atoms

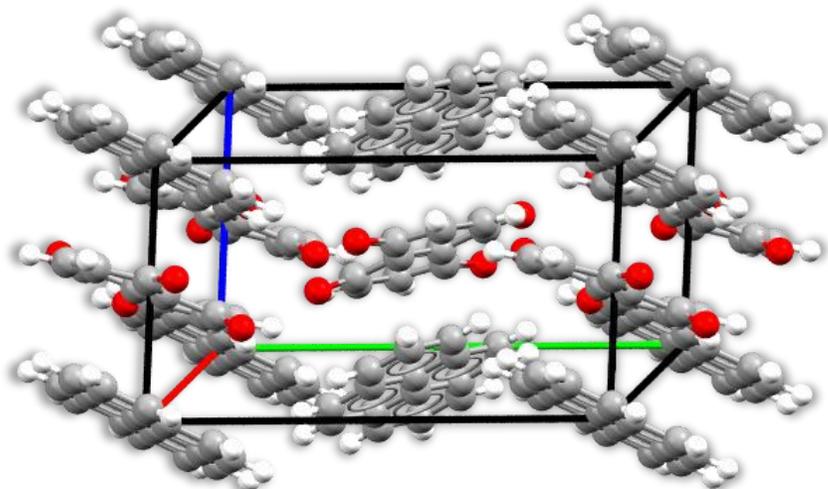
We have found pyrene and multicomponent substances containing pyrene.



CSD-ELUGUP



CSD-PYRENE11



CSD-CEKBUP



Tip

Do you need only pyrene structures?
Use
Combine Queries
to combine this
query with a query
for
Number of
Residues = 1
to just find pyrene.

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

Drag Query Icons into Boxes

Find entries that:

must have (boolean AND)

Query 3 Query 6

must not have (NOT)

Query 3 Number of Residues = 1 Edit... Delete

Query 6 C1=CC=C2C=CC(=C1)C=C3C=CC(=C2)C=C3 Edit... Delete

2D

What if I only want a certain number of molecules in the crystal structure?

Build Queries **Combine**

Draw

Peptide

Author/Journal

Name/Class

Elements

Formula

Space Group

Unit Cell

Z/Density

Experimental

All Text

Recode (entry ID)

Z/Density (1) - New

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

Search Store **=** not = > < <= >= range Reset

You can use the **Z/Density** search to set the number of molecules found in the structure.

Then use the **Combine Queries** tab to create a search that includes both the Pyrene molecule and the Z/Density restriction.

Combine Queries

Build Queries Combine Queries Manage Hitlists View Results

Allows you to combine different queries using boolean AND, NOT, OR

Query 1: Pyrene query

Query 2: Number of residues = 2

Must have: Query 1 AND Query 2

The screenshot shows the CCDC ConQuest (1) software interface. The window title is "CCDC ConQuest (1)". The menu bar includes "File", "Edit", "Options", "View Databases", "Results", and "Help". The main interface has four tabs: "Build Queries", "Combine Queries" (which is active and highlighted in red), "Manage Hitlists", and "View Results".

Under the "Combine Queries" tab, there is a section titled "Find entries that:" with three options:

- must have (boolean AND):** This section contains two query icons labeled "Query 1" and "Query 2".
- must not have (NOT):** This section is currently empty.
- must have at least one of (OR):** This section is currently empty.

At the bottom of the "Find entries that:" section are "Search" and "Reset" buttons.

The main workspace on the right shows two query boxes:

- Query 1:** Contains a chemical structure of pyrene, a polycyclic aromatic hydrocarbon. A small box labeled "2D" is visible near the structure. To the right of the structure are "Edit..." and "Delete" buttons.
- Query 2:** Contains the text "Number of Residues = 2". To the right of the text are "Edit..." and "Delete" buttons.

Explore More: Tips & Tricks

- Using and creating templates
- Finding structures with different functional group positions
- Customising your searching
- Searches for structures on the web
- Finding similar structures of interest

Making Drawing Easier: Templates

Click and drag to create a bond.
Drag to an existing atom to make a connection.

File Edit Atoms Bonds 3D Options Help

DRAW
EDIT
ERASE
ADD 3D
CONTACT

RingMaker

Templates...

View...

List

Custom

Groups... C Bond: Single

View Templates

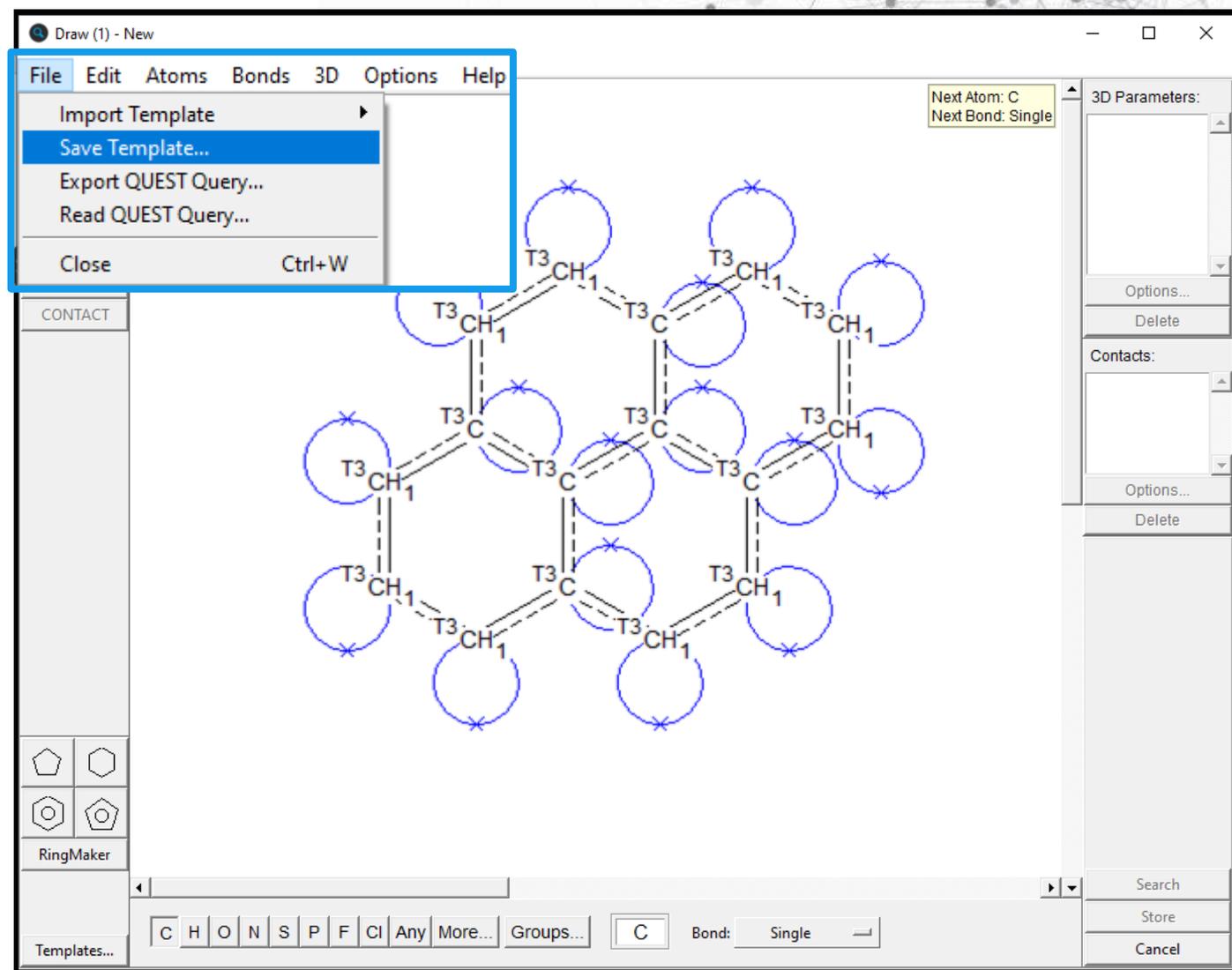
12-crown-4

Name	Category
12-crown-4	Cyclic-ethers
15-crown-5	Cyclic-ethers
18-crown-6	Cyclic-ethers
Acetylacetonato	O-ligands
Adamantane	C-ring-systems
Adenine	Nucleic-acid-bases
Alpha-cyclodextrin	Carbohydrates
Beta-cyclodextrin	Carbohydrates
Bicyclo-octane	C-ring-systems
Bipyridyl	N-ligands
Bis-bis-ethylenedithio-tetrathiafulvalene	Charge-transfers
Bis-diphenylphosphinoethane	P-ligands
Bis-diphenylphosphinomethane	P-ligands
Bis-diphenylphosphinopropane	P-ligands
Bis-salicylidenediminato	N,O-ligands
Calix(4)arene	Calixarenes
Calix(6)arene	Calixarenes

Cancel Load Store

Create your own Templates

1. Draw molecule
2. Save the Template
 - To access through the ConQuest custom template list, save the file in the [csds_data\templates](#) folder in your User Area.



