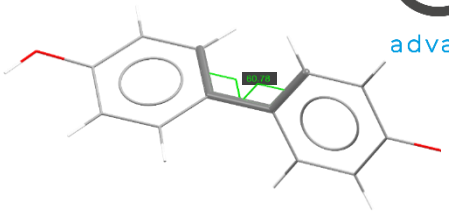
In-depth analysis of molecular geometries in Mogul (MOG-002)

2024.1 CSD Release



CCDC
advancing structural science

Table of Contents

Introduction 2
Learning Outcomes 2
Pre-required Skills 2
Materials 2
Example 1. Using Mogul to assess intramolecular geometry: torsion angles 3
Conclusions for Example 1 8
Example 2. Using Mogul to explain activity data 9
Exercise: Learn more about this unusual torsion using ConQuest. 12
Conclusions for Example 2 17
Summary 18
Next Steps 18
Feedback 18
Glossary 19

Introduction

Structural chemists can use the Mogul Geometry Check tool to validate the three-dimensional *conformation* of a particular molecule. The data from structures in the CSD can be used to show the most likely conformation a particular bond, angle, *torsion* or ring would 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 examples from this workshop demonstrate approaches to gain more insights from your data and to critically interpret the results obtained from Mogul.

Learning Outcomes

In this workshop you will learn:

- How to apply filters for the Mogul Geometry Check search.
- How to restrict the visualisation of the results to specific parts of the histogram and why you would do that.
- What to look for when analysing not unusual or unusual torsion angles.
- How to complement Mogul Geometry Check with ConQuest 3D search and Mercury analysis.

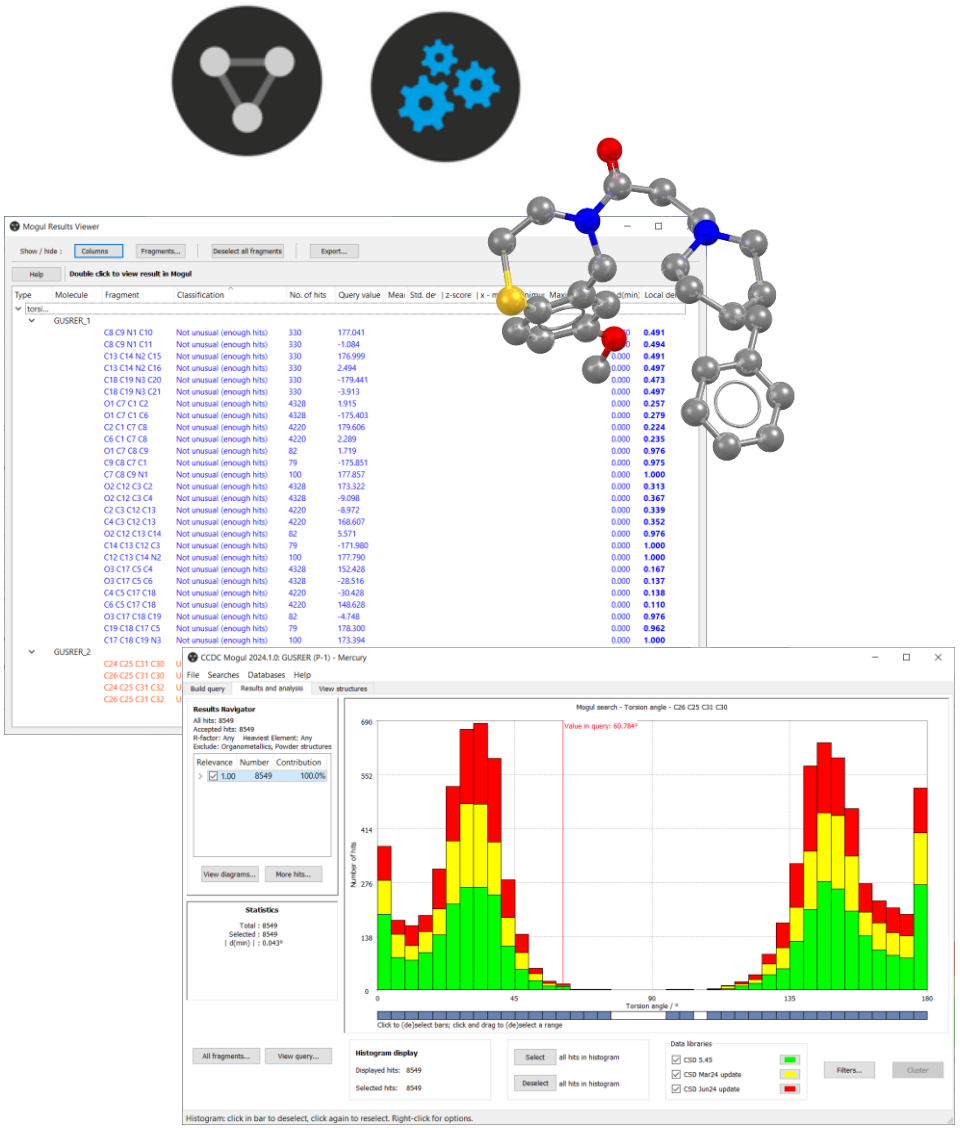
The exercises in this handout will take approximately ***45 minutes*** to be completed. The words in *Blue Italic* in the text are reported in the Glossary at the end of this handout.

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, and basic knowledge of Mogul, namely, how to run a search and how to access and read the results table. You can find the basics of visualization in Mercury and the basics of Mogul as CSDU modules (videos, exercises, and final test) or as self-guided workshops in the CSD-Core workshop area on our website.

Materials

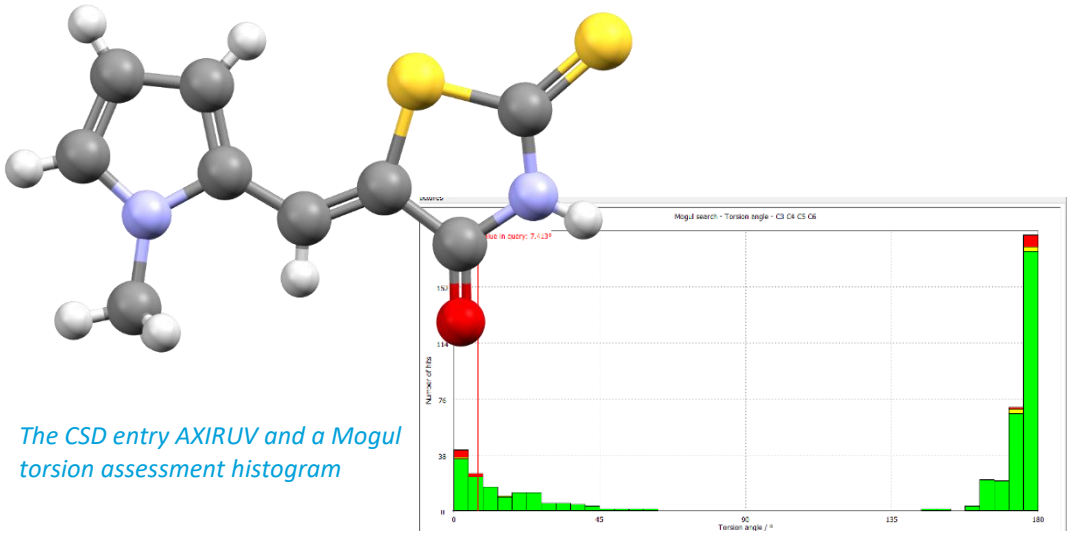
There are no additional materials required for this workshop.



Example 1. Using Mogul to assess intramolecular geometry: torsion angles.

By using the statistical distributions of similar *fragments*, Mogul can provide insights on whether your 3D geometry is appropriate, and flag values that are too far outside the norm.

In this example, you will use Mogul to assess the torsion angles of a given molecule and investigate patterns in the distribution. You will find results for four different torsion angles, all resulting as "Not unusual" and will be guided to further analyze one of them and use additional filtering options for a complete view.



The CSD entry AXIRUV and a Mogul torsion assessment histogram

1. Launch Mercury and type AXIRUV in the Structure Navigator toolbar.

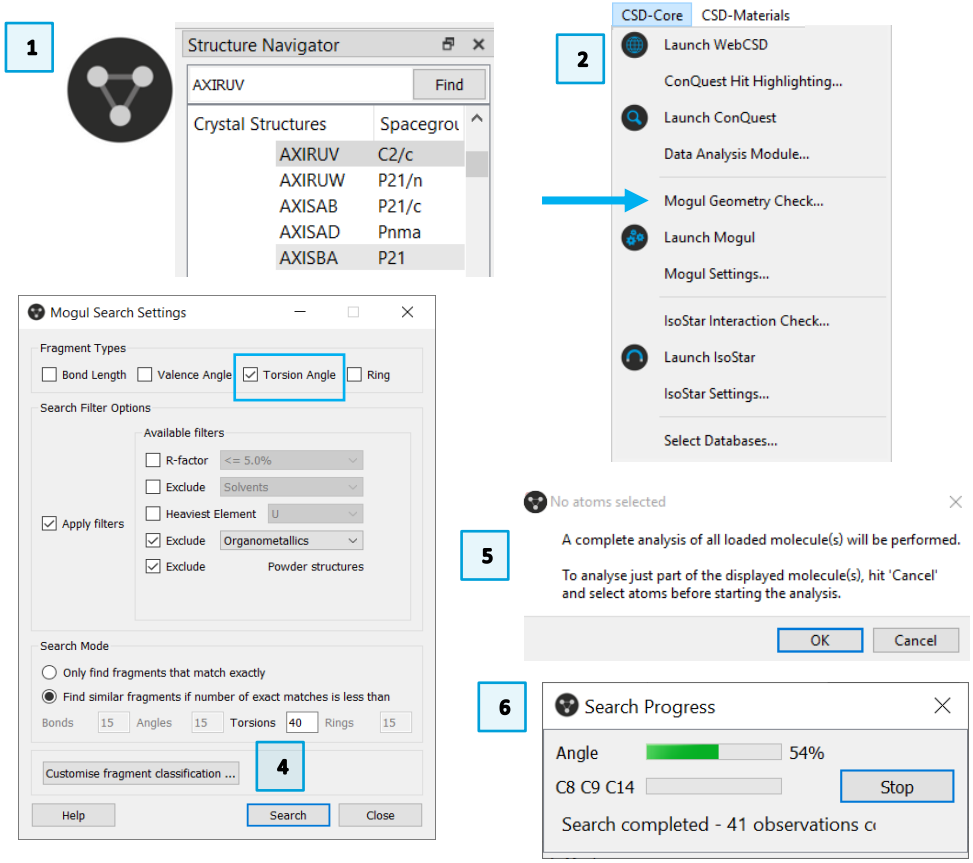
2. To start a geometry analysis, click the CSD-Core menu and then click *Mogul Geometry Check* from the dropdown menu.

3. This brings up the *Mogul Search Settings* dialogue box. You can typically use the defaults in this window, but we are interested in the torsion angle between the rings. To streamline our search, we will select Torsion Angle in the Fragment Types option (untick the remaining types) and tick the boxes for *Apply Filters*, *Exclude Organometallics*, and *Exclude Powder structures*. We will leave the remaining default options as is.

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.



