Mining Crystallographic Databases using CSD-CrossMiner

CCDC Virtual Workshop Spring 2021 – Session 2

Francesca Stanzione, Abhik Mukhopadhyay, Ilenia Giangreco, Suzanna Ward, Ilaria Gimondi, Eva Myers

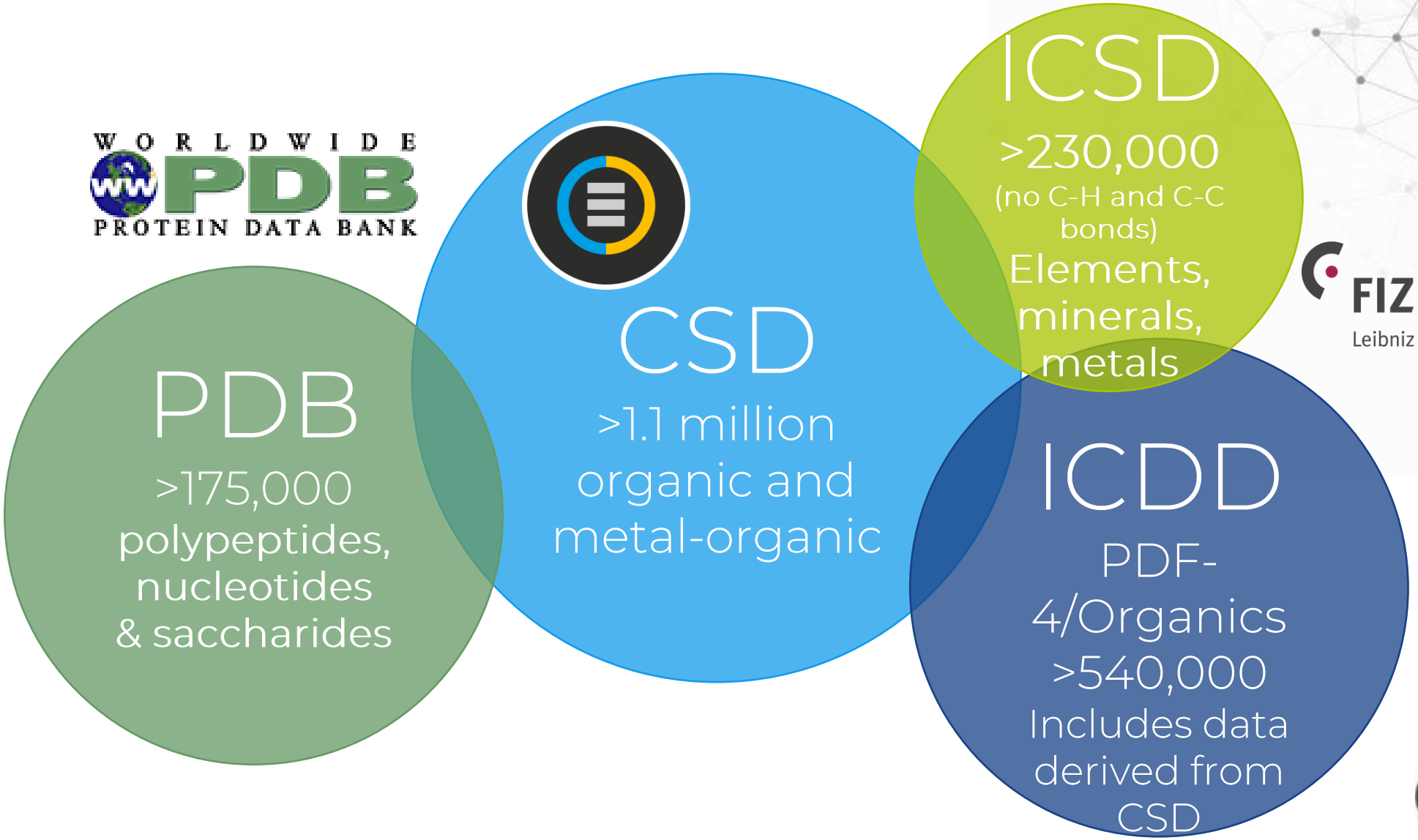
14 April 2021

CCDC
advancing structural science

Learning outcomes for today

- Familiarising yourself with CSD-CrossMiner.
- Learning how to interact with a pharmacophore query.
- Learning how to perform pharmacophore searches.
- Learning how to analyse and interact with your results.

Structural databases

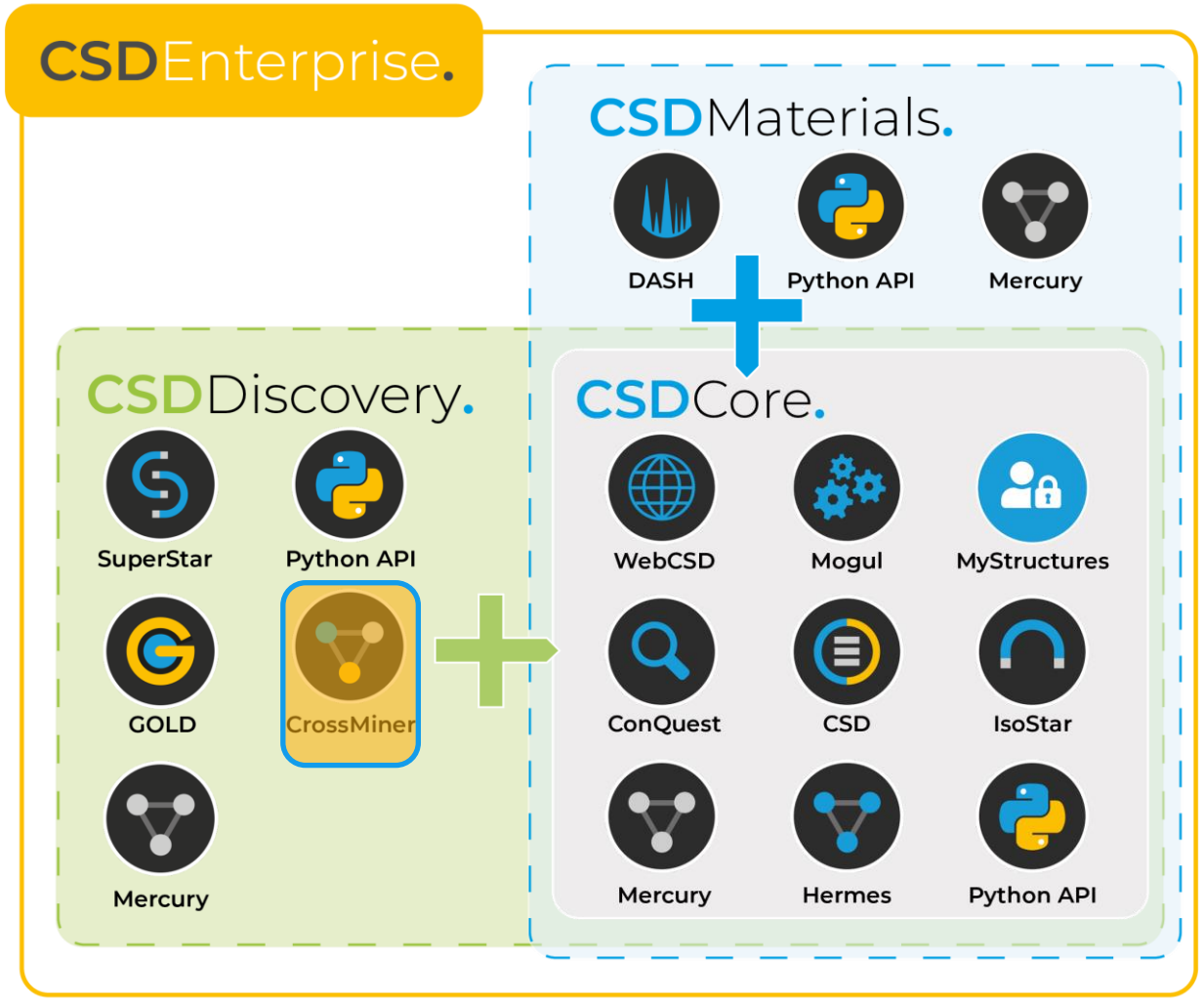


FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

ICDD

CCDC

The CSD Portfolio



Professional Services

Research & Knowledge partnerships

CCDC

CSD-CrossMiner

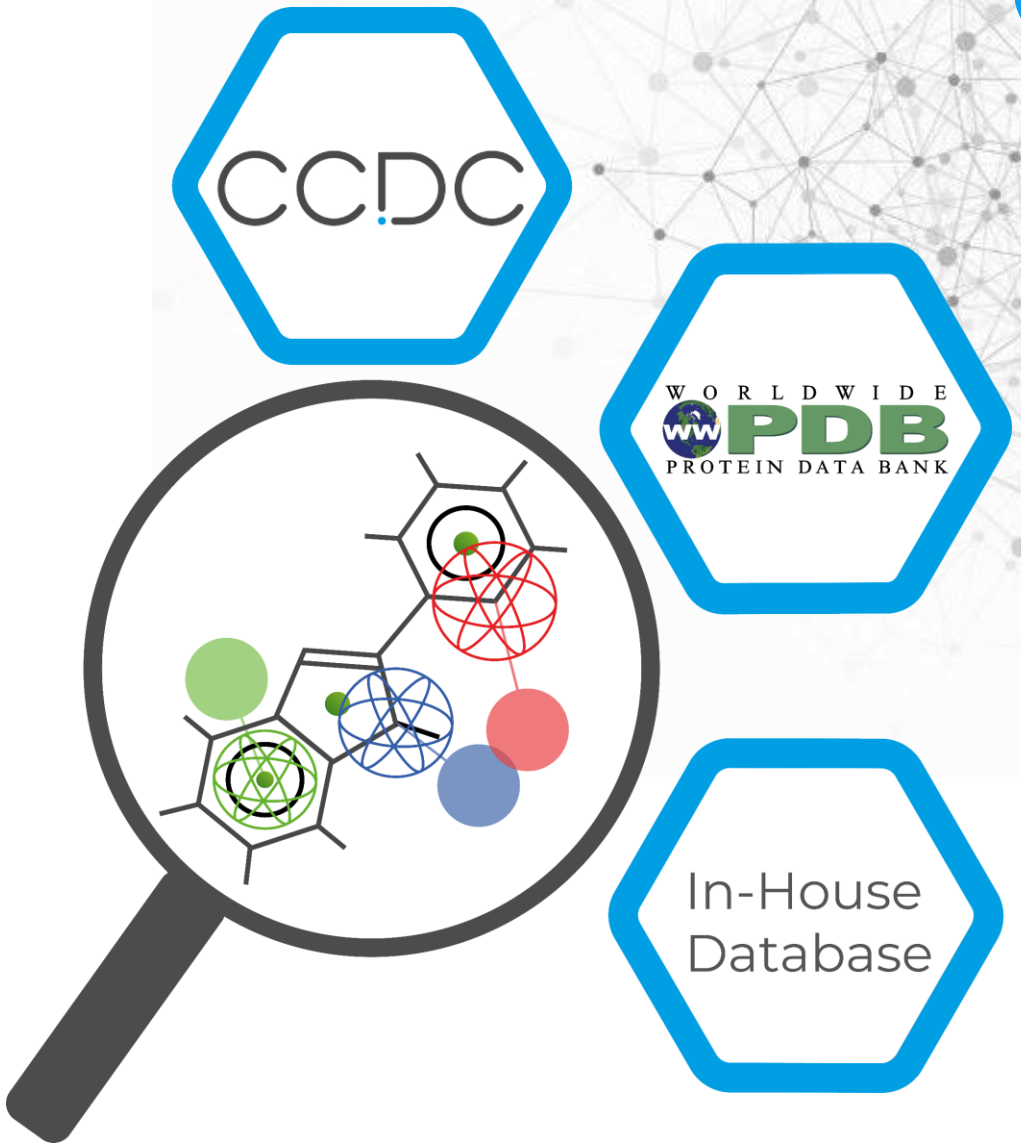
- Tool for searching structural databases by pharmacophore.
  - From IUPAC: “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions
  - Wikipedia: “abstract description of molecular features”

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

CCDC

CSD-CrossMiner

- Simultaneously search the PDB, CSD and your in-house database.
- Designed for speed – modify hypotheses on the fly.
- Structures are annotated for easy filtering of hits