Create your own Templates

1. Draw molecule
2. Save the Template
 - To access through the ConQuest custom template list, save the file in the [csds_data\templates](#) folder in your User Area.
3. Template can now be accessed through **Templates** button in either 'List' or 'View'

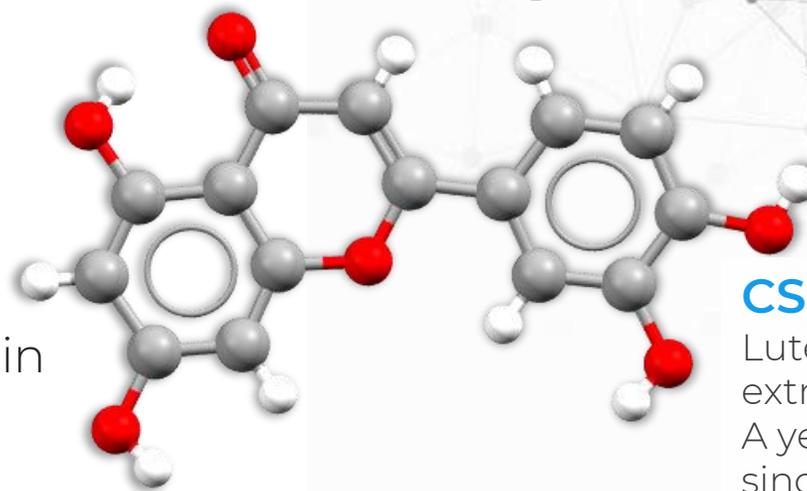
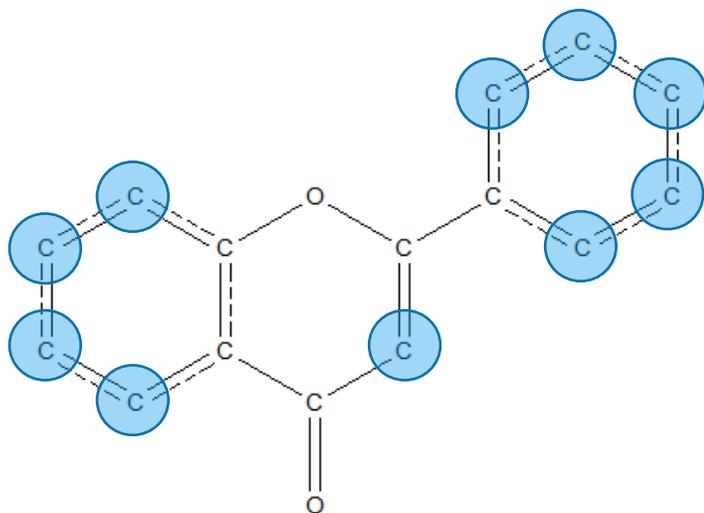
The screenshot shows the ConQuest software interface. The main window is titled 'Draw (1) - New' and contains a drawing tool with a menu (File, Edit, Atoms, Bonds, 3D, Options, Help) and a toolbar with buttons for DRAW, EDIT, ERASE, ADD 3D, and CONTACT. A tooltip indicates 'Click and drag to create a bond. Drag to an existing atom to make a connection.' The drawing area shows a pyrene molecule structure. The 'View Templates' dialog box is open, displaying a list of templates with columns for Name and Category. The 'Pyrene_template' is highlighted in red. The 'List' button is selected in the bottom left corner.

Name	Category
Cytidine	Nucleic-acid-bas
Diethylenetriamine	Ethylenediamine
Dimethylglyoxime	N-ligands
Dimethylpyrazolylborate	N-ligands
Dodecaborane	Boron-cages
Edta	Ethylenediamine
Fluorene	C-ring-systems
Furanose	Carbohydrates
Guanine	Nucleic-acid-bas
Hexaborane	Boron-cages
Nitrilotriacetato	N,O-ligands
Norbormane	C-ring-systems
Pentamethylcyclopentadienyl	C-ring-systems
Phenanthroline	N-ligands
Porphyrin	Porphyrins
Pyranose	Carbohydrates
Pyrene_template	Custom
Salicylideniminato	N,O-ligands
Squarate	O-ligands

Do you want to search for structures that can have functional groups in different places?

Example: Flavones

- A class of flavonoids, having a 2-phenylchromen-4-one backbone with functional groups (e.g. OH, OMe).
- Many naturally occurring flavones found in plants and fungi (with **hydroxy** and/or **methoxy** substituents) – but synthetic flavones have also been created.

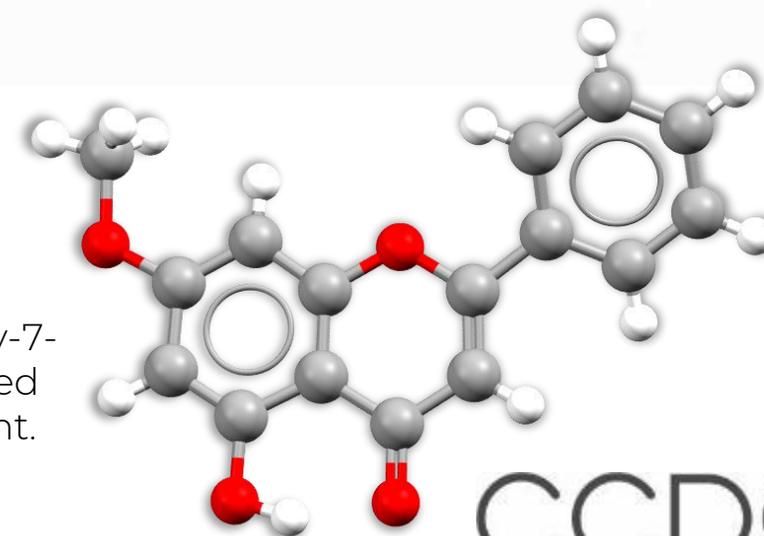


CSD-OJEQUP

Luteolin (tetra-hydroxy flavone), extracted from *Reseda luteola*. A yellow dye that has been used since 1st millennium BC.

CSD-BAZGOB

Teichochrysin (5-hydroxy-7-methoxyflavone) extracted from the sour cherry plant.



CCDC

Draw (1) - New

File Edit Atoms Bonds 3D Options Help

Click
Draw

connection.

Next Atom: O
Next Bond: Double

3D Parameters:

Options...
Delete

Contacts:

Options...

Variable Points of Attachment

Select Atom to Variably Attach O18

Select Points of Attachment

C1 C2 C3 C4 C9 C15 C16 C14 C12 C

Bond type: Single

Reset Cancel OK

Variable Attachments

To set a variable attachment:

- First select the atom that can be variably attached.
- Then select the places it could be attached to.