7. When the search is complete, your results will be displayed in the **Mogul Results Viewer.**

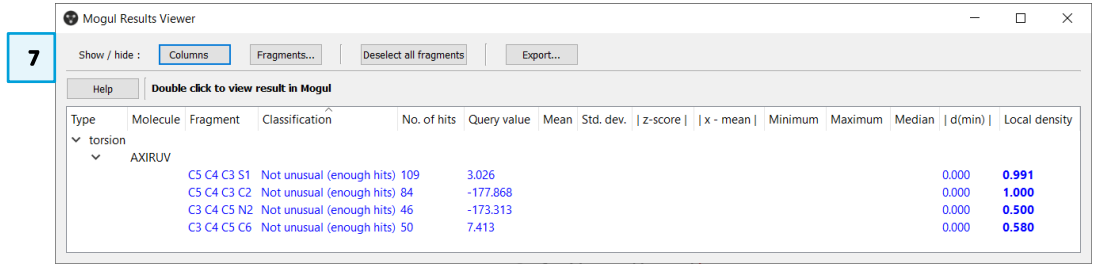
8. The results are colour-coded. Unusual values are typically flagged in orange, however there are no unusual values here. *Please, be aware that the number of hits might differ based on your version of the database. The data in this workshop are obtained with the 2024.1 + 2 Data updates version of the CSD.*

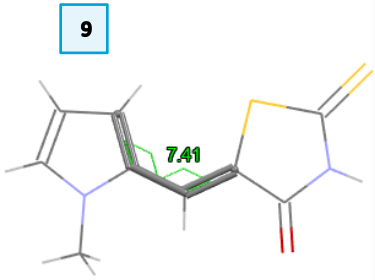
9. Clicking on each entry in the **Mogul Results Viewer** will highlight the corresponding value in the Mercury display window. Although the entries are not unusual, it is useful to review the torsion angle distribution. Double-click on the last entry C3 C4 C5 C6 to bring up the data from the Mogul library.

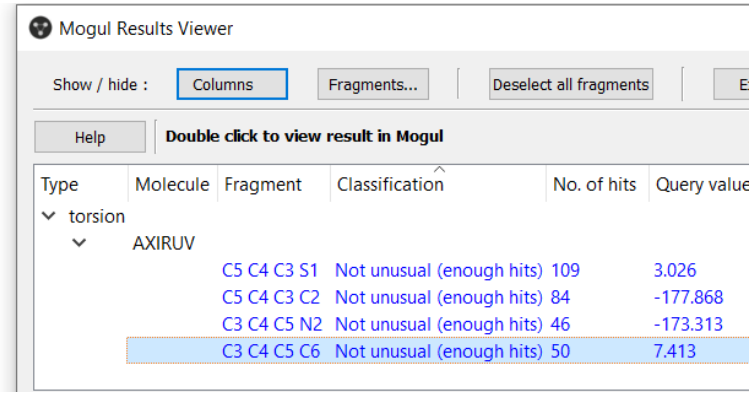
10. The red line marks the value of the torsion angle from your molecule (the query).

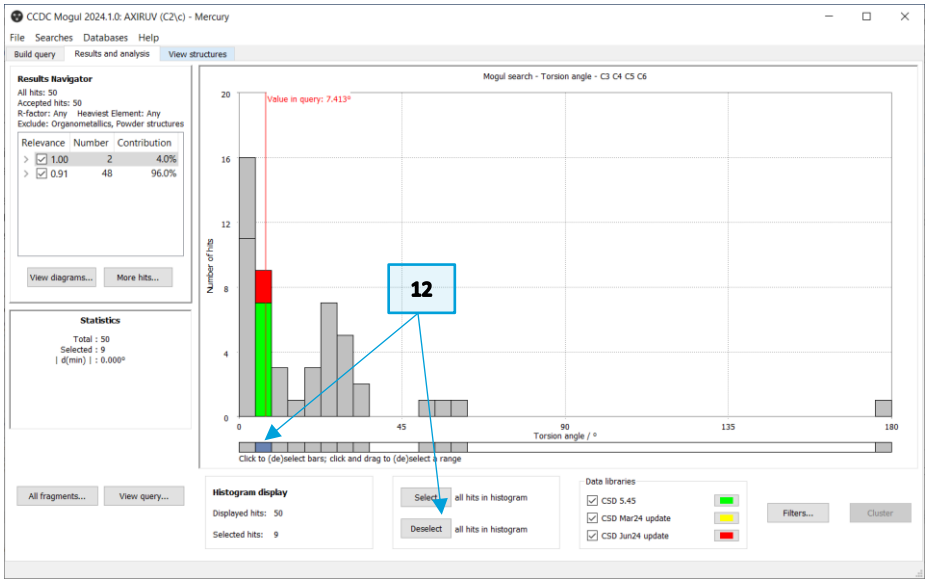
11. The histogram shows the data from the CSD, colour coded by update. Note: you can click the colour swatches under Data libraries to change the colour.

12. To see which structures contribute to a certain bar on the histogram, first click **Deselect** (all the bars will turn grey), and then click the bar directly under the red query line. This will highlight that specific bar of the histogram in colour.







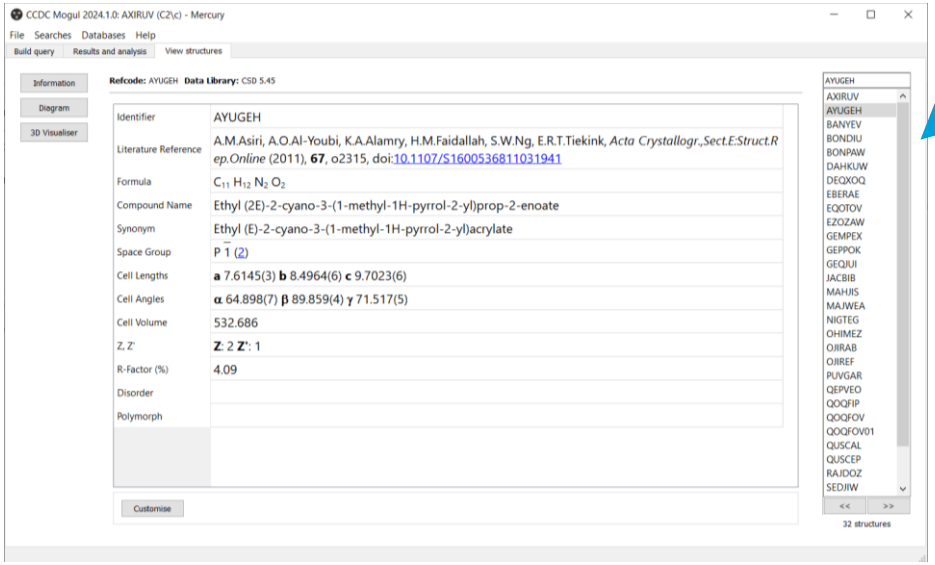


13. Return to the **Mogul Results Viewer** window (seen in **Step 7**) and double click on each entry to view the histogram in Mogul. You will notice that the query falls within the major distribution of observed hits for all torsion angles. We will take a closer look at the distribution of one of the torsion angles to understand more about how frequently its observed.

14. You should have the C3 C4 C5 C6 torsion angle histogram in the Mogul windows, but if you do not, then double-click again on the entry in the **Mogul Results Viewer**. Click **Deselect** all hits in histogram then select the torsion angles in the populous region. There are 46 of the total 50 structures in this region representing 92% of the total hits.

15. Now click the *View structures* tab, near the top of the window, to see a list of refcodes included in this region. The default view for this window is the 2D diagram. Scroll through the refcodes on the right side of the window to view different structures.

16. Click the **Information** button on the left to see further details about the structure

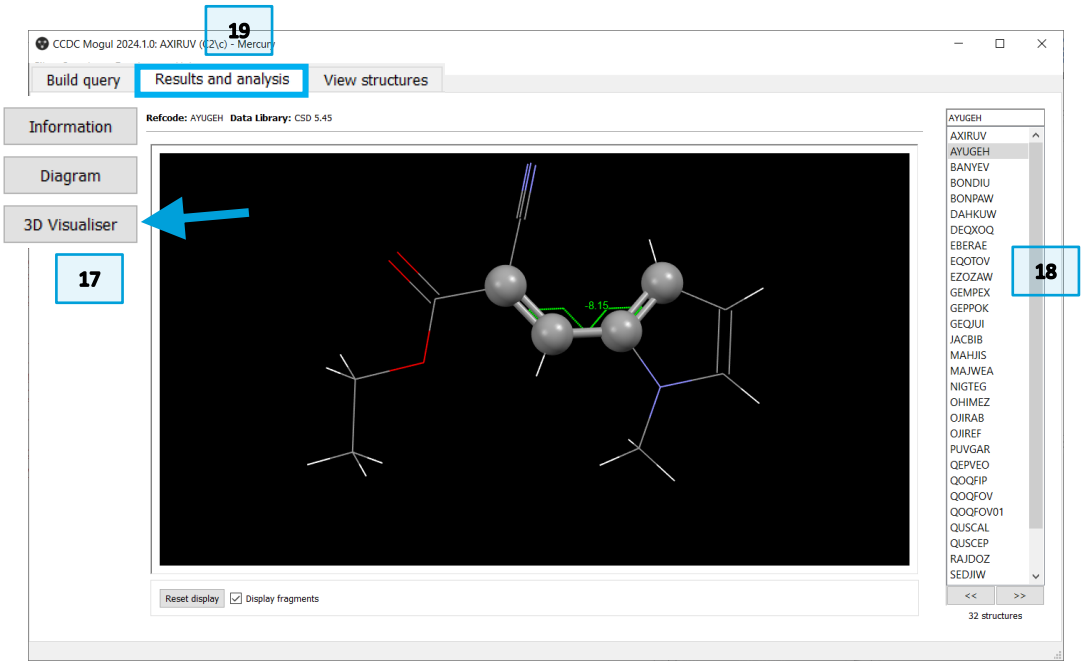


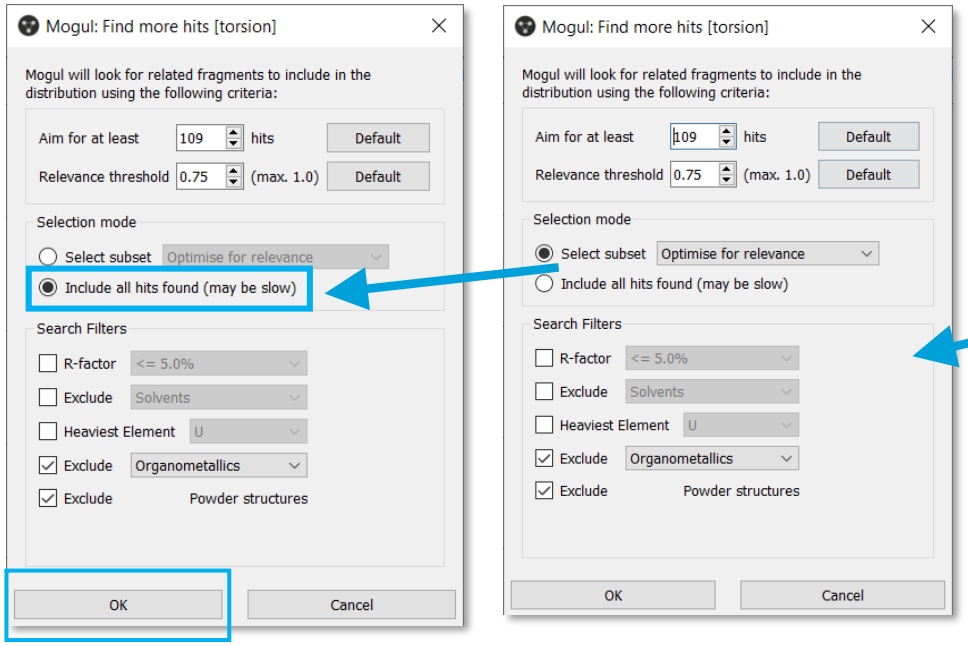
17. Click the **3D visualizer** button to see a 3D rotatable view of the structure.

