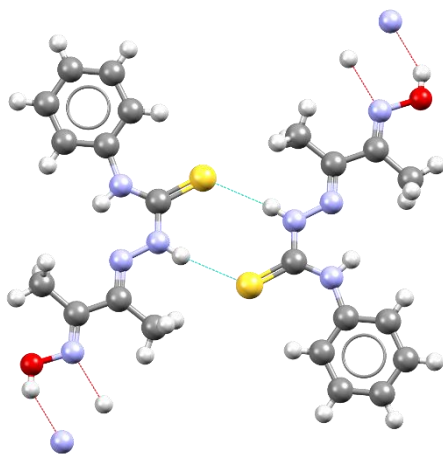


Search (Motifs, Crystal Packing Feature, Crystal Packing Similarity MAT-006)

Developed using
2020.0 CSD Release



CCDC
advancing structural science

Table of Contents	
Introduction.....	2
Objectives.....	2
Pre-required skills.....	2
Materials.....	2
Example 1. Motif searches for Primary Amides	3
Conclusion	8
Example 2. Crystal Packing Feature - Investigating & analysing intermolecular interactions.....	9
Conclusion	13
Example 3. Crystal Packing Similarity	14
Further exercises	17
Conclusions.....	18
Next steps.....	18
Feedback	18

Introduction

This workshop shows you how to analyse hydrogen-bonding interactions and packing trends in crystal structures using the features under the CSD-Materials>Search. The motifs, packing feature and packing similarity searches will allow you to easily interpret packing trends by quick comparison with data available in the CSD.

Objectives

In this workshop you will learn how to:

- Use the Motifs feature and search for hydrogen bond motifs for a specified set of functional groups
- Analyse unusual H-bond interaction using Crystal Packing feature
- Use Crystal Packing Similarity tool to analyse structures that form solvates

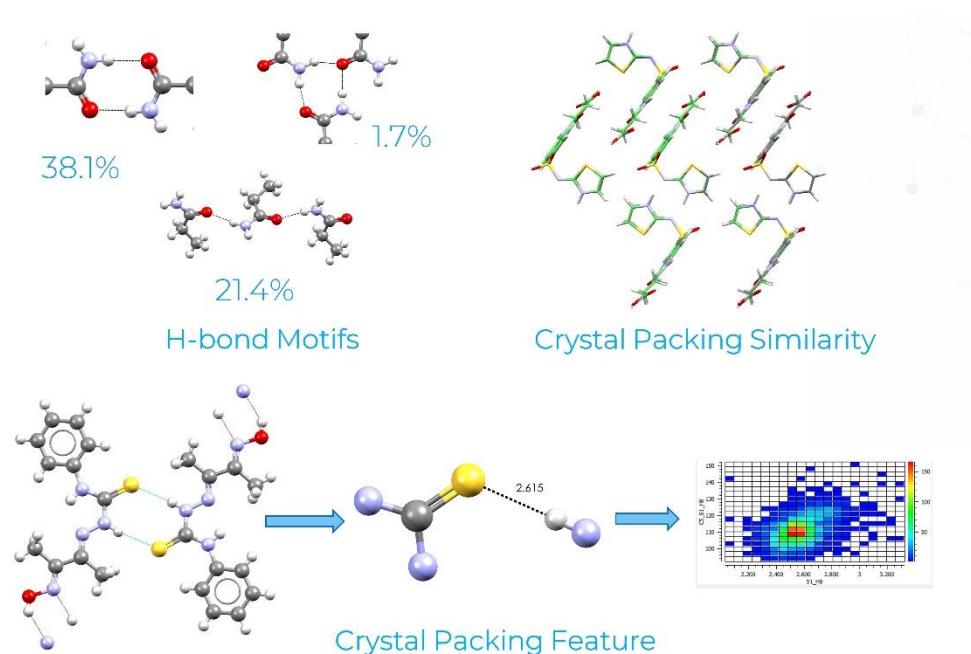
This workshop will take approximately **2.5** hours to be completed for all three examples. The examples are not connected, feel free to work on one example and return later for the other examples.

Pre-required skills

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. You can find the Visualisation and other Mercury workshops in the [CSD-Materials workshop area on our website](#).

Materials

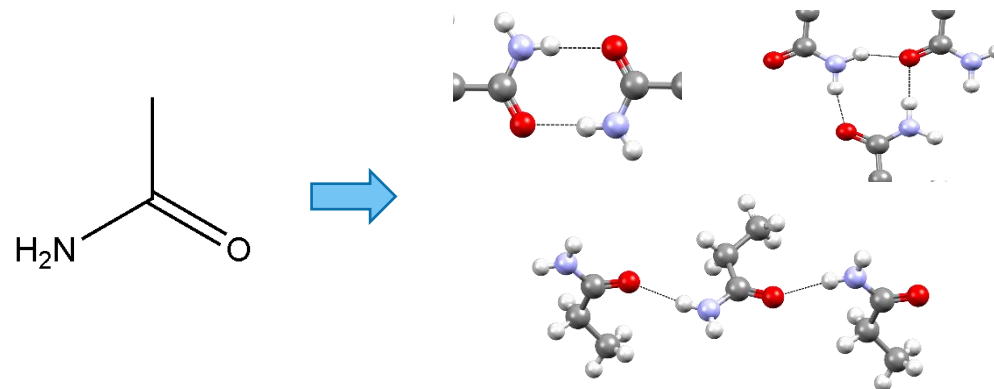
There are no additional materials required for this workshop.



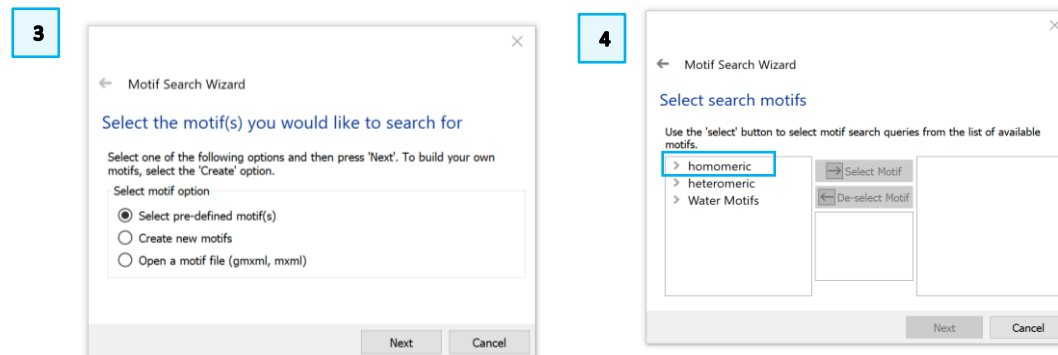
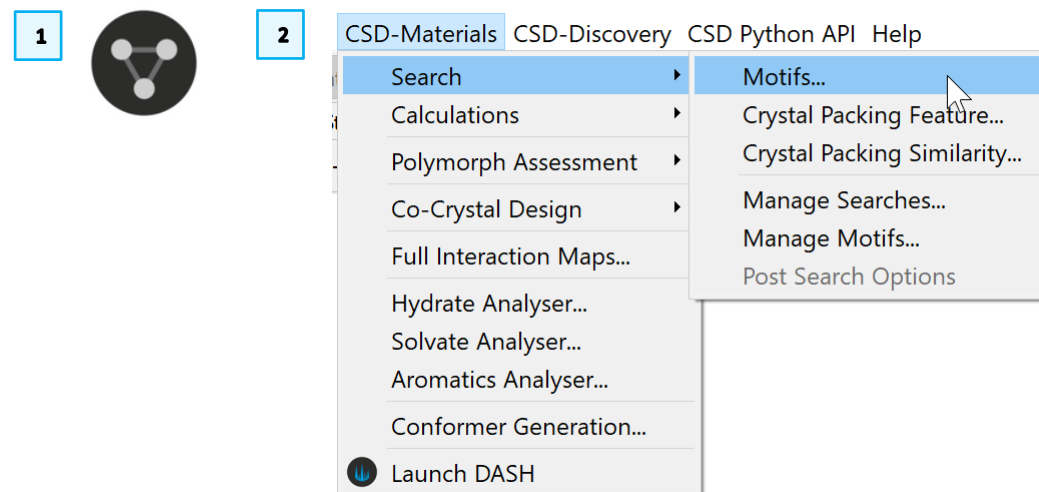
Example 1. Motif searches for Primary Amides