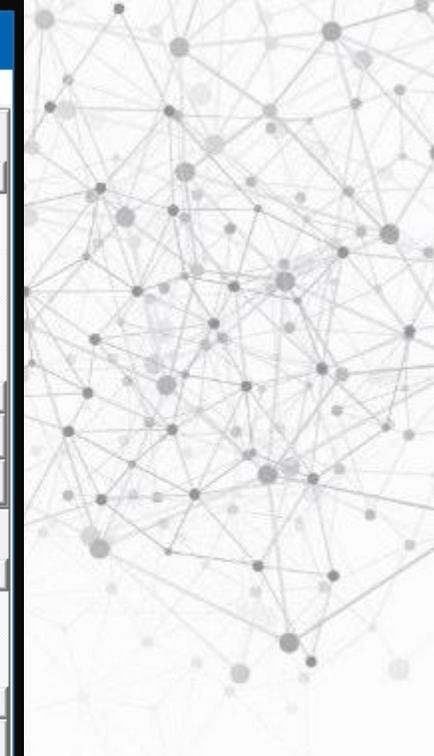
Variable attachments shown by \longrightarrow

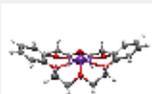
Search
Store
Cancel

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

Templates... RingMaker

DRAW
EDIT
ERASE
ADD 3D
CONTACT

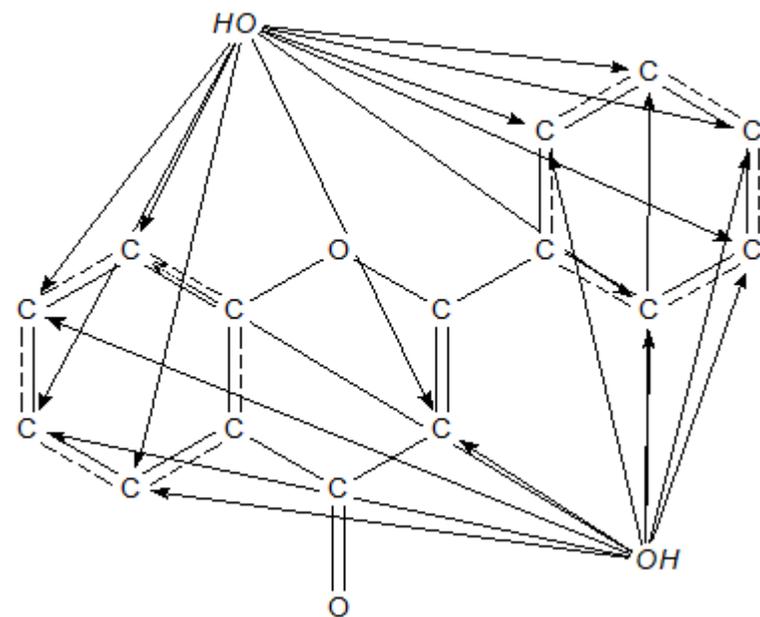




Click and drag to create a bond.
Drag to an existing atom to make a connection.

Next Atom: O
Next Bond: Double

- DRAW
- EDIT
- ERASE
- ADD 3D
- CONTACT



3D Parameters:

Options...

Delete

Contacts:

Options...

Delete



RingMaker

Templates...

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

Search

Store

Cancel



- How do I **only** find di-hydroxy flavones?

Combine Queries

Query 1: 2 variably attached OH
Query 2: 3 variably attached OH

Must have: Query 1

Must not have: Query 2

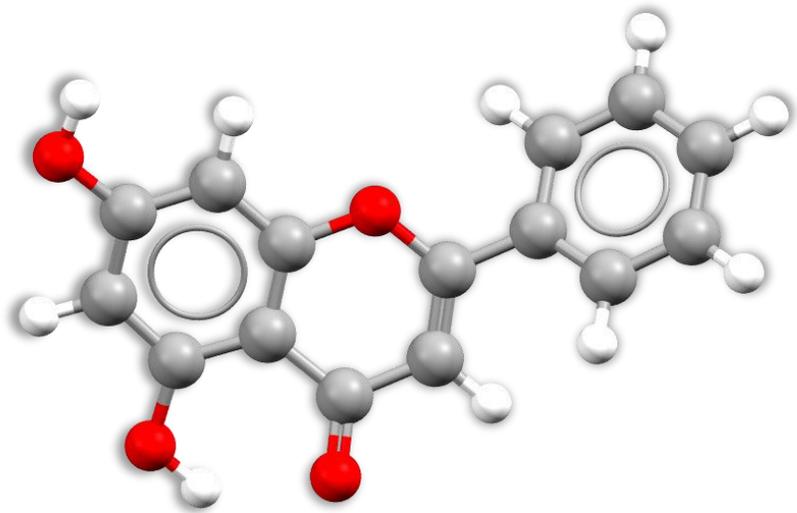
The screenshot shows the CCDC ConQuest (1) software interface. The window title is "CCDC ConQuest (1)". The menu bar includes "File", "Edit", "Options", "View Databases", "Results", and "Help". The main interface has four tabs: "Build Queries", "Combine Queries" (which is active), "Manage Hitlists", and "View Results".

Under the "Combine Queries" tab, there is a section titled "Find entries that:" with three options:

- must have (boolean AND)**: A box labeled "Query 1" with a green question mark icon.
- must not have (NOT)**: A box labeled "Query 2" with a green question mark icon.
- must have at least one of (OR)**: An empty box.

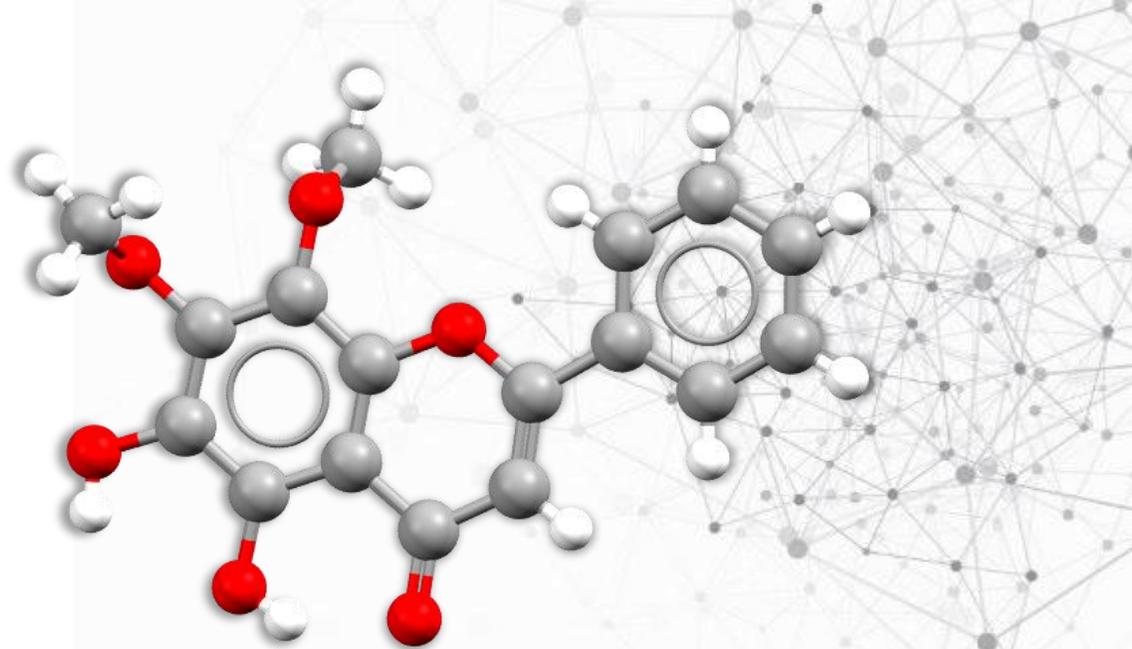
At the bottom of this section are "Search" and "Reset" buttons.

The right side of the interface shows two panels, each containing a chemical structure diagram of a flavone skeleton with arrows indicating attachment points for hydroxyl groups. The top panel is labeled "Query 1" and the bottom panel is labeled "Query 2". Each panel has "Edit..." and "Delete" buttons.



