CCDC

advancing structural science

What's Up

Customer Update Webinar

18th March 2021

CCDC

advancing structural science

Today's presenters

Ioana Sovago

Applications Scientist

Alex Eyes

Customer Success Manager

Seth Wiggin

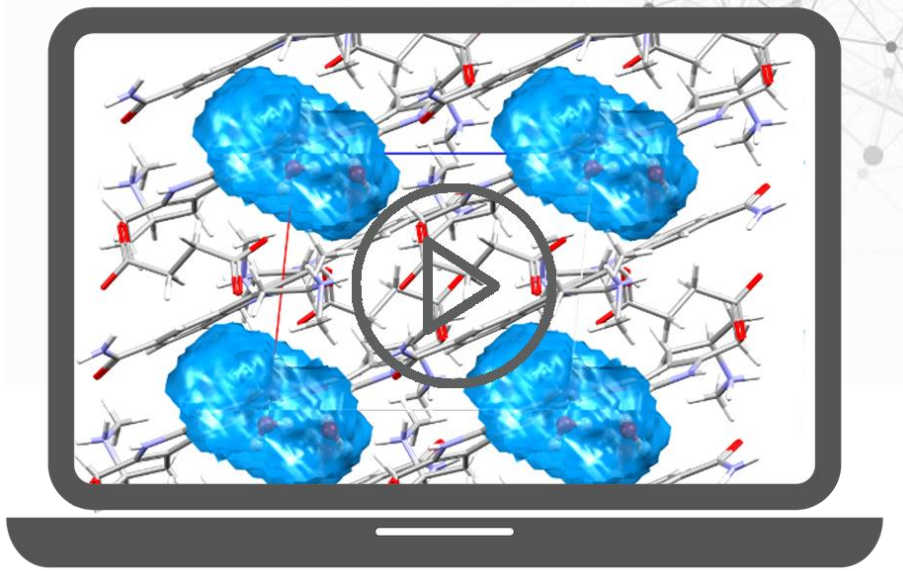
Senior Scientific Editor

CCDC

Overview

In this webinar we will discuss:

- Latest updates and news
- Designing Co-Crystals with CSD-Materials
- CSD MOF collection
- Q&A: the floor is yours



CCDC

Latest updates and news from CCDC >Events

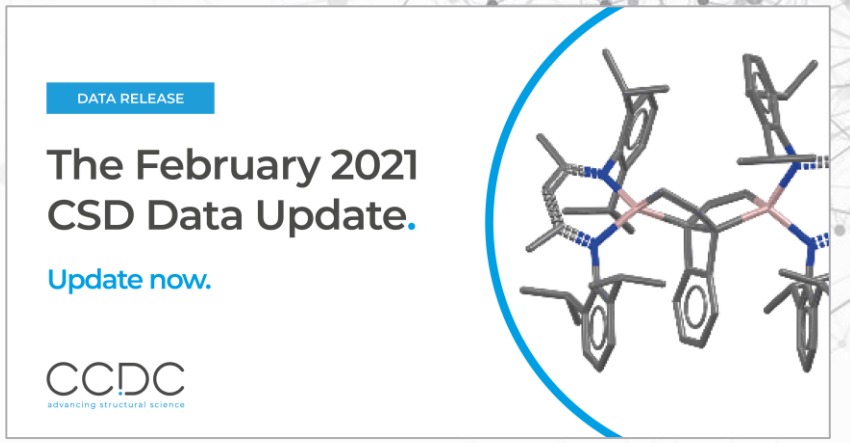
- CCDC **User Group Meetings 2021 – all virtual**:
  - CCDC Educators UGM – 16th and 17th March (watch on demand on our YouTube channel)
  - **CCDC Discovery Science UGM** – 9th and 10th June
  - **CCDC Materials Science UGM** – 7th and 8th September

- **Virtual workshops** series starting soon
  - Visualisation and Analysis in Mercury (Intermediate) – 7th April
  - Pharmacophore Searching with CSD-CrossMiner – 14th April
  - Programmatic search and analysis using the CSD Python API – 21st April

CCDC

Latest updates and news from CCDC

- CSD data update – February 2021
  - **20,058 new structures** and increasing the total size of the CSD to over 1.1 million structures.



- Latest **CSD Release**
  - 2020.3.1 CSD Release – patch release for users working on macOS 11.0 Big Sur update
  - Find out more about what's new in our website https://www.ccdc.cam.ac.uk/solutions/whats-new/

CCDC

The CCDC team is growing…

- **Recently joined CCDC**
  - Susan Reutzel-Edens - Head of Science
  - Joseph Rossetto – UX/UI Specialist

Just a reminder…

- CSD-KNIME and Pipeline Pilot are still available, if you haven't tried them yet, please do – we welcome your suggestions.
- To get the latest version of the 2020.3 you need to do a full uninstall/reinstall.
- CSD-Materials applications: DASH will be migrating to open source in late 2021/ early 2022