18. Scroll through the refcodes to view the 3D structures that populate this region. You will notice that the torsion angles are in structures similar to our query structure, *i.e.* not part of a macrocycle (large ring), which influences the torsion angle.

19. Let's take a closer look at the structures that may have been omitted from our histogram and see if it changes our results. Click on the *Results and analysis* tab to return to the histogram view.

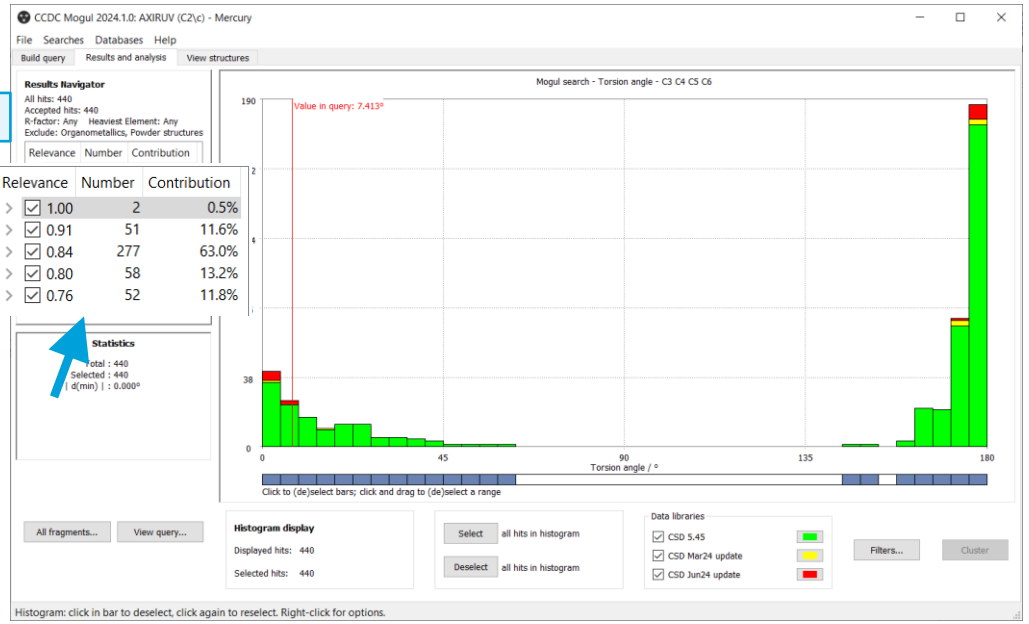
20. In the *Results and analysis* window, click **Select** all hits in histogram to return to the default view, then click *More hits…* to bring up the **Mogul: Find more hits [torsion]** window. You will notice that the default selection mode is *Select subset (Optimise for relevance)* and take note of the other default values here. Change the Selection mode to *Include all hits found (may be slow)* and click **OK**.

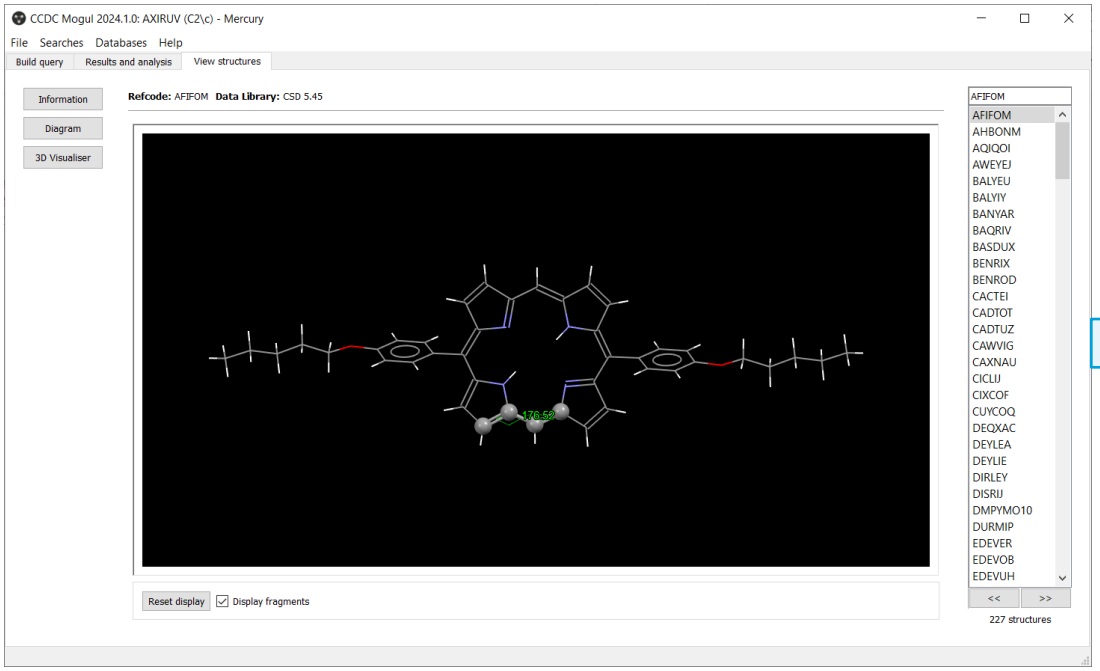


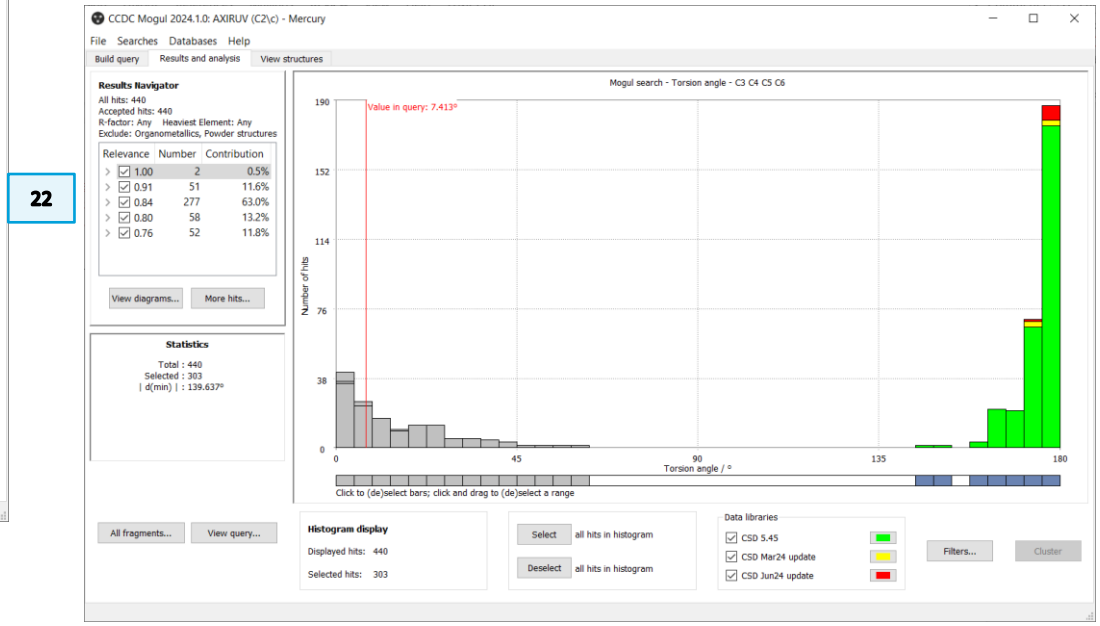


21. Note that the total number of hits has increased to 440 hits from 50 and now includes structures with relevance values of 0.84, 0.80 and 0.76. We can also see that our query value no longer falls within the major distribution of observed hits.

22. Let's take a closer look at the structures in the major distribution now. First click **Deselect** and then select the structures in the major distribution. Click the *View structures* tab and select **3D visualizer**. Scroll through the refcodes and you will see that the structures that populate this region are in most instances part of a macrocycle and thus it is reasonable that while there are more hits in this region, they were omitted in our default results and have a low relevance value in our search.



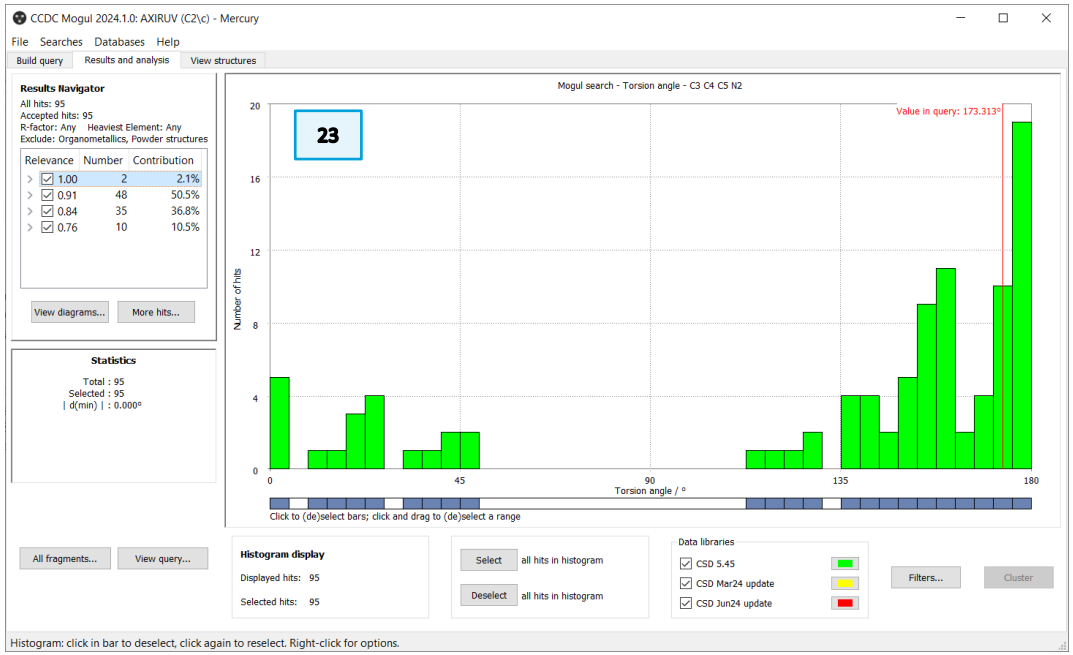




23. Repeat steps 20 and 22 for the other 3 torsion angles. You will notice that increasing the number of hits shown does not change the histogram significantly and does not change where the query value falls.

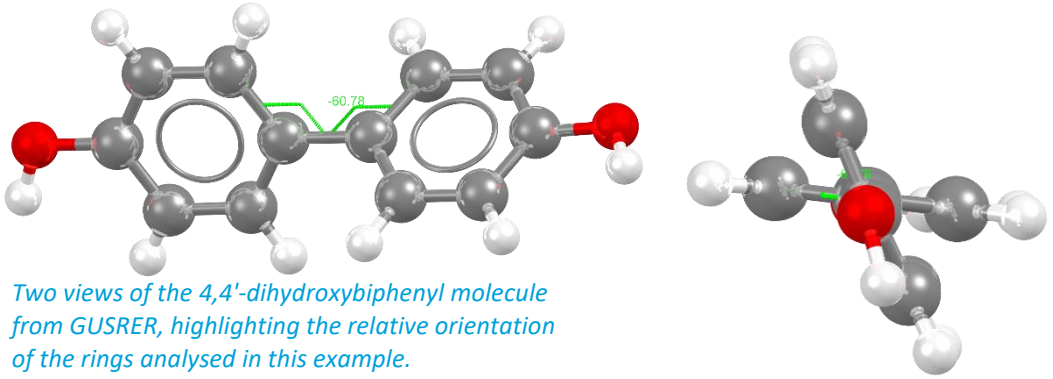
Conclusions for Example 1

In this exercise we have performed a Mogul Geometry Check on a molecule from the CSD. We have learnt how to analyse our results, reviewing relevant and less relevant structures and how those determinations are made.



