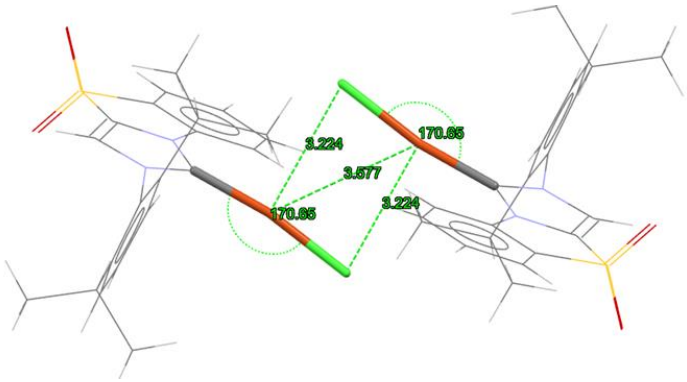
Exploring Structure Correlations Using ConQuest and Mercury (CQ-007)

Developed using
2024.3 CSD Release



CCDC
advancing structural science

Table of Contents

Introduction

ConQuest is the desktop search interface to the Cambridge Structural Database (CSD). It provides an extensive range of flexible search options including substructure searching with the tandem capability of retrieving defined geometrical parameters, subject to structural and chemical constraints. ConQuest can export results directly to Mercury, which is the desktop crystal structure visualization software. Mercury additionally acts as the interface to many other CSD applications, including a comprehensive data analysis package which is dynamically linked to the structure visualizer. We will harness this functionality in this workshop to explore structure correlations in metal complexes.

Learning Outcomes

After completing this workshop, you will be able to:

- Draw a substructure query in ConQuest
- Add defined 2D and 3D parameters to the search query
- Run a combination search and export results to Mercury
- Perform calculations in the Data Analysis Module
- Select, plot and filter data according to specified criteria

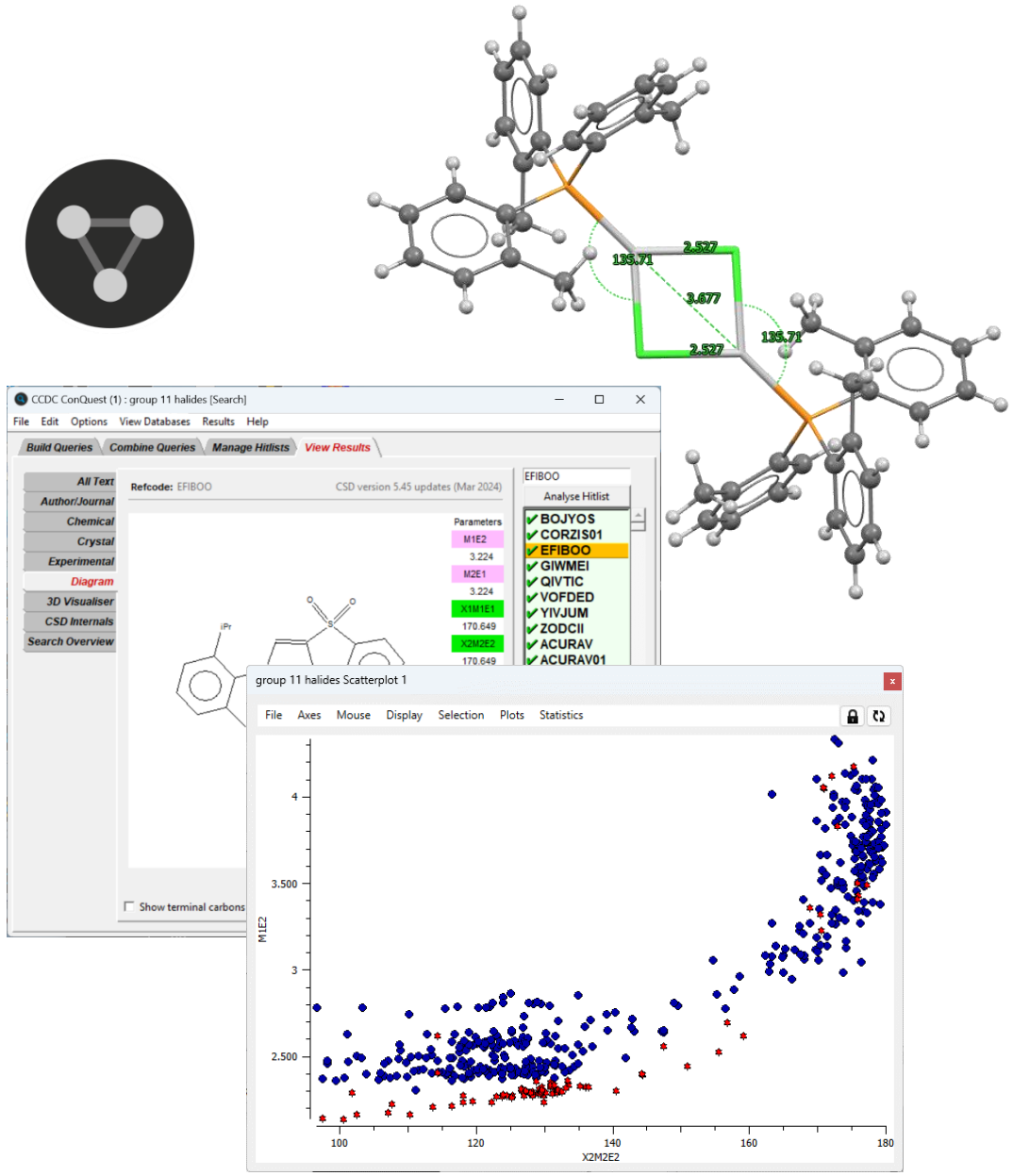
This workshop will take approximately ***40 minutes*** to be completed. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout.

Pre-required Skills

Basic familiarity with the ConQuest and Mercury interfaces is desirable, however, summary guides are provided at the end of this document.

Materials

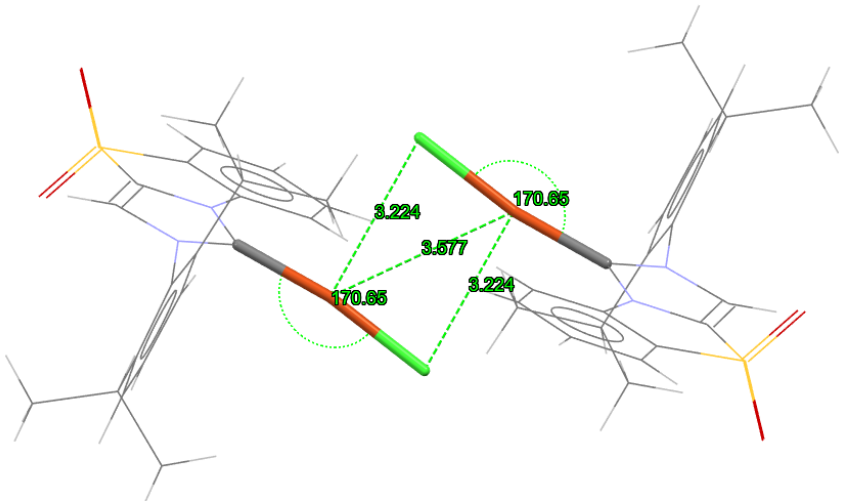
No additional materials are required.



Example 1. The Dimerization of Group 11 L–M–X Halo complexes

Interpreting the correlation between two structural parameters within a family of related compounds as providing snapshots of chemical reaction pathways is one of the fields of study that has benefitted most from the wealth of data contained in the CSD. One such system, investigated by Echeverría and Alvarez[1], is the dimerization of Group 11 and Group 12 metal halides. The abundance of structures bearing a dimeric L–M–X motif provides hundreds of measurements of bond distances (as well as formally non-bonded contact distances) and valence angles which can be used as the starting point for structure correlation studies.

In this example, we will search the CSD for $(MX)_2$ fragments formed by the association of L–M–X units, where L is any ligand, M is one of Cu, Ag or Au, and X is Cl, Br or I. In so doing, we will collect the M–X···M distance and the L–M–X valence angles, and search for correlations between them.



Part 1. Building search queries in ConQuest

We will build two queries in such a way that we can deploy them simultaneously in a structure search. The use of consistent labels means the data retrieved can be combined in the second half of this tutorial.

Important: it is vital that you keep track of the labels you use for atoms and parameters in the following steps. We strongly recommend that you use the labelling scheme shown in Figure 1.

Tip: When defining parameters, it is advisable to use more memorable symbols in the labels; we suggest M, E and X for metal, halogen and ligand, respectively. The numbers 1 and 2 in Figures 1a and 1b refer to the fragment of the dimer.

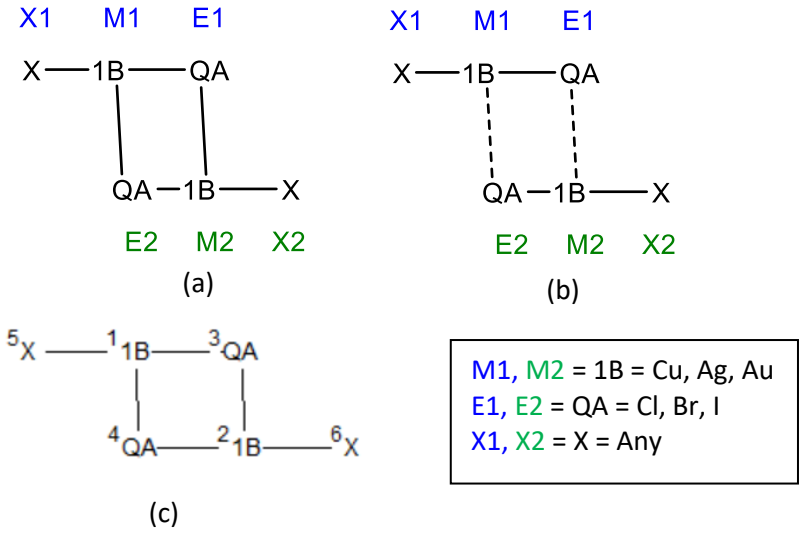


Figure 1. Suggested atom labelling schemes for (a) bonded and (b) non-bonded dimers. The labels 1B, QA and X are used in ConQuest to refer to *atom type*. The superscript numbers visible in (c) when defining 3D parameters refer to the order in which the atom was added to the draw window. **If you edit the structure, these numbers will change.**

[1] J. Echeverría and S. Alvarez, *Cryst. Growth Des.*, 2024, **24**, 4743–4747. https://doi.org/10.1021/acs.cgd.4c00335

1. Open ConQuest from the Start Menu or by clicking on the desktop icon 

2. Click on the **Draw** button from the **Build Queries** tab to bring up the Draw Window.

3. From the bottom toolbar, click on **More…**and then **Other Elements** to bring up the periodic table.

4. First, we will select the Group 11 metals. Click on **1B** at the top of the group and then click **Apply**.

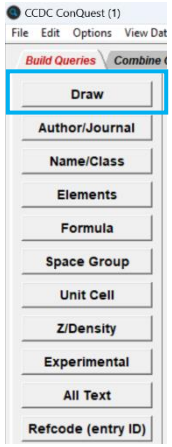
5. Click in the **Draw** window twice to add “1B” atoms, corresponding to M1 and M2 in Figure 1a.

6. Reopen the periodic table as in **Step 3** (click **Reset** if 1B is still selected). Select **Multi Pick** from the *picking mode* section, then click on **Cl**, **Br** and **I** and click **Apply**.

