CCDC

advancing structural science

What's Up

Customer Update Webinar

18th November 2021

CCDC

advancing structural science

Today’s presenters

Pete Wood

Director of Product Management

Sophie Bryant

Marketing Manager

Francesca Stanzione

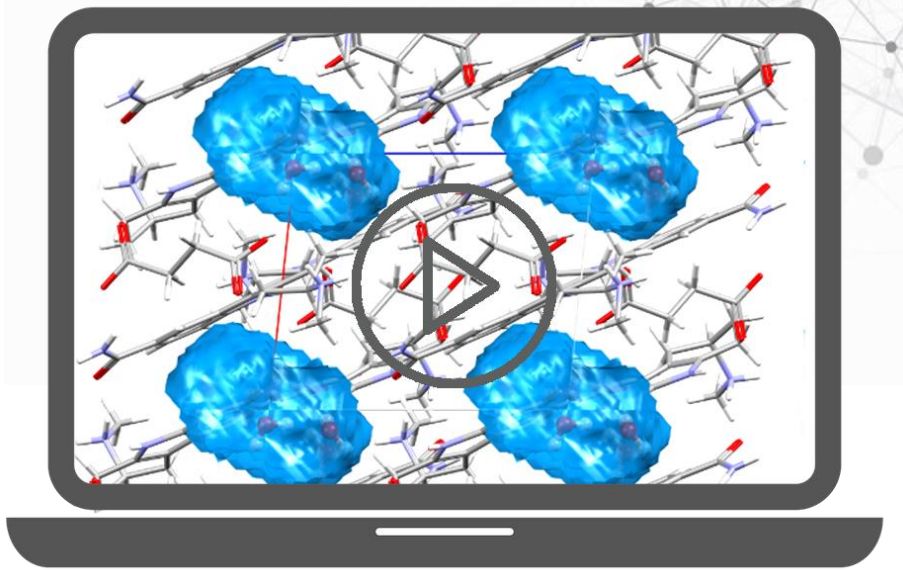
Research and Applications Scientist

CCDC

Overview

In this webinar we will discuss:

- Latest updates and news
- DASH: What has changed?
- CSD-CrossMiner: New features from the 2021.2 release
- Q&A: the floor is yours



CCDC

Latest updates

- Upcoming 2021.3 CSD release in December: DASH migration, new CSD subsets, and usability improvements for CSD-Core solutions.

- 7th CSP Blind Test: we released the 2D chemical structures of the target compounds.
  Learn more www.ccdc.cam.ac.uk/Community/initiatives/CSPBlindTests/

- CSDU new module available: Analysing molecular geometries 101 – basics of Mogul. www.ccdc.cam.ac.uk/Community/educationalresources/CSDU/

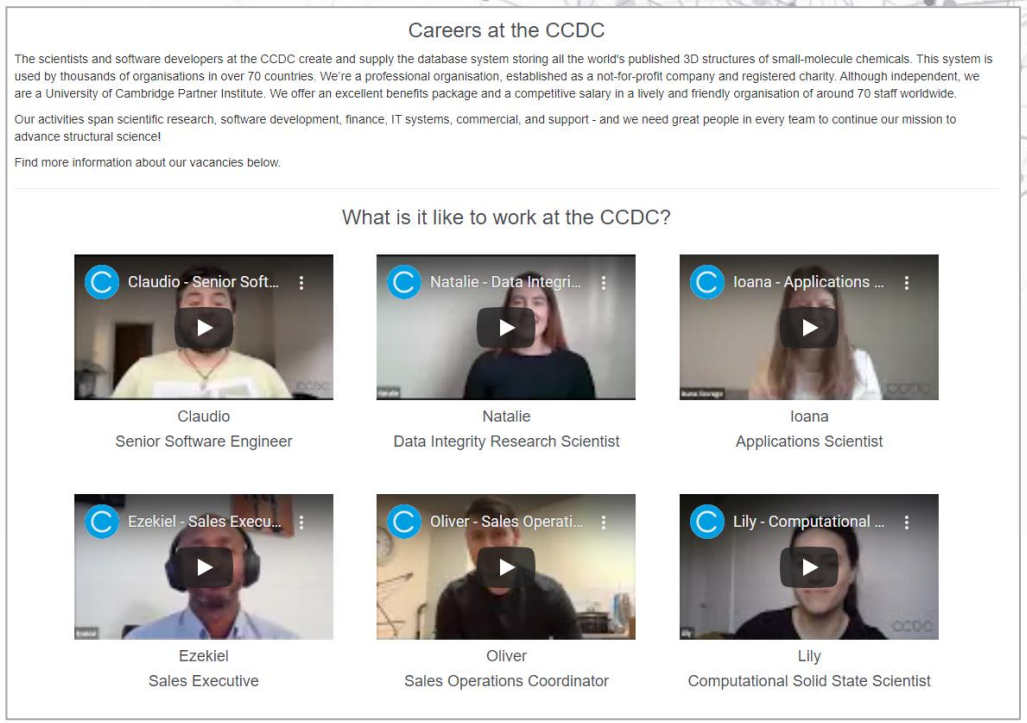
CCDC

Time for a change?

- We are hiring!
  - Postdoctoral Research Associate – Interactions
  - Research and Applications Scientist – Materials Science
  - Product Manager
  - Technical Manager
  - USA – Research and Applications Scientist – Discovery Science

- Work for a registered charity, active research centre and partner institute of the University of Cambridge.
- Advancing structural science for the public benefit.

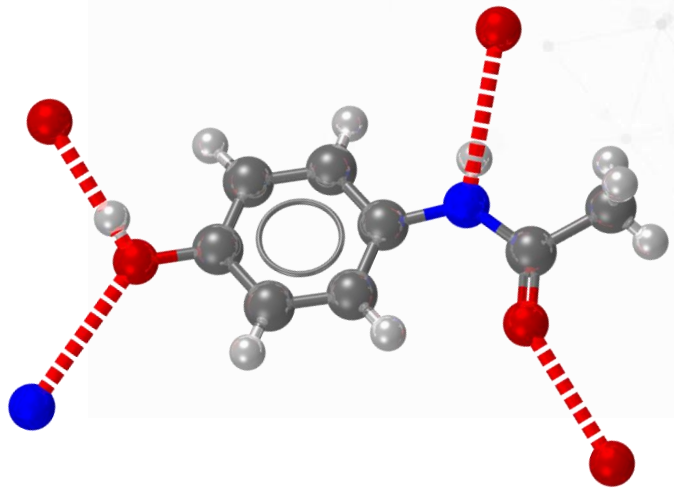
https://www.ccdc.cam.ac.uk/theccdcprofile/careers/



We need you

Are you interested in helping us to evolve WebCSD and our 3D visualiser?

- Let us know in the question box!
- Just drop us an email hello@ccdc.cam.ac.uk
- User research sessions should take < 1 hour



CCDC

DASH: What has changed?

Structure solution from powder data – *for the community*

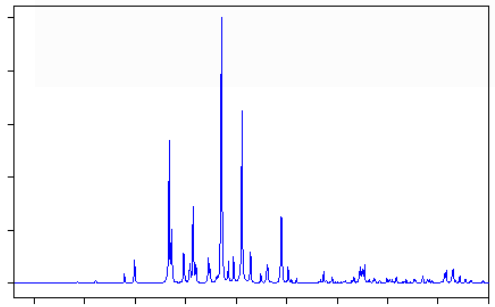
Pete Wood

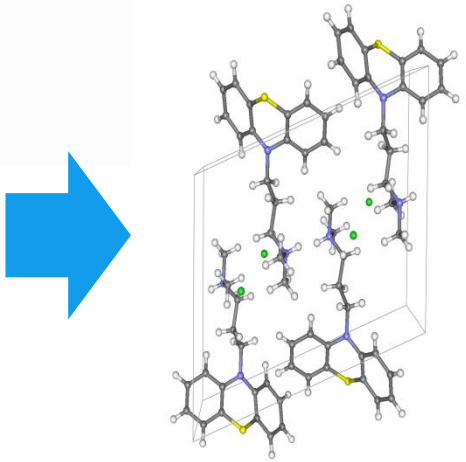
Director of Product Management

CCDC

What is DASH?

