StarDrop™ Worked Example:

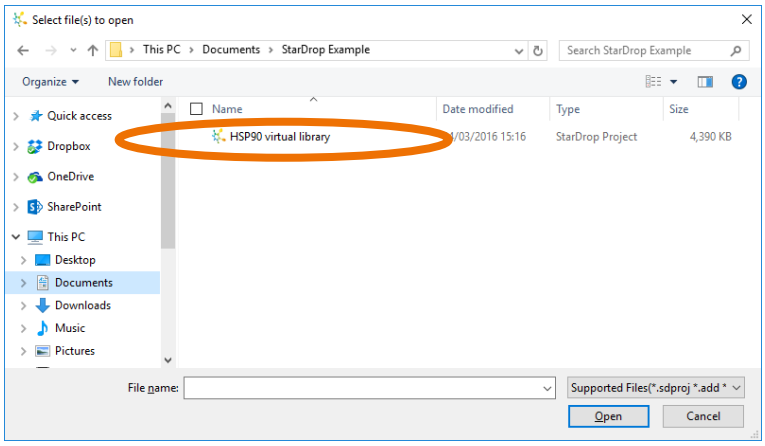
Guided Multi-Parameter Optimisation of 2D and 3D SAR

This worked example uses a combination of 2D and 3D methods to understand and optimise a virtual library of Heat Shock Protein 90 (HSP90) inhibitors. The library, created by a *de novo* design process, is based around an amide substitution on a beta resorcylic acid core. The objective in this example is to use the SeeSAR™ View module to develop an understanding of the 3D structure-activity relationships (SAR) and then use multi-parameter optimisation to further develop the absorption, distribution, metabolism and excretion (ADME) and physicochemical properties of a potent inhibitor without losing efficacy.

During this exercise, we will use a variety of StarDrop's capabilities to explore the data in order to understand the SAR and design compounds with a good balance of properties. Step-by-step instructions for all the features you will need to use in StarDrop are provided, along with screenshots and examples of the output you are likely to generate. If you have any questions, please feel free to contact stardrop-support@optibrium.com.

Exercise

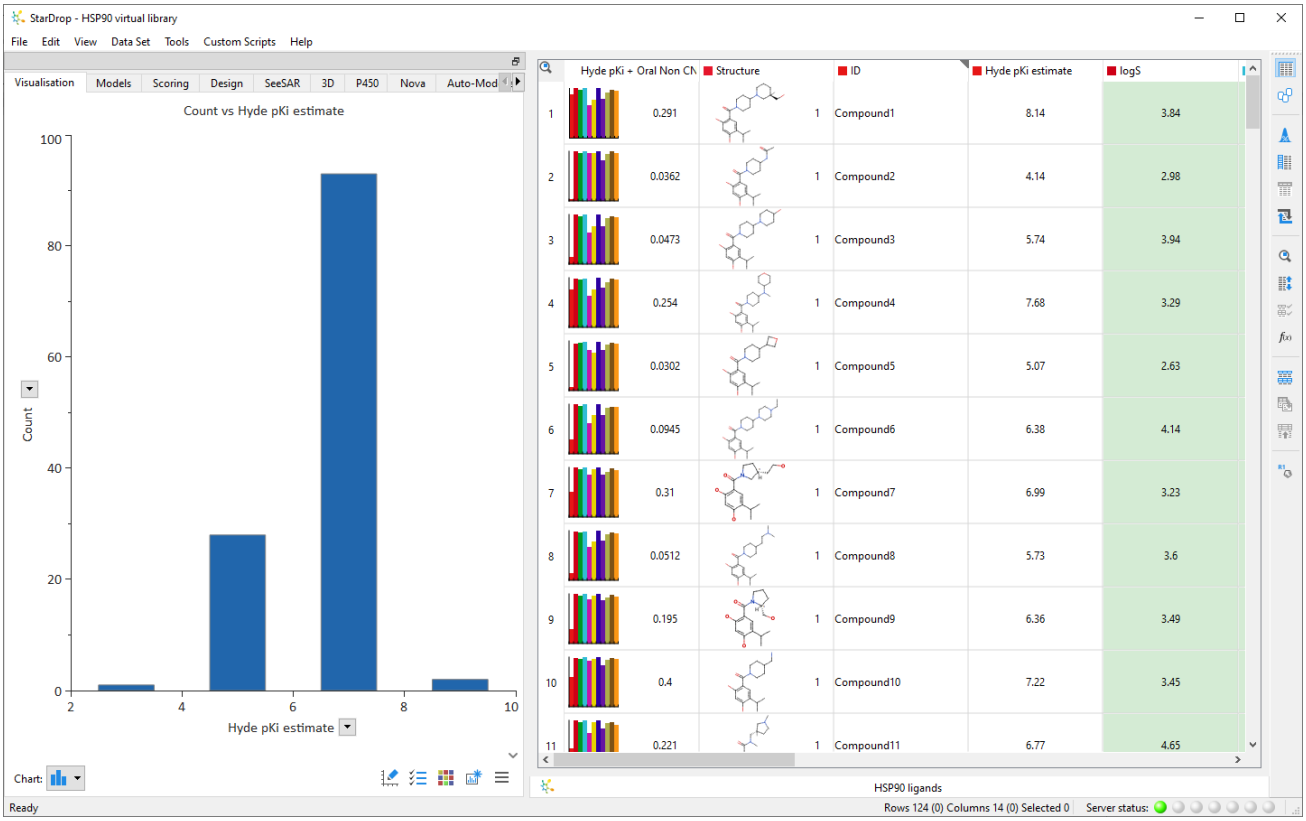
- In StarDrop, open the project file **HSP90 virtual library.sdproj** by selecting **Open** from the **File** menu.