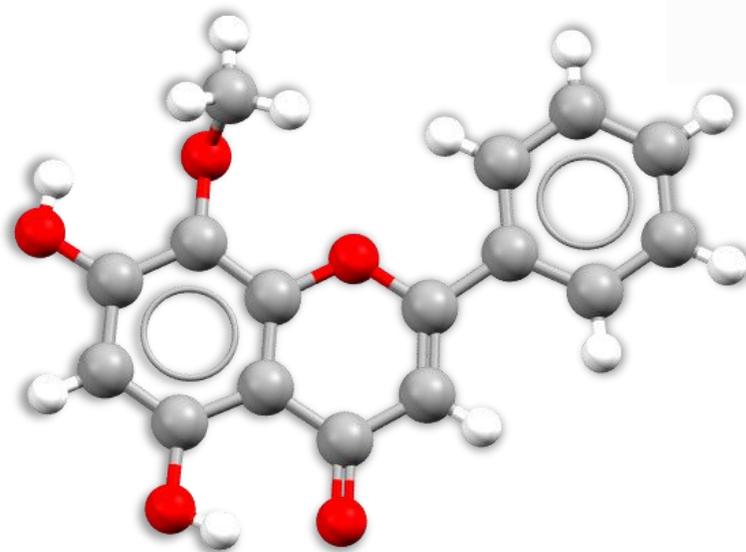
CSD-IYIWEU

Chrysin, 5-7-dihydroxyflavone, is found in honey, propolis and passionflowers.



CSD-AFOCUV

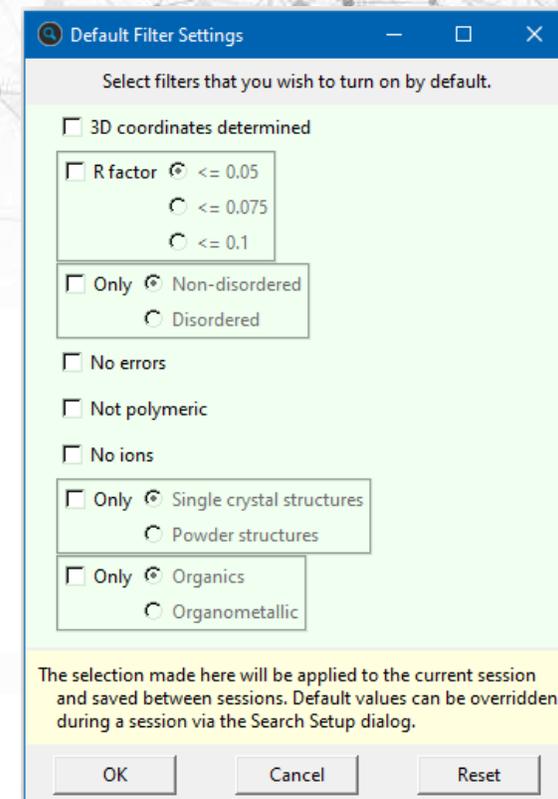
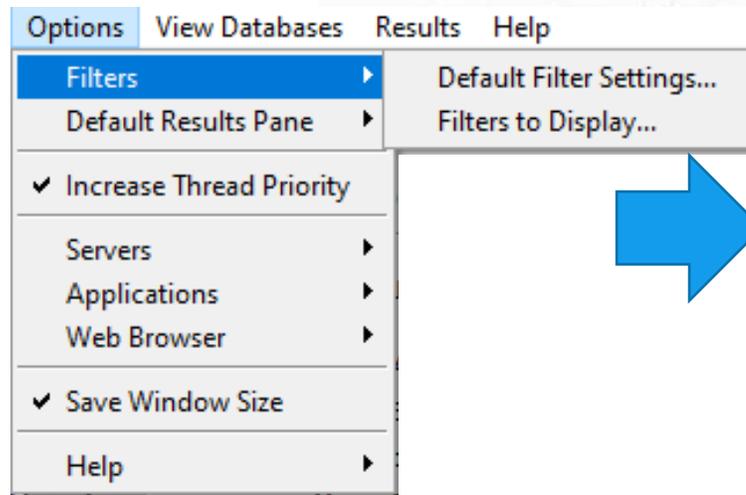
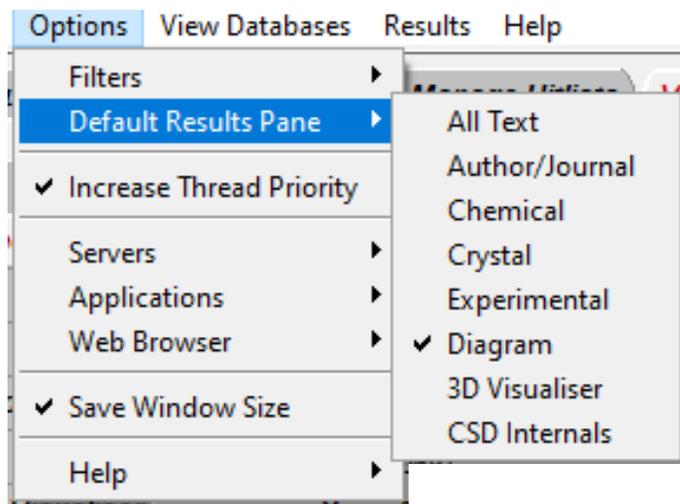
5,6-Dihydroxy-7,8-dimethoxyflavone, Isolated from a rare traditional Chinese medicinal herb, *Saussurea involucrate*



CSD-PUZLII

Wogonin, 5,7-Dihydroxy-8-methoxyflavone, extracted from *Scutellaria baicalensis*. It has been investigated for many different biological activities.

Customise your searching



Have a results panel you prefer to see first in ConQuest?

Select it in:

Options>Default Results Pane

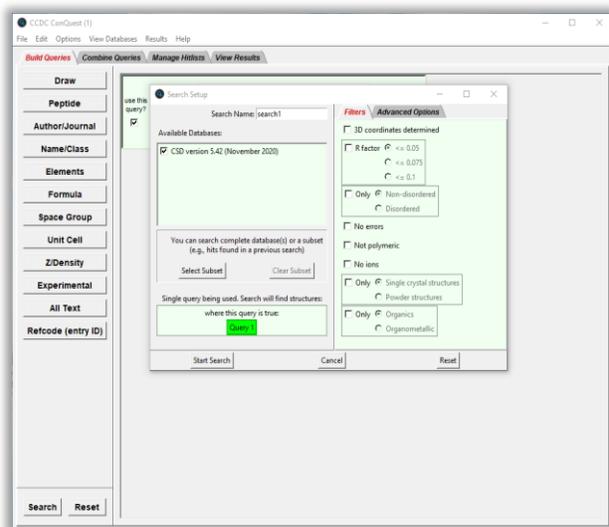
Always using the same search filter settings?

Change your default settings in: Options>

Filters> Default Filter Settings...

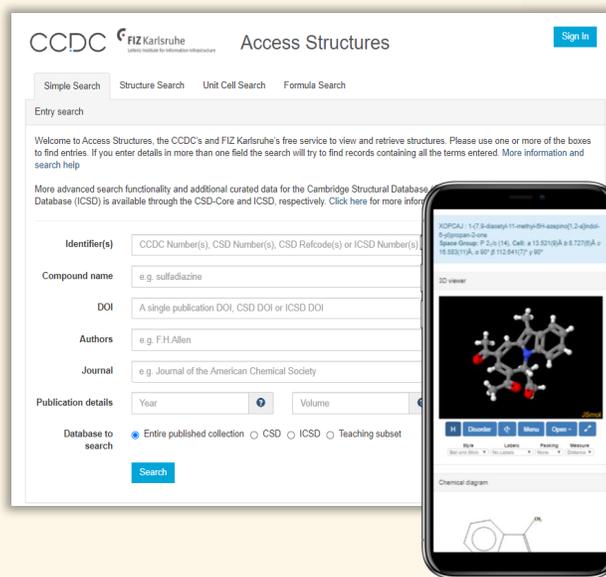
Finding fantastic structures

Desktop



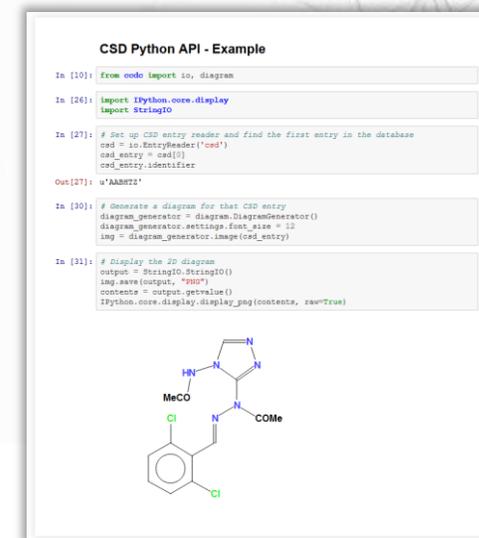
ConQuest

Web Browser



WebCSD /
Access Structures

Programmatic



CSD Python API

Access Structures

- Online portal to access the CSD and ICSD
- Free to view and retrieve individual structures
- Basic searching
- Accessed through any standard internet browser – no local installation of software required