- One of the most well-known and effective programs for crystal structure solution from power diffraction data
- DASH is named after the program's original authors Bill David and Kenneth Shankland
- First presented in a paper in 1998, then distributed by CCDC since April 2001 under licence from STFC Innovations Limited
- Designed with a user-friendly interface and extensive manual
- Powerful integration with Mogul for search space reduction



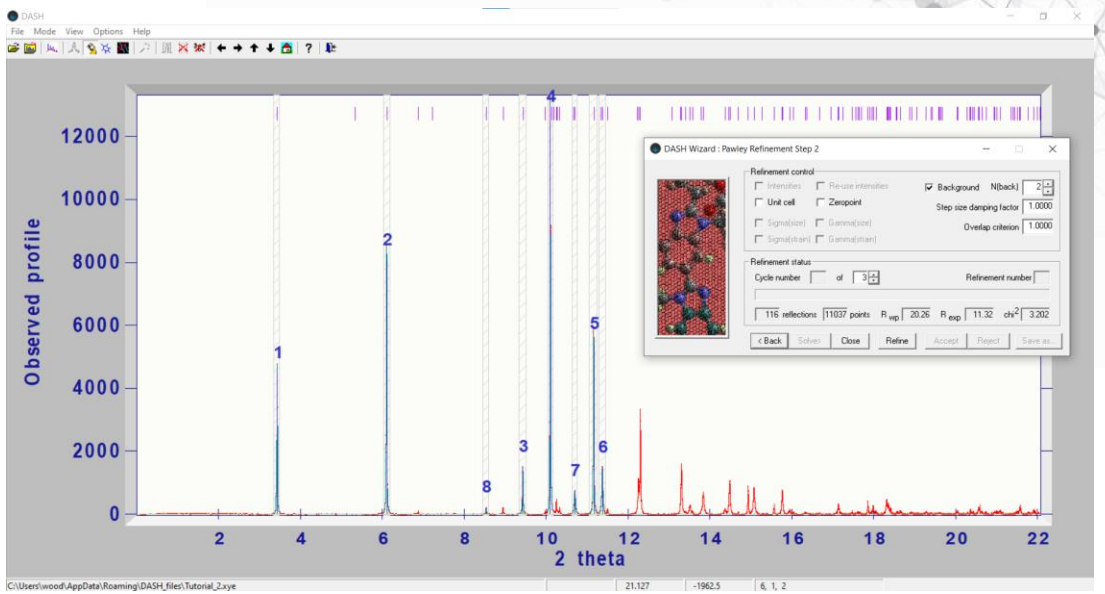


W. I. F. David, K Shankland, J. van de Streek, E. Pidcock, W. D. S. Motherwell, J. C. Cole. J. Appl. Cryst., 39, 910-915.

CCDC

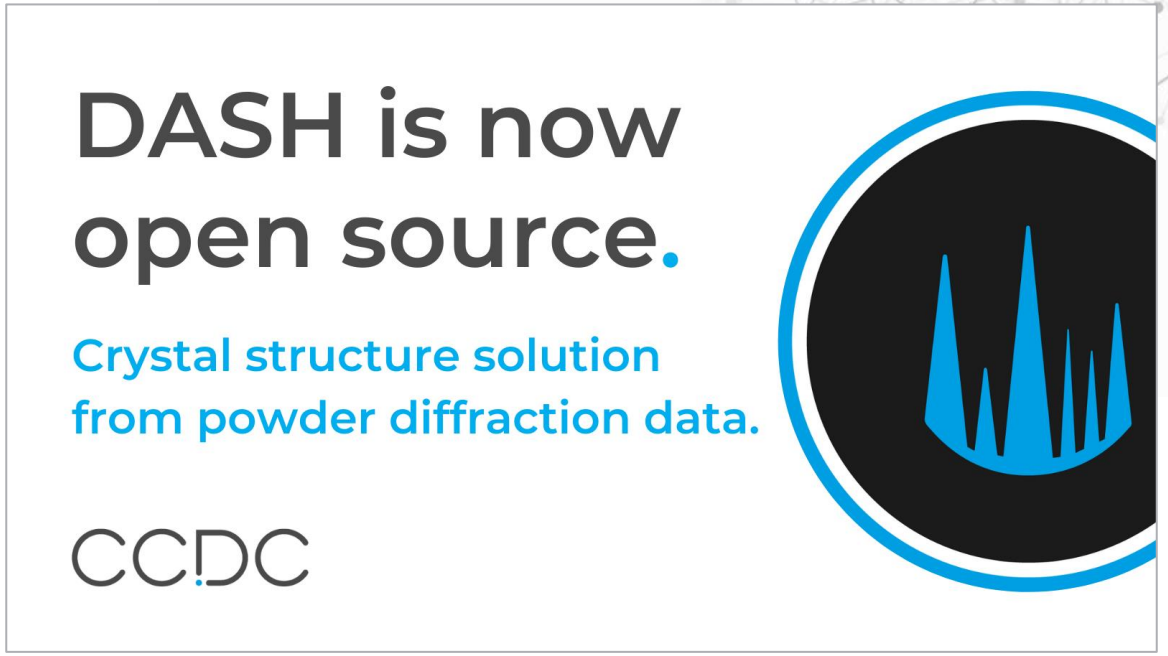
How does DASH work?

- DASH solves crystal structures from powder diffraction data
- The approach assumes you know the molecular formula of the material – you input this 3D model at the start
- The unit cell and space group are determined by indexing the pattern
- DASH generates many trial solutions and uses simulated annealing to adjust the molecular position and conformation until it agrees well with the pattern
- Mogul can help by dramatically reducing conformational search space
- All steps are guided using a UI Wizard



Why is the situation with DASH changing?

- As the CSD software portfolio has developed and expanded, DASH is now at the periphery
- The CCDC has not been able to spend much time developing its functionality in recent years
- There are still interesting directions that could be explored with DASH though
- To ensure DASH remains available to use, and can grow and change, the CCDC, STFC and DASH's co-authors have agreed to make it open to the community
- This will also allow our team more time to focus on other areas of CSD-Materials



https://www.ccdc.cam.ac.uk/solutions/open-source/dash-open-source/

What does that mean for me as a user?

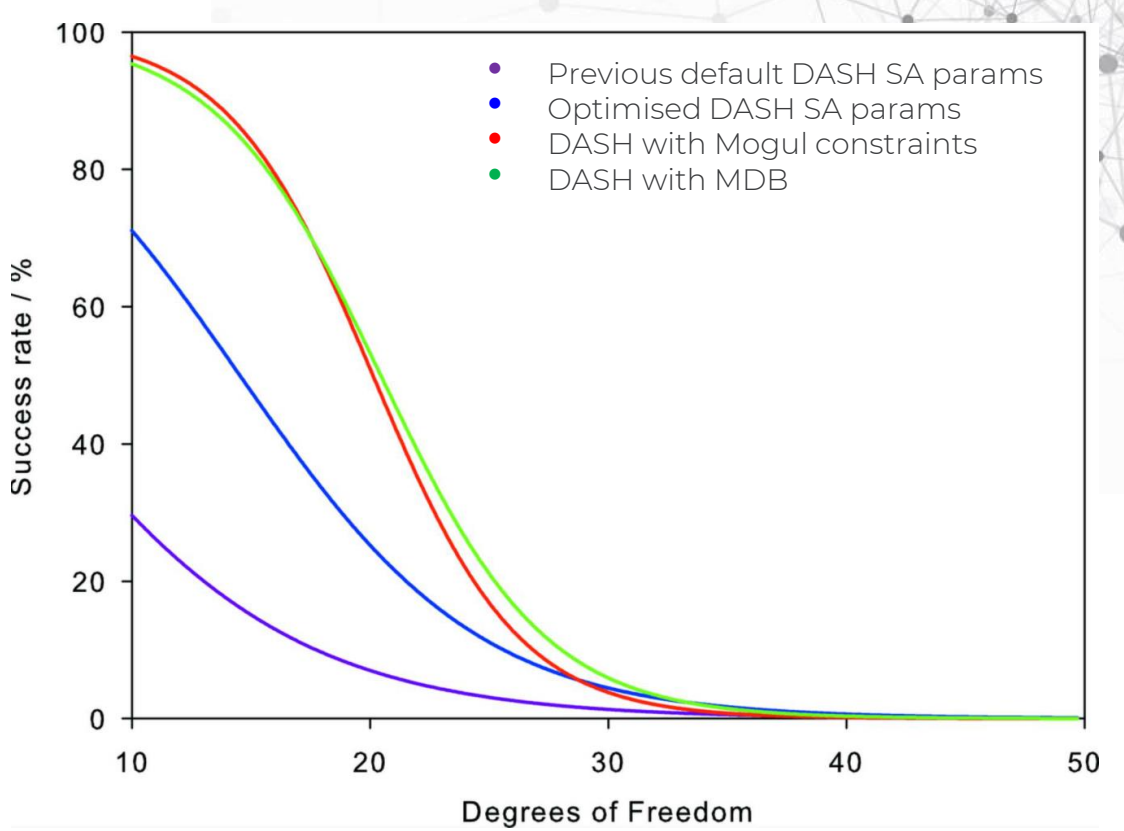
- To install DASH, you will need to get the installer from the new DASH GitHub repository
- From the 2021.3 CSD Release (December 2021), DASH will no longer be installed as a part of the CSD software portfolio
- In March 2022, older versions of DASH will also stop working
- DASH is now available for free, for anyone to use and adapt, under the MIT licence

https://github.com/ccdc-opensource/dash

CCDC

Will DASH still integrate with Mogul?

