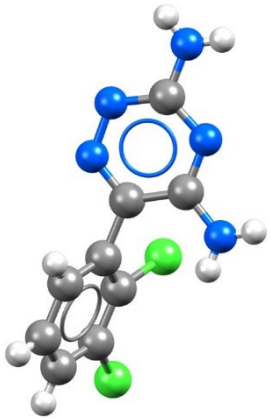
Introduction to IsoStar
(ISO-001)

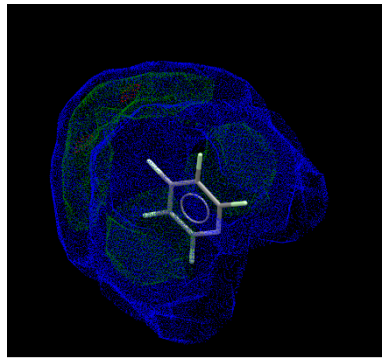
Developed using
2022.3 CSD Release
CSD version 5.43 (November 2021) + 4 data updates

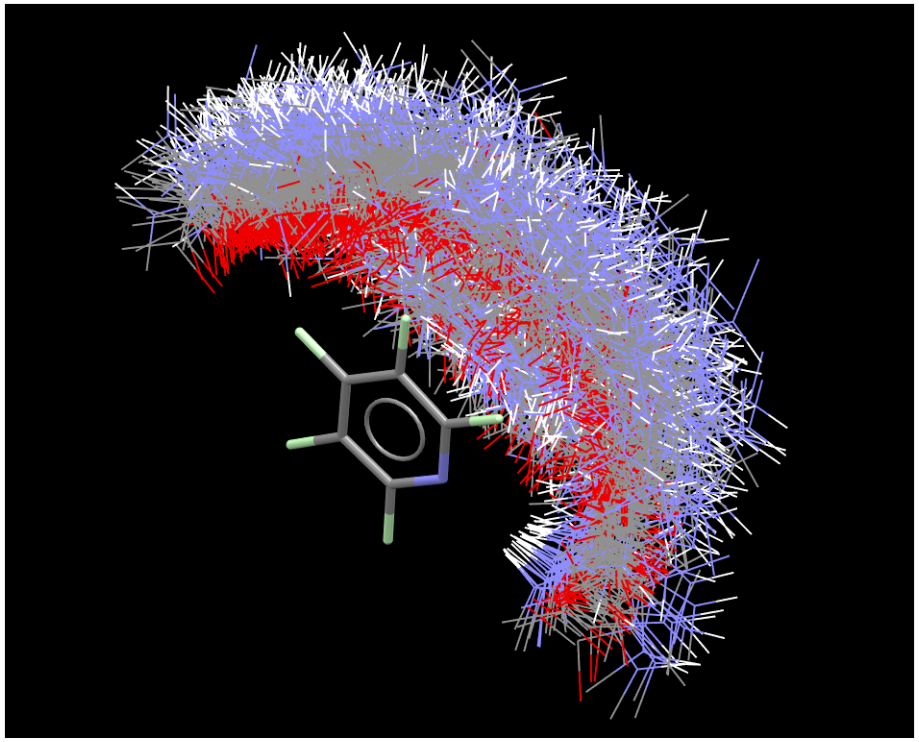


CCDC
advancing structural science

Table of Contents

Introduction





Entries in the CSD and PDB contain thousands of interactions, between all kinds of different functional groups. Extracting these out and mapping the space of preferred interactions around functional groups can lend to predictive properties of query molecules. This can help with ligand design, understanding of interactions, identifying bioisoteres, and many other components along the drug discovery pipeline.

IsoStar is a library of information about the intermolecular interactions formed by a wide variety of chemical groups. It contains data on intermolecular interactions in small molecule crystal structures taken from the CSD, as well as protein-ligand interactions in X-ray structures from the PDB. The library is presented in a web application for download based on the query molecule and *contact group* type. These files are then viewed and analysed in the IsoStar client, which is included with the CSD Suite. You can find video resources on how to use IsoStar here.