CCDC FIZ Karlsruhe Access Structures

Simple Search Structure Search Unit Cell Search Formula Search

Entry search

Welcome to Access Structures, the CCDC's and FIZ Karlsruhe's free service to view and retrieve structures. Please use one or more of the search options to find entries. If you enter details in more than one field the search will try to find records containing all the terms entered. More information is available in the search help.

More advanced search functionality and additional curated data for the Cambridge Structural Database (CSD) and the Inorganic Crystal Structure Database (ICSD) is available through the CSD-Core and ICSD, respectively. Click here for more information.

Identifier(s) CCDC Number(s), CSD Number(s), CSD Refcode(s) or ICSD Number(s)

Compound name e.g. sulfadiazine

DOI A single publication DOI, CSD Refcode or ICSD Number

Authors e.g. F.H.Allen

Journal e.g. Journal of the American Chemical Society

Publication details Year

Database to Entire published collection

Download deposited CIF

- Deposited CIF(s)
- Deposited CIF(s) without structure factor data
- Deposited file(s) with any available structure factor data and checkCIF reports included
- Include checkCIF reports as a PDF when available

Additional details

Deposition Number 200783

Date Citation I.D.H. Oswald, W.D.S. Motherwell, S. Parsons, C.R. Pulham CCDC 200783: Experimental Crystal Structure Determination, 2003, DOI: 10.5517/ccdcgroup

Synonyms **Propacetamol** morpholine, Acetaminophen morpholine

Deposited on 16/12/2002

Associated publications

I.D.H. Oswald, W.D.S. Motherwell, S. Parsons, C.R. Pulham, Acta Crystallographica Section E: Structure Reports Online, 2002, 58, 1290, DOI: 10.1107/S1605360202018111

WebCSD

Simple Search

Structure Search

Unit Cell Search

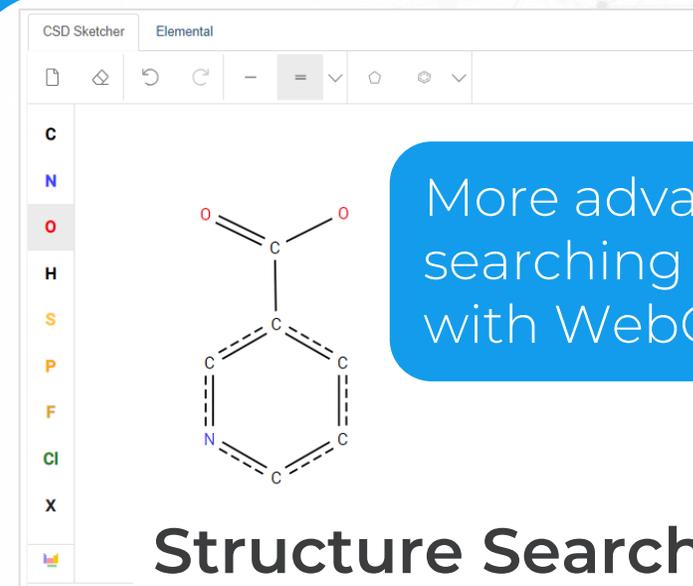
Formula Search

Access Structures and WebCSD

Search form for Simple Search:

- Identifier(s): CCDC Number(s), CSD Number(s), CSD Refcode(s) or ICSD Number(s)
- Compound name: e.g. sulfadiazine
- DOI: A single publication DOI, CSD DOI or ICSD DOI
- Authors: e.g. F.H.Allen
- Journal: e.g. Journal of the American Chemical Society
- Publication details: Year, Volume, Page
- Database to search: Entire published collection CSD ICSD Teaching subset
- + Add New Search Field
- Search

Simple Search



Unit Cell Search

Search form for Unit Cell Search:

- Lattice centring: Primitive (P)
- a: e.g. 10.0
- b: e.g. 10.0
- c: e.g. 10.0
- α : e.g. 90.0
- β : e.g. 90.0
- γ : e.g. 120.0

Formula Search

e.g. C8 H9 N1 O2

Structure Search

Simple Search Structure Search Unit Cell Search Formula Search

Chemical structure searching

Please draw your diagram or add a SMARTS string using the advanced section below.
You can help us improve the current version of the CSD Sketcher by telling us what you think.

[Leave your feedback](#)

The screenshot shows the CSD Sketcher interface. At the top, there are four tabs: 'Simple Search', 'Structure Search', 'Unit Cell Search', and 'Formula Search'. Below the tabs is a search bar containing the text 'Chemical structure searching'. A message prompts the user to draw a diagram or add a SMARTS string. A 'Leave your feedback' button is present. The main drawing area shows a toolbar with various icons and a search query '15-crown-5'. On the left, a vertical list of elements (C, N, O, H, S, P, F, Cl, X) is visible. On the right, there are zoom and pan controls. At the bottom, a 'Match condition' section is highlighted with a blue border, showing three radio buttons: 'Exact', 'Substructure', and 'Similarity', with 'Similarity' selected.

Version: 0.1.1318

Match condition: Exact Substructure Similarity

Exact

Returns structures that contain the exact molecule as it is drawn.

Substructure

Returns hits where the draw query is a part of any molecule (similar to conquest).

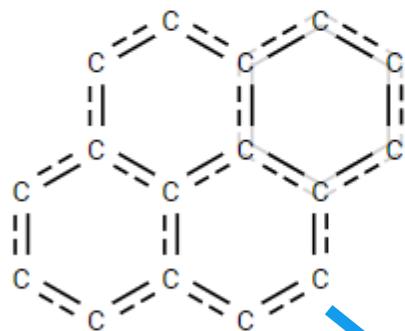
Similarity

Calculates a [molecular fingerprints using Tanimoto](#) for the drawn molecule and compares that to pre-calculated fingerprints for structures in the CSD.

Required: Full molecule drawn (like exact search)

Returns: Similar structures and a similarity score (where 1 is an identical molecule).