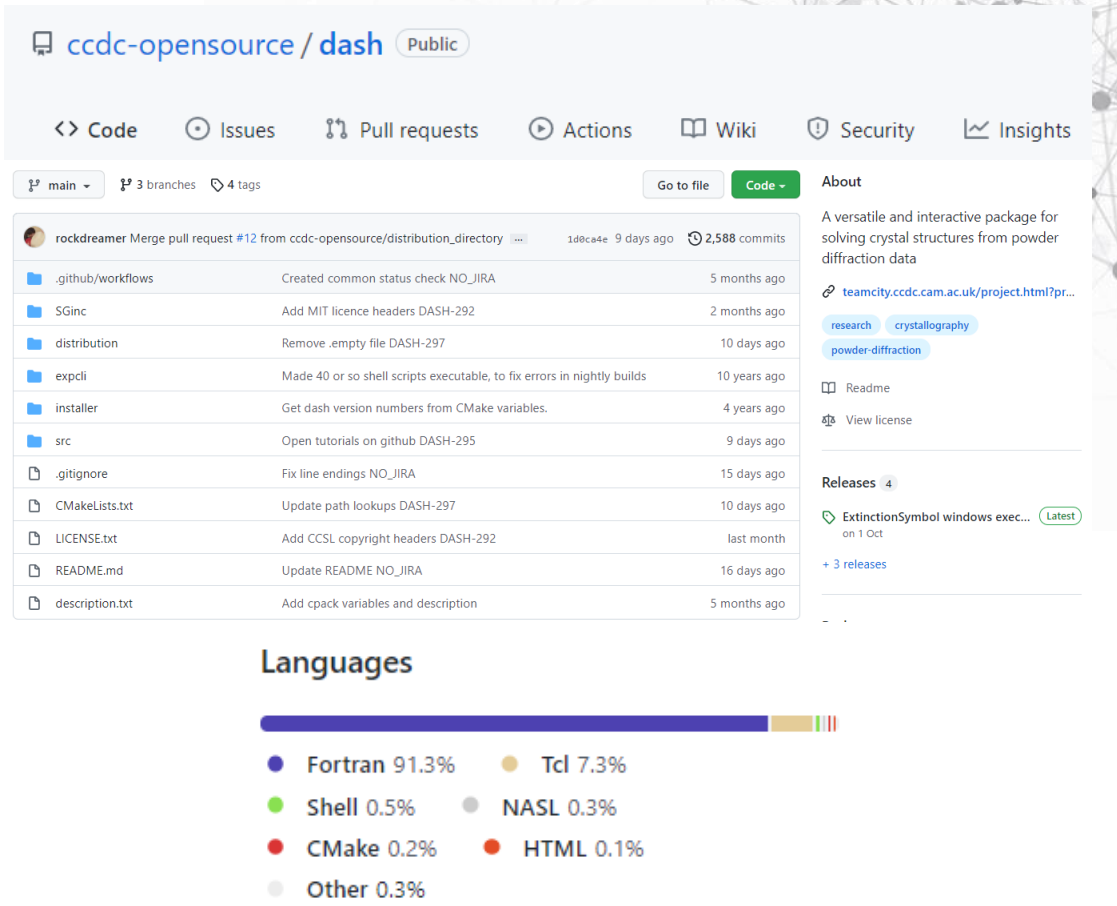
- Yes! This is a key integration that the CCDC will continue to test and maintain
- Users will need a valid licence for Mogul to make use of it
- Mogul will still have a major impact in improving the effectiveness of crystal structure solution
  - See Kabova et al., 2017



E. A. Kabova, J. C. Cole, O. Korb, A. C. Williams, K. Shankland, *J. Appl. Cryst*. (2017). 50, 1421-1427

How can the community progress DASH?

- All the code for DASH is now openly accessible under the MIT licence in GitHub
- This means that *anyone* can now work on improving or extending the DASH code
- The community can also report concerns, or improvement ideas, within GitHub
- To build DASH, you will need:
  - Intel Fortran compiler
  - CMake
  - Licence for Winteracter (the UI library)



https://github.com/ccdc-opensource/dash

CCDC

What directions might DASH take next?

- Further methods development
  - Use of GPUs for acceleration
  - Improved support for Cloud or Cluster usage
- Handling of alternative data sources
  - Neutron diffraction data
  - Electron diffraction data
- Support for alternative platforms
  - macOS version
  - Linux version
  - Command-line version

Where do *you* want to see DASH go now?

https://github.com/ccdc-opensource/dash

CCDC

CSD-CrossMiner

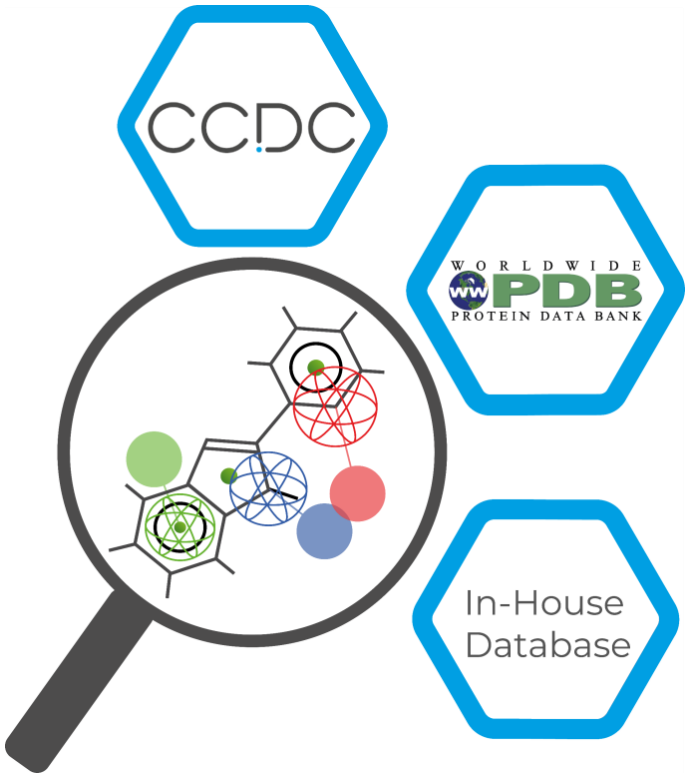
Improving user experience

Francesca Stanzione

Research and Applications Scientist

CCDC

Background

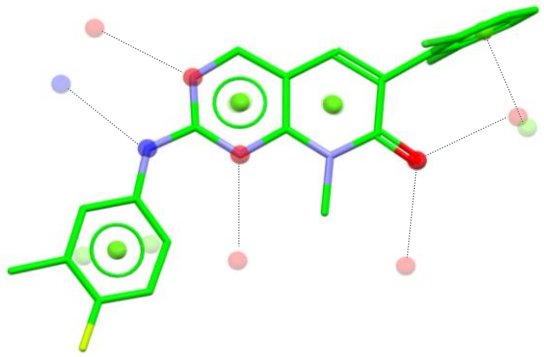


- Pharmacophore-based searches of structural databases (CSD & PDB & any in-house databases, simultaneously) for early stage drug discovery
- Modify a hypothesis/results on the fly: interactive tool!
- Annotated for easy filtering of hits

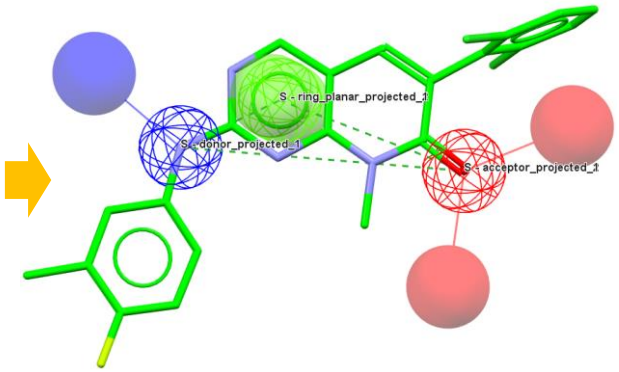
Korb *et al,. J. Med. Chem.* 2016, 59, 4257-4266. DOI: 10.1021/acs.jmedchem.5b01756