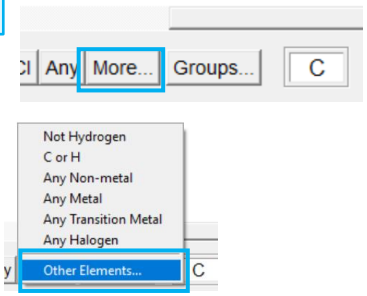
7. Ensure that *Single* is selected from the *Bond:* drop down menu. Click, drag and release on the 1B atoms sequentially in the **Draw** window to add “QA” atoms, corresponding to E1 and E2 in Figure 1a.

8. From the bottom toolbar, click **Any** and click, drag and release to add “X” atoms, corresponding to X1 and X2 in Figure 1a.

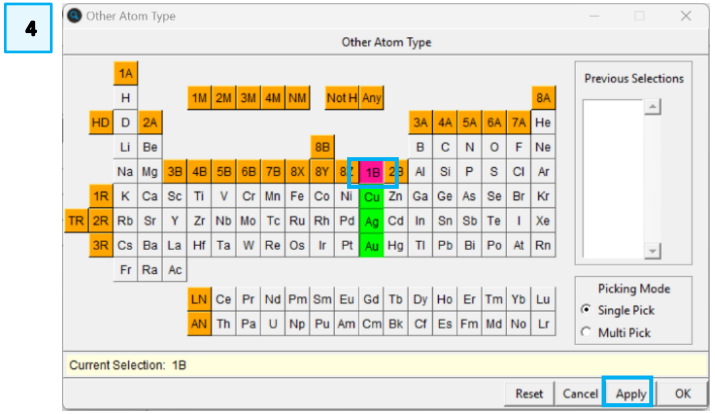
2



3



4



5

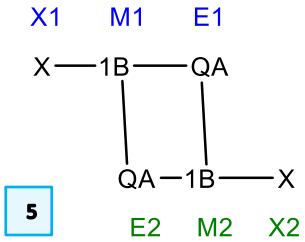
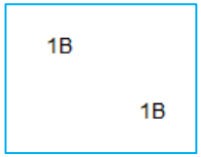
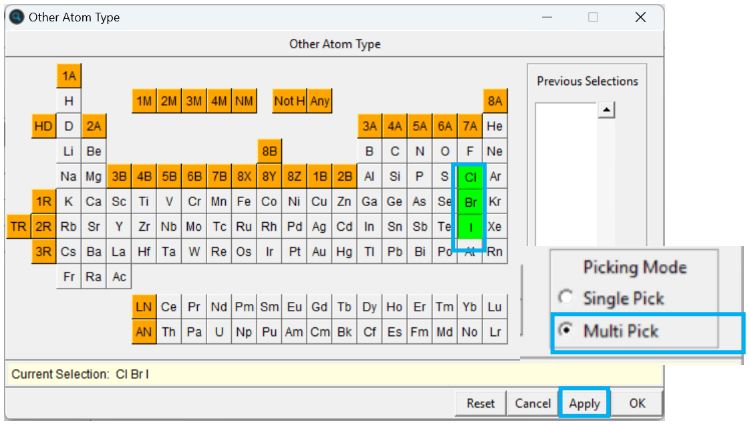


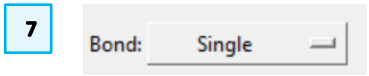
Figure 1a.

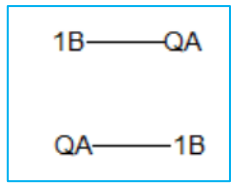


6

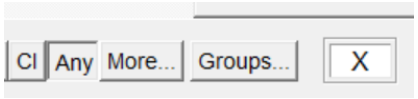


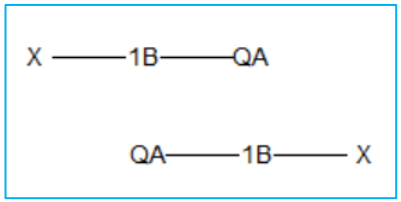
7





8





9. Draw single bonds between the QA and 1B atoms

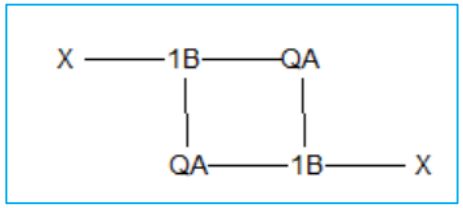
We will next define 3D parameters. All parameters except *Atom Labels* can be renamed, so we recommend that these are defined in the order suggested.

10. From the left-hand toolbar, click **Add 3D** to bring up the **Geometric Parameters** window.

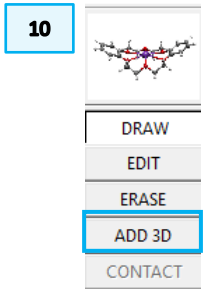
11. In the **Draw** window, click on the atoms corresponding M1 and E2 in Figure 1a (in the image to the right, these are labelled as ^{1}1B and ^{4}QA, respectively). Next to *Distance* in the define parameters window, click **Define**. Tip: the labels ^{n}atom (like ^{1}1B) depend on the order they were drawn in; this is why we recommend relabelling according to Figure 1a.

12. In the Geometric Parameters window, click **Options…**and in the *RENAME:* field, type “M1E2” and click OK.

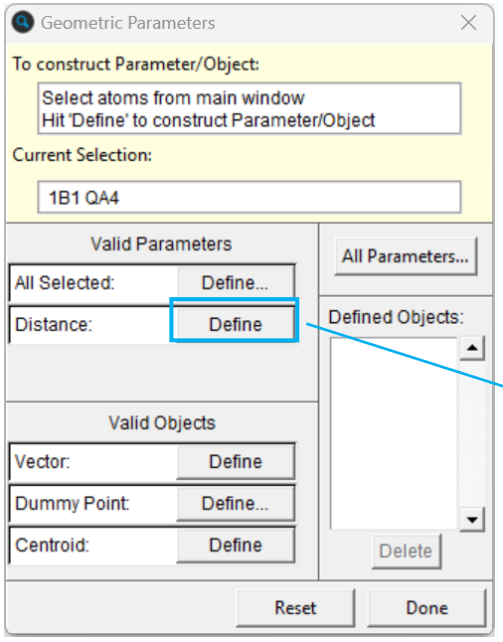
9



10



11



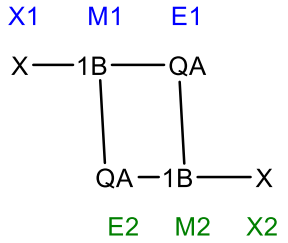
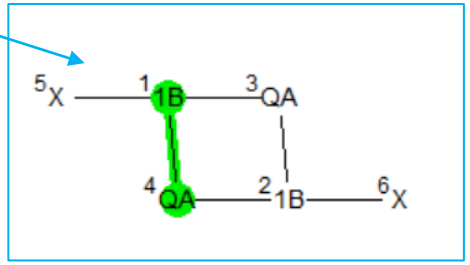
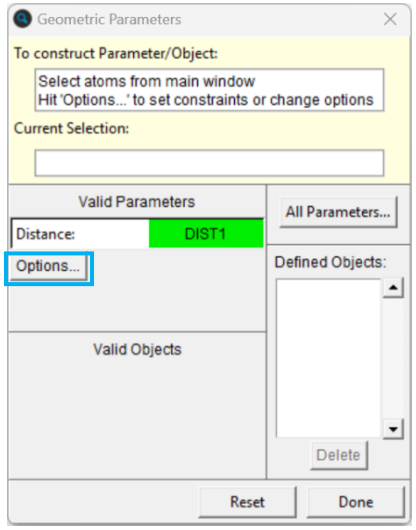
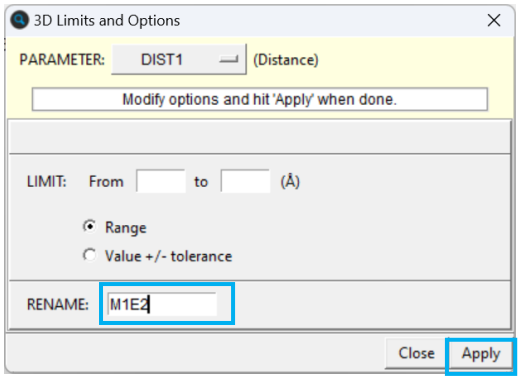


Figure 1a.



12





13. Repeat **Steps 11–12** for the atoms corresponding to M2 and E1 in Figure 1a (these are labelled ^{2}1B and ^{3}QA in the image to the right). Note you may need to deselect the previous atoms first. Edit the name to “M2E1”.

14. Click on the atoms corresponding to X1, M1 and E1 in Figure 1a, in that order. In the image to the right, these are labelled ^{5}X, ^{1}1B and ^{3}QA. Next to *Angle:* click **Define**.

15. Use the procedure in **Step 12** to rename ANG1 to “X1M1E1”.

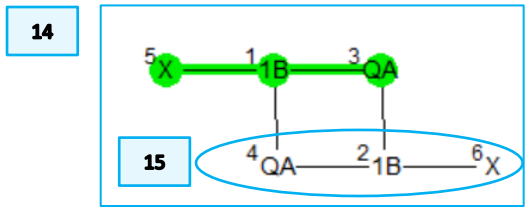
16. Repeat **Steps 13–15** for the atoms corresponding to X2, M2 and E2 in Figure 1a (in the diagram to the right, these are labelled ^{6}X, ^{2}1B and ^{4}QA, respectively). Rename this angle “X2M2E2”.

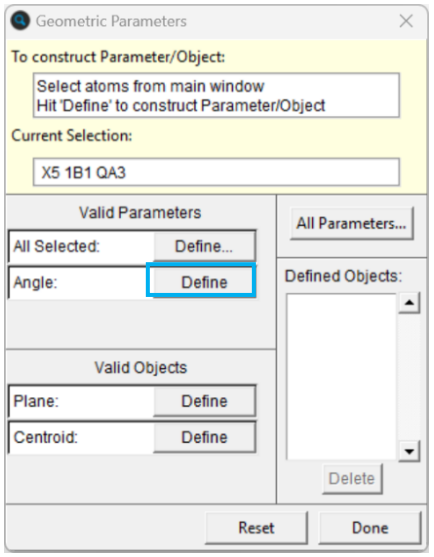
17. Define the distance between the two metal atoms (M1 and M2 in Figure 1a, and ^{1}1B and ^{2}1B in the figure to the right, respectively). Name this parameter “M1M2”.

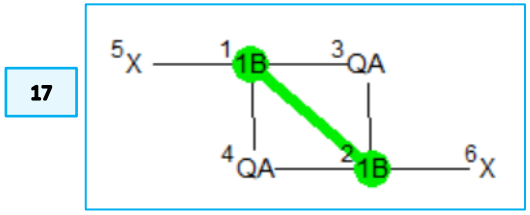
18. Select only the atom corresponding to M1 in Figure 1a (this is ^{1}1B in the image to the right, highlighted green). In the Geometric Parameters window, click **Define** next to *Atom Properties:* . In the All Parameters window that appears, tick **Atom label** and **OK**.

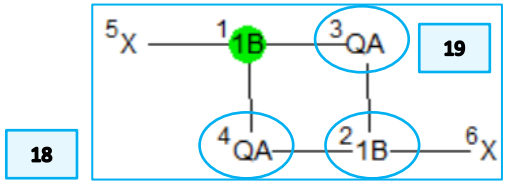
19. Repeat the procedure for the remaining atoms in this order (as in Figure 1a): E1, M2, E2 (these are labelled ^{3}QA, ^{2}1B and ^{4}QA in the image on the right). **Do not attempt to change the default labels, as this will cause an error.** When you are finished, click **Done** in the **Geometric Parameters** window.