Learning Outcomes

After completing this workshop, you will be able to:

- Navigate, understand, and download data from the IsoStar website.
- View and manipulate data in the IsoStar desktop client.
- Understand IsoStar scatterplots and generate contour surfaces for interactions.

This workshop will take approximately ***30 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

There are no pre-required skills for this workshop.

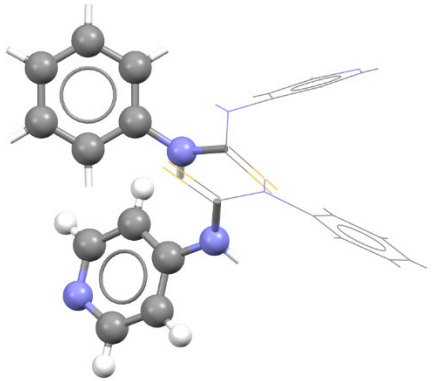
Materials

No extra materials are required to complete this workshop.

Example 1. IsoStar

Crystallographic information in IsoStar is presented as 3D scatterplots, or contoured density surfaces. Each data point in these scatterplots has been calculated by searching the CSD or PDB for nonbonded interactions between a pair of functional groups A and B, and then least-square superimposed.

This example will walk you through downloading information from the IsoStar server, and then using it to compare the interaction preferences of an organic molecule in the CSD vs. the PDB. The context of each of these is quite different, even though the query molecule is the same.



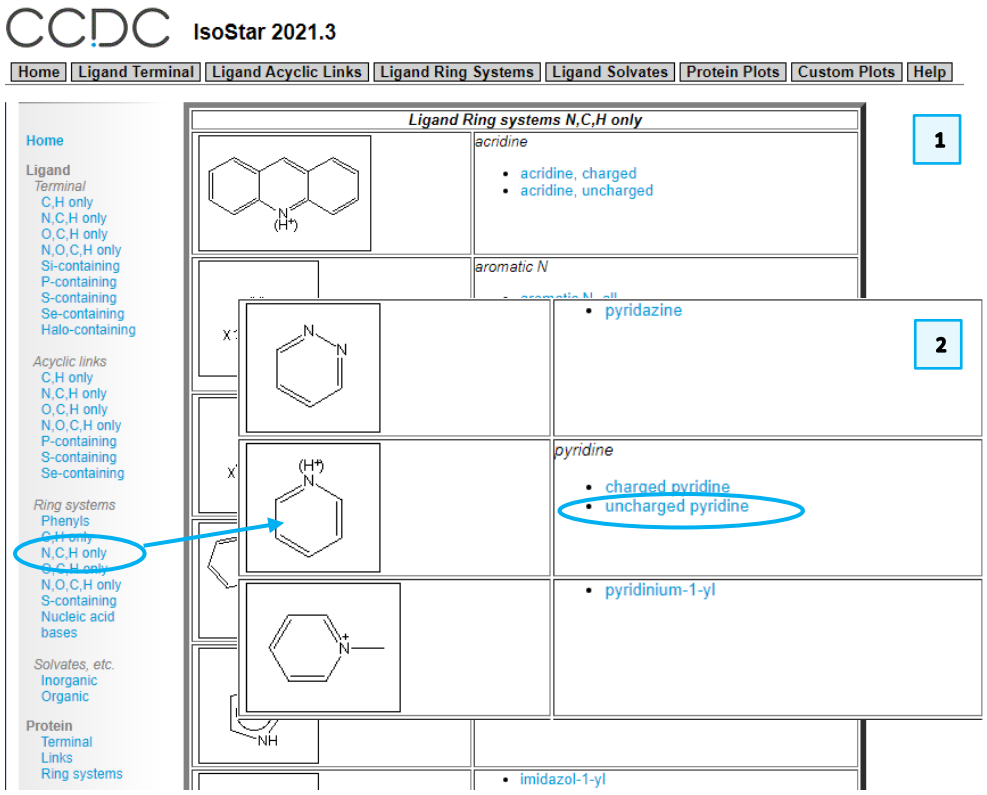
Example of an intermolecular interaction between an uncharged pyridine and a phenyl group in N-(4-pyridyl)-N'-phenylthiourea, refcode TIRMUD

This example will focus on the interactions of an uncharged pyridine molecule with phenyl groups.