You will see a spreadsheet containing 124 structures with their estimated affinities for HSP90 (in the column **Hyde pKi estimate**) displayed in a histogram. The data set has been scored using a profile of properties designed to assess compounds that need to be active against an oral non-CNS target.

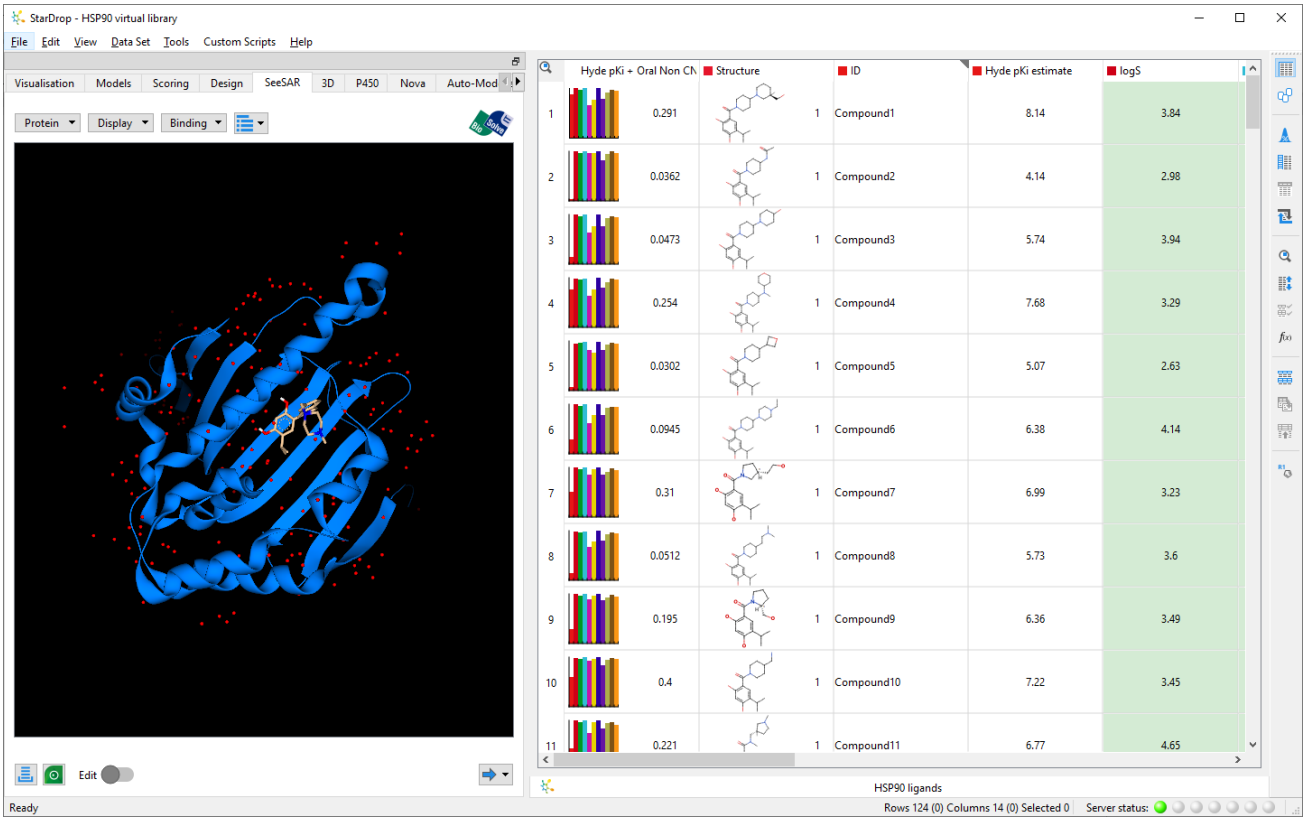
Optibrium™, StarDrop™, Card View® and Glowing Molecule™ are trademarks of Optibrium Ltd. SeeSAR™ is a trademark of BioSolveIT GmbH.

