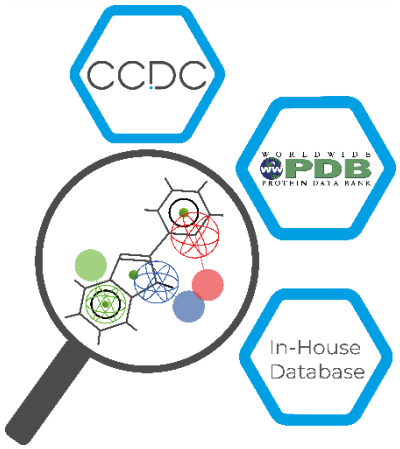
Modifying a Pharmacophore Query in CSD-CrossMiner (CROSS-003)

Developed using 2024.1 CSD Release



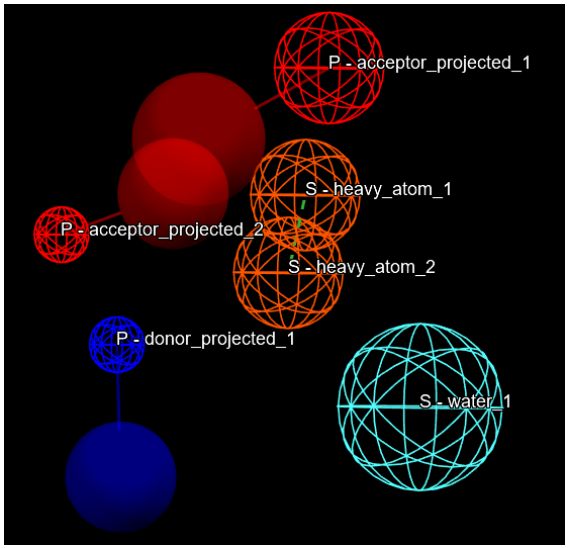
CCDC
advancing structural science

Table of Contents

Introduction 2
Learning Outcomes 2
Pre-required Skills 2
Materials 2
Example 1. Modifying a Pharmacophore Query 3
Conclusion 7
Summary 8
Next Steps 8
Feedback 8
Glossary 9
CSD-CrossMiner Terminology 10
Features and Pharmacophore Representation 11

Introduction

CSD-CrossMiner can be thought of as a pharmacophore-based query tool. However, it is much more powerful than traditional pharmacophore query tools as it allows you to query not only databases of ligands, but also proteins and protein-ligand interactions. CSD-CrossMiner includes a preconfigured database of biologically relevant subsets of the Cambridge Structural Database (CSD) and the Protein Data Bank (PDB). The pharmacophore used in the query is interactive, allowing you to easily edit it through a simple user interface. This delivers an overall interactive search experience with application areas such as interaction searching, scaffold hopping or the identification of novel fragments for specific protein environments, for example.



Further introductory information can be found in workshop CROSS-002 (https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-discovery-workshops/)

Learning Outcomes

After completing this workshop, you should be able to edit an existing pharmacophore query to create (and run) a new query.

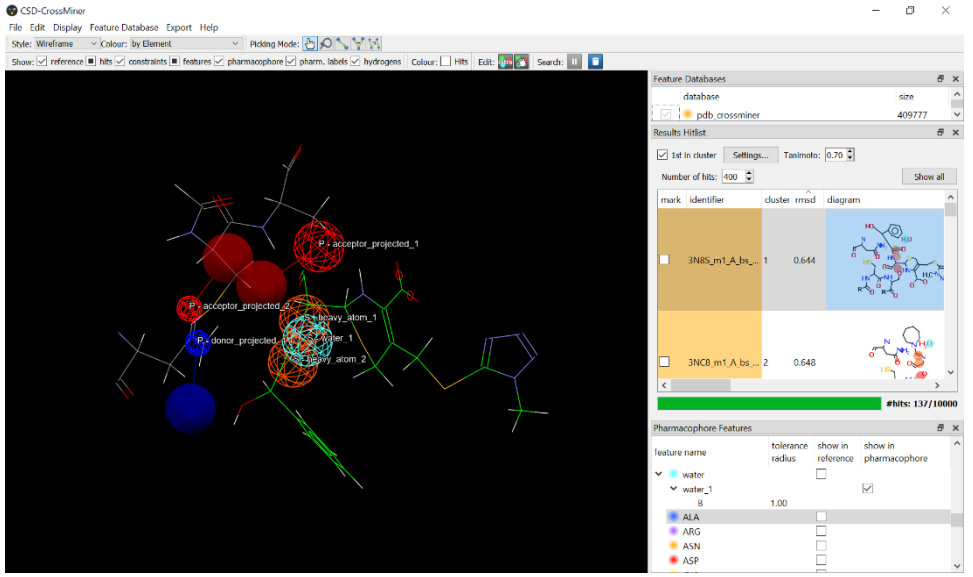
This workshop will take approximately ***30 minutes*** to be completed. The words in *Blue Italic* are reported in the Glossary at the end of this handout.

