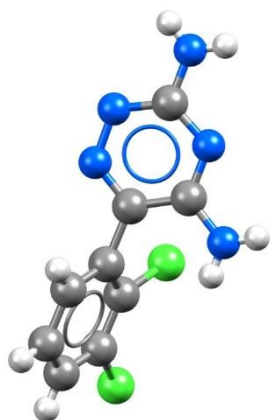


Full Interaction Maps Mogul Geometry Check Crystal Packing Feature Search

2020.1 CSD Release



CCDC
advancing structural science

Table of Contents

Introduction.....	2
Example 1. Using Mogul to assess intramolecular geometry	3
Example 2. Using Mogul to explain activity data	5
Conclusions.....	8
Further Exercises	8
Example 3. Using Full Interaction Maps to assess polymorph stability	9
Conclusions.....	11
Example 4. Investigating the potential for co-crystallisation.....	12
Conclusions.....	13
Example 5. Investigating a series of halogenated molecules.....	14
Conclusions.....	16
Further Exercises	16
Example 6. Investigating & analysing intermolecular interactions	17
Conclusion	21

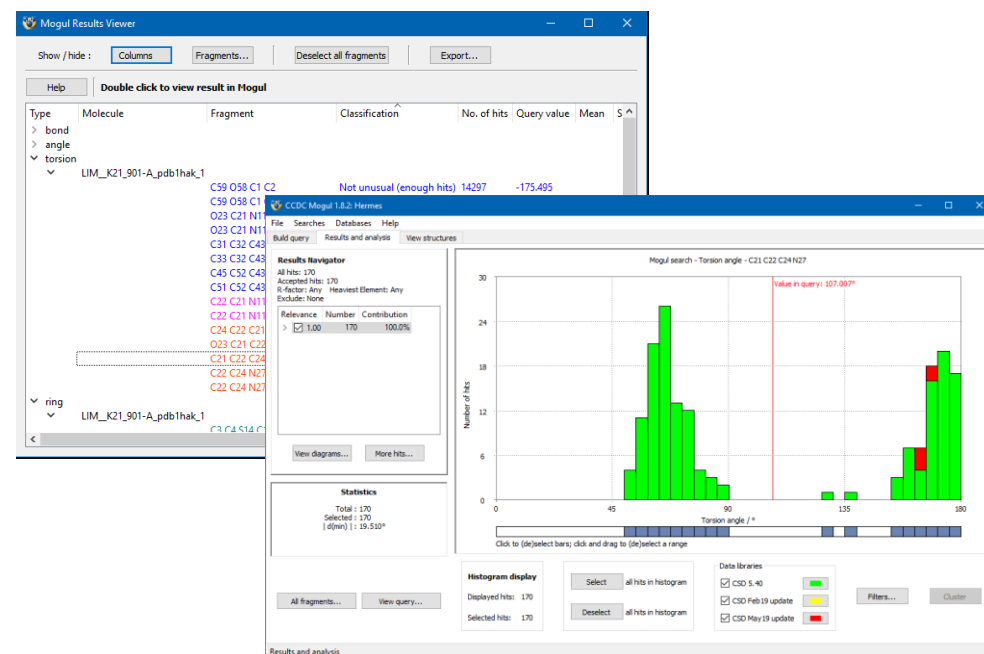
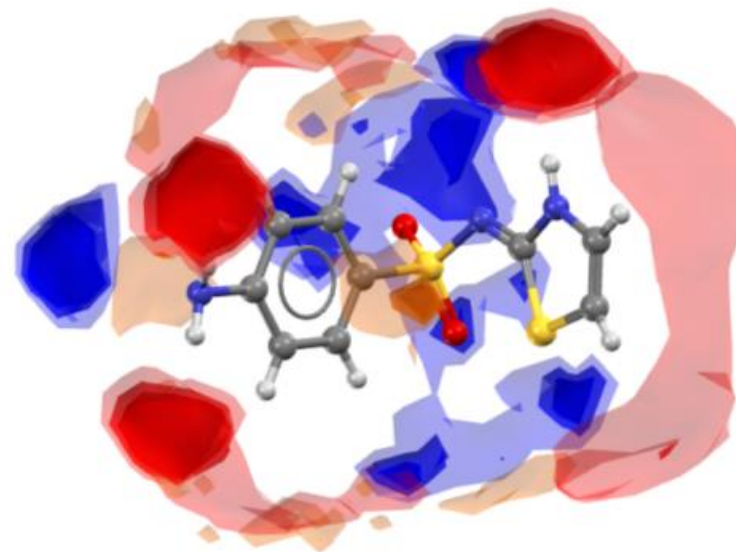
Introduction

This tutorial will introduce you to the Full Interaction Maps and Mogul Geometry Check features included with Mercury under the CSD-Materials/CSD-Discovery and CSD-System toolsets.

The Full Interaction Maps tool will generate a picture of the interaction landscape of your molecule from its three-dimensional coordinates. Using statistical distributions from the hundreds of thousands of structures included in the CSD, we can predict the most likely locations for a variety of functional groups. By comparing this distribution against a 3D packing diagram, we can determine whether a crystal structure fulfills the desired interactions of a particular conformation of a particular molecule. The Full Interaction Maps tool is instrumental in highlighting the potential for polymorphism of a given compound, assisting in the development of co-crystals, and understanding solid form stability.

Structural chemists can use the Mogul Geometry Check tool to validate the three-dimensional conformation of a particular molecule. The metric data from structures in the CSD can be used to show the most likely values a particular bond, angle, torsion or ring should adopt. The information obtained from a Mogul check can help identify inconsistencies within a crystal structure and can suggest values to be used for restraints during refinement. In addition, Mogul can also be run on 3D coordinates generated *in silico* as a validation of calculated structures.

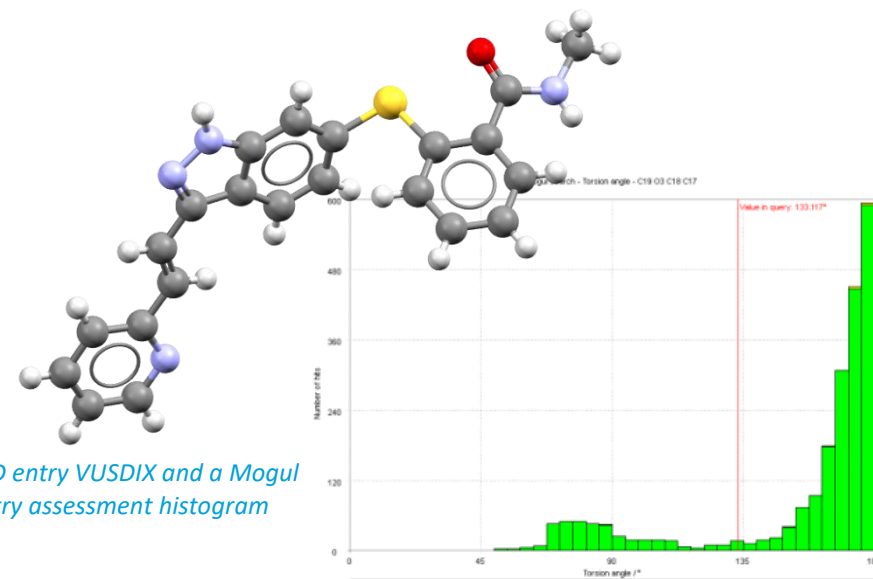
The following exercises assume you have a working knowledge of the program Mercury, namely how to display and manipulate structures from a 3D coordinates file. There are many other features included with Mercury that are beyond the scope of this tutorial. For assistance with any aspect of Mercury, please ask your workshop instructor, or view the help guides included with the application.



Example 1. Using Mogul to assess intramolecular geometry

Mogul is able to provide an assessment of a given structure's conformation by comparing it to the data from the hundreds of thousands of structures already in the CSD. By using the statistical distributions of similar fragments, Mogul can confirm your 3D geometry is appropriate, or flag values that are too far outside the norm.

In this example, you will see how to use Mogul to assess the geometry of a given molecule. Mogul can be run as a stand-alone application or from the Mercury or Hermes interfaces. For this tutorial, we will use Mercury to run Mogul.



The CSD entry VUSDIX and a Mogul geometry assessment histogram

1. Launch Mercury and type VUSDIX in the Structure Navigator toolbar.
2. To start a geometry analysis, click the CSD-System menu and then click *Mogul Geometry Check* from the dropdown menu.
3. This brings up the *Mogul Search Settings* dialogue box. You can typically use the defaults in this window, but we can streamline our search by unticking the box for rings and ticking the boxes for *Apply Filters*, *Exclude Organometallics*, and *Exclude Powder structures*.
4. Click **Search** to start.
5. A dialogue box will pop up to warn you that you are going to check the entire molecule. Click **OK** to continue.
6. The search will begin to run. You can follow its progress in the *Search Progress* dialogue box.

1 Mercury interface with VUSDIX in the Structure Navigator.

2 CSD-System menu open, showing 'Mogul Geometry Check...'.

3 'Mogul Search Settings' dialog box. Available filters: R-factor (≤ 5.0%), Exclude Solvents, Exclude Heaviest Element (U), Exclude Organometallics, Exclude Powder structures. Search Mode: Find similar fragments if number of exact matches is less than. Bonds: 15, Angles: 15, Torsions: 40, Rings: 15.

4 'Search' button highlighted.

5 'No atoms selected' warning dialog box. A complete analysis of all loaded molecule(s) will be performed. To analyse just part of the displayed molecule(s), hit 'Cancel' and select atoms before starting the analysis.