© 2022 Optibrium Ltd.



We can see a large difference between the estimated affinities for the first two compounds in the data set. To investigate what may be causing this, we are going to look at the 3D compound and protein structures.

- Click on the **SeeSAR** tab at the top.



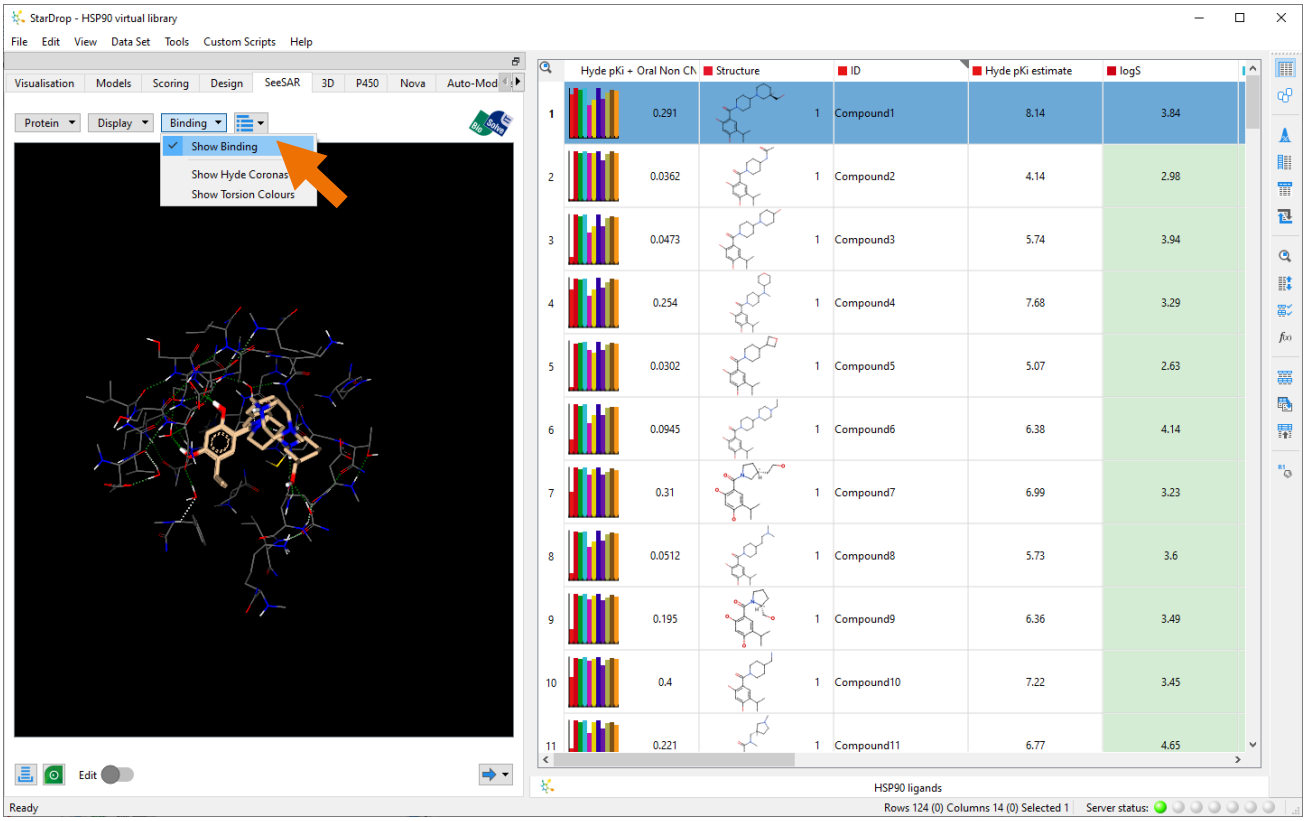
The PDB file 2XJX has already been loaded into this StarDrop project, and we can see this within the SeeSAR viewer. Here we can also see Onalespib, an HSP90 inhibitor that was included in the PDB file.