1. On the IsoStar webpage (isostar.ccdc.cam.ac.uk), under *Ring systems,* click *N,C,H only.*

2. About halfway down the page, click *uncharged pyridine.* This takes you to a page presenting the interacting groups for uncharged pyridines. Note that the tables are broken into element-based functional groups, with guides at the top of each table (databases are updated frequently so the numbers can be different).

While most entries have statistics associated with them, some also have entries in the *Theory* column. These are a series of calculations for various energies contributions, all done at the 6-31G** basis set level using CADPAC. More information can be found by following the *Theory* link in the tables and clicking on *Methodology*.



3. In the **C,H only** table, click the entry for *phenyl*, under *CSD*. This downloads an IsoStar file (.istr).

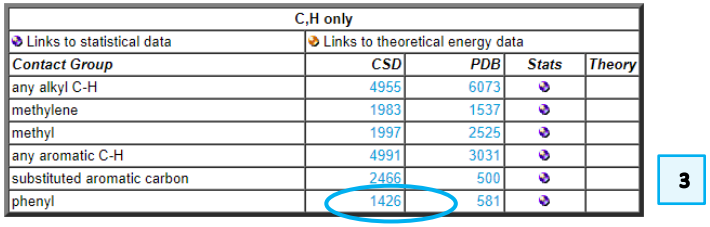
4. Click the downloaded IsoStar file to launch the IsoStar client. The IsoStar window has visualizers at the bottom of the screen, which by default shows all entries. Click the **vdW Overlaps** button to hide all contact groups except for those within van der Waals contact of any target atom in the central group.

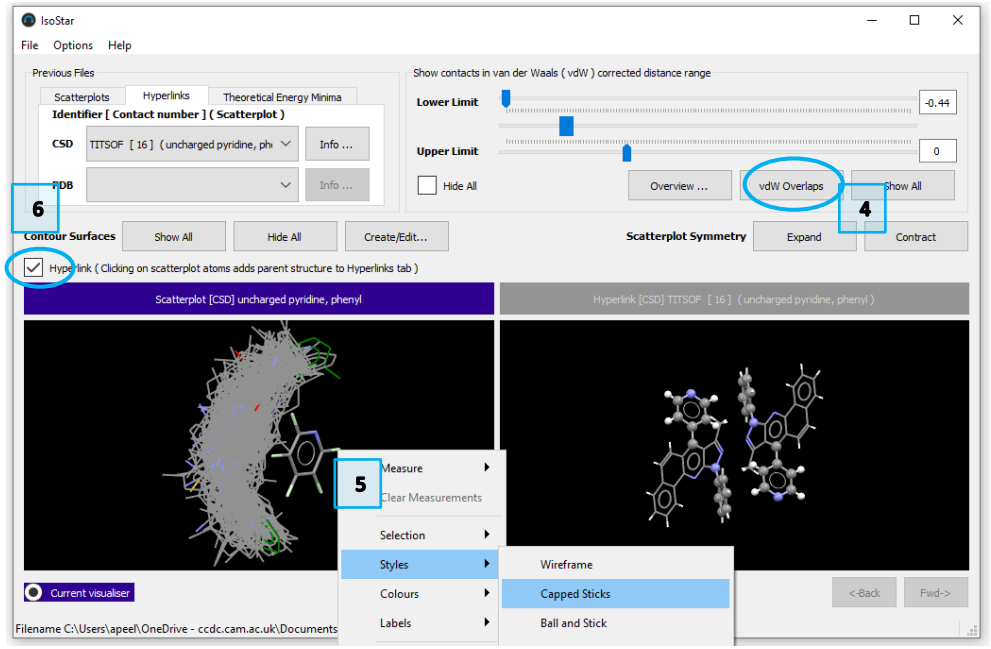
5. The visualizer in IsoStar is the same as Mercury, so the functionality is the same. Take a moment to rotate and zoom the scatterplot to get an appreciation for the π-π stacking and edge interactions around the uncharged pyridine. You can change the display style of pyridine to capped sticks for clarity by right clicking *Styles >> Capped Sticks*.

