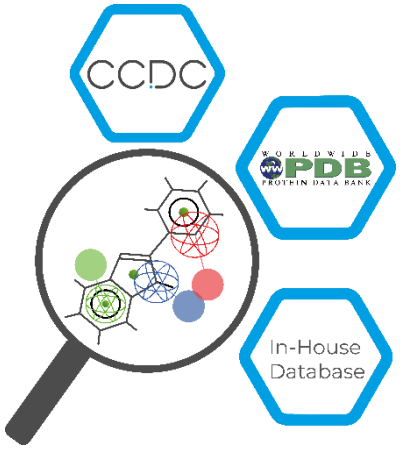
Further Pharmacophore Searching in CSD-CrossMiner (CROSS-006)

Developed using 2024.2 CSD Release



CCDC
advancing structural science

Table of Contents

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).

This tutorial will introduce you to building a pharmacophore query in CSD-CrossMiner using a reference molecule and searching a PDB-derived Feature Database.

Learning Outcomes

After completing this tutorial, you will:

- Understand how pharmacophore features are represented in CSD-CrossMiner
- Be able to construct a pharmacophore query from a reference molecule
- Perform a pharmacophore search of a preconfigured feature database and interpret search results

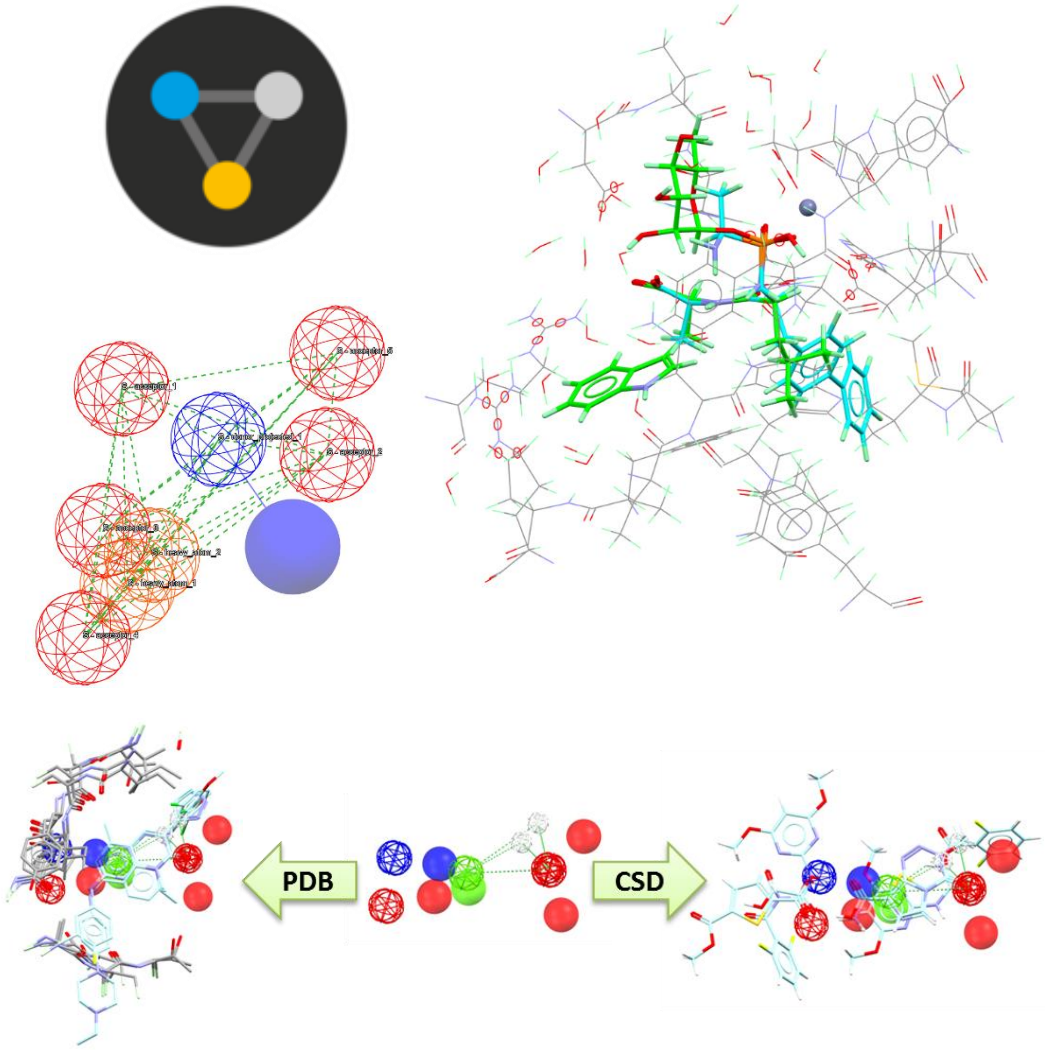
This workshop will take approximately ***45 minutes*** to be completed. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout. Summaries of CSD-CrossMiner terminology and pharmacophore representation are also provided.

Pre-required Skills

There are no pre-required skills, however some familiarity with life sciences is assumed.

Materials

No additional materials are required.

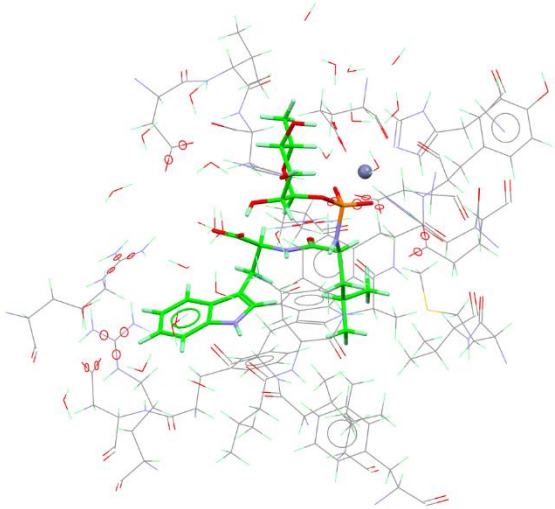


CSD-CrossMiner allows CSD and PDB databases to be searched in terms of pharmacophore queries.

Uncovering Neprilysin Inhibitors

Neprilysin is a zinc-dependent metalloprotease enzyme which is expressed in a wide variety of tissues and is particularly abundant in kidneys. Notably, it is a common acute lymphocytic leukemia (ALL) antigen that is an important cell surface marker in the diagnosis of the disease.

In this workshop, we will search for potential inhibitors of neprilysin using CSD-CrossMiner. We will do so by using the structure of neprilysin complexed with phosphoramidon (PDB code 1DMT), whose binding site region is incorporated into the pre-configured CSD-CrossMiner Feature Database, to create a pharmacophore query.



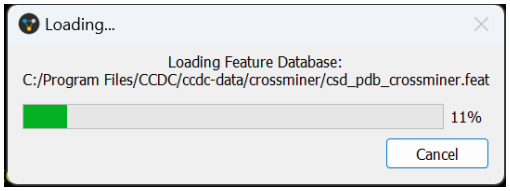
The binding site of a neprilysin-phosphoramidon complex.

