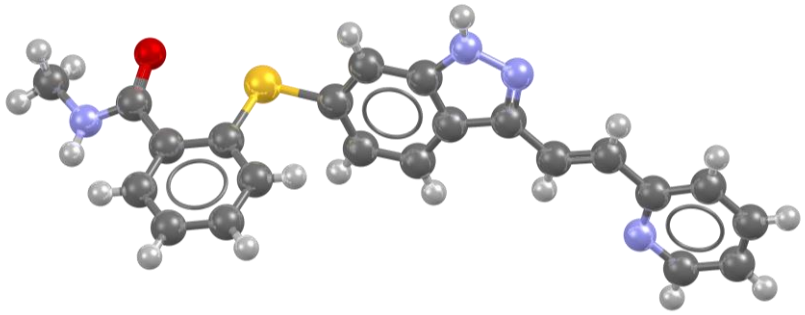
In-Depth Comparison of Polymorphic Structures Using Mercury



advancing structural science

Table of Contents

Introduction

Molecules have the potential to adopt multiple different packing arrangements in the solid state, a phenomenon known as polymorphism, which can have significant influence on a material's performance. Exploring the polymorphic landscape and understanding the relative stability of polymorphs is an important process, especially in early-stage formulation within the pharmaceutical industry. A comparison of polymorphic structures and assessment of relative stabilities should take into account molecular conformation, and intra- and intermolecular interactions. Mercury offers functionality to probe these features in a data-driven approach, using tools such as Mogul for molecular geometry, Hydrogen Bond Statistics (HBS), Hydrogen Bond Propensities (HBP) and Full Interaction Maps (FIMs) to explore hydrogen bonding, and Aromatics Analyser to assess aromatic interactions.

Before beginning this workshop, ensure that you have a registered copy of CSD-Materials or CSD-Enterprise installed on your computer. Please contact your site administrator or workshop host for further information.

Outcomes

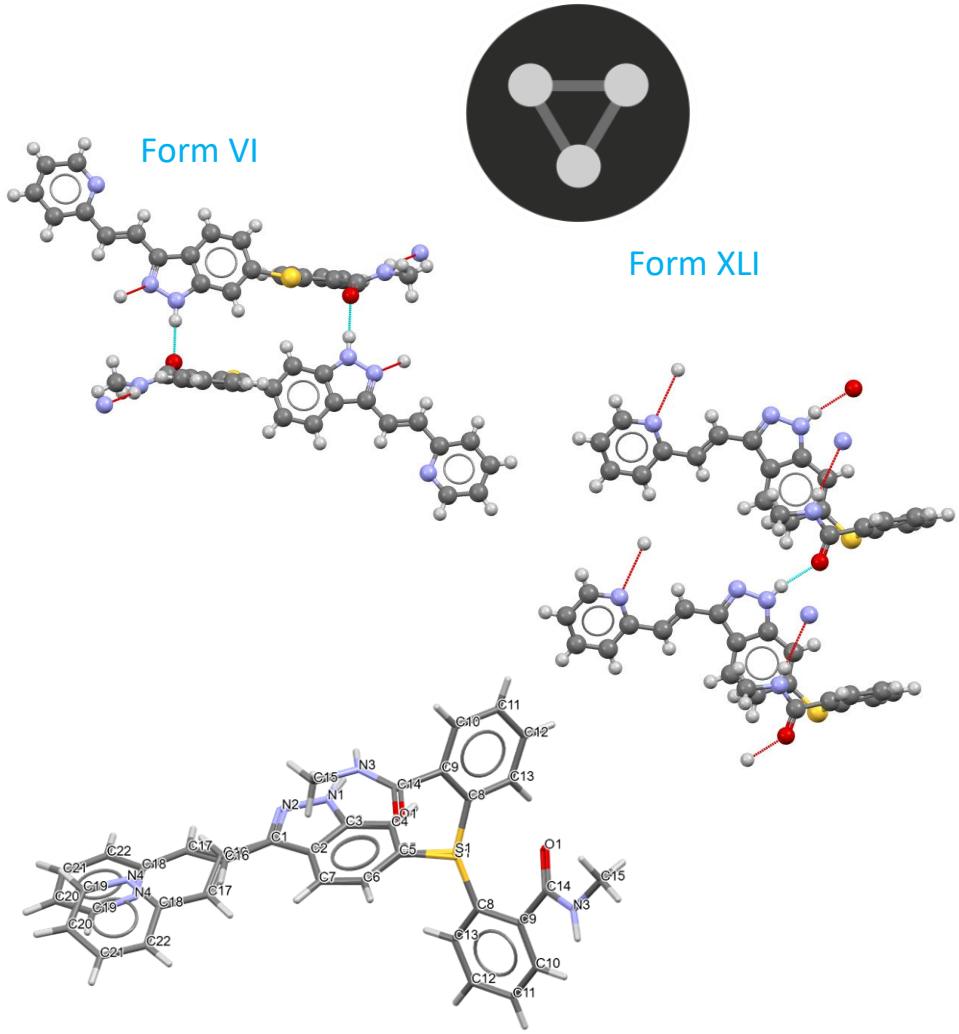
At the end of this workshop, you will:

- Be familiar with structure comparison using Mogul.
- Be able to calculate Hydrogen Bond Statistics and perform a Hydrogen Bond Propensities analysis and know how to read and interpret the results.
- Be able to run and interpret the results of Aromatics Analyser.
- Be able to use Full Interaction Maps as a complementary approach to assess solid forms.

Typically, several of the tools presented in this workshop would be used in a complementary manner for polymorph assessment. In the interest of time, we suggest you choose one or two of Parts 1-4 to complete during the "Hands-On" time. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout.

Materials

There are no additional materials required for this workshop.

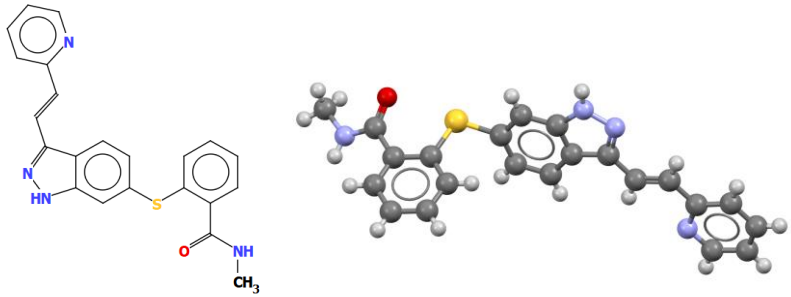


Pre-required skills

Familiarity with the Mercury interface is important; you can access the Visualization in Mercury self-guided workshop here (https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/csd-community-workshops/).

Examining polymorphs of Axitinib

N-methyl-2-((3-(2-(pyridin-2-yl)vinyl)-1*H*-indazol-6-yl)sulfanyl)benzamide, or Axitinib, is a tyrosine kinase inhibitor developed by Pfizer, approved for the treatment of renal cell carcinoma. It has been reported to exist in five *polymorphic* forms (I, IV, VI, XXV, XLI).[1] We will compare two of these polymorphs, Form VI and Form XLI using tools in Mercury.



N-methyl-2-((3-(2-(pyridin-2-yl)vinyl)-1*H*-indazol-6-yl)sulfanyl)benzamide *(refcode family **VUSDIX**).*

Part 1: Assessing conformation

We will begin by looking at the conformations of the two polymorphs of interest at a high level (structure overlay) and lower level (torsional data).

Overlaying structures

1. Start Mercury by double-clicking the icon on your Desktop or navigating from the Start Menu (*Start > CCDC > Mercury*).

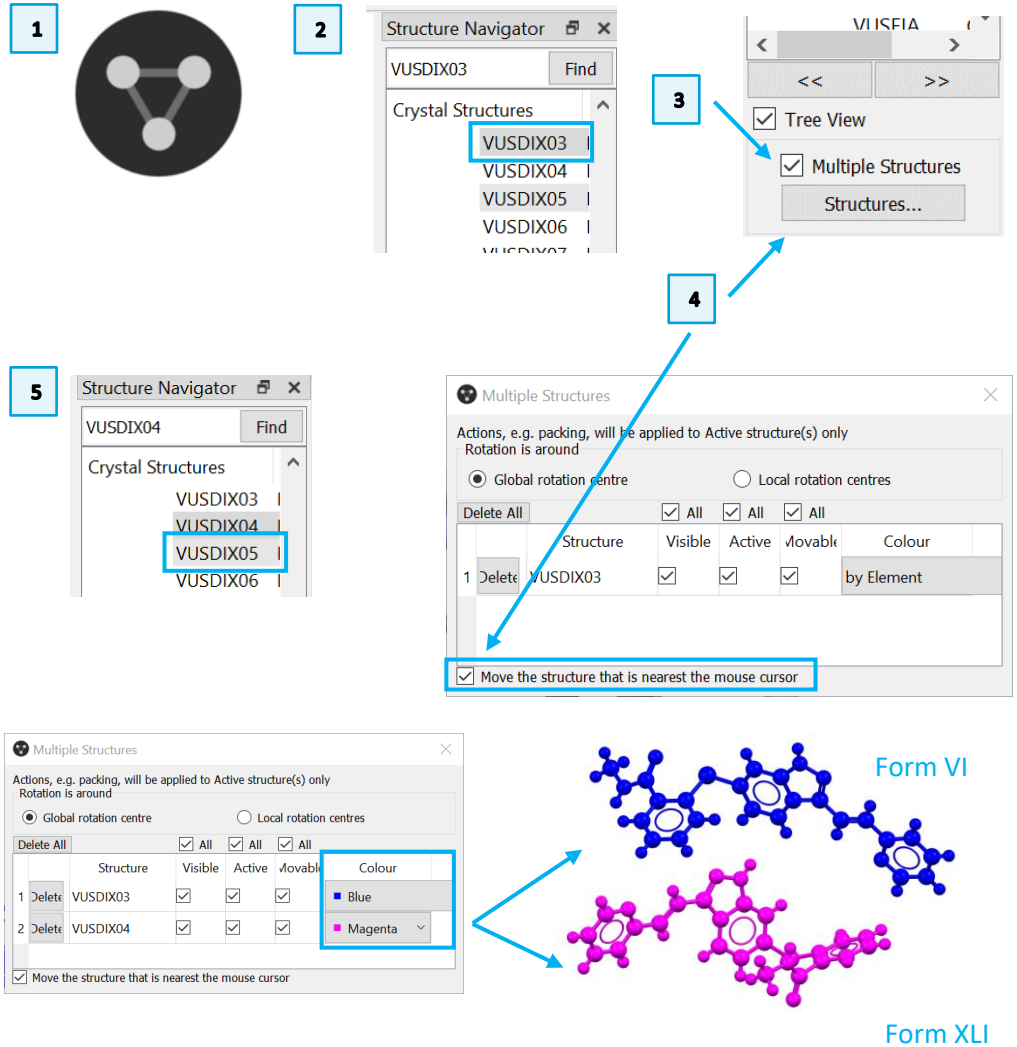
2. In the **Structure Navigator** window, type the refcode “VUSDIX03”, to bring up the structure of the first polymorph (Form VI).

3. At the bottom of the **Structure Navigator**, tick **Multiple Structures**

4. Click **Structures** to bring up the *Multiple Structures* Dialogue and Tick **Move the structure that is nearest the mouse cursor**.

5. In the Structure Navigator, select “VISDUX04” (Form XLI)

6. Both molecules will be displayed in the visualizer window. To help distinguish the two structures, you can select a colour for each from the *Colour* dropdown menus. Here we use *Blue* and *Magenta* for VUSDIX03 and VUSDIX04, respectively. You can change back to “by Element” at any time.



[1] M. Vasileiadis, C. C. Pantelides and C. S. Adjiman, *Chem. Eng. Sci.*, 2015, **121**, 60-76.

7. Tick **Show Labels for** *Non-hydrogen atoms* **with** *Atom Label* from the top tool window.

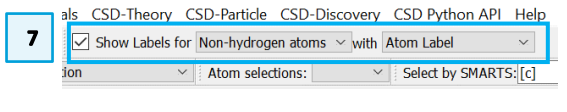
8. First, we will overlay the two molecules to get a feel for the *conformational* differences. You may prefer to use the *Capped Sticks* style for clarity.

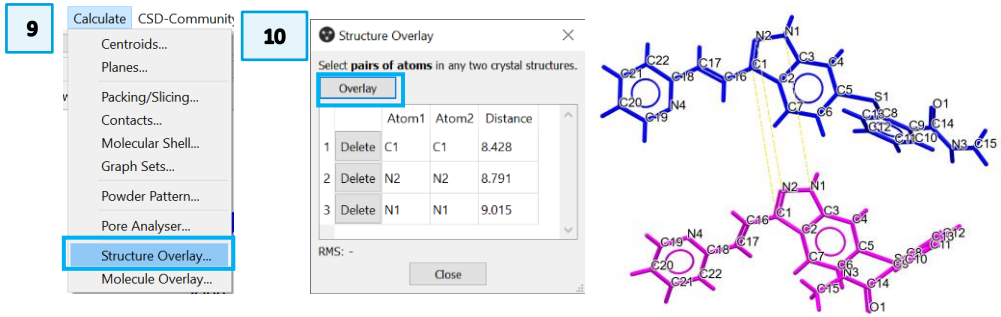
9. Select *Calculate > Structure Overlay* from the top menus to bring up the **Structure Overlay** window.

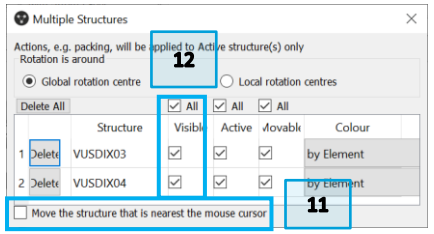
10. Select C1 in VUSDIX03 and then C1 in VUSDIX04. Repeat this process for N1 and N2. You should have three pairs. Click **Overlay**. This aligns the indazole ring which is rigid and common to both molecules.