Example 2. Using Mogul to explain activity data

In this example you will see how to investigate unusual torsions in Mogul and how to complement the Mogul analysis with ConQuest and Mercury to gain more insights into your structure**.**



Two views of the 4,4'-dihydroxybiphenyl molecule from GUSRER, highlighting the relative orientation of the rings analysed in this example.

1. Launch Mercury and type GUSRER in the Structure Navigator toolbar. You will see that the structure has two components.

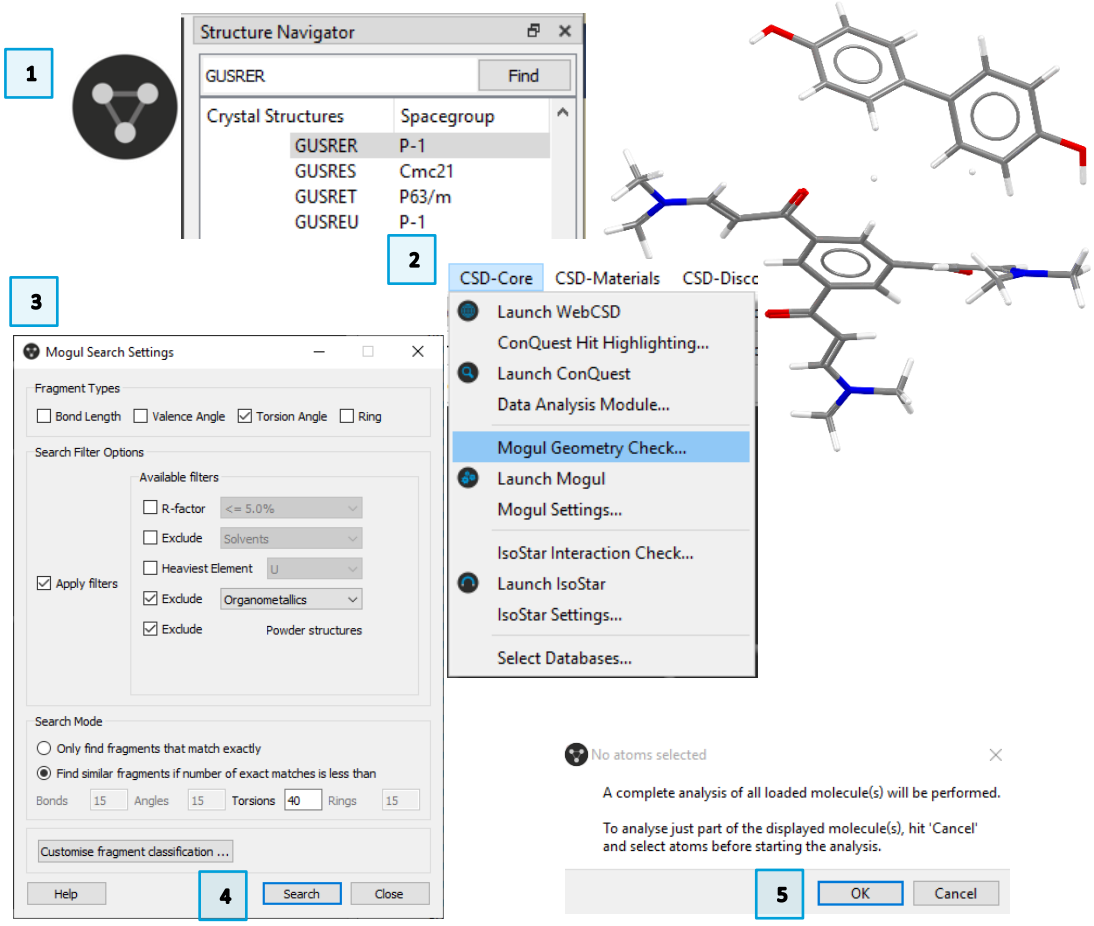
2. To start a geometry analysis, click the CSD-Core menu and then click *Mogul Geometry Check* from the dropdown menu.

3. This brings up the *Mogul Search Settings* dialogue box. For this example, we are interested in only torsion angles. To streamline our search, we will select only Torsion Angle in the Fragment Types option and tick the boxes for *Apply Filters*, *Exclude Organometallics*, and *Exclude Powder structures*. We will leave the remaining options as the default.

4. Click **Search** to start.

5. A dialogue box will pop up to warn you that you are going to perform the check for all loaded molecules. Click **OK** to continue.

6. The search will begin to run. You can follow its progress in the *Search Progress* dialogue box.



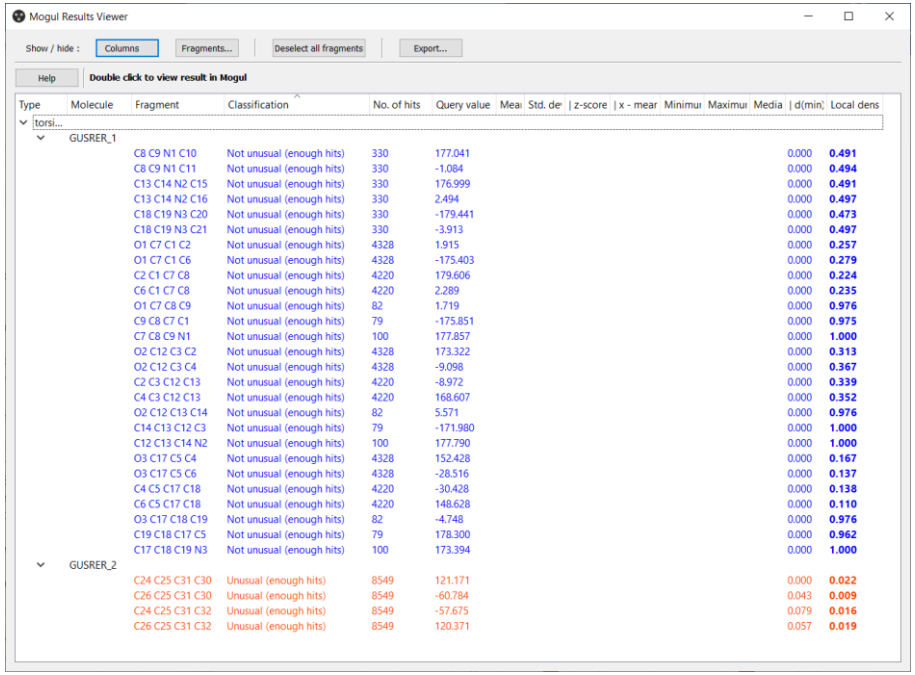
7. When the search is complete, your results will be displayed in the **Mogul Results Viewer.**

8. The results are colour-coded. Unusual values are typically flagged in orange. In this example we can see that the unusual values here are found only in one of the two molecules, the 4,4'-dihydroxybiphenyl (listed as GUSRER_2 in the **Mogul Results Viewer**). *Please be aware that the number of hits might differ based on the version of the data update. The data in this workshop are obtained with the CSD 2024.1 + 2 data updates version of the CSD.*

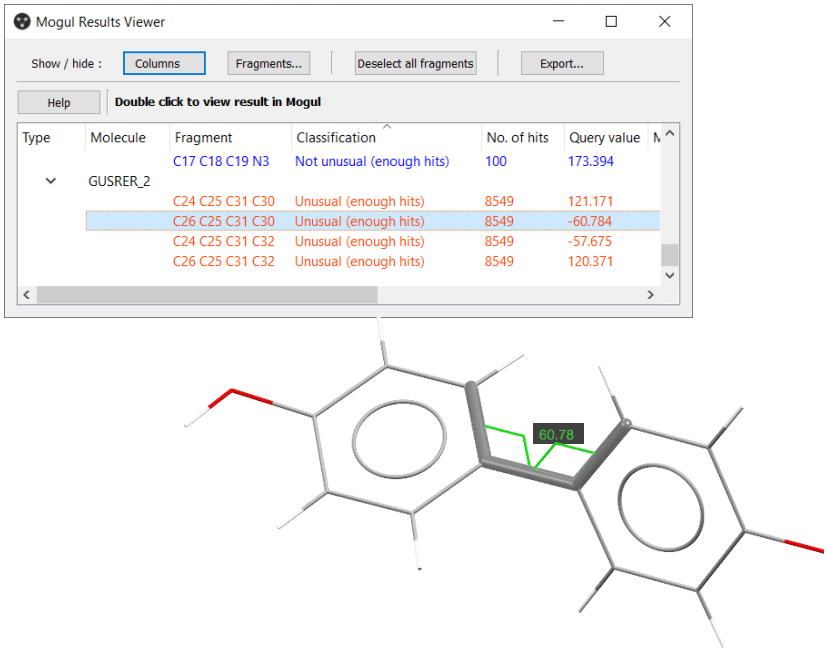
9. The four torsion angles analysed for the 4,4'-dihydroxybiphenyl are all representative of the relative rotation between the two aromatic rings. We will investigate C26 C25 C31 C30. Click once on the entry in the **Mogul Results Viewer** to highlight the corresponding value in the Mercury display window. Then double-click to open the Mogul interface with the histogram.

10. From the histogram we can see that the orientation in our molecule (red line, value 60.784°) is very unusual, compared to data in the CSD that show a preferred orientation at around 35° and 145°.

7



9



10

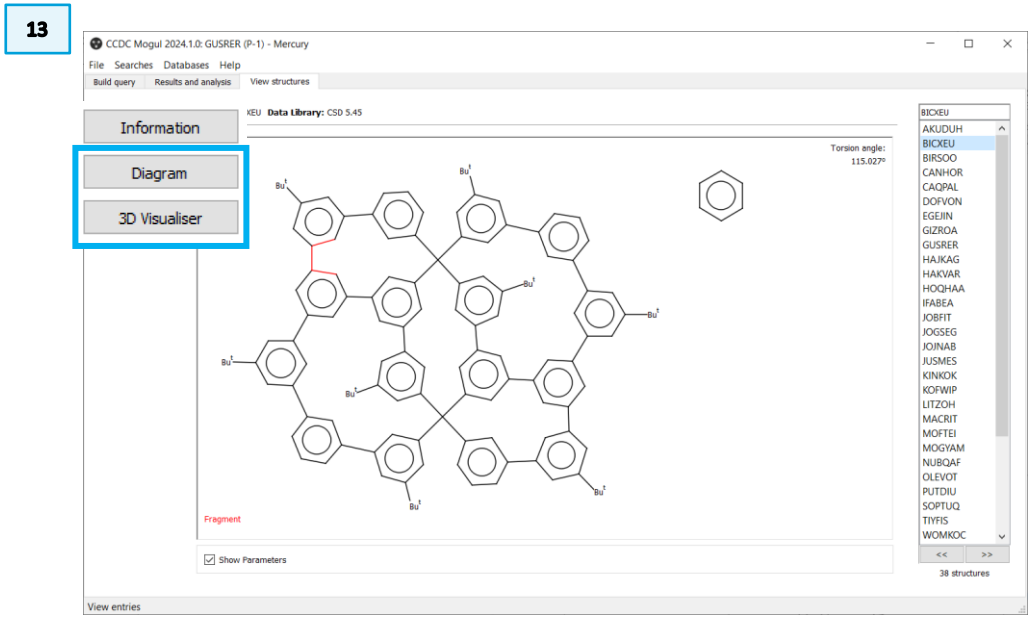
11. To see structures that contribute to the central part of the histogram (~60° to ~120° (please keep in mind that by symmetry these are equivalent) and assess whether these are similar to the query molecule, deselect all the bars that are not in this area by clicking and dragging the mouse arrow on the squares under the range of histogram bars to deselect. The deselected bars will turn grey.