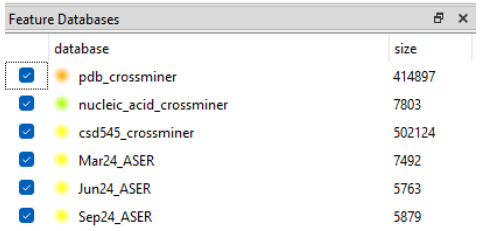
Launching CSD-CrossMiner

1. Launch CSD-CrossMiner by clicking the desktop icon  . You should see a *Loading…* dialogue box as the *Feature Database* loads. It may take some time to load. Once it has finished loading, the database components will be listed under **Feature Databases**.

2. *Only do this step if do not see the databases indicated in **Step 1***. First check that the **Feature Database** Window is displayed by going to *Display > Toolbars Feature > Databases* from the top menus. If still no databases are visible, from the top menus, go to *File > Load Feature Database…*, then navigate to "*<path to CCDC installation folder>*\ccdc-data\crossminer\csd_pdb_crossminer.feat" and click open.

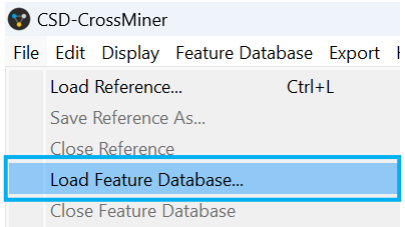
1

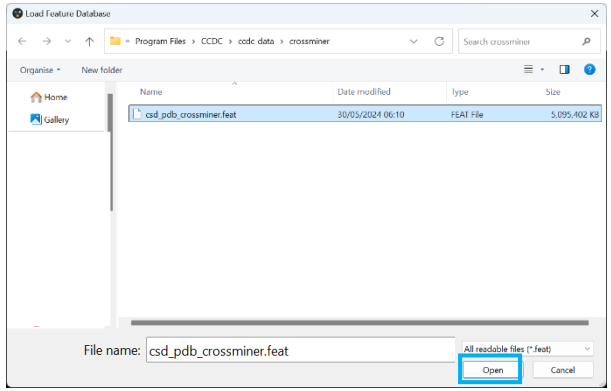




Only do Step 2 if you do not see these databases.

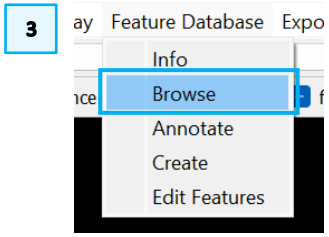
2



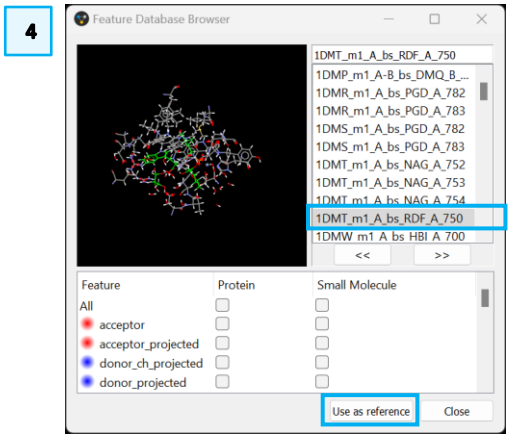


Loading and Preparing a Reference Structure

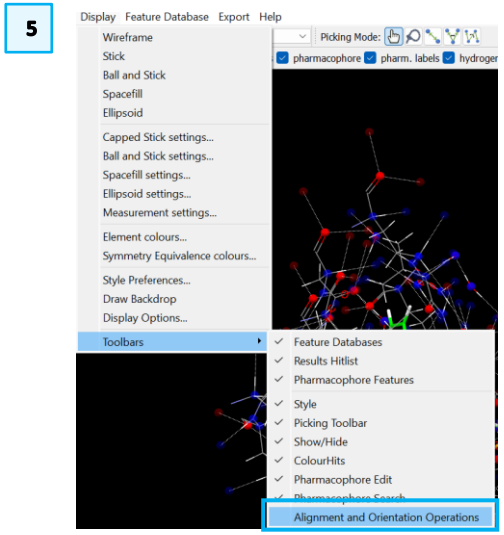
3. From the top menu, select *Feature Database* > *Browse* to bring up the **Feature Database Browser**



4. In the browser, scroll down until you find "1DMT_m1_A_bs_RDF_A_750" (or type it in the search bar) and click **Use as reference**, followed by clicking **Close**.

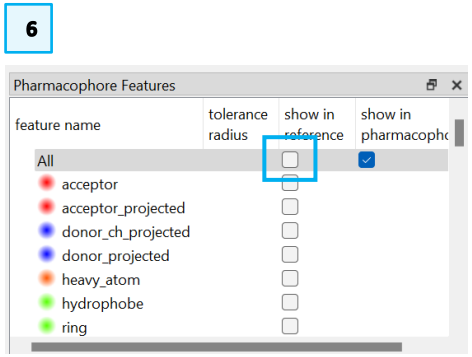


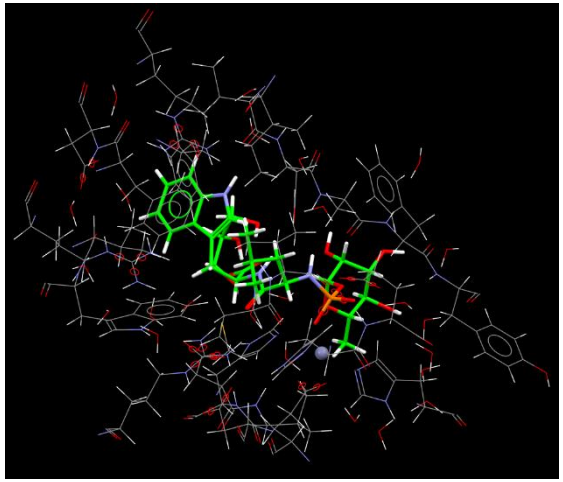
5. The *binding site* containing the ligand will be loaded into the main visualizer. Take a moment to inspect the structure. To change the view in the visualiser using a mouse:
    - Left-click + drag to rotate the structure
    - Right-click and move up/down to zoom in/out
    - Use the mouse wheel to move the structure vertically
    - Shift + left-click + drag to rotate in the plane
    - Shift+ right-click + drag to change depth cue



You may also enable the **Alignment and Orientation Operations** toolbar by choosing *Display* > *Toolbars* > *Alignment and Orientations Operations* from the top menu.

6. You will notice that there are many red and blue spheres at atomic positions, and thin white lines extending from them. These are potential *pharmacophore points*. To make the view clearer, these can be temporarily disabled form the ***Pharmacophore Features*** window by clicking the button in the *show in reference* column next to *All* twice, so that it is completely unticked.