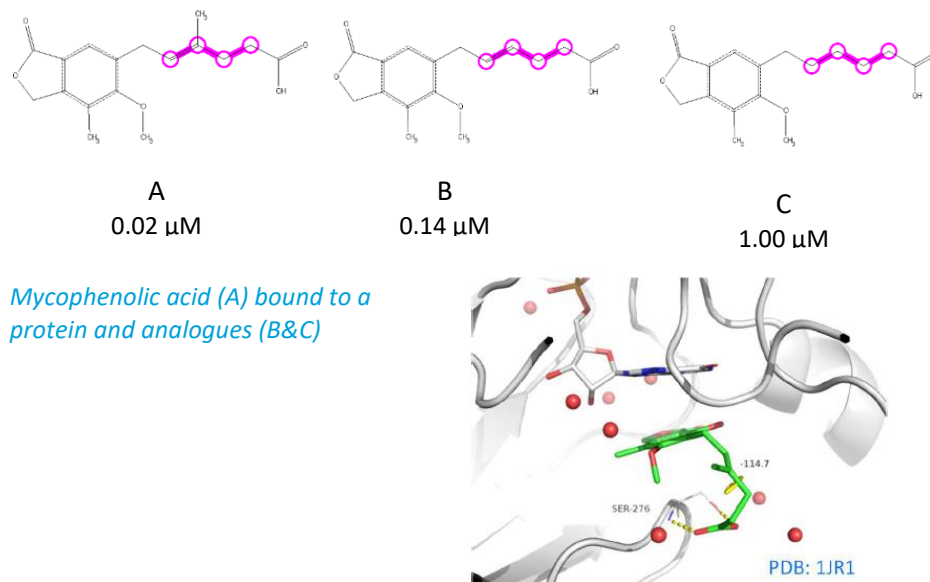
6 'Search Progress' dialog box. Torsion C11 C12 C13 C8: 98% completed. Search completed - 239 observations collected.

Example 2. Using Mogul to explain activity data

Mycophenolic acid (A) is an immunosuppressant that binds to inosine monophosphate dehydrogenase (IMPDH) with an IC_{50} of $0.02\mu M$. There are two close analogues (B and C) which are less active, with C much less so. How can Mogul help explain this decrease in activity?

In this example you will see how to use Mogul to correlate improved function with geometric preferences. You will need the file `1JR1_ligand.pdb` for this example

A successful target molecule will be able to form two favourable hydrogen bonds in the protein binding site, as shown. The angle of the allylic torsion angle should be in the range of $110-115^\circ$ for this interaction to be achieved. We know that molecule A is a successful target but molecules B and C are not. What does Mogul have to say about these structures?

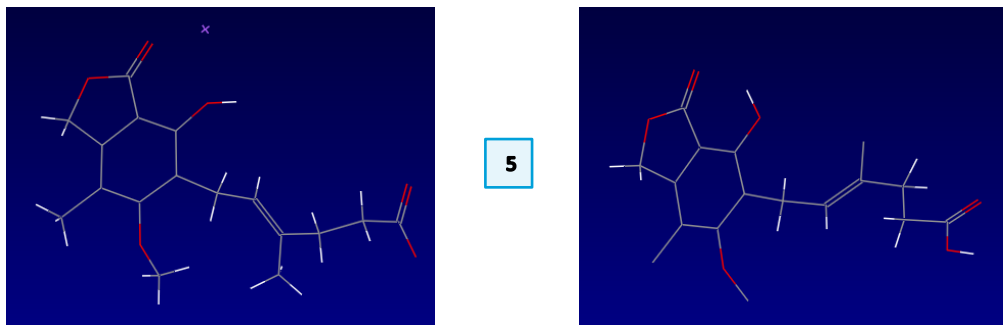


1. Launch ConQuest and click on the **Name/Class** query type.
2. In the box for Compound Name type "mycophenol" and then click **Add**.
3. Click **Search** and then **Start Search** for the *Search Setup* dialogue box.
4. The search should return 8 results. MYCPHA and MYCPHA01 are the neutral species. KIWQUC is a sodium salt. WAJYUC is a neutral derivative. The others are co-crystals.
5. Use the *3D Visualiser* tab to view the MYCPHA and KIWQUC structures noting the different conformations. Which one is the preferred orientation?

1

2

3



6

6. Launch Mogul (*via* the desktop icon, or *via* Mercury). We will first do a torsion search on the ligand as it appears in the protein crystal structure.

7. Click **Load...** and select the file 1JR1_ligand.pdb (provided for you along with this example). When the file is imported, Mogul will prompt you to edit the structure to assign bond types and add hydrogen atoms if necessary. Click **OK** to accept these changes.

8. Click to select the four carbon atoms that define the C2=C3-C4-C5 torsion angle. These will appear in the *Current Selection* box as you click.

9. Click **Search** when all atoms are selected.

10. The resulting histogram shows that the query value (114.654°) falls in a highly populated region of the histogram.

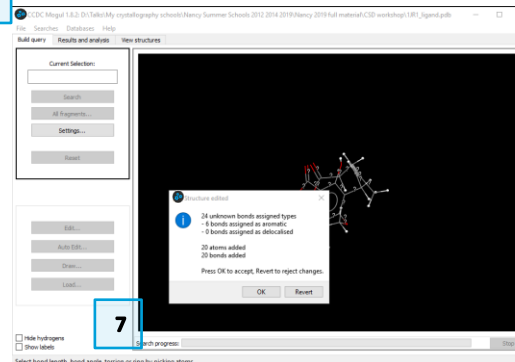
11. Now let's investigate the other two analogues in comparison to the bound ligand. We can do this by editing the ligand structure in the Mogul interface.

12. In the *Build query* tab, click **Edit...** This will launch the *Edit Structure* dialogue box. At the bottom of this box, under "Remove" click **Atoms & Bonds**. Then click on the methyl atoms (C and 3H) to delete them from the structure.

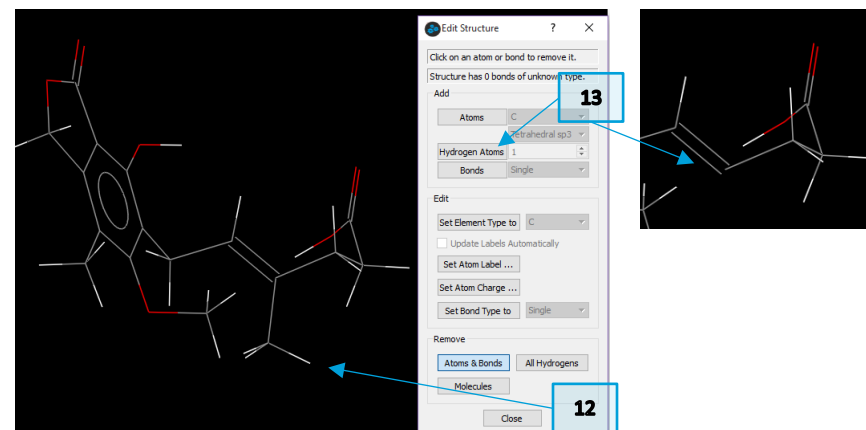
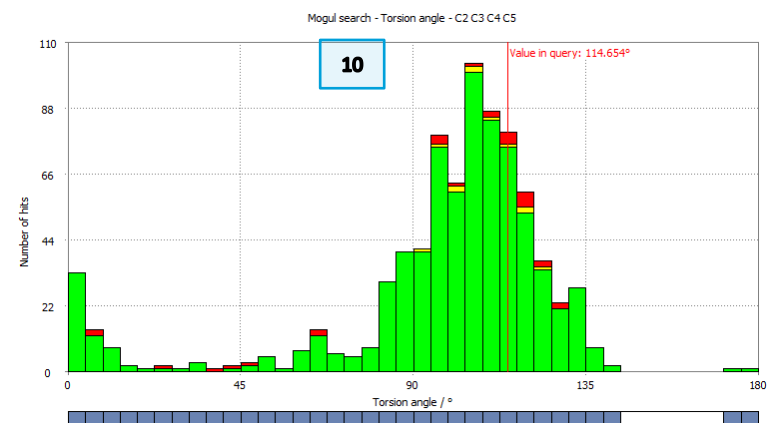
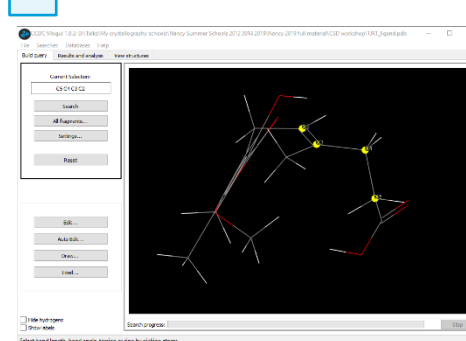
13. Once the atoms are deleted, at the top of the *Edit Structure* dialogue box under "Add" click **Hydrogen Atoms**. Make sure the value is set to 1 and then click the carbon atom from which you just deleted the CH₃. This will add an idealized hydrogen at this position. Click **Close** when you are finished.

14. As in step 8 above, select the four carbon atoms of the C=C-C-C group, and then click **Search** to start Mogul running.

