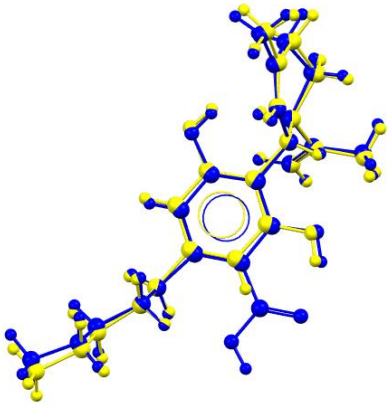
Overlaying Molecules in Mercury (MER-008)

Developed using
2025.2 CSD Release
(CSD 6.00 + 1 data update)



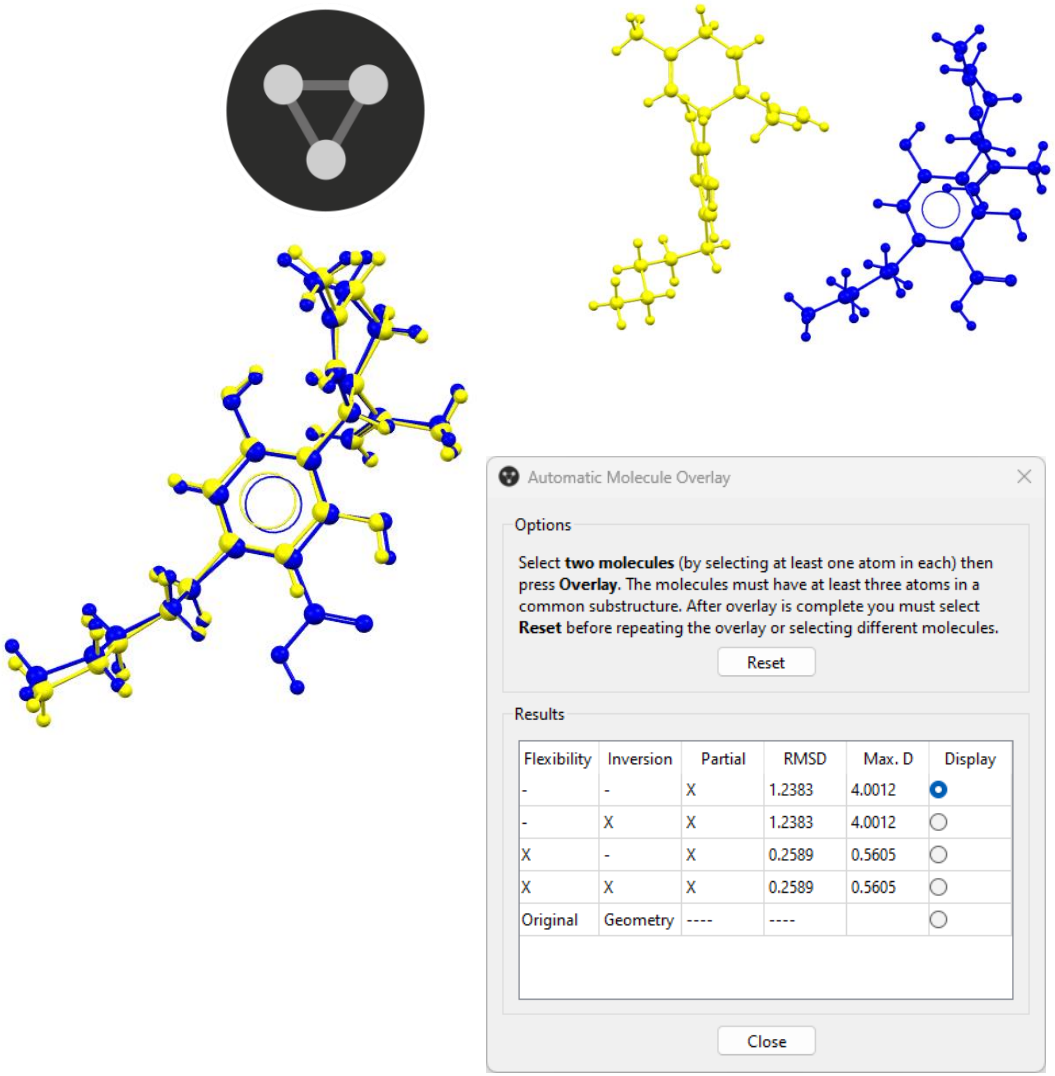
CCDC
advancing structural science

Table of Contents

Introduction

Mercury is the visualisation and analysis software of the Cambridge Structural Database (CSD). Mercury has numerous features for structure comparison, including tools for geometrical analysis as well as for assessment of intermolecular interactions. In this workshop we will explore one of the tools for direct structure comparison: Molecule Overlay. Molecule Overlay performs least squares superposition of a pair of molecules, which may be from the same or different structures. The tool can compare both chemically identical and non-identical molecules (provided they share a common substructure). In addition to the direct overlay, the calculation also applies inversion and rotation of flexible torsions before overlay to generate different solutions.

In this workshop, we will see how to apply Molecule Overlay in Mercury and how to interpret the output.



Learning Outcomes

After completing this workshop, you will be able to:

- Overlay molecules in Mercury from within the same structure
- Overlay molecules from different structures in Mercury

This workshop will take approximately ***25 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

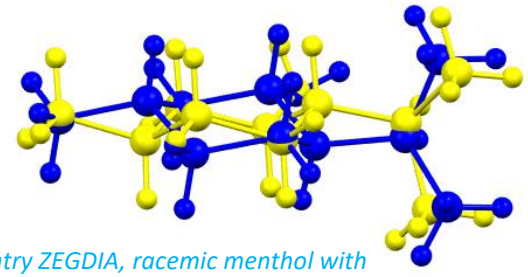
Basic familiarity with Mercury is required. You can find a brief summary guide at the end of the handout.

Materials

No additional materials are required.

Example 1. Overlaying Stereoisomers in Mercury

In this example, we will attempt to overlay stereoisomers of menthol using the Molecule Overlay tool in Mercury to see the effect of inversion and flexibility on the overlay.



CSD entry ZEGDIA, racemic menthol with two stereoisomers superimposed.

---

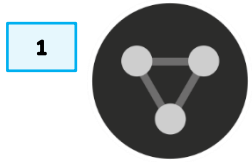
1. Begin by opening Mercury from the Start menu or by clicking on the desktop icon.