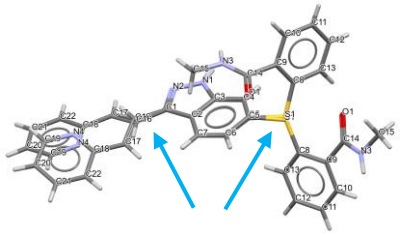
11. In the *Multiple Structures* dialogue untick **Move the structure that is nearest the mouse cursor**. You can now move the pair of overlaid structures around to examine the principal differences in their conformations.

12. Change the *Colour* back to *by Element* so that it is easier to distinguish the functional groups. The most obvious differences are the *s-cis* and *s-trans* relationships of the N2=C1 and C16=C17 double bonds, and the rotation about the C5–S1 bond. You can toggle on or off the visibility of each structure in the *Visible* column.

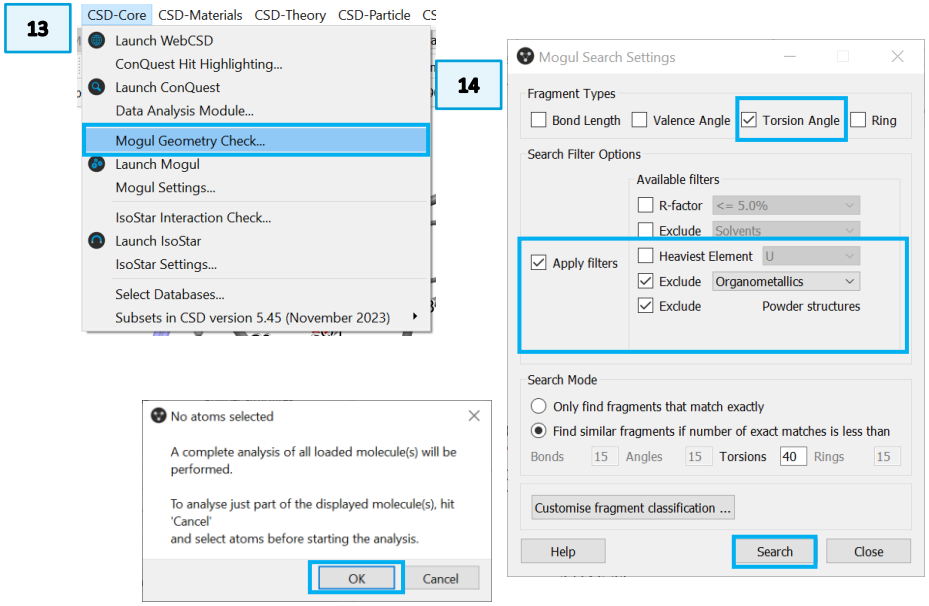








Large conformational differences



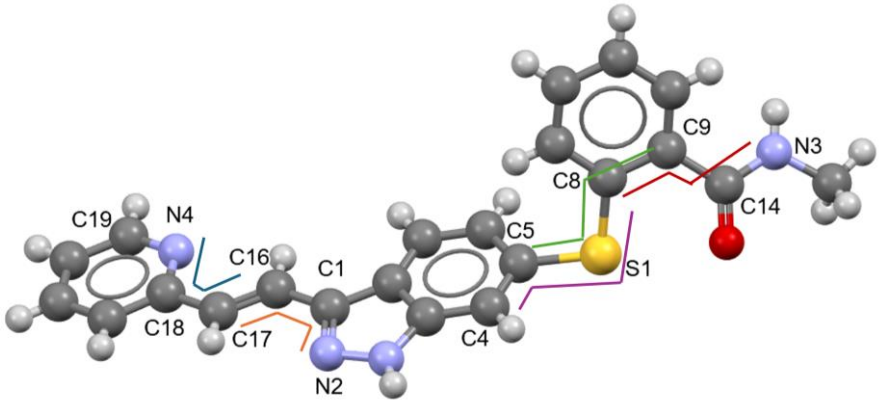
Running a Mogul Geometry Check

We can explore how usual or unusual the conformations are by looking at the torsions about rotatable bonds using Mogul Geometry Check.

13. Keeping both structures in the visualizer window, from the top menus select *CSD-Core > Mogul Geometry Check*.

14. Untick all except **Torsion Angle** in the *Fragment Types*, tick **Apply filters**, and **Exclude Organometallics** and **Exclude Powders**. Leave the other settings as the defaults and click **Search**. A warning will appear announcing that a comlete analysis of all molecule(s) will be performed. Click **OK** to let the calculation run.

15. In the **Mogul Results Viewer**, examine the following. Use the diagram and table below as a guide. You may also refer to the table for a summary.



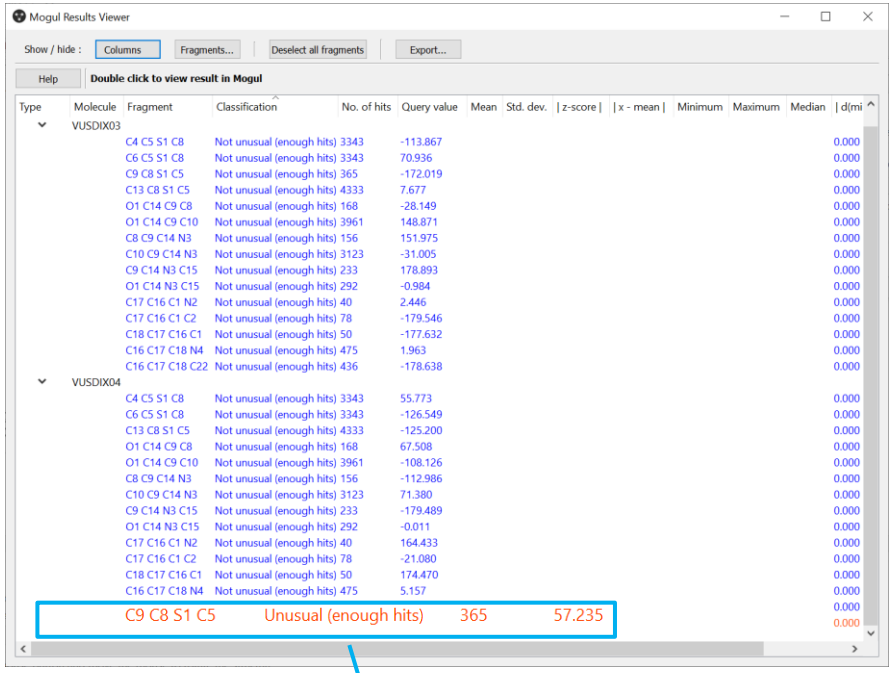
Torsion	VUSDIX03 / °	VUSDIX04 / °
C16-C17-C18-N4	1.963	5.157
C17-C16-C1-N2	−197.564	−21.080
C4-C5-S1-C8	−113.867	55.773
C9-C8-S1-C5	−172.019	57.235
C8-C9-C14-N3	−31.005	−112.986

16. Click on the torsions listed in **Step 15** in **the Mogul Results Viewer**. This will bring up a histogram of the values extracted from the CSD, together with the query value. Note that C9 C8 S1 C5 is highlighted in red in the results, indicating that this value is unusual. Notice, however, that the range of torsions observed is quite large.

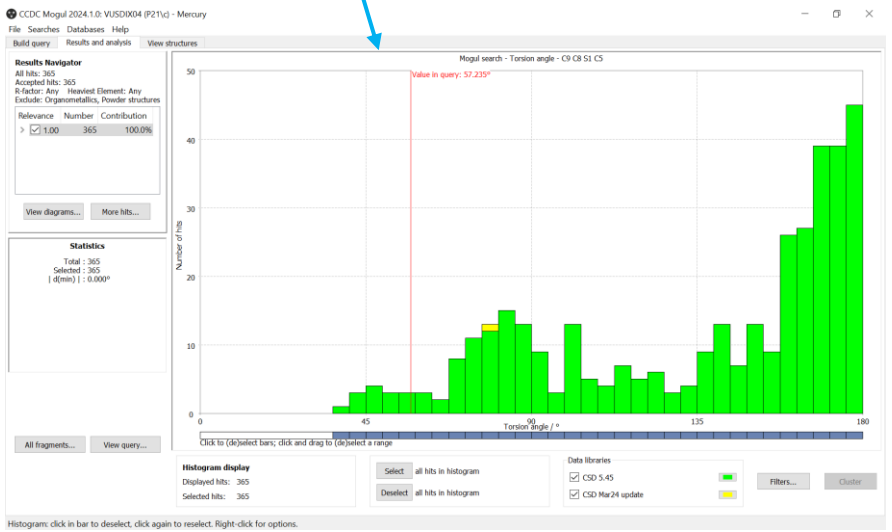
Conclusion

From the torsional data of VUSDIX03 and VUSDIX04, we can see that there is potential for considerable molecular flexibility. The occurrence of polymorphs with different conformations is perhaps unsurprising and whilst these conformations alone do not suggest differential stability, they do open the door to different intermolecular interactions, such as hydrogen bonding.

15



16



Part 2: Assessing hydrogen bonding

Here we will compare the use of simple Hydrogen Bond Statistics with more sophisticated Hydrogen Bond Propensities to assess hydrogen bonding in the polymorphs.

Calculating Hydrogen Bond Statistics

Hydrogen Bond Statistics offer a quick way to assess whether *hydrogen bonds* are unusual or not unusual.

1. Hydrogen Bond Statistics must be run on individual structures. In the *Multiple Structures* dialogue, click **Delete** next to VUSDIX04, so that only VUSDIX03 remains. Then, untick **Multiple Structures** in the **Structure Navigator**.

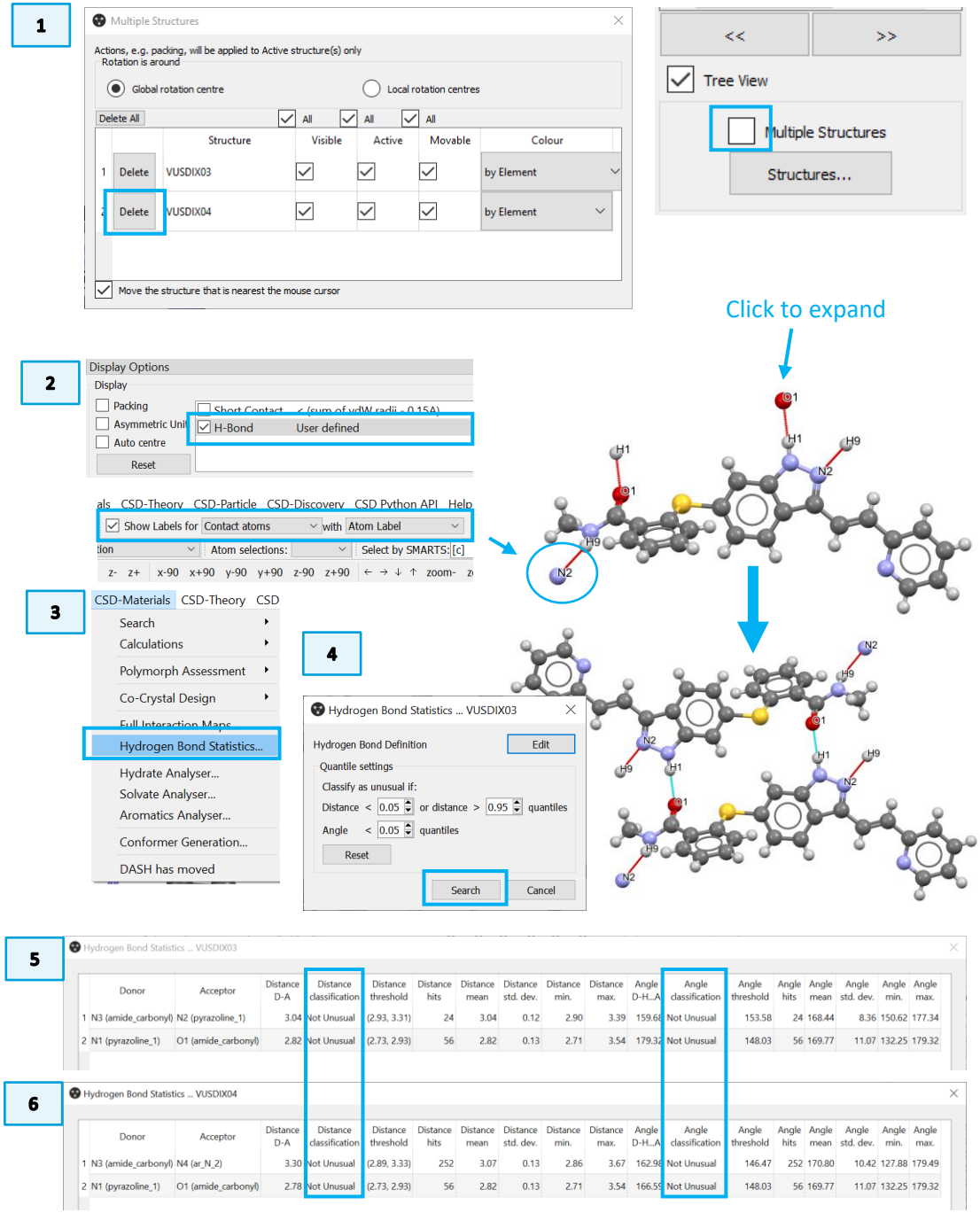
2. To view the hydrogen bonds in the structure, tick **H-Bond** in the **Display Options**. You may wish to change the labels to **Show Labels for** *Contact atoms* to make the display clearer. Red lines with hanging atoms will appear indicating intermolecular hydrogen bond interactions. Click on the lines to expand the hydrogen bonds.