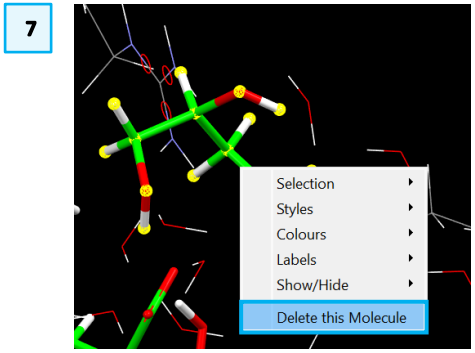


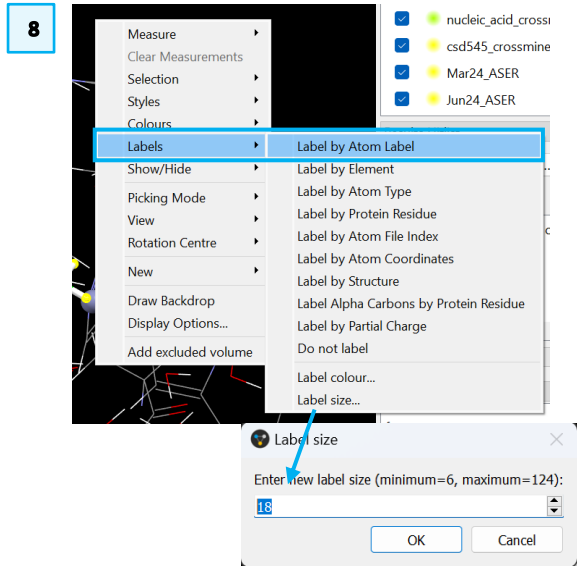


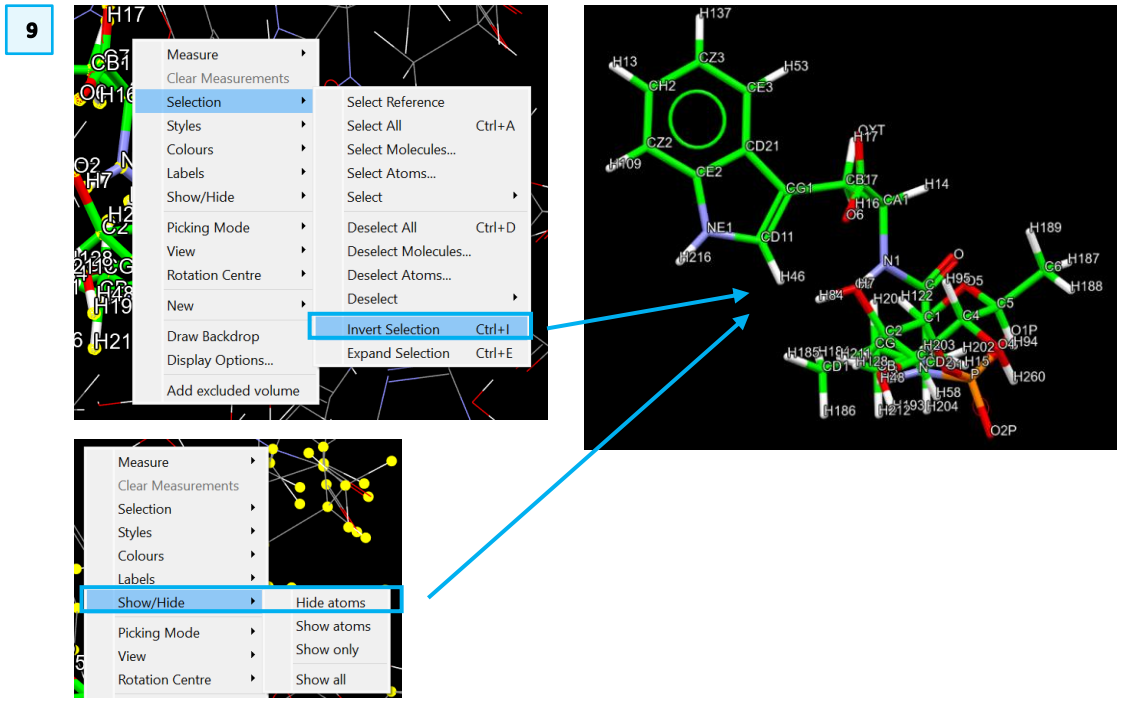
7. The binding site incorporates a glycerol molecule which will not be used to construct the pharmacophore query. Shift + left click on the molecule to select it (the atoms will be highlighted in yellow). Then right-click and from the drop-down menu, select *Delete this Molecule*.

8. Select the phosphoramidon molecule by shift + left clicking on the molecule. Then, in an empty region of the screen, away from the structure, right-click to bring up a drop-down menu. From here, select *Label > Label by Atom Label*. You can increase the label size by selecting *Label size…* from the same menu. We recommend a label size of at least 18pt.

9. To further clarify the display, you can hide all except the phosphoramidon molecule. Shift + left click on the phosphoramidon molecule. Right-click and from the dropdown menu, choose *Selection > Invert Selection*. Right-click again and from the dropdown menu, choose *Show/Hide > Hide atoms*. You should now see only the reference molecule.



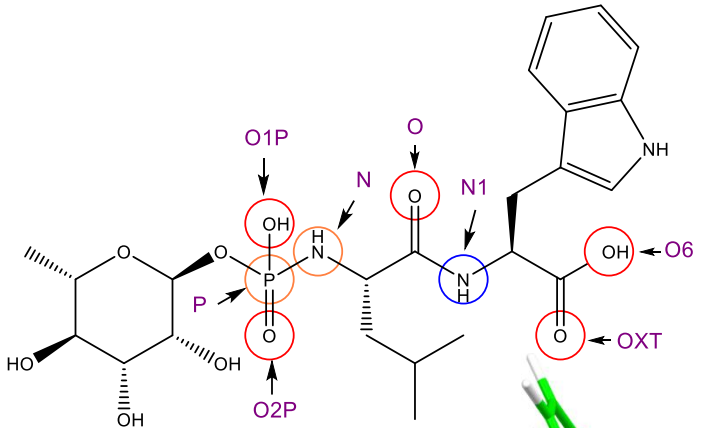




Building the Pharmacophore Query

We can now proceed to add features to the reference molecule to create a pharmacophore query. Detailed information on Features and Pharmacophore representation is provided at the end of this handout. Here we will discuss only pertinent features.

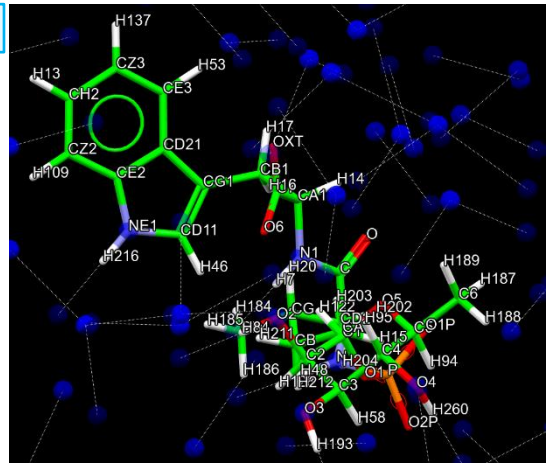
The structural diagram on the right shows the atoms that will be used to create the pharmacophore query.