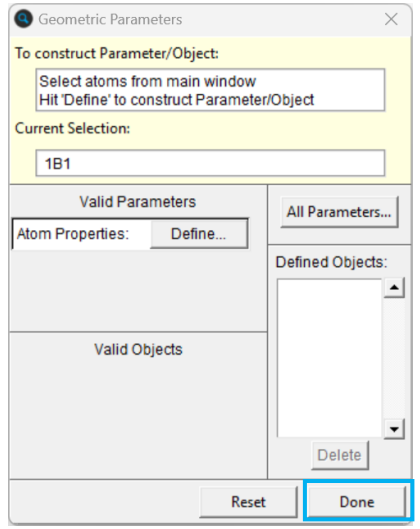
20. In the 3D parameters box at the top right-hand side of the Draw window, you should see a list of all the parameters you have defined. There should be nine defined parameters; you may need to scroll to see them all.

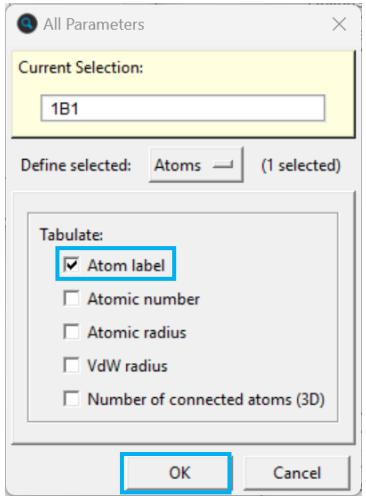


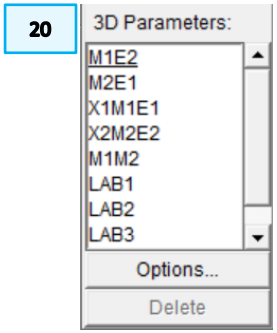












21. Right click on the 1B atoms (which correspond to M1 and M2 in Figure 1a) and from the drop-down menu, select *Number of Bonded Atoms > 3*. The superscript "T3" will appear next to these atoms.

22. Click **Store** at the bottom right-hand side of the **Draw** window.

We will now create a second query, as shown in Figure 1b, which involves the same fragments but with *non-bonded contacts* between them.

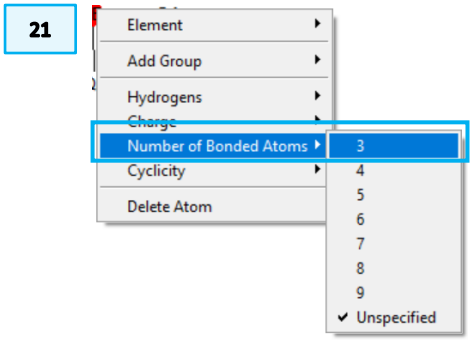
23. In the **Build Queries** area, next to *Query 1*, click **Edit…**. This brings up another instance of the query that we constructed earlier.

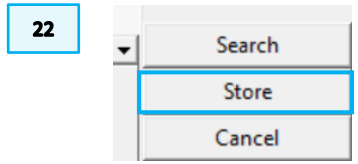
24. Click on **Erase** in the left-hand toolbar, then click on the bonds between the 1B and QA atoms.

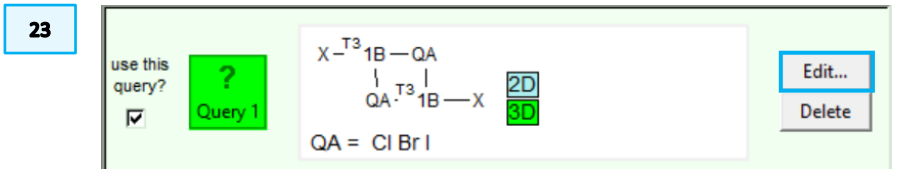
25. Click on Draw, then right click on the 1B atoms and select *Number of Bonded Atoms > 2*.

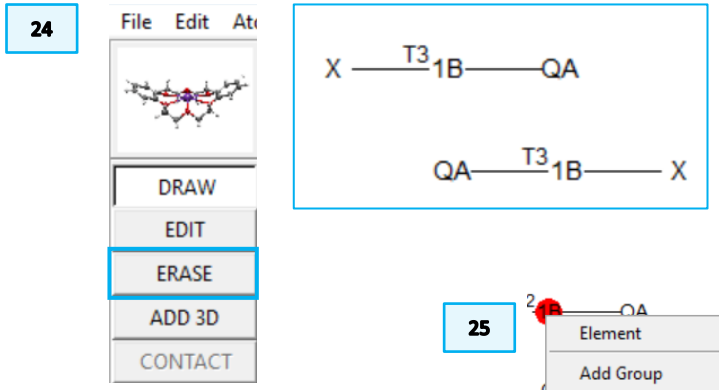
26. From the 3D Parameters section on the top right-hand side of the **Draw** window, click on M1E2 and click **Options**.

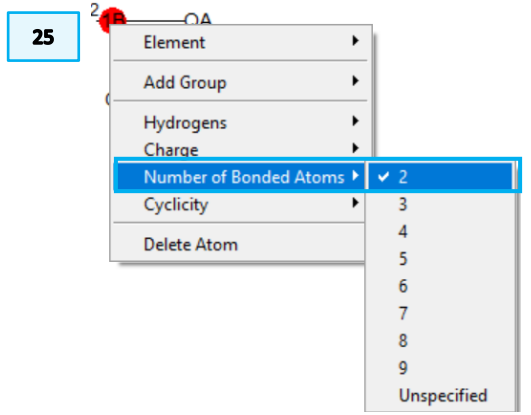
27. Next to *CONTACT:* click **Create**

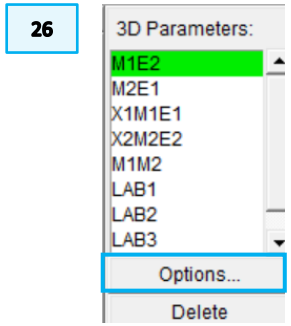


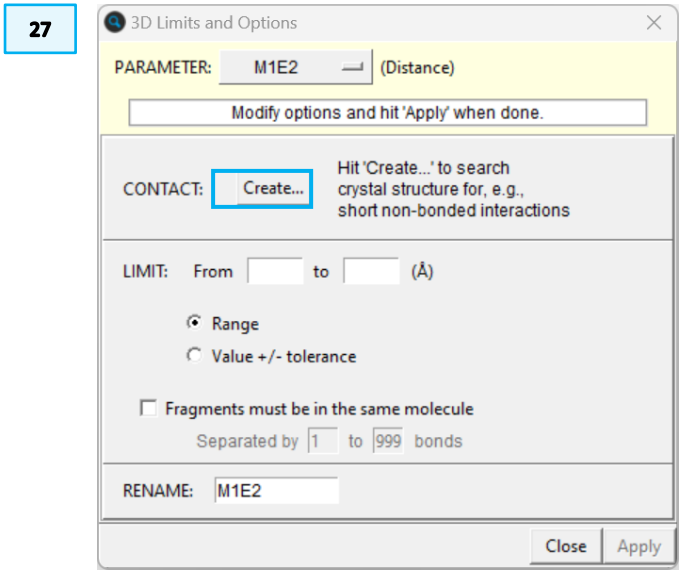












28. Click **Define** to bring up the Non-bonded Contact definition.

29. Another window with more options to customise the contact settings will appear. Click **OK**. Note: the van der Waals radii of the elements have been updated in the 2024.3 CSD Software release. If you are on an earlier software version, you will need to update to 2024.3, or manually update the van der Waals radii by clicking **Edit** and changing the relevant values (2.38; Ag, 2.53; Au, 2.32; Cl, 1.82; Br, 1.86; I, 2.04).

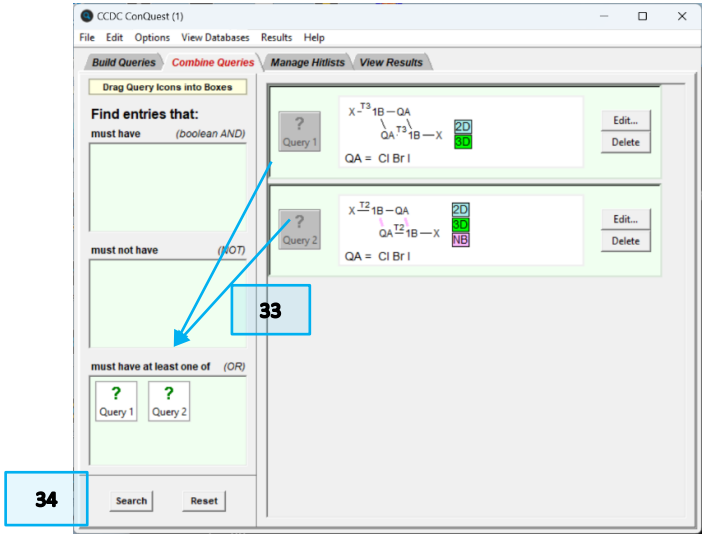
30. Click on M2E1 in the *3D Parameters* box and repeat **Steps 26–29** for the parameter M2E1.

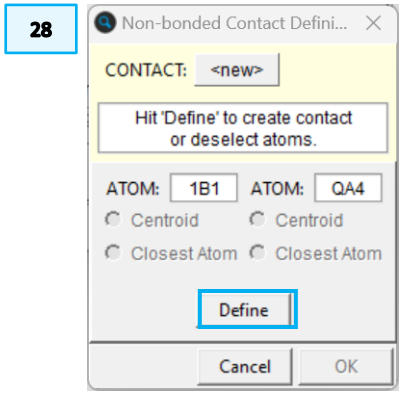
31. Click **Store** from the bottom right-hand side toolbar. When prompted if you wish to overwrite the query, press **No**. This will ensure that a separate query is created.

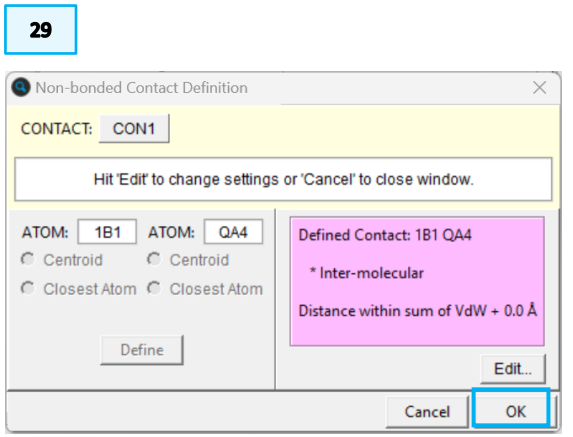
32. At this point, it is recommended that you save these queries, by selecting *File > Save Queries> All Queries…* and save it in a convenient location. These can be re-opened in another ConQuest session if needed.