CCDC

CSD Materials

Designing Co-Crystals with CSD-Materials

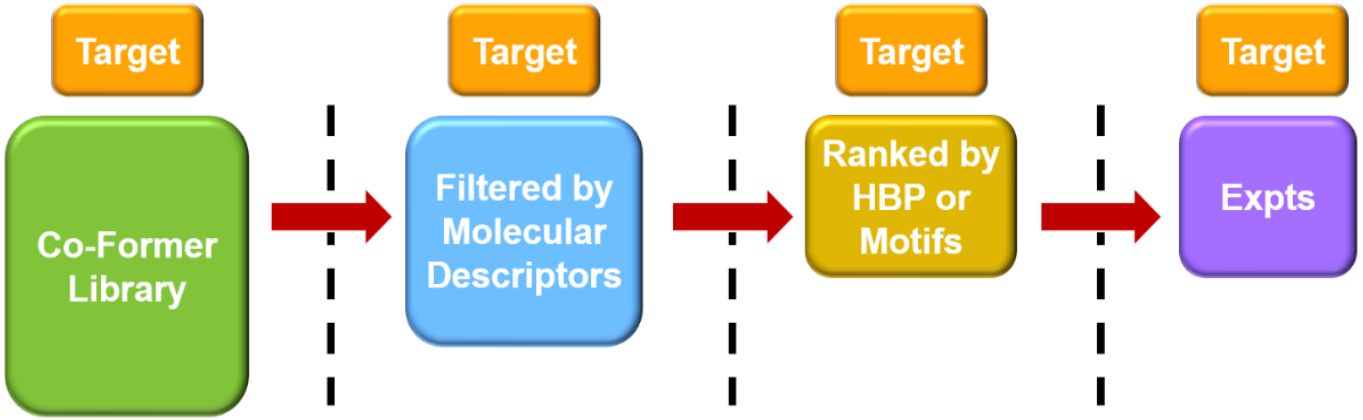
Ioana Sovago

Applications Scientist

CCDC

Co-crystal design approach

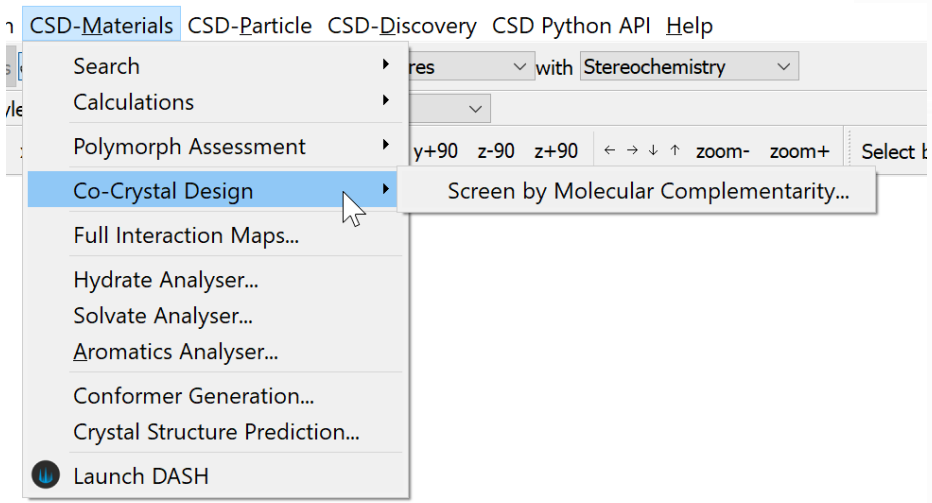
- The CCDC recommends a two-stage screening approach:
  - Stage 1: Molecular complementarity analysis – this is extremely fast and an ideal method for filtering out very unlikely co-formers from a list
  - Stage 2: Motif search or HBP analysis – this is more time-consuming, but is effective for ranking co-formers to improve enrichment

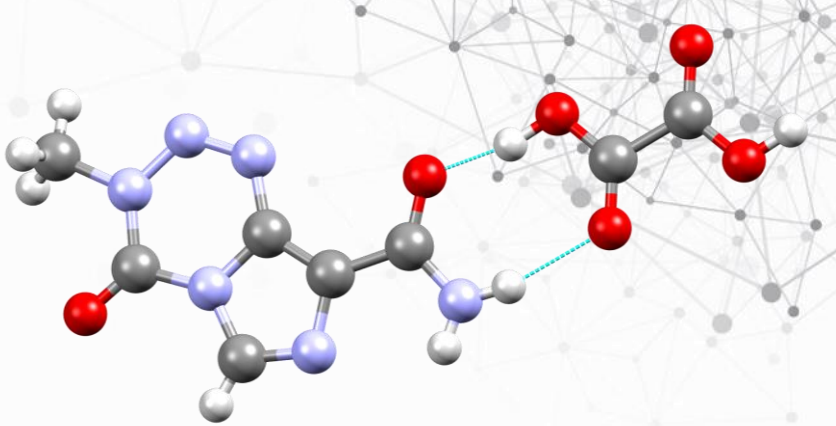


P. A. Wood, N. Feeder *et al.*, *CrystEngComm* (2014) 16, 5839-5848.

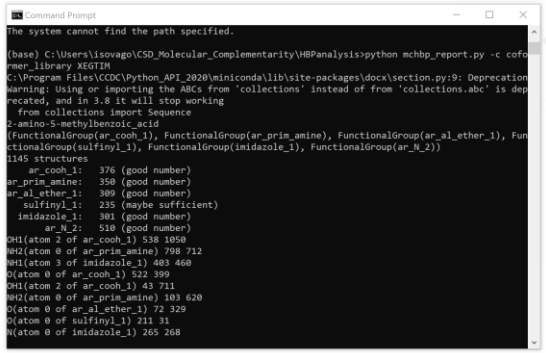
Designing Co-Crystals with CSD-Materials

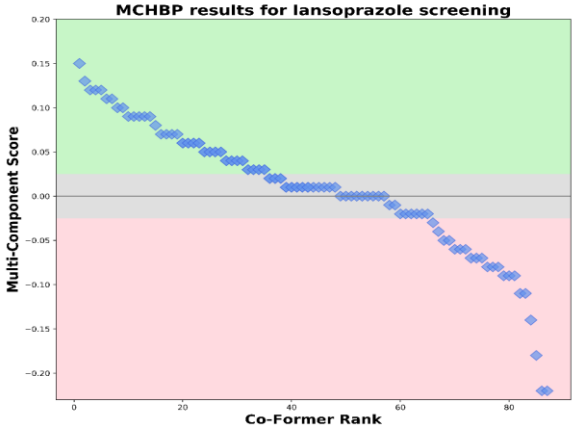
Stage 1:

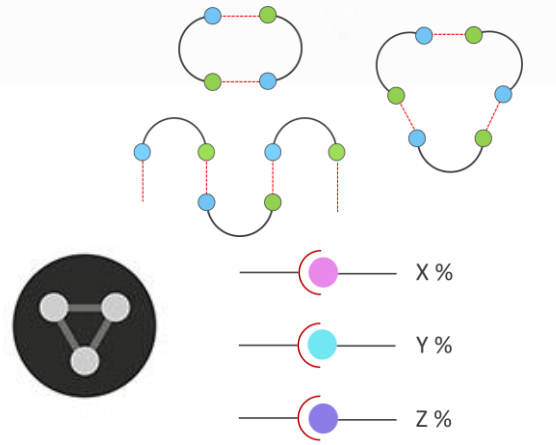




Stage 2:



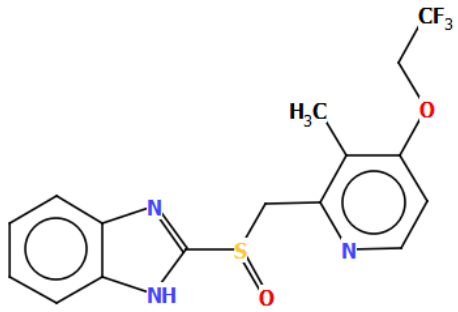


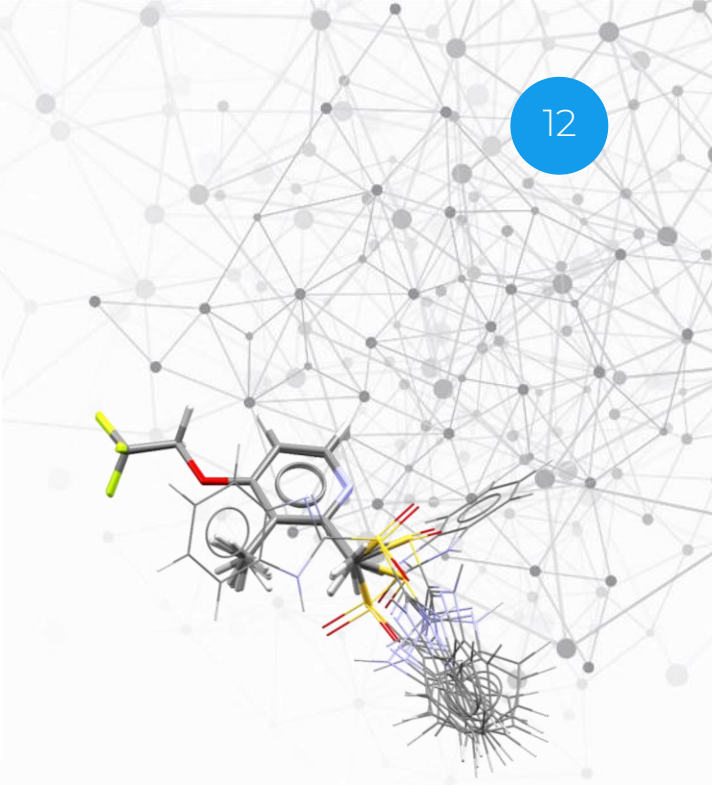


CCDC

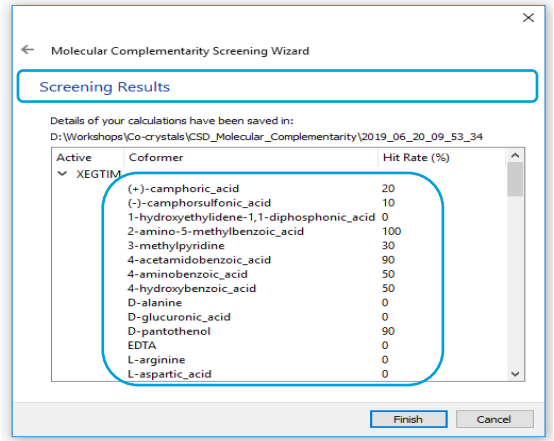
Co-crystal design approach

- Stage 1
  - Identify the target molecule (or “active”)
  - Generate an ensemble of likely molecular conformations for the target molecule
  - Identify the full co-former library to be considered
  - Run *Screen by Molecular Complementarity* (in CSD-Materials)





(-)-camphorsulfonic_acid
(+)-camphoric_acid
1-hydroxyethylidene-1,1-diphosphonic_acid
2-amino-5-methylbenzoic_acid
3-methylpyridine
4-acetamidobenzoic_acid
…

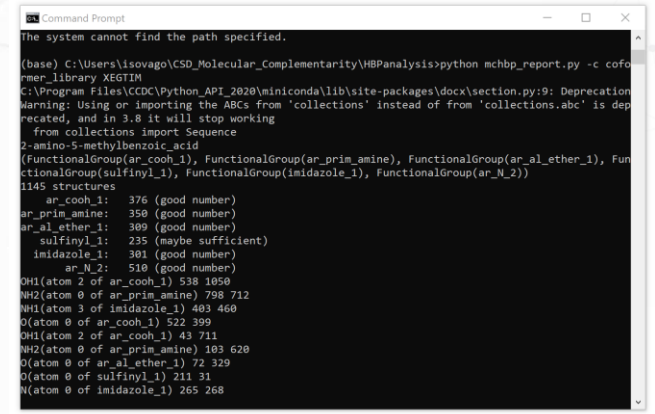


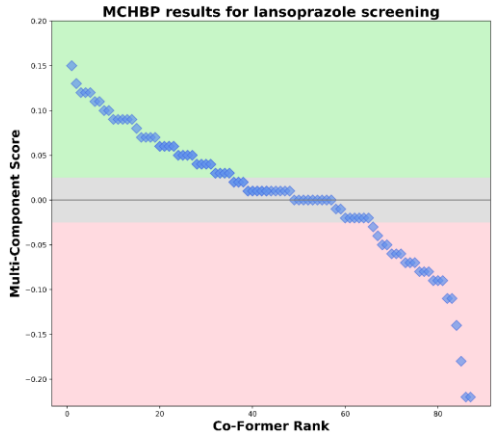
CCDC

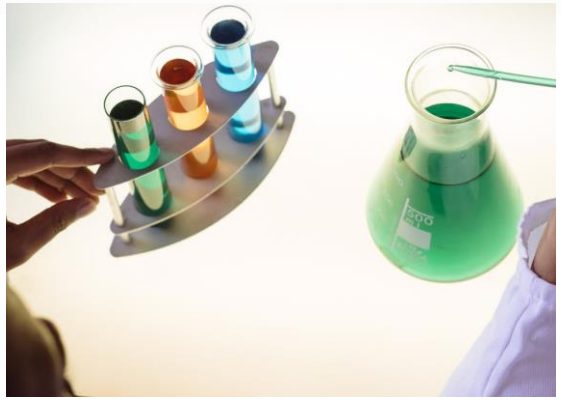
Co-crystal design approach

- Stage 2
  - Identify the co-formers that have passed stage 1 screening
  - Use one of the these approaches to rank co-formers that passed stage 1:
    - *Hydrogen-bond motif search* (in CSD-Materials)
    - *Screen by hydrogen-bond propensity* (in CSD-Materials)
  - Start experiments from the top of the ranked list of co-formers and work down depending on how many experiments you choose to do

	A	B	C
1	Results Summary		
2			
3	Coformer	Hit Rate %	
4	(+)-camphoric_acid	20	
5	(-)-camphorsulfonic_acid	10	
6	1-hydroxyethylidene-1,1-diphosphonic_acid	0	
7	2-amino-5-methylbenzoic_acid	100	
8	3-methylpyridine	30	
9	4-acetamidobenzoic_acid	90	
10	4-aminobenzoic_acid	50	
11	4-hydroxybenzoic_acid	50	
12	D-alanine	0	
13	D-glucuronic_acid	0	
14	D-pantothenol	90	
15	EDTA	0	
16	L-arginine	0	
17	L-aspartic_acid	0	