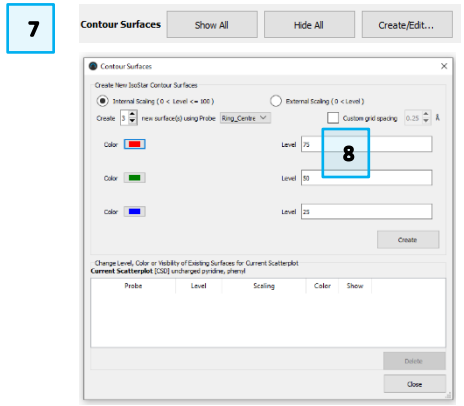
6. The scatterplot is made of 1426 entries at time of publication. While this lends vital information about interaction preferences, it is important to be able to drill down into each case to see the interaction. To enable this ability, click the tick box next to *Hyperlink*. Now click on any interaction in the scatterplot to see the individual CSD entry in the previously empty quadrant.

7. Scatterplots can be converted to contoured density surfaces, which provides richer information that looking at the *probes* overlaid. Click the **Create/Edit** button next to *Contour Surfaces* to launch the **Contour Surfaces** window.

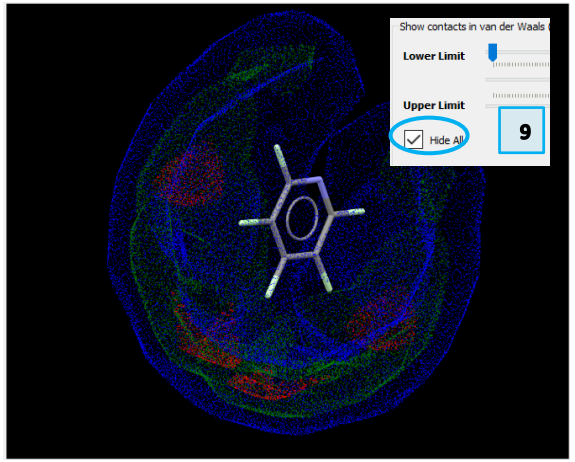
8. Note that the *Probe* is by default set to *Ring_Centre*. This means that the centroid of a ring will be taken as the representative point. Click **Create** to create the surface. The default values create a density surface with red being the highest concentration, to blue, and then green for lower threshold. Each of these thresholds can be edited. Click **Close** to close the window.