>1,000 hits
Includes structures
where yrene structure is
part of a larger molecule



Substructure

Exact

Simple Search Structure Search Unit Cell Search Formula Search

Search Complete **1000 Results Found**

100%

Modify Search New Search

Results

Database Identifier	Deposition Number
<input checked="" type="checkbox"/> ABALIZ	244177
<input checked="" type="checkbox"/> ABIBEW	2072185
<input checked="" type="checkbox"/> ABIBIA	2072186
<input checked="" type="checkbox"/> ABIBOG	2072187
<input checked="" type="checkbox"/> ABICAT	2072188
<input checked="" type="checkbox"/> ABICEX	2072189
<input checked="" type="checkbox"/> ABIKOP	2080990
<input checked="" type="checkbox"/> ABINIM	2091443
<input checked="" type="checkbox"/> ABUHEN	1434681
<input checked="" type="checkbox"/> ABUJIT	1448617
<input checked="" type="checkbox"/> ABUJIU	2076486
<input checked="" type="checkbox"/> ABUREW	830174
<input checked="" type="checkbox"/> ABURIA	830175

ABALIZ : 2,5,8,10,15,17-hexakis(*t*-Butyl)-(bc,ef,hi,uv)tetrabenzoalene hexane solvate
Space Group: P c a b (61), Cell: a 17.2333(6)Å b 24.8954(8)Å c 29.6704(9)Å, α 90° β 90° γ 90°

3D viewer

Ball and Stick No Labels

Chemical diagram

View group symbols key

Simple Search Structure Search Unit Cell Search Formula Search

Search Complete **327 Results Found**

100%

Modify Search New Search

Results

Database Identifier	Deposition Number
<input checked="" type="checkbox"/> ABUJIU	2076486
<input checked="" type="checkbox"/> ABUTEY	823544
<input checked="" type="checkbox"/> AGORAS	1847617
<input checked="" type="checkbox"/> AGORAS01	1847613
<input checked="" type="checkbox"/> AGOREW	1847618
<input checked="" type="checkbox"/> AGOREW01	1847616
<input checked="" type="checkbox"/> AGOREW02	1847615
<input checked="" type="checkbox"/> AKACUM	1451536
<input checked="" type="checkbox"/> AKACUM01	1451537
<input checked="" type="checkbox"/> AKECEA	1449783
<input checked="" type="checkbox"/> AKIVEX	998244
<input checked="" type="checkbox"/> ASULOS	1893487
<input checked="" type="checkbox"/> ATAMAM	2062107

ASULOS : (2-[[3-(pyridin-2-yl)-1H-pyrazol-1-yl]methyl]-3-[[3-(pyridin-2-yl)-1H-pyrazol-1-yl]methyl]quinoxaline)-silver perchlorate hemikis(pyrene)

Space Group: P 2₁/c (14), Cell: a 10.771(2)Å b 19.080(4)Å c 15.000(3)Å, α 90° β 93.45(3)° γ 90°

3D viewer

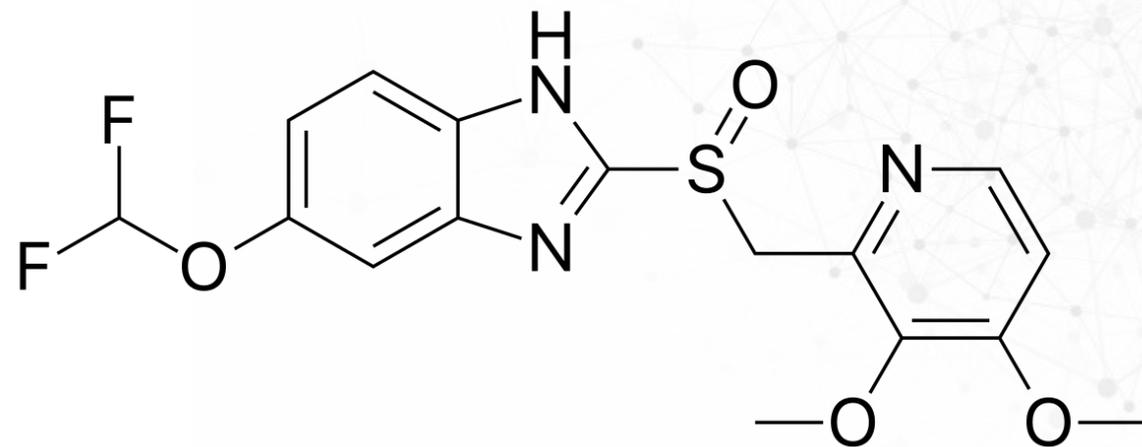
Ball and Stick No Labels

Chemical diagram

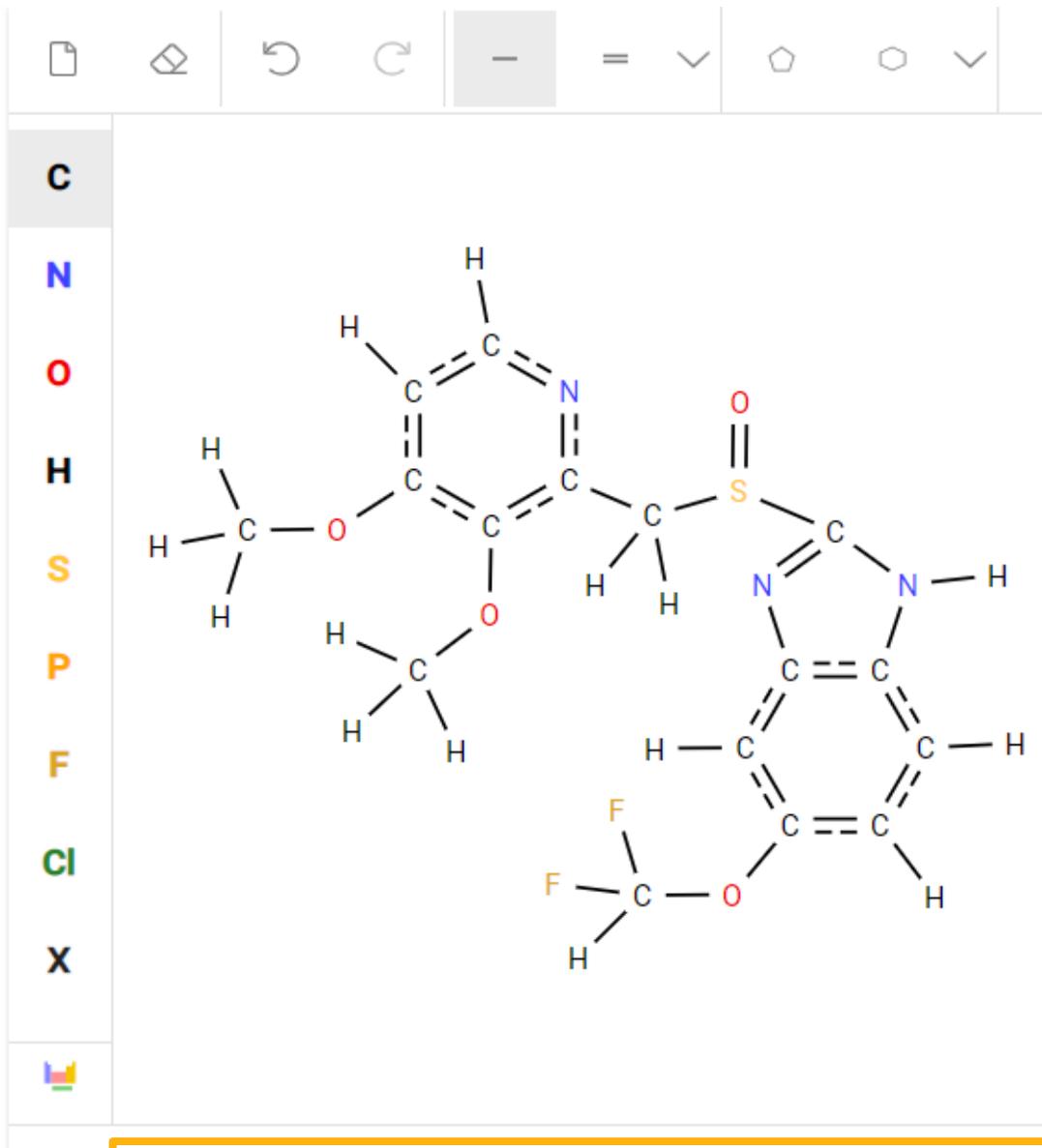
View group symbols key

327 hits
Only structures with
pyrene on its own

Similarity



Example: Pantoprazole, a medication to treat stomach ulcers
Not currently in the CSD