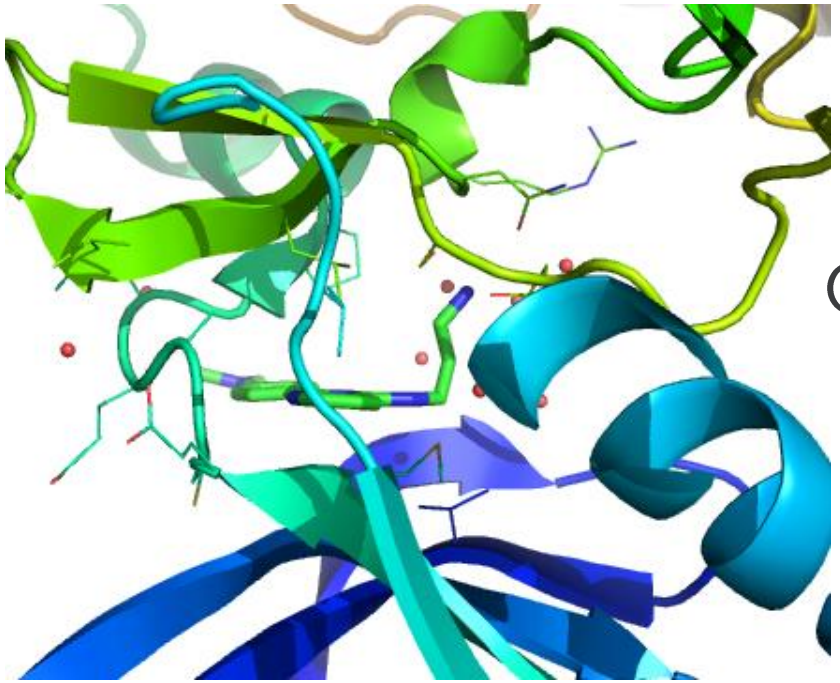
Korb et al,. J. Med. Chem. 2016, 59, 4257-4266. DOI: 10.1021/acs.jmedchem.5b01756

CCDC

What can you do with CSD-CrossMiner?

CCDC
advancing structural science

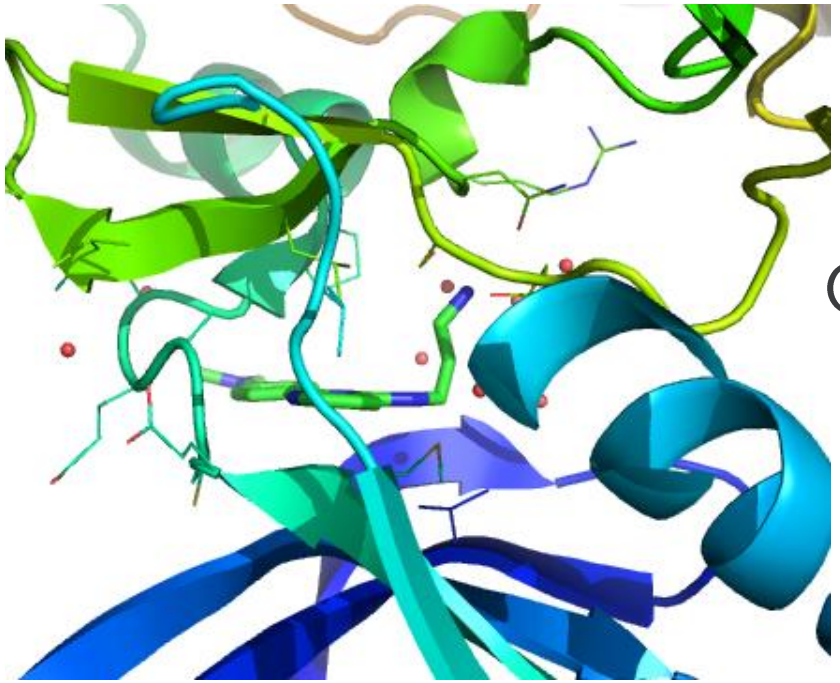
Understand protein-ligand interactions.



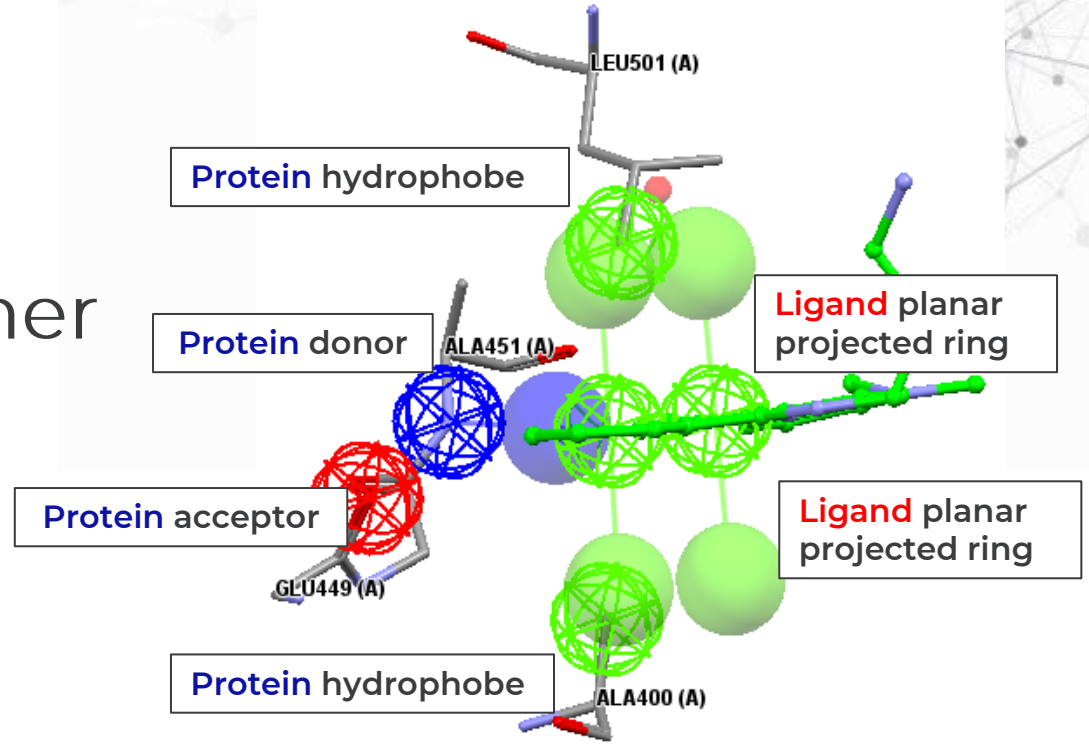
CSD-CrossMiner

CCDC

Understand protein-ligand interactions.

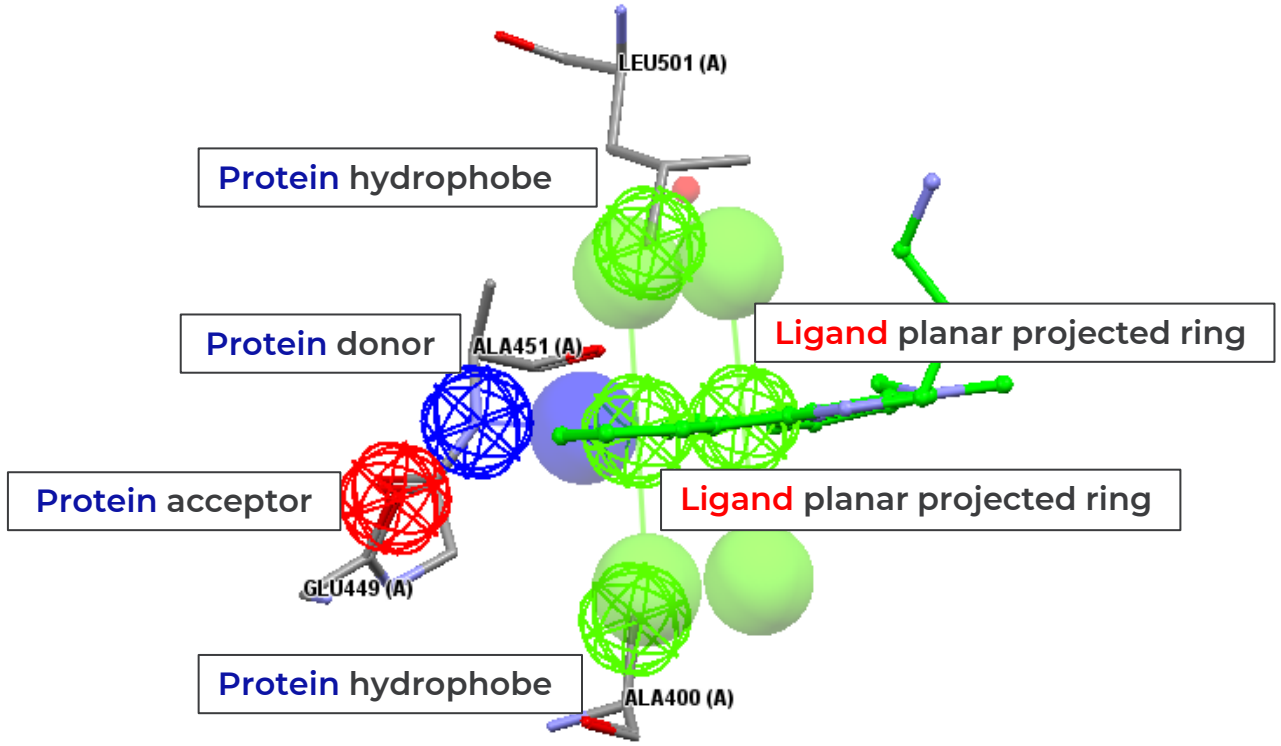


CSD-CrossMiner



CCDC

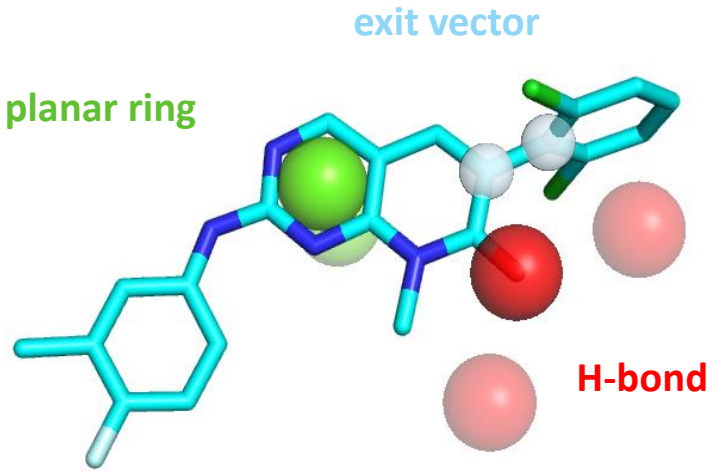
Understand protein-ligand interactions.



- Determine common protein binding sites in PDB structures
- Determine structural motifs that bind in similar environments
- Shed light into cross-pharmacology between protein targets

CCDC

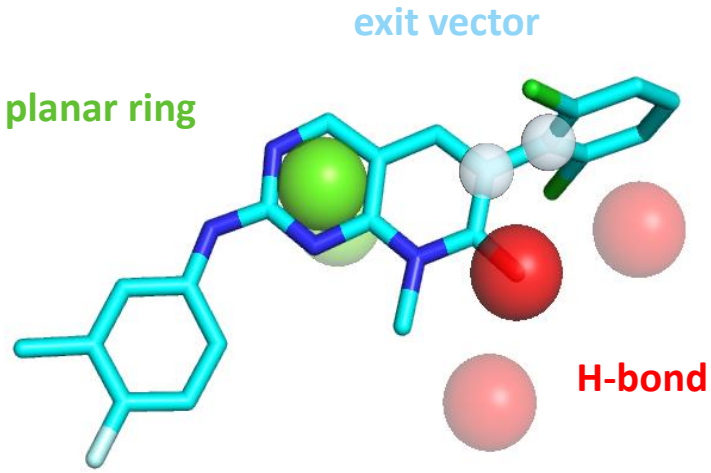
Generate new ideas.