Pre-required Skills

There are no pre-required skills for this workshop, however some knowledge of the representation of pharmacophores and features in CSD-CrossMiner is expected. A summary can be found at the end of this handout.

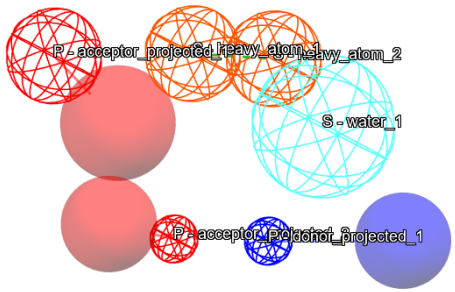
Materials

There are no additional materials required for this workshop.



Example 1. Modifying a Pharmacophore Query

One of the powers of CSD-CrossMiner is the ability to manually interact with the pharmacophore query and edit a *pharmacophore point* at any time (even while a search is running). For this example, you will be editing the cathepsin L pharmacophore that is provided in the CSD-CrossMiner installation folder.



The pharmacophore derived from protease cathepsin L.

Launch CSD-CrossMiner clicking on the CSD-CrossMiner icon: . Wait a few minutes for loading and initialising. If you already have work in CSD-CrossMiner, close it by clicking *File > Close Pharmacophore* and/or *File > Close Reference.*

1. Load the cathepsin L pharmacophore by clicking on **File** > **Load Pharmacophore**… and select:

<CCDC installation folder>\ ccdc-software\csd-crossminer\example_pharmacophores\catl_s3.cm

2. Right click on the *hydrophobe* feature (green in the 3D view) to bring up the pharmacophore context menu. Click *Hydrophobe > Morph Into > water*. You will notice that the molecule type also changed from *Protein (P)* to *Small Molecule (S),* as water is not considered part of the protein.

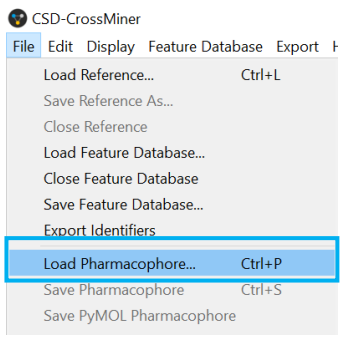
Through the pharmacophore context menu, for a feature it is possible to:

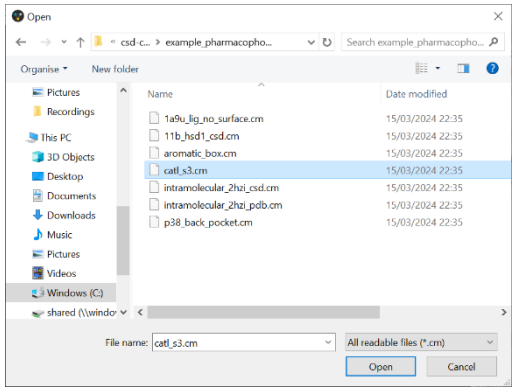
- Define where the feature belongs: a **Protein**, a **Small Molecule**, or **Any**.
- Change the feature type (**Morph Into**).
- Change the label of the feature (**Change Description**).
- Change the tolerance radius of the pharmacophore point (**Change Tolerance Radius).**
- Delete the pharmacophore point (**Delete Pharmacophore Point**).

Note that in the context menu of small molecule (S) pharmacophore points, there is the additional option to:

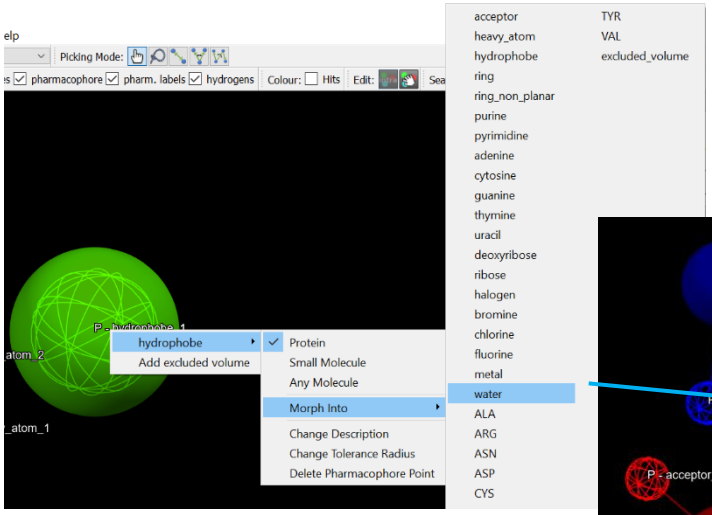
- Add constraints (**Constrain To**).