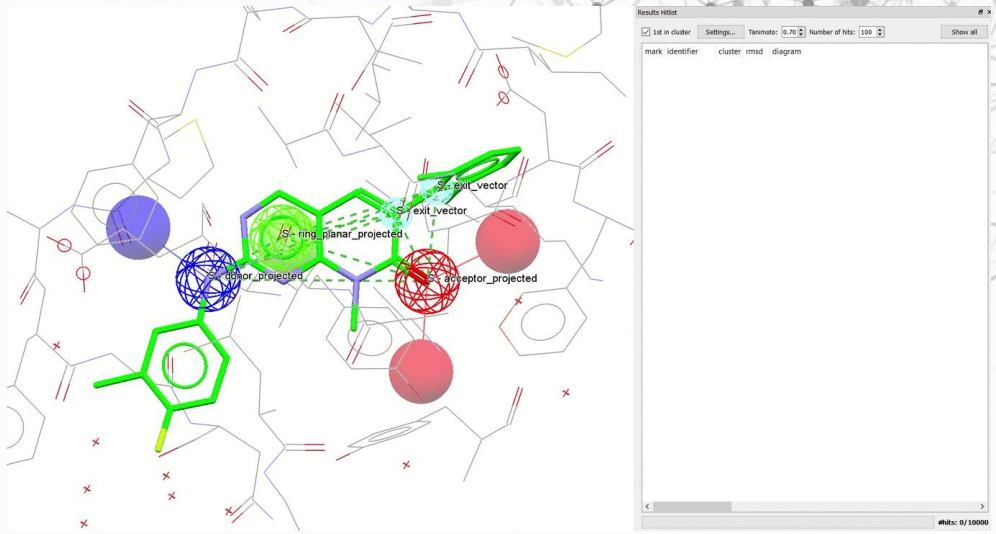
CCDC

How does it work?



donor
acceptor
planar ring
hydrophobe





Molecules annotated with customisable SMARTS feature definition

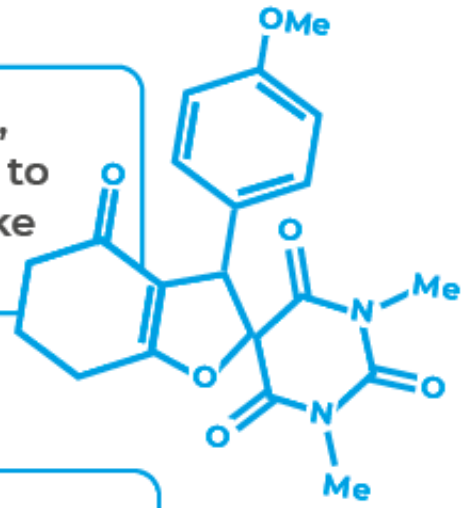
Pharmacophore points are defined using the same types of features but with a tolerance sphere

Entries matching the pharmacophore query are overlaid in the 3D view and listed with the 2D diagram showing the features matching the pharmacophore query

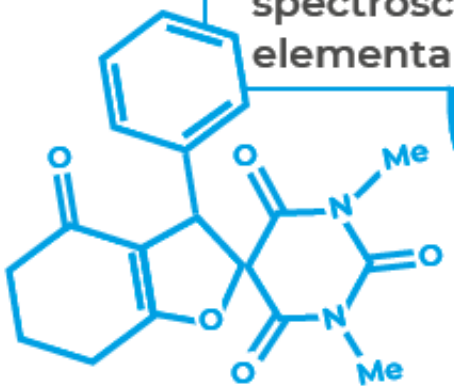
CCDC

Elinson, M.N. *et al.*
J. Heterocycl. Chem., 2021, 1–12.
doi.org/10.1002/jhet.4274

Multicomponent, one-pot reaction to generate drug-like scaffold



Confirm structure with NMR, IR and Mass spectroscopy and elemental analysis



Scaffold-based search using CSD-CrossMiner of about 300K potential target proteins

Identify a protein (aldose reductase) in complex with a molecule of similar size with similar features

Note relevance of H-bonds and π-π interactions for designing future scaffolds

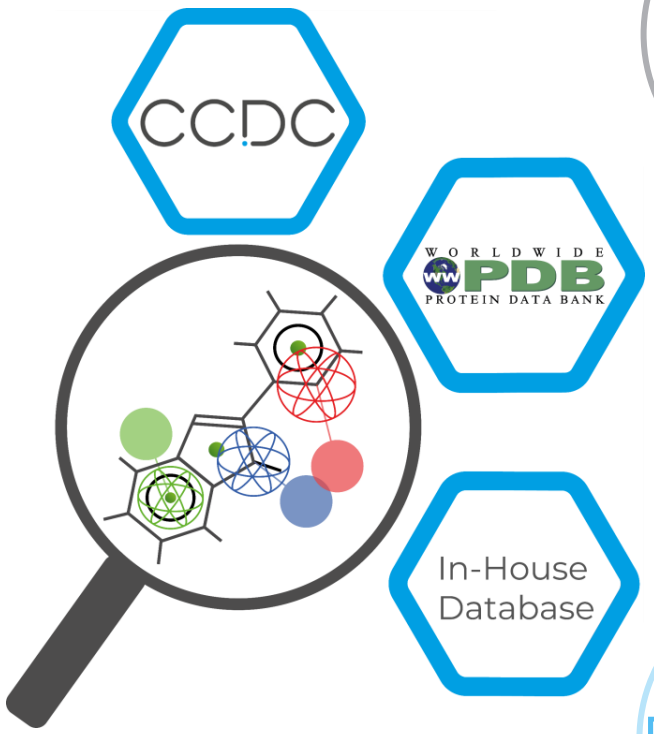
Further investigate use as aldose-reductase inhibitors via docking studies

CCDC
advancing structural science

Why improving CSD-CrossMiner?

- From user feedback:
  - Difficult to distinguish between protein and ligand
  - Do I have access to CSD-CrossMiner?
  - Better integration with other tools
  - I don’t really need organo-metallic structures
  - Flexibility in the search – OR/NOT Boolean operators
  - Don’t let me think about SMARTS
  - Reusable sessions
  - More available databases
  - …

What?



Visibility

- One installer for CSD-Discovery suite
- Integration with Hermes (protein-ligand visualiser)

Usability

- Style & colour of different components
- Style & colour reference and hits

Flexibility

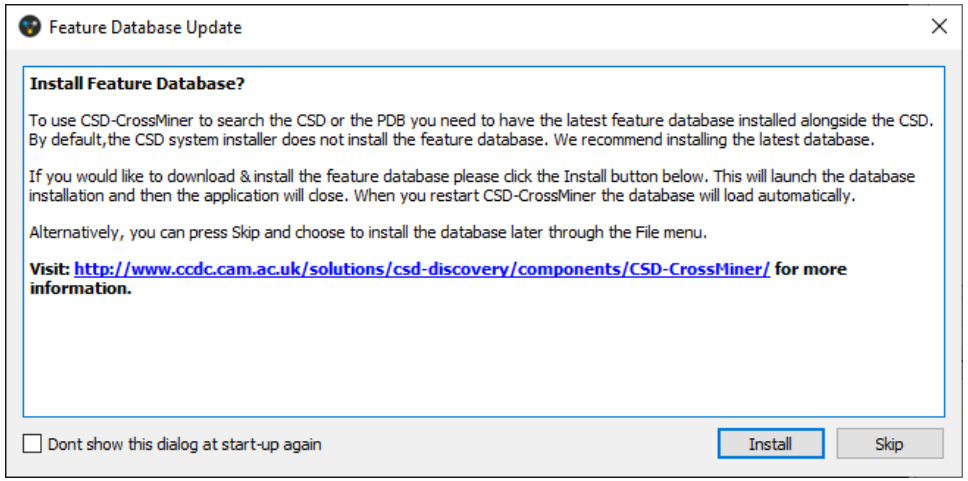
- Customisable excluded volume
- Exclude organometallic structures