The Motif searches functionality in CSD-materials determines the likelihood of hydrogen bond motifs for a specified set of functional groups. Using this functionality, you can assess different motifs by their relative frequency of occurrence in the CSD. The tool allows to search for auto-generated and bespoke motifs. Also, motif searches can be used to analyse the results of crystal structure prediction runs by identifying the range of predicted motifs.

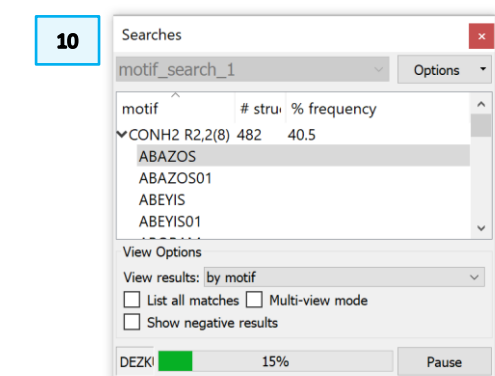
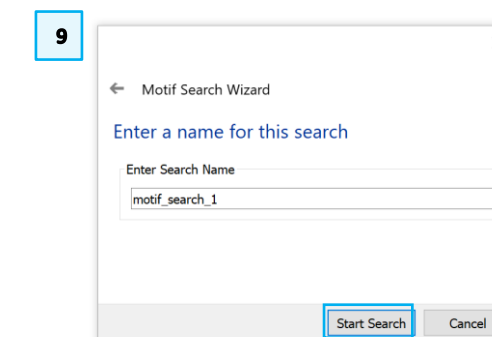
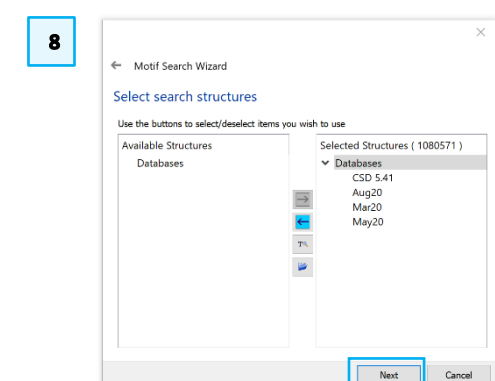
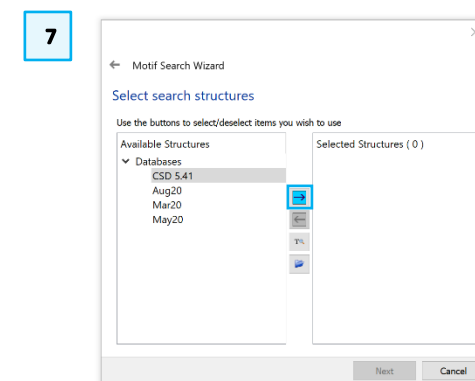
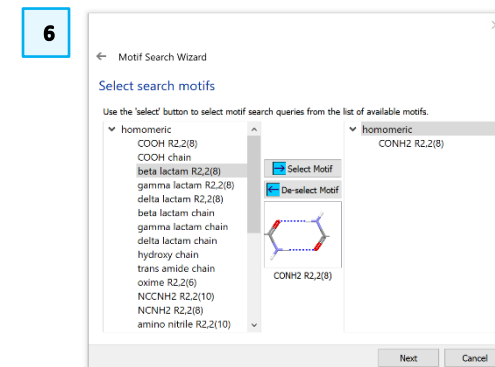
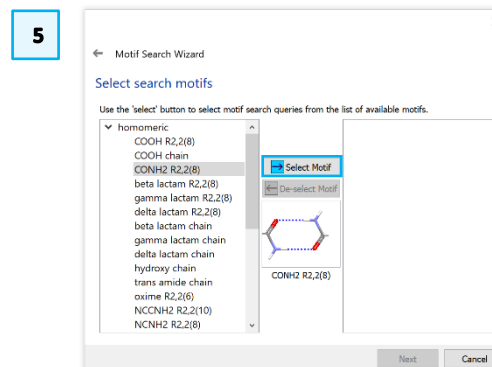
In this example we will use the Motifs functionality to search for Primary Amides and identify their frequencies occurrence in CSD.



1. Open Mercury by double-clicking the Mercury icon on the desktop
2. From the top-level menu select **CSD-Materials > Search > Motifs...** to launch the *Motif Search Wizard* dialog box
3. In the *Motif Search Wizard* dialog box select the **Select pre-defined motif(s)** option to choose a pre-defined motif then click Next.
4. Click on **homomeric** to select the pre-defined motif.

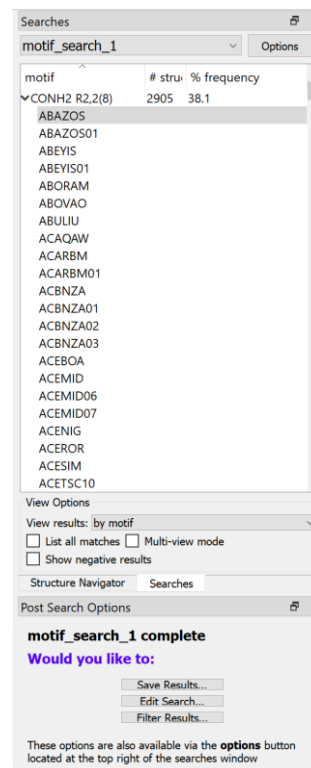


5. Under the **homomeric** list select the CONH2 R2,2(8), then click on Select Motif button.
6. The CONH2 R2,2(8) motif should be listed on the right-hand side column. Select **Next** to continue.
7. Select the *Databases* available and then click on the black arrow of the blue box button.
8. The selected Databases should appear on the right-hand side list. Select **Next** to continue.
9. You can enter a search name if you wish, for this example we will leave the default name as motif_search_1. To start the search, click the **Start Search** button.
10. The Searches have now started, and you can notice a progression bar in the bottom right corner. Wait for the bar progression to reach 100% in order to complete the search.