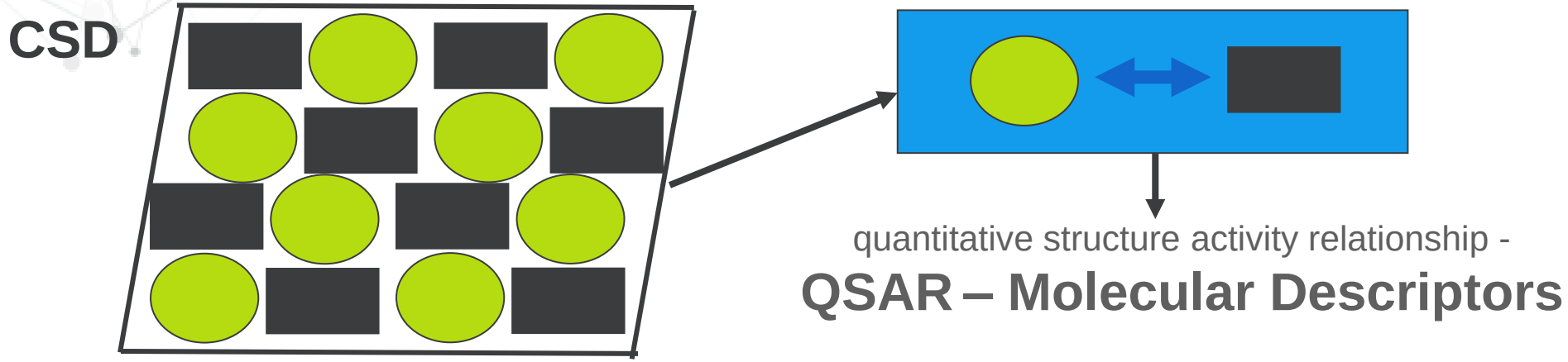




CCDC

Molecular complementarity

- Molecular Complementarity (Fábián) Method*



- Calculated QSAR-type molecular descriptors (131 in total) for each pair of molecules in each co-crystal

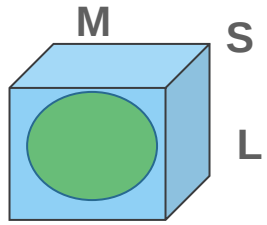
- Molecules that form co-crystals tend to have *similar* properties

CCDC

*L. Fábián, *Cryst. Growth Des.* (2009) 9, 1436-1443.

Molecular complementarity

- Surveyed the CSD – used QSAR principles to select 5 most statistically significant descriptors + cut-offs
  - Shape: S axis length
  - Shape: S/L axis ratio
  - Shape: M/L axis ratio
  - Fraction of N/O atoms
  - Dipole moment



Shape Descriptors

FNO= (#O+#N)/(#Heavy atoms)

- Co-crystal formation is predicted to be likely if all descriptors differ by less than the cut-off values among the molecules*

CCDC

***L. Fábián, *Cryst. Growth Des.* (2009) 9, 1436-1443.**

Illustrative example - MC

- Considerable enrichment can be achieved experimentally*

	API-CF pairings	Co-crystal observed	No co-crystal observed	Percentage co-crystals
Total	218	48	170	22%
Predicted to be Likely	107	40	67	37%
Predicted to be Unlikely	111	8	103	7%

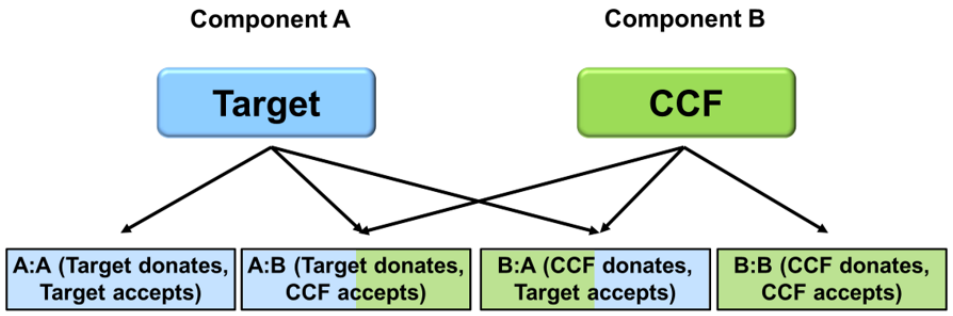
- 49.1% of the experiments, but a retention of 83.3% of the hits
- Can remove 111 experiments which only have a 7% hit rate

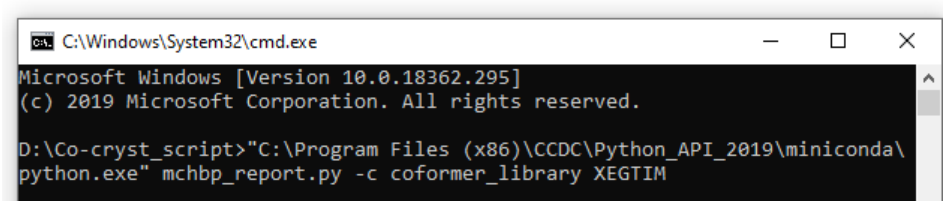
CCDC

*L. Fábián, unpublished results.

HBP co-former selection

Hydrogen-bond propensity permutations





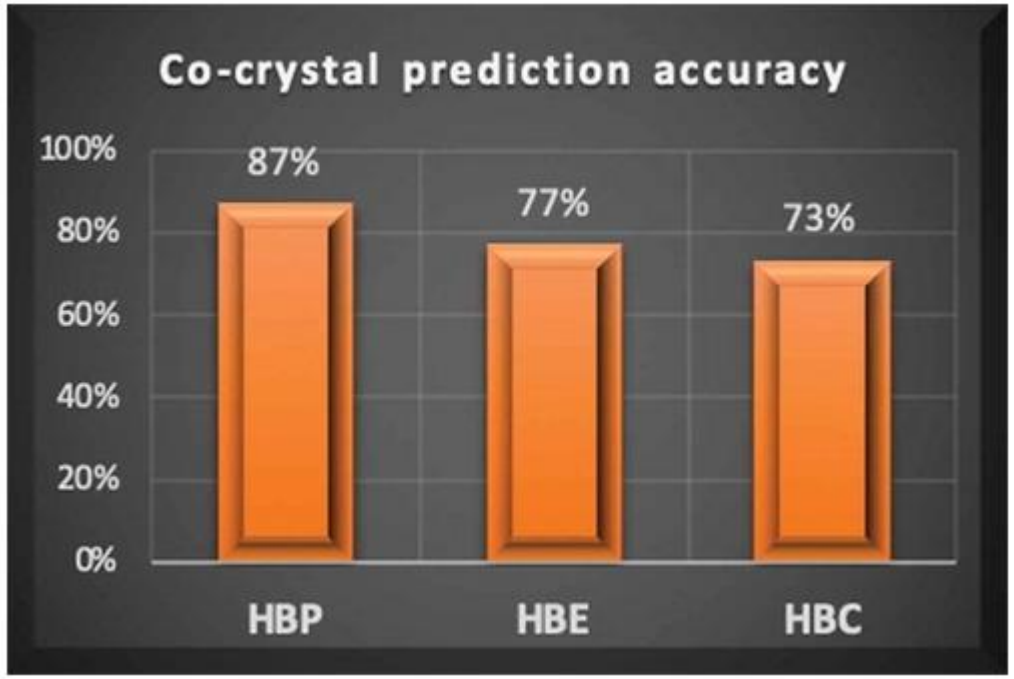
$$\text{MC Score} = \text{Best Hetero Interaction} - \text{Best Homo Interaction}$$