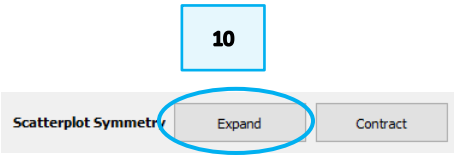




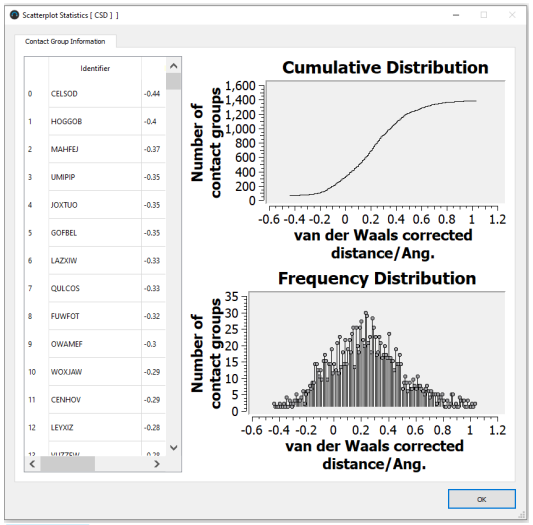
9. The new contour surface is showed overlaid with all of the probes. Hide the probes by clicking the *Hide All* tick box in the **Show contacts in *van der Waals (vdW) corrected distance*** **range** box. Inspect the contour surface. Note the hotspots (red areas) and the uneven distribution of them around the ring. There is also a higher density of probes underneath the aromatic ring, for the π-π stacking interactions.

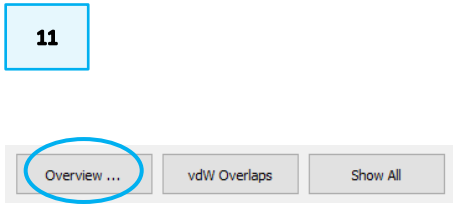


10. In IsoStar, scatterplots are symmetrised. This means that all contacts are reflected into an area of space based on the symmetry of the central group (e.g. only one quarter of an aromatic iodine group with its contact group is shown since the central group has two planes of mirror symmetry, whereas half of all contacts to an azide group are shown since it has one plane of mirror symmetry). The symmetry can be expanded and contracted as desired by clicking on **Expand** and **Contract** next to *Scatterplot Symmetry*. Expand the symmetry. While this does show <u>all</u> of the information, it can get visually cluttered.



11. Click the **Overview** button to see a drill down of all contact group information. Note the bell curve type distribution of the vdW corrected distances between the pyridine and contact groups.

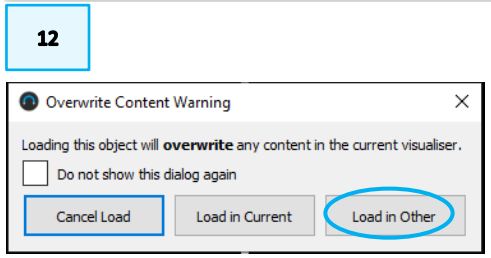




Remember that this data is based only on small molecule interactions (databases are updated frequently so the results can be different) How would this compare with protein-ligand data from the PDB?

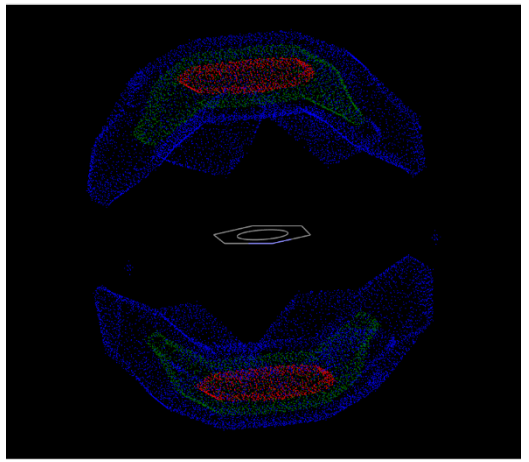
12. Go back to the IsoStar website and download the *phenyl* group information from the PDB (521 entries). Open this in IsoStar as you did previously. Choose to **Load in Other** to load in the second visualizer.

13. Click **vdW Overlaps** to show only contacts within vdW distances. Create contour surfaces again by clicking **Create/Edit** and clicking **Create**. You may be warned that this will result in a statistically insignificant data set. Proceed by clicking **Yes**. Close the window and click *Hide All* to hide the probes.



14. Note that the hotspots are exclusively at π-π stacking regions, showing a strong preference for this interaction in the PDB. Given that the PDB is full of protein-based interactions, and that π-π stacking is a common bioactive motif, it is not surprising to see the difference in CSD vs. PDB hotspot distributions for uncharged pyridine. Compare this contour to the contour of the CSD from step 9 above. What other PDB vs. CSD reasons might account for this difference?