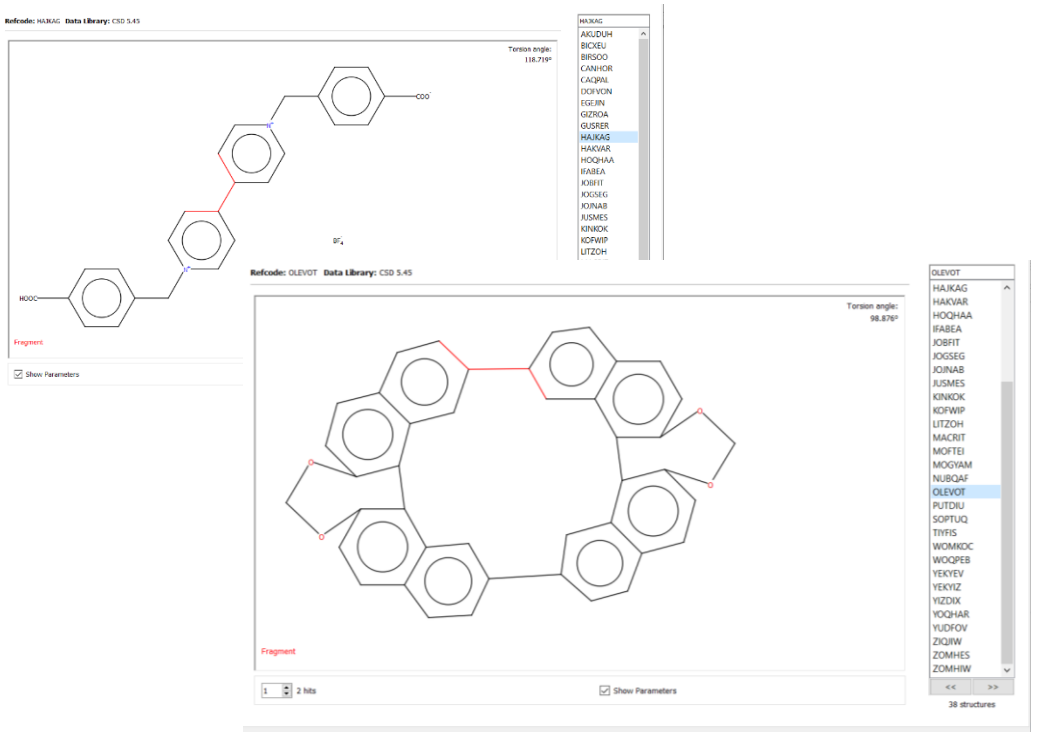
12. Click *View Structures* to access the visualisation tab and inspect the structures that contribute to this part of the histogram.

13. Scroll through the structures and **reflect on the following questions**: Are these molecules similar to the one we are studying, e.g., do they have oxygen substituents on the two aromatic rings? Can you identify a pattern or any similarity among these structures that might explain why they are adopting this unusual orientation of the rings? You can switch between the Diagram and 3D visualisation if needed.

14. We observe different types of structures (macrocycles, different substituents attached to the aromatic rings), but there is not an obvious pattern that might explain the unusual torsion angle.

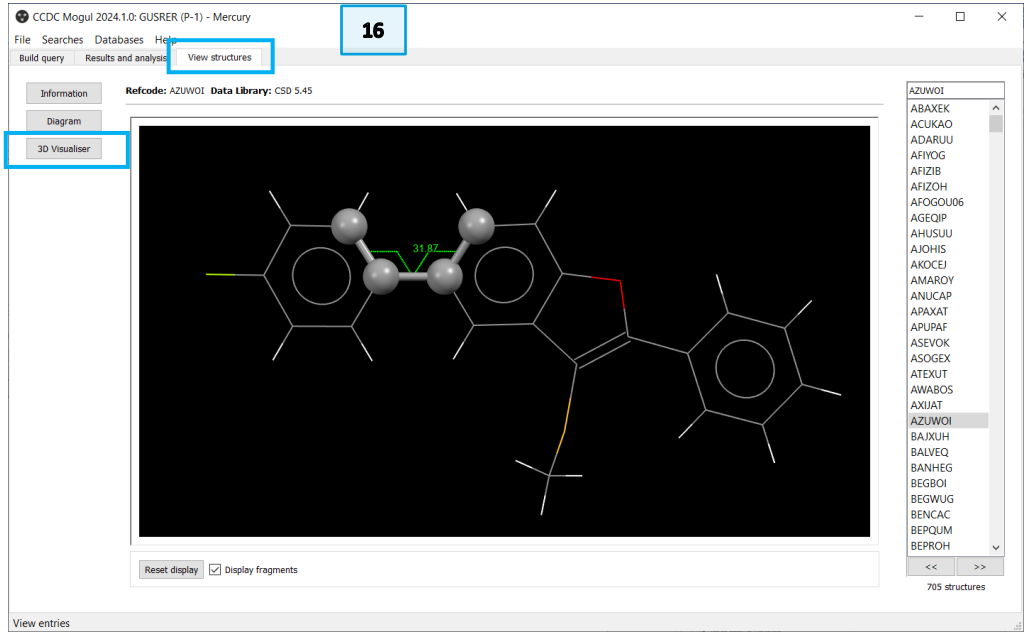
13





15. To see what the most common orientation looks like, go back to the *Results and analysis* tab, deselect all the bars using the **Deselect** button, and then click the two highest populated columns (~35° and ~145°) to select them.

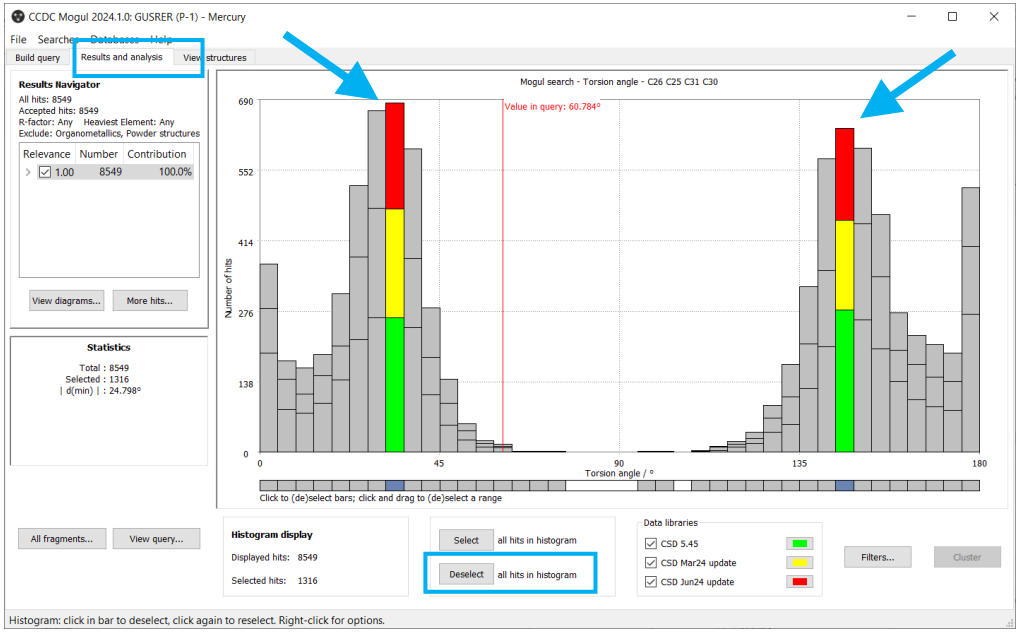
16. Go to the *View structures* tab, and investigate these structures in the **3D Visualiser** to assess what the most common torsion angle looks like.

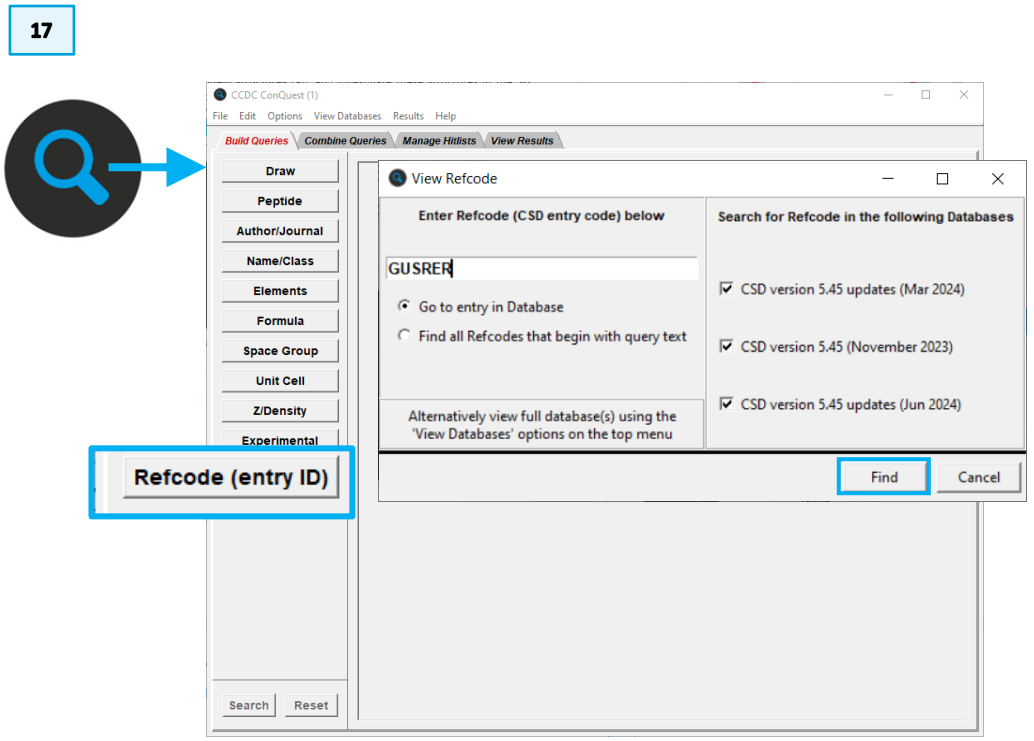


Exercise: Learn more about this unusual torsion using ConQuest.

To investigate further how unusual this torsion is, we will perform a search in ConQuest of 4,4'-dihydroxybiphenyl and analyse the value of the torsion in our reference structure compared with other structures.

17. Launch ConQuest  and in the **Build Queries** tab click **Refcode (entry ID)**. In the textbox enter the refcode of the structure we are investigating, GUSRER, then click **Find**.