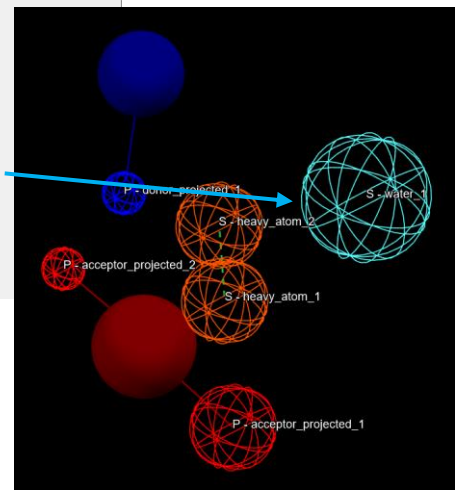
1





2





If a reference molecule is loaded, you have the additional option to:

- Move the pharmacophore point to the nearest atom (**Snap To Atom**).

3. Double-click on the radius size of virtual *acceptor_projected_1 (V)* in the **Pharmacophore Features** window or use *Change Tolerance Radius* from the pharmacophore context menu to change the radius of virtual *acceptor_projected_1 (V)* from 1.20 to 1.00. Then change the *water* radius from 1.50 to 1.00. This will reduce the uncertainty in the position of the water pharmacophore point.

4. Be sure that the **1st in cluster** tick-box is checked, this will show only the cluster representatives in the **Results Hitlist** window. Then start the search by clicking on the **Play** button:  in the **Search:** toolbar, the matched hits will populate the **Results Hitlist** window, as well as the 3D view. The progress bar at the bottom of the **Results Hitlist** window show the total number of hits.

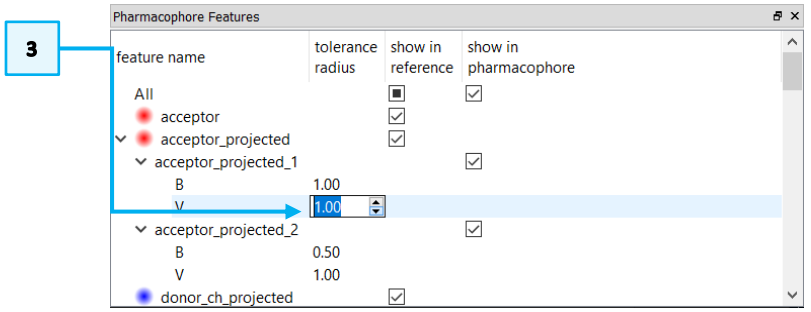
The search may require several minutes; however, let it go to completion. It is possible to change the number of processors dedicated to the pharmacophore search by changing the Number of threads in *Edit > Options*.

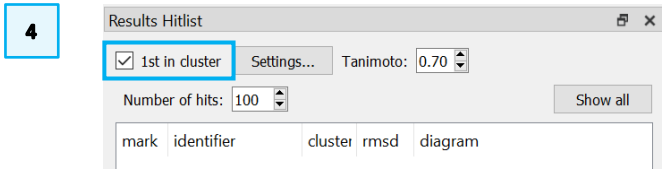
Note that the pharmacophore search options are not available to be changed (greyed-out) when the pharmacophore search is running or paused.

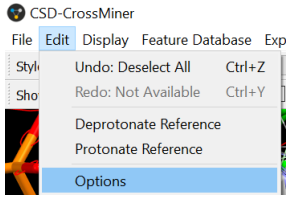
5. Locate and visualise the result with the lowest *RMSD* by clicking on the *rmsd* column in the **Results Hitlist** window, to show ascending order. If not present, you may wish to show the 2D diagram of the matched hits by right-clicking in the **Results Hitlist** window and then ticking **diagram**.