6

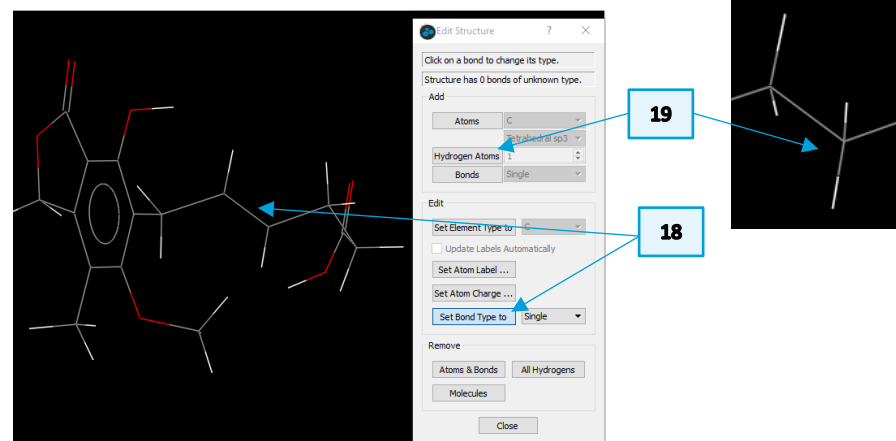
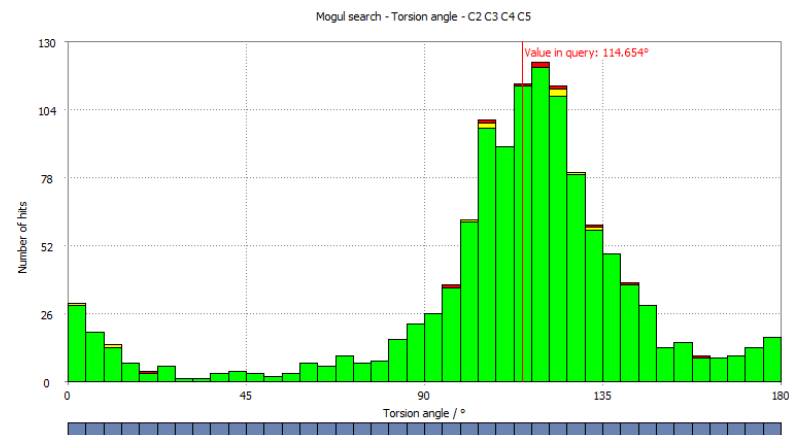


8

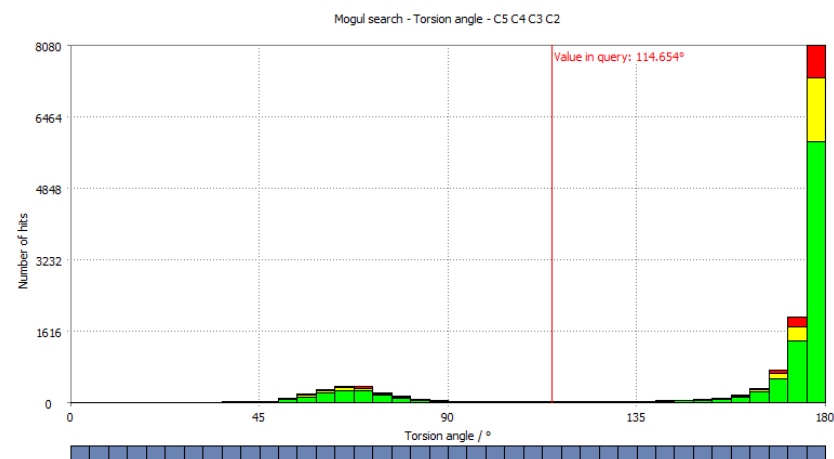


15. You will see the resulting histogram has shifted slightly, indicated the change in *chemistry* of the query fragment (des-methyl vs. methyl). However, the query *value* stays the same because we did nothing to alter the geometry of our query fragment. This shift in the predicted geometry indicates a slight loss of potency in this target molecule (B, above).
16. Finally, to investigate molecule C above, we need to make one more change. This molecule lacks the double bond that is present in molecules A and B.
17. Once again, return to the *Build query* tab and click **Edit...**
18. This time in the *Edit Structure* dialogue box click **Set Bond Type to** and choose *Single* from the drop-down menu beside it. Then click on the double bond to change it to a single bond.
19. Now, as in step 13 above, in the “Add” section, click **Hydrogen Atoms** and make sure the value is set to 1. Click the two carbon atoms on either side of what had been the double bond, to ensure they are fully saturated. Click **Close** to exit the dialogue box.
20. Once again, click to select the same four carbon atoms and then click **Search** to start Mogul running.
21. Now you will see the resulting histogram for this alkyl chain has shifted far away from the query value for mycophenolic acid. This means that this chemical change will not satisfy the necessary orientation for effective binding and explains the loss of potency in molecule C.

15



21



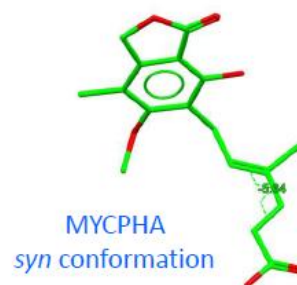
Conclusions

The skewed conformation for the crucial torsion in mycophenolic acid is the preferred orientation, although the neutral form of the molecule itself adopts a planar conformation for this torsion in the small molecule crystal structure. An analysis of conformational preferences rationalized the observed activity data for similar compounds.

Mogul can be an effective tool in explaining pharmacophore activity in relation to changes in substituents and bonding patterns. You should now be able to use Mogul to make changes to your own structures to see what functional groups can be altered in order to produce desired conformations or geometries.

Further Exercises

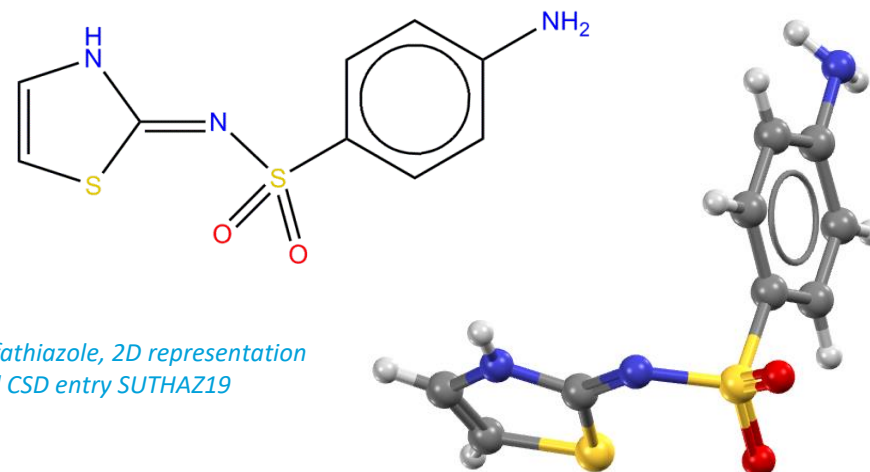
- Use Mogul to investigate subtle changes to your own molecules, or others that you find in the CSD.
- Using what you know of Full Interaction Maps, is there an explanation for the planar torsion angle of MYCPHA?




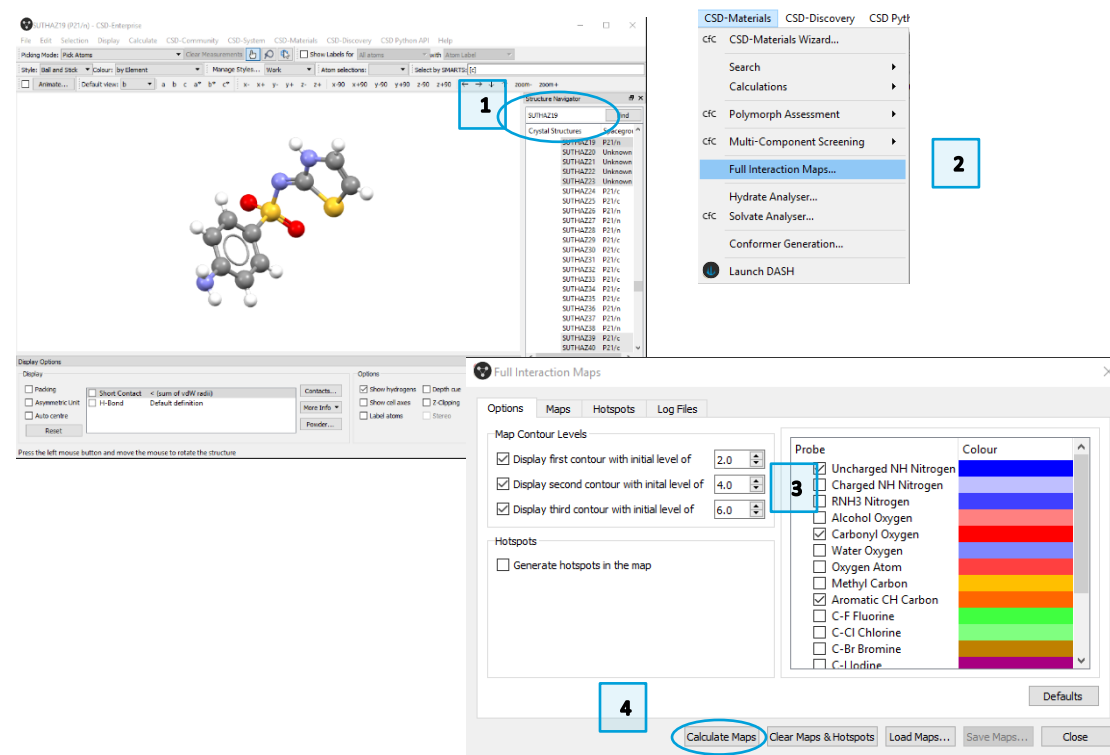
Example 3. Using Full Interaction Maps to assess polymorph stability

The stability of a given crystal structure is a balance between the intramolecular conformation and the intermolecular packing of the molecules in the crystalline state. One method for understanding the relative stability of crystal structures is to compare the observed intermolecular interactions with preferred geometries for that type of interaction.

This example will look at two known polymorphs of sulfathiazole to answer the question: **How do the interactions in each polymorph compare with what is expected?** You will learn how to produce a Full Interaction Map for a given structure and how to interpret this map.