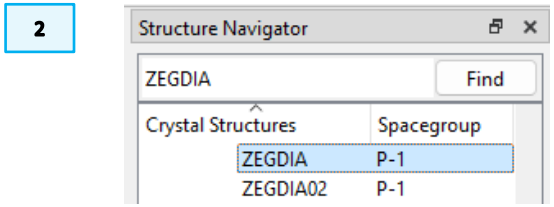
2. We need to select an appropriate structure. In the *Structure Navigator*, type *refcode* "ZEGDIA". This is racemic menthol.

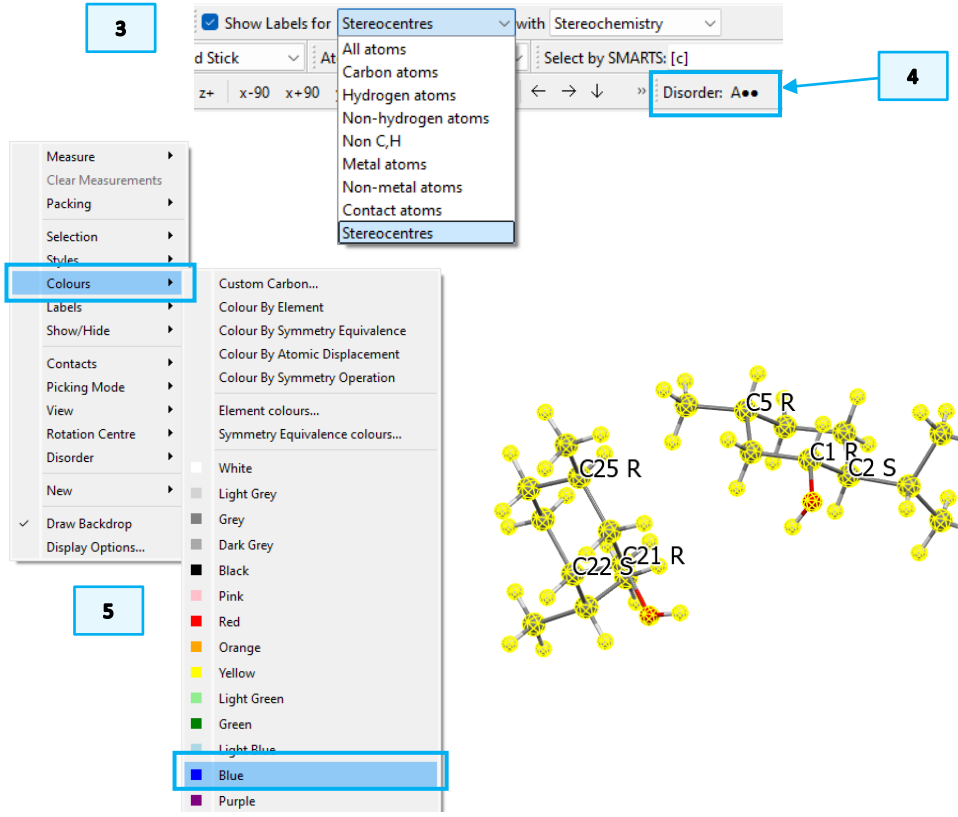
3. There are three molecules in the asymmetric unit of ZEGDIA, two of which are (−)-menthol ((1*R*,2*S*,5*R*)-2-isopropyl-5-methylcyclohexanol) and third is the enantiomer (+)-menthol ((1*S*,2*R*,5*S*)-2-isopropyl-5-methylcyclohexanol). To see which is which, tick the *Show Labels for* box in the top toolbar and from the drop-down menus pick **Stereocentres** for the first and **Stereochemistry** for the second.

4. There is *disorder* of the alcohol hydrogen atoms in this structure. It will not affect the analysis, but you can remove it from the display by clicking the two black dots next to *Disorder* in the top toolbar.

5. Study the structures to identify the two stereoisomers. To make them easier to identify, you might wish to colour them differently. We will colour (−)-menthol blue and (+)-menthol yellow for clarity. To select a complete molecule, hold down the Shift key and left-click on the molecule of interest. Then right-click in the visualiser area and from the drop-down menu, select your preferred colour. You can now turn off the stereochemistry labels if you wish.



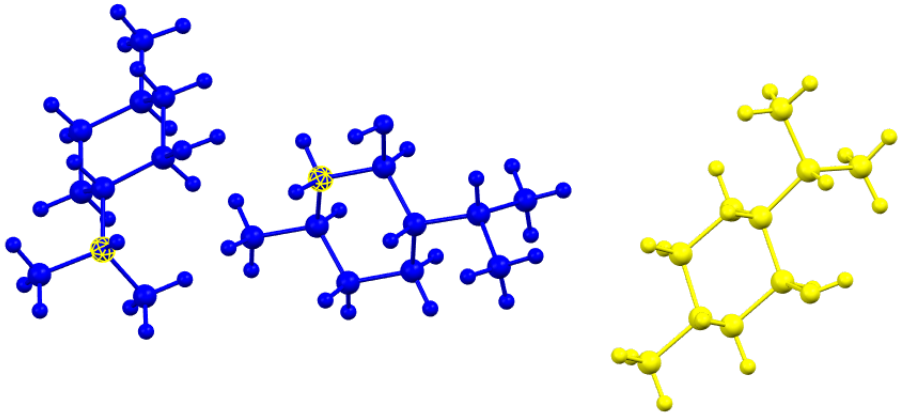




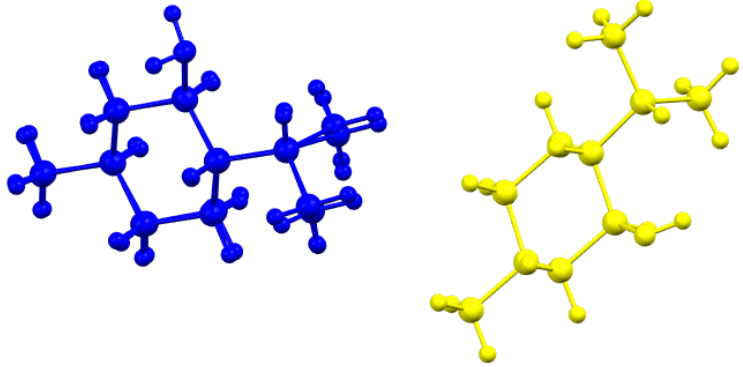
6. From the Calculate menu, click **Molecule Overlay…** to launch the *Automatic Molecule Overlay* dialogue. There is some useful information in the Options section of the dialogue:
   - Two molecules must be selected
   - The molecules must have at least three atoms in a common *substructure*
   - You need to press **Reset** before analysing different pairs of molecules

   Note: you do not need to select equivalent atoms in the two molecules.