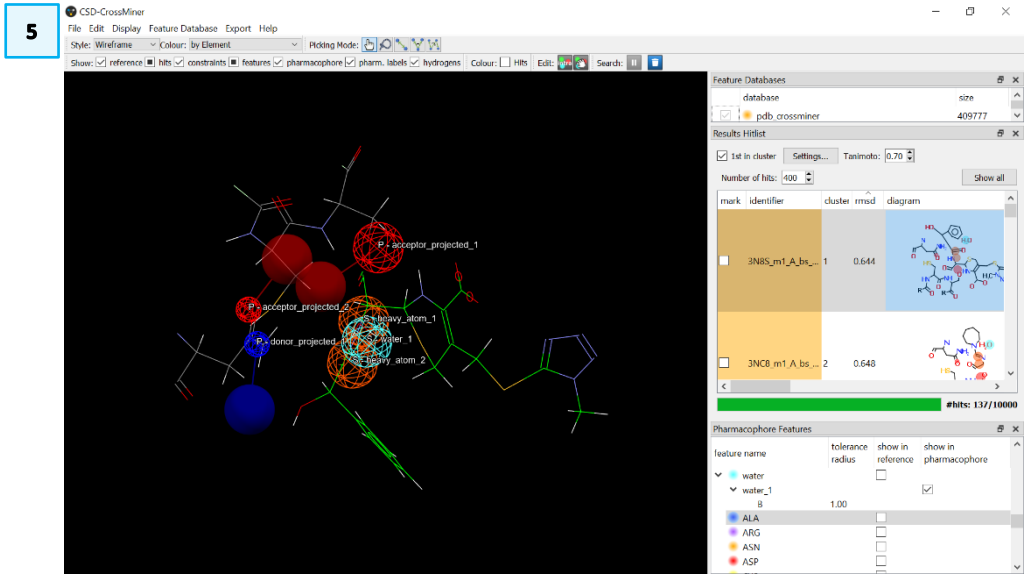
CSD-CrossMiner allows you to visually interact with the pharmacophore:

- You can edit the pharmacophore points in the **Pharmacophore Features** window or in the 3D view (changing the pharmacophore tolerance,









molecule type, pharmacophore type, etc.).

- You can move and resize pharmacophore points from the 3D view. All changes are in real time and have immediate feedback on your query and on the hits found.

  To do so:

6. Switch to interactive editing by click on the **Pharmacophore editing** button  in the *Edit:* toolbar (an open hand means editing is off, and a closed hand  means editing is on). This will turn the mouse cursor in a small hand.
7. If you have a three-button mouse: hover over the water feature, then click and drag using the middle mouse button (MMB). This will change the radius of the water sphere and automatically restart the search.

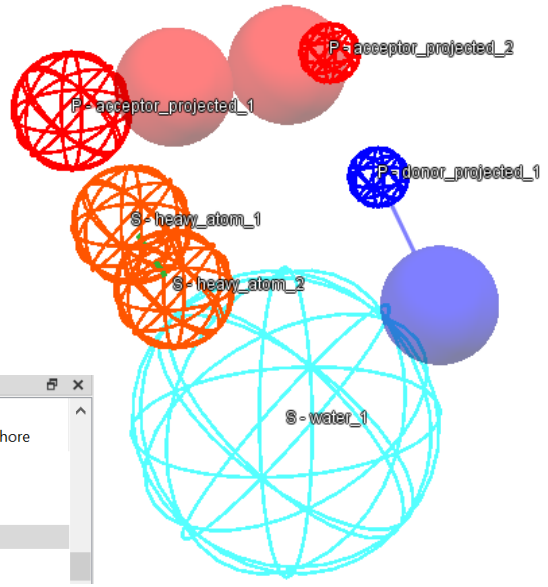
   If you do not have a three-button mouse or you have trouble getting this to work, you can use the *Pharmacophore Features* window (See point 3) to change the radius of the water pharmacophore sphere.
8. Change the water radius back to 1.00 using **Pharmacophore editing** button or from the **Pharmacophore Features** window.

Whenever you edit the pharmacophore, the current results in the **Results Hitlist** window and in the 3D view will disappear (as a new search will start). If you want to edit a pharmacophore overlaid with a molecule from the hits, you will need to first set the molecule as a reference.

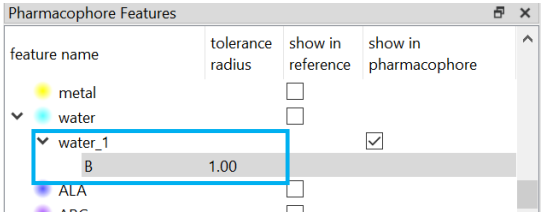
9. Right-click on one of the hits in the **Results Hitlist** window and click **Use as reference**. This will load the molecule into the 3D view and will show the donor and acceptor features associated with the molecule.

Note that if a different choice of displayed features was made during the CSD-CrossMiner session, those features (if present in the reference molecule) will be displayed instead.