14



Conclusion

We have seen how to examine IsoStar data using the desktop client to view the interaction preferences of an uncharged pyridine central group with a phenyl contact group. IsoStar allowed us to inspect individual interactions, view statistical plots, generate contour surfaces, and facilitated comparison between interaction data from the CSD and from the PDB.

Exercises

- Download the IsoStar CSD data for the interaction of a *charged pyridine* ring with a *phenyl* contact ring. Generate a symmetrised contour plot and compare this to the map obtained from steps 1–10.
- How do the two maps compare?

Tips and Tricks

Although IsoStar contains data for many different functional groups, users may wish to generate their own scatterplots for contacts not currently included in the library. This may be done by performing a nonbonded search of the CSD with ConQuest, and then converting the output to a scatterplot using IsoGen. IsoGen is a Linux-only program and is not covered in this tutorial. However, the IsoStar documentation explains how to do this.

Summary

In this workshop we used IsoStar to view and analyse intermolecular interactions. You should now be able to:

- Locate and download CSD, PDB and Theory IsoStar files from the IsoStar web server.
- Load data and customise the IsoStar display.
- Create contour surfaces.
- Retrieve statistical information for interactions.

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

Next Steps

If you would like to learn more about analysing intermolecular interactions using data from the CSD, you can explore Full Interaction Maps in Mercury. The self-guided workshop on Full Interaction Maps (MER-002) is available on this page: https://www.ccdc.cam.ac.uk/community/training-and-learning/workshop-materials/csd-materials-workshops/. You may also wish to explore more advanced workshops from this webpage on tools such as the Hydrate Analyser (MAT-002) and Aromatics Analyser (MAT-005).

If you would like to try a free on-demand training module on intermolecular interactions you can access the Analysing intermolecular interactions 101 – Full Interaction Maps CSDU module here: https://www.ccdc.cam.ac.uk/community/training-and-learning/csdu-modules/fims-101/

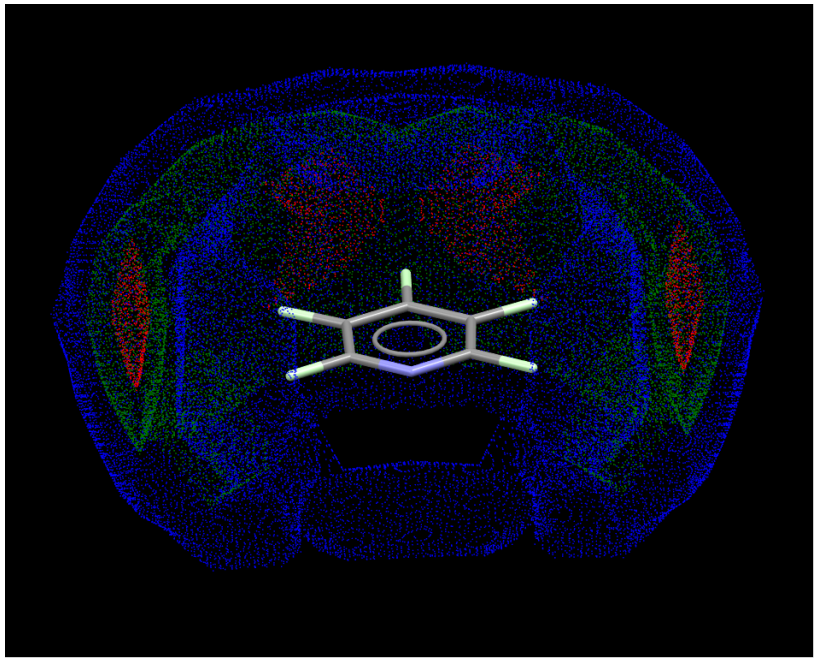
Feedback

We hope this workshop improved your understanding of IsoStar 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 ISO-001. It will only take 5 minutes and your feedback is anonymous. Thank you!