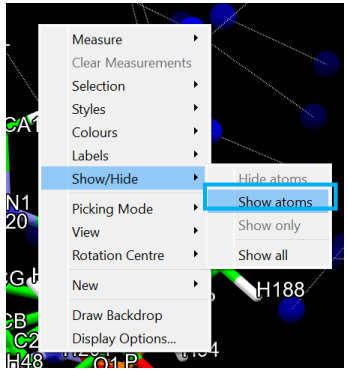


10. In the Pharmacophore Features window, tick *donor_projected* in the *show in reference column*. Blue spheres will appear in the visualizer. These correspond to *hydrogen bond donors*, and the projection indicates the direction in which the X-H group points. There are floating blue spheres which correspond to donors in the parts of the structure we have hidden. You can verify this by right-clicking and from the drop-down menu, selecting *Show/Hide > Show atoms*. Hide the atoms again as in **Step 9** afterwards.

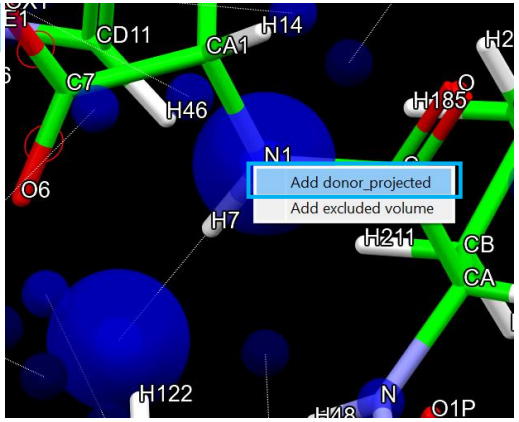
11. Locate atom *N1* in the reference molecule and right click. A drop-down menu will appear; from here, select *Add donor_projected*. Once this is done, a blue mesh sphere, called the base point, will appear around the N1 atom. A second sphere with solid colour will appear as well. This is the virtual point. The vector from the base to the virtual point follows the direction of the N-H group (from N to H).

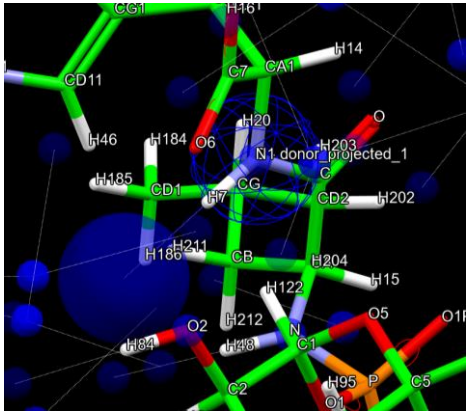
10





11



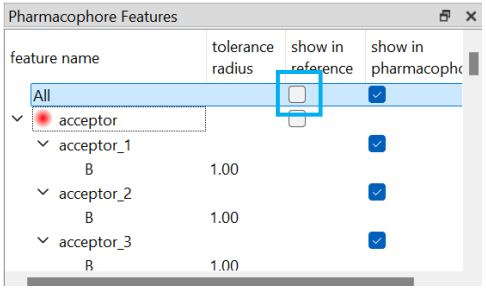


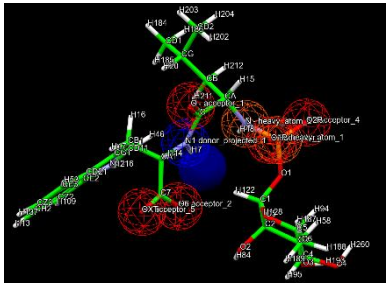
12. In the **Pharmacophore Features** window, untick *donor_projected* and tick *acceptor*. This time, you will see red spheres appear in the structure. For each of the following atoms: OXT, O6, O, O1P and O2P, add acceptors by right-clicking on a highlighted atom and clicking *Add acceptor*. You will notice that only mesh spheres appear in this instance. This is because it is not a projected feature and therefore consists of a base point only.

13. Untick *acceptor* in the **Pharmacophore Features** window and tick *heavy_atom* (a heavy atom is a non-H atom). Orange spheres will appear. Add heavy atoms features to *N* and *P*.

14. Once you have finished, turn off the display of pharmacophore features using the tick box in the *show in reference* column and *All* row. The result should look as shown below.

14

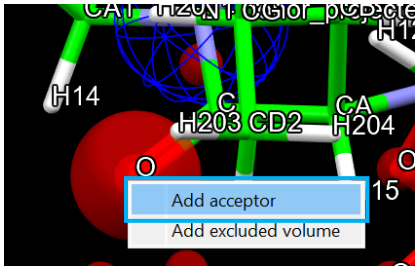


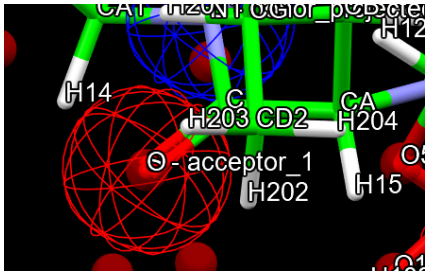


15. Click the **intra** button  . This will ensure that all pharmacophore features belong to the same molecule. Green dashed lines will appear between the base features (mesh spheres).

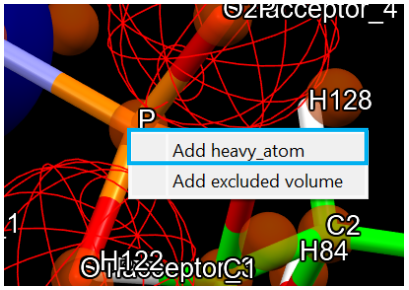
16. Save the pharmacophore query by going to the top menu and selecting *File > Save Pharmacophore*. Choose an appropriate location and name e.g. “1DMT.cm”and click **Save**. If you wish to continue this workshop at a later stage, you can re-open the pharmacophore from *File > Load Pharamcophore…*