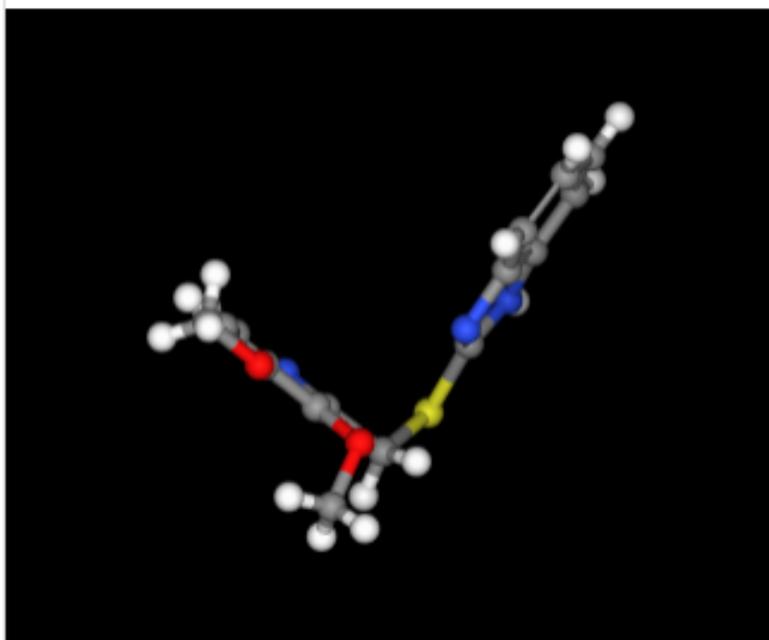
Match condition: Exact Substructure Similarity

WETBIJ : 2-[[[(3,4-dimethoxyphenyl)methyl]sulfonyl]-1H-benzimidazole

Space Group: $P 2_1 2_1 2_1$ (19), Cell: $a 9.1828(16)\text{\AA}$ $b 11.625(2)\text{\AA}$ $c 13.463(2)\text{\AA}$, $\alpha 90^\circ$ $\beta 90^\circ$ $\gamma 90^\circ$

3D viewer

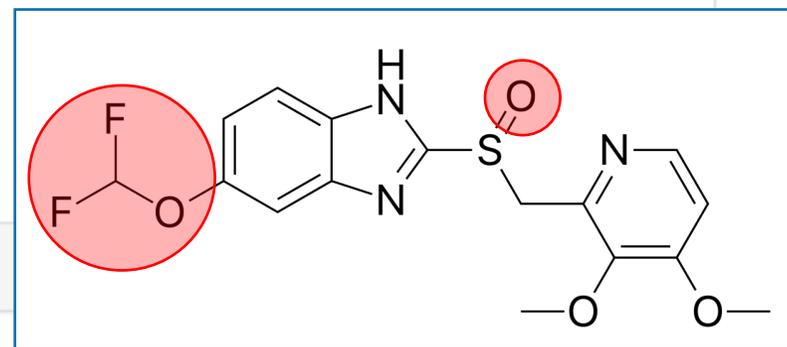
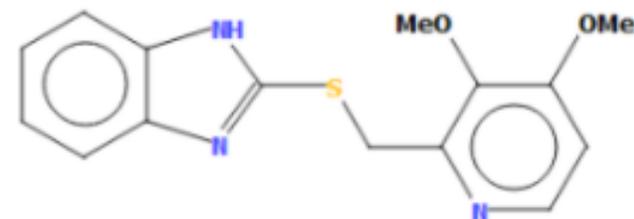
Ball and Stick No Labels    



 No Packing **H**  

DISORDER 

Chemical diagram



Results

<input checked="" type="checkbox"/>	Database Identifier	Deposition Number	Similarity Score More Info
<input checked="" type="checkbox"/>	BOMQOK	1113696	0.998
<input checked="" type="checkbox"/>	EYEJUP	1511140	0.775
<input checked="" type="checkbox"/>	WETBIJ	1563921	0.763
<input checked="" type="checkbox"/>	CENBIJ	1556811	0.745
<input checked="" type="checkbox"/>	FOPYOZ	692400	0.745
<input checked="" type="checkbox"/>	UDAVIF	633383	0.745
<input checked="" type="checkbox"/>	UDAVIF01	633384	0.745
<input checked="" type="checkbox"/>	UDAVIF02	633385	0.745
<input checked="" type="checkbox"/>	UDAVIF03	633386	0.745
<input checked="" type="checkbox"/>	VAYXOI	1280849	0.745
<input checked="" type="checkbox"/>	VAYXOI01	1280850	0.745
<input checked="" type="checkbox"/>	VAYXOI02	159050	0.745
<input checked="" type="checkbox"/>	VAYXOI03	633382	0.745

 1 - 13 of 13 items

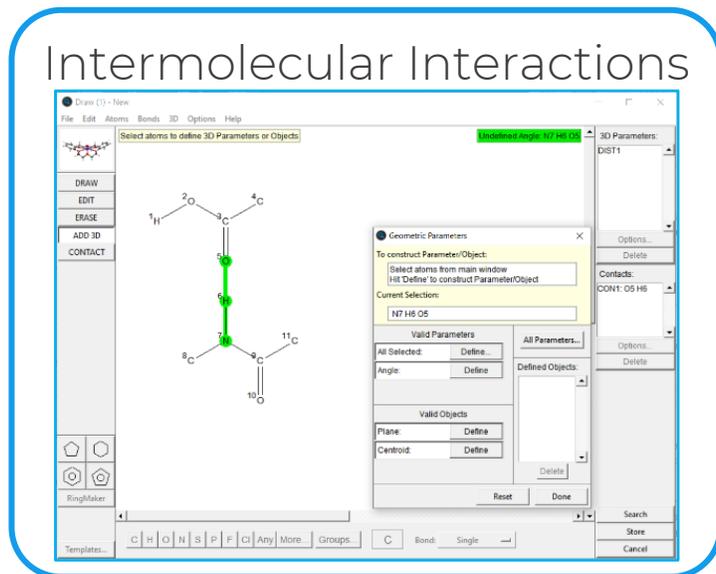
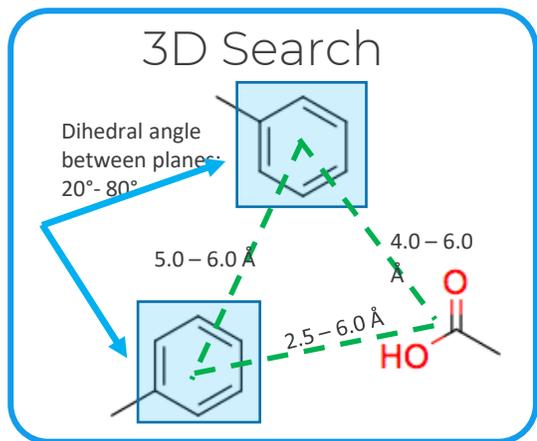
Download 

CCDC

What have we learnt?

- How to search over a million published crystal structures using our desktop software
 - Building queries from the Draw Window
 - Different ways to build queries (Name, Journal, ...)
 - Edit queries
 - View hitlists
- How to refine your results to structures in targeted subsets allowing you to explore drugs, pesticides and MOFs in more detail
 - Restrict search to a subset
 - Bonus: Load and combine subsets

What else can you explore in ConQuest?



Search
In-House
Databases

Export & Analyse Data

CCDC ConQuest (1) : search1 [Search]

File Edit Options View Databases Result

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

ACANIL07

Analyse Hitlist

✓ ACANIL07

✓ ARCLAM04

✓ BOTBEU

✓ BOTBEU0

✓ BOTBIY

✓ BOTBIY01

✓ BOVSUD

✓ BOVYOD

✓ BOVYOD01

✓ BOWJUV

✓ BOXPOW

✓ BOXPUC

✓ BOXQAJ

✓ BOXQIR

✓ BOXRAK

✓ BOXYOF

✓ BOYPEN

✓ BOYPOX

✓ BOYREP

✓ BOYRIT

✓ BOYWIY

✓ RO7VAO

1610 hits

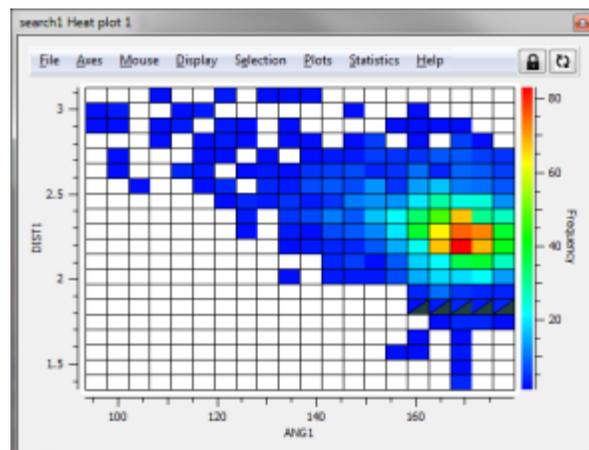
... and more!