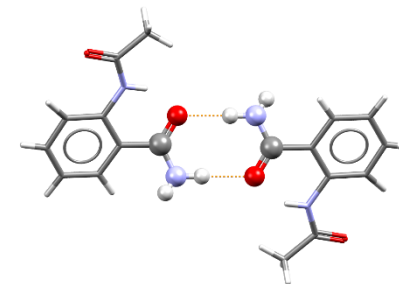


11. Once the search is completed all refcodes identified are listed in the window wizard. For the CONH2 R2,2(8) motif, at least 2905 structures were identified which represent an overall frequency of 38.1%.
12. Scroll down through the refcode list to analyse the type of molecules that contain the CONH2 R2,2(8) motif. We are illustrating here the presence of CONH2 R2,2(8) motif in o-Acetamidobenzamide, refcode ACBNZA.
13. Repeat steps 2 to 4 above and select the CONH2 R2,3(8) motif to perform a search for this type of interactions. Click on Select Motif button and the motif should be listed on the right-hand side column. Select **Next** to continue. In the next step ensure CSD is selected (as in steps 8-9 above) and select **Next** to continue.
14. You can enter a search name if you wish, for this example we will leave the default name as motif_search_2. To start the search, click the **Start Search** button
15. The Searches have now started, and you can notice a progression bar in the bottom right corner. Wait for the bar progression to reach 100% in order to complete the search.

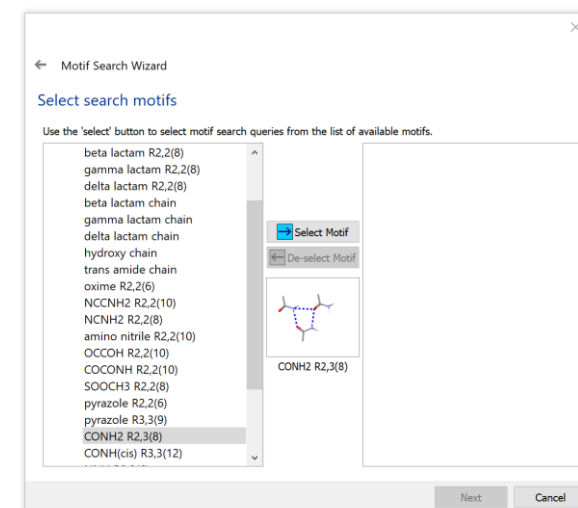
11



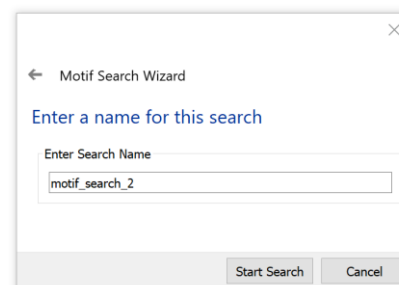
12



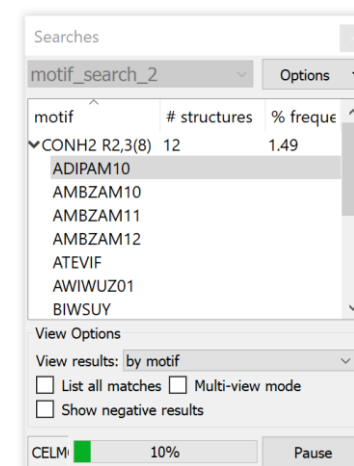
13



14

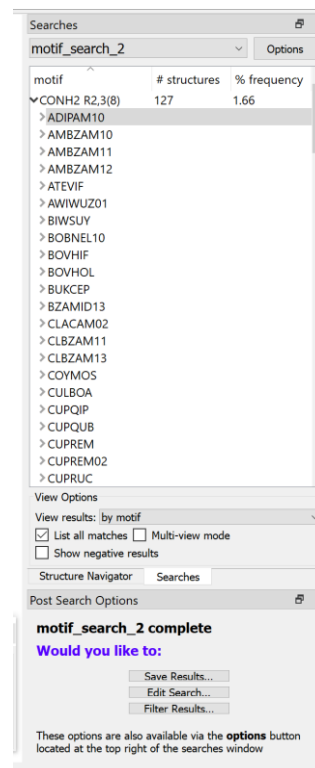


15

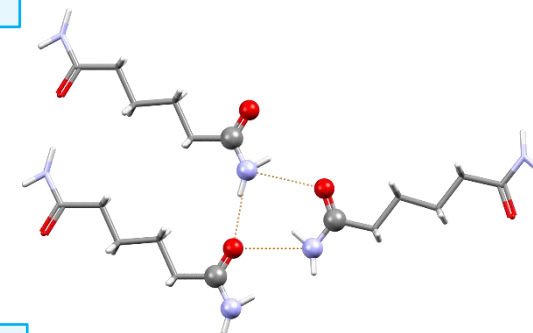


16. Once the search is completed all refcodes identified are listed in the window wizard. For the CONH2 R2,3(8) motif at least 127 structures were identified which represent an overall frequency of 1.66%.
17. Scroll down through the refcode list to analyse the type of molecules that contain the CONH2 R2,2(8) motif. We are illustrating here the first refcode from the list ADIPAM10 which is the crystal structure for Adipamide.
18. In the next search we will identify the frequency in the CSD of trans amide chain. Repeat step 2 above and then select Create new motifs.
19. To generate the trans amide motif select the **Add...** button
20. In the Select substructure window click on the Substructure menu and select the carbamoyl_2 substructure the click **OK** button to continue.
21. The carbamoyl substructure should be displayed in the Motif Search Wizard window, click **Next** to continue