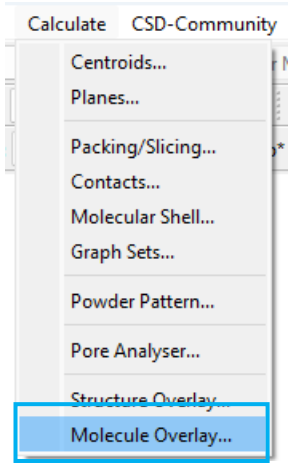
7. Click on an atom in each of the (−)-menthol structures (blue in the image below, following step 5) and click **Overlay**.

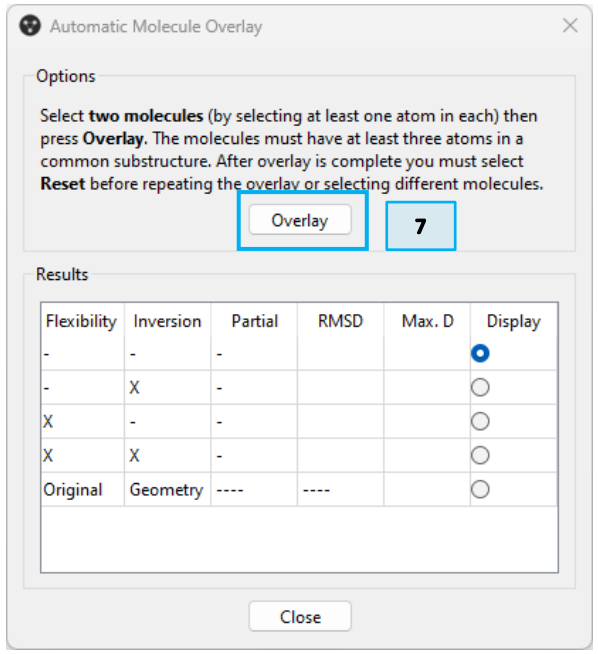


8. The molecules will be overlaid. Since the (−)-menthol molecules are conformationally almost identical, the initial overlay looks very good.