1. Launch Mercury by clicking its icon . In the Structure Navigator toolbar, type SUTHAZ19 to bring up the structure for the Form V polymorph of sulfathiazole.
2. Click on the *CSD-Materials* menu or *CSD-Discovery* menu and select *Full Interaction Maps...* Note: If the CSD-Materials menu bar is inactive, click on "Help" and "Activate CSD-Materials" before proceeding further.
3. In the *Full Interaction Maps* dialogue box, 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.
4. Click the **Calculate Maps** button to start.



5. The generated map will now be displayed in the main Mercury window. Notice the three different colours 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 hydrogen bond **donors**, and **brown** regions indicate **hydrophobic** preferences.
6. Now we want to see how the overall packing of this polymorph fits with the Full Interaction Map we have generated. Tick the box for *H-Bond* in the Display Options toolbar.
7. Double click the *H-Bond* line to launch the *Define H-bonds* dialogue. In this dialogue, tick the box for “Require hydrogen atom to be present”. Click **OK** to apply the change.
8. Now you will see dashed **red** lines in the Mercury window that indicate where hydrogen bonding interactions/contacts are present.
9. Click on these contacts to generate nearby molecules. You will see that in each case, the interaction falls within the contour range for the expected type. This indicates that the packing of Form V satisfies the expected interaction landscape of this conformation of sulfathiazole.
10. Now let's look at the Form I polymorph. In the Structure Navigator toolbar, type SUTHAZ16 to load this structure.
11. Click the **Reset** button in the Display Options toolbar to remove all the hydrogen bonding interactions.
12. Note that $Z' = 2$ for Form I; meaning there are two unique molecules in the asymmetric unit. We will focus on one unique molecule in the Form I structure.
13. Tick the box to Label atoms in the Display Options toolbar. Locate which molecule contains N1. This is the molecule we will focus on. Once you have located it, remember which one it is, but untick the Label atoms box to keep the display clean.

5

7

6

8

9

10

11

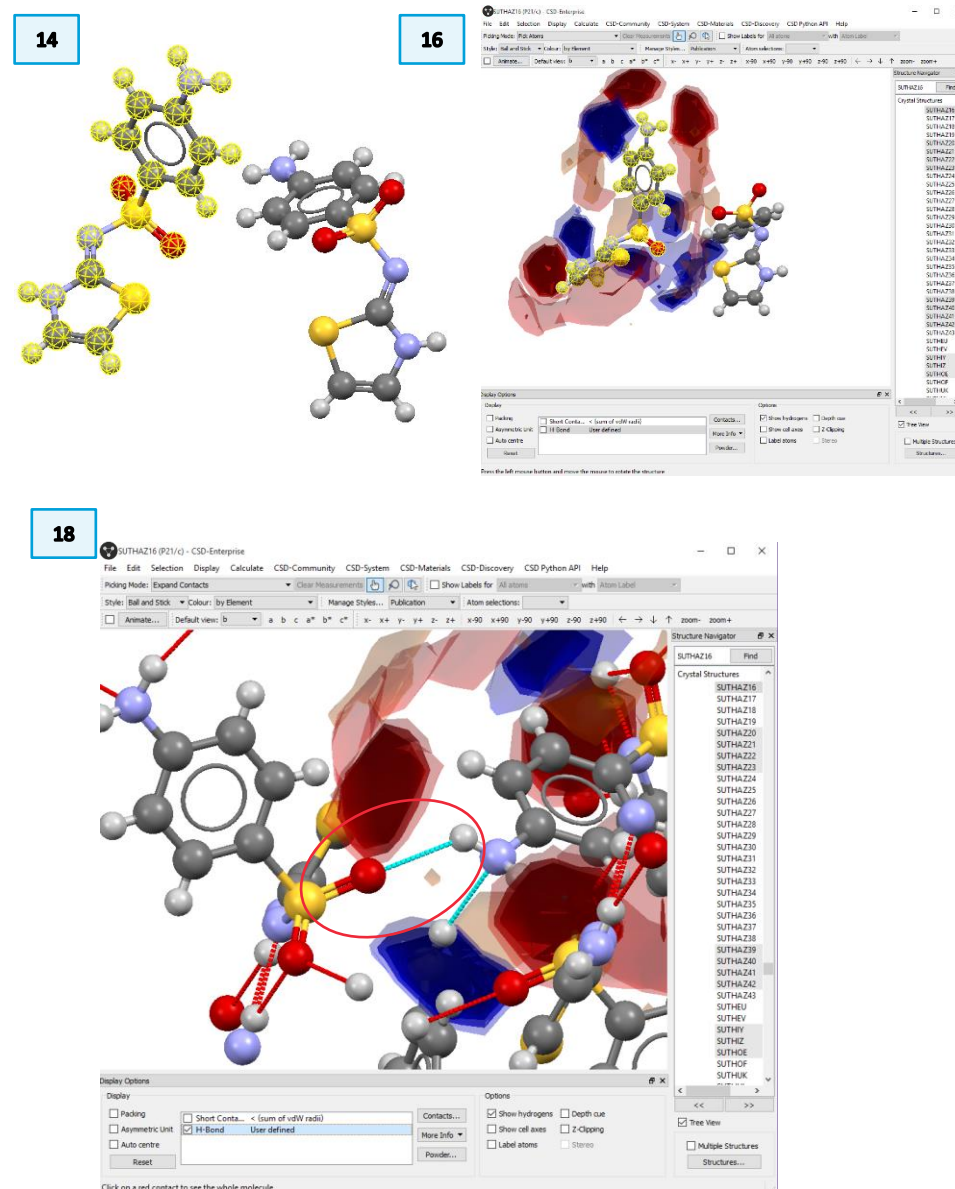
13

14. Hold the Shift key and click the molecule that contains N1. This will select the entire molecule.
15. If the *Full Interaction Maps* dialogue is still open, simply click **Calculate Maps** again. Otherwise, follow steps 2-4 to re-open the dialogue.
16. You should now have a Full Interaction Map surrounding molecule 1.
17. Following steps 6-9 above, turn on the hydrogen bonding interactions and click to expand the interaction around N1.
18. Notice that one of the three interactions falls well outside the predicted region for a hydrogen bond acceptor. This suggests that this interaction has a non-ideal geometry and is likely to be significantly less stabilising than the interactions in Form V.

Conclusions

The observed polymorphs of sulfathiazole exhibit different interactions, but also noticeably different geometries for those interactions. We can use knowledge-based approaches to compare the observed intermolecular interactions in two polymorphs with the preferred geometries for these interaction types. Full Interaction Maps indicate that Form V (known to be the most stable of the sulfathiazole polymorphs) has interactions which are near to ideal, whereas Form I (known to be metastable) has non-ideal interactions.

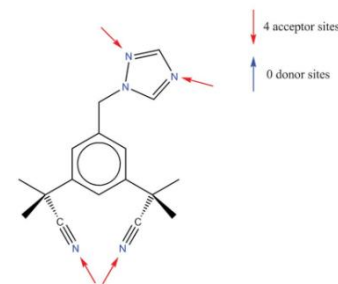
You should now know how to generate a Full Interaction Map of a molecule through Mercury and how to explore hydrogen bonding interactions in a crystal structure in relation to the predicted interaction landscape.



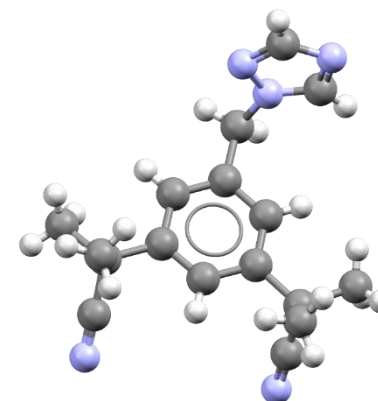
Example 4. Investigating the potential for co-crystallisation

For this tutorial, we will investigate the molecule anastrozole. This compound is a non-steroidal aromatase inhibitor used to treat breast cancer. The only published crystal structure is of the pure form. The molecule also contains no conventional hydrogen bond donors.

In this example, you will generate a Full Interaction Map for this molecule and investigate the interaction landscape. You will see how this information can be used to design a co-crystal.



2D Diagram of anastrozole and its CSD entry SATHOL



1. In the Mercury window, type SATHOL in the Structure Navigator toolbar to load the structure of anastrozole.
2. Click the *CSD-Materials* menu or *CSD-Discovery* menu and select *Full Interaction Maps...* from the dropdown menu.
3. Keep the default settings.
4. Click **Calculate Maps** to generate the Full Interaction Map for SATHOL.
5. The resulting map shows several areas for hydrogen bond donors (blue regions); two high-probability areas near the nitrogen atoms of the triazole ring, and two slightly lower probability regions near the cyano groups.
6. While there are no true hydrogen bond donors in this structure, the C-H groups in the triazole ring could interact with possible hydrogen bond acceptors (red regions).
7. There is also a small brown region indicating the possibility for a hydrophobic or π - π interaction.