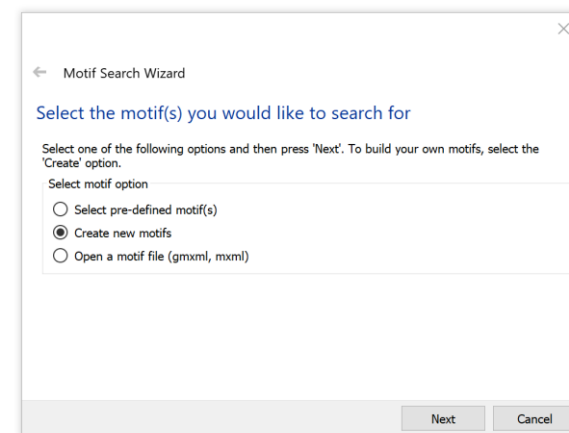
16



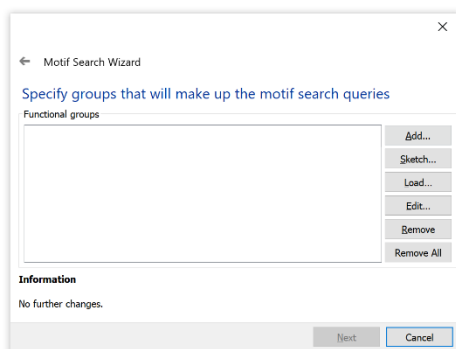
17



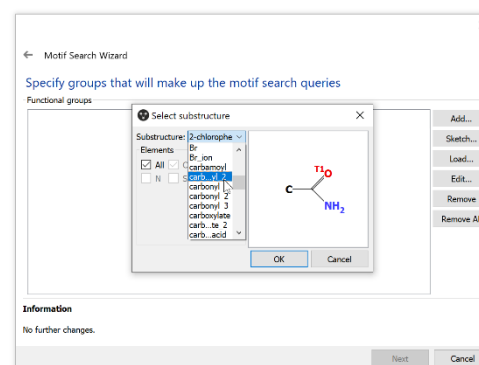
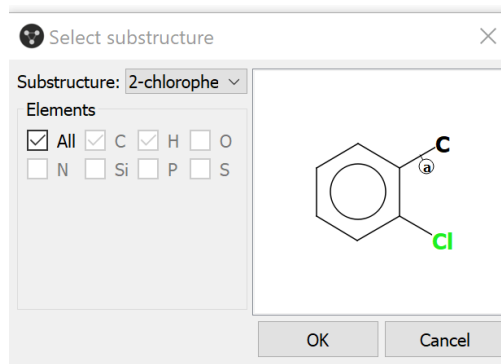
18



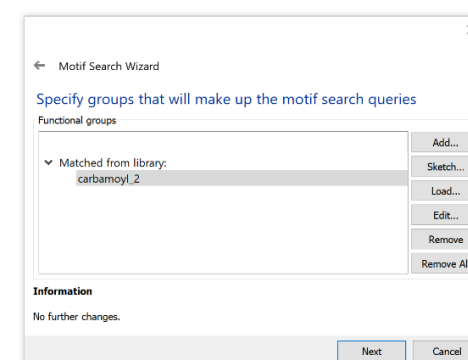
19



20



21



22. Select **Add...** to select the contacts present in the motif.

23. In the From column select carbamoyl_2 in the Group drop down menu and N atom will be displayed in the Atom drop down menu. In the To column select carbamoyl_2 in the Group drop down menu and O atom. The interaction will be illustrated in red in the display window. Click **OK** to continue

24. The specified contact will be illustrated in the wizard window. Select **Next** to continue

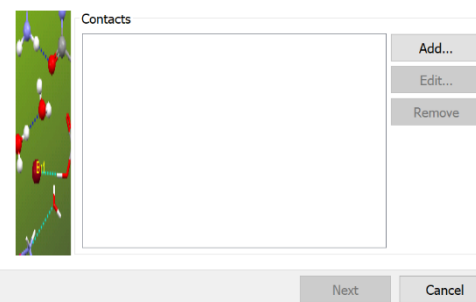
25. Select infinite chains that repeat every 1, 2 and 3 unit the select **Next** to continue.

26. Note that 6 motif search queries were generated containing the chains to search for. Select **Next** to continue.

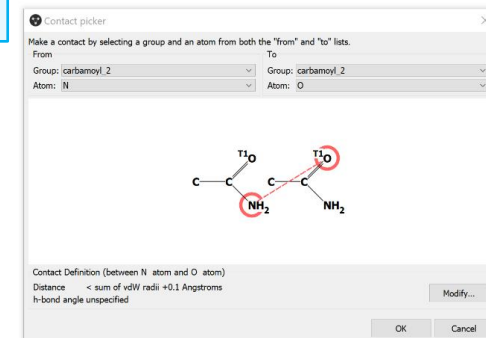
27. Ensure that the CSD is selected as shown in steps 7-8 above. You can enter a search name if you wish, for this example we will leave the default name as motif_search_3. To start the search, click the **Start Search** button.

22

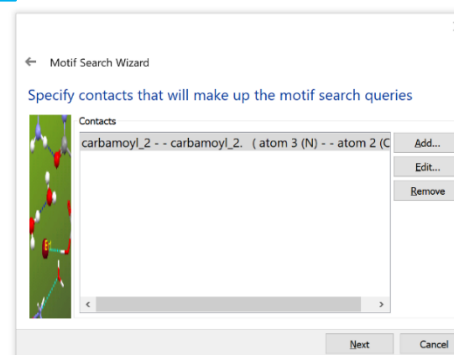
Specify contacts that will make up the motif search queries



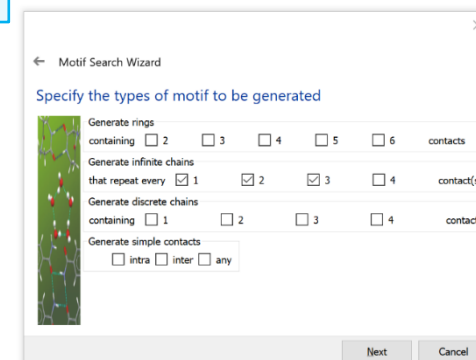
23



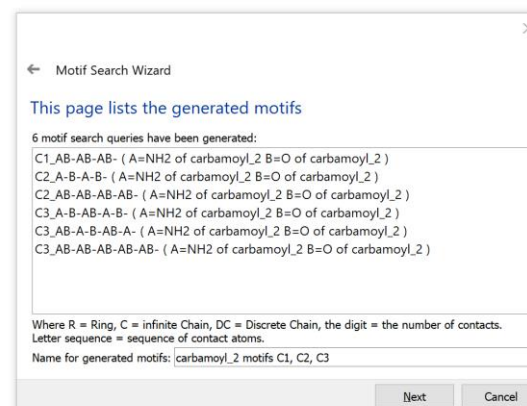
24



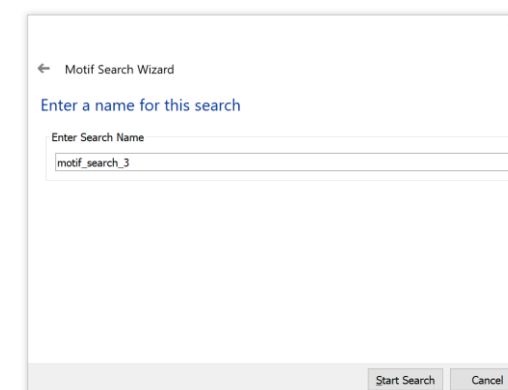
25



26



27

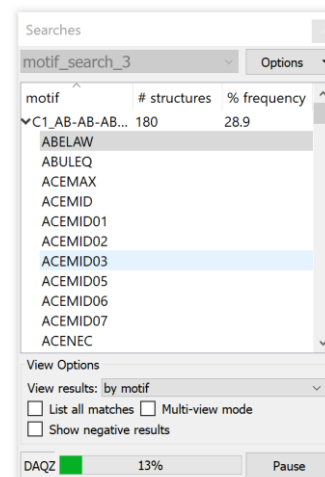


28. The Searches have now started, and you can notice a progression bar in the bottom right corner. Wait for the bar progression to reach 100% in order to complete the search.
29. Once the search is completed all refcodes identified are listed in the window wizard. For the trans amide chain motif at least 1008 structures were identified which represent an overall frequency of 21.4%.
30. Scroll down through the refcode list to analyse the type of molecules that contain the trans amide chain motif. We are illustrating here the first refcode from the list ABELAW.