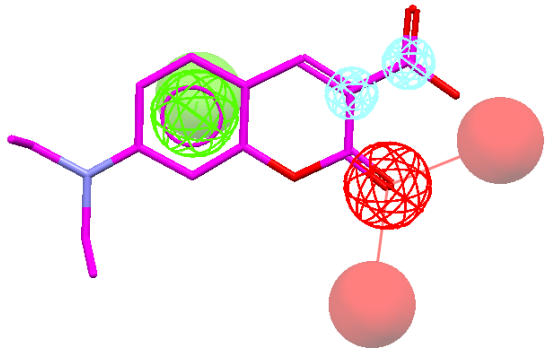
CSD-CrossMiner

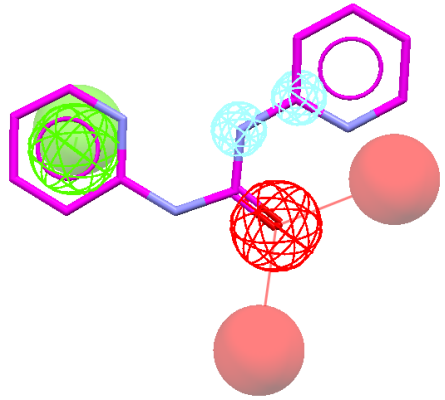
CCDC

Generate new ideas.



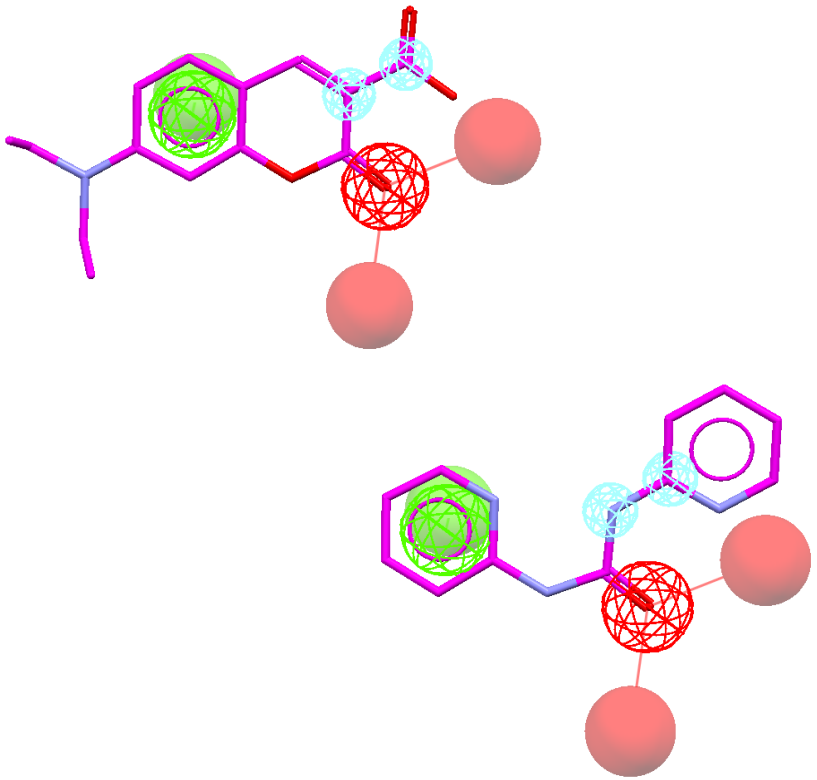
CSD-CrossMiner





CCDC

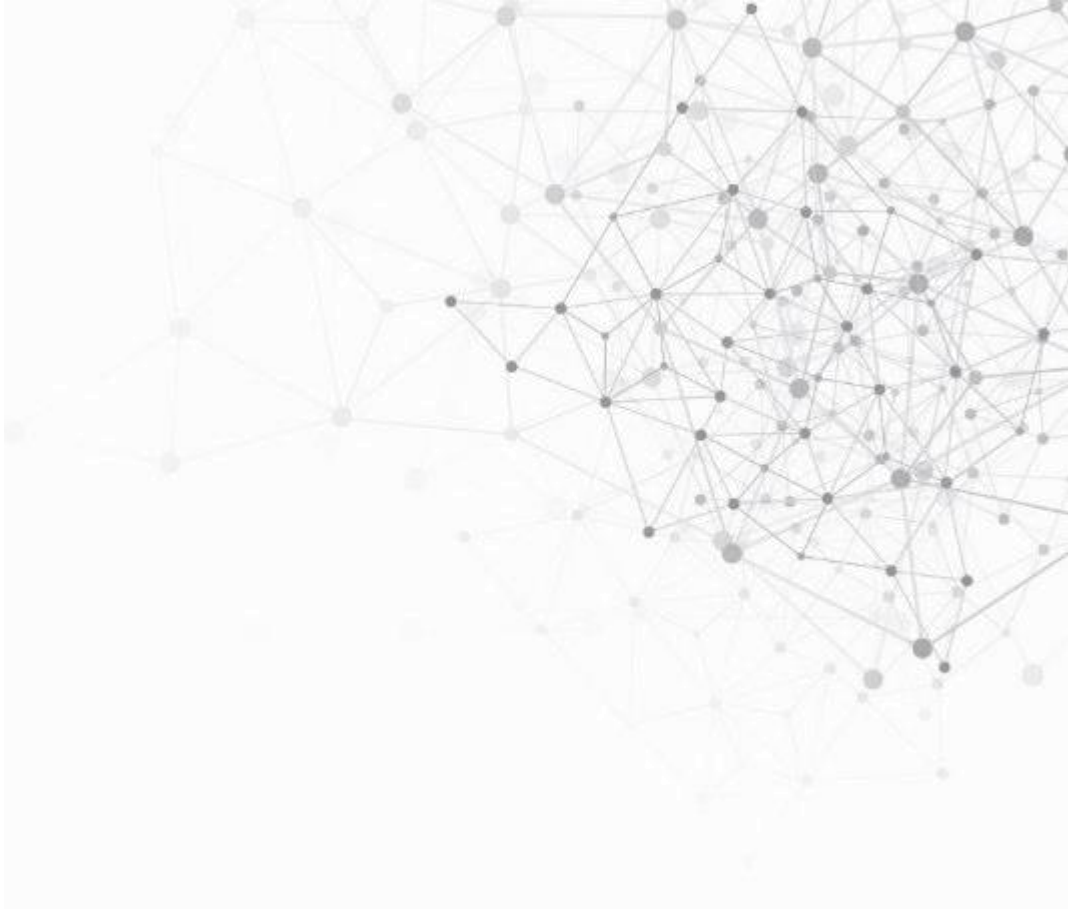
Generate new ideas.



- Design novel motifs that mimic extablished ligands
- Scaffold-hopping: retrieve a diversity of ligand topologies that can be used as scaffolds
- Look at chemistries that interact with unexplored parts of the binding site.

CCDC

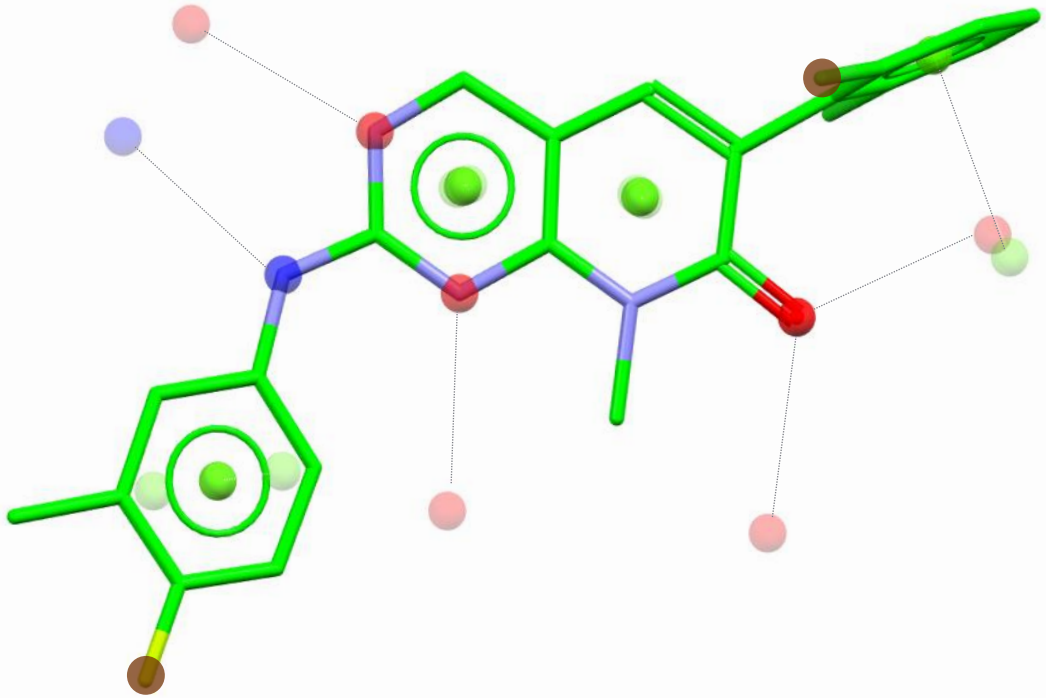
How does it work?



CCDC
advancing structural science

Look at the features.

Donor Planar ring
Acceptor Halogen

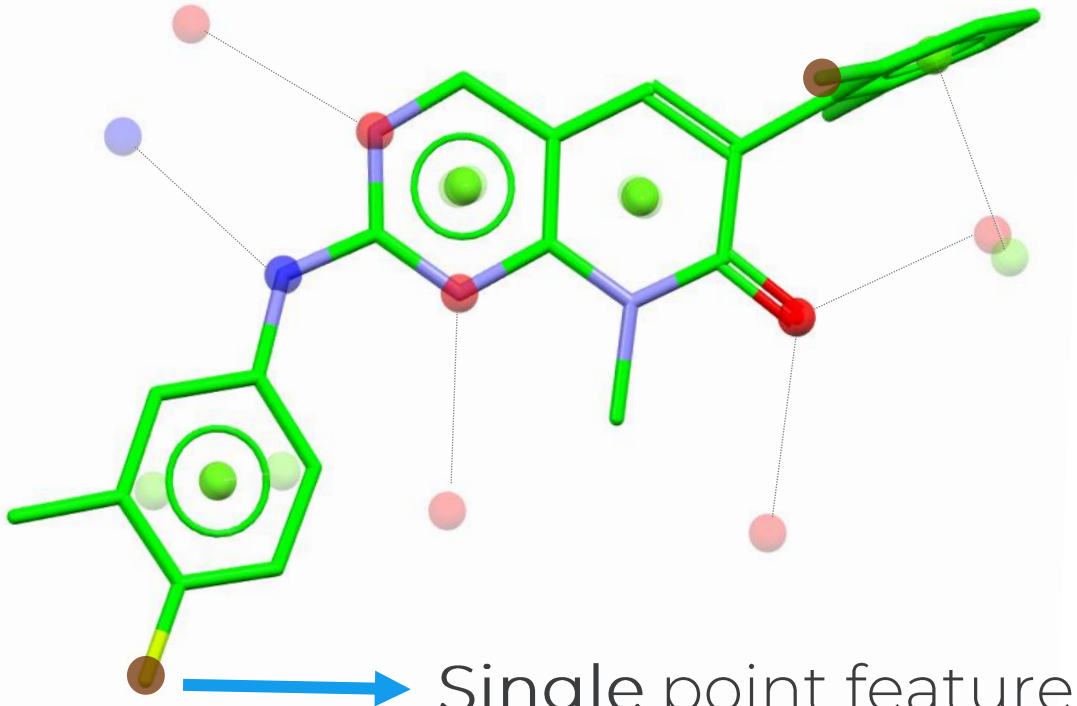


- Molecules annotated with steric and electronic features.
- Based on SMARTS patterns – user editable.

CCDC

Look at the features.