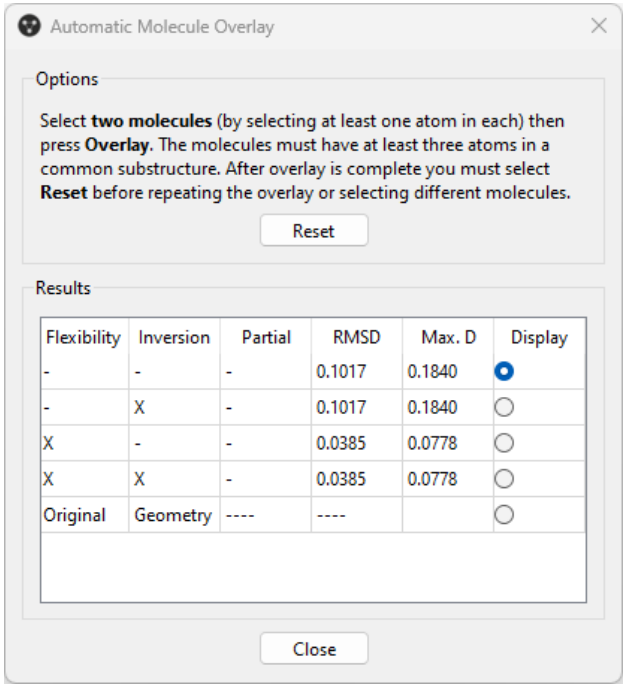


6





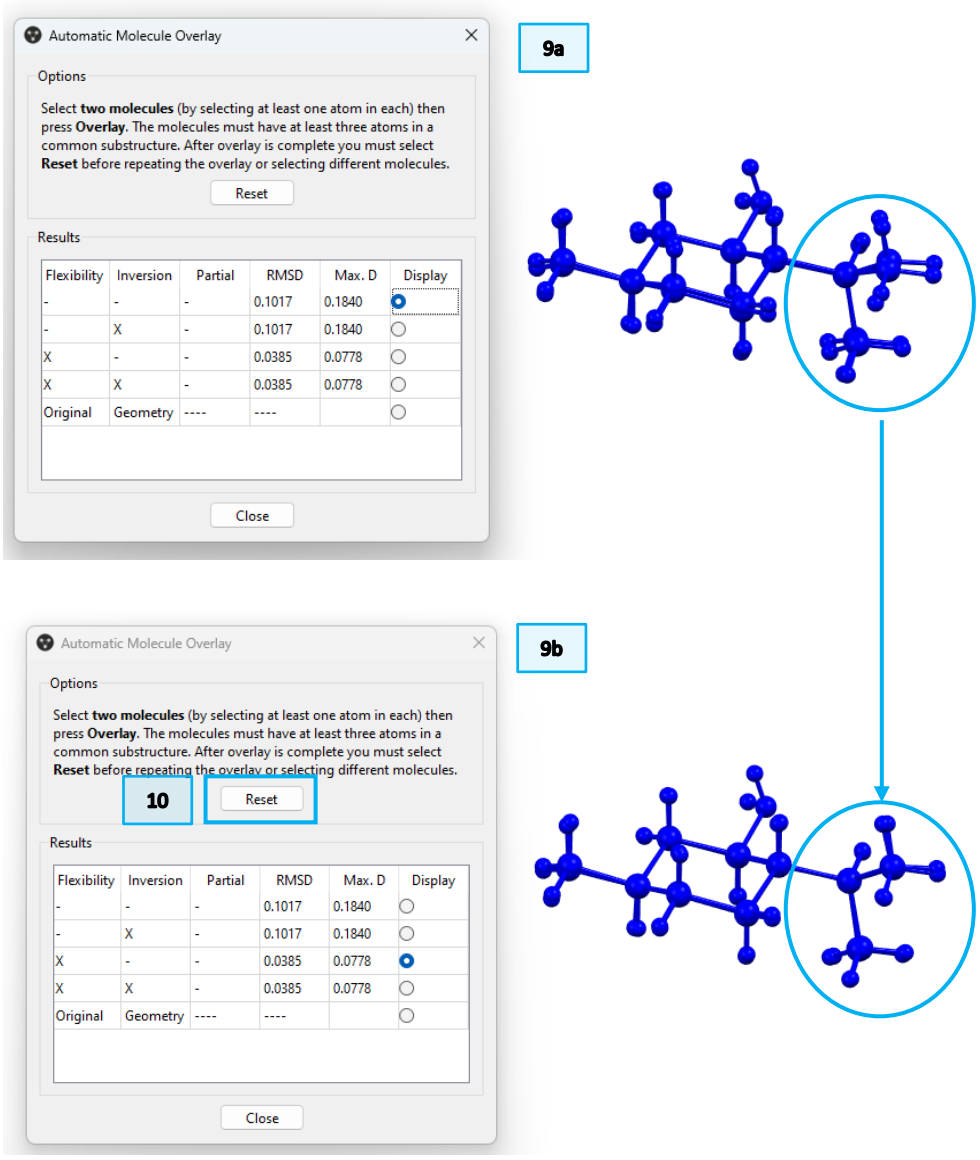
8



9. Examine the information in the *Results* section of the *Automatic Molecule Overlay* dialogue.
   - The *Flexibility* and *Inversion* columns indicate whether the respective operation have been applied by an "X" if they have or "-" if not. If inversion does not improve the overlay of the two molecules, it is not applied and the RMSD reported is of the non-inverted overlay. This is the case for the current pair of molecules (we already know they have the same stereochemistry; therefore, inversion would make the overlay worse and is not applied).
   - The *Partial* column is applicable if the molecules are not chemically identical and successful overlay of a common substructure is indicated by an "X" in this column.
   - The *RMSD* and *Max. D* columns indicate the root mean squared deviation of the least squares fit and the maximum distance found between pairs of equivalent atoms, respectively. Lower numbers here generally indicate a better overlay; however, it is possible to have a low RMSD and high Max. D if the molecules have very similar structure except for in one small region.
   - The *Display* column allows you to visualise the superposition corresponding to the different overlay calculation result by clicking the radio button for the desired row in the *Results* table. The final row of the table allows you to switch to the non-overlaid view without resetting the calculation.

   In the current results, you can see that the initial overlay (9a) is very good and is further improved by rotating flexible torsions (9b).

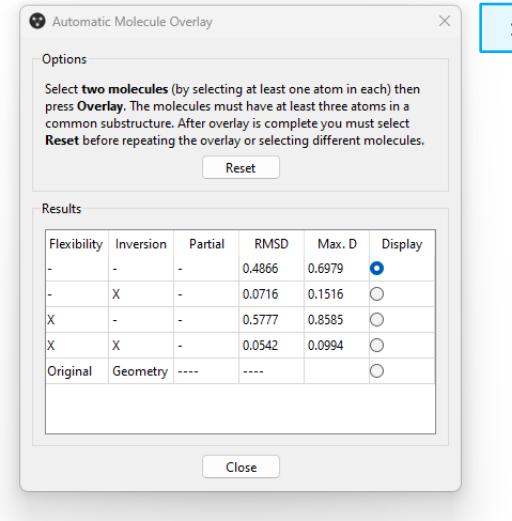
10. Click Reset in the *Automatic Molecule Overlay* dialogue.



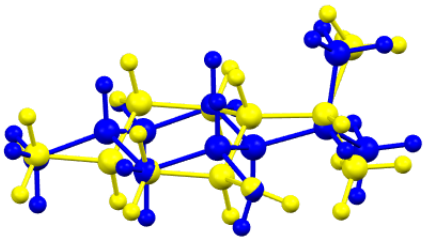
11. Repeat the overlay procedure for different enantiomers of menthol by picking an atom from one of the blue molecules and an atom from the yellow one (assuming you have used the suggested colour scheme) and clicking **Overlay**. The results you get will be slightly different depending on which of the two possible pairs you choose. In this case, the direct overlay (11a) has a poor fit. This is expected – enantiomers of this kind can never be superimposed. With inversion, the fit is greatly improved and rotation on top of inversion gives an incremental improvement (11b).

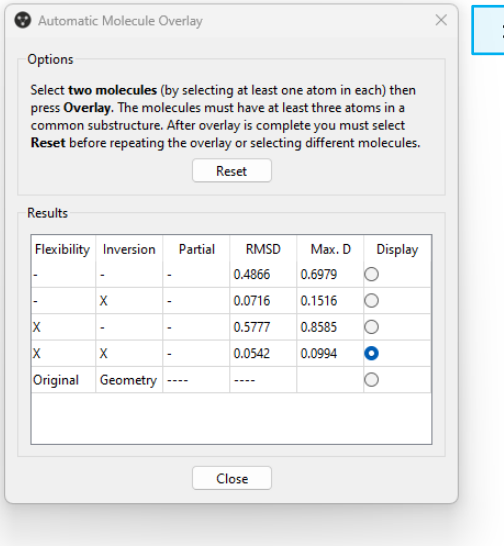
Conclusion

In this example we have primarily seen how the inversion feature of the molecule overlay works. In so doing, we have established that the conformations of the two enantiomers of menthol in CSD entry ZEGDIA are virtually identical once stereochemistry is accounted for.

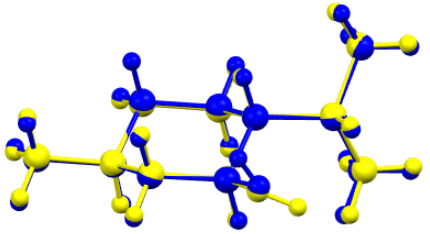


11a



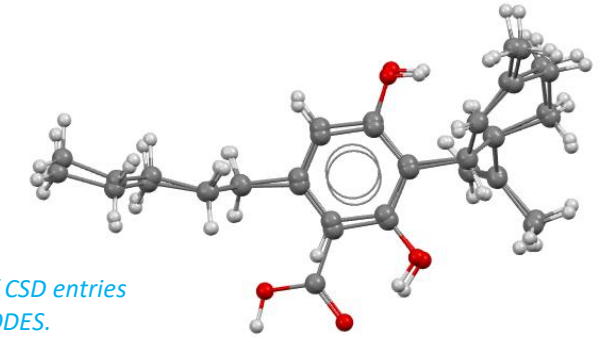


11b



Example 2. Overlaying Non-Identical Molecules

In this example, we will look at two related but non-identical molecules in different structures, cannabidiol and cannabidiolic acid. Cannabidiolic acid has disorder in its crystal structure, and we shall see how selection of disorder components can be achieved via the Multiple Structures tool.