The screenshots illustrate the workflow in Mercury:

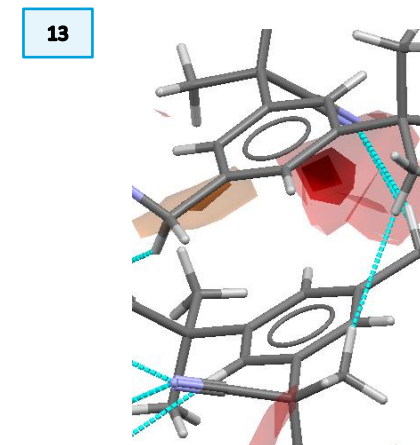
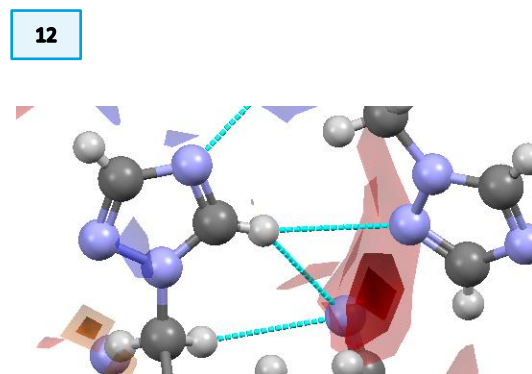
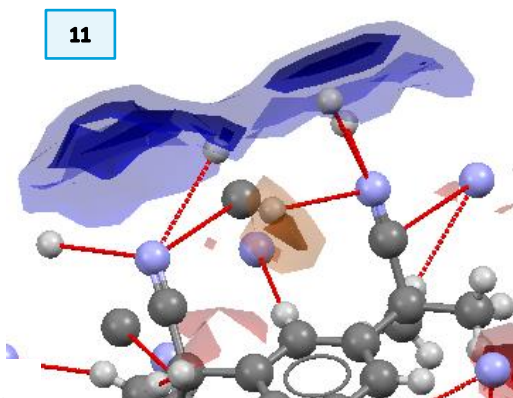
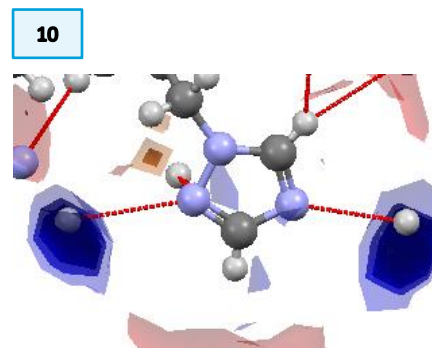
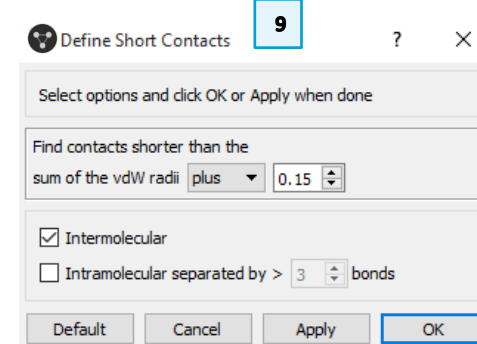
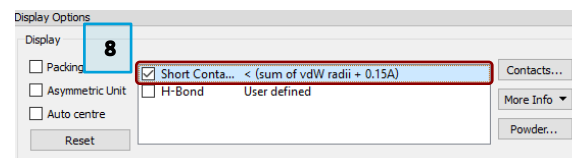
- Step 1:** The Mercury window shows the anastrozole structure loaded. The Structure Navigator toolbar at the top has 'SATHOL' entered.
- Step 2:** The 'CSD-Materials' menu is open, showing 'Full Interaction Maps...' selected under the 'cfc' (Crystallographic Frameworks) section.
- Step 3:** The 'Full Interaction Maps' dialog box is shown with default settings. The 'Map Contour Levels' section has 'Display first contour with initial level of 2.0', 'Display second contour with initial level of 4.0', and 'Display third contour with initial level of 6.0' all checked. The 'Hotspots' section has 'Generate hotspots in the map' unchecked. The 'Probe' list includes 'Uncharged NH Nitrogen', 'Charged NH Nitrogen', 'RNH3 Nitrogen', 'Alcohol Oxygen', 'Carbonyl Oxygen', 'Water Oxygen', 'Oxygen Atom', 'Methyl Carbon', 'Aromatic CH Carbon', 'C-F Fluorine', 'C-Cl Chlorine', 'C-Br Bromine', and 'C-I Iodine'. The 'Colour' scale is set to 'Defaults'.
- Step 4:** The 'Calculate Maps' button is highlighted at the bottom of the dialog.
- Step 5:** The resulting Full Interaction Map is displayed over the 3D model of anastrozole. Blue regions indicate potential hydrogen bond donor sites, red regions indicate potential hydrogen bond acceptor sites, and a small brown region indicates a potential hydrophobic or π - π interaction site.

8. As there are no conventional hydrogen-bond donors in this structure, there will be nothing to display by ticking the H-Bond box in the Display Options toolbar. Instead, tick the Short Contact box.
9. Double click this bar to open the *Define Short Contacts* dialogue box. Change this so that we find contacts shorter than the sum of the vdW radii **plus** 0.15Å. Click **OK** to save your changes.
10. Let's first look at the region around the triazole ring. We can see that there are two C-H donors that will satisfy the interaction preferences of the two N acceptors in the triazole ring. Click on the two contacts (dashed red lines) to generate the neighbouring molecules.
11. Looking at the region around the cyano groups, there are some short contacts available for interactions, but these are not directed toward the high-probability areas.
12. From the acceptor probe maps (red contours), we can see that the main region of acceptor preference is only weakly satisfied by the cyano group from one molecule and the triazole ring from another. Click the contacts to expand nearby molecules. The other acceptor regions are not satisfied at all.
13. Finally, there is a weak π — π interaction between the two phenyl rings that matches with one of the hydrophobic regions.

Conclusions

The Full Interaction Maps for SATHOL show that the highest probability regions for interactions, the donor regions for the triazole N atoms, are quite nicely satisfied. However, this comes at the cost of satisfying the interactions that would fulfill the acceptor sites. Since anastrozole has no classical H-bond donors of its own to fulfill those interactions, a molecule with one or more H-bond donors would be a good choice for a potential co-crystallisation agent.

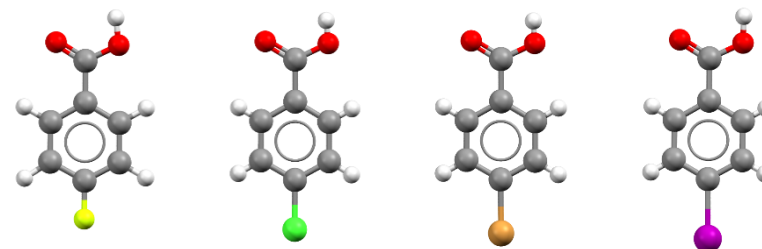
There are no co-crystals or solvates of anastrozole yet found in this CSD.



Example 5. Investigating a series of halogenated molecules

When designing molecules to fit a specific purpose, chemists will often make subtle changes to a single functional group; for example, replacing fluorine with chlorine, bromine or iodine. We can use Full Interaction Maps to investigate the differences in the interaction preferences of a molecule when these substitutions are made across a series.

This example will look at the series of halogenated benzoic acids to compare what differences there are across the series.



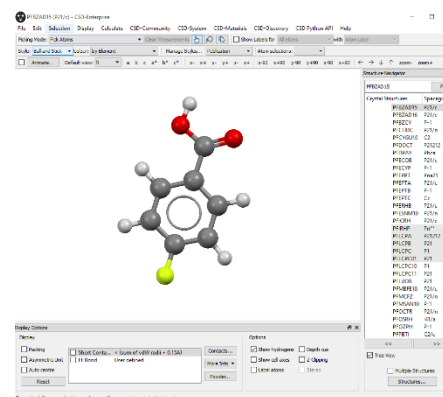
A series of halogenated benzoic acids

1. Launch Mercury and type PFBZAD15 in the Structure Navigator tool bar to bring up one of the parafluorobenzoic acid entries.
2. Click the *CSD-Materials* menu or *CSD-Discovery* menu and choose *Full Interaction Maps...* from the drop-down menu.
3. We will again use the default settings.
4. Click **Calculate Maps** to generate the Full Interaction Map for parafluorobenzoic acid.
5. Repeat steps 1-4 for the other molecules in this series. Open a new instance of Mercury each time so you can easily compare the results. Use refcodes CLBZAP10 for the Cl analogue, BRBZAP01 for the Br analogue, and BENMOW07 for the I analogue. These particular entries were chosen because they are relatively recent crystal structure determinations at low temperatures. To open multiple Mercury windows on a Mac, open a terminal window and run this command:

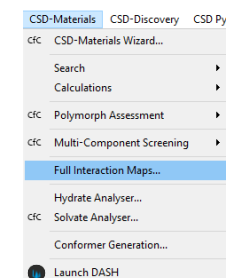
```
open -n -a "/Applications/CCDC/CSD_2019/mercury.app"
```

 This command tells the system to open a new instance (the "-n" flag) of an application (the "-a" flag) that is the given name in quotes.
6. Now let's compare the maps for each molecule in the series.