3. From the top menus select *CSD-Materials > Hydrogen Bond Statistics*.

4. In the *Hydrogen Bond Statistics* dialogue, leave the settings as default and click **Search**.

5. Examine the results. Clicking on each row will display distance and angle histogram distributions and the corresponding heat plot. No distances or angles are classified as unusual for Form VI (VUSDIX03), though note that there is not much data for the N3···N2 interaction.

6. In the **Structure Navigator**, select “VUSDIX04” and repeat **Steps 2-5** and compare the results.



Calculating H-bond propensity

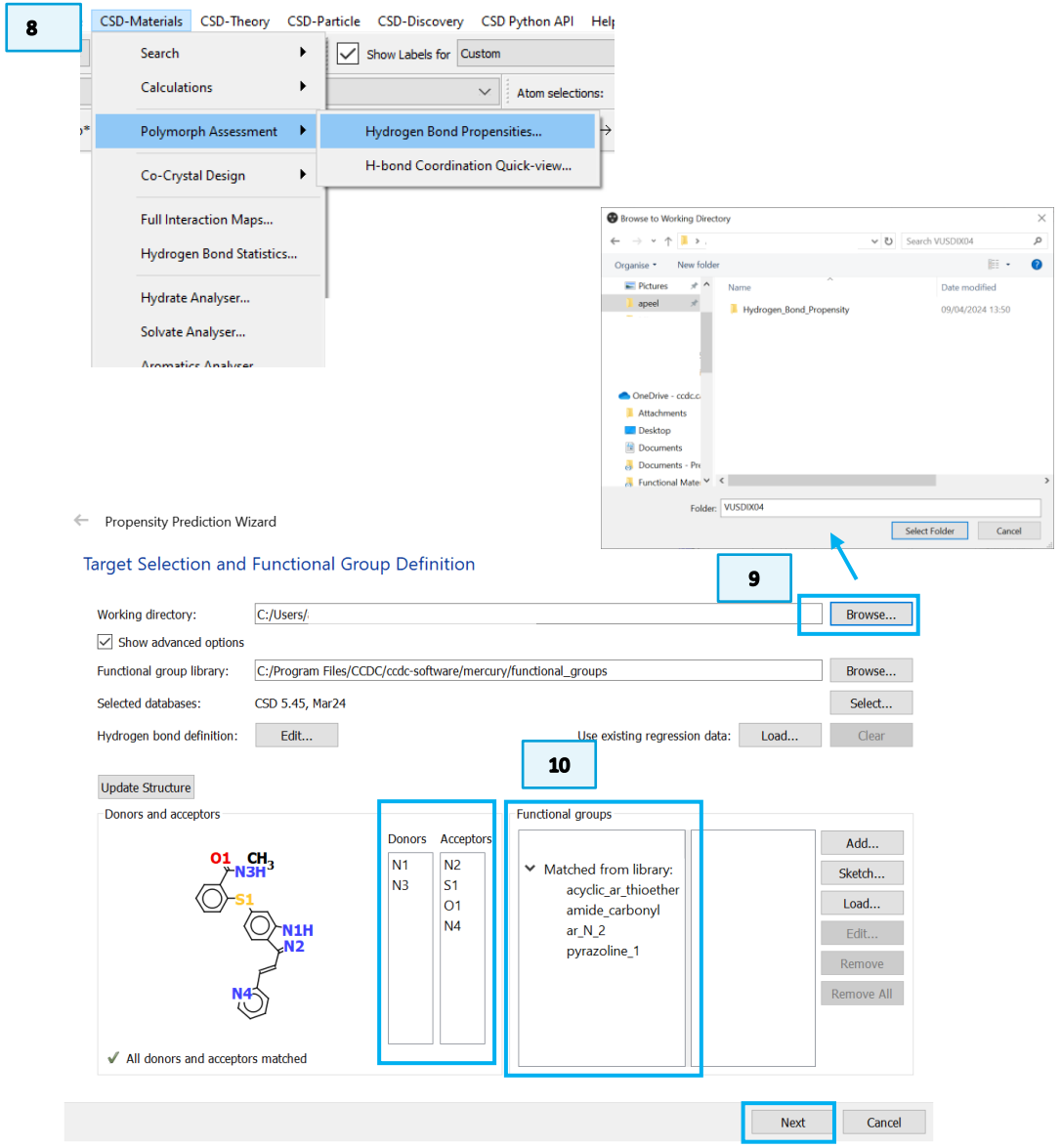
Hydrogen Bond Propensities offers an assessment of potential hydrogen bond networks based on the likelihood of functional groups interacting in a particular way, together with the observation that higher hydrogen coordination numbers lead to more stable networks.

7. You should still have VUSDIX04 selected from **Step 6**. If not, select it using the **Structure Navigator**.

8. From the top-level menu select *CSD-Materials > Polymorph Assessment > Hydrogen Bond Propensities…*

9. In the *Propensity Prediction Wizard* select a working directory by clicking on **Browse**. The potential *hydrogen bond donor* and *acceptor atoms* are automatically identified and linked to their functional groups. Two donors have been identified: N1 as pyrazoline_1 and N3 as amide_carbonyl. Four acceptors have also been identified: N2 as pyrazoline_1, S1 as acyclic_ar_thioether, O1 as amide_carbonyl, and N4 as ar_N_2.

10. The *Donors* and *Acceptors* atoms can be highlighted in the 2D chemical diagram by selecting them from the list. You can also highlight a functional group from the *Matched from library* list; the corresponding atoms will be automatically highlighted in the *Donors/Acceptors* lists. The functional group as defined will appear in the second window of the *Functional groups* dialogue box. You can adjust the functional groups if desired by using the buttons on the right-hand side **Add…**, **Sketch…**, etc. We will leave all the default settings for this example and click **Next**.



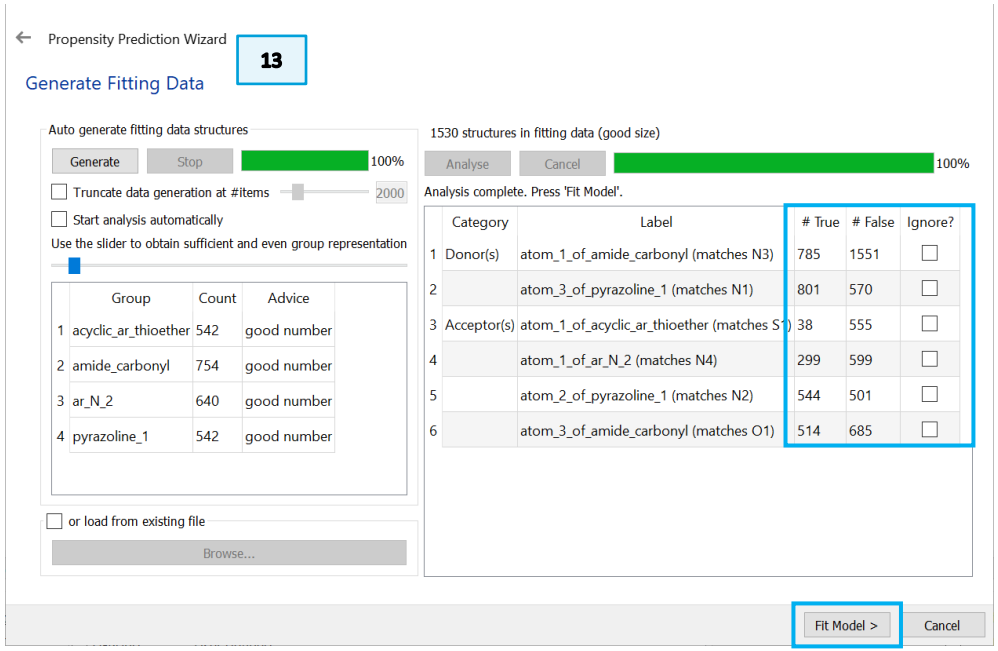
Tips and tricks

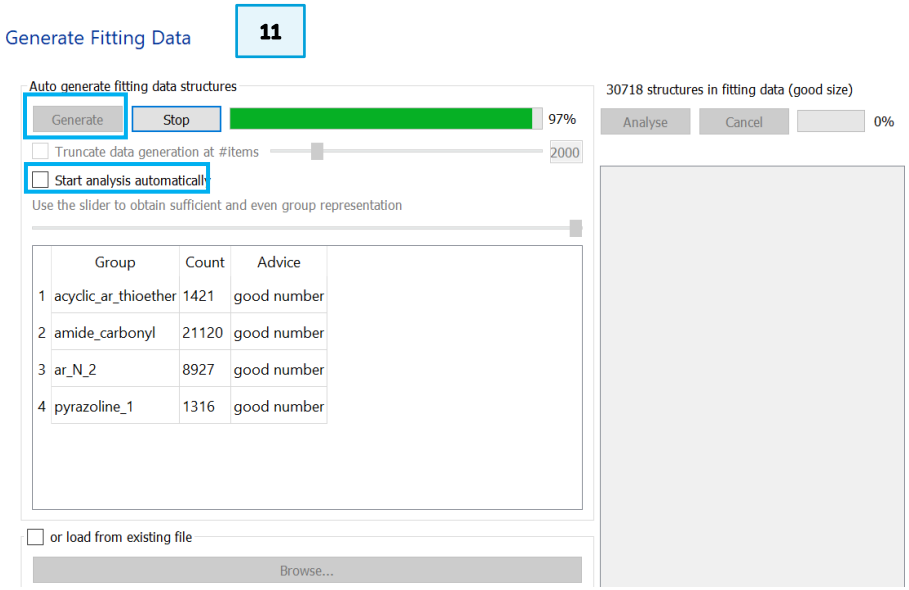
If you want to adjust the atoms involved as donors or acceptors, you can use the advanced settings: toggle on the **Show advanced options** check box and click **Edit…**

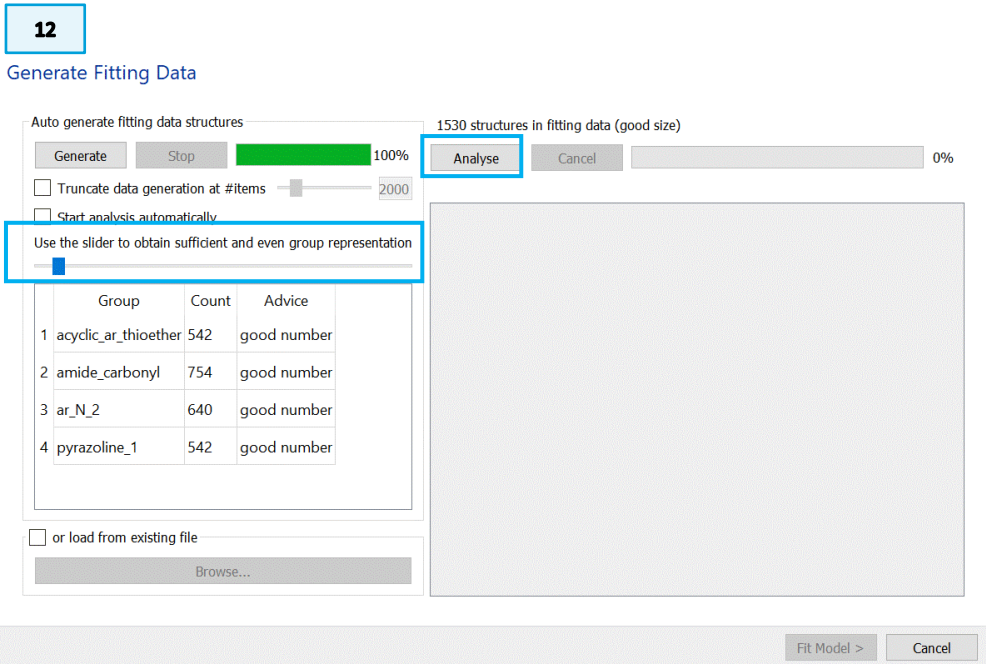
11. Ensure that the **Start analysis automatically** check box is unchecked and click **Generate**. As the training set (generated fitting data) starts to be populated with CSD structures, the functional groups and an indication of their *Count* and *Advice* can be seen.

12. When the run is finished, it attempts to automatically select a sufficient number of hits (count) per functional group with fairly even representation across the groups. In general, around 500 structures per functional group should be enough. The group numbers can be adjusted by using the slider highlighted in blue. This allows you to remove or add structures until a more even set of data, or more appropriate number of groups, is obtained. Select 1530 structures in total using the slider (your number may be slightly different depending on which CSD version you have installed). Click **Analyse**.

13. When the analysis is finished, the number of the True and False outcomes will be listed. If there are very low numbers for True or False, you should check that they are ticked in the **Ignore?** checkboxes. There are no very low values in this example. Click the **Fit Model >** button to continue.