The superposition of CSD entries CANDOM14 and IHODES.

1. If Mercury is not already open from Example 1, open it from the Start menu or by clicking on the desktop icon.

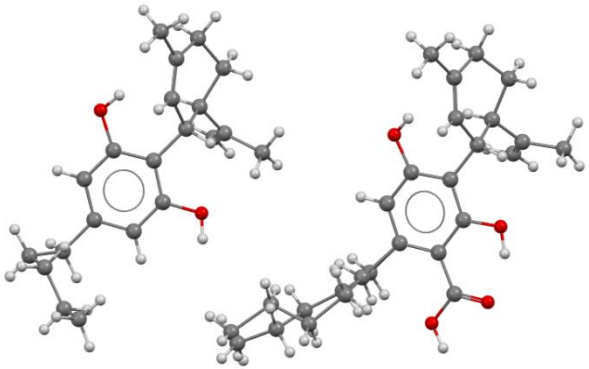
2. In the *Structure Navigator*, type "CANDOM14" to load a structure of cannabidiol (there are two polymorphs, we only use one of these).

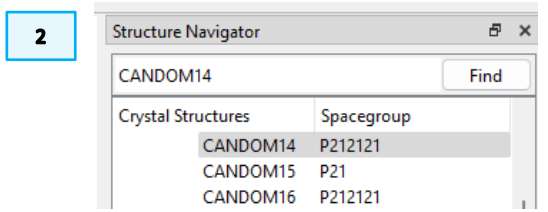
3. Tick **Multiple Structures** underneath the *Structure Navigator*.

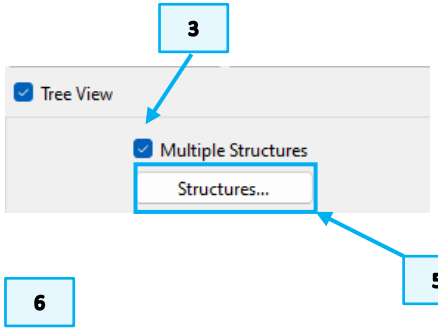
4. Type "IHODES" in the *Structure Navigator* to load in the second structure, which is cannibidiolic acid.

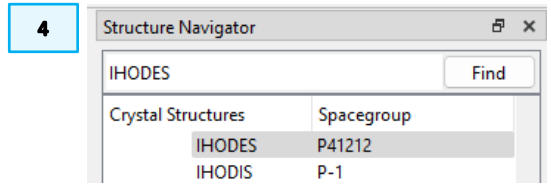
5. Click **Structures** to open the *Multiple Structures* dialogue.

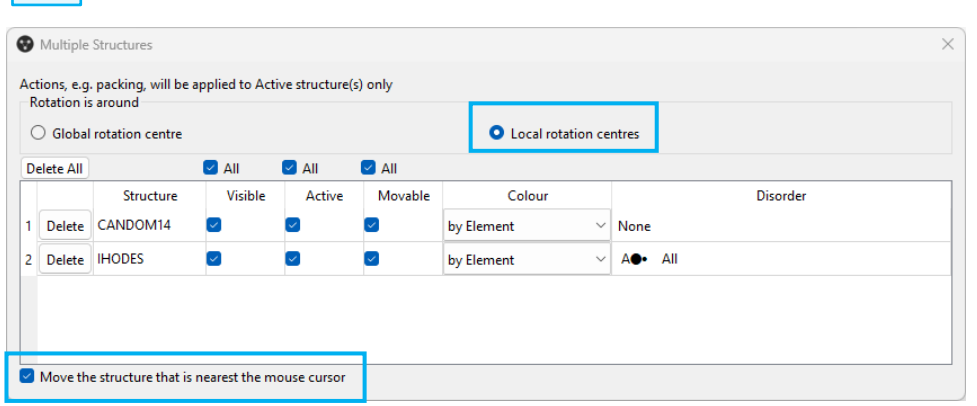
6. In the *Multiple Structures* dialogue, tick the box next to **Move the structure that is nearest the mouse cursor** and click the **Local rotation centres** radio button. You should now be able to reposition the structures in the visualiser area more easily, using the regular mouse/keyboard operations.