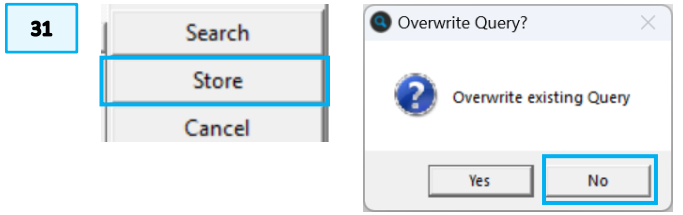
33. Click on the **Combine Queries** tab and drag both Query 1 and Query 2 to the box *must have at least one of (OR)*

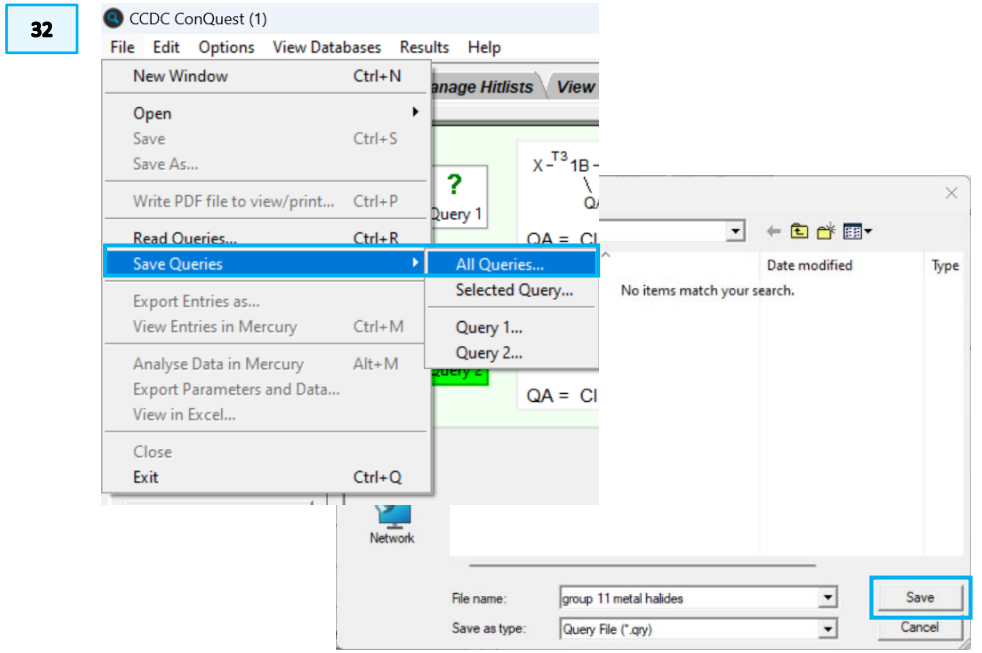
34. Click **Search** to launch the **Search Setup** window.









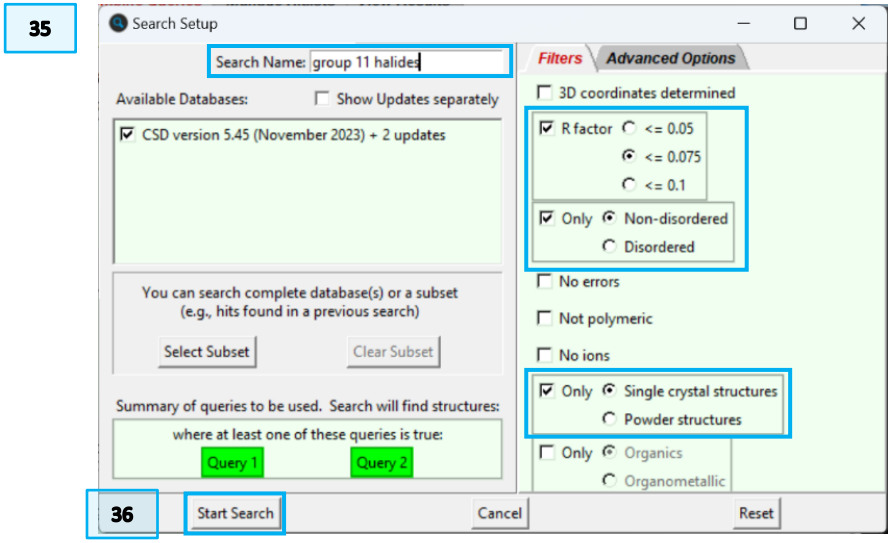


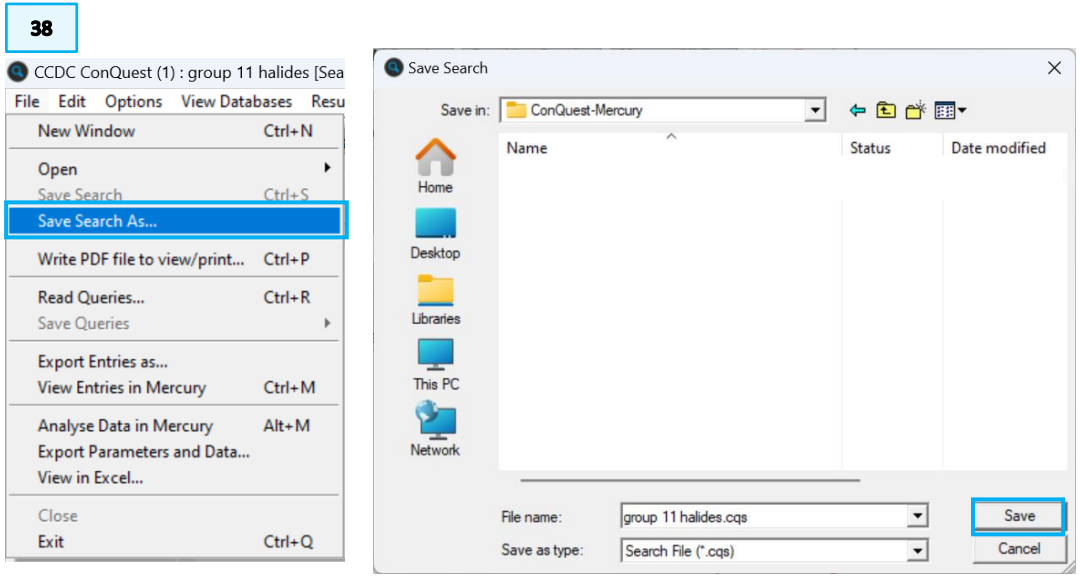
35. Choose a suitable search name, such as “group 11 halides” and tick *R factor* **<=0.075**, *Only* **Non-disordered** and *Only* **Single crystal structures**. Leave all other settings as default.

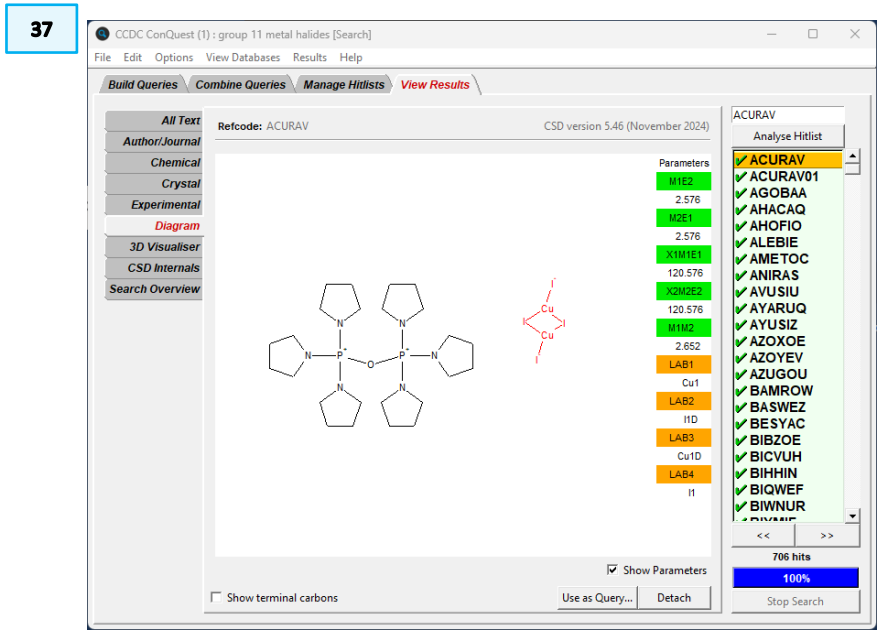
36. Click **Start Search**.

37. The search should return a *hitlist* of at least 706 *refcodes* based on CSD 5.46 (2024.3 release); your value may differ depending on your version. You can scroll through the hits if you wish; you will notice that where a non-bonded dimer has been identified, the 3D parameter is highlighted pink, whereas it is highted green for the directly bonded dimers.

38. You can save the search results by going to *File > Save Search As…*







Conclusion

In this section, we have seen how to set up a substructure query in ConQuest to measure distances and angles in metal halide dimers where there are formal bonds between the constituent monomers. We have also seen how to adjust the query to capture cases where these formal bonds are replaced by intermolecular non-bonded contacts. Lastly, we combined both queries in a ConQuest search to retrieve all relevant structures in a single set of results.

We will now make use of the data gathered in Part 1 to uncover structure correlations using Mercury's Data Analysis Module in Part 2. Our aims here are to:

- Calculate new structural parameters with the Calculator tool to filter out irrelevant structures
- Plot the data to identify distance-angle relationships
- Apply filters to isolate subsets of the data relating to specific metal/halide combinations

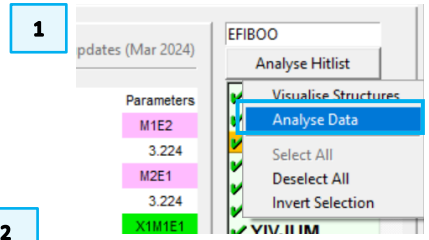
Part 2. Data analysis in Mercury

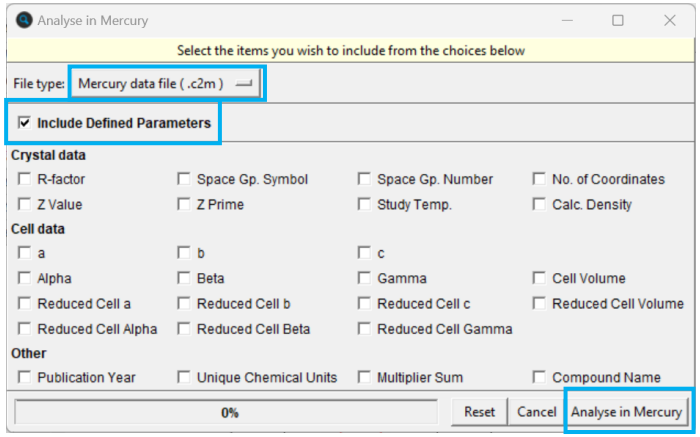
In this section we will analyse ConQuest search data in Mercury.

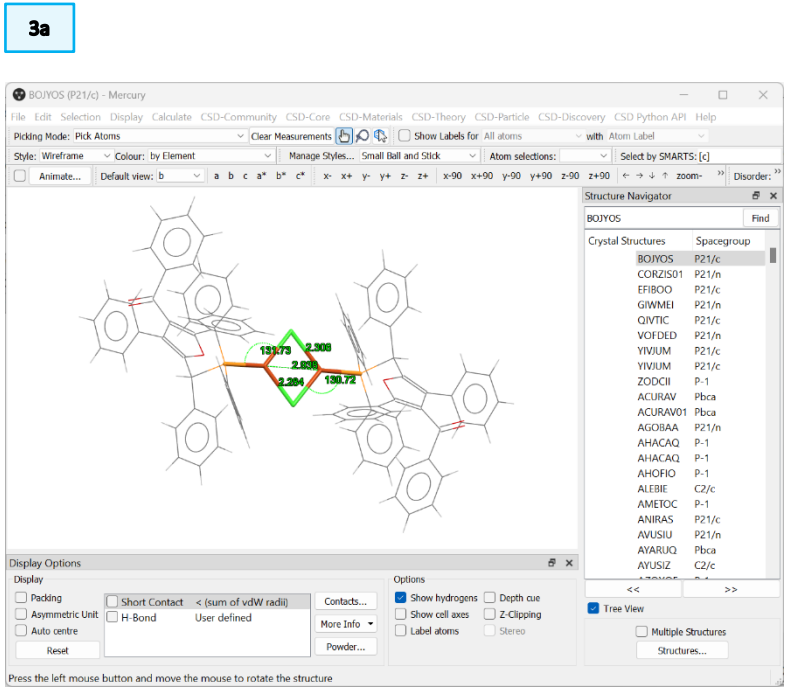
1. At the top of the results list in the **View Results** tab, click *Analyse Hitlist > Analyse Data*.