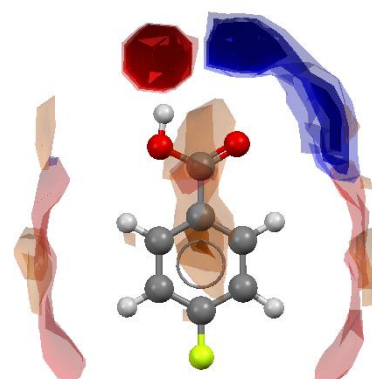
1



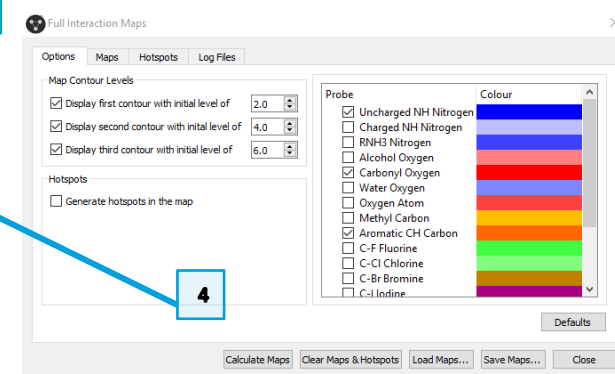
2



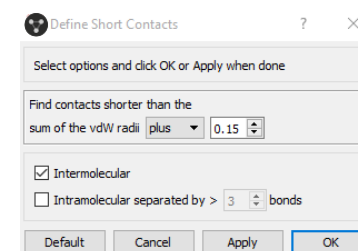
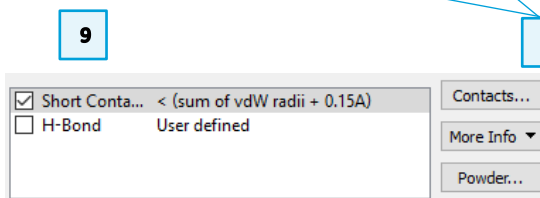
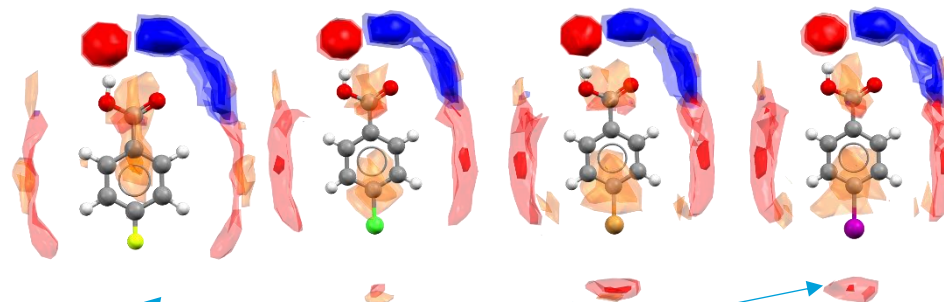
3



4



7. Look at the maps generated across the series. You should notice that as the size of the halogen increases across the series, there is an increased likelihood of finding an acceptor interacting with the halogen. This interaction is non-existent for the fluoro species, but fairly well-defined for the bromo and iodo analogues.
8. Now we will investigate how this affects the packing in each crystal structure.
9. Start with the iodo analogue. In the Display Options make sure the Short Contact settings are for “sum of vdW radii **plus** 0.15Å.” If not, double click that line to open the Define Short Contacts dialog and change the settings. Click **OK** to accept these changes. Then tick the box to display Short Contacts.
10. With the short contacts displayed, you should see that most of the interactions are satisfied. The two high-probability donor and acceptor sites near the carboxylic acid group are satisfied, as are the weaker acceptor sites near the aromatic hydrogen atoms on the phenyl ring. The iodine from a neighbouring molecule serves as the acceptor for a halogen-halogen interaction.
11. Click the interactions highlighted in Figure 10 to expand these neighbouring molecules. You will notice that while the acceptor (red) contours on either side of the phenyl ring are roughly equivalent, only one region is satisfied well by an interaction. Iodine atoms from neighboring molecules fill this region on the opposite side.



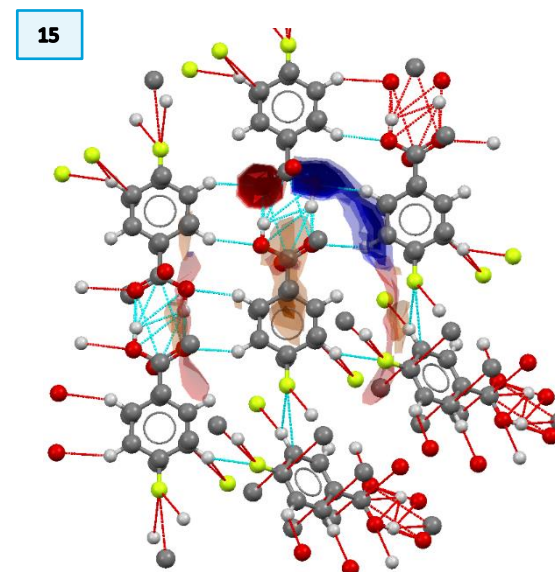
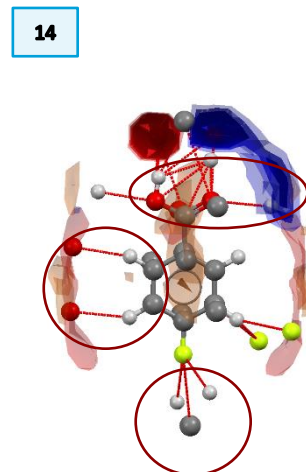
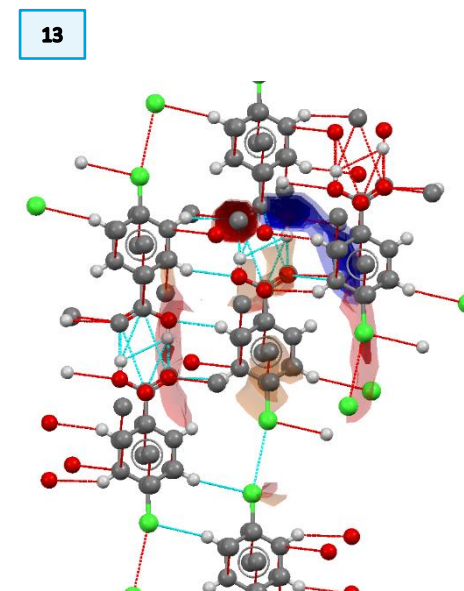
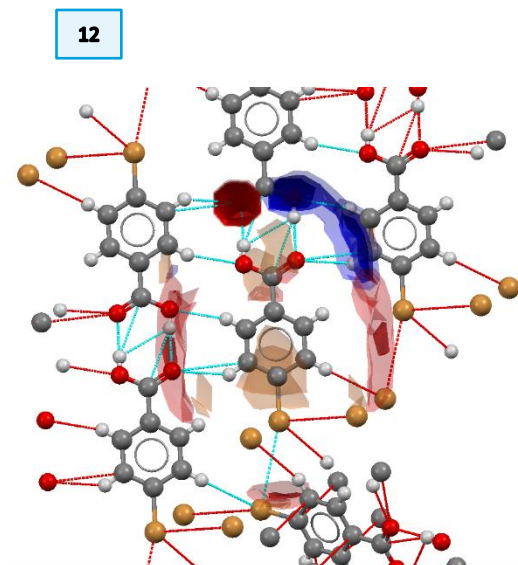
12. Repeat steps 9-11 for the Br and Cl analogues. What similarities and differences do you see across the series so far? What do you expect to see for the fluoro analogue?
13. For the fluoro analogue, follow step 9 to generate the short contacts.
14. You will see that the interactions for the carboxylic acid group and for the two weak C-H donors on the phenyl ring are satisfied. However, there is no probability of an acceptor for the fluorine, and consequently there is only C and H in this region.
15. Click these contacts to expand the neighbouring molecules. Now you can see that there is no halogen-halogen interaction involving the fluorine.

Conclusions

It is well understood that iodine forms halogen-halogen interactions more so than the other halogen atoms. Using Full Interaction Maps, we can clearly see how the probability of these interactions affects the overall packing in a crystal structure.

Further Exercises

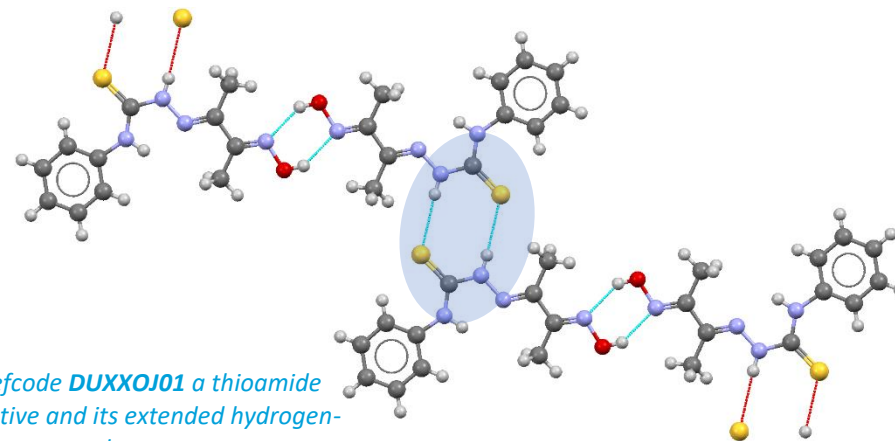
- Try calculating Full Interaction Maps using only the C-F and C-Cl probes. Do the results match your expectations?
- Investigate the hydrophobic interactions among the halogenated benzoic acid molecules in this series.
- Try looking at ortho- or meta- substituted halogenated benzoic acid species. Do your results match your expectations?
- Run a Full Interaction Maps calculation on your own molecules.
- **Advanced:** Try designing a ConQuest search to return just the molecules we used for this example. How close can you get? *Tips: You may want to try using the **Formula** and **Author/Journal** query types, too.*



Example 6. Investigating & analysing intermolecular interactions

Hydrogen-bonding interactions can be used in the design of extended solid-state structural arrangements, such as in co-crystals. In order to use intermolecular interactions in crystal engineering, it is important to understand the relative geometric preferences of those interactions. The CSD is a valuable source of information on intermolecular interactions and their geometries.