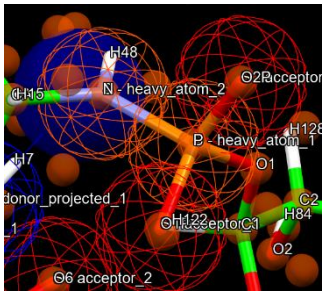
12



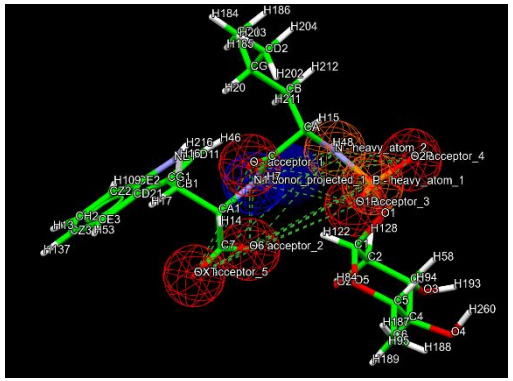


13

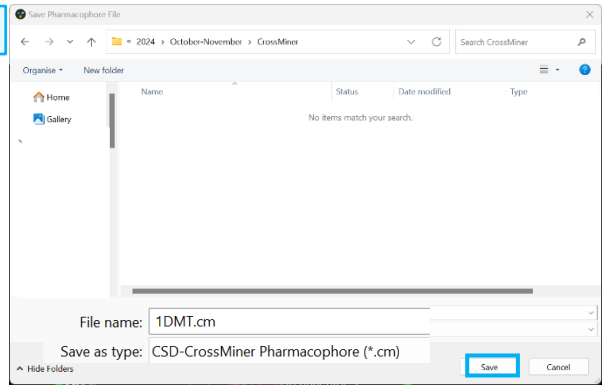




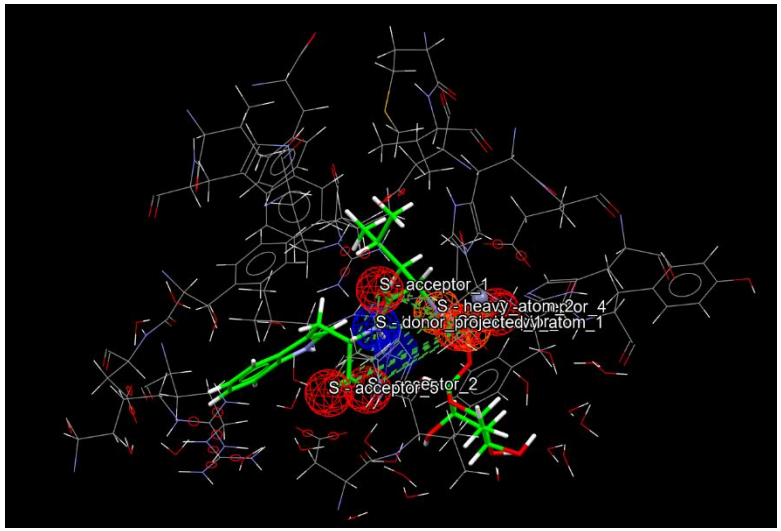
15



16



17. Unhide the remaining atoms in the structure by right-clicking in the visualizer (in an empty region) and selecting *Show/Hide > Show Atoms*. Additionally, turn off the atom labels by right clicking in the visualizer and selecting *Label > Do not label*. The result should look as shown below.



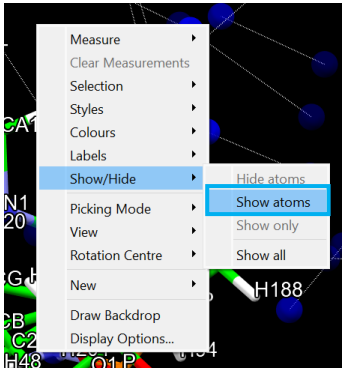
Running the CSD-CrossMiner Search

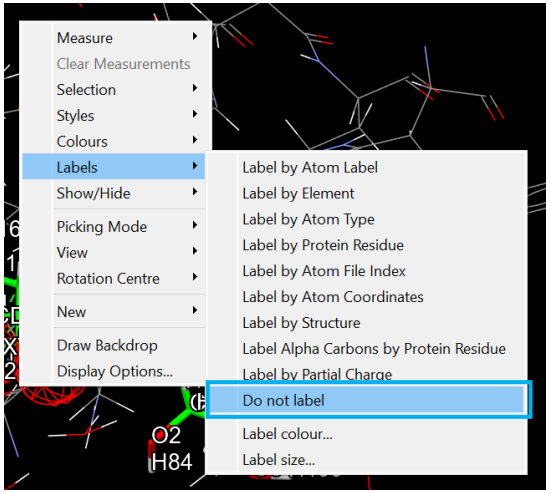
18. In the **Feature Database** window, untick all except the *pdb_crossminer* feature database.

19. From the top menu, Select *Edit > Options*. This brings up the search options window. Check that the options selected are as shown to the right. In particular, ensure that *Keep top (by rmsd) n matches per database entry* is enabled and *n* is set to **1,** and that Skip organometallic structures is selected. Press **OK**.

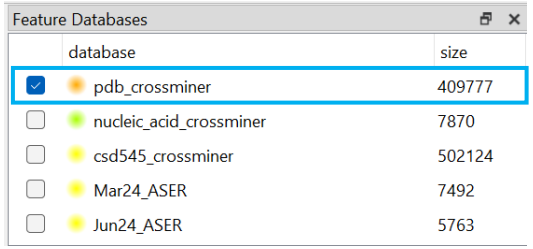
20. Press play  to begin the search. It may take several minutes to complete.