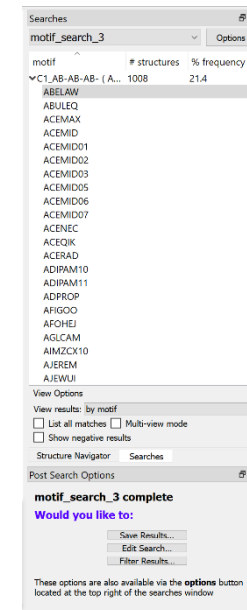
Conclusion

The Motif Search based on intermolecular interaction was used here to identify the abundance of different H-bond arrangements present in the CSD. We identified the frequencies of occurrence in CSD for primary amides and noted that the CONH2 R2,2(8) motif is the most commonly observed (38.1%, *as of the 2020.0 CSD release*). In this motif the amide utilises its donors and acceptors to form dimers. The trans amide forming a chain H-bonding motif is the second commonly observed (21.4%) and the CONH2 R2,3(8) is the least observed with only 127 structures found (1.7% frequency).

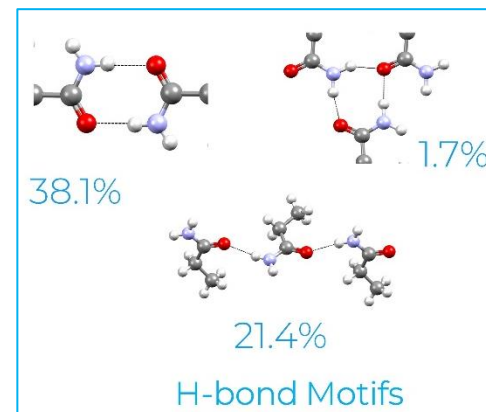
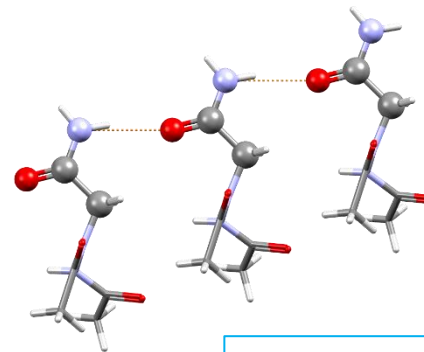
28



29



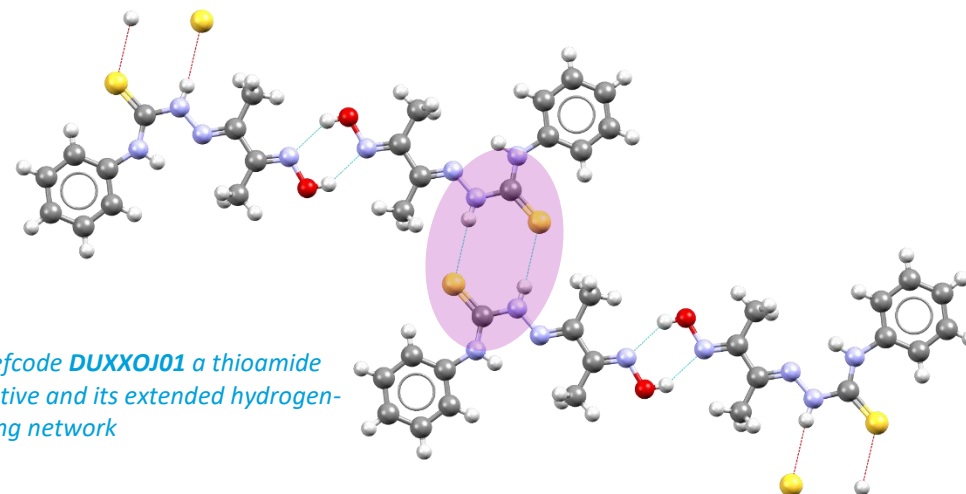
30



Example 2. Crystal Packing Feature - Investigating & analysing intermolecular interactions.

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 to determine the most preferred geometries.



1. Open Mercury by double-clicking the Mercury icon on the desktop.
2. In the Structure Navigator tool bar type the DUXXOJ01 refcode.
3. 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.
4. 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.
5. 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
6. Expand the hydrogen bonds around the C=S group by clicking on the atoms at the end of the dashed lines.

7. You should now see two molecules linked by the hydrogen bonds between C=S groups. This is the interaction that we will be investigating.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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**.
13. 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.

7

8

9

10

11

12

13

14. Click **Next** in the wizard to continue.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

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

15 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

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: S1, H8
Add Distance >
Parameter List: S1, H8
Delete Rename
Next Cancel

18 Packing Feature Search Wizard
Select Parameters
Select atoms in the visualiser to select distance, angle or torsion parameters. Or simply press Next.
Current Selection: R3, H8, S1
Add Angle >
Parameter List: S1, H8
C5, S1, H8
R3, H8, S1
Delete Rename
Next Cancel

19 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

20 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

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

22 Mercury
File Edit Selection Display Calculate CSD-Community CSD-System CSD-Molecules CSD-Dictionary CSD-Utilities Help
Display: Ball and Stick Colour: by Element
Style: Ball and Stick Colour: by Element
Animate... Default view: b a b c a* b* c*
Atoms: Default view: b a b c a* b* c*
Display: ☐ Hitting ☐ Show Hydrogens ☐ Show negative results ☐ Show packing feature ☐ Show packing feature
☐ Parameters only ☐ Show contacts ☐ Show contacts ☐ Show contacts ☐ Show contacts
☐ Auto centre ☐ Auto centre ☐ Auto centre ☐ Auto centre
Reset
Click on a cell number to see the chosen molecule

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

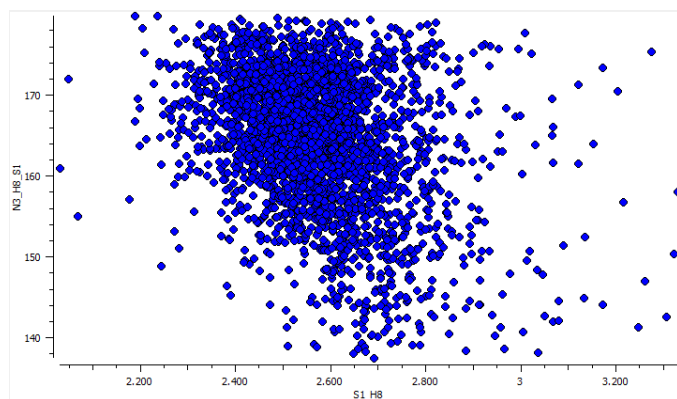
24. To plot the data, select the **Plots** menu, and the type of plot you wish to use. For this example, we will start with a heat plot. Choose **Plots / Heat plot**.