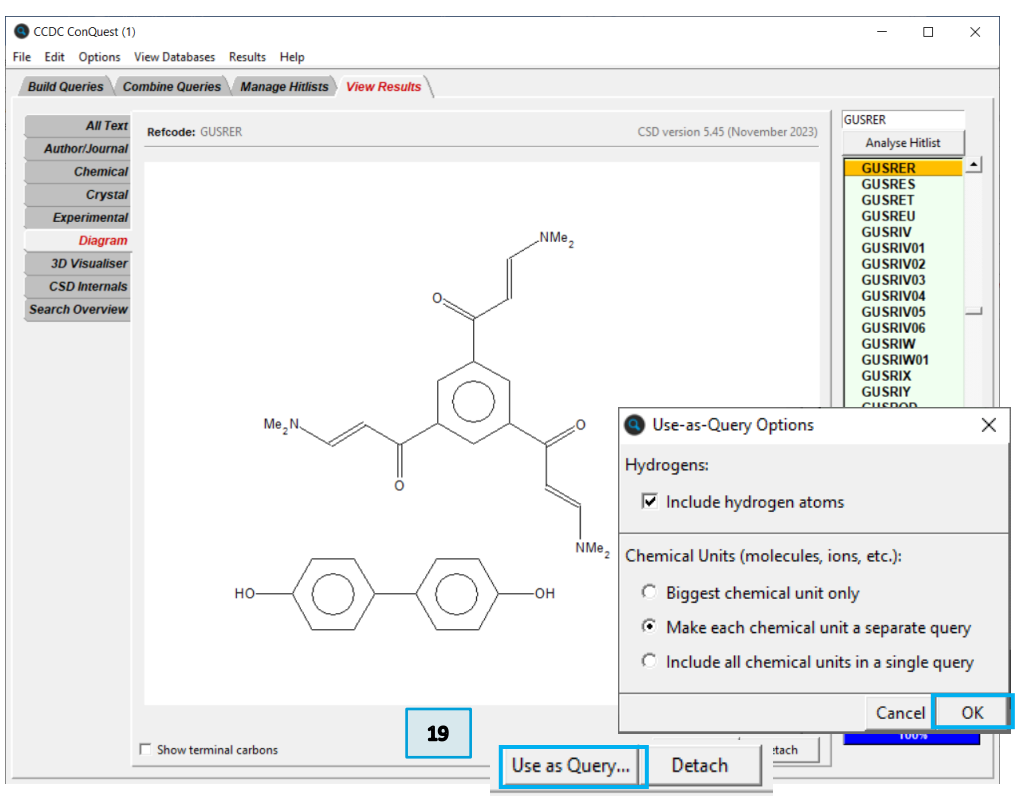
18. This will bring you to the **View Results** tab, where the first result is the structure of interest.

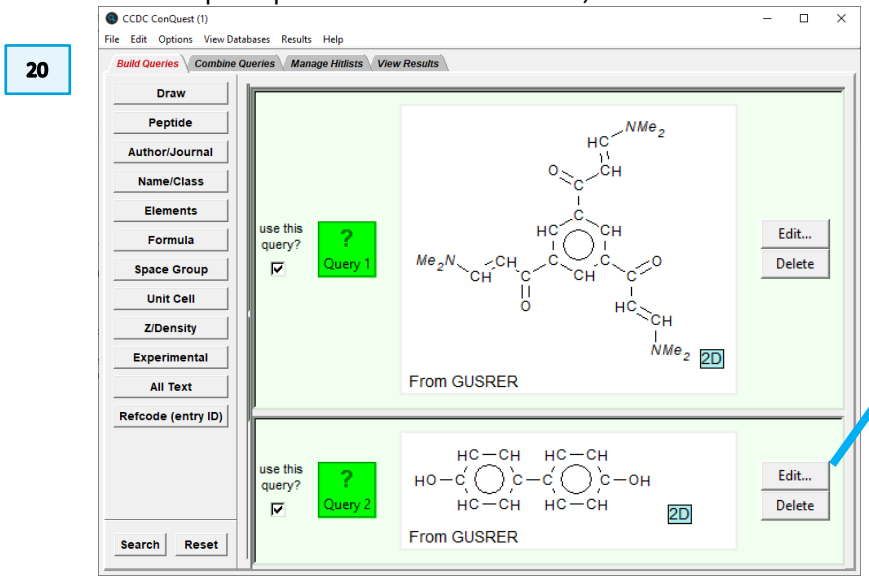
19. We want to use the 4,4'-dihydroxybiphenyl molecule as the starting point for our next search. We click on **Use As Query** and select the options *Include hydrogen atoms* and *Make each chemical unit a separate Query*, then click **OK**.

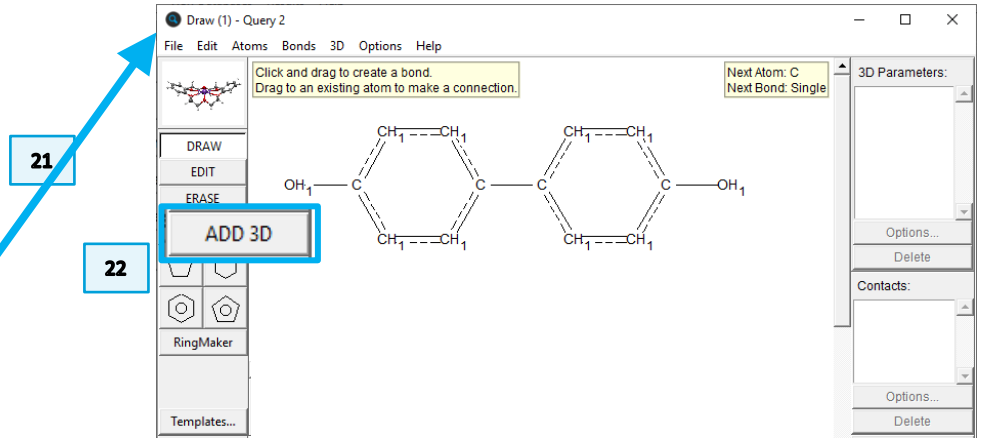
20. This will bring us back to the Build Queries tab, where we can see the two molecules as Query 1 and Query 2. 4,4'-dihydroxybiphenyl is Query 2. Untick *use this query?* for Query 1.

21. We now need to set up a search to analyse the torsional angle for molecules in structures in the CSD like the one we are studying. Click **Edit** for Query 2 to access the **Draw** window. **Steps 17 to 21** enabled us to have the diagram of the molecule we need in a fast and efficient way. Otherwise, you could have drawn it directly from the Draw tab at **Step 17**.

22. To access the measurements of the torsional angle during our search, we need to set up the parameter in 3D. To do so, click **Add 3D**.



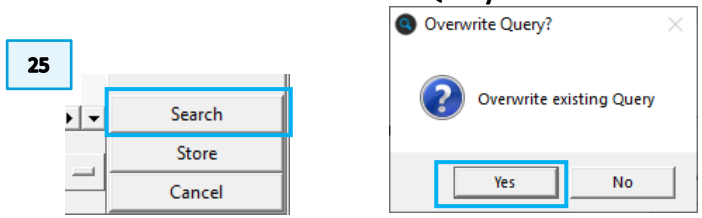




23. Now click on the four carbon atoms incolved in the torsion, in order: C5, C2, C1, C3. In the **Geometric Parameters** pop-up, click **Define** next to *Torsion*, then click **Done**.

24. You could include all four torsional angles relevant to the relative rotation of the two rings. If you wish to do so, repeat **Step 23** for the C5 C2 C1 C4 torsion, the C6 C2 C1 C3 torsion and the C6 C2 C1 C4 torsion. For this example we will not add any more torsions.

25. Now, click **Search**. When asked "**Overwrite Query?**" click **Yes**.

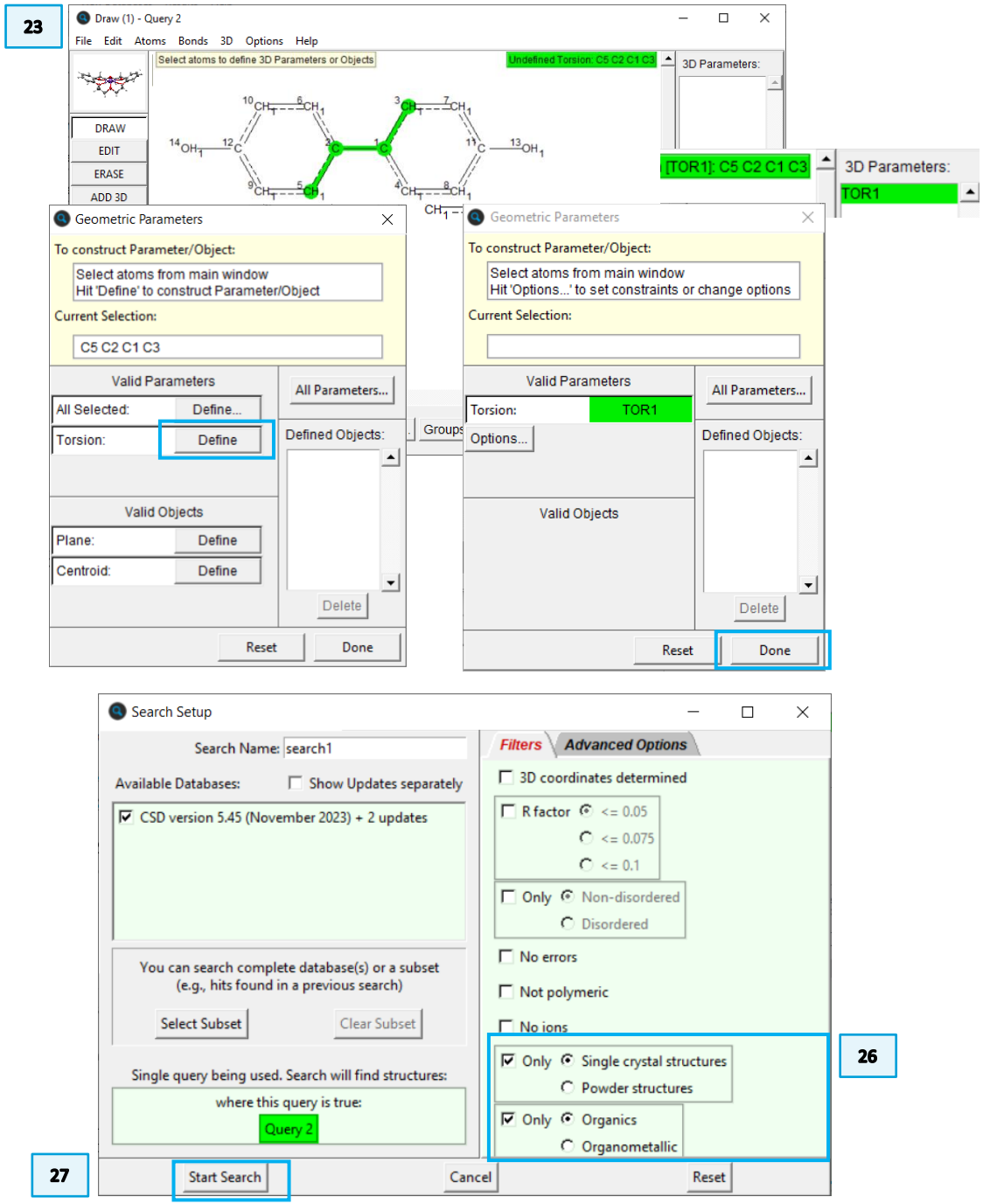


26. In the **Search Setup** window, tick as **Filters** *Only – Single crystal structures* and *Only – Organics*. This will make the search consistent with the Mogul analysis performed before.