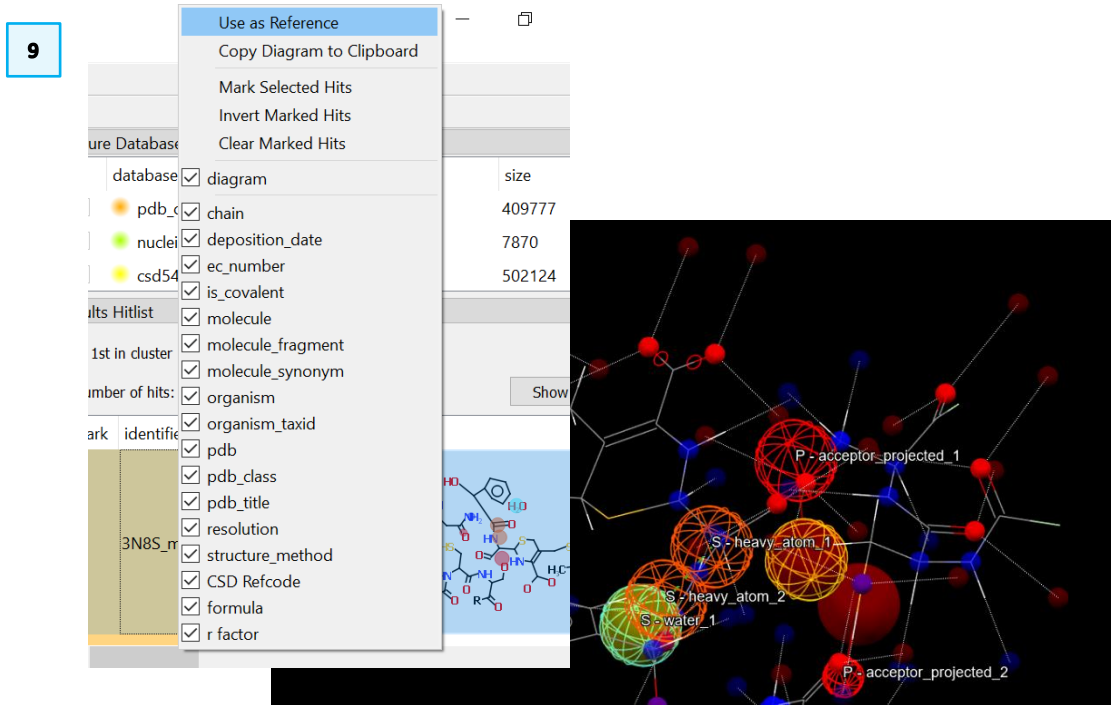
7



8



9



10. Display all the features associated with the reference structure by clicking on the **feature** ☐ tick-box in the CSD-CrossMiner *Show:* toolbar (which will turn to ☑) or alternatively by ticking the tick-box for *All* features in the *show in reference* column in the *Pharmacophore Features* window.

11. Hide all features by clicking on features ☑ (which will turn to ☐) or alternatively by unticking the **All** tick-box in the **Pharmacophore Features** window.

12. Change the style of the small molecule in the 3D view by left-clicking on one of the small molecule atoms while pressing **Shift** key. This will select all the atoms of the same molecule (the selected atoms are highlighted with small yellow spheres). Right-click on one of the selected atoms and pick **Styles**, from the pull-down menu and select the desired style.

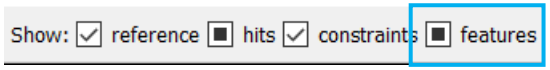
13. With **Pharmacophore editing** mode on (), drag one of the heavy atom pharmacophore points (in orange in the 3D view) to a nearby carbon atom using the left mouse button (LMB).

Note that the search starts as soon as you let it go. The reference molecule stays visible even while the search runs.

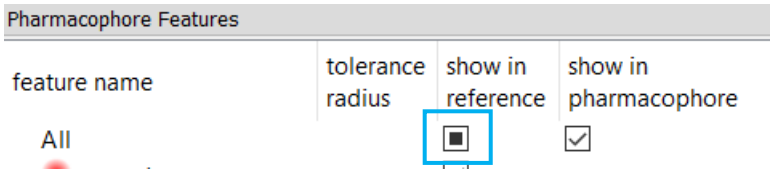
14. To undisplay the reference molecule, untick the **reference** tick-box in the *Show:* toolbar. This *Show:* toolbar controls the visibility of the reference molecule, hits, constraints, features, pharmacophore, pharmacophore labels and hydrogens.