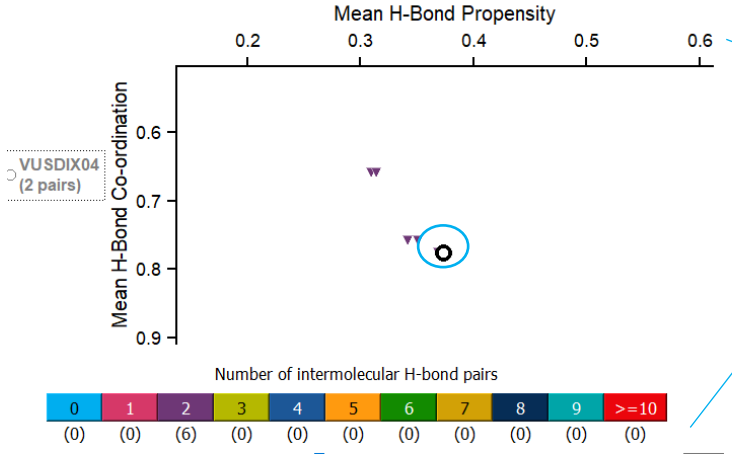


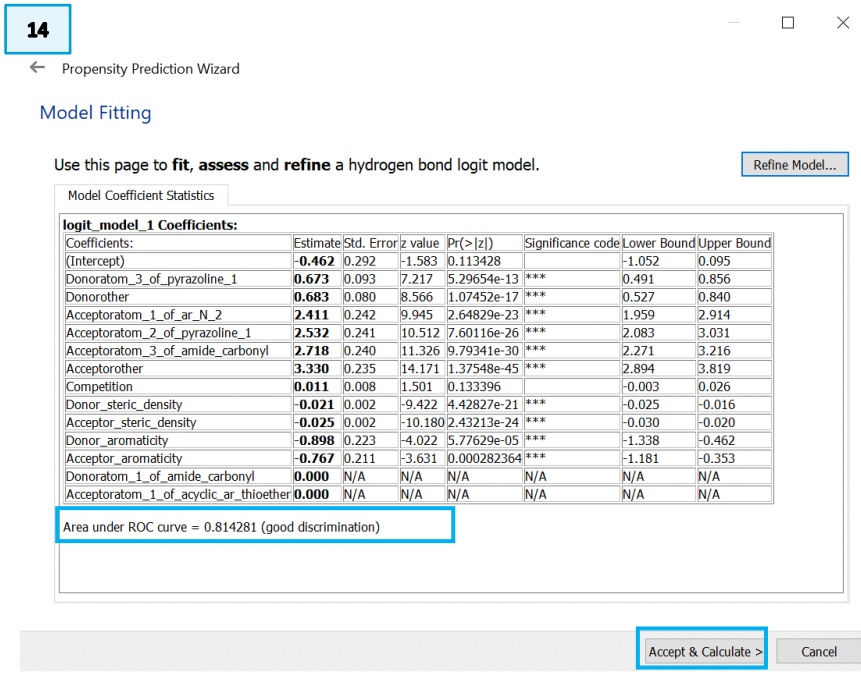
14. For this example, the Area under *ROC* (Receiver Operating Characteristic) curve should be around 0.81. To achieve a good H-bond propensity calculation you should always aim for a value of around 0.75 or above. Click **Accept & Calculate** to continue.

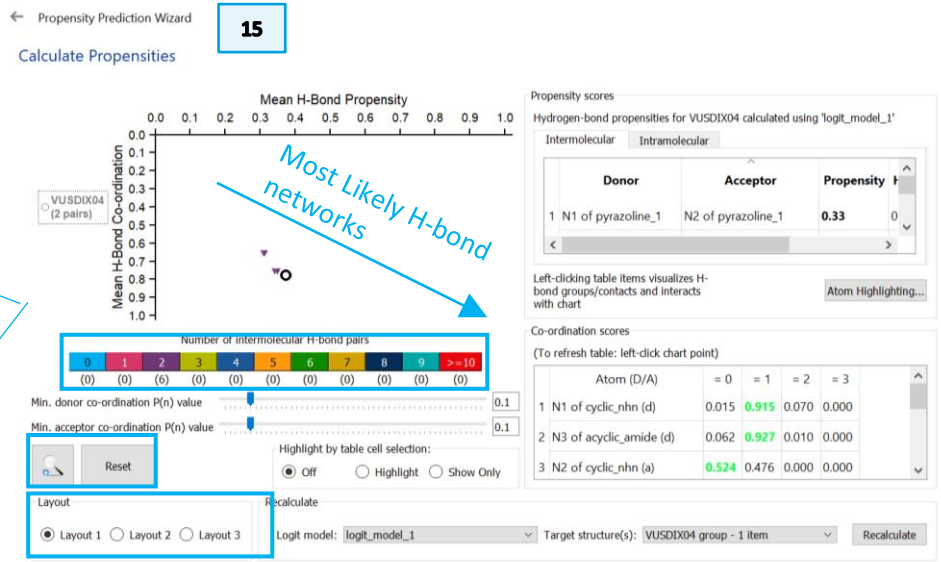
Summary of HBP results

We have now obtained the results of our HBP calculation, and we can analyse them in the graph and tables displayed. Three different layouts are available in the *Layout* section and can be selected by pressing the radio buttons; layout 1 is more convenient for viewing the graph, whereas layout 3 is preferable for viewing the tables, and layout 2 is a nice balance for viewing both the graph and tables. The overall size of the window can be changed by dragging the corners or edges.

15. *The Chart:*
    - Plots *Mean H-bond Propensity* *vs* the *Mean H-Bond Co-ordination*.
    - The target structure is represented as a black circle with a white interior.
    - To zoom, use the magnifying glass icon in the lower left-hand corner of the wizard, left click and drag on the area to zoom in on it. To go back to the default view, press **Reset**.
    - To filter the chart for a given number of H-Bond pairs, use the colour legend.
    - The most likely H-bonding network is displayed towards the lower-right corner, the outcome should be read along the diagonal.
    - VUSDIX04 has the most likely H-bonding network.
    - Click on the points to highlight the network in the *Propensity score* table.
    - Hover over a point to display the mean propensity and mean co-ordination values.







16. *Propensity Scores Table:*
    - Select **Layout 2** or **Layout 3** to see the full Propensity Scores table. The table can be expanded horizontally if needed by dragging the double-headed arrow that appears when hovering over the border between the propensity and co-ordination tables.
    - The most likely H-bond pair will score the highest propensity.
    - The H-bonds present in the targeted structure are marked as observed.
    - The table is interactive, clicking on **observed**, which is located at the far right-hand side of the table will highlight the donor and acceptor group in the 3D visualizer, while clicking on an atom label, in either the *Donor* or *Acceptor* columns, will highlight the functional group and label the atom in the 3D visualizer.
    - The *Propensity scores* table shows all possible H-bond interactions for the molecule, with N1-H1···O1 giving the highest propensity. You can see this interaction is observed in the VISDUX04 structure.

17. *Co-ordination Scores Table:*
    - (d) stands for donor and (a) for acceptor.
    - =0, = 1, = 2 etc. denote the number of times a functional group donates or accepts.
    - The numbers that are coloured relate to the outcome present in the selected H-bonding network; if this is green it indicates that the outcome is optimal, whereas if it is red that indicates the outcome is sub-optimal.
    - For VISDUX04, for the observed H-bonds, the coordination numbers are all optimal.

18. To summarise our observations, Form XLI (VUSDIX04) is represented as a white circle in the propensity chart; and the propensity and coordination number scores indicate that the observed hydrogen bonding network is optimal. We can now add VUSDIX03 and compare the two.

16

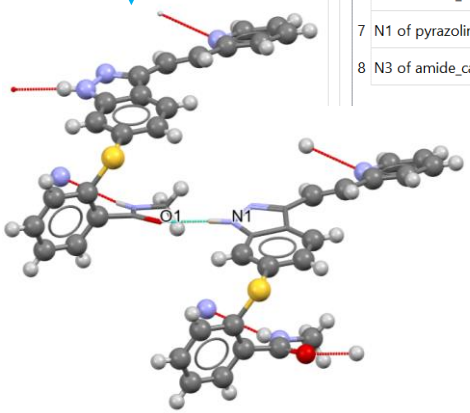
Propensity scores

Hydrogen-bond propensities for VUSDIX04 calculated using 'logit_model_1'

Intermolecular | Intramolecular

	Donor	Acceptor	Propensity	H-Bonds	Donor ranking	Acceptor ranking	VUSDIX04
1	N1 of pyrazoline_1	N2 of pyrazoline_1	**0.33**	0	1	3	
2	N3 of amide_carbonyl	N2 of pyrazoline_1	**0.24**	0	2	3	
3	N1 of pyrazoline_1	N4 of ar_N_2	**0.39**	0	1	2	
4	N3 of amide_carbonyl	N4 of ar_N_2	**0.28**	1	2	2	observed
5	N1 of pyrazoline_1	O1 of amide_carbonyl	**0.46**	1	1	1	observed
6	N3 of amide_carbonyl	O1 of amide_carbonyl	**0.35**	0	2	1	
7	N1 of pyrazoline_1	S1 of acyclic_ar_thioether	**0.02**	0	1	4	
8	N3 of amide_carbonyl	S1 of acyclic_ar_thioether	**0.01**	0	2	4	

Observed



17

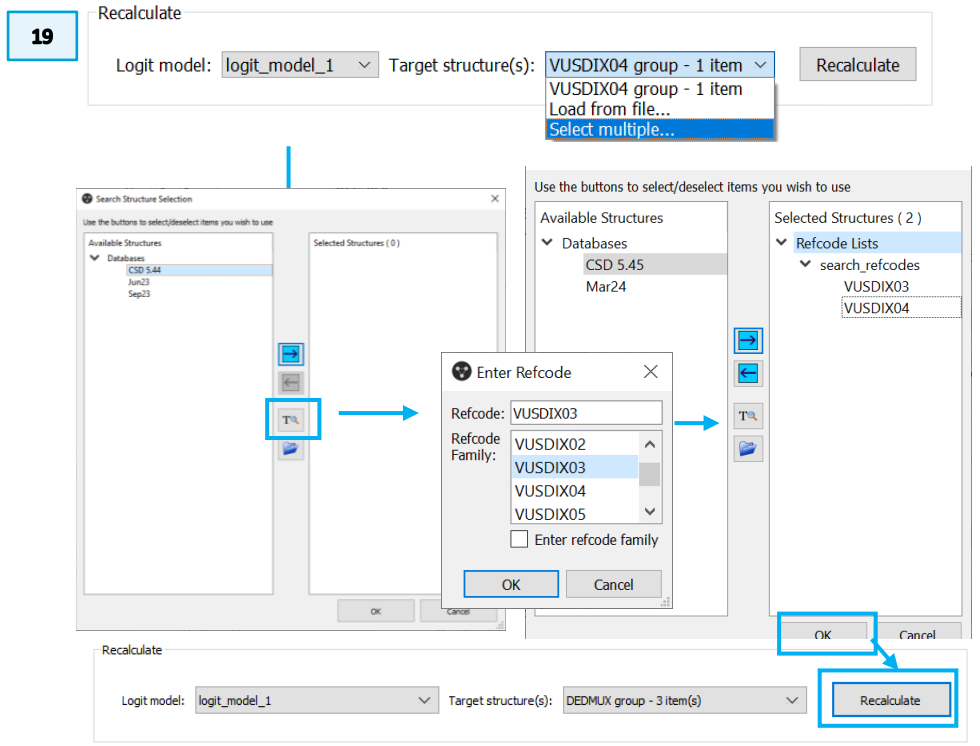
Co-ordination scores

(To refresh table: left-click chart point)

	Atom (D/A)	= 0	= 1	= 2	= 3
1	N1 of cyclic_nhn (d)	0.015	**0.915**	0.070	0.000
2	N3 of acyclic_amide (d)	0.062	**0.927**	0.010	0.000
3	N2 of cyclic_nhn (a)	**0.524**	0.476	0.000	0.000
4	N4 of ar_n (a)	0.463	**0.519**	0.019	0.000
5	O1 of acyclic_amide (a)	0.148	**0.803**	0.048	0.001
6	S1 of acyclic_ar_thioether (a)	**0.965**	0.035	0.000	0.000

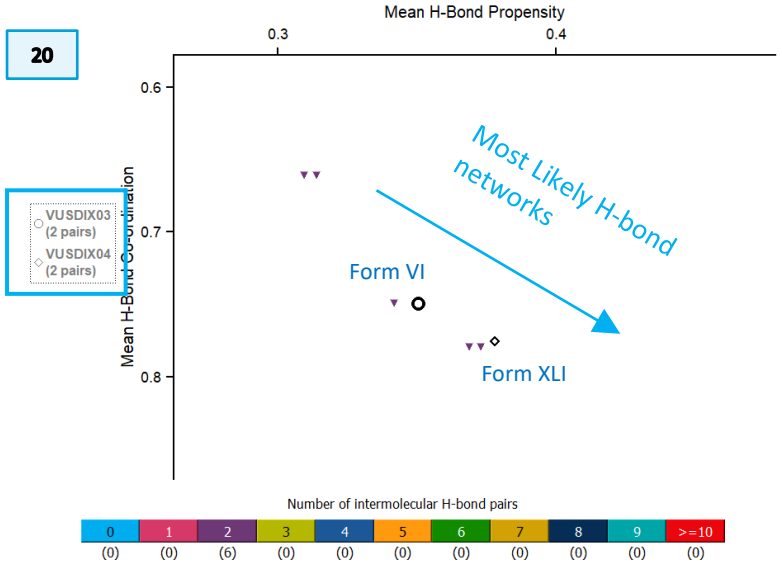
19. To see where Forms VI (VUSDIX03) and XLI (VUSDIX04) are located in the chart, you can load them by clicking *Target structure(s)* drop-down menu in the *Recalculate* section and then click *Select multiple…* In the *Search Structure Section* dialog box, click the **T** icon, click VUSDIX03, then click **OK**. Repeat this to add VUSDIX04. You can see the two VUSDIX refcodes in the **Selected Structures** pane. Click **OK**, then click **Recalculate**.