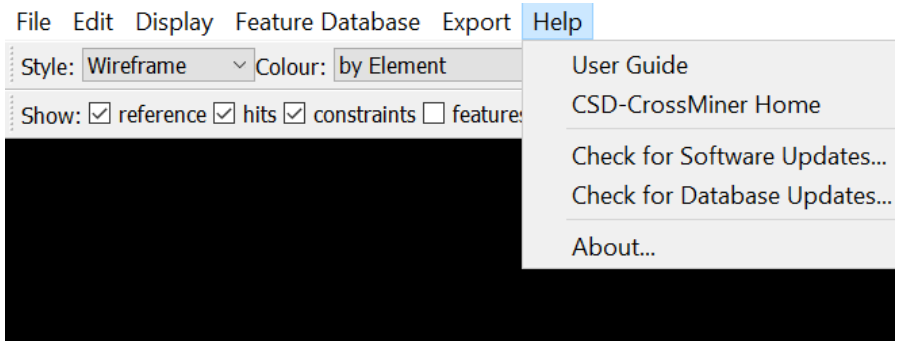
https://www.ccdc.cam.ac.uk/solutions/whats-new/

CCDC

Visibility - Include CSD-CrossMiner in the CSDS installer

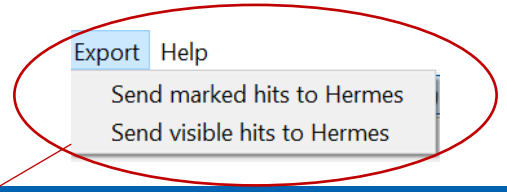
- Include CSD-CrossMiner software in the CSDS installer:
  - CSD-CrossMiner software added to the Discovery component of the CSD software portfolio
  - CSDS software update will also update CSD-CrossMiner software
  - When opening CSD-CrossMiner a pop-up window will guide you to download the feature database
  - Built-in auto-update mechanism that makes updating both the software and the database easy