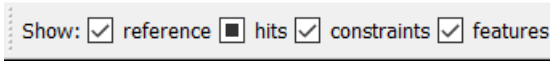
Every CSD-CrossMiner user has accidently grabbed a pharmacophore point and dragged it across the screen when they intended to rotate the molecule viewer. In such cases **Ctrl + z** (or *Edit* > *Undo*) is very helpful. It will undo the last change made to the pharmacophore. However, the search will start over again. Therefore, be sure to turn off the **Pharmacophore editing** mode when you are

10

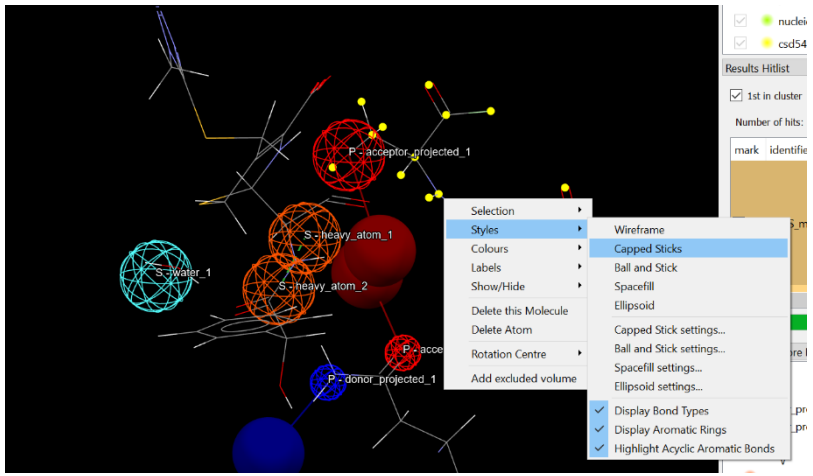


11

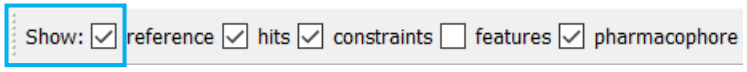




12



14



not using it (open hand  button).

15. Press  to stop and erase the pharmacophore search.

16. You can also add new features from the **Pharmacophore Features** window. Scroll to fluorine in the feature browser and right click, then click **Create fluorine**. This will drop the new feature into the 3D view.

17. Align the new fluorine near to the closest atom by right-clicking on the fluorine feature and select *Snap To Atom*.

18. Start the search by clicking on  . This new search probably won't find any results. This step was just to demonstrate how to add and manoeuvre new features.

19. Save your edited pharmacophore by clicking *File > Save Pharmacophore*. Name the file "pharmacophore_edited.cm".

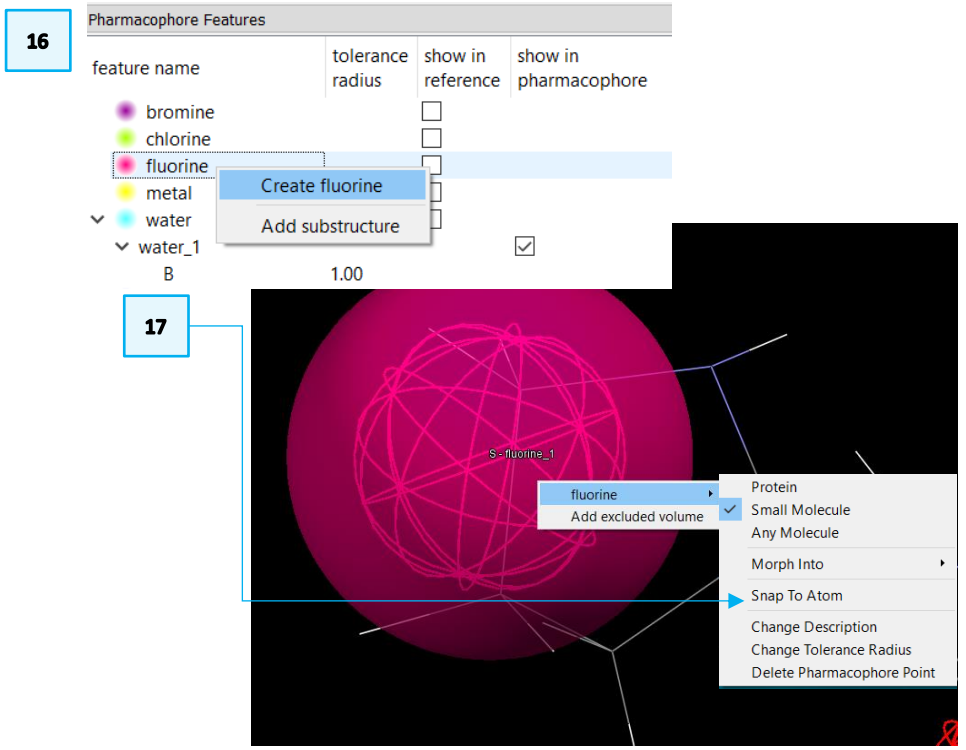
There are two options for saving a pharmacophore underneath the **File** menu, as defined below:

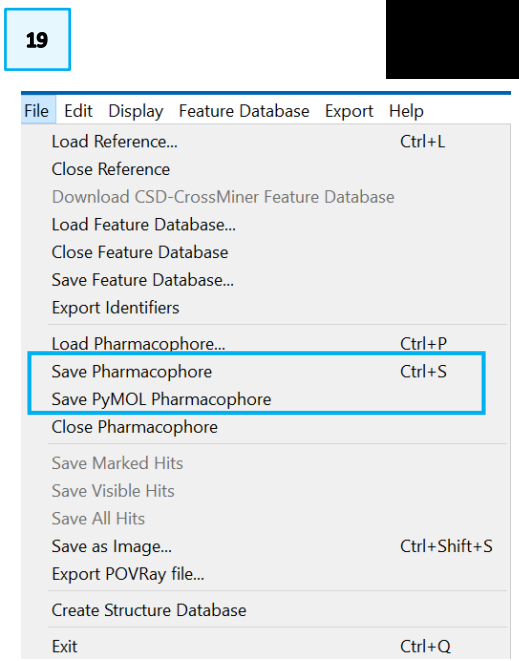
Save Pharmacophore: Will save the current pharmacophore in a CSD-CrossMiner file format (*cm*). This is a text file with all feature definitions and the (x, y, z) of the instances used.

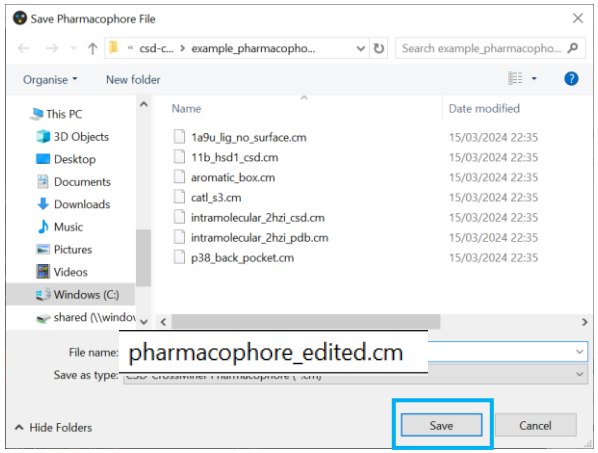
Save PyMOL Pharmacophore: Will save the current pharmacophore as a python script to be run in PyMOL. This will create graphic objects in PyMOL in the same coordinate frame as the saved hits.

Conclusion

This example demonstrated how an existing pharmacophore can be modified, including on-the-fly, to create a new search.







Summary

In this workshop, you have seen how to edit/modify an existing pharmacophore query. You should now be able to:

- Modify existing features from a given pharmacophore
- Run a pharmacophore search interactively
- Add new features to a pharmacophore
- Save pharmacophore queries

For your reference, you can find the user manual at this link.

Next Steps

After this workshop, you can continue learning about CSD-CrossMiner with more exercises available in the self-guided workshops available in the CSD-Discovery workshops area on our website.

https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/csd-discovery-workshops/

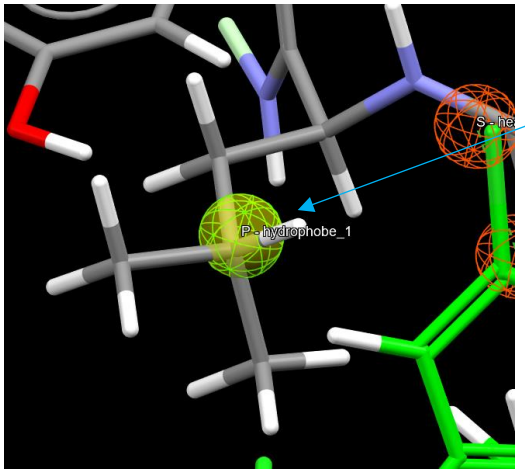
Feedback

We hope this workshop improved your understanding of *[tool name and activity]* and you found it useful for your work. As we aim to continuously improve our training materials, we would love to hear your feedback. Follow the link on the workshop homepage and insert the workshop code, which for this self-guided workshop is CROSS-003. It will only take 5 minutes and your feedback is anonymous. Thank you!

Glossary

Hydrophobic/hydrophobe
Hydrophobic molecules effectively “repel” water and thus have a tendency to self-aggregate in aqueous media, excluding water in so doing. On a structural level, these are non-polar groups such as alkyl or aryl moieties. If these functional groups or molecular fragments are also pharmacophore features, then they are called *hydrophobes* in CSD-CrossMiner.

Root Mean Square Deviation (RMSD)
The root mean square deviation (RMSD) is a commonly used measure of the difference between two sets of values (usually comparing observed data to estimated data). The RMSD is defined as the square root of the mean squared error.



Hydrophobe pharmacophore point.