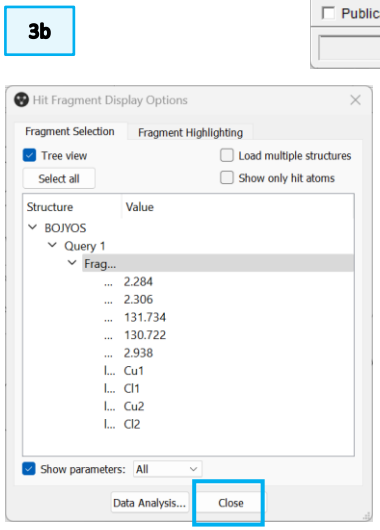
2. The Analyse in Mercury window will appear. There are many details that you can export, but all we require is **Include Defined Parameters**. Ensure this is selected and that the *File type* is set to *Mercury data file (.c2m),* and press analyse in Mercury.

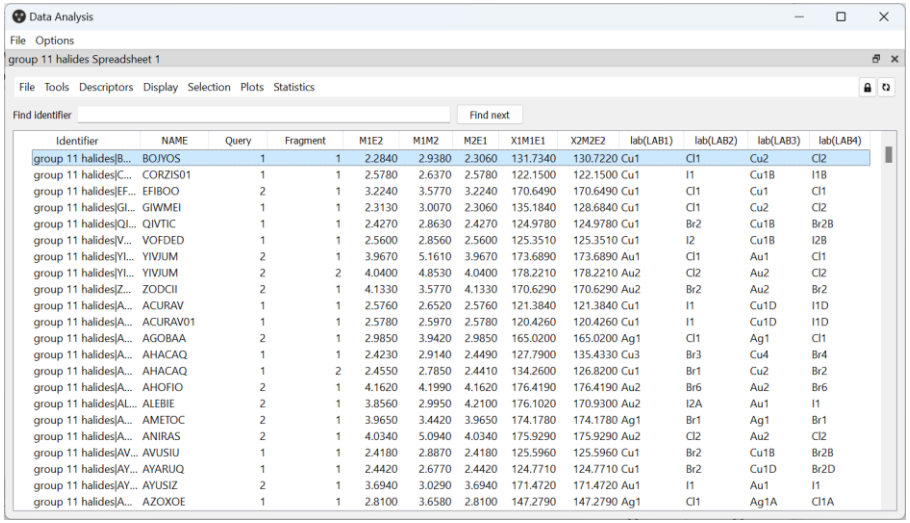
3. Three windows will appear: (a) the **Mercury visualizer** window, with the search results displayed in the **Structure Navigator** (b) the **Hit Fragment Display Options** and (c) **Data Analysis**. You can experiment with the Hit Fragment Display options if you wish, but we do not require it, so it may be closed. The remainder of the workshop will focus on the **Data Analysis** window.











4. The headers of the columns in the spreadsheet correspond to the parameters that we defined in our search queries. You may need to expand the columns to read the headings properly (click and drag on the column dividers).

Dimers of the type we have searched for are known to form M···M interactions in some circumstances. The presence of such bonding may invalidate the bonding model we are using in describing M···Hal dimers. To filter these out, we can disregard examples where the M1M2 distance is less than 0.1 Å greater than the M1E2 and M2E1 distances.

5. Click *Tools* > *Calculator* from the top menus of the spreadsheet.

6. From the available heading in the *Name* section, click *M1M2*. It will appear in the lower box as 'group 11 halides'.'M1M2'. In the box, type "–" (or click the minus button from the right hand) and then click on M1E2. In the *New descriptor* box give a suitable name, such as "M1M2 – M1E2" and click **Calculate**.

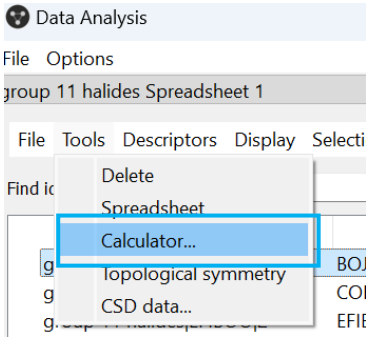
7. The name of this calculation will appear in the left-hand box of the calculator and also in the spreadsheet as a new column.

8. Click **Clear** and repeat the procedure to calculate M1M2-M2E1. **Close** the calculator.

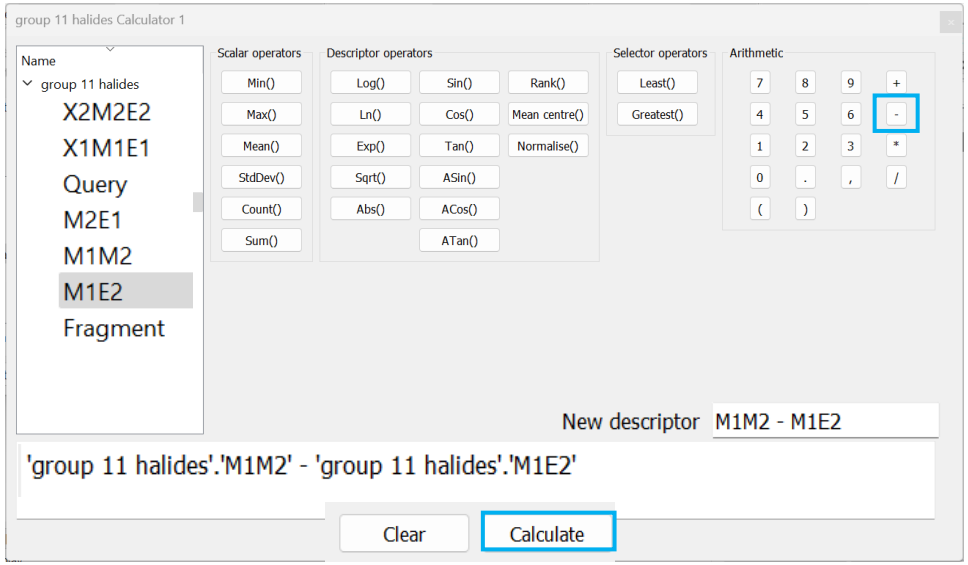
4

Identifier	NAME	Query	Fragment	M1E2	M1M2	M2E1	X1M1E1	X2M2E2	...(LAB1)	...(LAB2)	...(LAB3)	...(LAB4)
group 11 halides|BOJYOS|0	BOJYOS	1	1	2.2810	2.9380	2.2970	127.0820	129.2040	Cu1	Cl2	Cu2	Cl1

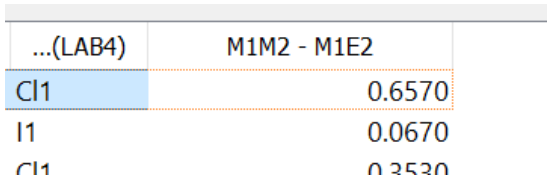
5



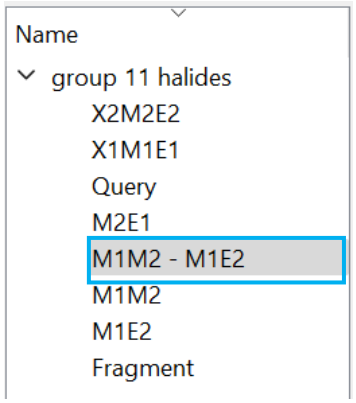
6



7



...(LAB4)	M1M2 - M1E2
Cl1	0.6570
I1	0.0670
Cl1	0.3530



9. Right-click on the M1M2-M1E2 column header and select Sort from the dropdown menu. This will sort the column by numerical value. You may need to click twice to get the value in descending order.

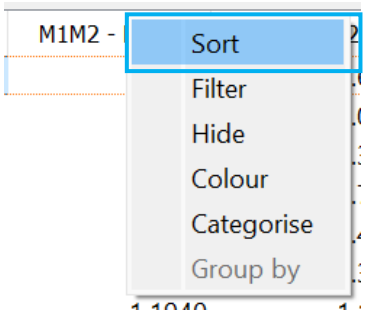
10. Scroll down the spreadsheet until you find the row at which M1M2-M1E2 is less than 0.1 Å. Left-click on the row and drag to select all remaining rows below 0.1 Å.

11. From the top menu, click *Selection > Hide selected*.

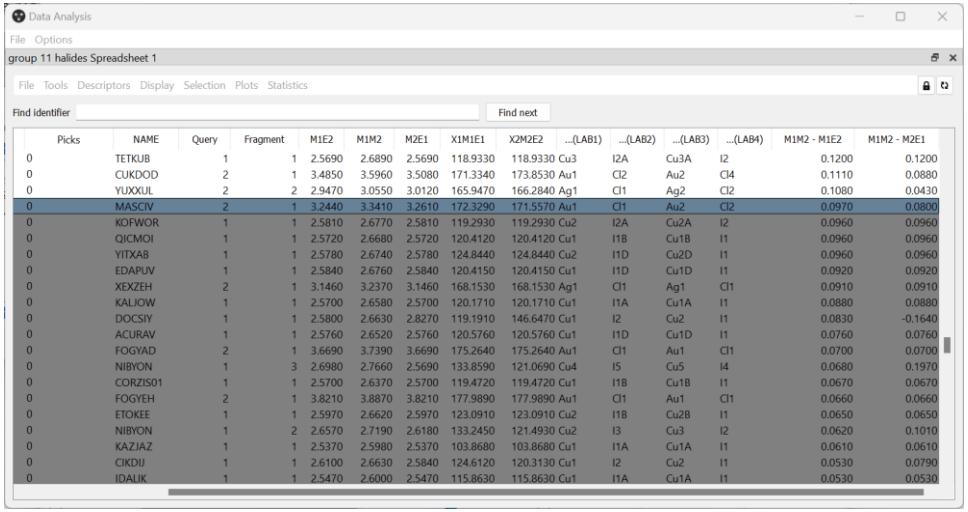
12. Repeat steps 9–11 for the column M1M2-M2E1 to remove the small number of additional entries for which this value is less than 0.1 Å.

13. Make sure that no columns are selected by clicking the header of any columns that are currently highlighted. Click Ctrl + A to highlight the selected entries. From the top menu, select *Plots > Scatterplot*. In the resulting **Choose descriptors** menu, from the drop-down menus, select *X2M2E2* for the *X axis* and *M1E2* for the *Y axis*. Leave the **Colour** and **Size** boxes unchecked and click OK.