Data analysis in Mercury

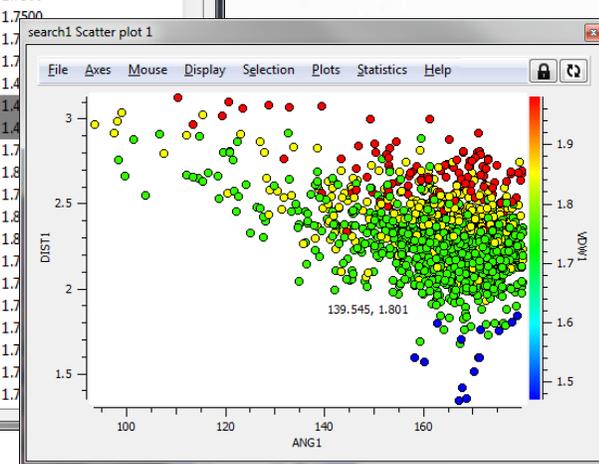
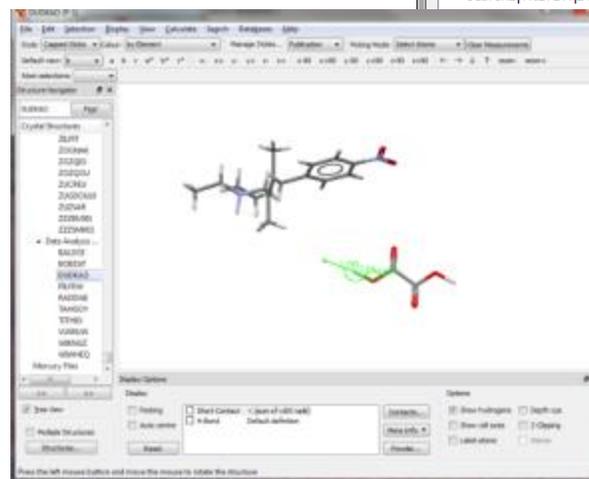
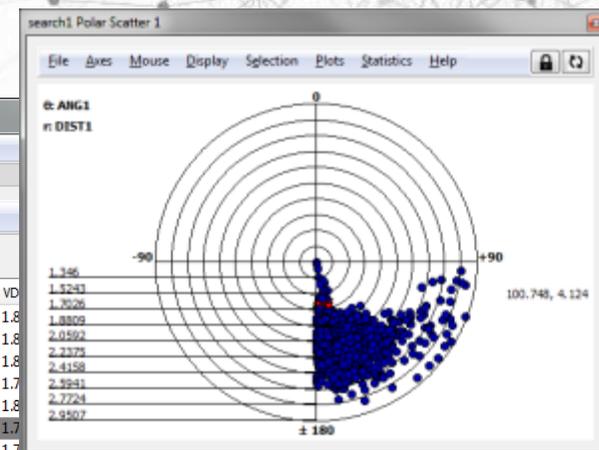
CCDC ConQuest (1) : search3 [Search]

File Edit Options View Databases Results

- 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



NAME	Query	Fragment	ANG1	ANG2	ANO1	DIST1	LAB1	VD
9 WIBLEY	1	1	157.1810	108.0870	35	2.5040	Br1	1.8
10 WIBZEM	1	1	148.3450	105.0420	35	2.0700	Br1	1.8
11 WIFCOE	1	1	157.6580	109.4580	35	2.4070	Br1	1.8
12 WIHVAK	1	1	103.7470	176.6930	17	2.5530	Cl1	1.7
13 WIKMAE	1	1	123.0220	91.0810	35	2.9160	Br1	1.8
14 WIKNUZ	1	1	163.4610	115.6550	17	1.8480	Cl1	1.7
15 WIKNUZ	1	2	142.0390	116.8170	17	1.9980	Cl2	1.7
search1 WILYEW 1456 WILYEW	1	1	136.4500	109.4520	17	2.3990	Cl1	1.7500
search1 WILYEW 1457 WILYEW	1	2	174.2010	109.4290	17	2.2010	Cl1	1.7500
WILYEW	1	3	167.8320	109.4460	17	2.3280	Cl1	1.7500
WISJOX	1	1	168.5380	107.6350	17	1.9930	Cl1	1.7
WISJOX	1	2	167.0650	108.3220	17	2.2280	Cl2	1.7
1 WIWHEQ	1	1	171.3820	109.5370	9	1.7630	F1	1.4
2 WIWHEQ	1	2	178.7710	109.5070	9	1.8470	F2	1.4
3 WIWHEQ	1	3	177.8390	109.4200	9	1.8100	F2	1.4
4 WIWSOK	1	1	167.5080	108.6550	17	2.2260	Cl1	1.7
WIZLOH	1	1	169.6270	117.2500	35	2.4670	Br1	1.8
WIZYAF	1	1	178.4360	106.2840	17	2.3930	Cl1	1.7
7 WOBMEF	1	1	171.8990	109.5670	35	2.3480	Br1	1.8
WOBMUJ	1	1	159.9410	109.4460	35	2.3770	Br1	1.8
69 WOBMUV	1	1	179.9310	113.1610	17	2.3300	Cl1	1.7
... WOBMUV...	1	1	164.1590	112.2520	17	2.3320	Cl1	1.7
1 WOCFOJ	1	1	142.4990	109.5570	17	2.6160	Cl1	1.7
WOJZEB	1	1	174.7420	109.4890	17	2.1980	Cl1	1.7
WOKRIX	1	1	172.7240	111.1030	17	2.1120	Cl1	1.7
74 WOMQUK	1	1	174.1710	111.1890	17	1.9840	Cl1	1.7
75 WOMQUK	1	2	148.4200	101.6230	17	2.0140	Cl1	1.7



Want to explore more?

Training and Educational Resources

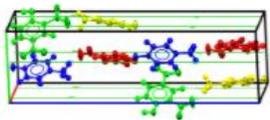
Register for
E&O newsletter

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](#).

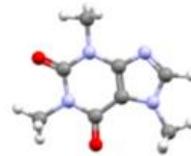


CSD-Core

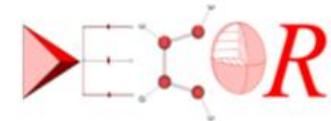
Essential crystallographic and structural chemistry capabilities.



Information on the Teaching Subset



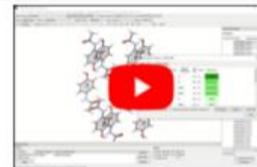
Access a series of teaching modules for use in the classroom



DECOR: Educational Resources for Teaching Crystallography



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



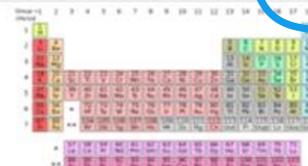
Watch software training and support videos



CSDU modules



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



Explore the Periodic Table through Crystal Structures

Self-guided workshops

On-demand modules with completion certificate

YouTube and LabTube channels

Want to explore more?

More Virtual Workshops in the coming weeks

<https://www.ccdc.cam.ac.uk/News/Events/>



Stability and
design of solid
forms

May 10th
9:30 BST



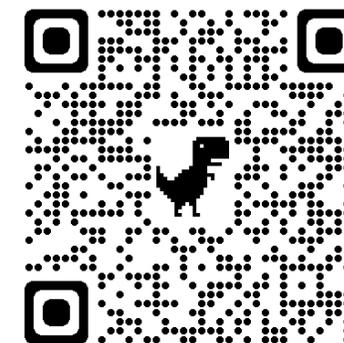
Stability and
design of solid
forms

May 12th
21:30 BST



First steps in
protein ligand
docking

May 18th
11:00 BST



REGISTER
HERE