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

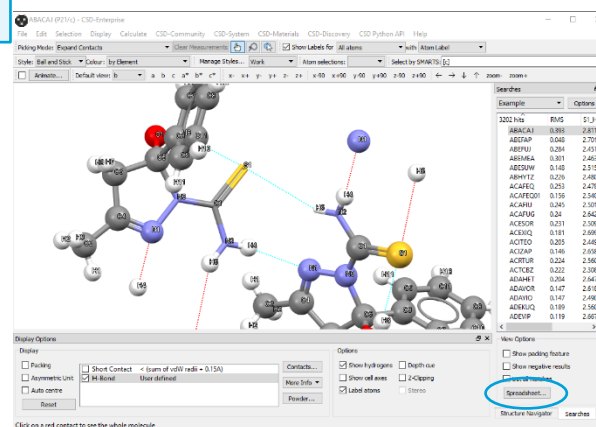
26. The heat plot shows a strong geometric preference for a H...S contact distance of 2.5 to 2.6 Å and a H...S=C angle of 106° to 114°. This H...S contact distance is considerably less than the sum of the van der Waals radii of the atoms that are involved in the interactions (1.8 Å for S + 1.2 Å for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

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

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



22



23

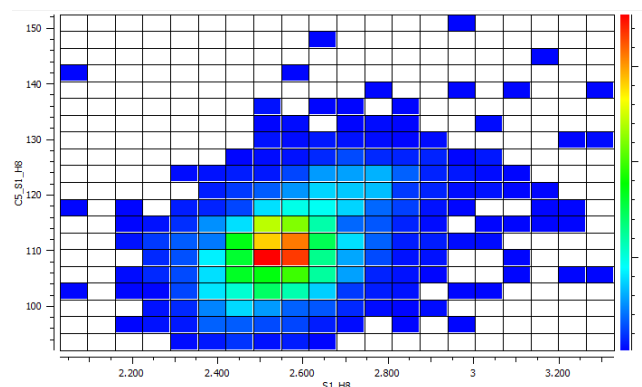
Identifier	NAME	rmsd	S1_H8	C5_S1_H8	N3_H8_S1
Example ABACAIJ0	ABACAI	0.3930	2.8119	95.6333	145.8700
Example ABEFAP1	ABEFAP	0.0480	2.7018	119.3410	158.8010
Example ABEFUJ2	ABEFUJ	0.2840	2.4512	101.9800	166.0500
Example ABEMEJ3	ABEMEA	0.3010	2.4632	100.5170	158.6540
Example ABESUW4	ABESUW	0.1480	2.5155	113.4160	169.7640
Example ABHYTZ5	ABHYTZ	0.2260	2.4805	105.0080	164.7620
Example ACAFQJ6	ACAFQJ	0.3330	2.4783	107.8340	162.8120
Example ACAFQJ7	ACAFQJ	0.1560	2.5407	111.6430	167.8220
Example ACAFUJ8	ACAFUJ	0.2430	2.5015	106.3400	159.8870
Example ACAFUJ9	ACAFUJ	0.2400	2.6423	104.7910	157.2130
Example ACESR10	ACESR	0.2310	2.3090	108.4700	164.1930
Example ACFQJ11	ACFQJ	0.1810	2.6993	111.4470	161.8570
Example ACFQJ12	ACFQJ	0.1460	2.6581	112.4840	165.1030
Example ACFQJ13	ACFQJ	0.2240	2.5610	105.8560	170.4520
Example ACFQJ14	ACFQJ	0.2220	2.3881	108.8620	168.9780
Example ADAFUJ15	ADAFUJ	0.2440	2.6474	109.8020	171.3300
Example ADAFUJ16	ADAFUJ	0.1470	2.6182	116.6920	170.4410
Example ADAFUJ17	ADAFUJ	0.1470	2.4909	114.4590	171.9710
Example ADAFUJ18	ADAFUJ	0.1890	2.5603	113.7590	169.0110
Example ADAFUJ19	ADAFUJ	0.1190	2.6072	110.8840	160.6410
Example ADEZAL20	ADEZAL	0.1870	2.5411	108.5680	165.9590
Example ADEZAL21	ADEZAL	0.1980	2.5230	108.2900	165.3380

24

Identifier	NAME	rmsd	S1_H8	C5_S1_H8	N3_H8_S1
Example ABACAIJ0	ABACAI	0.3930	2.8119	95.6333	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

25

26



27

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.

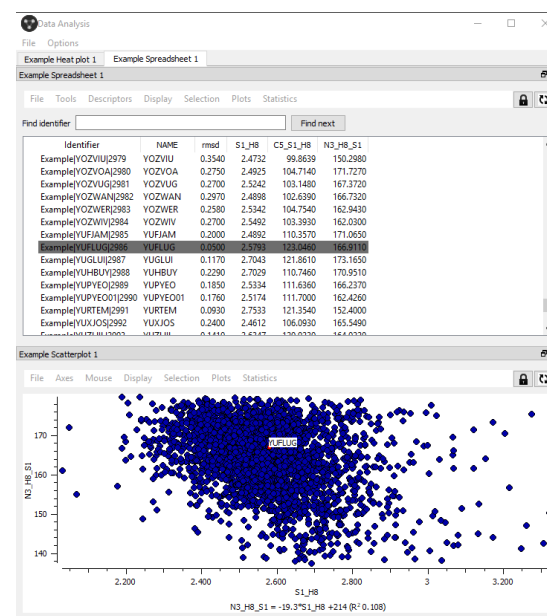
Conclusion

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

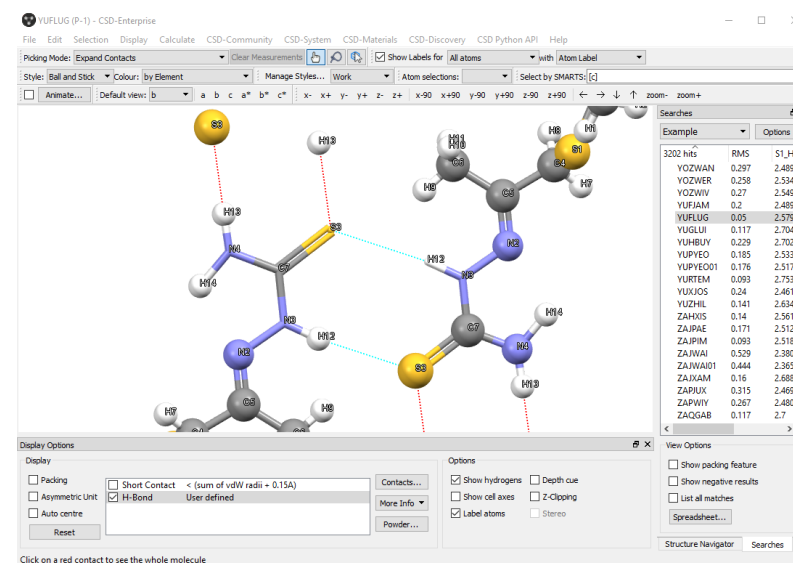
This example shows a high frequency of hits with a particular geometry. This clustering suggests that C=S groups can act as hydrogen-bond acceptors with clear directional preferences. Most observations of C=S as a hydrogen-bond acceptor are in systems where there are one or two nitrogen atoms covalently 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.

29



30



Example 3. Crystal Packing Similarity

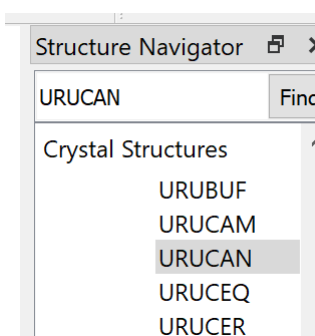
Crystal Packing Similarity functionality in CSD-Materials can be used to examine the packing similarities within a family of polymorphic structures, as well as solvated/anhydrous no solvated/hydrated pairs. This example examines the crystal structure similarity between an anhydrous-hydrated pair. The similarity between the anhydrous-hydrated pair can give further insights about the hydration process and the ease of incorporating or removing water within the crystal structure of a material.