An isobutyl group is hydrophobic. The green mesh sphere indicates the position at which such a feature (functionally a hydrophobe) must be found.

CSD-CrossMiner Terminology

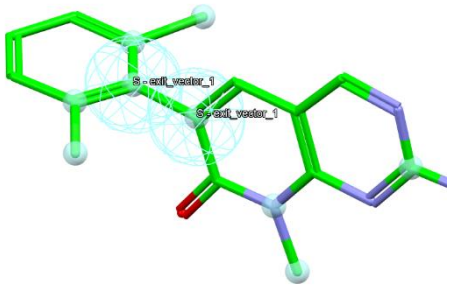
Exit vector
A two-point feature that represents a single, non-ring bond between two heavy atoms features; and it will be represented as two mesh spheres. In the case of CSD-CrossMiner, directionality in an exit vector does not matter.

Features
An ensemble of steric and electronic features that characterise a protein and/or a small molecule. In CSD-CrossMiner a feature is defined as point(s), centroid or vector which represent a SMARTS query and, in the case of a vector, this includes geometric rules.

Pharmacophore point
A feature that has been selected to be part of a pharmacophore because its presence is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger or block its biological response.

Structure database
Is a database containing the 3D coordinates of small molecule structures and/or protein-ligand binding sites. This database is used to create a feature database.

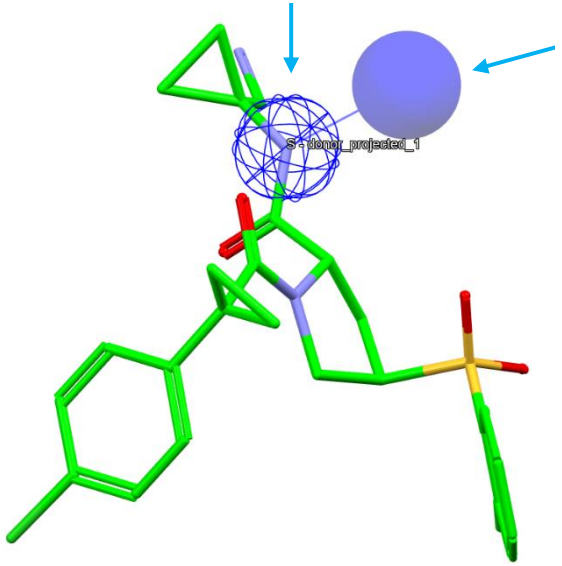
Feature database
A database containing the structures from the structure database, indexed with a set of feature definitions provided by CSD-CrossMiner and any additional features defined by the user. This is the database that CSD-CrossMiner uses to perform the actual 3D search against a pharmacophore query.



An exit vector (light blue mesh spheres) defined by the position of two carbon atoms.

Base point (the heavy atom of the donor group)

Virtual point - defines the direction the X-H group should point (Base point →Virtual point)



A molecule with a donor_projected pharmacophore point defined.

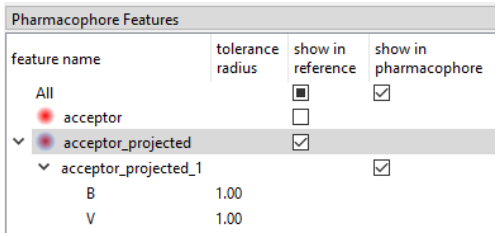
Features and Pharmacophore Representation

In the CSD-CrossMiner 3D visualiser, features are represented as small translucent spheres coloured as defined in the *Pharmacophore Features* window. A pharmacophore point is represented as a mesh sphere which reflects the uncertainty in the position of the pharmacophore point. In the 3D view:

- **P**: Protein pharmacophore point
- **S**: Small molecule pharmacophore point
- **A:** Either a small molecule or protein pharmacophore point
- **Dashed line**: intra and intermolecular constraints. Constrained features must belong to either the same molecule as each other (*intra*, dashed green line) or different molecules (*inter*, dashed red line).
- **Mesh sphere**: the actual feature itself, where the sphere size represents the radius of tolerance of the pharmacophore point.
- **Solid sphere**: the projected virtual point to represent the directionality of e.g. a hydrogen bond acceptor/donor. A feature can have more than one projected point. For example, a H bond acceptor can have multiple potential lone pair preferred projections.

Note that the colour coding of the pharmacophore points is defined in the *Pharmacophore Features* browser; e.g. hydrophobe features are green, hydrogen bond acceptors are red, and so on.

In the directional pharmacophore, the mesh sphere (the actual feature itself) is defined as *B* in the *Pharmacophore Features* window (**B**ase feature), and the projected virtual point representing the directionality of the feature is defined as *V* (**V**irtual point).



Protein (P): H bond acceptor feature (mesh) with projected directionality (solid)