27. Then click **Start Search**.

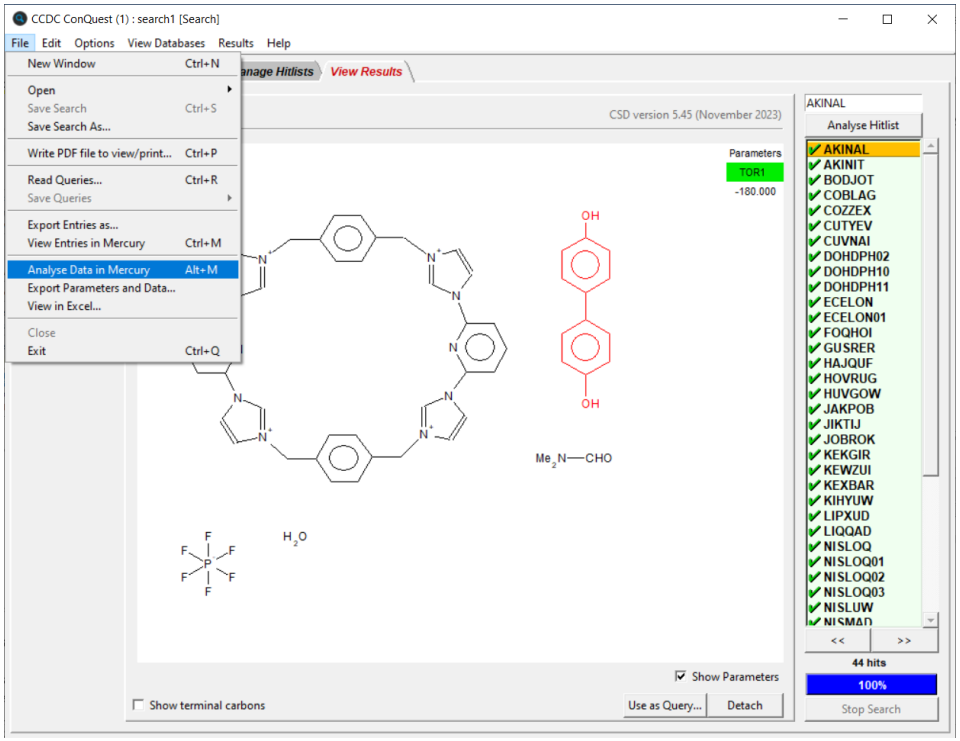
28. With CSD version 5.45 and 2 data updates (Mar, Jun 2024), we obtain 44 hits that we can see in the *View Results* tab. You can scroll through the hitlist and see which entries are in the list and in the 2D Diagram. You will see the fragment we looked for in red and the value of the torsion at the top right of the visualisation window.

29. We want to analyse more systematically the value of the torsion. To do so, we will need to export the data to Mercury. Go to *File > Analyse Data in Mercury*.

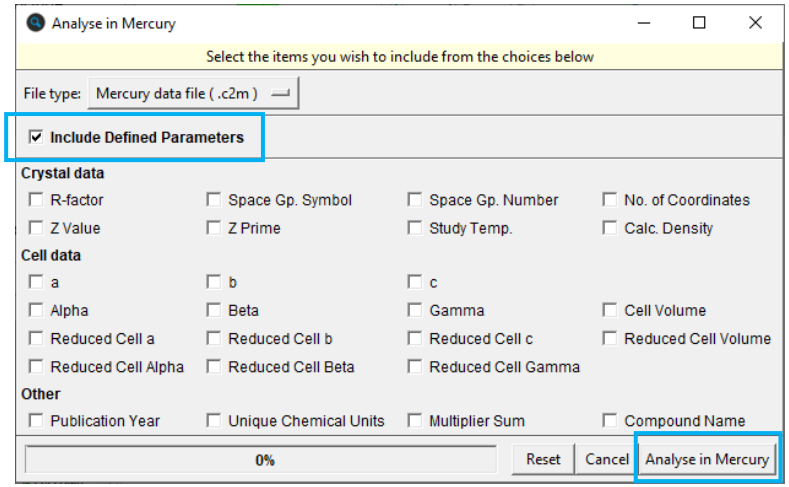


28

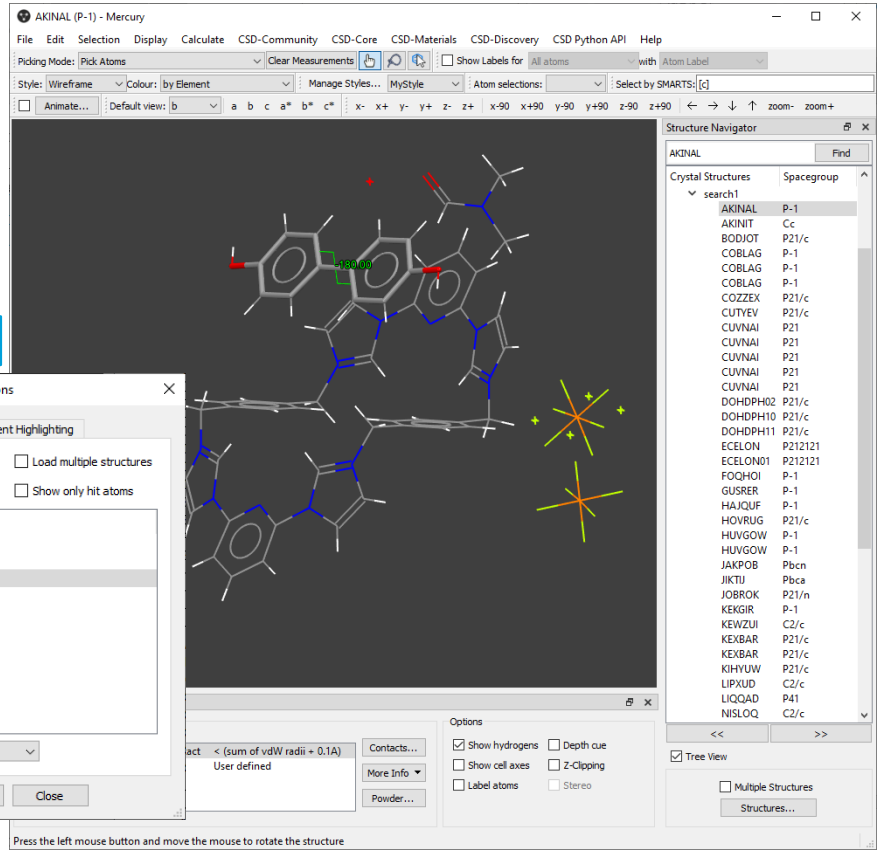
29



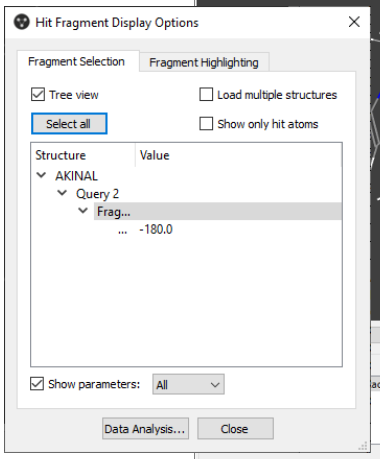
30



31a



31b

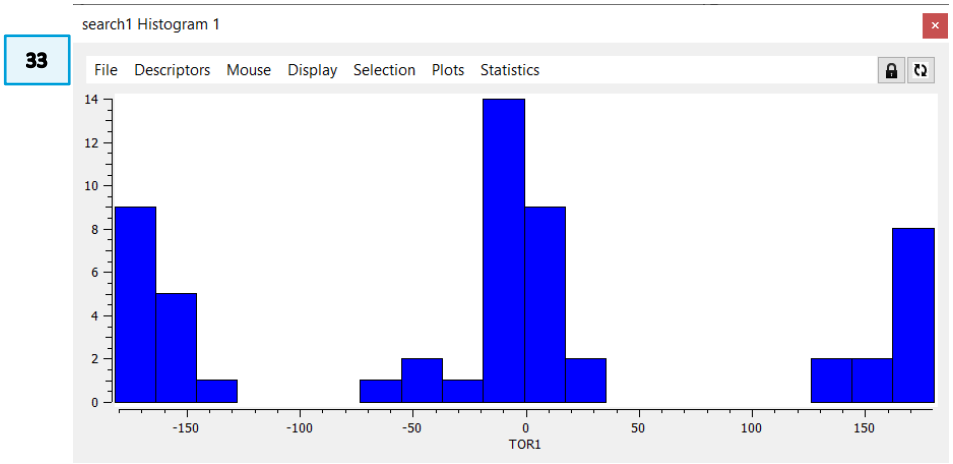


30. In the **Analyse in Mercury** pop-up ensure the option *Include Defined Parameters* is ticked. You may select other parameters if you want to include them in your analysis. Click the **Analyse in Mercury** button at the bottom right to proceed. This will automatically bring you to a new Mercury window.

31. In Mercury you have three new windows open:
    a. The Mercury interface with the list of refcodes from the search in the **Structure Navigator**. The torsion angle value is highlighted on the structure in the 3D viewr.
    b. The **Hit Fragment Display Options**. This is not of interest for this example, so we can close it.
    c. The **Data Analysis** window, which we will use to analyse the data.

32. In the **Data Analysis** table, select the column for the torsion angle under investigation, clicking on *TOR1*. To generate a histogram with the measurements go to *Plots > Histogram*.

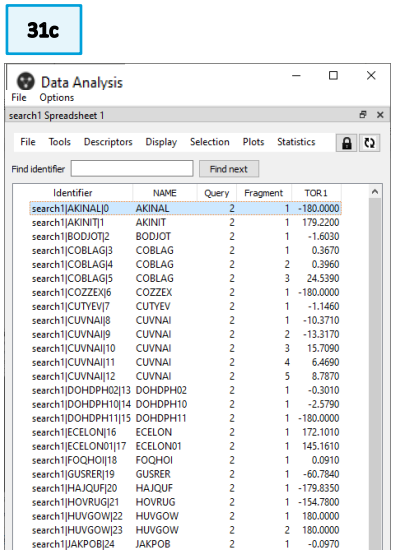
33. The histogram will appear at the bottom of the **Data Analysis** window. **Observe the data**: what do you notice? Where are the clusterings for this distribution? How does it compare to the one obtained from Mogul?