20. Both polymorphs are now plotted on the chart. To identify where each polymorph is represented on the chart, check the legend shown on the left-hand side of the dialogue indicating the structures displayed. You can see that Form VI (VUSDIX03) has worse Mean H-Bond Propensity and Mean H-Bond Co-ordination. In the *Propensity scores*, whilst Form VI and XLI share the highest propensity pair, Form VI has a second pair whose propensity is lower than the second of Form XLI. Similarly, in the *Co-ordination scores* table, N2 of cyclic_nhn (a) has coordination score of 1, whereas it would prefer to be uncoordinated (=0) and is therefore sub-optimal.



Conclusion

The Hydrogen Bond Statistics have revealed nothing unusual about the two polymorphs under investigation. However, based on the Hydrogen Bond Propensity calculations, the hydrogen bonding in Form XLI (VUSDIX04) is optimal whereas that in Form VI (VUSDIX03) is not; it would be predicted to be less stable on this basis.



Hydrogen-bond propensities for VUSDIX03 calculated using 'logit_model_1'

Intermolecular | Intramolecular

	Donor	Acceptor	Propensity	H-Bonds	Donor ranking	Acceptor ranking	VUSDIX03	VUSDIX04
1	N1 of pyrazoline_1	O1 of amide_carbonyl	**0.46**	1	1	1	observed	observed
2	N1 of pyrazoline_1	N4 of ar_N_2	**0.39**	0	1	2		
3	N3 of amide_carbonyl	O1 of amide_carbonyl	**0.35**	0	2	1		
4	N1 of pyrazoline_1	N2 of pyrazoline_1	**0.33**	0	1	3		
5	N3 of amide_carbonyl	N4 of ar_N_2	**0.28**	0	2	2		observed
6	N3 of amide_carbonyl	N2 of pyrazoline_1	**0.24**	1	2	3	observed	
7	N1 of pyrazoline_1	S1 of acyclic_ar_thioether	**0.02**	0	1	4		
8	N3 of amide_carbonyl	S1 of acyclic_ar_thioether	**0.01**	0	2	4		

Co-ordination scores

(To refresh table: left-click chart point)

	Atom (D/A)	= 0	= 1	= 2
1	N1 of cyclic_nhn (d)	0.015	**0.914**	0.070
2	N3 of acyclic_amide (d)	0.015	**0.972**	0.013
3	N2 of cyclic_nhn (a)	0.598	**0.402**	0.000
4	N4 of ar_n (a)	**0.498**	0.484	0.018
5	O1 of acyclic_amide (a)	0.230	**0.745**	0.025
6	S1 of acyclic_ar_thioether (a)	**0.965**	0.035	0.000

Part 3: Assessing aromatic interactions

Another factor affecting polymorph stability in aromatic compounds is the interactions between the aromatic groups. There are two phenyl groups that could form potential *aromatic interactions* in this molecule.

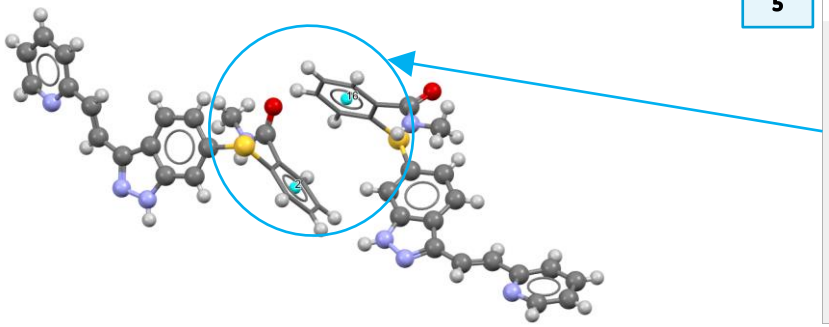
1. If it is not already loaded, select "VUSDIX04" in the **Structure Navigator**.

2. From the top menus, select *CSD-Materials > Aromatics Analyser*.

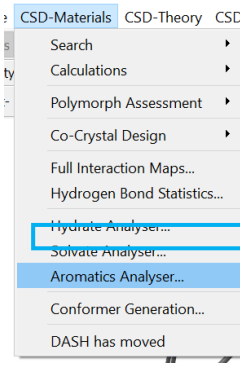
3. Select all of the atoms in the molecule (shift + left click) then click **Calculate** in the *Aromatics Analyser*.

4. The results will be displayed in a table, which lists:
   - the centroids of the interacting aromatic rings (numbered according to the display in the visualizer)
   - the distance between these centroids
   - the relative orientation (angle between the mean planes of two aromatic rings)
   - a classification of the interaction as either inter- or intramolecular (if such pairs are included)
   - a score from 0–10, with 10 being the best possible score for a phenyl-phenyl interaction
   - an assessment of the overall strength of the interaction as Weak (score 0–3), Moderate (score 3–6) or Strong (score 7–10).

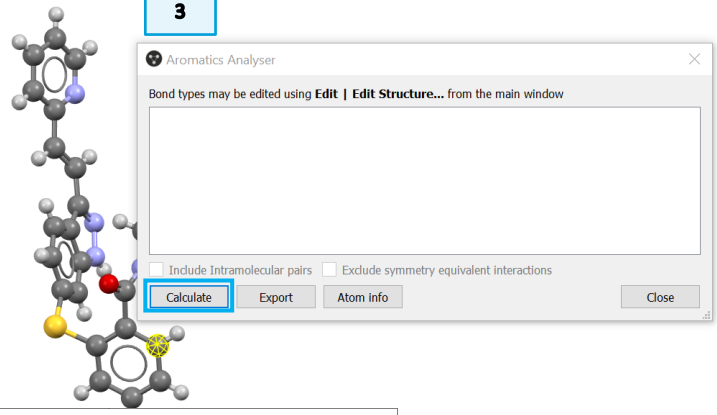
5. The results indicate two strong interactions and four moderate interactions, together with a number of weaker interactions. Click on a row of the table to visualize the interaction. The centroids of the participating aromatic groups will be highlighted blue. The table is interactive and will update the visualizer with the selected row.



2



3



4

Aromatics Analyser... VUSDIX04

Select atoms in just **one** molecule

	Centroid1	Centroid2	Distance	Relative Orientatior	Inter-molecular	Score	Assessment
1	2	16	4.88	50.33	Yes	7.2	Strong
2	2	18	4.88	50.33	Yes	7.2	Strong
3	1	18	5.91	55.5	Yes	5.4	Moderate
4	2	15	5.91	55.5	Yes	5.4	Moderate
5	1	26	5.67	55.5	Yes	5.2	Moderate
6	2	27	5.67	55.5	Yes	5.2	Moderate
7	1	21	7.24	0	Yes	2.1	Weak
8	1	16	7.61	55.5	Yes	1.6	Weak
9	2	17	7.61	55.5	Yes	1.6	Weak
10	2	26	7.89	50.33	Yes	1.3	Weak

Include Intramolecular pairs Exclude symmetry equivalent interactions

Calculate Export Atom info Close

5

Aromatics Analyser... VUSDIX04

Select atoms in just **one** molecule

	Centroid1	Centroid2	Distance	Relative Orientatio	Inter-molecular	Score	Assessment
1	2	16	4.88	50.33	Yes	7.2	Strong
2	2	18	4.88	50.33	Yes	7.2	Strong
3	1	18	5.91	55.5	Yes	5.4	Moderate

Include Intramolecular pairs Exclude symmetry equivalent interactions

Calculate Export Atom info Close

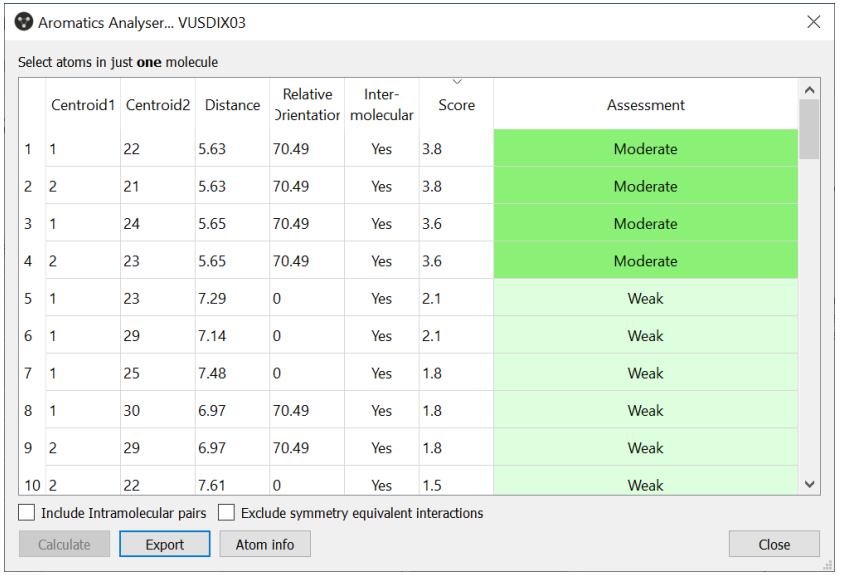
6. Click Reset in the **Display Options** in the main Mercury window, then select "VUSDIX03" in the **Structure Navigator**

7. Repeat **Step 2-5** to compare the results. You should find that only moderate and weak aromatic interactions have been identified.

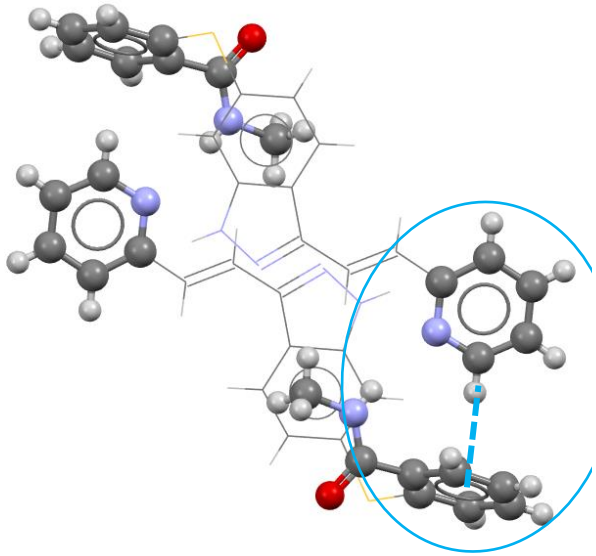
Conclusion

Aromatics Analyser has identified strong and moderate aromatic interactions in the case of Form XLI (VUSDIX04) whereas only moderate and weak interactions are found for Form VI (VUSDIX03). This may contribute towards the relative stability of the different polymorphs. Considerable care is required in this interpretation however, because Aromatics Analyser does not consider heterocycles (the pyridyl group and pyrazole part of the indazole group) and phenyl···heterocycle interactions may exceed phenyl···phenyl interactions in strength. In fact, you can find such interactions in these structures (see the T-shaped pyridyl···phenyl interaction below). Future implementations of Aromatics Analyser will incorporate heterocycles in the analysis.

7



	Centroid1	Centroid2	Distance	Relative Orientation	Inter-molecular	Score	Assessment
1	1	22	5.63	70.49	Yes	3.8	Moderate
2	2	21	5.63	70.49	Yes	3.8	Moderate
3	1	24	5.65	70.49	Yes	3.6	Moderate
4	2	23	5.65	70.49	Yes	3.6	Moderate
5	1	23	7.29	0	Yes	2.1	Weak
6	1	29	7.14	0	Yes	2.1	Weak
7	1	25	7.48	0	Yes	1.8	Weak
8	1	30	6.97	70.49	Yes	1.8	Weak
9	2	29	6.97	70.49	Yes	1.8	Weak
10	2	22	7.61	0	Yes	1.5	Weak



T-shaped interaction

Part 4: Using Full Interaction Maps to assess intermolecular interaction geometry

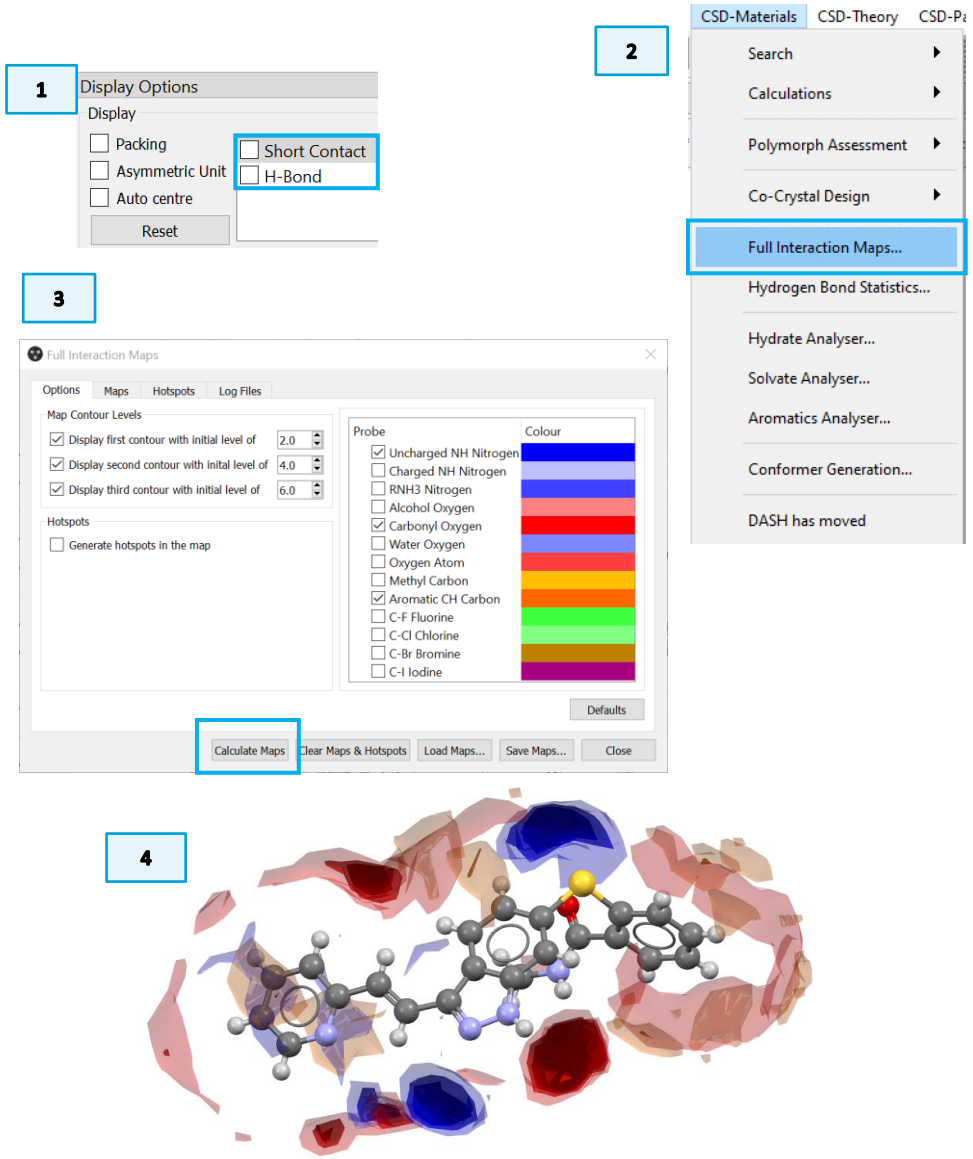
Full Interaction Maps are a complementary approach to Hydrogen Bond Propensities and Aromatics Analyser for assessing the intermolecular interactions. FIMs uses *IsoStar* data to create a probability density plot based on the likelihood of finding a particular probe in a particular region in the structure.