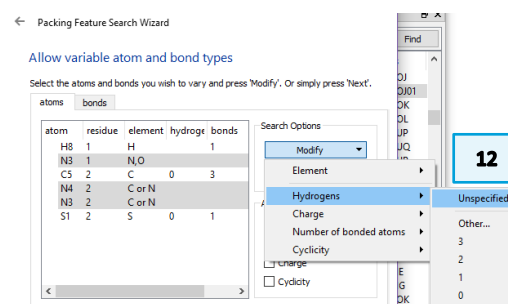
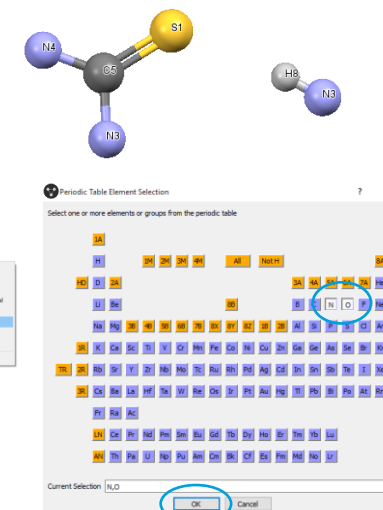
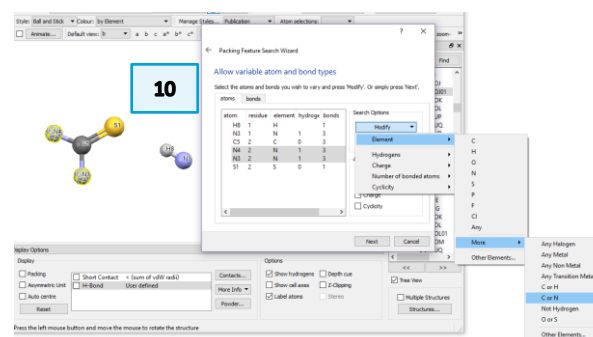
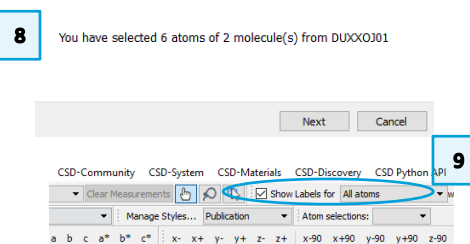
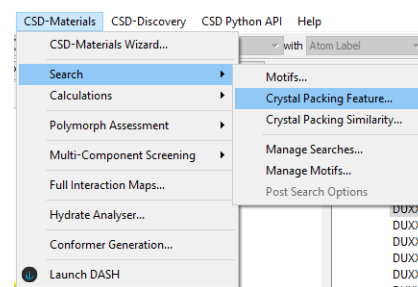
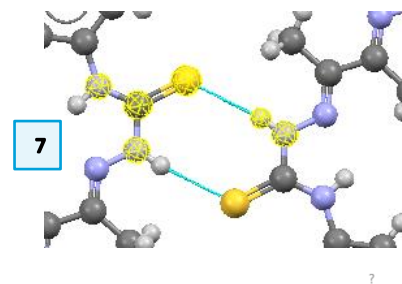
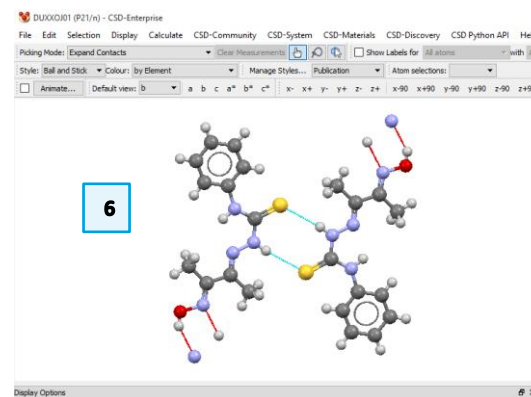
This example uses a thioamide derivative to demonstrate how to determine if the C=S moiety is a genuine hydrogen-bond accepting group, and if so, what geometries the hydrogen bonds to the C=S group might adopt. The crystal packing feature search tool in CSD-Materials allows users to search the CSD for a specific molecular feature in order to determine the most preferred geometries.



CSD refcode **DUXXOJ01** a thioamide derivative and its extended hydrogen-bonding network

1. Open your copy of Mercury and type the refcode DUXXOJ01 into the Structure Navigator toolbar.
2. The structure will appear in the Mercury interface. Edit the H-bond definition by double-clicking on the *H-Bond* line in the Display Options toolbar.
3. This will launch the *Define H-bonds* dialogue. Tick the box next to “Require hydrogen atom to be present”. Click **OK** to close the dialog box and return to the Mercury interface.
4. Tick the box next to “H-Bond” in the Display Options toolbar to turn on hydrogen-bond displays. This will show hydrogen bonding interactions as dotted red lines in the display.
5. Expand the hydrogen bonds around the C=S group by clicking on the atoms at the end of the dashed lines.

6. You should now see two molecules linked by the hydrogen bonds between C=S groups. This is the interaction that we will be investigating.
7. Click to select the atoms forming the interaction of interest in the structure, *i.e.* the thiourea – N-H donor group. Selected atoms will be highlighted in yellow. To de-select an atom, simply click it again.
8. Open the Crystal Packing Feature Wizard, by clicking *CSD-Materials* from the top menu, then *Search*, then *Crystal Packing Feature...* The wizard will tell you there are 6 atoms selected from 2 molecules of DUXXOJ01. Click **Next** to continue.
9. The Mercury display will now only show the atoms you selected; these may be hidden behind the dialogue box. In the top-level tool bar of the Mercury interface, tick the box to “Show Labels for All atoms” to label the atoms in the display. This will help in modifying atom properties in the following steps.
10. The wizard now allows you to define variable atom or bond types for the fragment interaction you are searching. Click to select the two nitrogen atoms bonded to the central carbon, N3 and N4 in the Mercury display. Then from the dropdown menu choose *Modify / Element / More / C or N*. This allows us to search for structures that have similar geometries but not necessarily the exact same elements.
11. Click to de-select the two nitrogen atoms from the previous step. Click to select N3 in the donor N-H fragment and then choose *Modify / Element / Other Elements...* and click **N** and **O** from the periodic table. Then click **OK**.
12. We also want to allow the number of hydrogen atoms and bonded atoms to vary on all three nitrogen atoms. To do this, click to select all three nitrogen atoms in the Mercury display. Choose *Modify / Hydrogens / Unspecified* and then *Modify / Number of bonded atoms / Unspecified* from the drop-down menu in the wizard.
13. Click **Next** in the wizard to continue.



14. The next step of the wizard allows you to change the tolerances or customise search parameters. If your search fragment is very common, you may want to reduce the tolerance level. However, for this example, we will use the defaults. Click **Next** to continue.

15. Now we need to specify the bond lengths and angles we want to investigate for this fragment. We want to define the H...S distance and the N-H...S and H...S=C angles.

16. Make sure all atoms are de-selected by clicking in the background of the Mercury display. Click to select the S atom, and then the H atom then click the button **Add Distance >** to add this contact distance to the query.

17. Click, in order, the C atom, the S atom and the H atom, and then **Add Angle >** to add this parameter to the search. Likewise, click, in order, the N atom, the H atom and the S atom, and then **Add Angle >** to add this parameter as well. Click **Next** to continue.

18. The next step allows you to choose which data to search. For this example, use all available data. To do this, click the headline **Databases** at the top of the list and then the right arrow, to move all data to the *Selected Structures* box. If you have your own in-house databases, you can add them to the search in this way. Click **Next** to continue.

19. The next window allows you to apply filters. These can be used to target specific structure types (e.g. only organic molecules) or to improve the structure quality of our hits (e.g. no disorder.) We will use the defaults (*i.e.* no filters) for this example. Click **Next** to continue.

20. Choose an appropriate *Packing feature name* and *Search name* and enter them in the boxes. Click **Start Search** to continue.

21. As the search runs, refcodes for each hit will appear in the *Searches* dialog box on the right of the Mercury interface. You can click to select one and use the arrow keys to search through them while the search is running.

14. Packing Feature Search Wizard. Allow tolerances on the geometry of the packing feature. Level of Geometric Similarity Required: ☐ Very High ☒ Medium ☐ Low ☐ Custom. Custom Tolerances: Distances must match those of selected feature to within + or - 30 %; Angles must match those of selected feature to within + or - 20 degrees. Match Options: ☐ All matched fragments must belong to same molecule. ☐ Show distances ☐ Show angles. **Next** Cancel.

15. Packing Feature Search Wizard. Select Parameters. Select atoms in the visualiser to select distance, angle or torsion parameters. Or simply press Next. Current Selection: S1, H8. Add Distance >. Parameter List: S1, H8. **Next** Cancel.

16. Packing Feature Search Wizard. Select Parameters. Select atoms in the visualiser to select distance, angle or torsion parameters. Or simply press Next. Current Selection: N3, H8, S1. Add Angle >. Parameter List: S1, H8; C5, S1, H8; N3, H8, S1. **Next** Cancel.

17. Packing Feature Search Wizard. Select Parameters. Select atoms in the visualiser to select distance, angle or torsion parameters. Or simply press Next. Current Selection: N3, H8, S1. Add Angle >. Parameter List: S1, H8; C5, S1, H8; N3, H8, S1. **Next** Cancel.

18. Packing Feature Search Wizard. Select search structures. Use the buttons to select/deselect items you wish to use. Available Structures: Databases. Selected Structures (1009141): CSD 5.40 Feb 19 May 19. **Next** Cancel.

19. Packing Feature Search Wizard. Select Search Options. Filters: ☐ 3D coordinates determined ☐ Not disordered ☐ No errors ☐ Not polymeric ☐ No powder structures. ☐ R factor ☒ <= 0.05 ☐ <= 0.075 ☐ <= 0.1. ☐ Only ☒ Organics ☐ Organometallic. **Next** Cancel.

20. Packing Feature Search Wizard. Enter search name. Packing feature name: C=S...Donor. Search name: Example. **Start Search** Cancel.