17

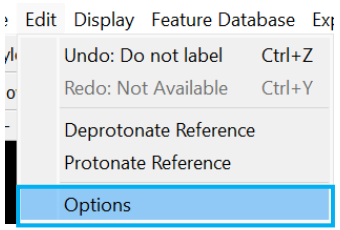


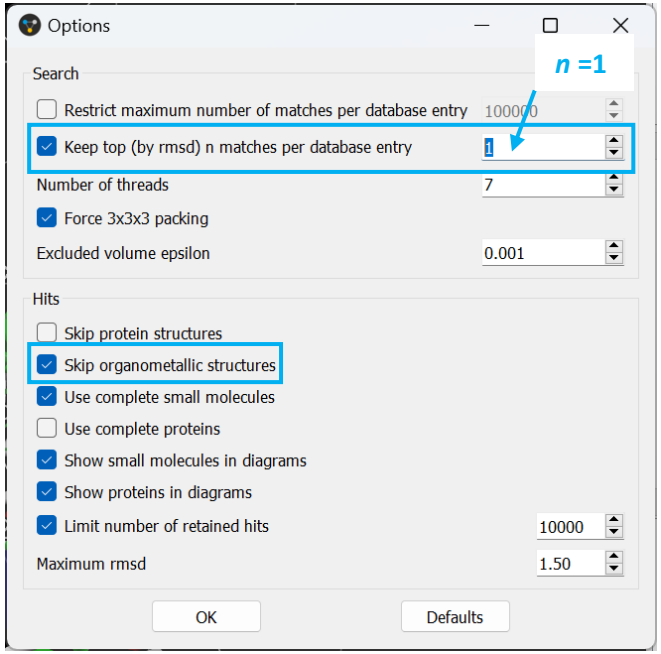


18



19





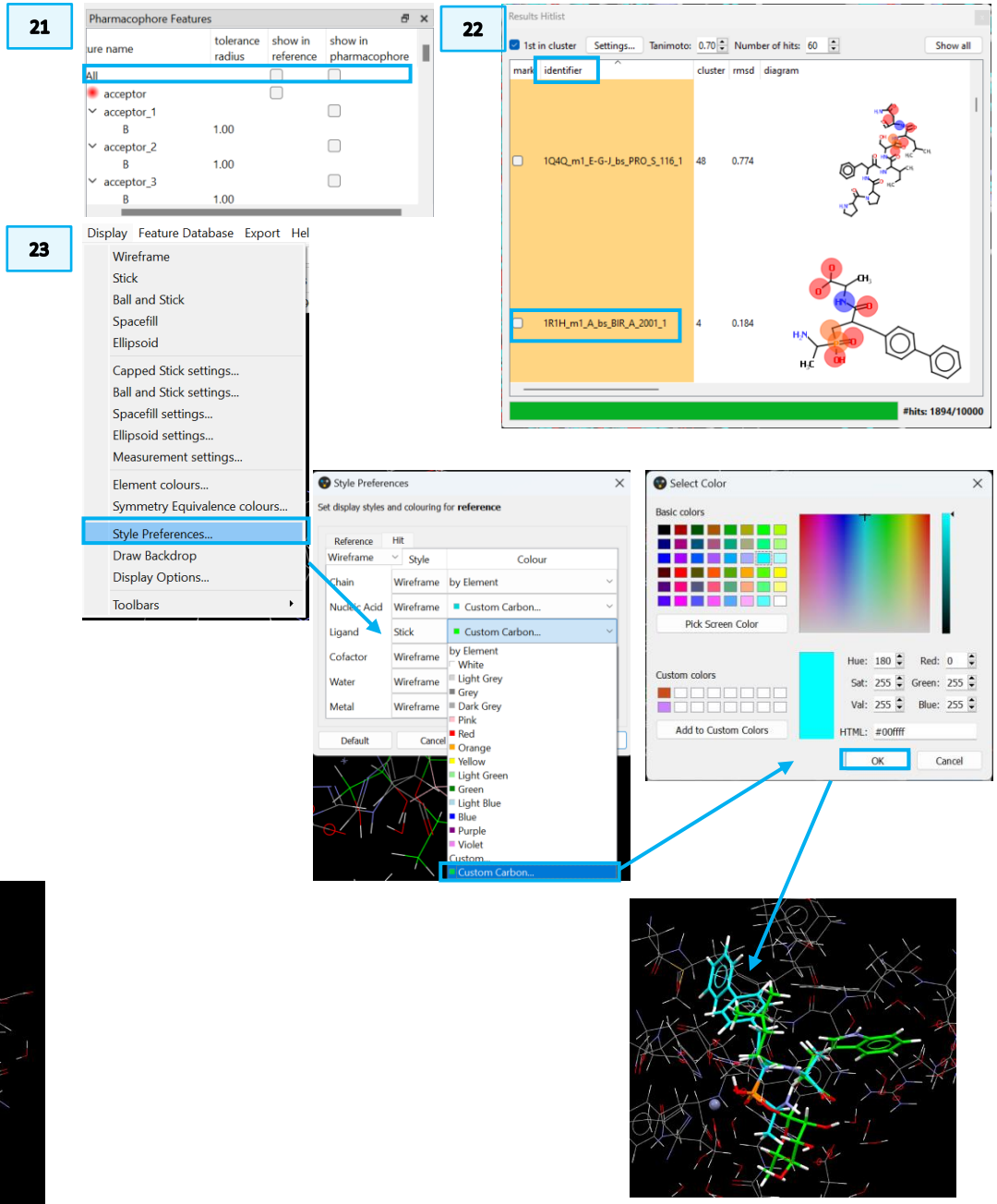
Inspecting the Search Results

21. The search results are displayed in the **Results Hitlist** window. There should be at least 1894 hits (based on CSD v. 5.45 + 3 updates). Before inspecting the results, untick *show in reference* and *show in pharmacophore* in the *All* row of the **Pharmacophore Features** window.

22. You can sort the entries in the results by clicking on the headings in the table. Click on the *identifier* column header to sort alphanumerically. Scroll through the results until you find "1R1H_m1_A_bs_BIR_A_2001_1" and click on it.

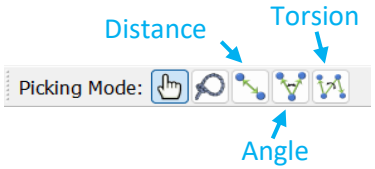
23. You will see this ligand overlaid with the reference structure. To make it easier to compare the structures, you can adjust the styles. From the top menu, select *Display > Style Preferences…*. Click the **Hit** tab in the **Style Preferences** window and next to *Ligand*, choose *Stick* from the *Style* dropdown menu, and *Custom Carbon* from the *Colour* dropdown menu. From the **Select Color** window choose your preferred colour. Here, we have picked cyan. Click **OK** and then in the **Style Preferences** window, click **Apply** and **OK** to close.

24. Explore the overlays to see how the molecules compare (numerically, their geometric fit to the pharmacophore is indicated by the *rmsd* in the **Results Hitlist**). If you wish, you can turn back on the pharmacophore display from the **Pharmacophore Features** window to see how closely matched the hit is. Which features common to these molecules are important for interaction with the protein? Which interactions are critical, given that both compounds bind to the target well?



25. Locate compound “1TMN_m1_E_bs_0ZN_E_317_1” in the **Results Hitlist**. This compound is similar to “1R1H_m1_A_bs_BIR_A_2001_1”. What structural patterns can you identify amongst these kinase inhibitors? To make it easier to see the amino acid residue, you can add labels by right-clicking and selecting *Labels > Label Alpha Carbons by Protein Residue*.

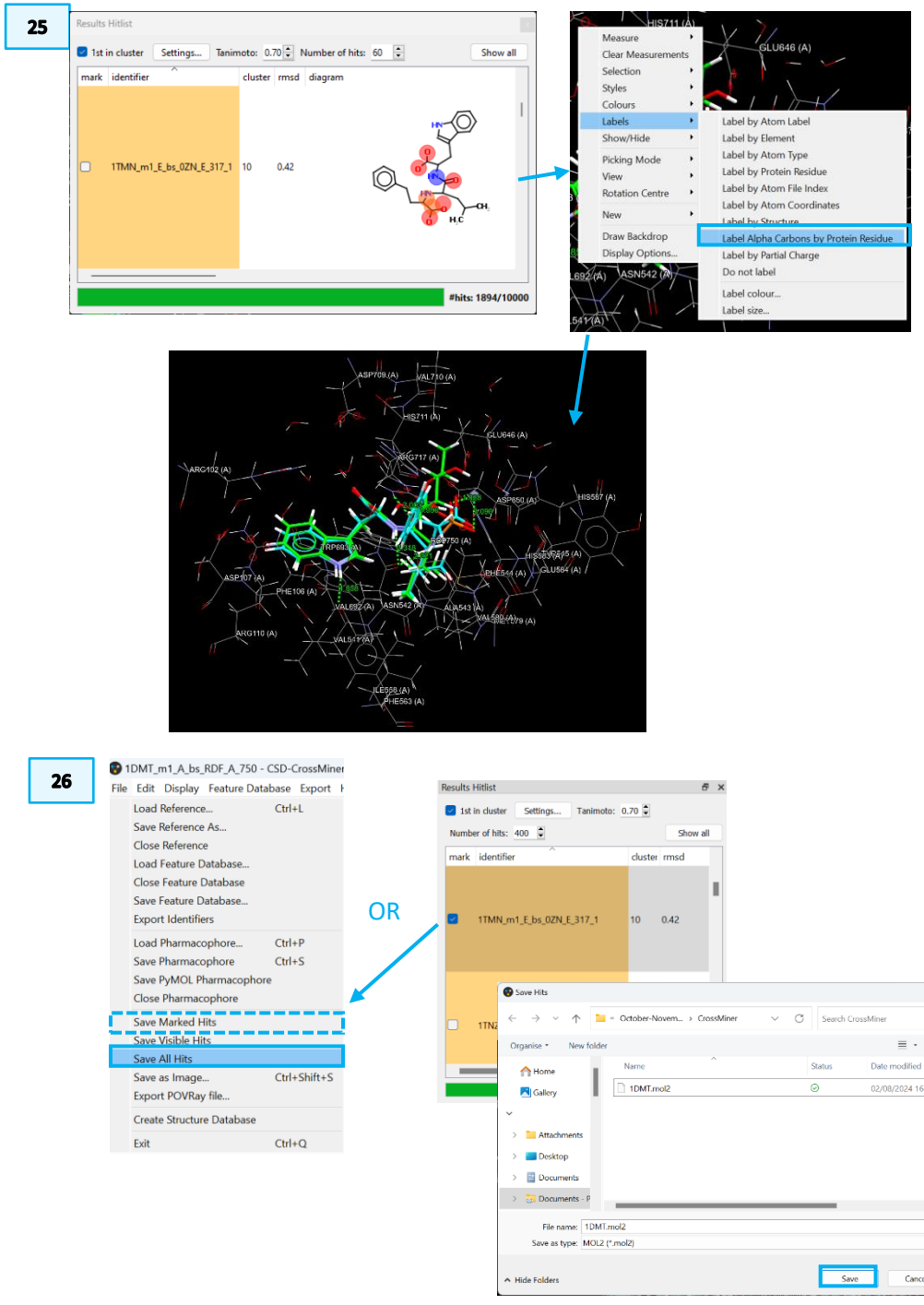
Tip: You can make measurements using the tools next to *Picking Mode* in the top toolbar. For example, click  and select pairs of atoms to measure distances between potential hydrogen bond donor/acceptor pairs.