(Highest propensity A:B or B:A interaction) (Highest propensity A:A or B:B interaction)

Rank	Component B	MC score	Max. Interaction	Max A:B or B:A propensity	Max A:A propensity	Max B:B propensity
1	L-arginine	0.15	B:A	0.92	0.77	0.7
2	folic_acid	0.13	B:A	0.91	0.78	0.72
3	glycine	0.12	B:A	0.9	0.78	0.75
4	biotin	0.12	B:A	0.92	0.8	0.71
5	t-butylamine	0.12	B:A	0.88	0.76	0.65
6	L-serine	0.11	B:A	0.85	0.74	0.63
7	D-alanine	0.11	B:A	0.89	0.78	0.69
8	L-methionine	0.1	B:A	0.88	0.78	0.62
9	L-leucine	0.1	B:A	0.88	0.78	0.61
10	L-glutamic_acid	0.09	B:A	0.88	0.79	0.64
11	alitame	0.09	B:A	0.87	0.78	0.69
12	L-glutathione	0.09	B:A	0.86	0.77	0.68
13	L-aspartic_acid	0.09	B:A	0.88	0.79	0.65
14	L-glutamine	0.09	B:A	0.89	0.8	0.8
15	4-aminobenzoic_acid	0.08	B:A	0.89	0.81	0.59
16	lactobionic_acid	0.07	B:A	0.83	0.76	0.66
17	acesulfame	0.07	B:A	0.86	0.79	0.6
18	2-amino-5-methylbenzoic_acid	0.07	B:A	0.88	0.81	0.54
19	lactose	0.07	B:A	0.9	0.83	0.76
20	maltitol	0.06	B:A	0.87	0.81	0.68

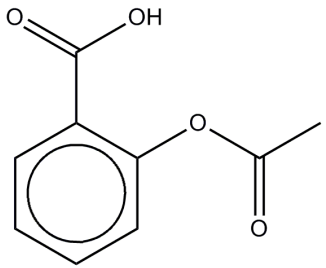
Illustrative example HBP

- Evaluation of hydrogen-bond propensity, hydrogen bond coordination and hydrogen-bond energy tools
- Two API's were investigated Nevirapine and Diclofenac
- The highs accuracy is observed for HBP 87%



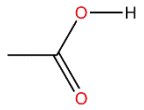
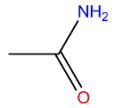
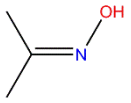
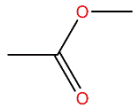
C. Aakeröy, *et al.*, *Supramolecular Chemistry*, (2019), 32:2, 81-90

Motif searching



- There are two polymorphs of aspirin which share significant structural similarity including a carboxylic acid H-bonded dimer

- Let's assume we want to break that interaction…

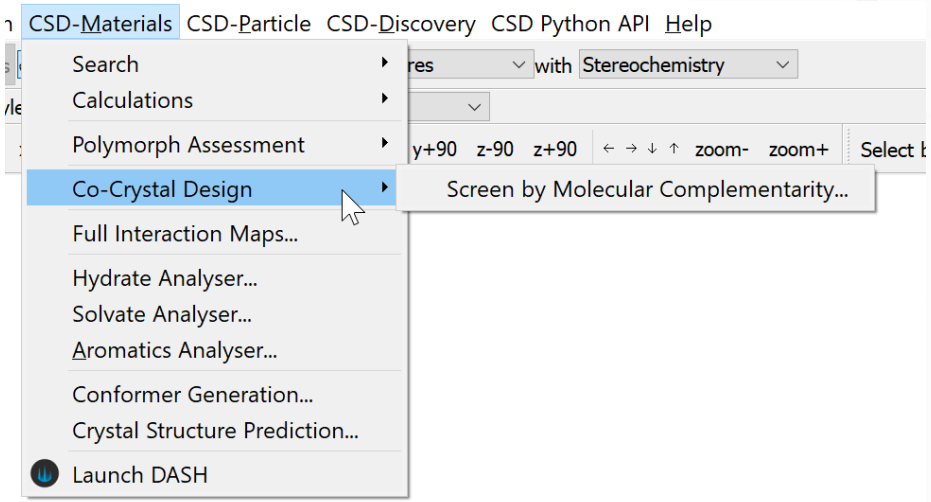
2D Diagram	Dimer	Freq. of Occurrence
	carboxylic acid···carboxylic acid	29.7 %
	carboxylic acid···carboxamide	50.3 %
	carboxylic acid···oxime	59.5 %
	carboxylic acid···hydroxy	33.8 %
	carboxylic acid···ester	15.0 %

Groups to try

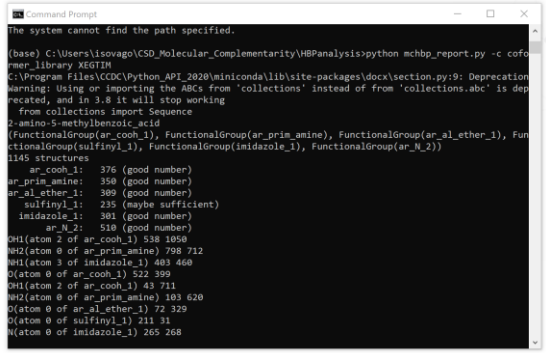
CCDC

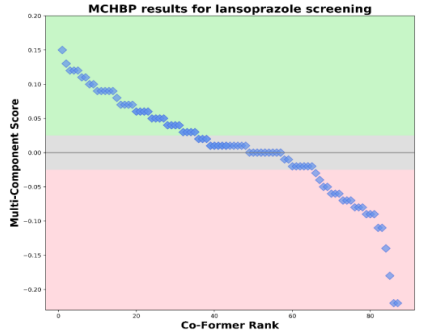
Live Demo

Stage 1:



Stage 2:





CCDC

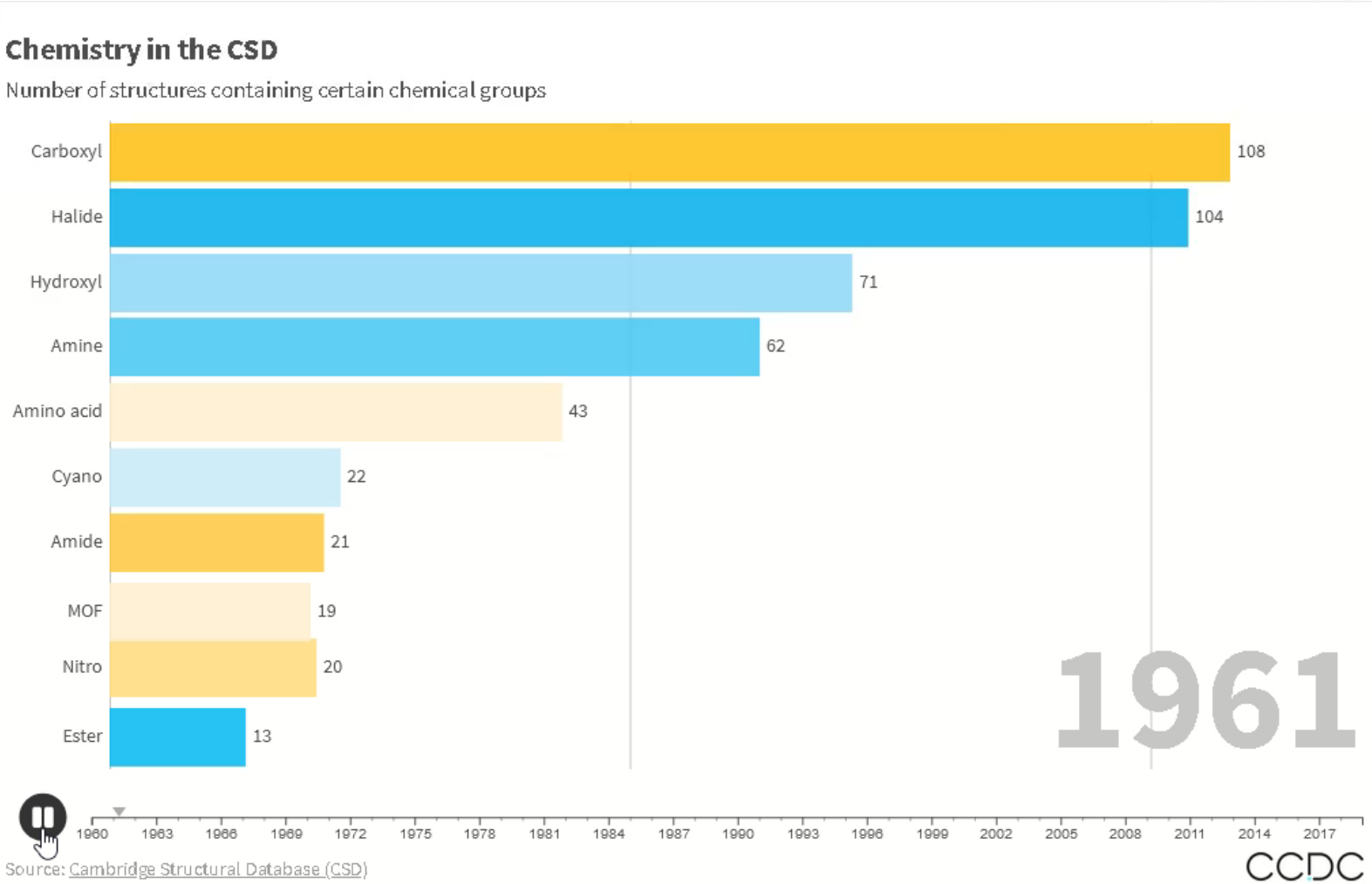
CSD

CSD MOF collection.

Seth Wiggin

Senior Scientific Editor

CCDC


Chemistry in the CSD
Number of structures containing certain chemical groups
Carboxyl
Halide
Hydroxyl
Amine
Amino acid
Cyano
Amide
MOF
Nitro
Ester
108
104
71
62
43
22
21
19
20
13
1961
1960
1963
1966
1969
1972
1975
1978
1981
1984
1987
1990
1993
1996
1999
2002
2005
2008
2011
2014
2017
Source: Cambridge Structural Database (CSD)
CCDC

CCDC

- CSD MOF subset
- Tools for MOF analysis
- Results from analysis of the MOF subset

Chemical Science, 2020, **11** 8373, **DOI:** 10.1039/D0SC01297A

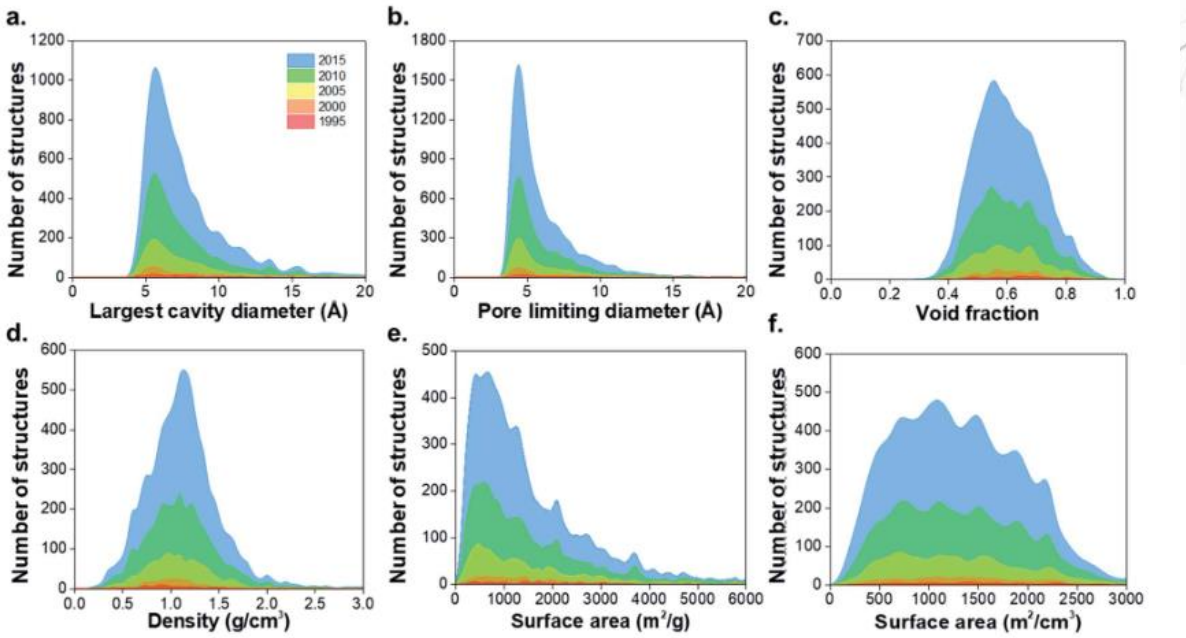
CrystEngComm, 2020, **22** 7152-7161, **DOI:** 10.1039/D0CE00299B