1. With "VUSDIX04" (Form XLI) loaded in **Structure Navigator**, check that the **H-Bond box** in the **Display Options** toolbar is not checked.

2. Click on *CSD-Materials* menu and then select *Full Interaction Maps…* from the dropdown menu.

3. In the **Full Interaction Maps** window, you will see several options. On the left you will find options to change the display contour levels. On the right, you will see a list of functional groups to be used as probes. For the purposes of this tutorial, we will keep the default options. These typically work well for most situations, but if you know you are looking for a specific functional group, or if you want to change the look of the map, you will want to change these settings. Click the **Calculate Maps** button to start. *Note: if you are working with multicomponent systems or structures with Z' > 1, you must select only one molecule in the visualizer before calculating FIMs.*

4. The generated map will now be displayed in the main Mercury window. Notice the three different colors in the map. Red regions of the map denote areas in which there is a high probability of locating a hydrogen bond acceptor. Blue regions denote areas in which there is a high probability of locating a hydrogen bond donor, and orange regions indicate hydrophobic pockets.



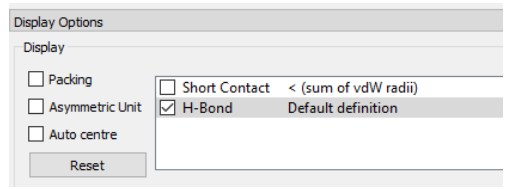
5. Now we want to see how the overall packing of this polymorph fits with the FIMs we have generated. Tick the box for **H-bond** in the **Display Options** toolbar.

6. Use a D-H…A angle of 120˚ to define the hydrogen bond criteria. To ensure this is the case double click the H-bond line to launch the **Define H-bonds** window. Then tick the box for **Require hydrogen atoms to be present**. Click **OK** to apply the change.

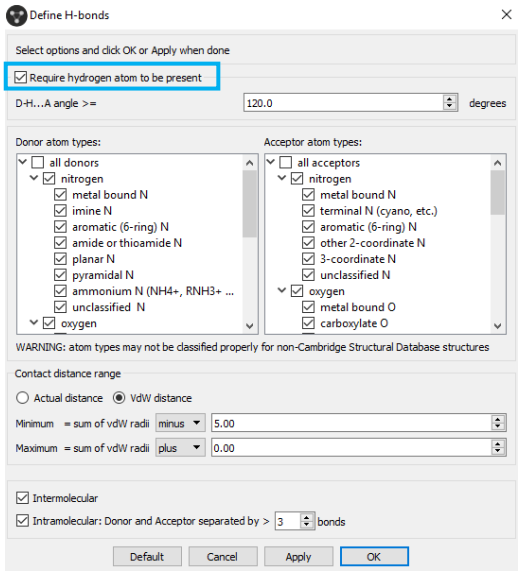
7. Now you will see dashed red lines in the Mercury window that indicate where hydrogen bonding interactions/contacts are present.

8. Click on these contacts to generate nearby molecules. Some interactions are moderately well fulfilled but some fall just outside the regions predicted by FIMs, or are absent, which indicates that the intermolecular H-bonds are not optimal.

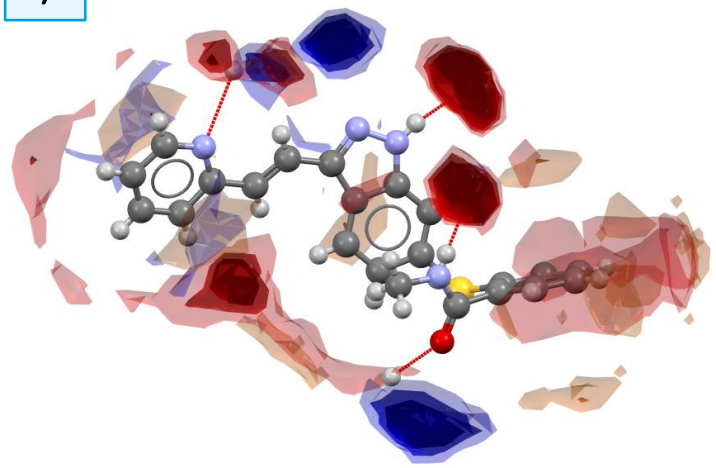
5



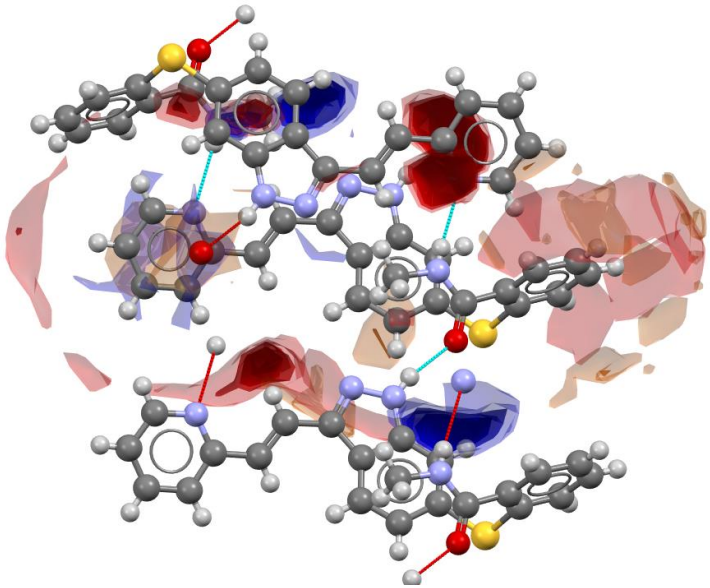
6



7



8



9. Untick **H-bond** in **Display Options** and click **Reset** to remove the molecules generated by expanding the hydrogen bonds.

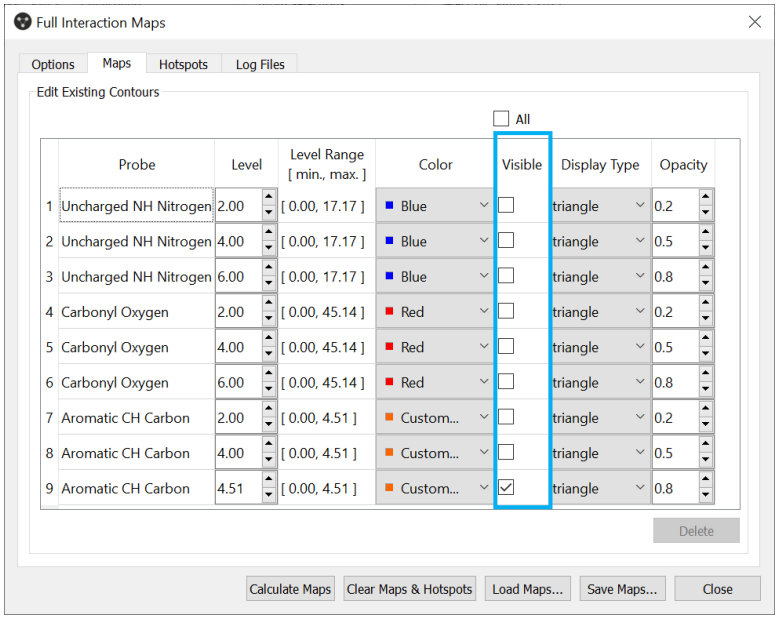
10. In **Full Interactions Maps** window, click on the **Maps** tab and untick the boxes in the *Visible* column for all rows **except row 9**, Aromatic CH carbon.

11. The value in the *Level Range* column indicates the minimum and maximum probability above random of finding a particular probe in a particular region. Change the value in the *Level* column, **row 9** to 2.75. You will see that there are a number of regions where there is some likelihood of finding aromatic interactions.

12. To see how well these interactions are fulfilled we can generate a *molecular shell* around the molecule for which FIMs have been calculated. From the top menus, select *Calculate > Molecular Shell…* to bring up the dialogue. To get a sufficient shell of molecules, select one atom in each of the six-membered aromatic/heterocyclic rings. In the *Molecular Shell* dialogue, set the *Maximum* **Actual distance** to 4 Å. Select the molecule (shift + left click on it), tick **Show contact lines** and then click apply.

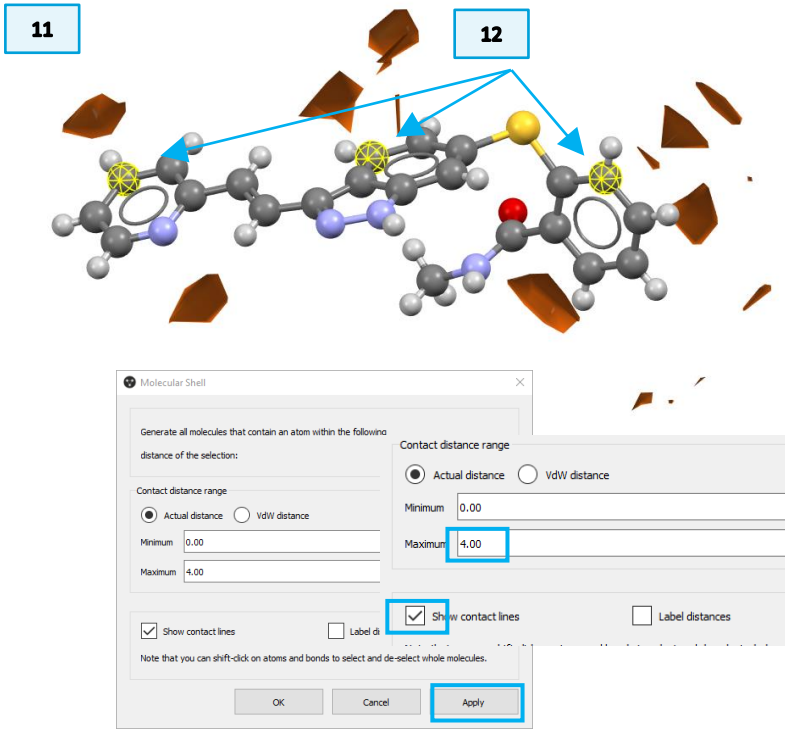
13. The atoms used to build the molecular shell should still be highlighted. Use these to locate the molecule at the center of the shell, shift + left click on it (use a non-highlighted atom) to select all of the atoms and change the colour by right clicking, then choose *Colours* and pick a colour (here we use green).