1. Open Mercury by double-clicking the Mercury icon on the desktop.
2. In the Structure Navigator tool bar type the URUCAN refcode.
3. Toggle on the Multiple structure
4. In the Structure Navigator tool bar type URUCER
5. Toggle on the Packing in the Display option dialog box
6. Select the view along the b axis from the top-level menu of Mercury
7. In the bottom right corner, Select Structures...and colour URUCAN as cyan and URUCER as magenta. Note the very similar packing between the two anhydrous and hydrated structures.

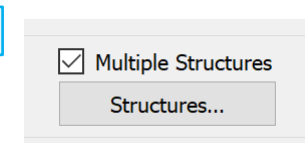
1



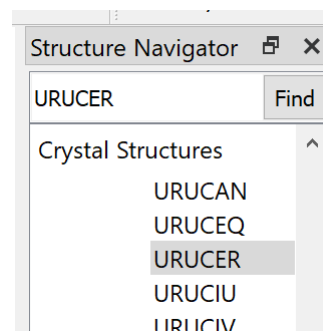
2



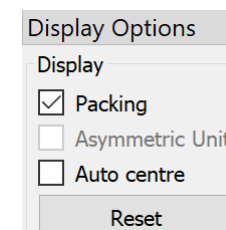
3



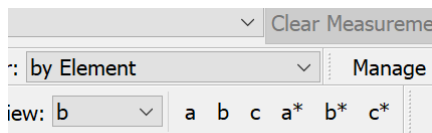
4



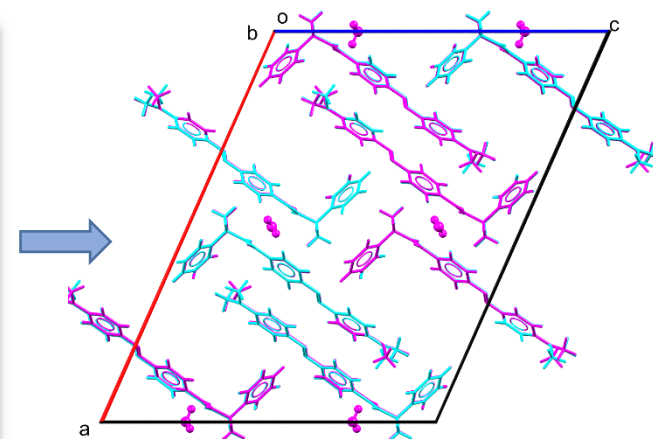
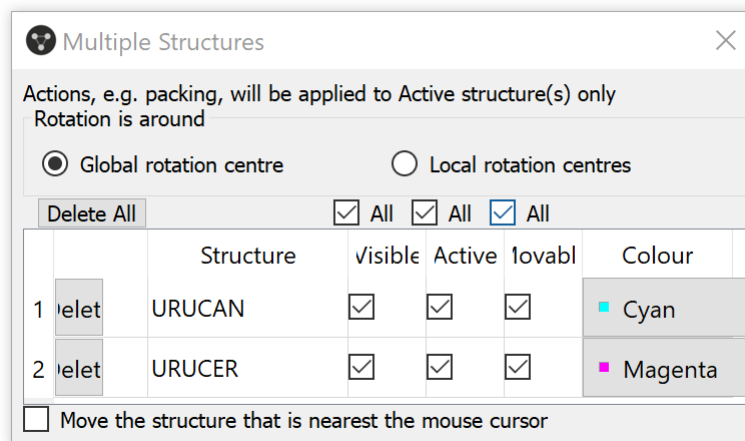
5



6



7



8. Toggle off multiple Structures

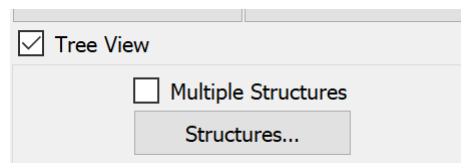
9. In the structure navigate select the URUCAN refile

10. From the top-level menu in Mercury select CSD-materials>Searches>Crystal Packing Similarity

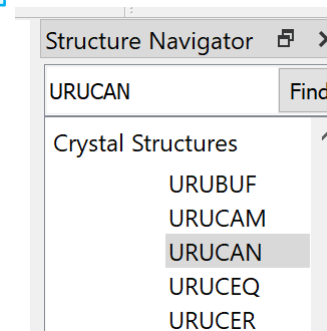
11. Select Add Refcode in the Reference Structure and type URUCAN then OK.

12. Select the Add Refcode in the Comparison Structure and select URUCER then Ok and click Next

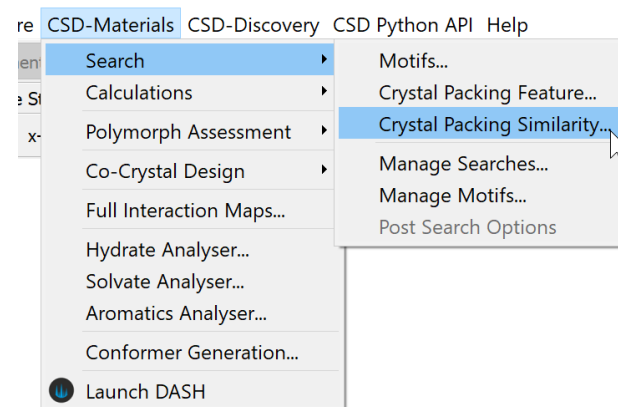
8



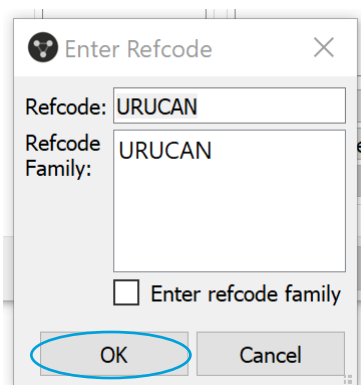
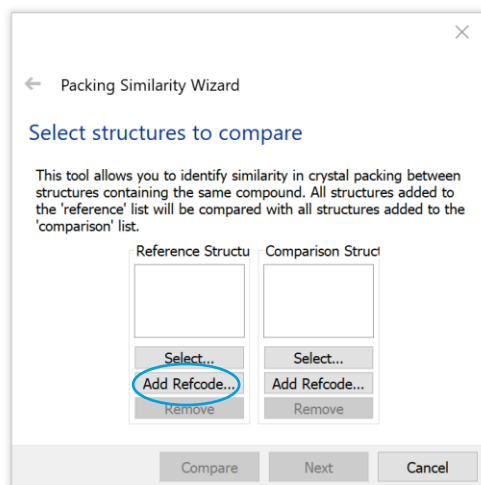
9