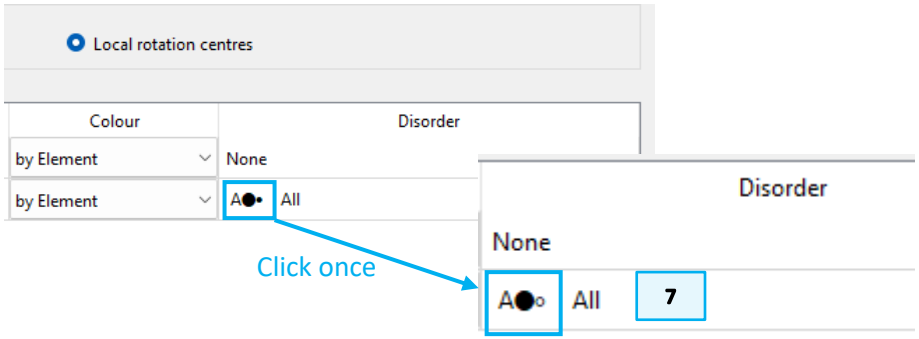




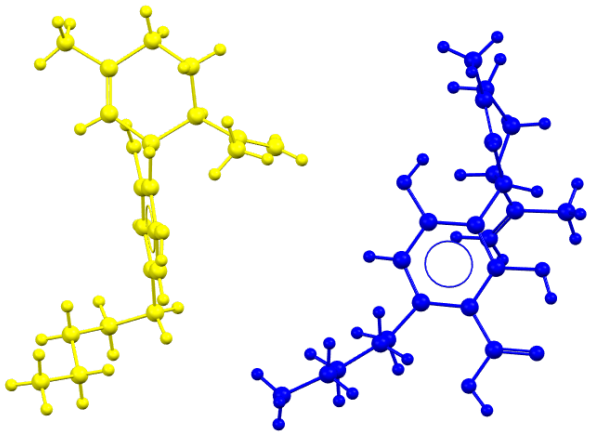


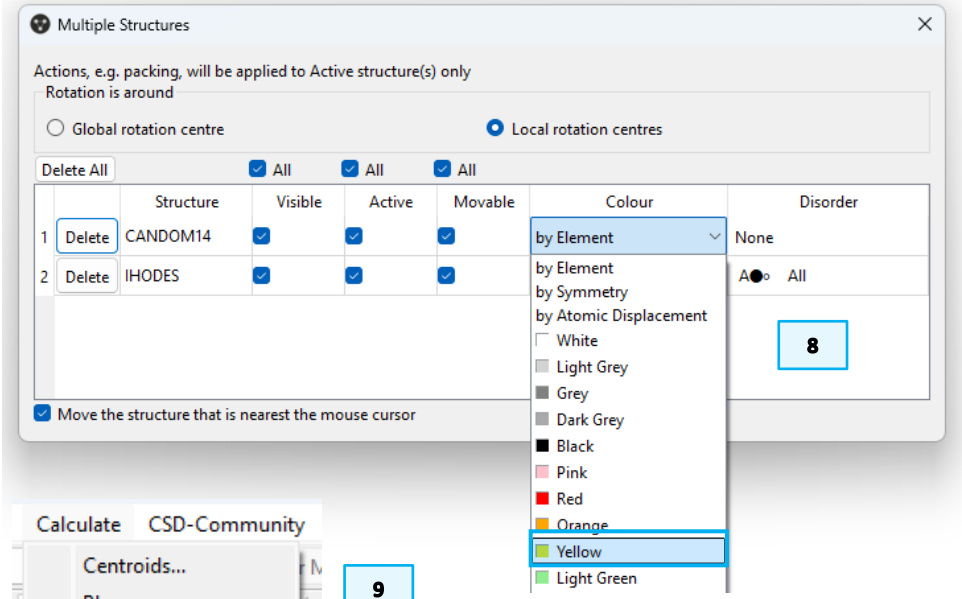


7. You will notice similarities and differences between the two molecules. Chemically, they differ only in the replacement of a hydrogen *ortho* to the pentyl group by a carboxylic acid group. On the other hand, the *conformation* of the pentyl group in cannabidiol (CANDOM14) is clearly quite different to that in cannabidiolic acid (IHODES), and in any case, the pentyl group is disordered in IHODES. In the Disorder column of the *Multiple Structures* dialogue, click the two black dots next to **A** once. This will select the major disorder group. You can subsequently click it to switch back and forth from major to minor disorder group. The larger dot indicates the major occupancy group. If a disorder group is activated, the indicator for that disorder group is filled in black, otherwise it is empty. If you want to view both disorder groups together again, click the **All** button.

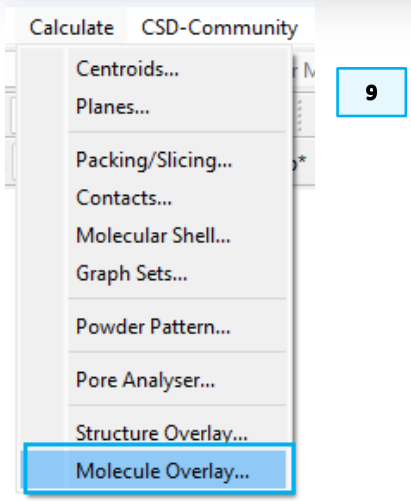


8. For ease of visualisation, we recommend colouring the two molecules differently. This can be achieved by clicking the drop-down menu in the *Colour* column of the *Multiple Structures* dialogue. In this case, we shall choose yellow for CANDOM14 and blue for IHODES.





9. From the *Calculate* menu in the top menus, click **Molecule Overlay** to open the *Automatic Molecule Overlay* dialogue. See step 9 of Example 1 for an explanation of the interface.

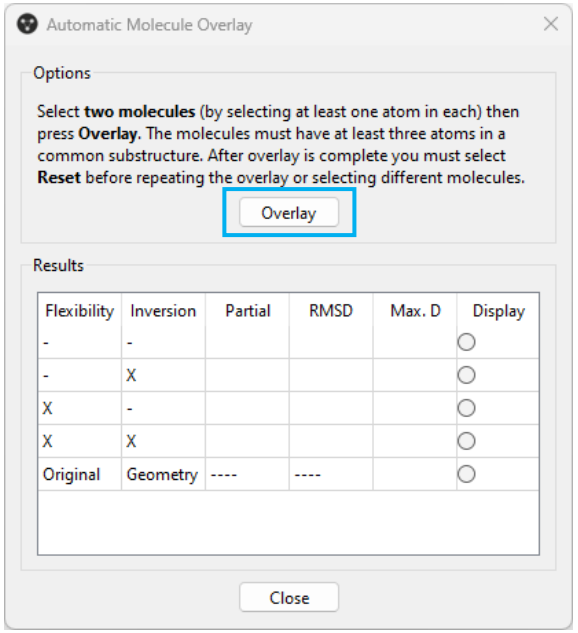


10. Click one atom of each molecule, then click **Overlay** to overlay the molecules. Caution: the results of a molecule overlay with *flexibility* can be affected slightly by which molecule is selected first. We have selected IHODES first.

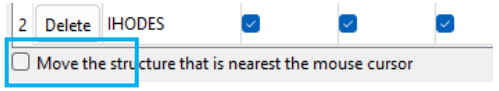
11. When the overlay is complete, inspect the results reported in the dialogue. Switch between different solutions using the radio button in the *Display* column. Notice that in the *Partial* column, there is an "X" for each overlay solution because the molecules differ by a carboxylate group and therefore a common substructure is being used for the overlay. The initial overlay is poor (RMSD = 1.1320, see 11a). Inversion will not improve the overlay because both molecules have the same stereochemistry at the equivalent chiral atoms in the pair of molecules. Therefore, inversion is not applied in the calculation and the metrics reported are as for the first solution (no inversion, no flexibility). On the other hand, the result with rotation of flexible torsion is a greatly improved overlay (RMSD = 0.1087, see 11b), which we can see is largely due to achieving the same conformation of the pentyl group by rotation around the C–C bonds. Tip: if you want to rotate the view of the molecules, untick **Move the structure that is nearest the mouse cursor** in the *Multiple Structures* dialogue.