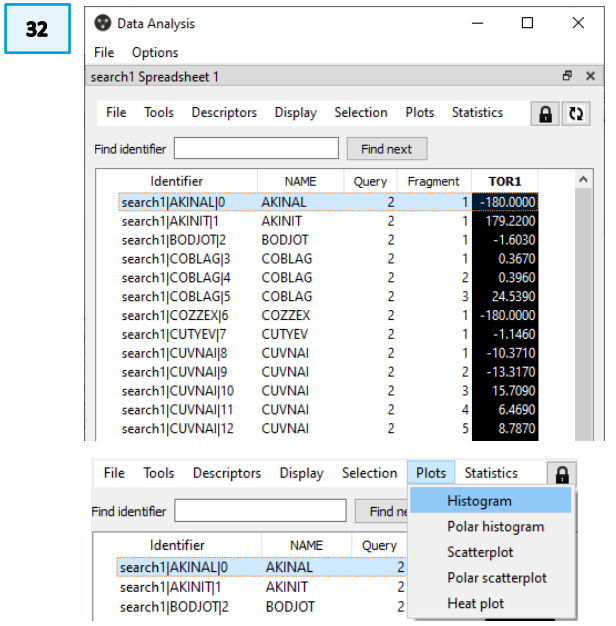


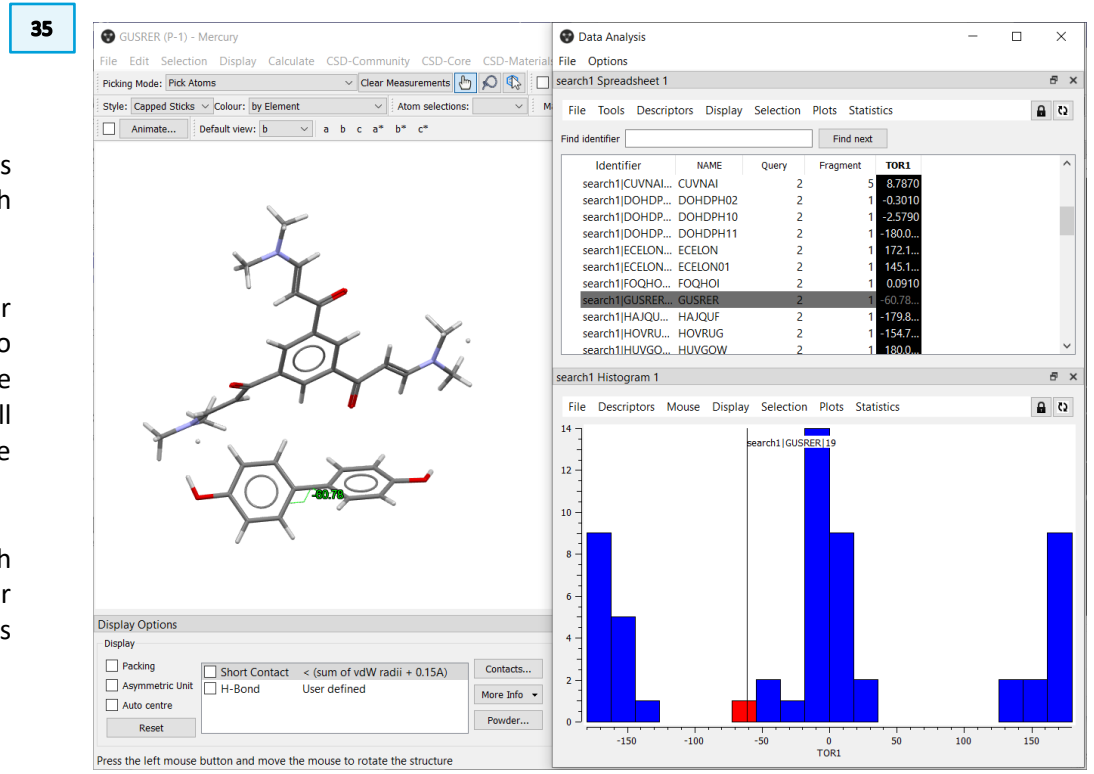
34. From this distribution, we can see clustering towards 180° and 0°, so it seems that this molecule prefers to have the two aromatic rings nearly in plane with one another, rather than twisted.

35. To compare the torsion in the molecule from the structure under investigation with this distribution, we can click on the row corresponding to the refcode GUSRER in the **Data Analysis** table and this will highlight the value on the histogram and show the structure in the Mercury interface. To see all structures in a bar of the histogram, click on the bar and you will see the corresponding rows highlighted in the table.

36. We can see that the 4,4'-dihydroxybiphenyl molecule in the structure with refcode GUSRER is the only result with a torsion angle of ~60°. Searching for the relative orientation of two aromatic rings with *para* OH substituents







shows that the conformation adopted in GUSRER is highly unusual in this scenario, even more unusual than indicated in the original Mogul analysis.

Conclusions for Example 2

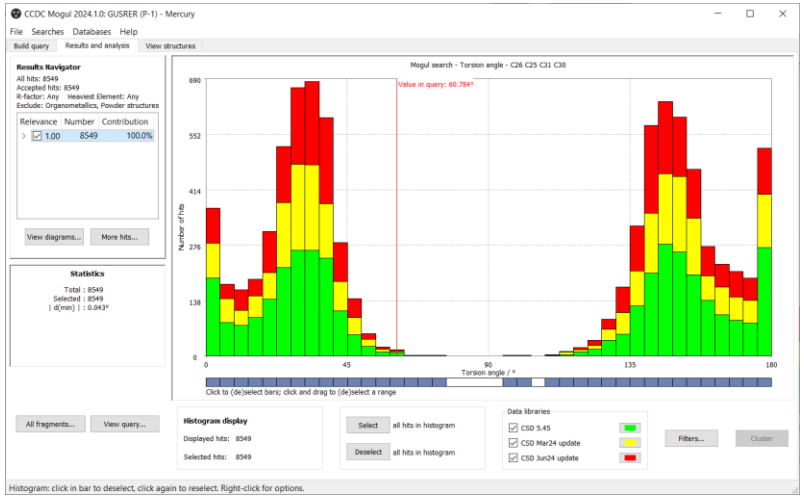
This example shows how to combine advanced searches technique in the CSD-Core suite to analyse and validate structures.

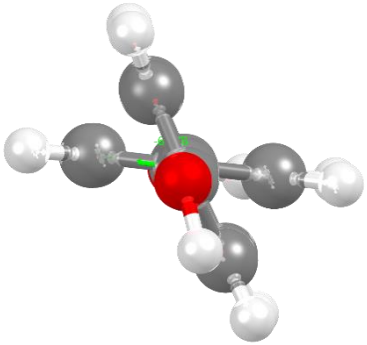
The analysis performed with Mogul and ConQuest raises two questions that you might want to consider if a similar situation arises in your research while analysing your structure:

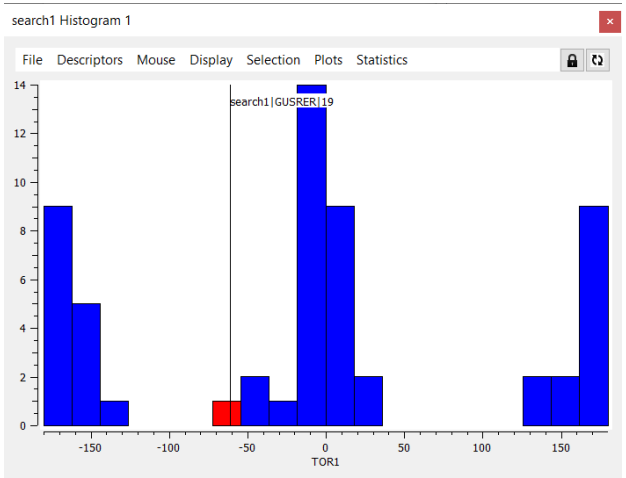
1. Is the refinement correct or is there anything that needs double-checking?
2. What is the reason driving such unusual geometry? E.g., a strong hydrogen bond network.

Hint!

When you export ConQuest results for analysis in Mercury, you can include more parameters. You can see the options available at **Step 30**. Including such information, if relevant, could complement the Mogul and ConQuest analysis performed and provide more insights.



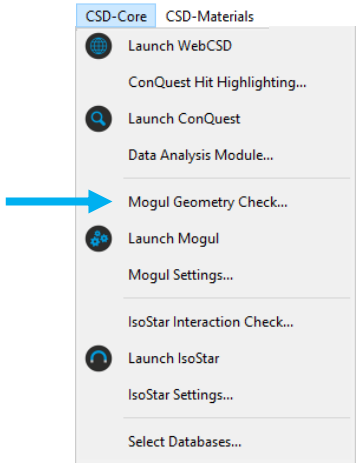




Summary

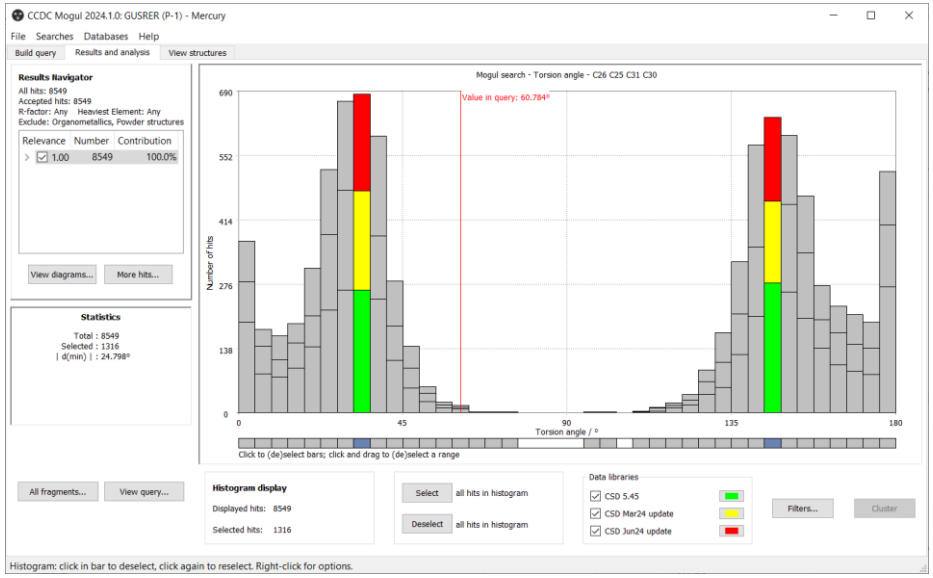
After this workshop you will be able to perform an in-depth analysis of the geometry of a molecule in a structure of interest using Mogul in Mercury, with particular focus on torsion angles. You should now know how to:

- Apply filters for the Mogul Geometry Check search.
- Select and deselect specific bars of the histogram to restrict the visualisation of the results to specific parts of the histogram.
- Look for patterns and similarities in structures contributing to the same area of Mogul results.
- Include or exclude structures contributing to the Mogul results based on their relevance.
- Complement Mogul Geometry Check with ConQuest 3D search and Mercury analysis.



Next Steps

After this workshop, you can explore more exercises in the self-guided workshops available in the CSD-Materials workshops area on our website. We suggest trying the Full Interactions Maps workshop, which complement the Mogul geometry Check in the assessing of stability.
https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/csd-materials-workshops/



Feedback

We hope this workshop improved your understanding of In-depth analysis of geometry in Mogul 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 MOG-002. Thank you!

Glossary

Conformation

The shape, or more precisely, the spatial arrangement of a molecule is called conformation. The same molecule can present different conformations (different shapes).

Fragment

Fragment is the generic word used in Mogul for a bond, valence angle or torsion.

Hydrogen Bonds

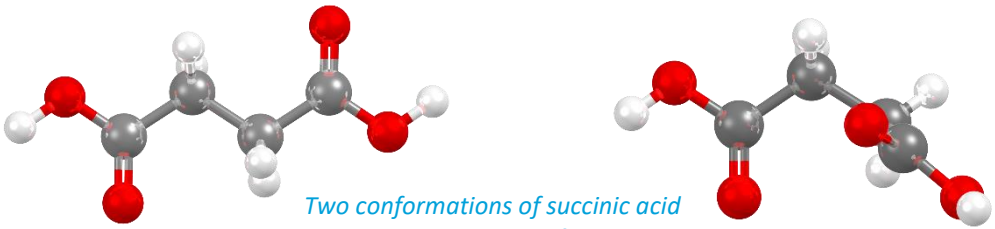
Hydrogen Bonding occurs between donor-acceptor interactions precisely involving hydrogen atoms. The H-bonds interactions are classified as: strong (mostly covalent), moderate (mostly electrostatic) and weak (electrostatic).
Their strength is observed to be between 12 and 30 kJ/mol.

In silico

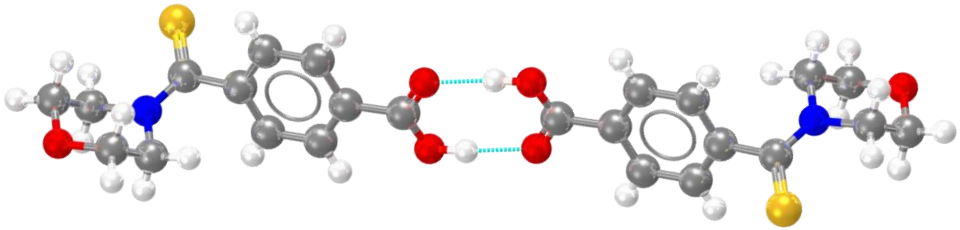
The expression *in silico* is used to refer to values that are generated via computer calculations or simulations.

Torsion Angles

Torsion angles are used to describe conformations around rotatable bonds. The torsion angle between 4 atoms A-B-C-D is the angle by which the vector A-B must be rotated in order to eclipse the vector C-D when viewed along the vector B-C.



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



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