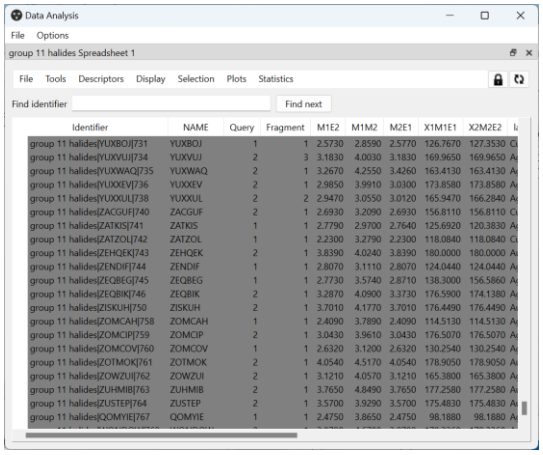
9

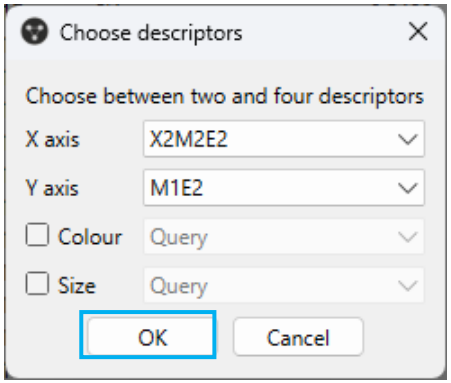


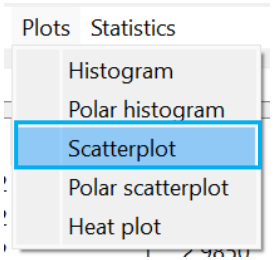
10



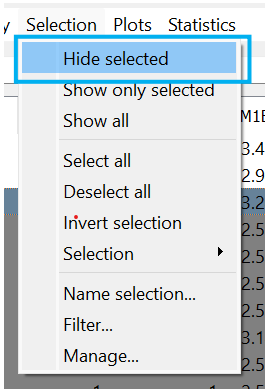
13







11

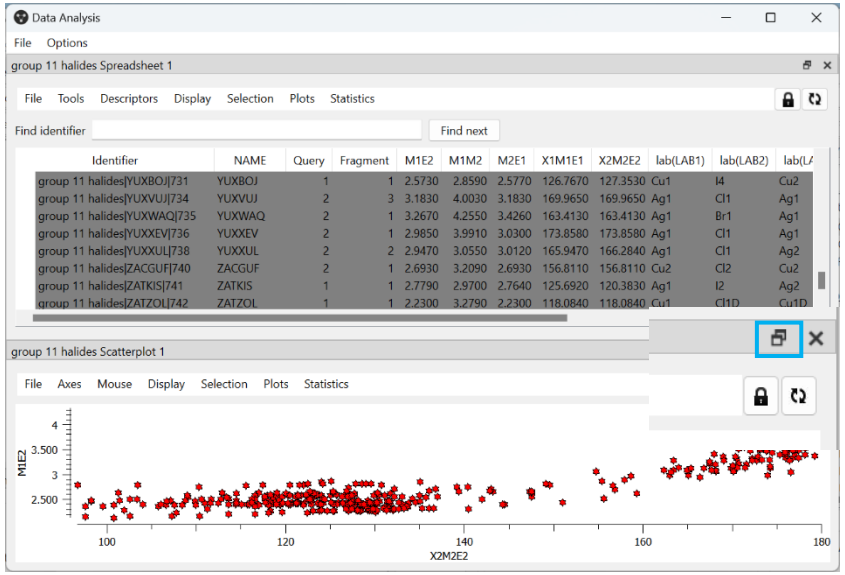


14. A plot will now appear, probably below the spreadsheet. You can detach this by clicking the two windows symbol at the top of the plot window to position it somewhere more convenient.

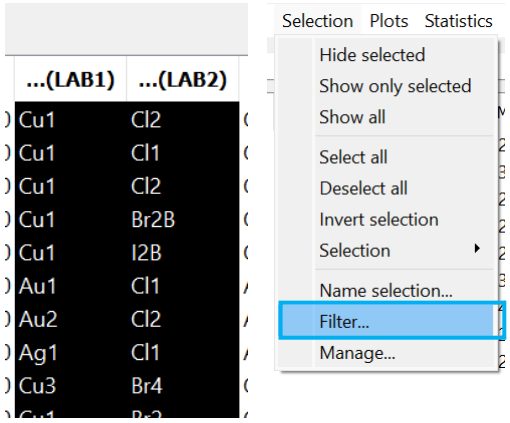
15. The graph contains entries for different metal-halide pairs e.g. Cu–Cl and Au–I. We can filter based on the atom labels defined in the search to highlight particular combinations. Left-click on the LAB1 and LAB2 headings; they will become shaded black. Then from the *Selection* menu at the top of the spreadsheet, choose *Filter*.

16. In the filter window (you can detach it) you will see filters for LAB1 and LAB2. From the dropdown menu, select *Starts with* for both labels and type Cu for LAB1 and Cl for LAB2. This will filter the results, so we see the Cu–Cl dimers highlighted red in the plot window.

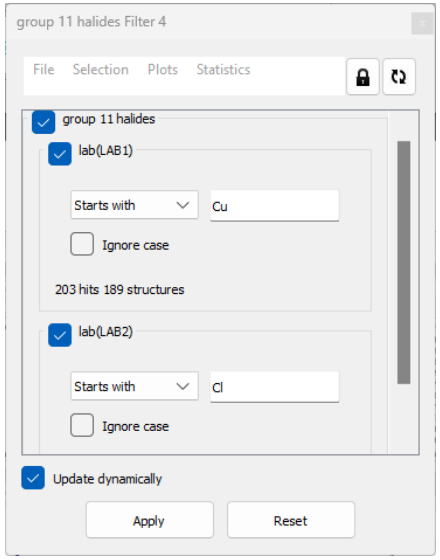
14

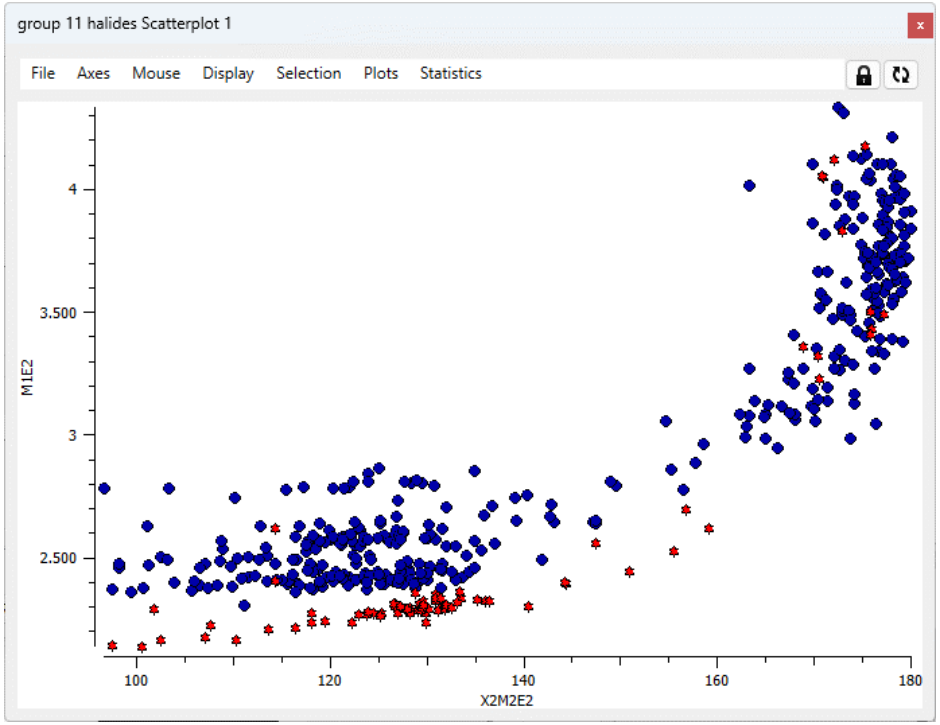


15



16



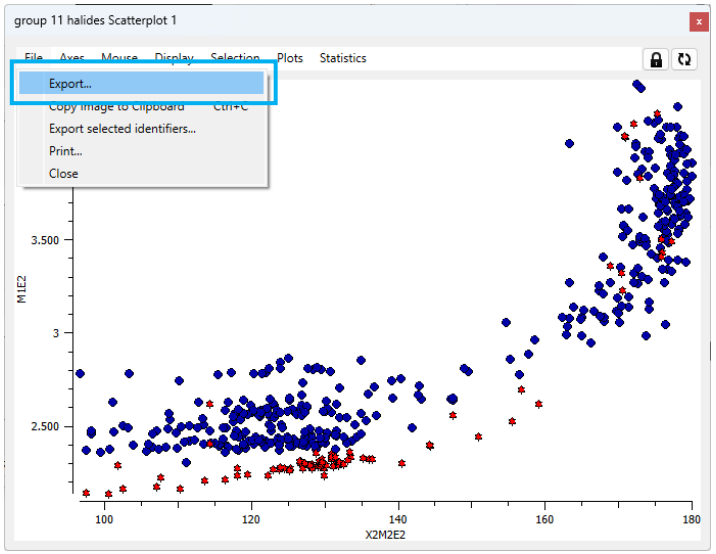


17. Explore other possible metal-halide combinations. You can save the plots at any point from the plot window by going to *File* > *Export* and selecting as appropriate.

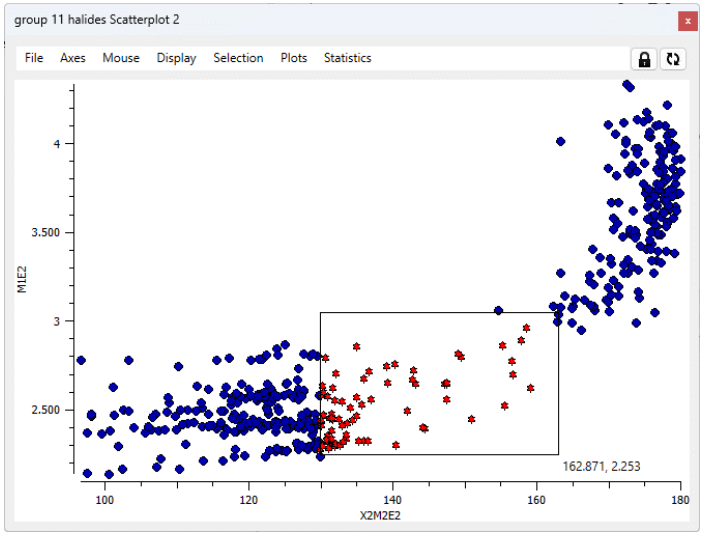
18. Untick the filters and close the filter window. You can reverse the filter process and select entries from the graph by clicking individually or dragging a box. The highlighted entries will turn red in the scatter plot and will be shaded grey in the spreadsheet.