12. Click **Reset** in the *Automatic Molecule Overlay* dialogue.

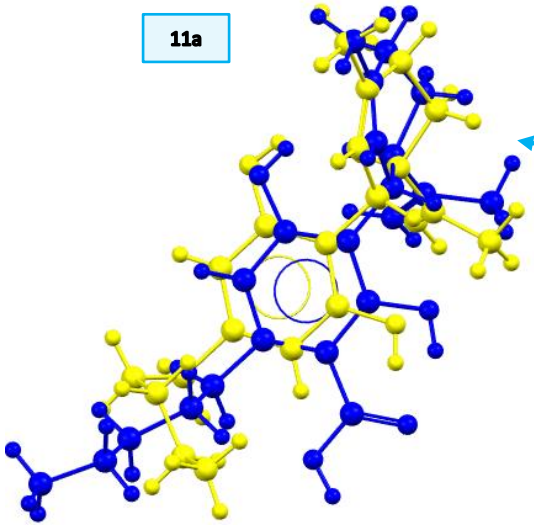
10



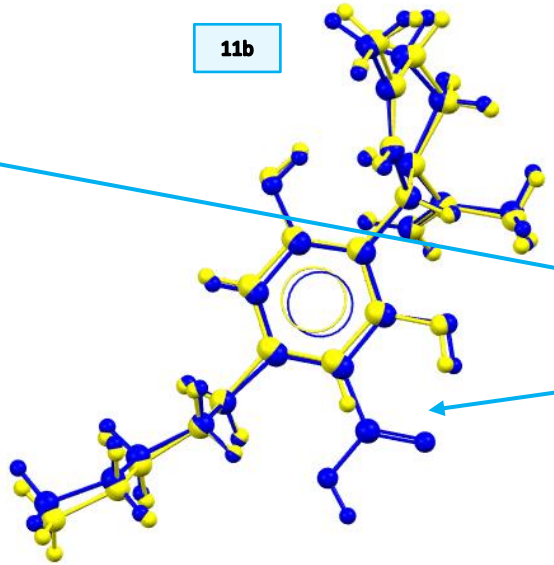
Tip:



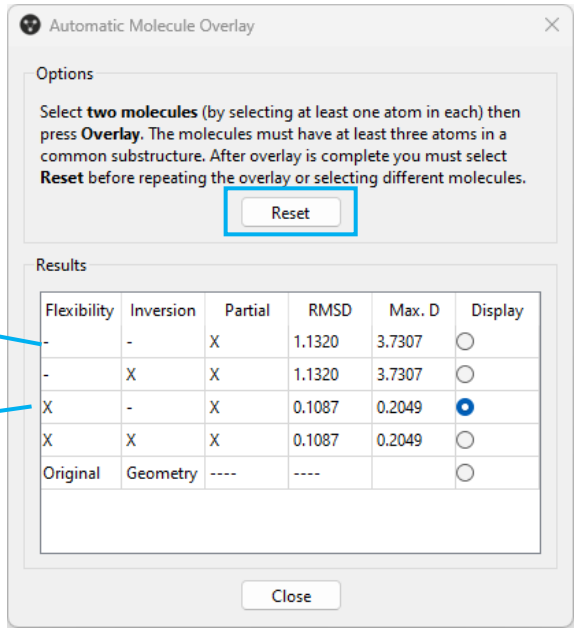
11a



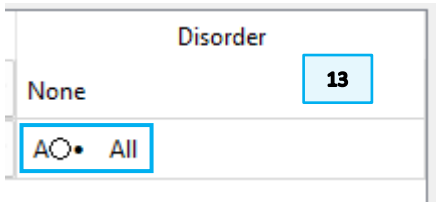
11b



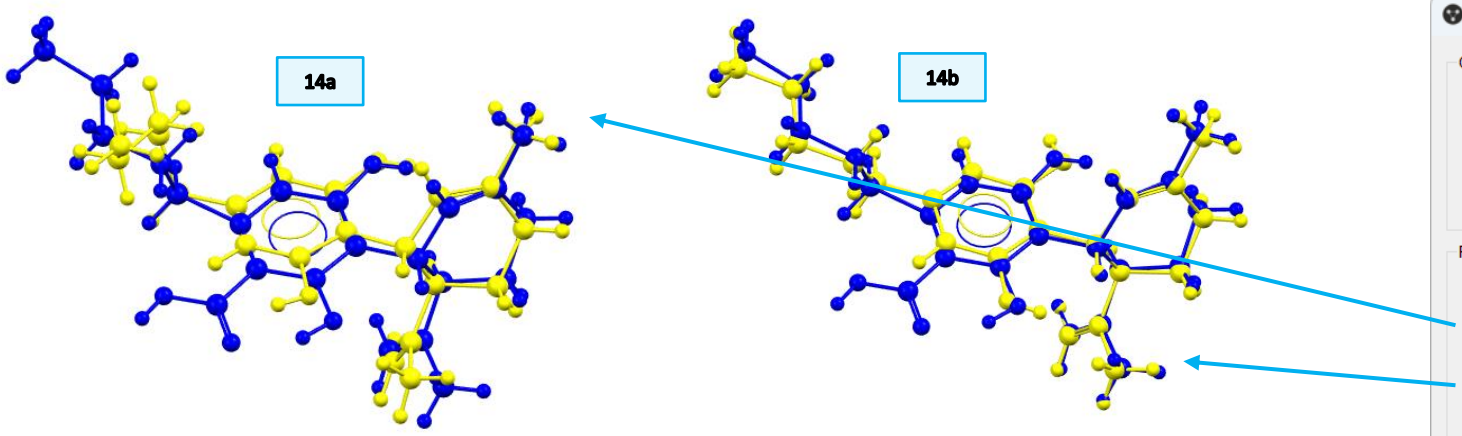
12

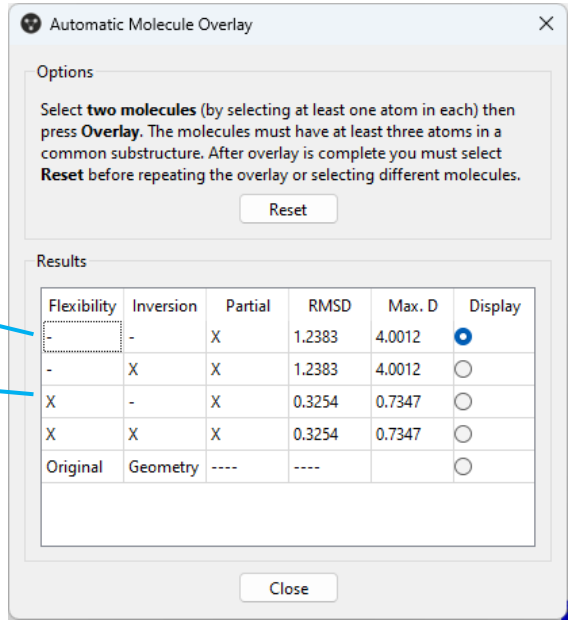


13. In the *Multiple Structures* dialogue, in the *Disorder* column, click the two dots next to "A" to activate the minor occupancy disorder group. The colour of the affected carbon atoms may revert to grey. This does not affect the overlay. You can change the colour of these atoms to your preferred colour again, as in step 8, if you wish.