Donor Planar ring
Acceptor Halogen



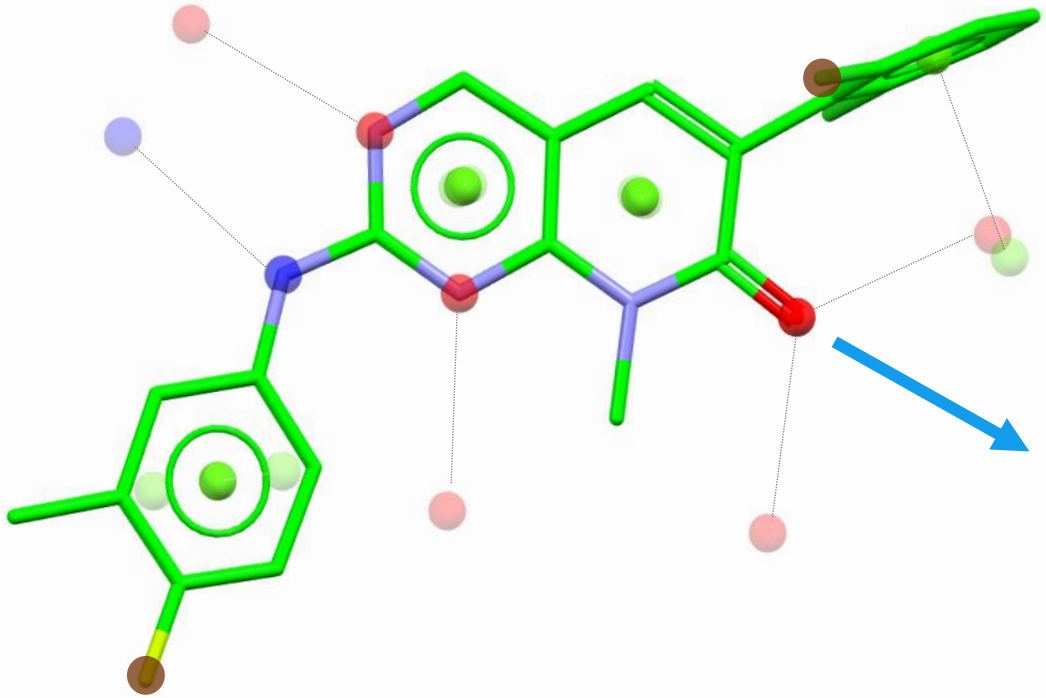
Single point feature

- Molecules annotated with steric and electronic features.
- Based on SMARTS patterns – user editable.

CCDC

Look at the features.

Donor
Acceptor
Planar ring
Halogen



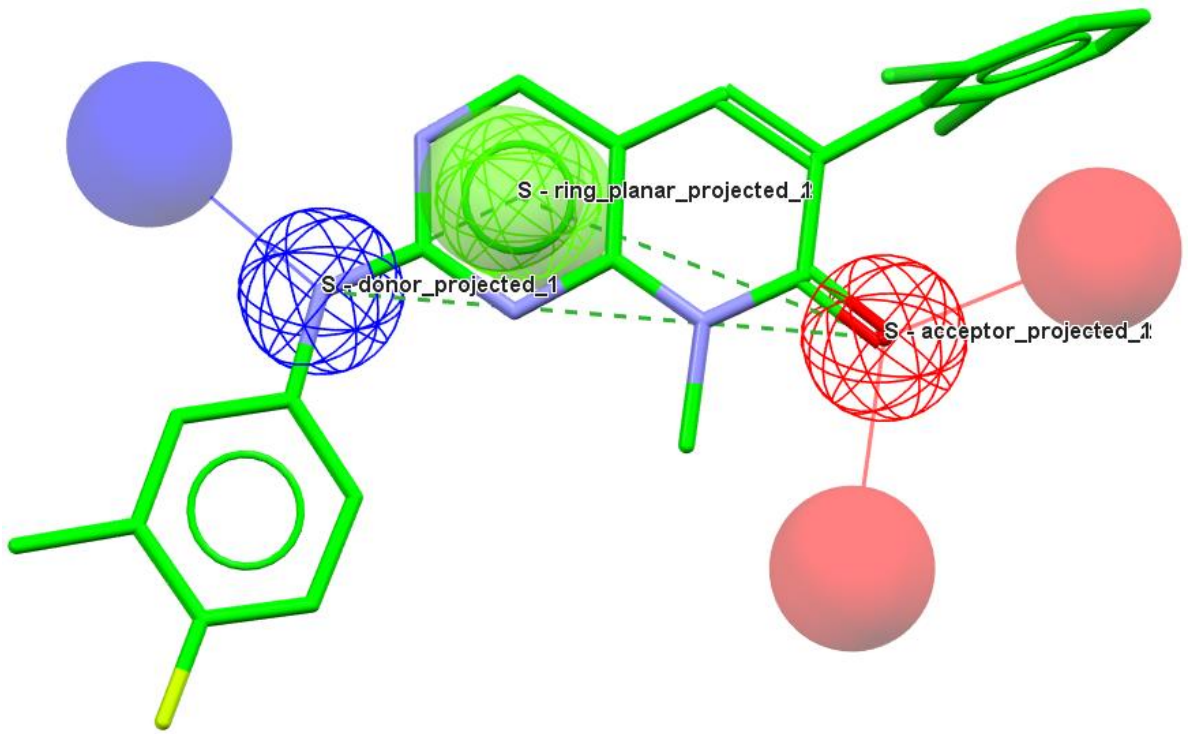
- Molecules annotated with steric and electronic features.
- Based on SMARTS patterns – user editable.

Directional two point feature

CCDC

Build the pharmacophore query.

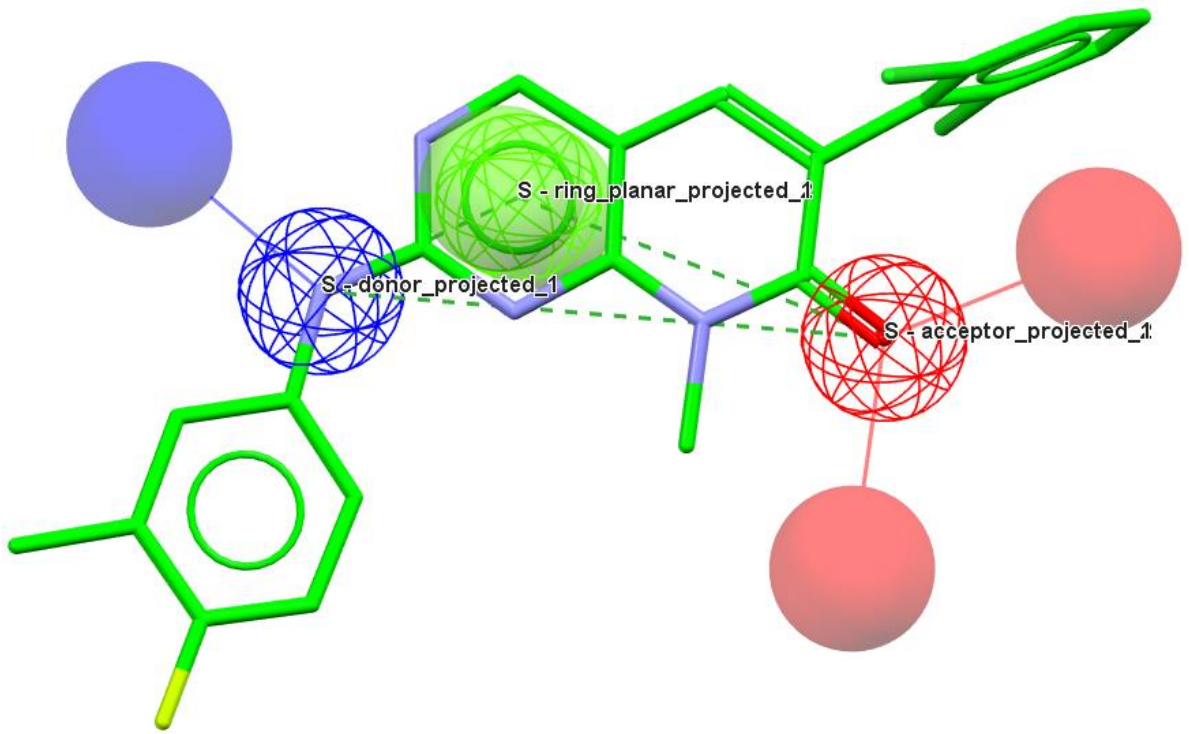
Donor Planar ring
Acceptor



CCDC

Build the pharmacophore query.

Donor
Acceptor
Planar ring



- Tolerance sphere around each feature.
- Radius reflects the uncertainty of position.

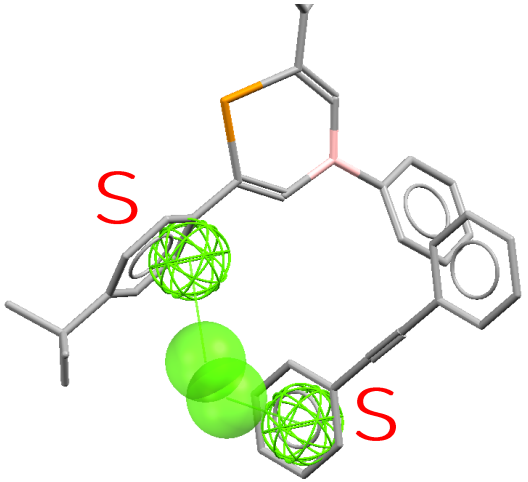
CCDC

Build the pharmacophore query.

- Customise the query – define the nature of the point.

Build the pharmacophore query.

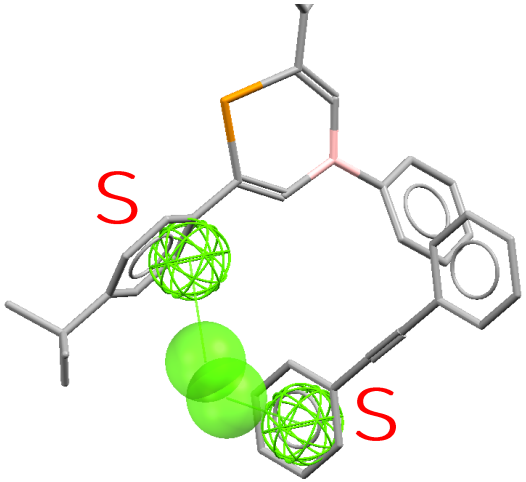
- Customise the query – define the nature of the point.



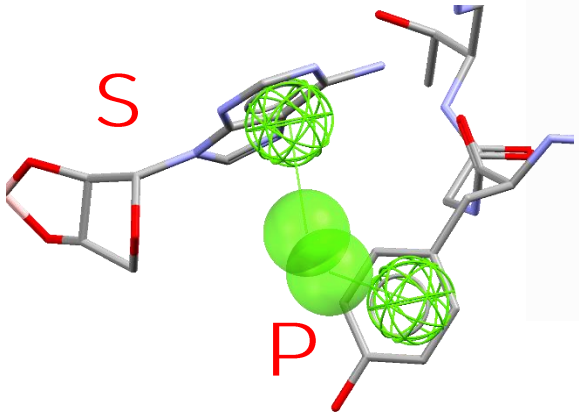
Small molecule (**S**)

Build the pharmacophore query.

- Customise the query – define the nature of the point.



Small molecule (**S**)

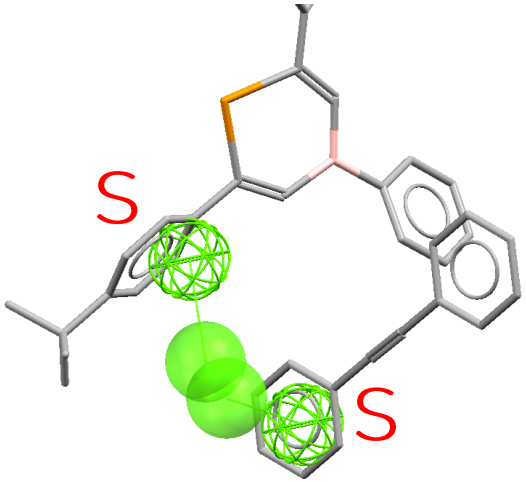


Protein (**P**)

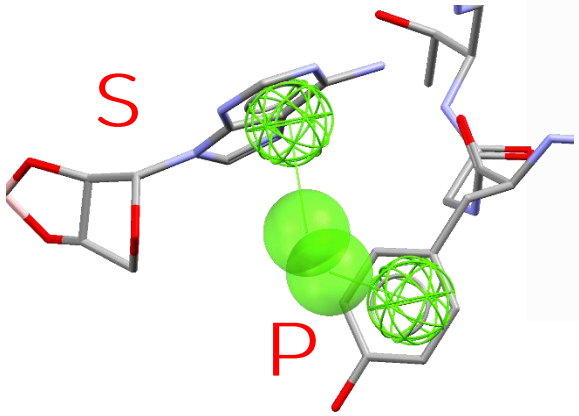
CCDC

Build the pharmacophore query.