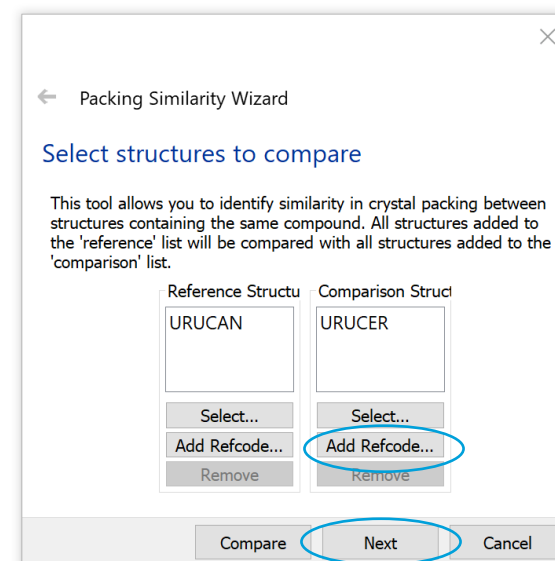
10



11



12



13. Select 30 molecules as Size of molecular cluster to compare and click Next.

14. Leave everything as default and select Next.

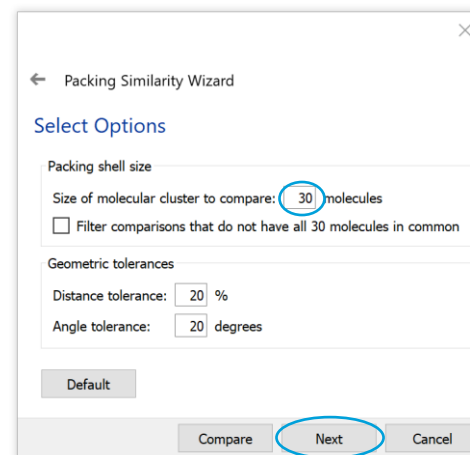
15. Note that the name of the search is given automatically as `crystal_comparison_1`, you can change this if you want. We will leave this as default for this exercise. Select Compare to begin the Crystal Packing Similarity Search

16. Once the search is completed the outcome displays how many molecules are similar, in this case 30 out of 30 and the RMS is 0.096 which is very low in this case. The lower the RMS the more similar the structures are.

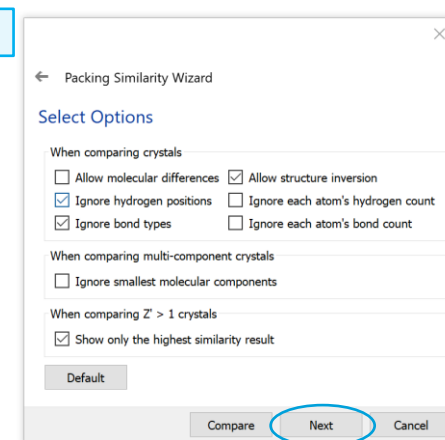
17. The 30 molecules compared are highlighted in the 3D display viewer of Mercury.

18. Toggle on the Show cell axes dialog box in the Display Option section in Mercury to see which molecules in the unit cell were compared.

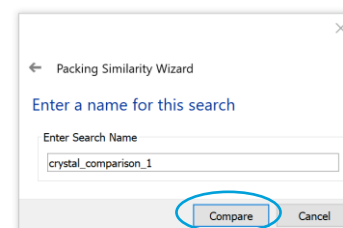
13



14



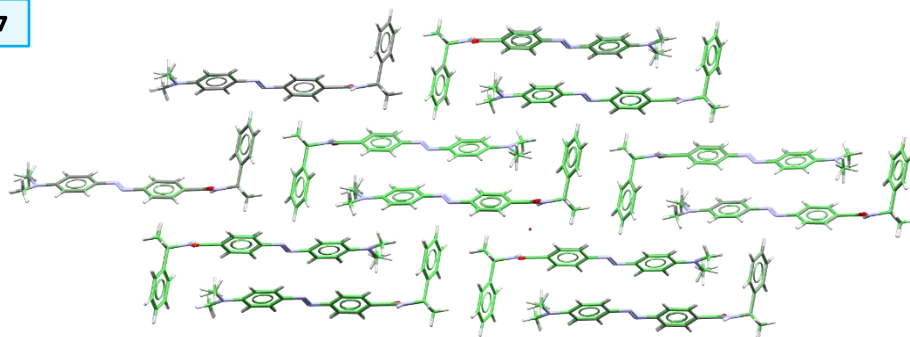
15



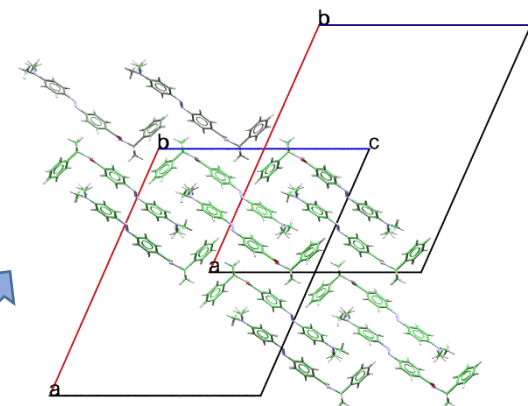
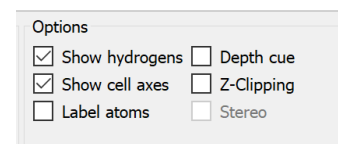
16

Searches				
crystal_comparison_1				
	refere	comp	molecul	RMS
	URU...	URU...	30 out...	0.096
				group1

17



18







19. Repeat step 10 to 15 for the Triamcinolone acetonide acetate pair, IJUTAJ02 and QANVAF refcodes. Note that not all the anhydrated, hydrated pairs are similar.

20. In this case only 1 out of 30 molecules are similar and the RMS is much higher 0.327

Further exercises

Repeat steps 10 to 15 in the Crystal Packing Similarity session for VAWWEX VAWWOH anhydrous-hydrated pair. What RMS do you notice? Is this pair very similar in packing?

20

Searches  				
crystal_comparison_2 				Options 
reference	compari	molecules in common	RMS	reference group
IJUTAJ02	QANVAF	1 out of 30	0.327	group1

Conclusions

- This workshop illustrated Searches that can be performed using CSD-Materials.
- You have learned how to identify the frequencies of occurrence of primary amides in different hydrogen bond network using Motif Search.
- You have learned how to apply Crystal Packing Feature to unusual Hydrogen Interactions such as S...H-N.
- You have learned how to apply Crystal Packing Similarity to anhydrous-hydrated pair.

To learn more about the Search functionality in Mercury, check Tutorials 9 and 10 under Help>Tutorials top level menu in Mercury.

Next steps

If you want to see how Crystal Packing Similarity was applied to identify similar packing of an optoelectronic system with structures in CSD check out this paper <https://pubs.acs.org/doi/abs/10.1021/acs.cgd.1c00177>

To see how many similar anhydrous and hydrated pairs are in CSD check out this publication.

<https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/chem.201404693>

Feedback

We hope this workshop improved your understanding of Searches under CSD-Materials and you found it useful for your work. As we aim to continuously improve our training materials, we would love to hear your feedback. Follow the link on the workshop homepage or click on [this link](#) to a survey, it will take less than 5 minutes to complete. The feedback is anonymous. You will be asked to insert the workshop code, which for this self-guided workshop is MAT-006. Thank you!