26. You can save the hits that you have found in the search. There are several options to do this available from the *File* menu. For this example, we will choose *Save All Hits*. If you wish to retain only specific hits, you can select these in the **Results Hitlist** and choose *File > Select Marked Hits*.

Conclusion

Phosphoramidon is known to form a number of stabilising interactions with neprilysin, and several hydrogen bond and metal-ligand interactions can be identified within the structure of the complex.[1] Constructing a pharmacophore query in CSD-CrossMiner based on these features uncovers a number of candidate inhibitors, and amongst those with low RMSD, other experimentally verified inhibitors of the same protein can be identified.[2]



[1] D. M. Ferraris, D. Sbardella, A. Petrera, S. Marini, B. Amstutz, M. Coletta, P. Sander and M. Rizzi, *J. Biol. Chem.*, 2011, **286**, 32475.

[2] C. Oefner, B. P. Roques, M. -C. Foournie-Zaluska and G. E. Dale, *Acta Crystallogr. D*, 2004, **D60**, 392.

Summary

In this tutorial, you have seen how CSD-CrossMiner can be used to uncover potential inhibitors of an enzyme by constructing a pharmacophore query based on a reference structure. You should now:

- Be familiar with the CSD-CrossMiner interface
- Know how to add pharmacophore features to a reference molecule to build a pharmacophore query
- Be able to prepare and run a pharmacophore search using selected databases
- Know how to view and save search hits

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

Next Steps

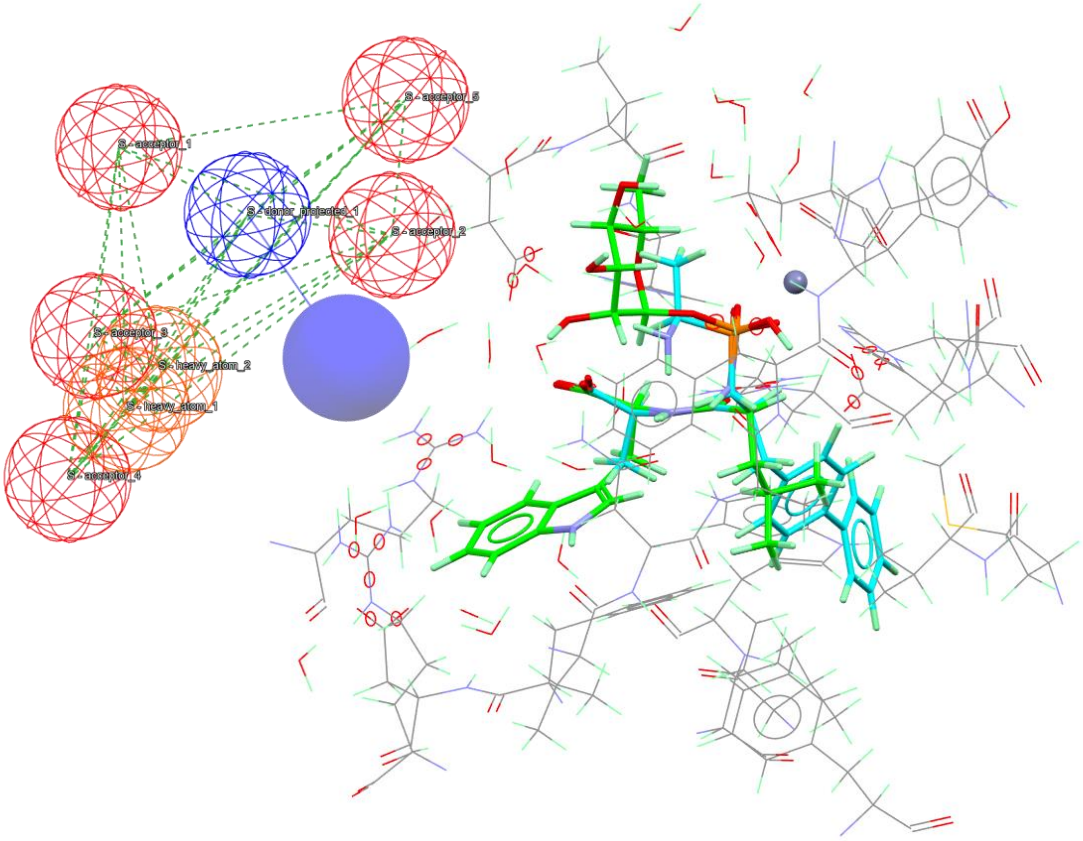
After this workshop, you can continue learning about CSD-CrossMiner with more exercises available in the self-guided workshops available in the CSD-Discovery workshops area on our website. You might also like to try our Pharmacophore Searching 101 – Introduction to CSD-CrossMiner on-demand module.

https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-discovery-workshops/

https://www.ccdc.cam.ac.uk/community/training-and-learning/csdu-modules/csd-crossminer-101/

Feedback

We hope this workshop improved your understanding of pharmacophore searching in CSD-CrossMiner 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-006. It will only take 5 minutes and your feedback is anonymous. Thank you!