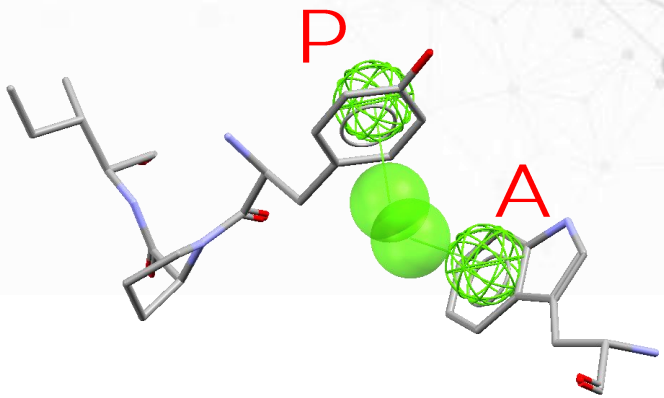
- Customise the query – define the nature of the point.



Small molecule (**S**)



Protein (**P**)

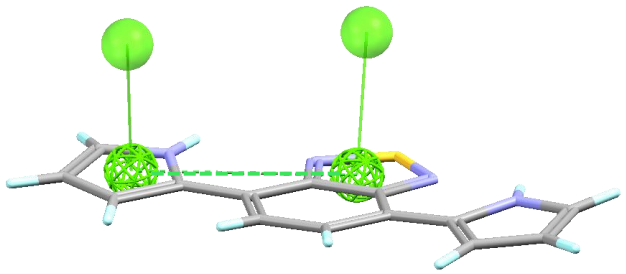


Any molecule (**A**)

CCDC

Build the pharmacophore query.

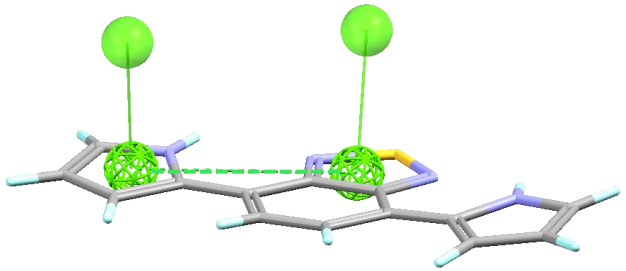
- Customise the query – define the nature of the point.



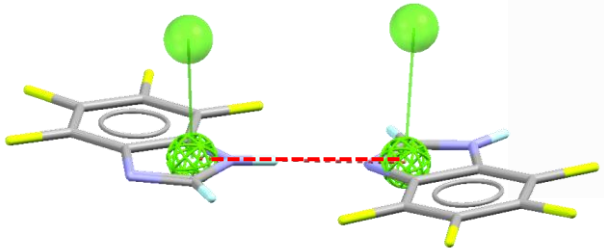
Intra-molecular constraint

Build the pharmacophore query.

- Customise the query – define the nature of the point.



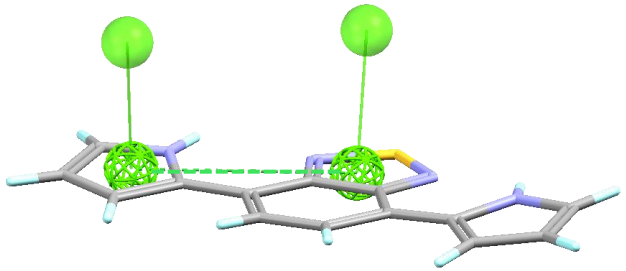
Intra-molecular constraint



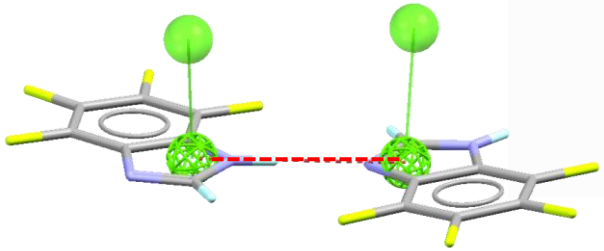
Inter-molecular constraint

Build the pharmacophore query.

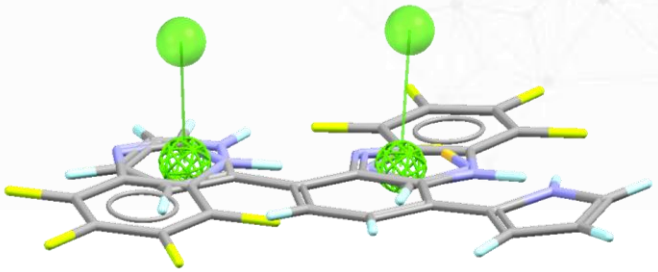
- Customise the query – define the nature of the point.



Intra-molecular constraint



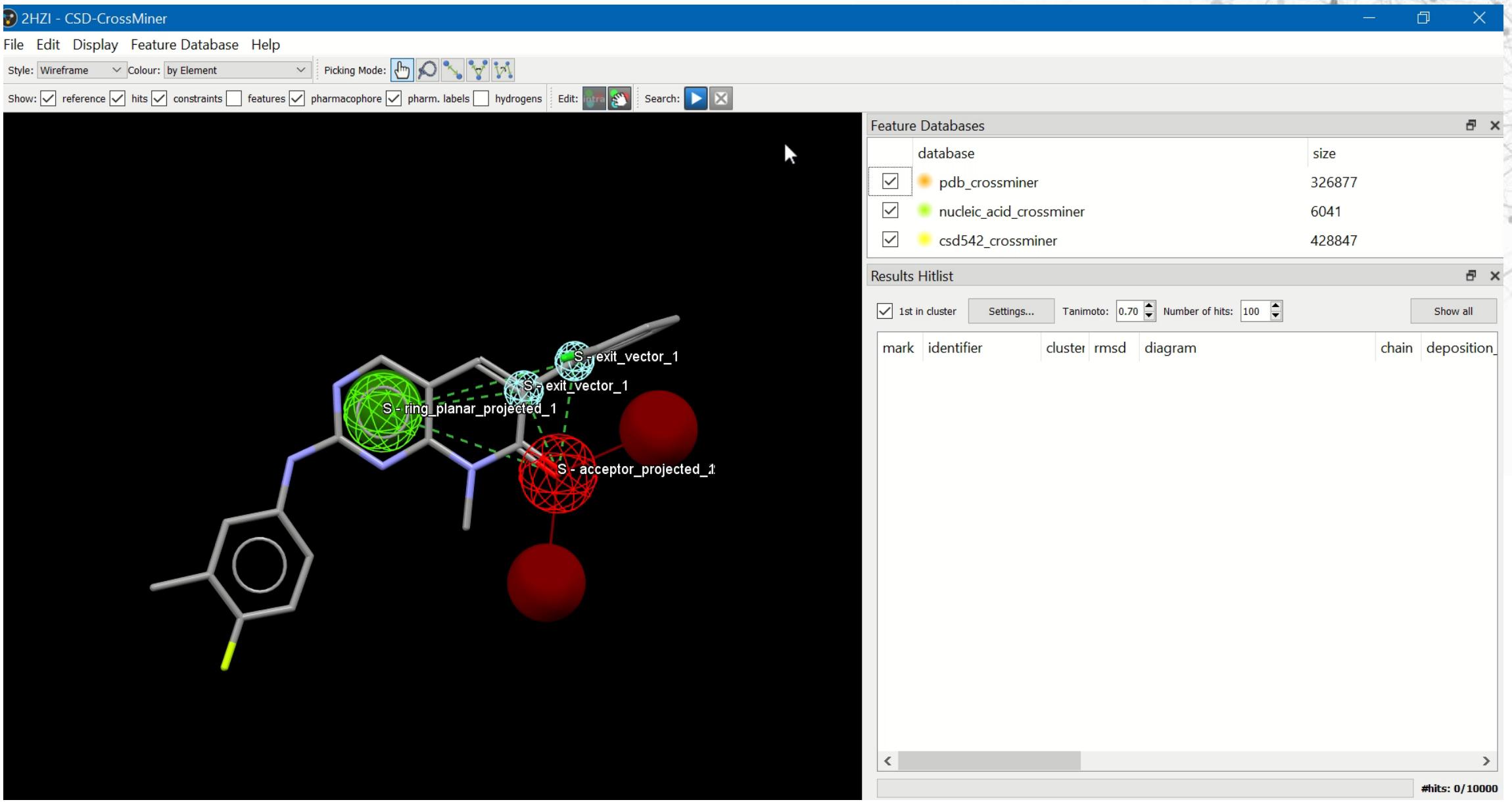
Inter-molecular constraint



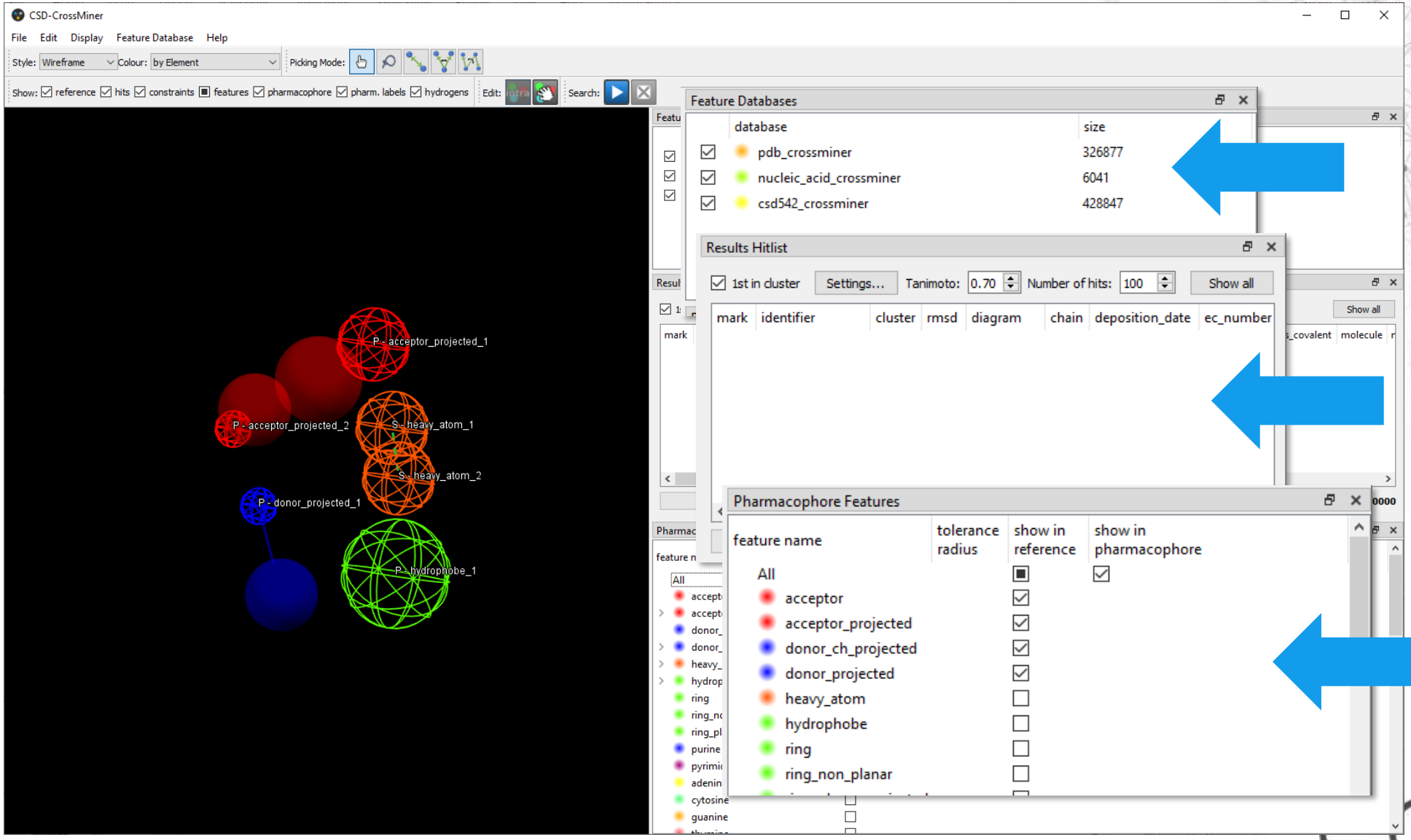
No constraint

CCDC

Run the search



Show One: CSD-CrossMiner Interface



The 3D window basics

- Left mouse button and move – rotate structures / pharmacophores
- Middle Mouse wheel –move structures / pharmacophores up and down