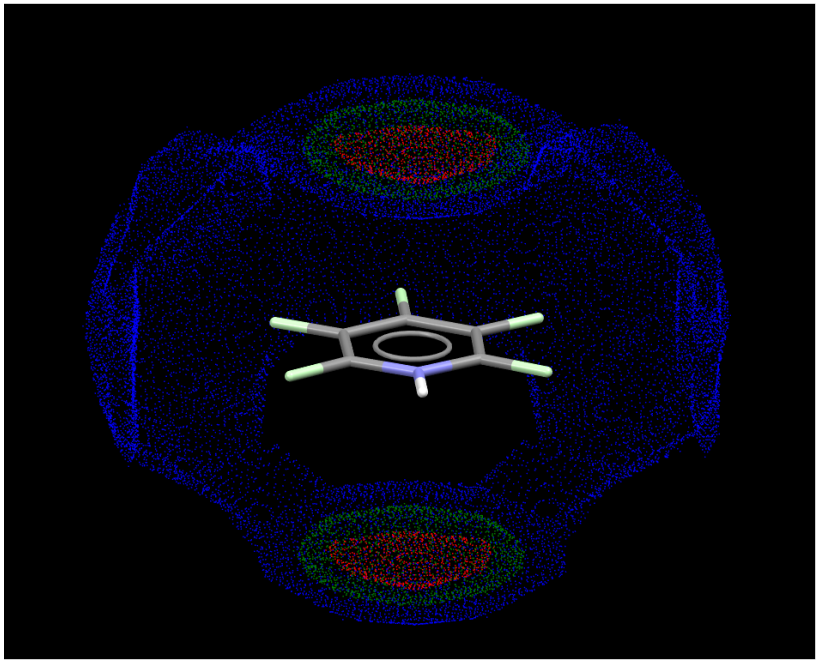
Exercise discussion

(a) and (b) show the contour surface plots for neutral and charged pyridine molecules, respectively, with a phenyl contact group. In (b) the most probable regions to find a phenyl contact group are above and below the aromatic ring. This is a reversal with respect to the uncharged case because the pyridinium ion is π-deficient.

a



b

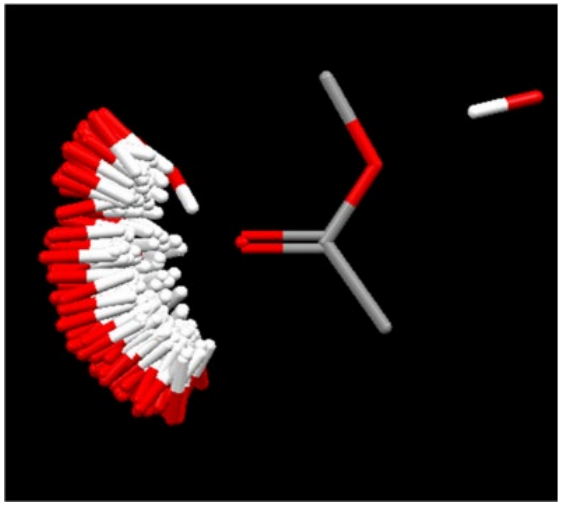


Glossary

Central/contact group
In IsoStar, crystallographic information is presented as a 3D scatterplot of interactions between functional group A (central group) and B (contact group). A…B contacts are transformed such that the A groups are least-squared superimposed. The scatterplot thus shows the distribution of B (contact group) around A (central group).

Contour probe
An atom of the contact group from which it is possible to compute a contoured density surface in IsoStar e.g. for carbonyl, the O or C atom are potential contour probes.

van der Waals corrected distance
The van der Waals corrected distance is the interatomic distance between two groups minus the sum of their van der Waals radii.



In the example above, the ester is the central group whilst hydroxyl is the contact group.