19. Clicking on a single row will update the structure in the visualizer.

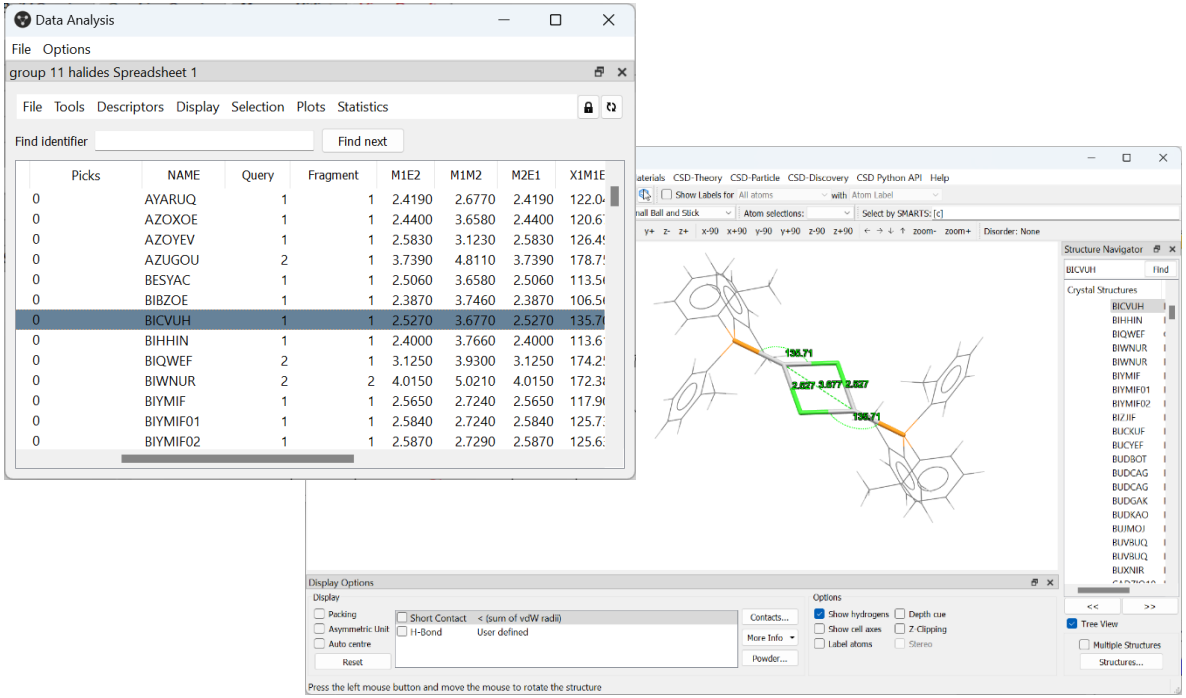
17



18



19

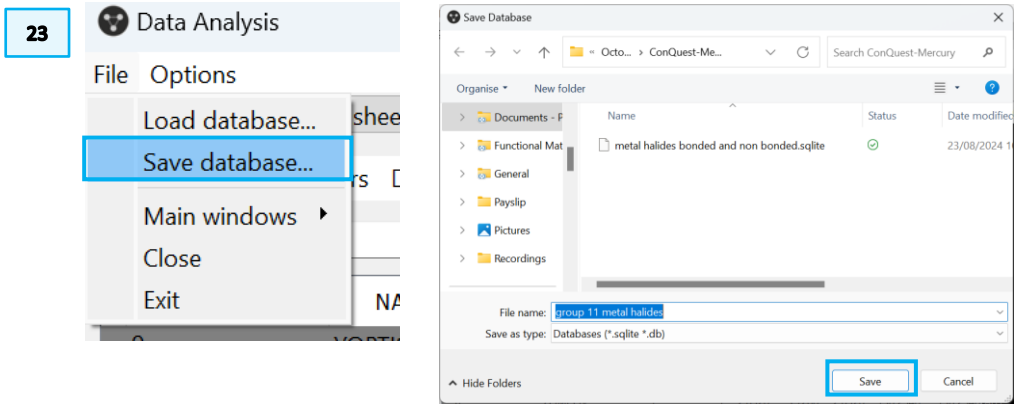


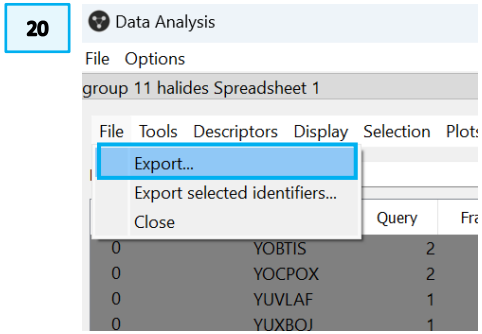
20. In the main spreadsheet, press Ctrl + A to select the entries then click *File > Export*. This will allow you to save out the data to a CSV file for further processing by other programs.

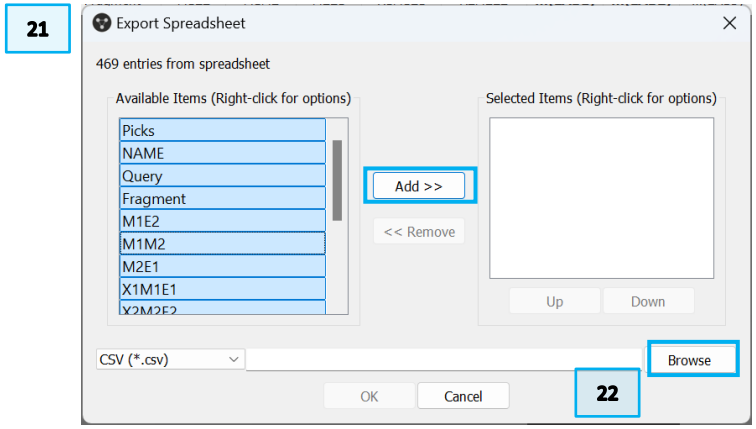
21. In the Export Spreadsheet window, you can select the columns you want to export by selecting from the *Available Items* area and clicking **Add >>** to move it to the *Selected Items*. We will select all with Ctrl+A and transfer them to the *Selected Items* using **Add >>**.

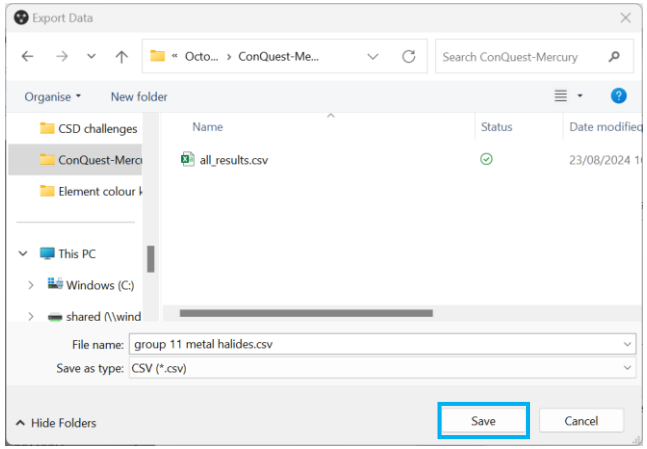
22. Select a suitable location using the **Browse** button and click **Save**.

23. To save the results in a format that can re-opened by Mercury, from the Data Analysis top menu (*not* the spreadsheet top menu). Click *File > Save database…*. Choose a suitable location and save the database. This can be reopened from *File > Load database* in the future.









Conclusion

In this example, we have seen how 3D geometric data from a large number of crystal structures can be used to elucidate structure correlations. We have verified that amongst group 11 metal halides, there is a correlation between the inter-dimer metal-halide bond length and the bond angle in the monomers. Furthermore, we have used the capabilities of the Mercury Data Analysis module to compare trends amongst different metal-halogen pairs and have established that there are indeed substantial differences.

Further exercises

If you have enjoyed this exercise, you can explore the results further:

- Plot M2E1 *vs* X1M1E1 using the Data Analysis module to check the same correlations are found.
- In the exported spreadsheet, combine the distance and angle data for all the fragments and plot graphs of the complete dataset.

Echeverría and Alvarez have proposed analysis of these data in terms of the *penetration index* which takes into account both van der Waals and covalent radii to give a more meaningful comparison of distances. The penetration index of two atoms, A and B, is defined as:

$$p_{AB} = 100 \times \frac{v_A + v_B - d_{AB}}{v_A + v_B - r_A - r_B}$$

where v_A and v_B are the van der Waals radii of A and B, r_A and r_B are the covalent radii of A and B, and d_{AB} is the interatomic distance.

- Calculate the penetration indices for the data, using the values in Table 1, and plot these against the L-M-X angles (see Figure 2).

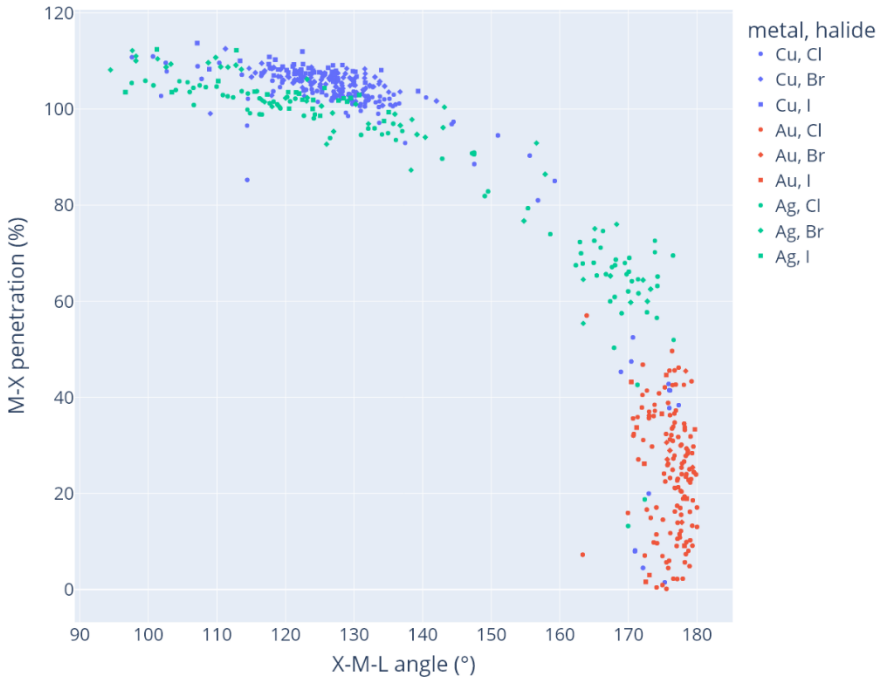


Figure 2. Plot of X-M-L angle vs M…X percentage penetration.

Atom	van der Waals radius	Covalent radius
Cu	2.38	1.32
Ag	2.53	1.45
Au	2.32	1.36
Cl	1.82	1.02
Br	1.86	1.20
I	2.04	1.39

Table 1. Selected van der Waals and covalent radii.

Summary

In this workshop we have seen how to search for substructures in the CSD and retrieve 3D data from the hits. We subsequently saw how to analyse these results in Mercury. You should now be able to:

- Set up substructure searches in ConQuest
- Define 3D search parameters, including distances, angles and non-bonded contacts
- Run a combination search in ConQuest, with suitable filters
- Export search results to Mercury
- Use the calculator tool in the Data Analysis Module
- Plot selected data and apply filters
- Export data in spreadsheet format and save data as a database

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

Next Steps

If you have enjoyed this workshop, you might like to explore some of our other self-guided workshops which use ConQuest and Mercury. In particular, we would suggest “Using 3D information in Searches in ConQuest”, “Intermolecular Interaction Searching in ConQuest” and “Searching and Analysing Metal-Organic Structures using ConQuest and Mercury”.

Feedback

We hope this workshop improved your understanding of searching with *ConQuest* and data analysis using *Mercury* and you found it useful for your work. As we aim to continuously improve our training materials, we would love to hear your feedback. Follow the link on the workshop homepage and insert the workshop code, which for this self-guided workshop is *CQ-007*. It will only take 5 minutes and your feedback is anonymous. Thank you!