- Right mouse button and move up and down – zoom in and out of structures / pharmacophores

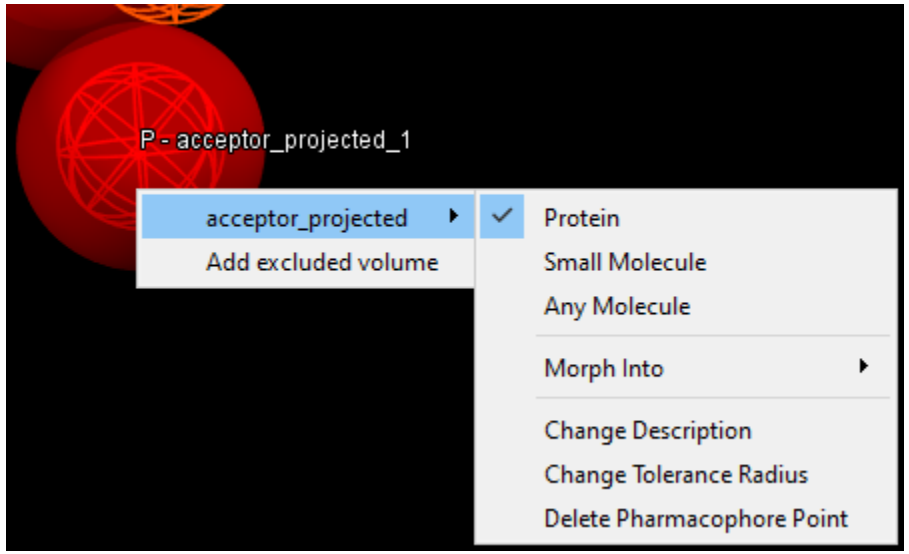
- Shift + Left mouse button and move - rotate in the plane structures / pharmacophores

- Ctrl + Left mouse button and move - translate structures / pharmacophores

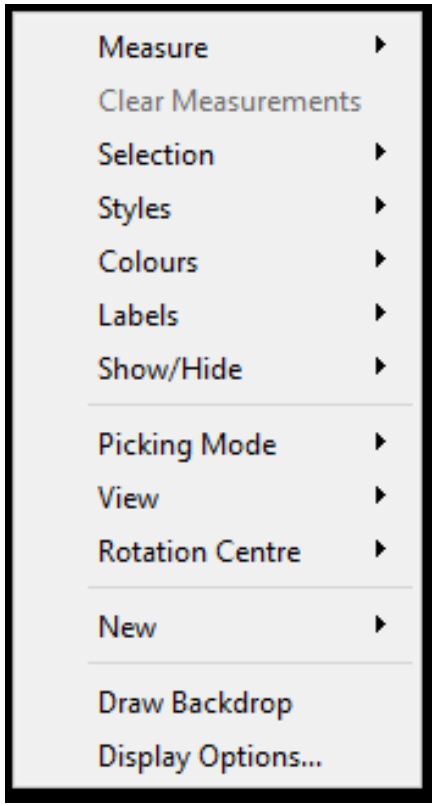
CCDC

The 3D window basics – Right Click

On a feature
(Pharmacophore context menu)

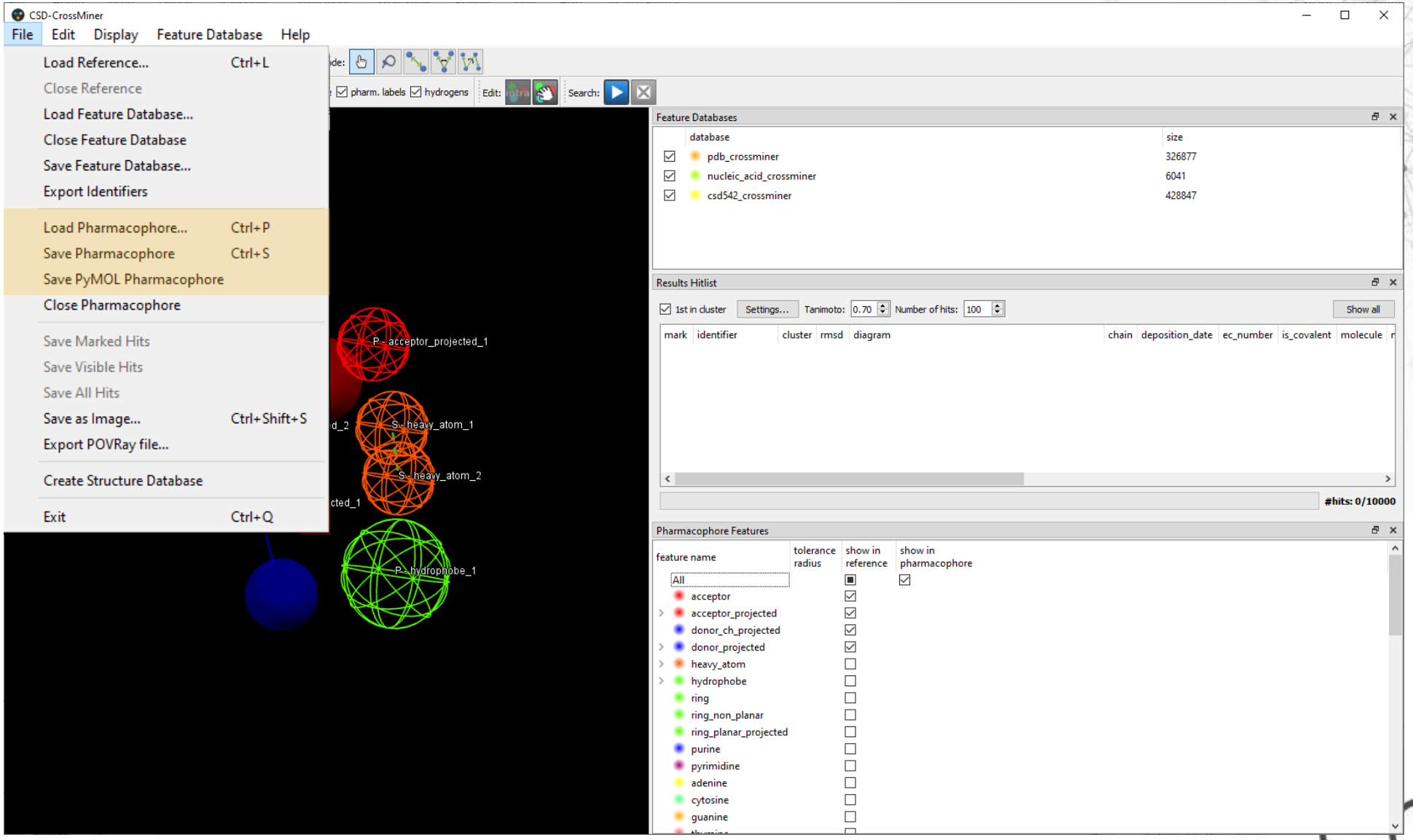


Away from a feature
(3D view menu)