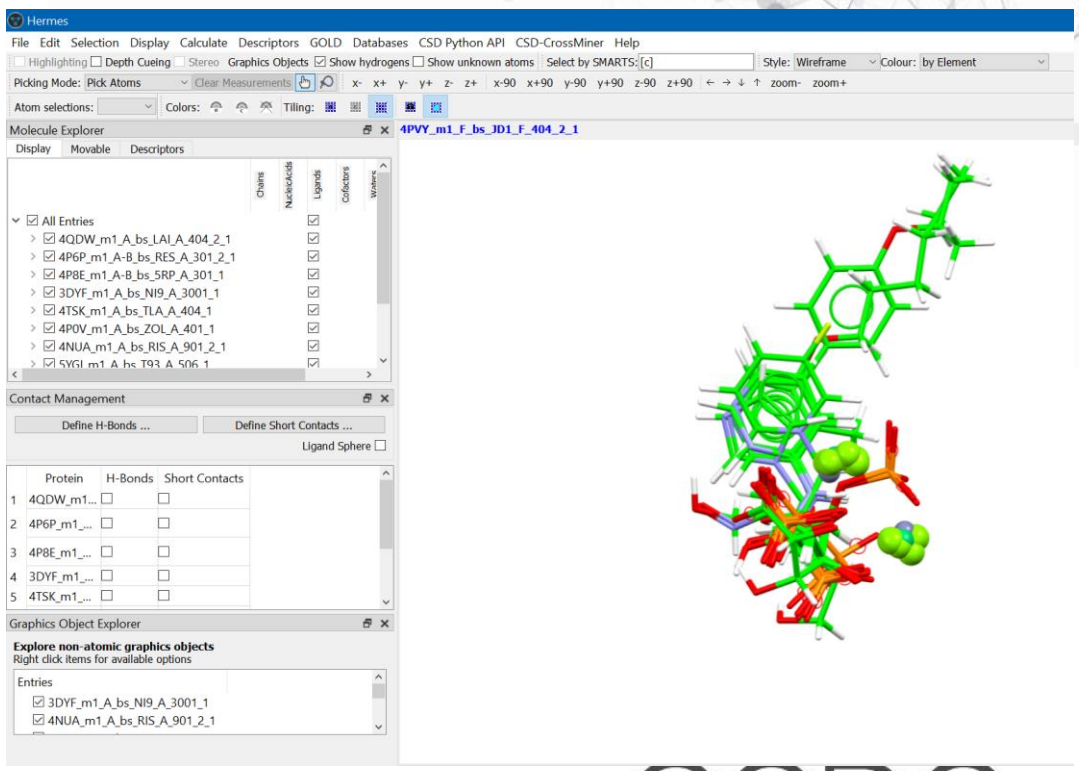




Visibility - Integration with Hermes

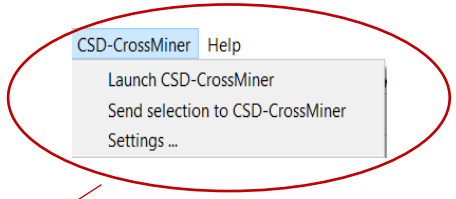
- Export hits to Hermes

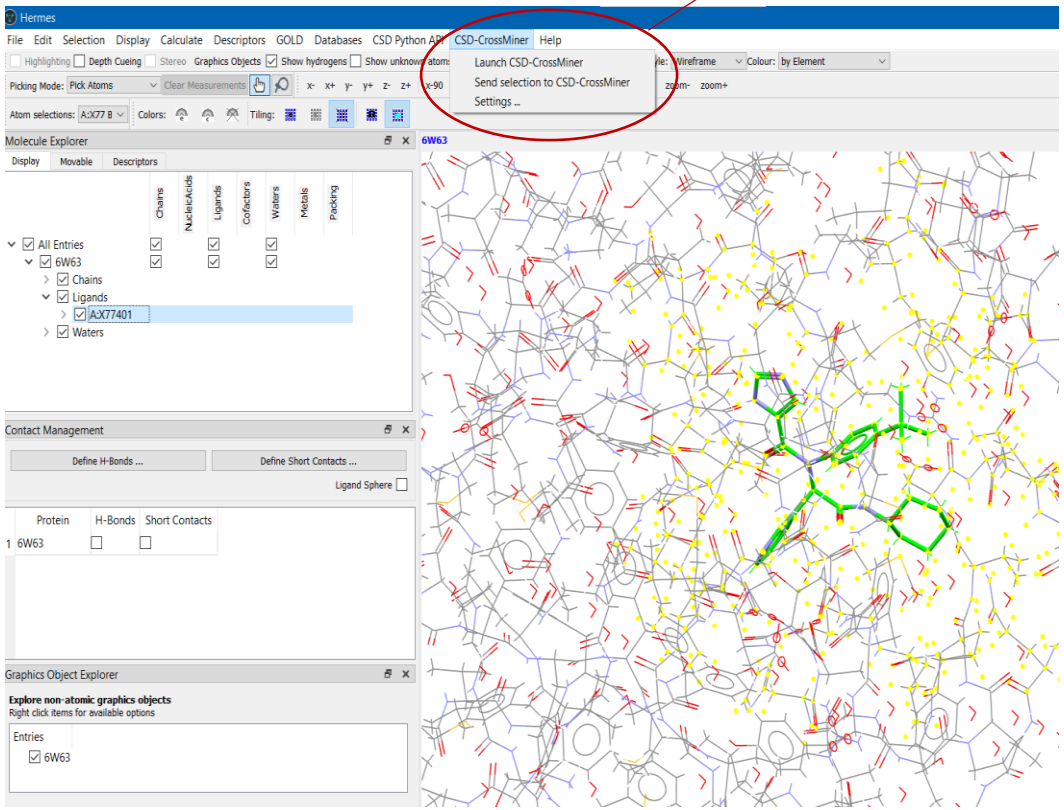


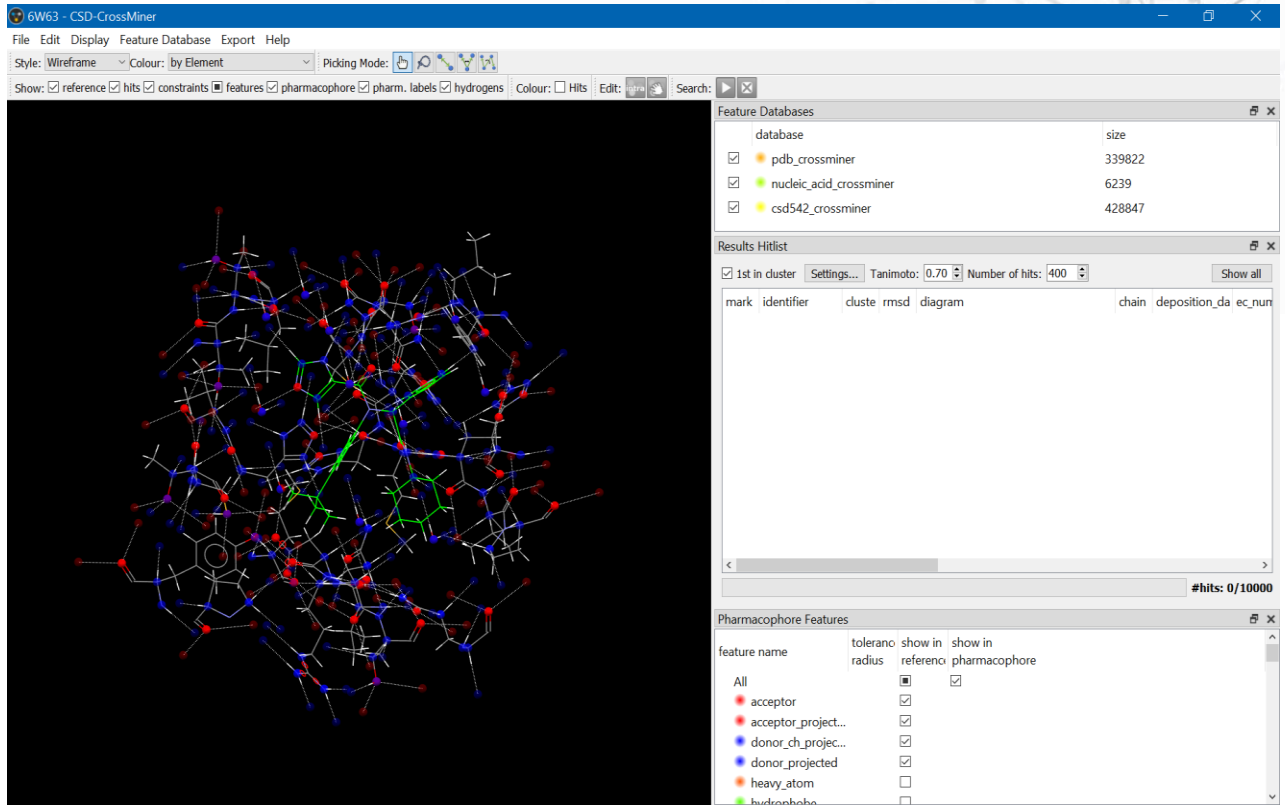


Visibility - Integration with Hermes

- Launching CSD-CrossMiner from Hermes

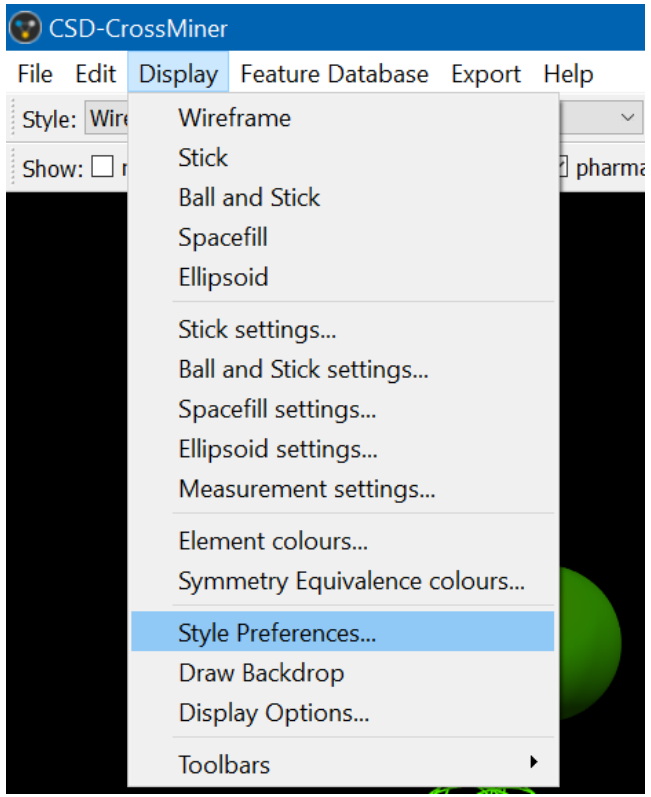


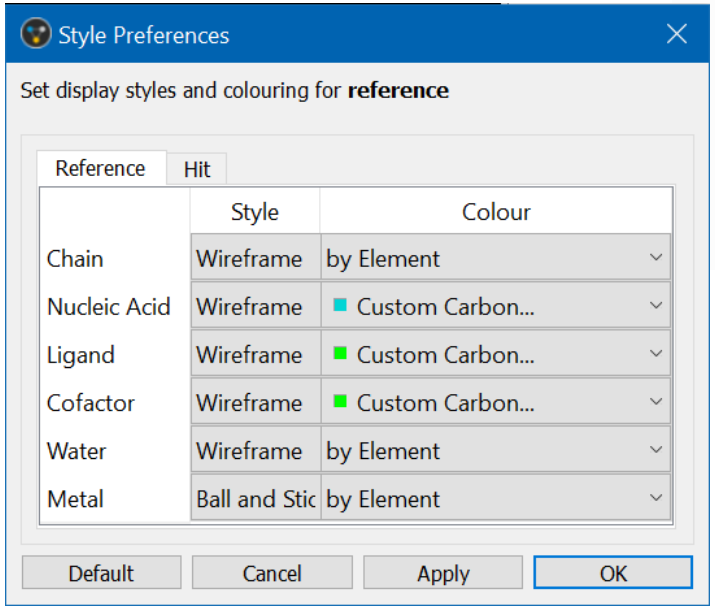


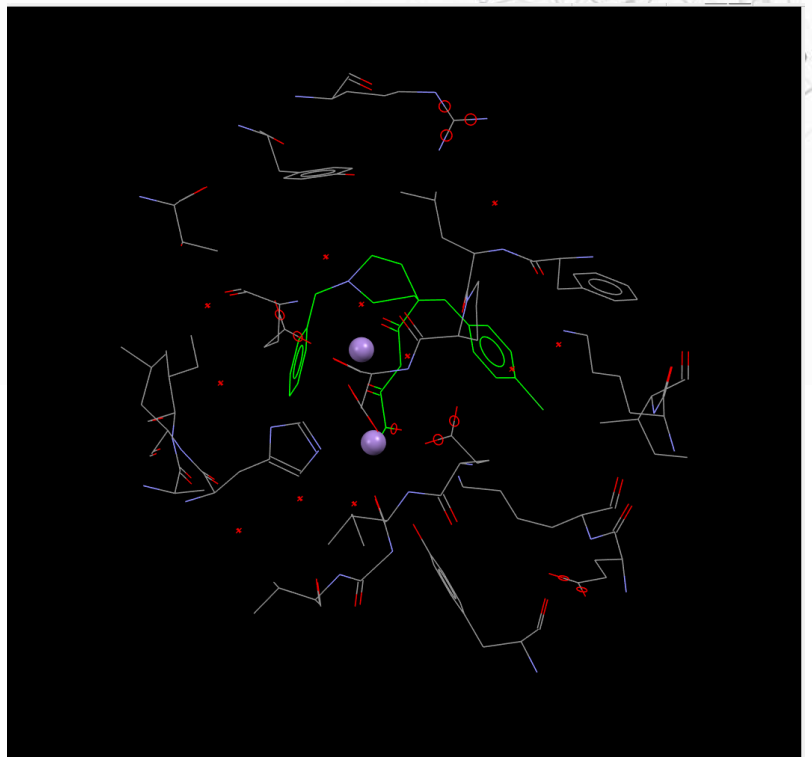


Usability – Style/colour of different components

- Style Preferences... menu to style and colour different components of both reference molecule and hits