Glossary

Binding Site

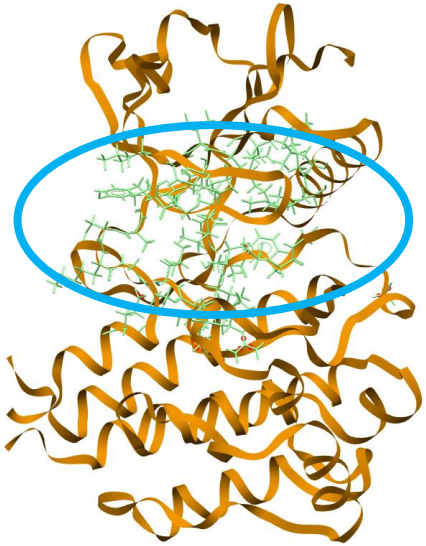
A specific region in a molecular entity that is capable of entering into a stabilising interaction with another molecular entity. An example of such an interaction is that of an active site in an enzyme with its substrate. Typical forms of interaction are by hydrogen bonding, coordination and ion pair formation. Two binding sites in different molecular entities are said to be complementary if their interaction is stabilizing. *Source*: PAC, 1994, 66, 1077. (Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)) on page 1089.

Hydrogen Bonds

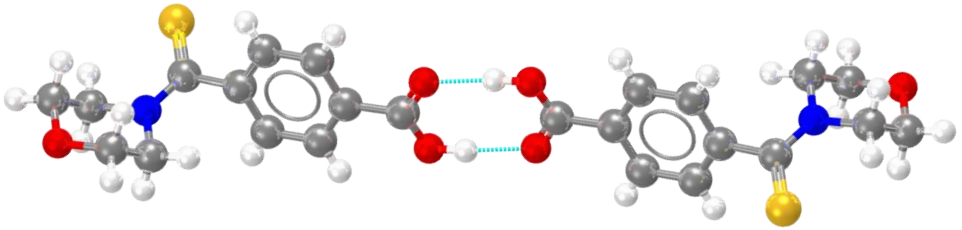
Hydrogen bonding occurs between donor-acceptor interactions precisely involving hydrogen atoms. The H-bonds interactions are classified as: strong (mostly covalent), moderate (mostly electrostatic) and weak (electrostatic). Their strength is observed to be between 12 and 30 kJ/mol.

Hydrogen Bond Donor/Acceptor

If a typical hydrogen bond is depicted as X—H···Y—Z, where the dots denote the bond, X—H represents the hydrogen bond *donor*. The *acceptor* may be an atom or anion Y, or a fragment of a molecule, Y—Z, where Y is bonded to Z. The acceptor is an electron-rich region such as, but not limited to, a lone pair on Y or a π-bonded pair of Y—Z. [Source: E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci and D. J. Nesbitt, *Pure Appl. Chem.*, 2011, **83**, 1637 – 1641.]



A protein displayed as a ribbon with its binding site represented in capped sticks style in green colour.



In light blue, example of hydrogen bonds for refcode MULWIC.

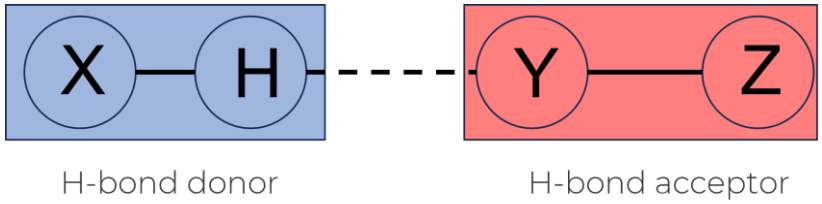


Illustration of a hydrogen bond interaction with between hydrogen bond donor X–H and hydrogen bond acceptor Y–Z.

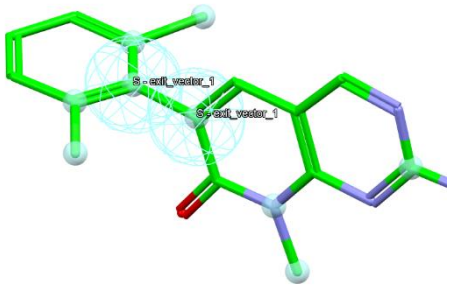
Root Mean Square Deviation (RMSD)

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.

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.



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

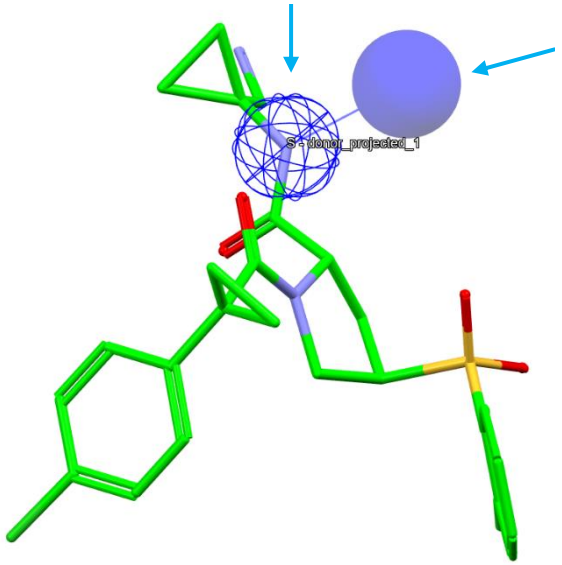
Features

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.

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.

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.

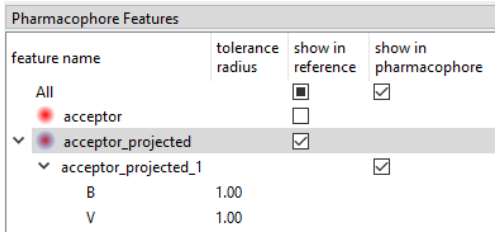
Features and Pharmacophore Representation

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:

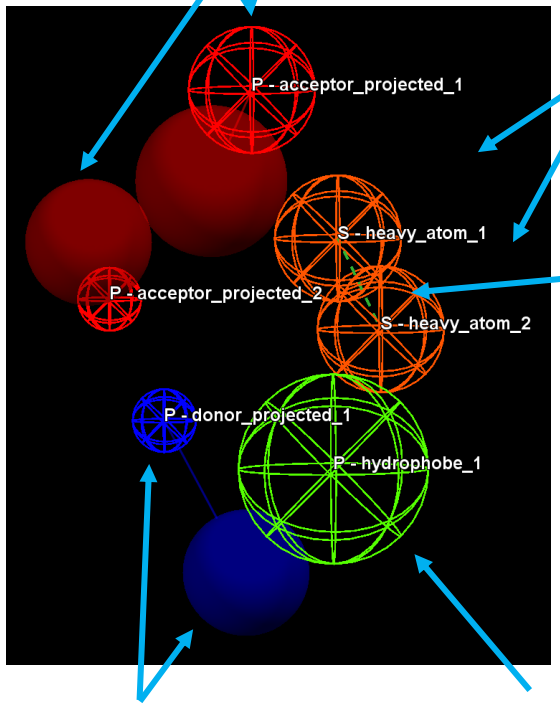
- **P**: Protein pharmacophore point
- **S**: Small molecule pharmacophore point
- **A:** Either a small molecule or protein pharmacophore point
- **Dashed line**: intra and intermolecular constraints. Constrained features must belong to either the same molecule as each other (*intra*, dashed green line) 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)



Small molecule (S): heavy atoms

Intramolecular constraint

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

Protein (P): hydrophobe feature