21. Mercury interface showing the search results. The Searches dialog box on the right lists refcodes for each hit. The main window shows the molecular structure with the packing feature highlighted.

22. When the search is done, click **Spreadsheet...** at the bottom of the Search toolbar window to open the **Data Analysis** window for analysing the results.

23. All of the parameters selected during the set-up procedure will be listed in the Data Analysis spreadsheet.

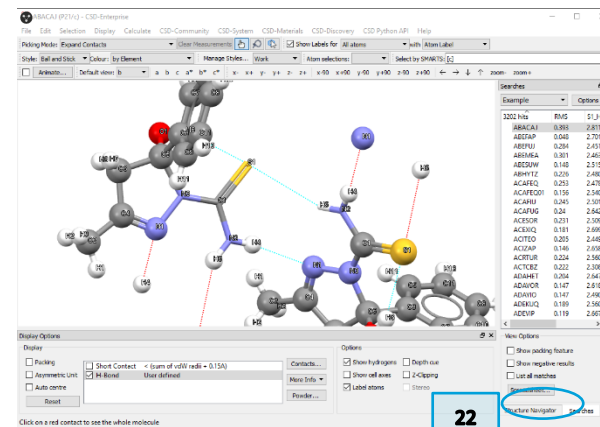
24. To plot the data, select the **Plots** menu, and the type of plot you wish to use. For this example, we will start with a heat plot. Choose **Plots / Heat plot**.

25. Choose the $S\cdots H$ distance for the X Axis and the $C=S\cdots H$ angle for the Y Axis. Click **OK** to close the dialog box.

26. The heat plot shows a strong geometric preference for a $H\cdots S$ contact distance of 2.5 to 2.6 Å and a $H\cdots S=C$ angle of 106° to 114° . This $H\cdots S$ contact distance is considerably less than the sum of the van der Waals radii of the atoms that are involved in the interactions (1.8 Å for S + 1.2 Å for H). This indicates a hydrogen bonding interaction.

27. Return to the spreadsheet view by clicking the appropriate tab. Choose **Plots / Scatterplot**. Choose the $S\cdots H$ distance for the X axis and the $N-H\cdots S$ angle for the Y axis. Click **OK** to continue.

28. There is a dense cluster of structures with $H\cdots S$ contact distances in the region of 2.5 to 2.6 Å and more linear $N/O-H\cdots S$ bond angles in the region of 160° to 175° indicating that the C=S group is a common and good hydrogen-bond acceptor.



23

Data Analysis

Example Spreadsheet 1

Identifier	NAME	rmsd	S1_H8	C5_S1_H8	N3_H8_S1
Example ABACAJ0	ABACAJ	0.3930	2.8119	95.6533	145.8700
Example ABEFAP1	ABEFAP	0.0480	2.7018	119.3410	158.8010
Example ABEFUJ2	ABEFUJ	0.2840	2.4512	101.9800	166.0500
Example ABEMEJ3	ABEMEA	0.3010	2.4632	100.5170	158.6540
Example ABESUW4	ABESUW	0.1480	2.5155	113.4160	169.7640
Example ABHYTZ5	ABHYTZ	0.2260	2.4805	105.0080	164.7620
Example ACAFEQ6	ACAFEQ	0.2530	2.4783	107.4340	162.8120
Example ACAFEQ017	ACAFEQ01	0.1560	2.5407	111.6430	167.9220
Example ACAFUJ8	ACAFUJ	0.2450	2.5015	106.3400	159.8970
Example ACAFUJ9	ACAFUG	0.2400	2.6423	104.7910	157.3130
Example ACESOR10	ACESOR	0.2310	2.5090	108.4700	164.1930
Example ACEXIQ11	ACEXIQ	0.1810	2.6993	111.4470	161.8570
Example ACIZAP12	ACIZAP	0.1460	2.6581	112.4840	165.1030
Example ACRTUR13	ACRTUR	0.2240	2.5610	105.8560	170.4520
Example ACTCBZ14	ACTCBZ	0.2220	2.3081	108.6620	168.9780
Example ADAHET15	ADAHET	0.2040	2.6474	109.4920	171.3500
Example ADAVOR16	ADAVOR	0.1470	2.6182	116.6920	170.4410
Example ADAYQI17	ADAYIO	0.1470	2.4809	114.4590	171.9710
Example ADEKUQ18	ADEKUQ	0.1890	2.5603	113.7590	169.0110
Example ADEVIP19	ADEVIP	0.1190	2.6672	110.9840	160.6410
Example ADEZAL20	ADEZAL	0.1870	2.5411	108.5680	165.9590
Example ADEZAL021	ADEZAL02	0.1980	2.5230	108.2900	165.3580

24

Example Spreadsheet 1

Identifier	NAME	rmsd	S1_H8	C5_S1_H8	N3_H8_S1
Example ABACAJ0	ABACAJ	0.3930			
Example ABEFAP1	ABEFAP	0.0480			
Example ABEFUJ2	ABEFUJ	0.2840			
Example ABEMEJ3	ABEMEA	0.3010	2.4632	100.5170	158.6540

Plots menu options: Histogram, Polar histogram, Scatterplot, Polar scatterplot, Heat plot (selected).

25

Choose descriptors

Choose two descriptors

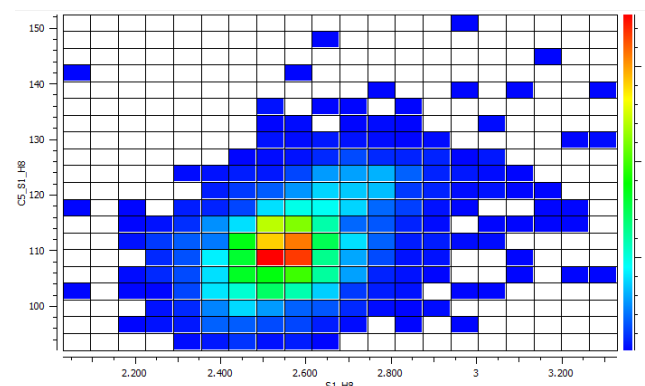
All descriptors...

X Axis: S1_H8

Y Axis: C5_S1_H8

OK Cancel

26



27

Plots menu options: Histogram, Polar histogram, Scatterplot (selected), Polar scatterplot, Heat plot.

Choose between two and four descriptors

All descriptors...

X axis: S1_H8

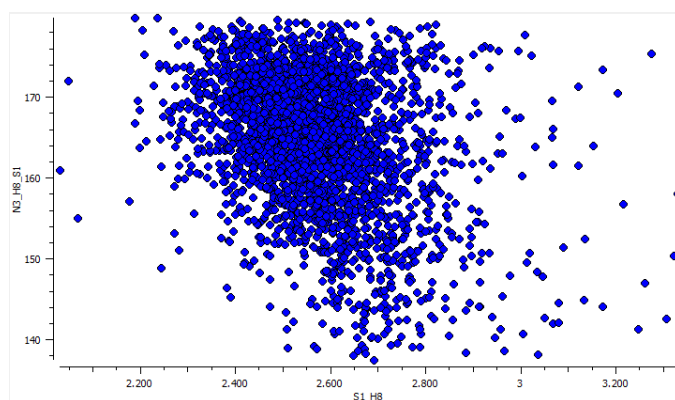
Y axis: N3_H8_S1

Colour: rmsd

Size: rmsd

OK Cancel

28



29. An individual structure in the cluster can be visualised in the Mercury display by clicking on a spot in the plot. The specific entry in the spreadsheet will also be highlighted.

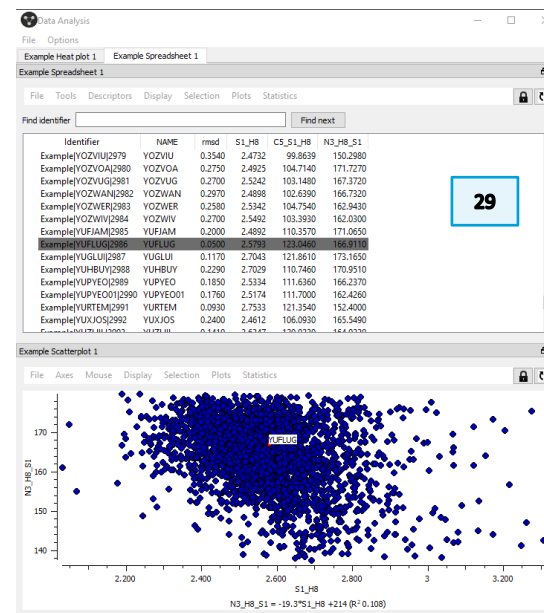
30. Visual inspection of the structures in the densest part of the distribution shows that most acceptors have one or two nitrogen atoms covalently bonded to the C=S carbon.

Conclusion

You should now be familiar with the *Crystal Packing Feature* search and analysing data through the *Data Analysis* tool.

This example shows a high frequency of hits with a particular geometry. This clustering suggests that C=S groups can act as hydrogen-bond acceptors with clear directional preferences. Most observations of C=S as a hydrogen-bond acceptor are in systems where there are one or two nitrogen atoms covalently bonded to the carbon (thiourea or thioamide derivatives) which significantly enhances the accepting capability of the sulfur atom. The preferred interaction geometry for this moiety is an H...S distance of 2.5-2.6 Å and an H...S=C angle of approximately 110°. This is quite different to the equivalent geometries of C=O acceptors, which show distances of 1.8-1.9 Å and angles around 130°.

When running a packing feature search on a compound or moiety of interest, look for hits that cluster in specific distance and angle ranges. This suggests the most likely geometry a particular feature will adopt. This information can help guide development of co-crystals and other supramolecular assemblies.



30