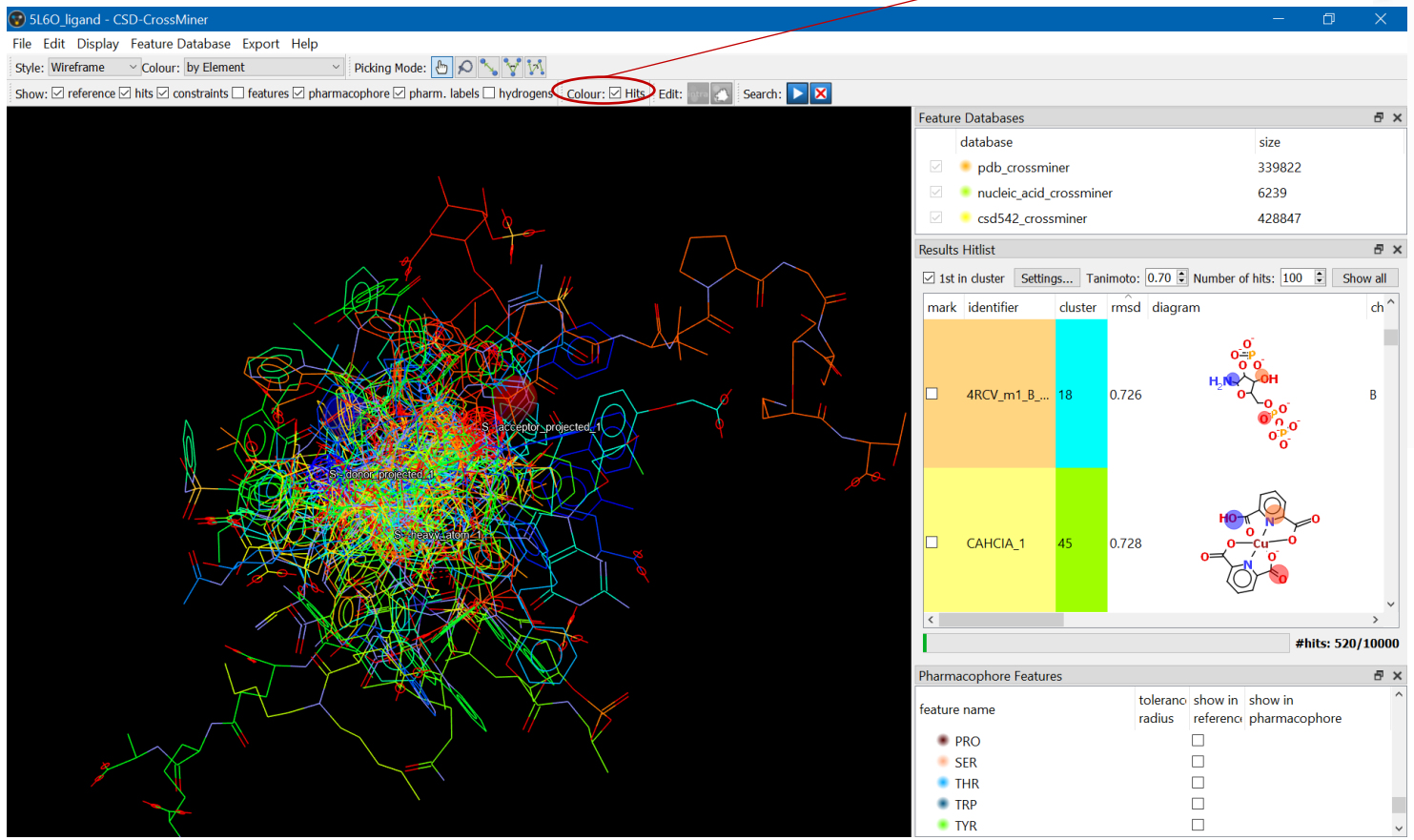






Usability – Colouring hits

- Colour hits by rainbow

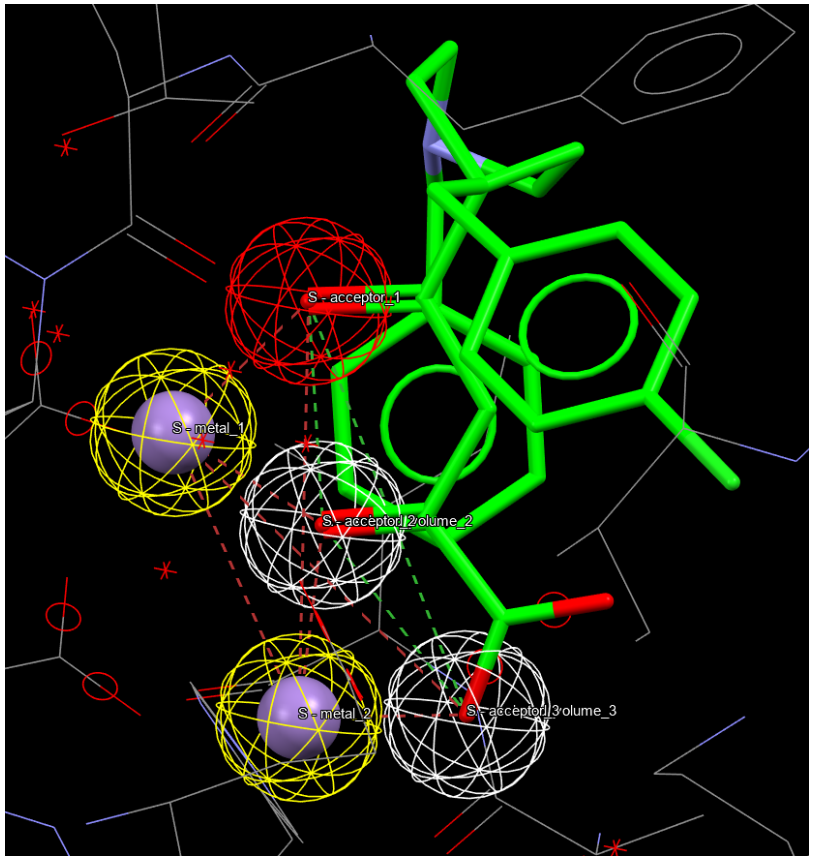


- Colouring is applied when the pharmacophore search is paused or completed
- Colouring is applied to the cluster
- Hits of the same cluster have the same colour

CCDC

Flexibility – Customisable excluded volume

- Customisable exclude volume (NOT feature) – Exclusion sphere(s) such that only defined moieties are excluded from a region.



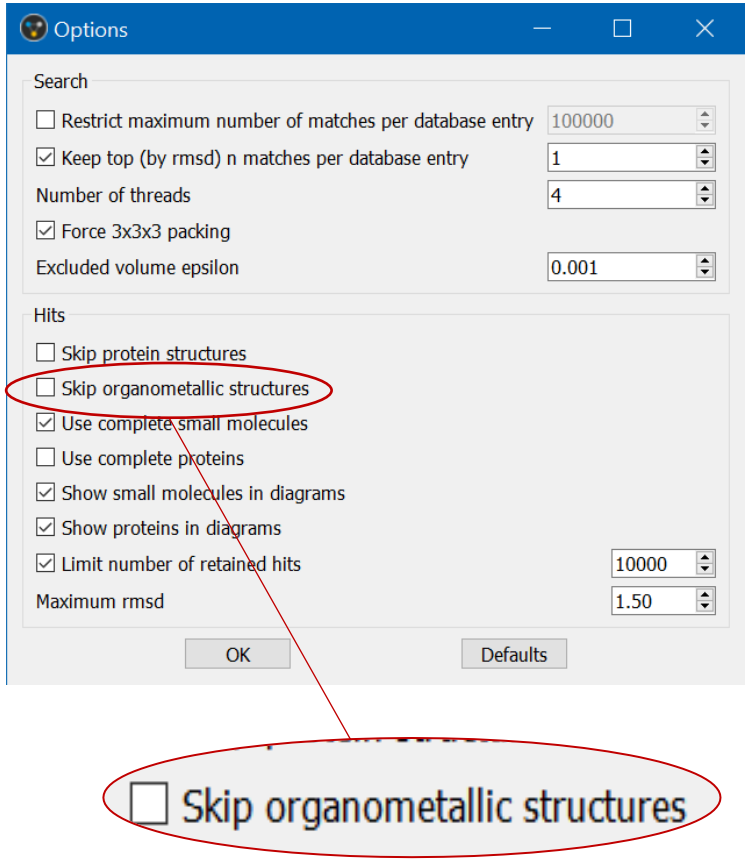
- Moieties are specified using SMARTS patterns.

Pharmacophore Features

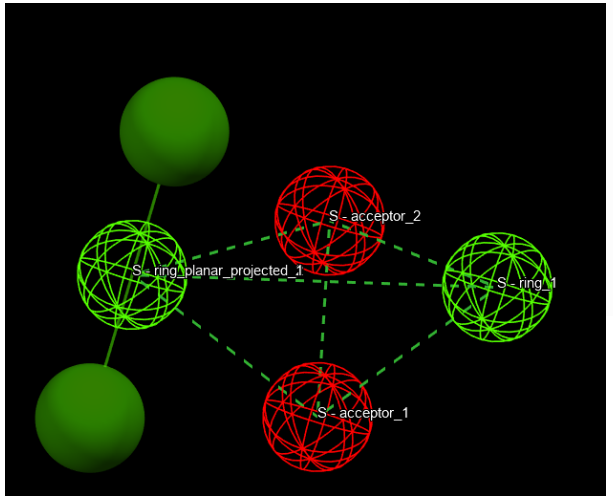
feature name	tolerance radius	show in referenc	show in pharmacophore
B	1.00		
Excluding atoms (SMARTS)	[#15]~[#8]		
excluded_volume_3			☑
B	1.00		
Excluding atoms (SMARTS)	[#15]~[#8]		
annotation_filter		☐	
substructure_filter		☐	