14. Repeat steps 10-11. You should find that the metrics are worse for the overlays involving the minor occupancy disorder group, though rotation of torsions still markedly improves the results.





Conclusions

In this example we have seen that good overlays can be achieved even in non-identical molecules and that molecules with flexible groups such as alkyl chains, which can easily adopt different conformations, can be overlaid well when the conformation is allowed to change via rotation of flexible torsions.

Summary

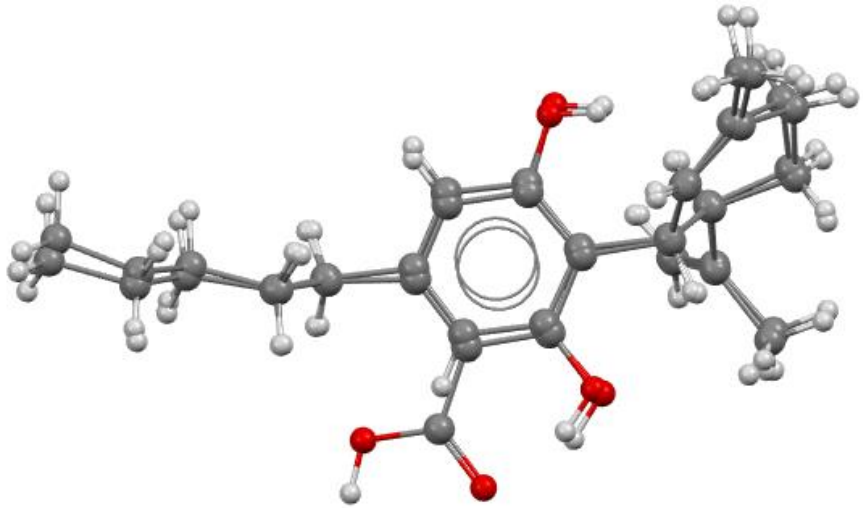
In this workshop, we have seen how to use the Molecule Overlay functionality in Mercury to perform a least-squares superposition of molecules. You should now:

- Be able to superimpose molecules using the Automatic Molecule Overlay tool in Mercury
- Be comfortable overlaying molecules from the same structures and from different structures by making use of the Multiple Structures tool
- Understand the results of a molecule overlay calculation

For your reference, you can consult the Mercury user guide.

Next Steps

If you are interested in overlaying structures, you might like to explore the example in In-Depth Comparison of Polymorphic Structures Using Mercury. You can find self-guided workshops on other tools in Mercury in the CSD-Core and CSD-Materials self-guided workshop pages.



Feedback

We hope this workshop improved your understanding of Molecule Overlay 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 MER-008. It will only take 5 minutes and your feedback is anonymous. Thank you!

Glossary

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

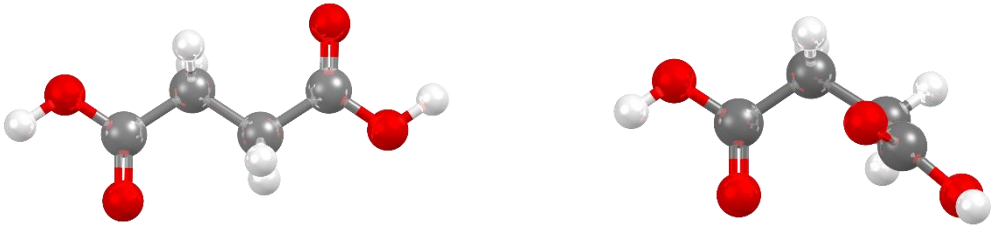
Disorder
A disordered structure lacks long range order, but it may show local order, which can be modelled by the crystallographer during a structure refinement. Disorder in the CSD is described with disorder assemblies and groups. A disorder assembly is a cluster of atoms that show long-range positional disorder but are locally ordered. Within each assembly a disorder group is used to identify sites that are simultaneously occupied (i.e. the atoms in the group all have the same occupancy).

Flexibility
Molecules with rotatable single bonds are typically flexible – they can easily adopt different conformations. Some single bonds can show limited rotation due to steric hindrance.

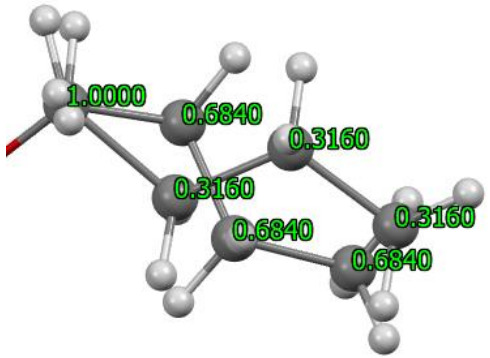
Inversion
Inversion maps coordinates (x, y, z) to (−x, −y, −z). This has the effect of swapping *R* and *S* configuration at stereocentres of a molecule.

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

Root Mean Square Deviation (RMSD)
The root mean square deviation (RMSD) is a commonly used measure of the difference between two sets of values. In the case of molecule overlay, this difference is the distances between equivalent pairs of atoms in two molecules.



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



A disordered butyl group with occupancies of the carbon atoms show.

The formula for the RMSD of two sets of coordinates after optimal superposition is:

$$\mathrm{RMSD}\left(\mathbf{x}, \mathbf{x}^{\mathrm{ref}}\right) = \min_{\mathbf{R},\mathbf{t}} \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left\| \left(\mathbf{R} \cdot \mathbf{x}_i(t) + \mathbf{t}\right) - \mathbf{x}_i^{\mathrm{ref}} \right\|^2}$$

Where $\mathbf{x}_i$ and $\mathbf{x}_i^{\mathrm{ref}}$ are the atomic coordinates of the pairs of atoms to be overlaid, **t** and **R** are the optimised translational vector and rotation matrix.

Substructure

A substructure of a structure is a specific smaller part of a structure. It does not need to be a complete molecule. For example, the C_6 ring of benzene is a substructure of both benzene and phenol but is not a complete molecule.

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