Chem. Mater., 2017, **29** 2618-2625, **DOI:** 10.1021/acs.chemmater.7b00441

CCDC

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

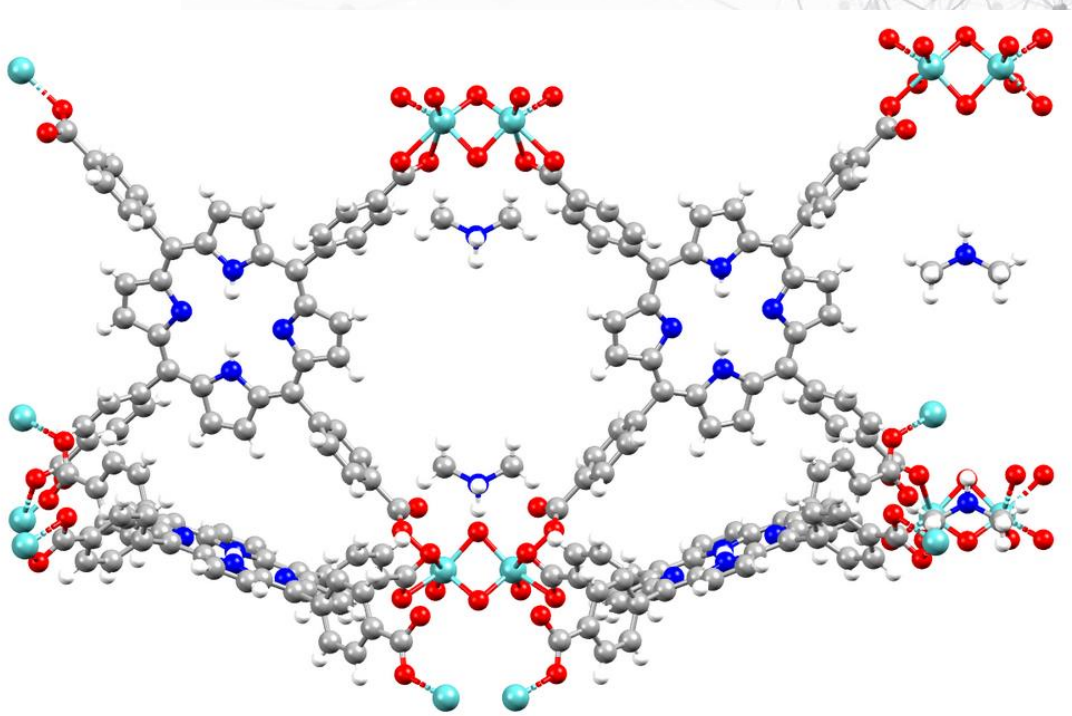
- Topology
- Pore size
- Void space
- Surface area



CSD MOF Collection

- 10,636 3D porous MOFs
  - All 3D MOFs containing over 10 % void space
  - Converted to space group P1
  - Non-bonded solvent removed
  - Missing hydrogen atoms added

CSD MOF Collection Entry



CCDC

CSD MOF Collection

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

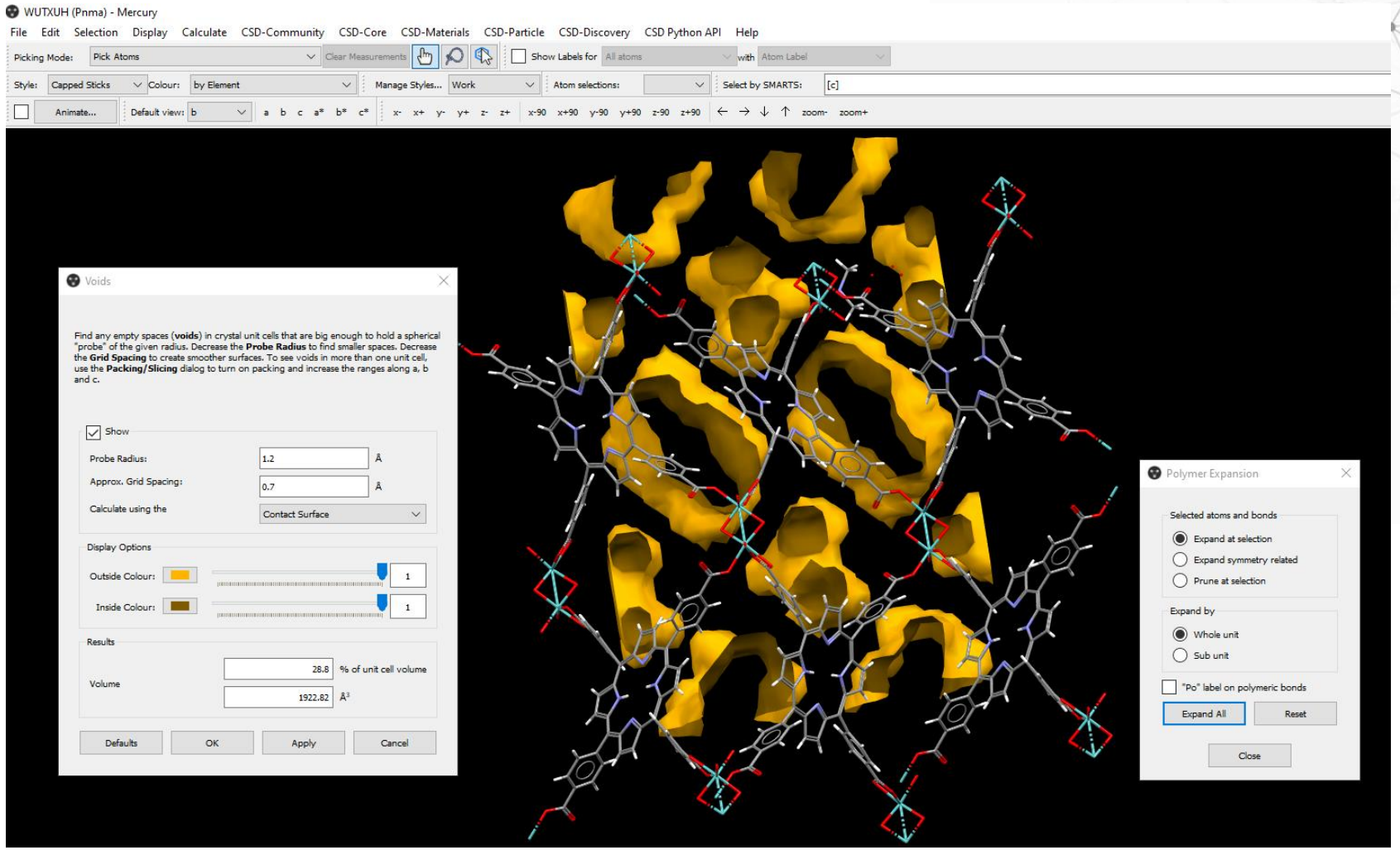
CSD refcode	CIF filename	original crystal system	Sohncke space group	Percentage void space	Charged framework	Hydrogen added	Unreliable chemistry
ABAVIJ	abavij_P1	monoclinic	-	15.538	-	-	-
ABAVOP	abavop_P1	monoclinic	-	13.946	-	-	-
ABAYIO	abayio_P1_charged	cubic	-	62.844	Yes	-	-
ABAZAF	abazaf_P1	monoclinic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