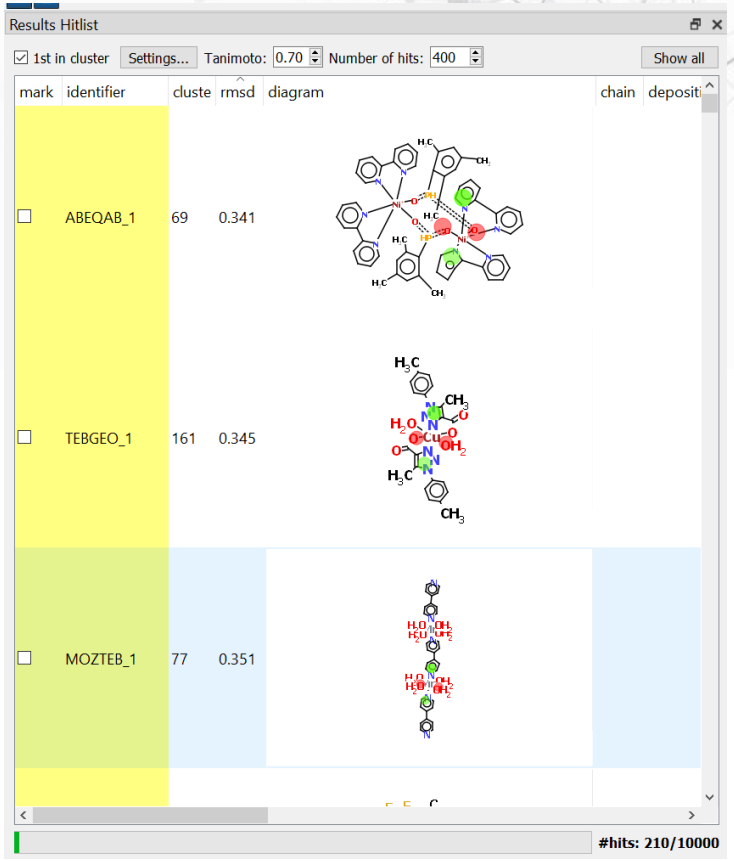
Flexibility – Excluding organometallic structures

- Provided the option to exclude organometallic structures from the search



- Hits containing at least one transition metal, lanthanide, actinide, or any Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po will be excluded from the search





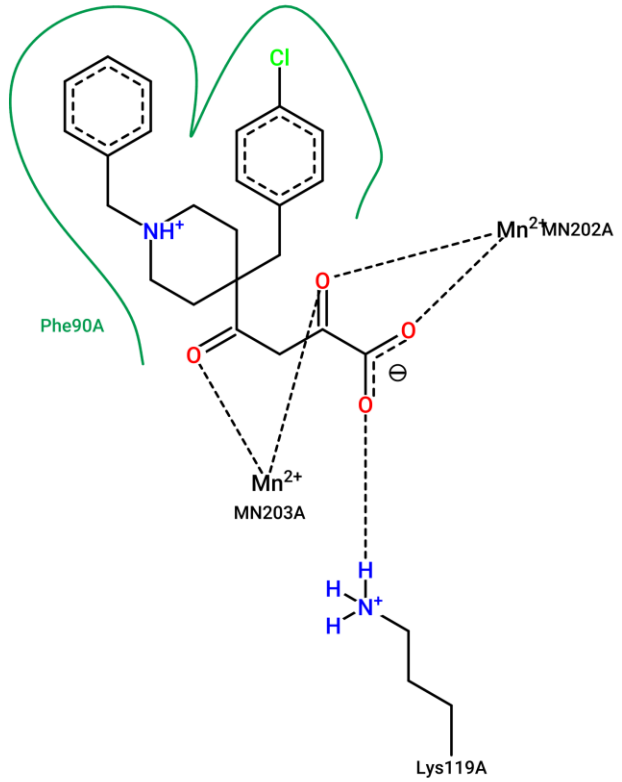
Additional new features

- Easy way to select a reference molecule in CSD-CrossMiner
- Optimised the Style/Colour style behaviour
- Improved the generation of the feature database
- Easy way to select protein-ligand binding site in Hermes

CCDC

Demo

- Performing a pharmacophore search using CSD-CrossMiner



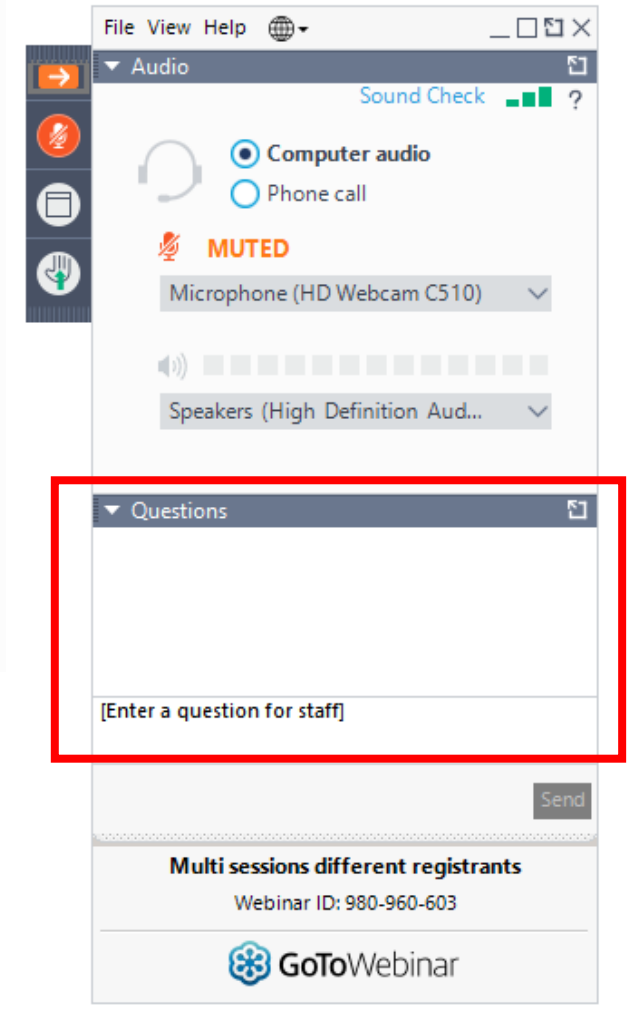
PDB code 5FDG

- Influenza polymerase acidic protein N-terminal domain (PAN) in complex with high potency endonuclease inhibitor (0N8)
- 0N8 chelate to the metal ions by its metal-chelation motifs
  - What other metal-chelation motifs are in the PDB?
  - What about in the organic subset of the CSD?

CCDC

Q&A

- Type your questions in the box as shown



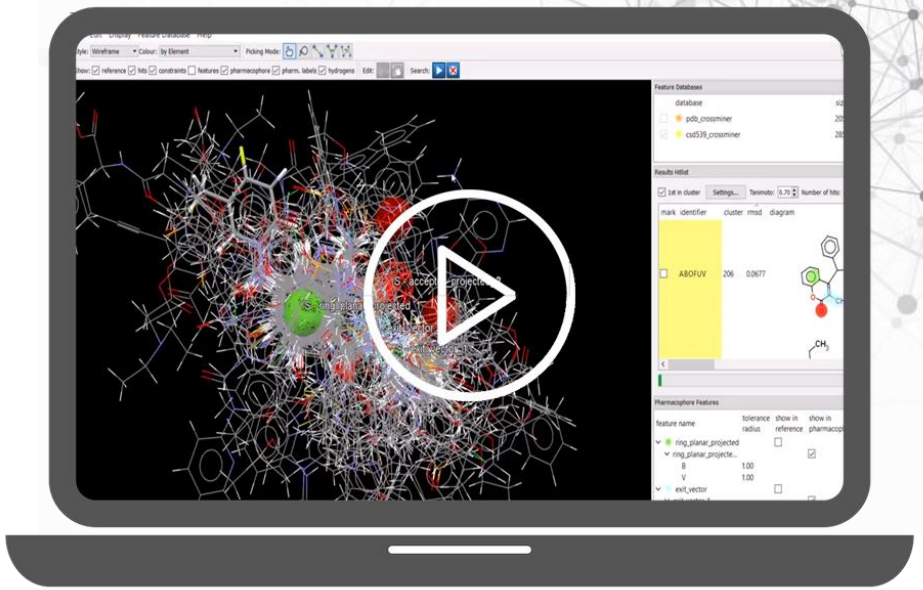
CCDC

Next What's Up Webinar

- Next webinar: **20th January 2022**

- Follow us on social media
- Send us your ideas and news

  hello@ccdc.cam.ac.uk



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