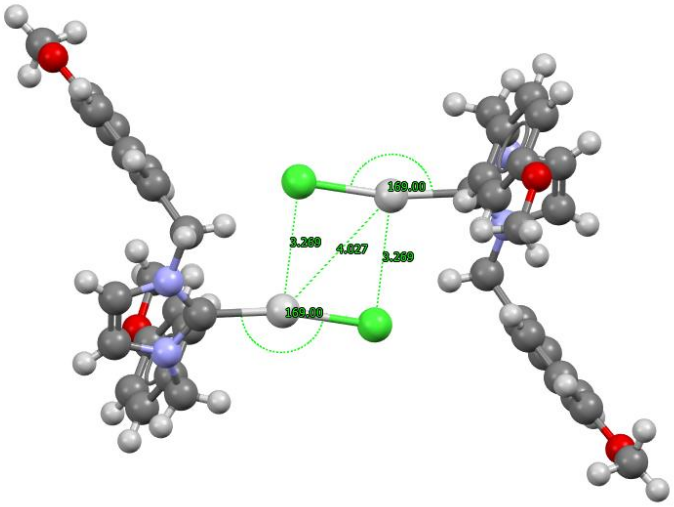
Glossary

Hitlist
A hitlist is a subset of CSD entries which can include search results, refcode lists, or the results of combining these.

Non-bonded contact
This type of search can be used for finding intermolecular interactions such as hydrogen bonds, halogen–halogen interactions, intramolecular non-covalent interactions and generating tables of geometries for non-bonded interactions. Non-bonded contacts between atoms are defined in terms of distance limits.

Refcode
A refcode is a CSD entry identifier comprising six letters e.g. ABACOF. Two digits identifying additional structure determinations e.g. ABACOF03.

Van der Waals radius
The van der Waals radius of an element is one half of the distance between the closest approach of two non-bonded atoms of a given element.



Non-bonded contacts between Ag and Cl in CSD entry UPAFAU.

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.

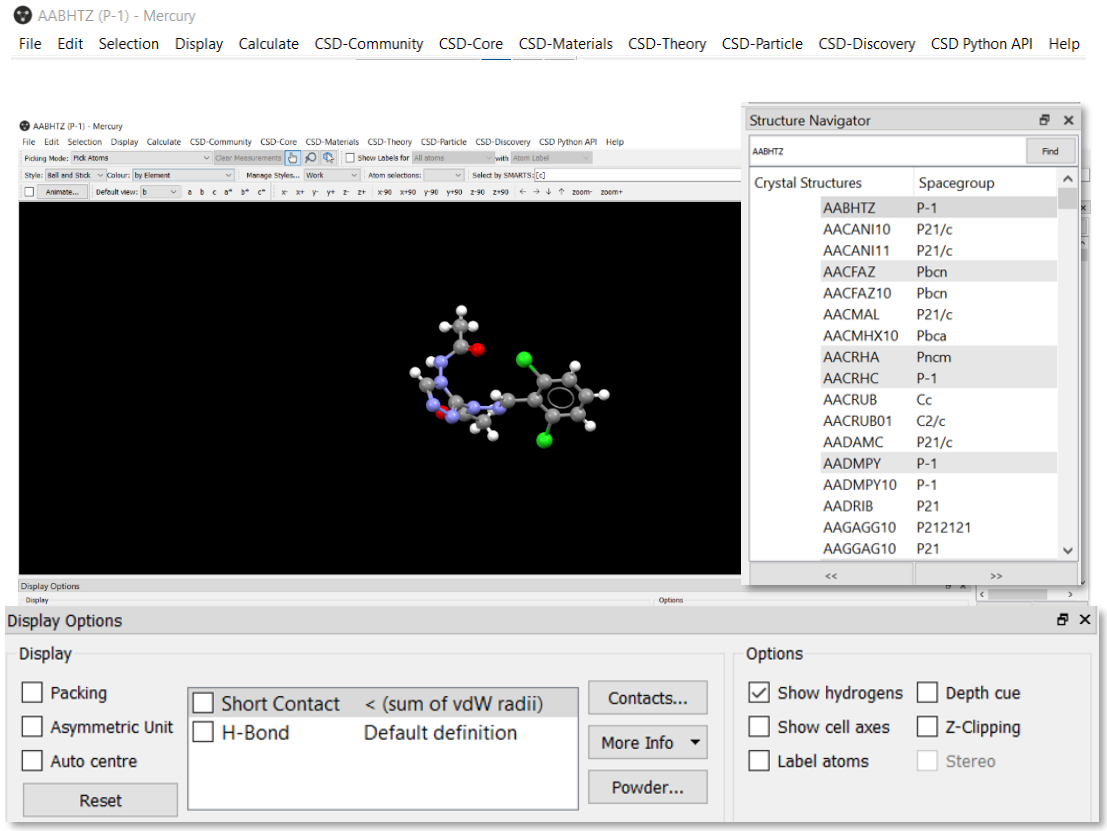
In the **Mercury interface** we find:

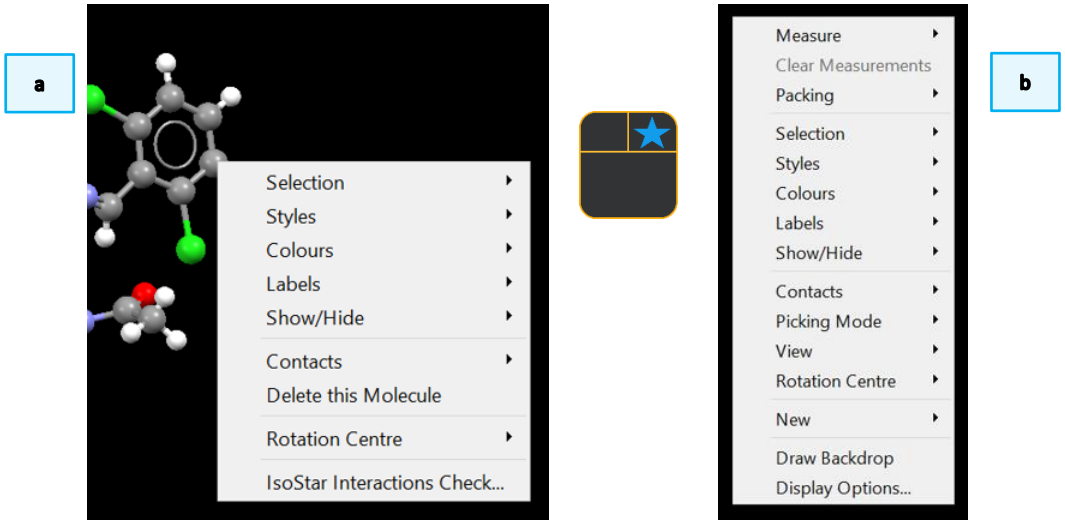
- **At the top**: list of menus from which we can access visualization and analysis options, and other CSD components such as CSD-Materials.
- **On the right-hand side**: the **Structure Navigator**, with the database loaded (depending on your licence). The Structure Navigator allows you to select a 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.

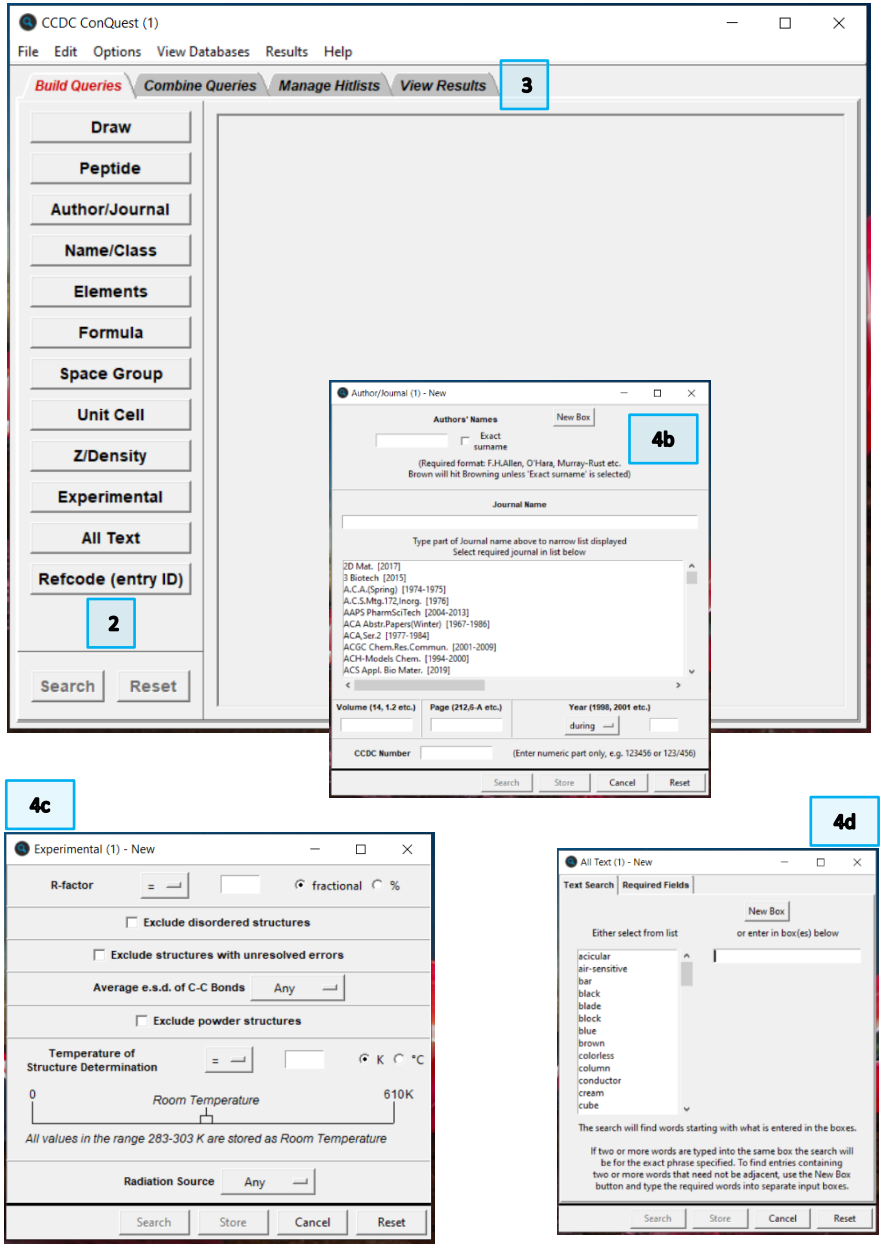
Right click:
a) Near a molecule and
b) Away from a molecule





Review. ConQuest Interface

1. Launch ConQuest by clicking the ConQuest Icon on your desktop or launching it from the Start or Applications menu.
2. The ConQuest main window shows all the search routines you can perform on the left-hand side of the window.
3. The row of tabs across the top of the window will guide you through the steps of the search process.
4. Some example searches are
    a. Draw – substructure and 3D information searching
    b. Author/Journal – bibliographic searching
    c. Experimental – experimental set up searching
    d. All Text – generic text-based searching
5. The majority of the searching we will do in these tutorials will be substructure searching, so we will focus on the Draw tab here.



Review. Draw Window

All drawing takes place in the central white area of the *Draw* window. In addition to creating 2D chemical structure sketches, the *Draw* window allows for the inclusion of 3D parameters for searching or for filtering.

ConQuest sketching conventions

- Left click in the sketcher to insert the selected atom type
- Left click and drag to sketch two bonded atoms
- Use the **Edit** button to modify properties of or delete atoms, bonds or entire substructures
- Right-click on atoms or bonds to modify their properties
- Use the **Templates…** button to pick from a list of CSD editor devised and drawn substructures
- Use the **More…** button to find less frequently used element types, or generic atom type groups (e.g. halogens), or define custom element combinations (e.g. C or N or O).