10

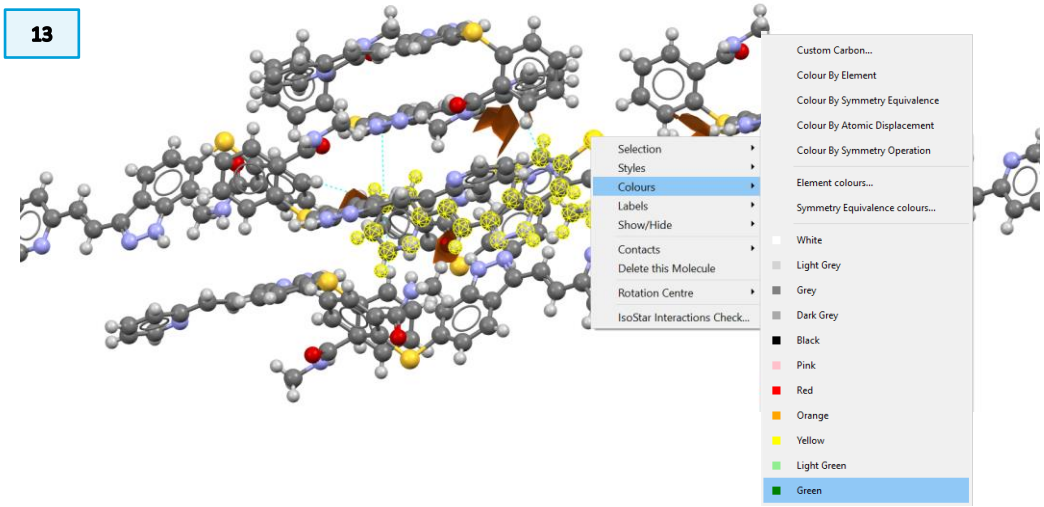


11

12



13

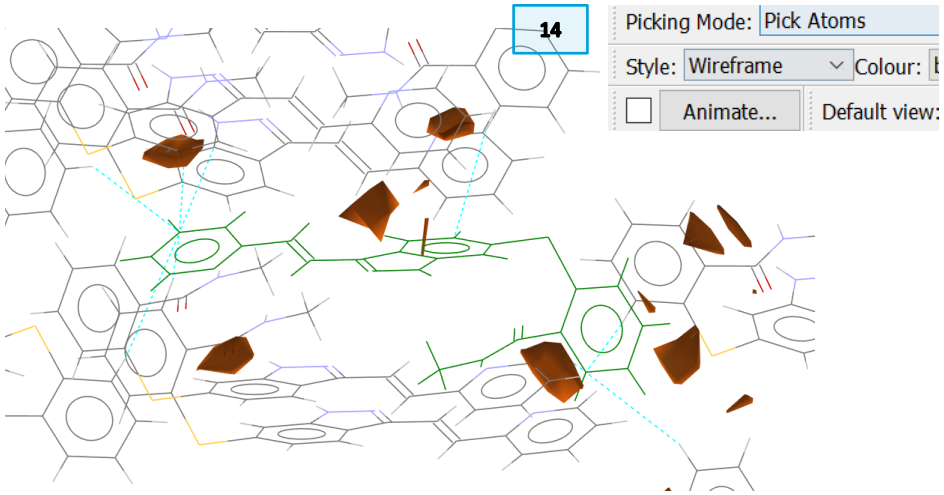


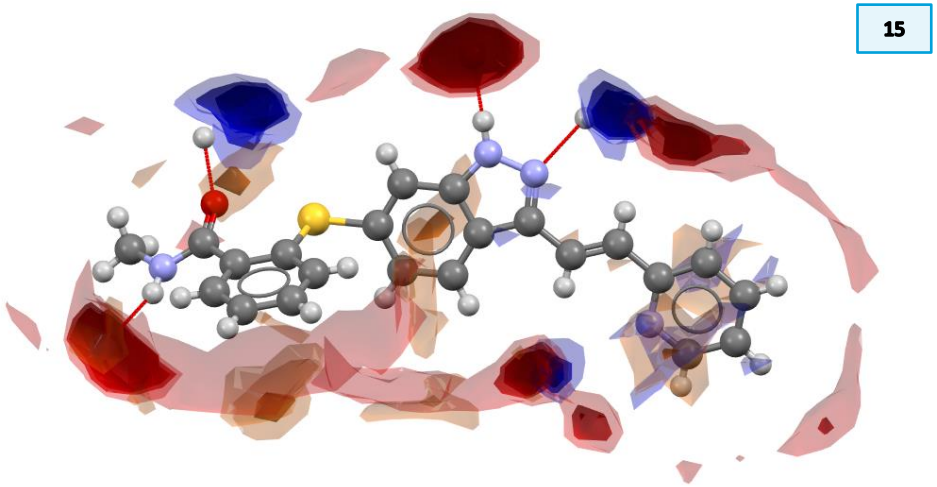
14. To make the display clearer, change the *Style* to *Wireframe* from the top toolbar. You will see that a number of the aromatic CH regions are quite well satisfied, however the best interactions *do* involve the pyridyl group, which could not have been uncovered using Aromatics Analyser.

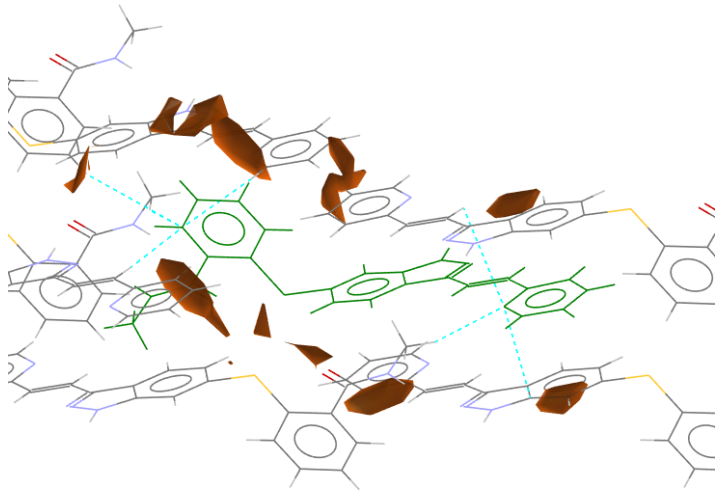
15. Repeat **Steps 2-14** for Form VI (VUSDIX03). You may wish to open a separate instance of Mercury to faciliate side-by-side comparison. Compare and contrast the results with those from Form XLI (VUSDIX04).

Conclusion

The FIMs of Forms VI and XLI of Axitinib indicate that predicted H-bonding interactions are only moderately well fulfilled in either form. Aromatic CH interactions do seem reasonably well fulfilled and although density levels are low compared to hydrogen bonding, they show a preference for the involvement of phenyl-pyridyl interactions. However, unlike Aromatics Analyser, a quantitative assessment of these interactions is not possible with FIMs.







Summary

To summarise, in this workshop you have learnt how to use several complementary tools to investigate polymorph stability in *N*-methyl-2-((3-(2-(pyridin-2-yl)vinyl)-1*H*-indazol-6-yl)sulfanyl)benzamide (Axitinib). Using structure overlays and Mogul, we assessed molecular conformation. We performed a detailed analysis of potential hydrogen bonding networks using Hydrogen Bond Propensities, based on the functional groups of the molecule, and explored potential aromatic interactions using Aromatics Analyser. Finally, we used Full Interaction Maps to assess and validate the results of the HBP and Aromatics Analyser calculations, prompting us to also consider the role of heterocyclic groups in stabilising the structure. Ultimately, Hydrogen Bond Propensities gave the clearest discrimination between the two crystal forms, suggesting that Form XLI (VUSDIX04) is the most stable. This is found to be the case experimentally (Form VI (VUSDIX03) is metastable). It is worth noting that Axitinib forms five “neat” polymorphs and sixty-six solvates which suggest, in line with our findings here, that self-self interactions are not particularly good.

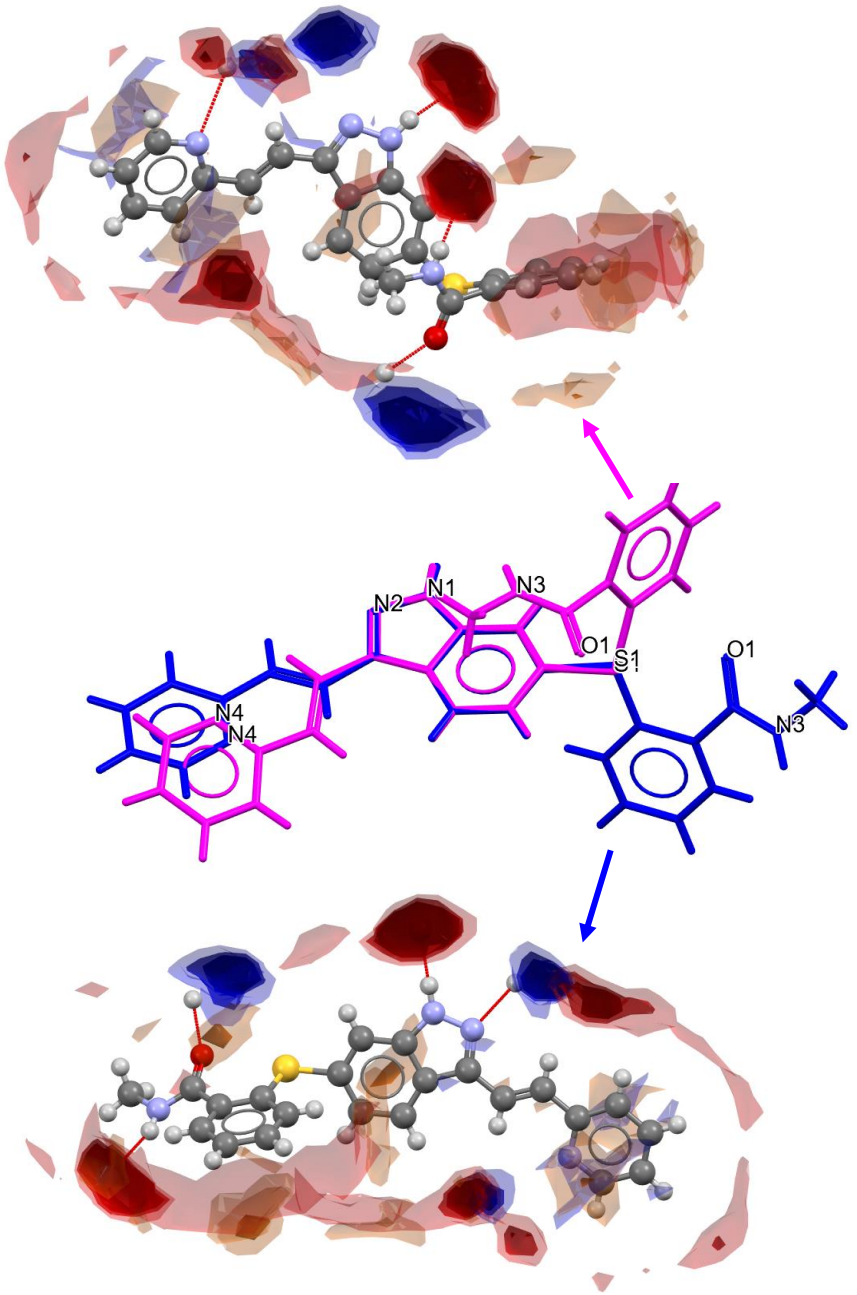
You should now be able to:

- Use Structure Overlay and Mogul Geometry Check to assess conformation.
- Calculate Hydrogen Bond Statistics and Hydrogen Bond Propensities, and interpret the results.
- Use Aromatics Analyser to search for and evaluate potential aromatic interactions.
- Calculate FIMs for a molecule and interpret the maps.

Next Steps

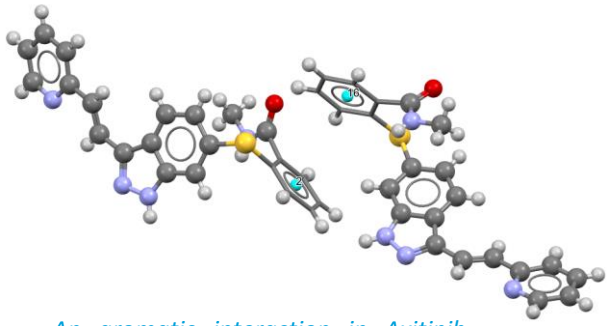
You can find out more about each of the individual tools presented in this workshop in self-guided workshops in the CSD-Materials and CSD-Core sections of the CCDC website. On-demand courses on structure visualization, Mogul and FIMs are also available and a certificate can be obtained upon completion.

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



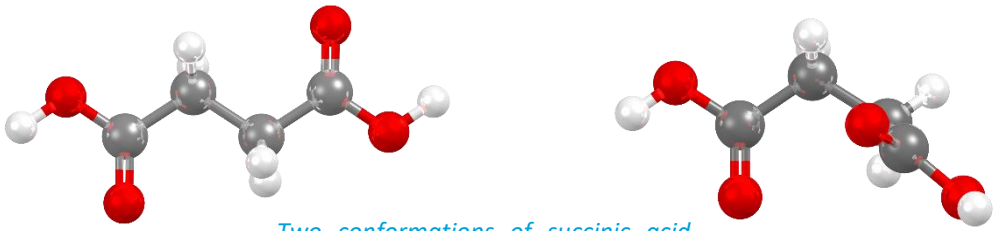
Glossary

Aromatic Interactions
Noncovalent interactions formed between aromatic rings. These interactions are important in material science since they will contribute to the overall crystal structure stability. The orientation of the aromatic ring can vary from parallel to T-shape, and we found during our DFT calculations that the T-shape interactions are very close in strength to the parallel displaced ones. Their strength is found between 0 and 16 kJ/mol based on DFT calculations.



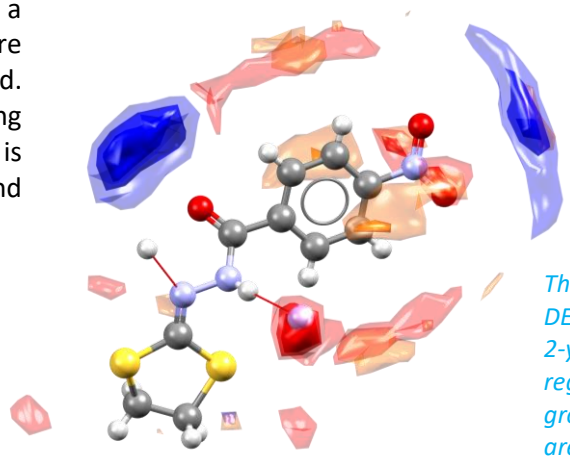
An aromatic interaction in Axitinib (refcode VUSDIX04).

Conformation
The spatial arrangement of the atoms affording distinction between stereoisomers which can be interconverted by rotations about formally single bonds. Some authorities extend the term to include inversion at trigonal pyramidal centres and other polytopal rearrangements. *Sources*: PAC, 1994, 66, 1077. (Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)) on page 1099.



Two conformations of succinic acid molecules, shown on refcodes SUCACB02 (left) and SUCACB19 (right).