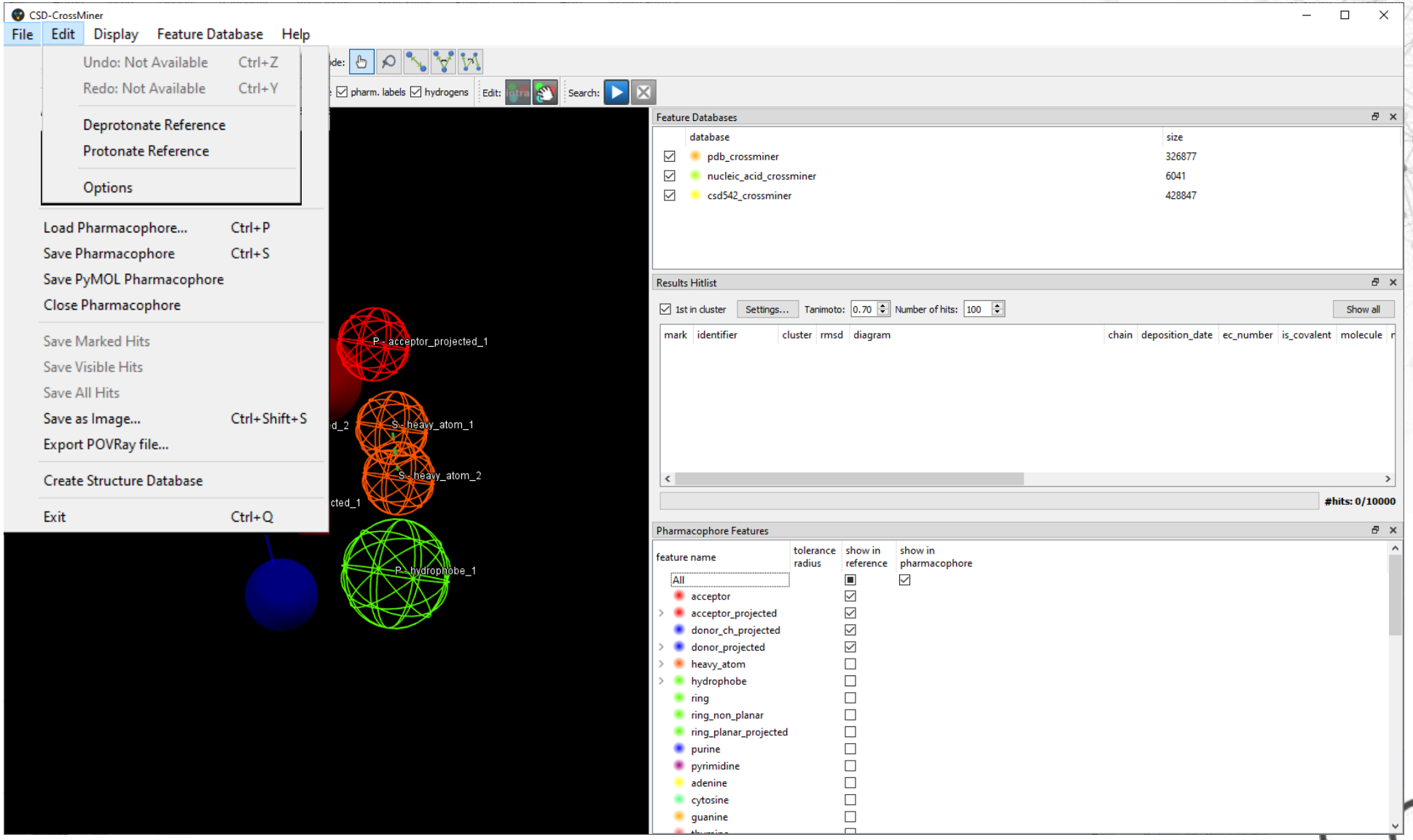


CCDC

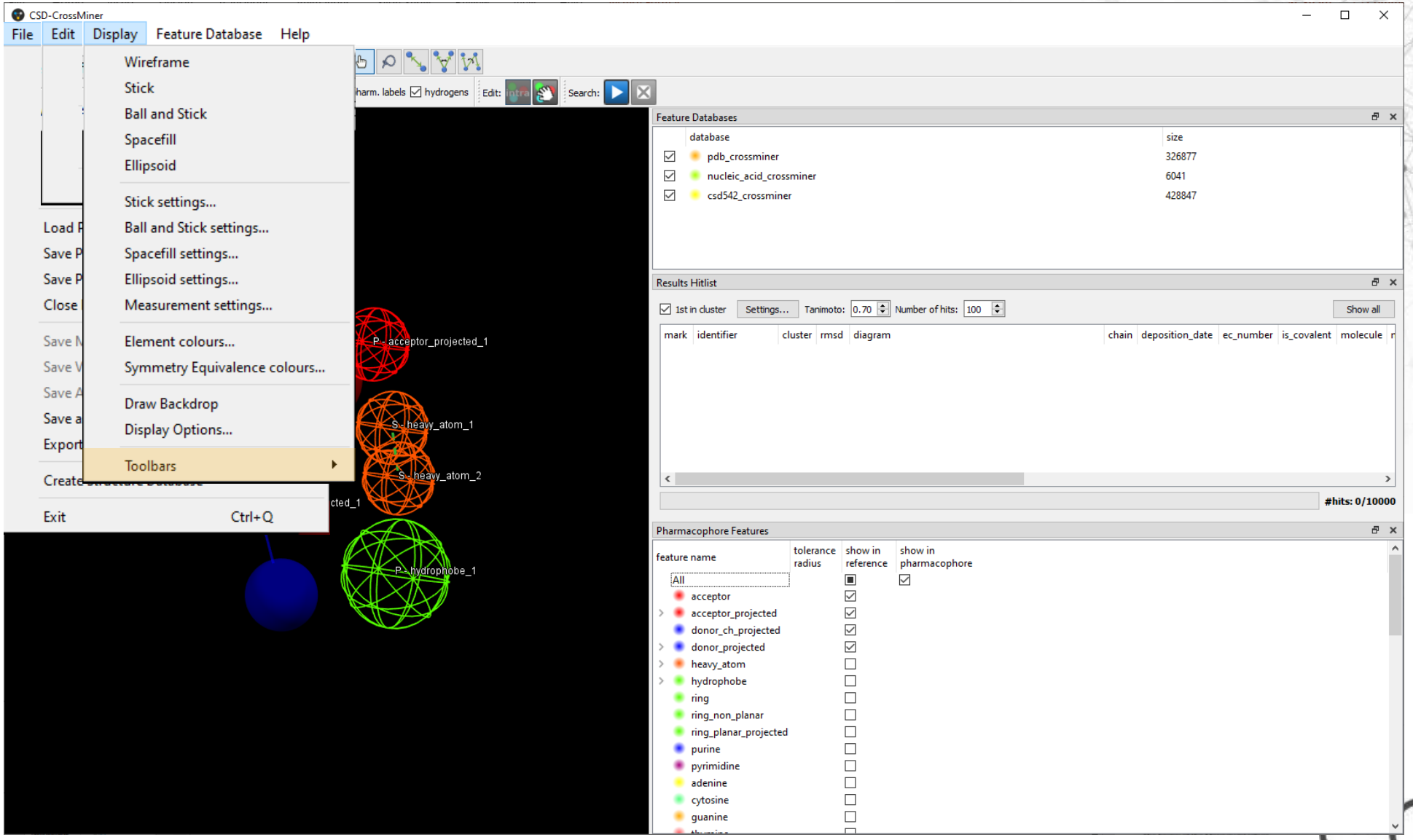
Show One: CSD-CrossMiner Menus



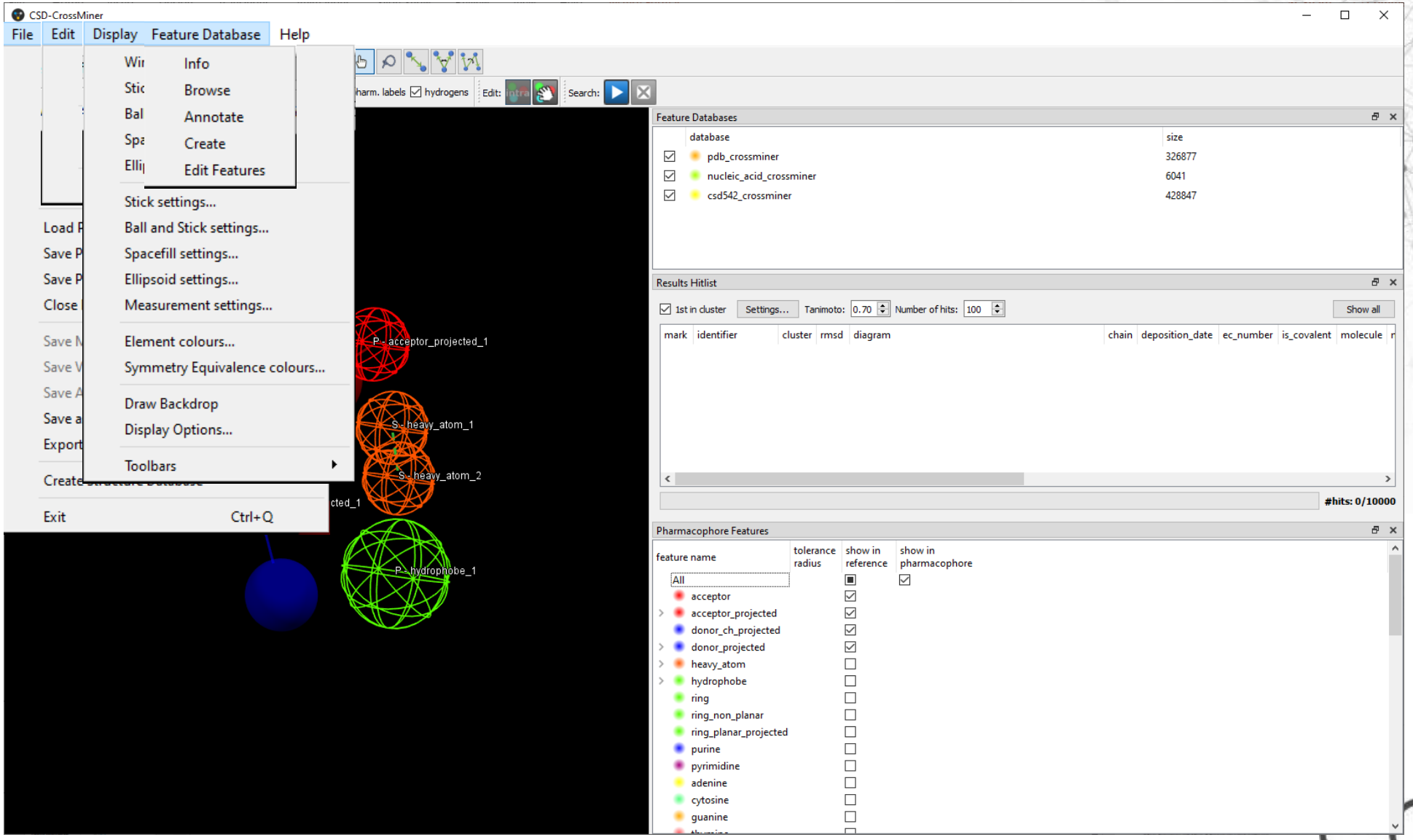
Show One: CSD-CrossMiner Menus



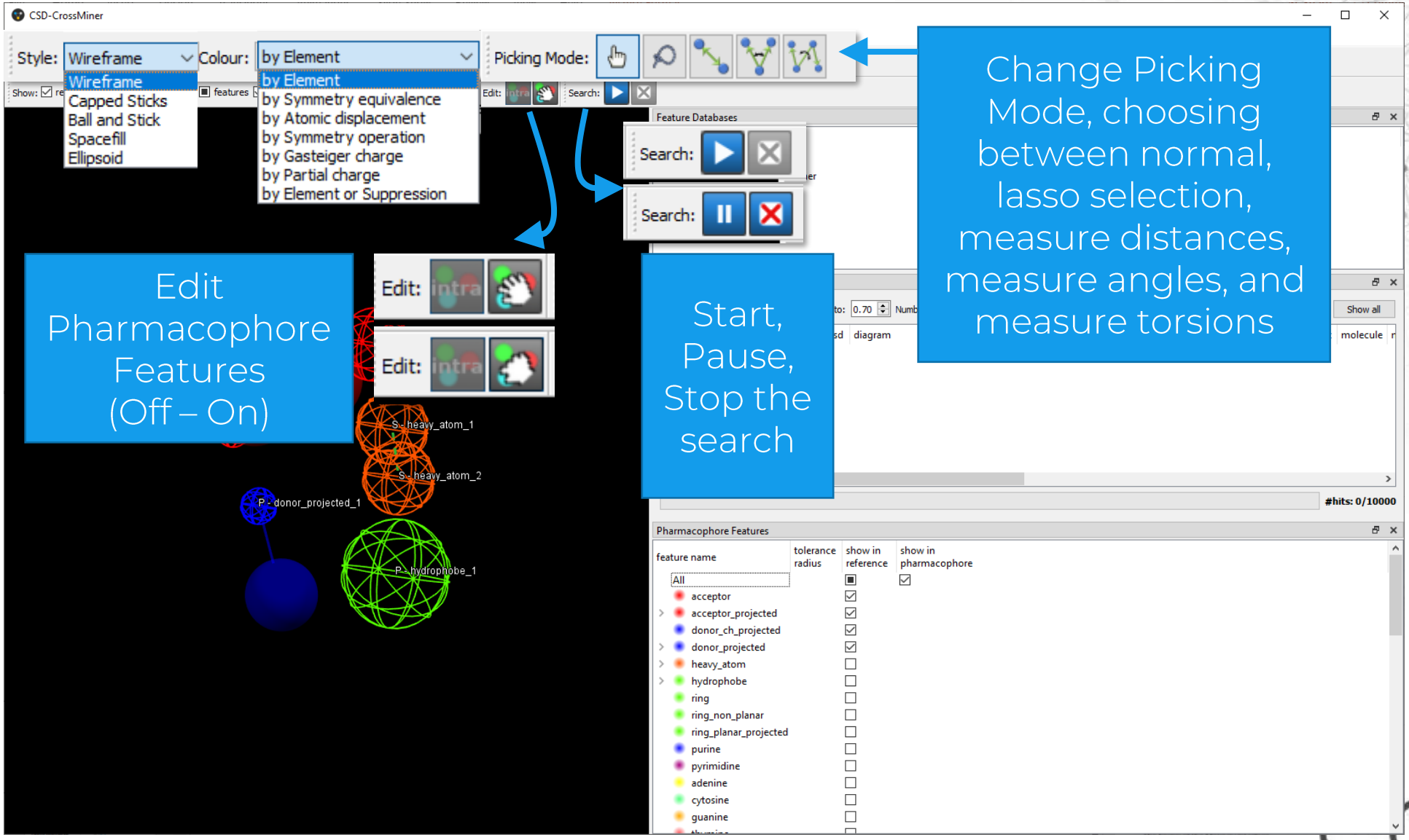
Show One: CSD-CrossMiner Menus



Show One: CSD-CrossMiner Menus

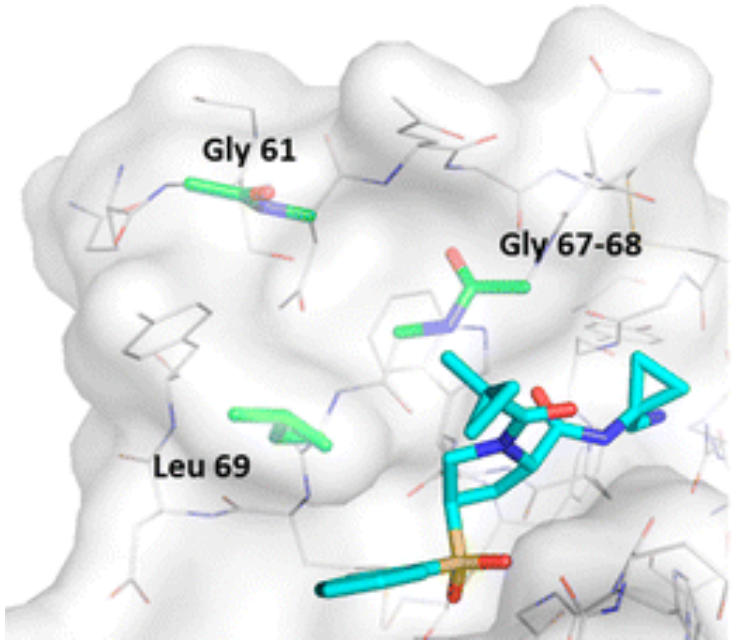


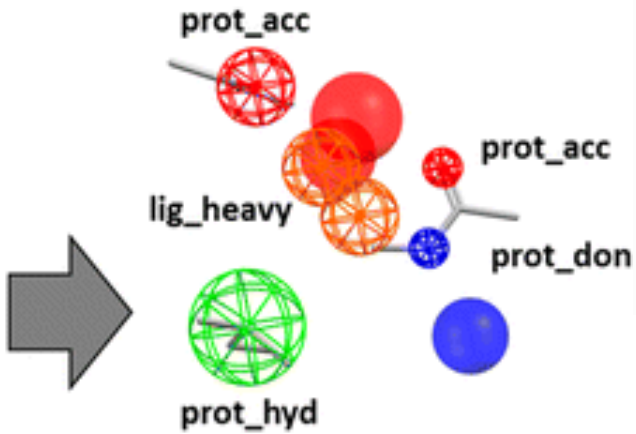
Show One: CSD-CrossMiner Interface



Show One: CSD-CrossMiner Demo

- Cathepsin L in complex with a nitrile inhibitor



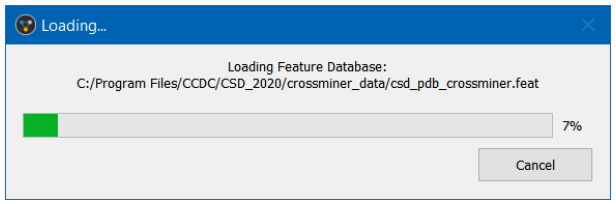


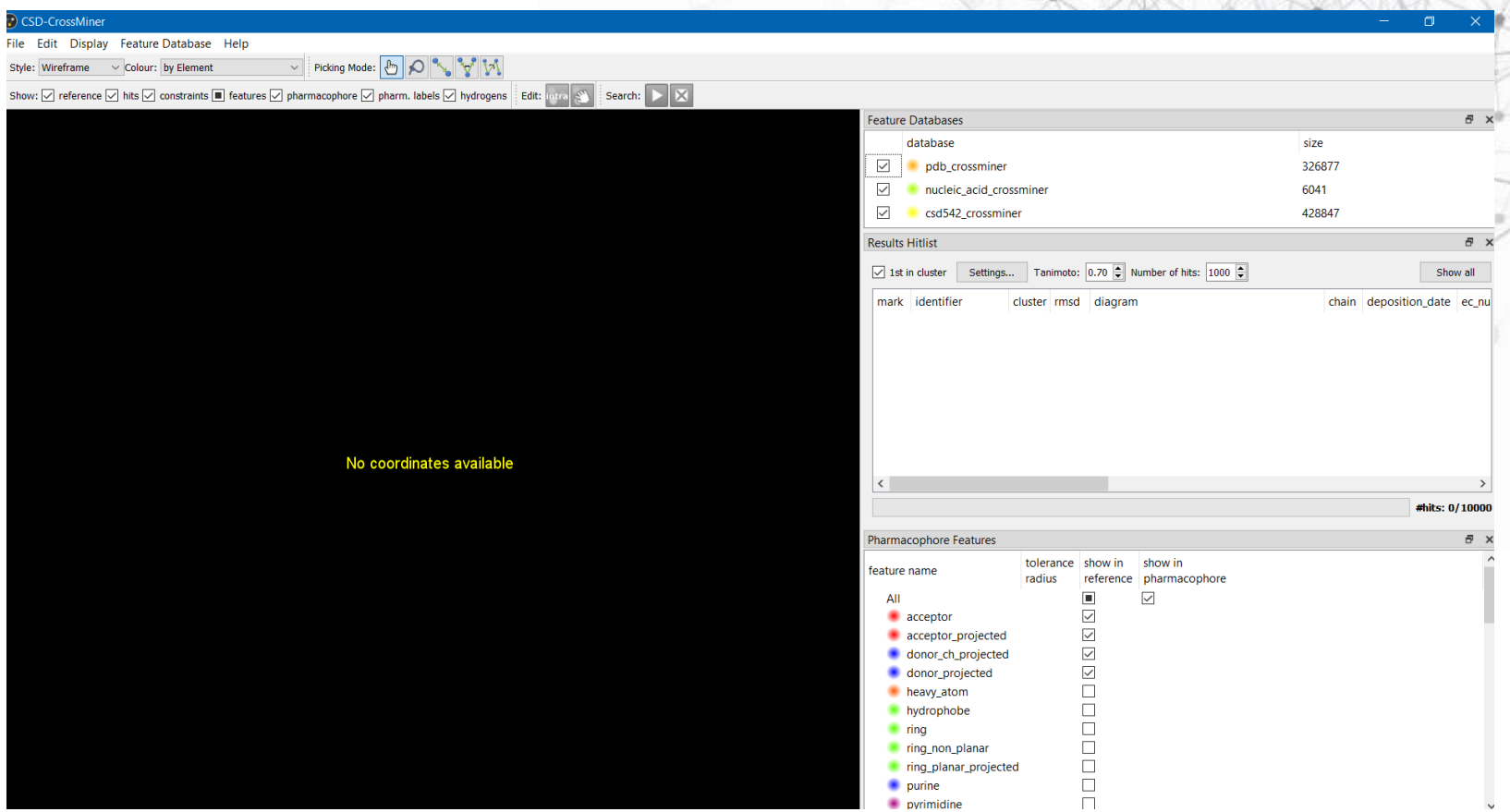
2XU1 PDB code

CCDC

Show One: CSD-CrossMiner Demo

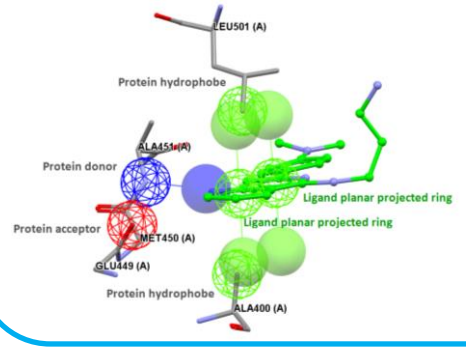
1. By default the provided *csd_pdb_crossminer.feat* feature database will be loaded

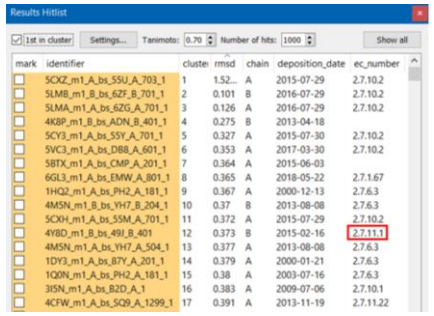




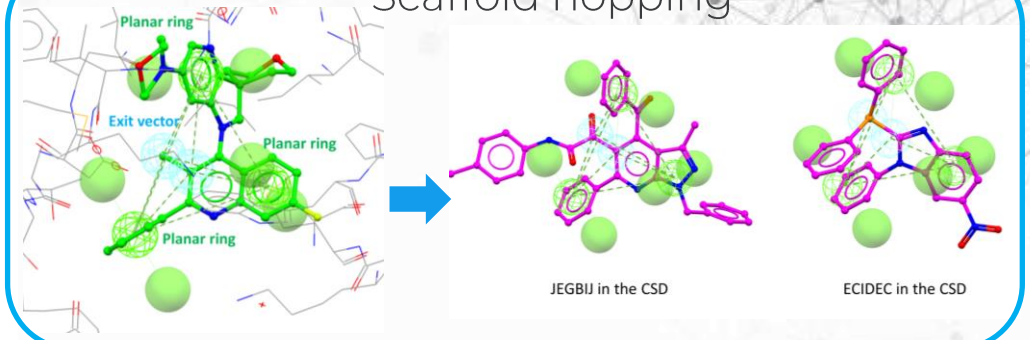
What else can you explore in CSD-CrossMiner?

Identify common protein-ligand pattern

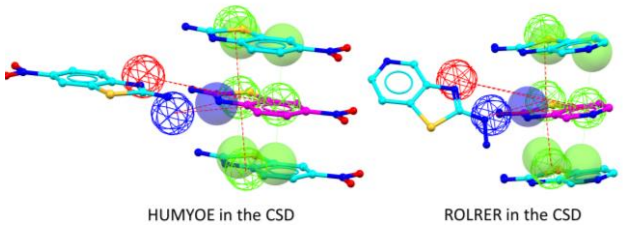




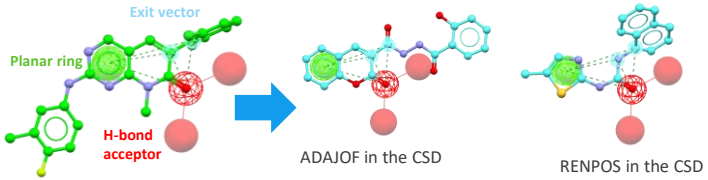
Scaffold hopping



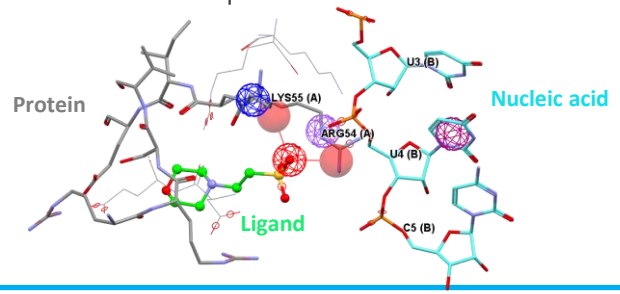
Mine co-crystals



Find bioisosters



Identify nucleic-acids-protein pattern



...and more!

CCDC

CSD-Discovery overview

CSD-Discovery.

GOLD: Protein-ligand docking and virtual screening

CSD-CrossMiner: Interrogate the CSD and the PDB for common interaction patterns

Ligand-based **virtual screening** workflow to find new hits.

CSD-Conformer Generator: Generation of molecular conformations.

CSD-Ligand Overlay: Flexible alignment of ligands.

CSD Python API: Create CSD-driven analyses and workflows.

Script-based interfaces to the **Field-Based Ligand Screener**; protein **cavity and sub-pocket search** and comparison; **protein-ligand substructure & interaction pattern mining.**

SuperStar: Analyse, predict and understand protein and ligand interactions

CCDC

Drug Discovery Pipeline

CCDC

Want to explore more?

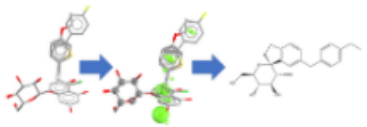
Training and Educational Resources

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