- The CSD MOF Collection has a Creative Commons licence: CC BY-NC-SA 4.0

- https://www.ccdc.cam.ac.uk/Community/csd-community/csd-mof-collection/

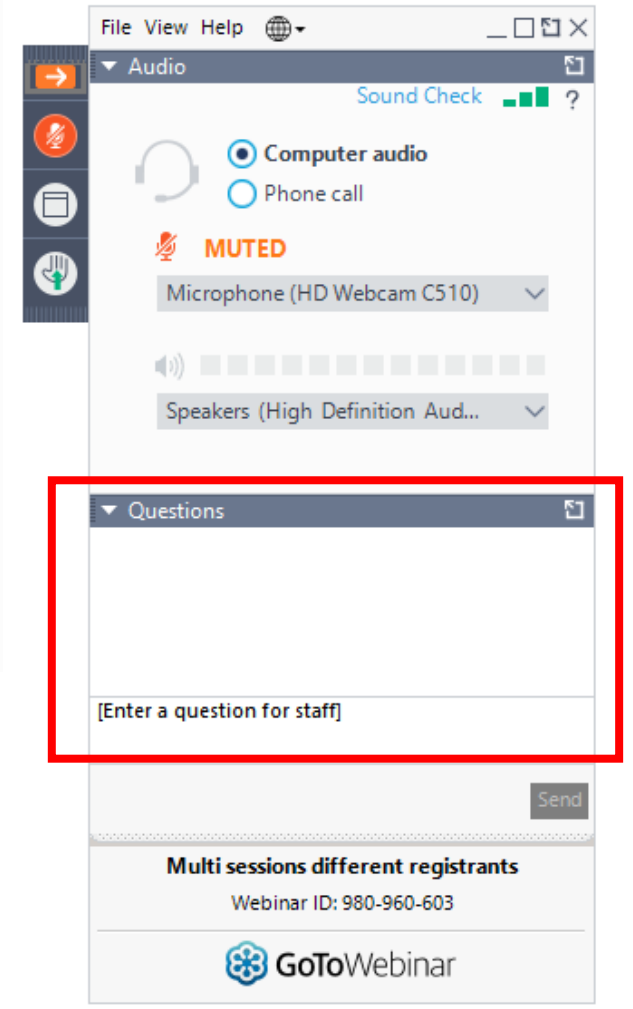
CCDC

Live Demo


WUTXUH (Pnma) - Mercury
File Edit Selection Display Calculate CSD-Community CSD-Core CSD-Materials CSD-Particle CSD-Discovery CSD Python API Help
Picking Mode: Pick Atoms
Clear Measurements
Show Labels for
All atoms
with
Atom Label
Style: Capped Sticks
Colour: by Element
Manage Styles...
Work
Atom selections:
Select by SMARTS:
[c]
Animate...
Default view: b
zoom-
zoom+
Voids
Find any empty spaces (voids) in crystal unit cells that are big enough to hold a spherical "probe" of the given radius. Decrease the Probe Radius to find smaller spaces. Decrease the Grid Spacing to create smoother surfaces. To see voids in more than one unit cell, use the Packing/Slicing dialog to turn on packing and increase the ranges along a, b and c.
Show
Probe Radius: 1.2 Å
Approx. Grid Spacing: 0.7 Å
Calculate using the Contact Surface
Display Options
Outside Colour:
Inside Colour:
Results
Volume 28.8 % of unit cell volume
1922.82 Å³
Defaults
OK
Apply
Cancel
Polymer Expansion
Selected atoms and bonds
Expand at selection
Expand symmetry related
Prune at selection
Expand by
Whole unit
Sub unit
"Po" label on polymeric bonds
Expand All
Reset
Close

Q&A

- Type your questions in the box as shown

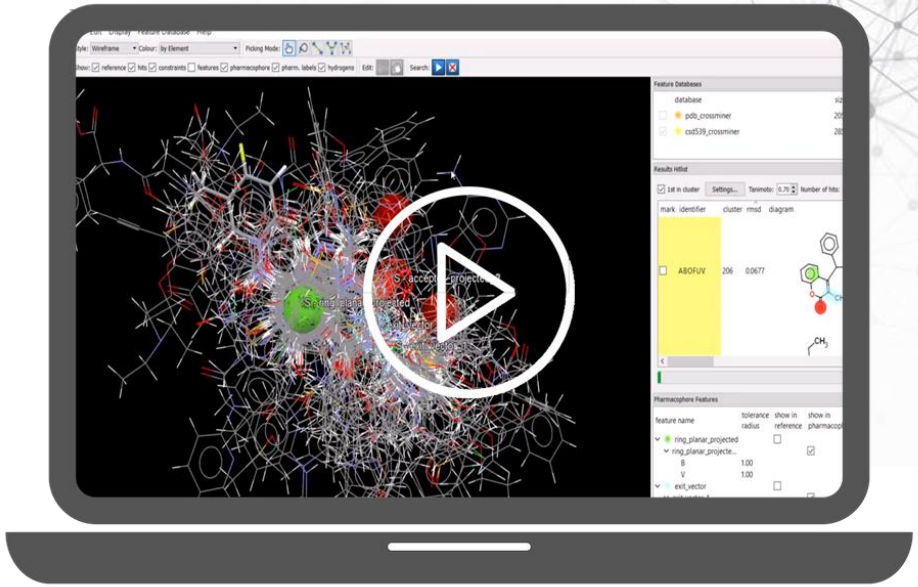


CCDC

Next What's Up Webinar

- Next webinar: 20th May
  - CSD-Discovery: Ligand Overlay
  - Agile Development at CCDC
- Follow us on social media
- Send us your ideas and news

  hello@ccdc.cam.ac.uk



CCDC

Thank you

hello@ccdc.cam.ac.uk

The Cambridge Crystallographic Data Centre
12 Union Road, Cambridge CB2 1EZ, United Kingdom
Registered Charity No. 800579

CCDC