Full Interaction Maps
Full Interaction Maps provide a way of visualising the preferred interactions of a molecule. Regions around the molecule (maps), where chemical probe groups are likely to be found, based on pre-extracted IsoStar data from the CSD, are calculated. The program works by first identifying distinct functional groups in the molecule being studied, then finds the relevant data in IsoStar. The group-based interaction data is evaluated, taking into account the environmental effects of combinative factors and steric exclusion to create a 3D picture of molecular interaction preferences.



The Full Interaction Maps for CSD entry DEDMUX02, a polymorph of N'-(1,3-dithiolan-2-ylidene)-4-nitrobenzohydrazide. The blue regions show the likely locations of donor groups, the red regions show where acceptors are expected to be found and the brown regions indicate potential hydrophobic interactions.

Hydrogen Bonds

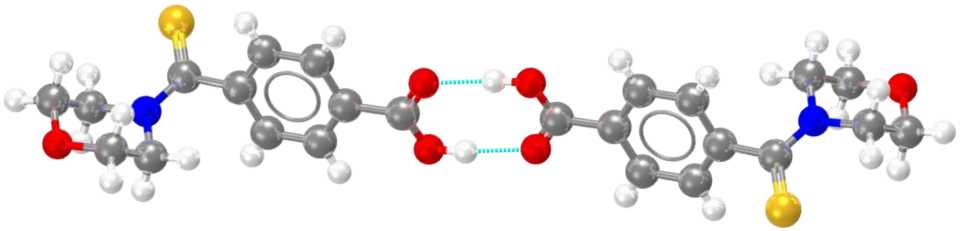
Hydrogen bonding occurs between donor-acceptor interactions precisely involving hydrogen atoms. The H-bonds interactions are classified as: strong (mostly covalent), moderate (mostly electrostatic) and weak (electrostatic). Their strength is observed to be between 12 and 30 kJ/mol.

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.]

Hydrogen Bond Propensity (HBP)

- The HBP tool in Mercury > CSD-Materials evaluates the relative likelihoods of possible H-bonding networks in any observed polymorphs of a target system.
- Probabilities for hydrogen bond pairings to form in the target system are calculated from a statistical model built from relevant structures in the CSD. The model encapsulates information regarding the environment of the functional groups, which ensures the prediction is specific to the target molecule.
- Combining probabilities of hydrogen bond formation with a statistical model that captures information regarding how often a functional group participates allows the generation of chemically sensible alternative structures.
- The view of the solid-state landscape of an active ingredient afforded through the combination of propensity and coordination addresses questions such as how likely polymorphism is and whether there is the possibility of a more stable form. Specifically, you can:
    - Predict likely hydrogen bonds for a given molecule.
    - Assess crystal forms, e.g., by identifying sub-optimal hydrogen bonding.
    - Calculate hydrogen bond propensities for individual donor and acceptor groups.
    - Perform a comprehensive analysis of hydrogen bonding on a set of structures.



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

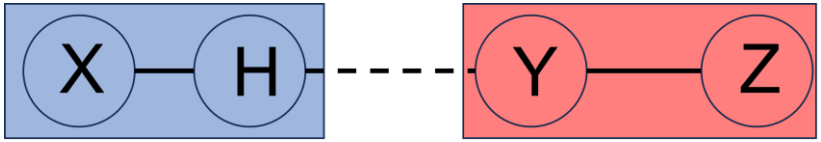
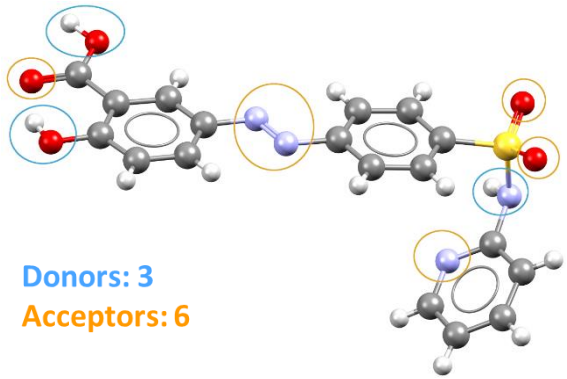


Illustration of a hydrogen bond interaction with between hydrogen bond donor X–H and hydrogen bond acceptor Y–Z.



Sulfasalazine exhibits 3 potential donors and 6 acceptors that might compete in forming H-bond interactions. HBP can be used to evaluate which of these potential interactions are more likely to form.

IsoStar

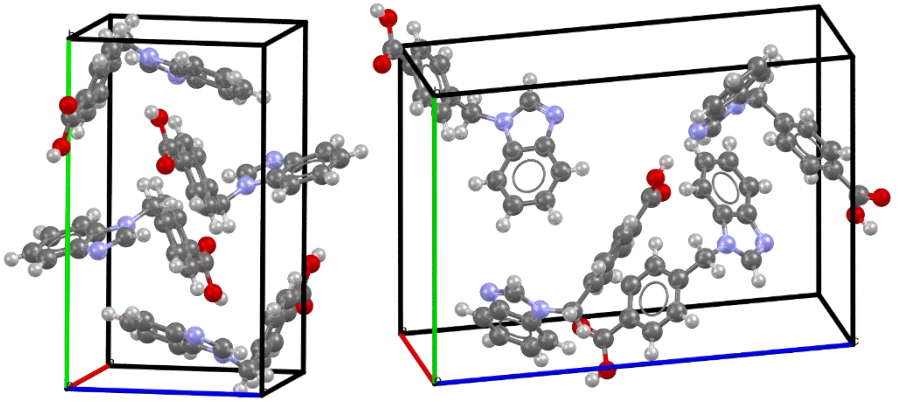
IsoStar is a knowledge-based library of intermolecular interactions. It contains data on intermolecular interactions in small molecule crystal structures taken from the CSD, as well as protein-ligand interactions in X-ray structures from the PDB.

Molecular Shell

A molecular shell in Mercury will display all molecules within a specified distance of a selected atom or atoms. In some fields this would be referred to as a "coordination sphere".

Polymorph

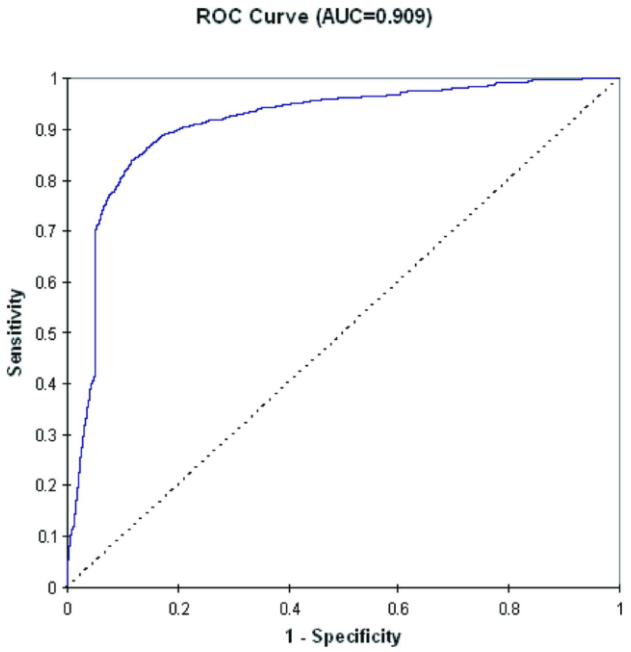
Polymorphism is the occurrence of two or more crystalline forms of the same substance. Where available, polymorph information can be displayed for Cambridge Structural Database (CSD) structures. Structures known to be polymorphic contain comments which include the word polymorph (when reported by the author), e.g., non-triboluminescent polymorph. There is also a CSD subset of polymorphic structures.



Example of polymorphic structures of 4-(1H-Benzimidazol-1-ylmethyl)benzoic acid: the monoclinic polymorph (CSD Entry ABADIS) at the top, and the orthorhombic polymorph (CSD Entry ABADIS01) on the right.

Receiver Operator Characteristics

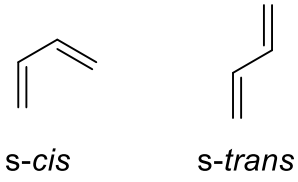
The ROC curve (receiver operating characteristics) gives a measure of how well classified the predictions are using the training data as a test. It calculates percentage classification using a variable cut-off, either side of which a propensity is considered positive or negative. The *sensitivity* is the fraction of correct positive predictions and the *specificity* is the fraction of correct negative predictions. The diagonal dotted line is the outcome of a purely random model as there is an equal number of correct and incorrect predictions. An AUC (area under the curve) greater than 0.5 indicates the model predictions are correct more frequently than a random choice. An AUC of 1 indicates a perfect model: correct every time. An AUC above 0.8 is considered excellent and above 0.9 indicates an outstanding model. The difficult middle section around sensitivity/specificity = 0.5 needs a well discriminating model in less extreme cases, and must be well described by the model parameters in order to obtain a high AUC. One may observe that the example LHP model gives an outstanding classification of the training data and achieves an AUC of 0.909. [Reference: P. T. A. Galek, L. Fábián, W.D. Sameul Motherwell, F. H. Allen and Neil Feeder, *Acta Crystallogr. B*, 2007, **63**, 768 – 782]



ROC curve using the model to predict the training set outcomes The diagonal dotted line indicates the curve of a model with no predictive power: there is equal likelihood of a correct and an incorrect prediction.

s-cis/trans

Two conjugated double bonds may commonly adopt one of two conformations; either the double bonds are on the same side of the single bond (s-*cis*) or on opposite sides (s-*trans*).



Demonstration of cis *and* trans *conformations.*

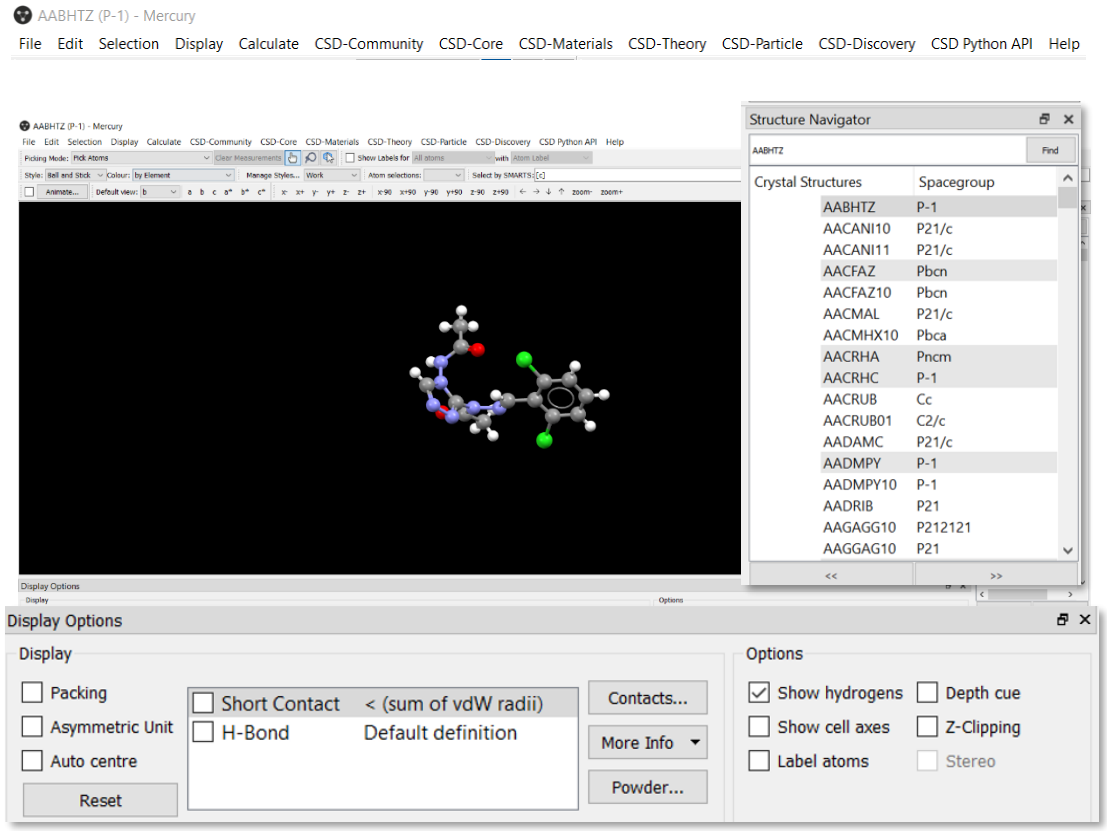
Basics of Mercury Visualization

Mercury is the CCDC's visualization software to view 3D structures of small molecules, generate images, and animations of molecules.

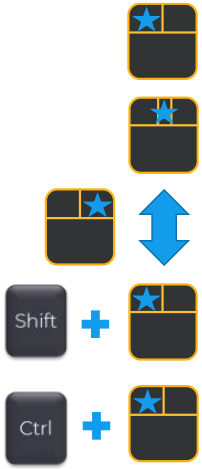
In the following we will see some of the basics of navigation and visualization in Mercury that you will find helpful to support your analysis.

In the **Mercury interface** we find:

- **At the top**: list of menus from which we can access visualization and analysis options, and other CSD components such as CSD-Materials.
- **On the right-hand side**: the **Structure Navigator**, with the database loaded (depending on your licence). The Structure Navigator allows you to select a refcode to visualize in the main Mercury window.
- Beneath the main display window: Display options toolbar. You can quickly view a packing diagram, display Hydrogen bonding and detailed information about the molecule using the More Info option.



Using the **mouse to enhance visualization**:



- Left mouse button and move – rotate molecules.
- Middle Mouse wheel – move molecules up and down.
- Right mouse button and move up and down – zoom in and out of molecules.
- Shift + Left mouse button and move - rotate in the plane molecules.
- Ctrl + Left mouse button and move - translate molecules.

Right click:
a) Near a molecule and
b) Away from a molecule