...r colleagues continually produce educational materials for use in classroom and computer lab settings, or as independent study modules. Many of these materials ...eaching 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 ...ase of over one million entries are available for free through our Access Structures portal.

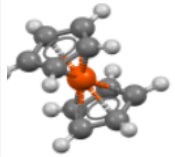
...ator looking for supplementary teaching materials, find out more about the Teaching Database **here**. If you have developed your own modules using the CSD and ...e them with the broader community, please contact us at education@ccdc.cam.ac.uk.

...e with the latest news from education and outreach at the CCDC, sign up for the Education and Outreach Newsletter **here**.

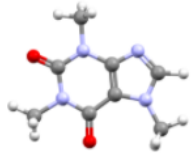


CSD-Discovery

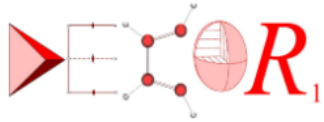
For pharmaceutical and agrochemical researchers, tools for discovering new molecules and performing protein docking studies.



Information on the Teaching Subset



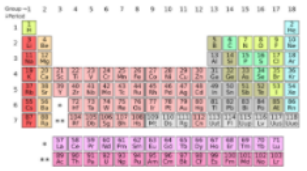
Access a series of teaching modules for use in the classroom



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



Explore the Periodic Table through Crystal Structures

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

Register for E&O newsletter

Self-guided workshops https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/

YouTube and LabTube channels links from https://www.ccdc.cam.ac.uk/Community/educationalresources/ccdc-videos/

CCDC

Want to explore more?

CCDC Virtual Workshops:

- 21st Apr - Programmatic search and analysis using the CSD Python API

How did it go today?
Let us know and help us improve filling in this short survey:
https://www.surveymonkey.co.uk/r/SLMYVC3

- Follow #CSDTopTipTuesday for weekly tips on using our software
- What's Up Webinars
  - Next one is 20th May
  - Featuring our Ligand Overlay tool
- Discovery Science Meeting
  - 9th – 10th June 2021
  - Registration and call for abstracts now open

CCDC