Any ligands we select in our data set will be displayed in the SeeSAR 3D viewer.

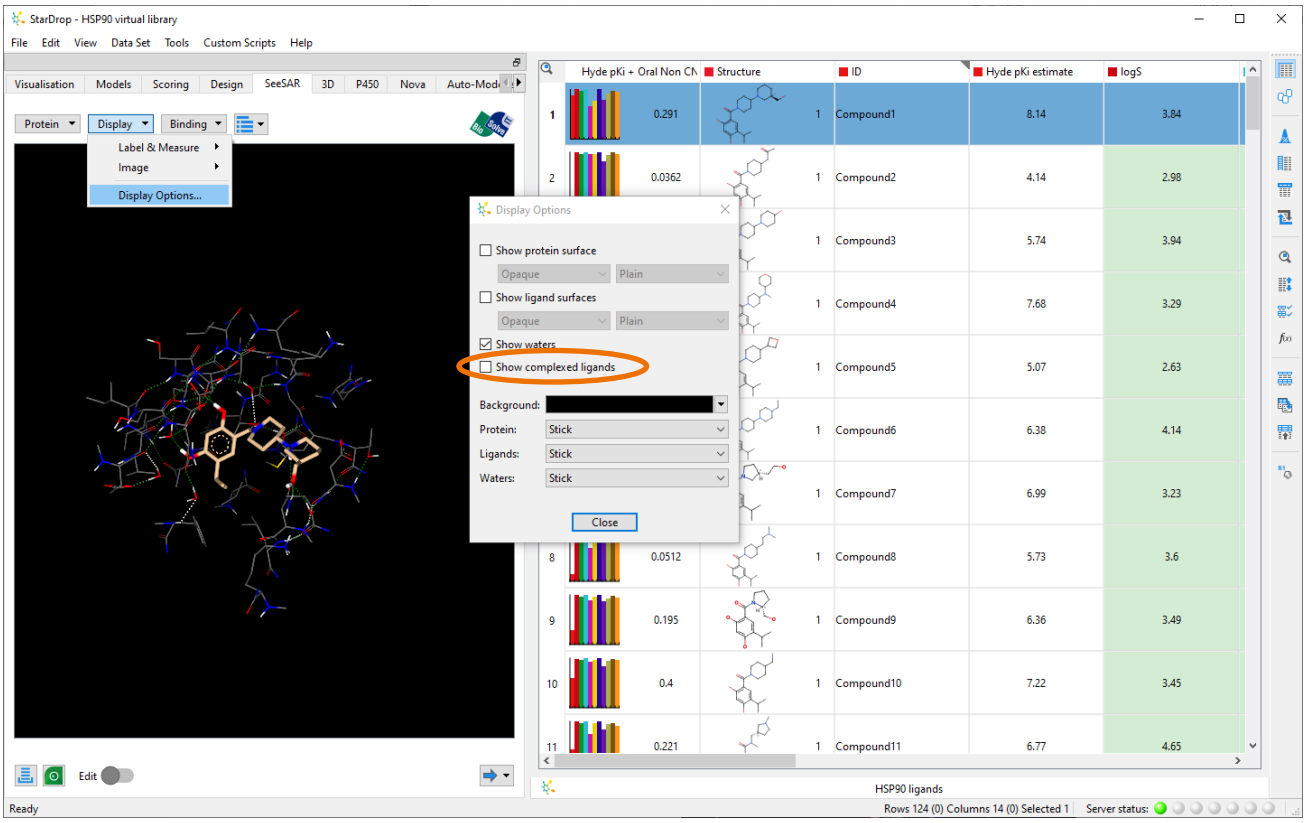
- Select the first row in the data set – this is the ligand we identified with good estimated activity.

Note: You can select multiple rows in the data set by holding the **Ctrl** key on Windows, or C**ommand** key on macOS, as you select them in turn.

- Now choose **Show Binding** from the **Binding** menu at the top of the SeeSAR area to change the view of the protein to focus on the binding pocket.

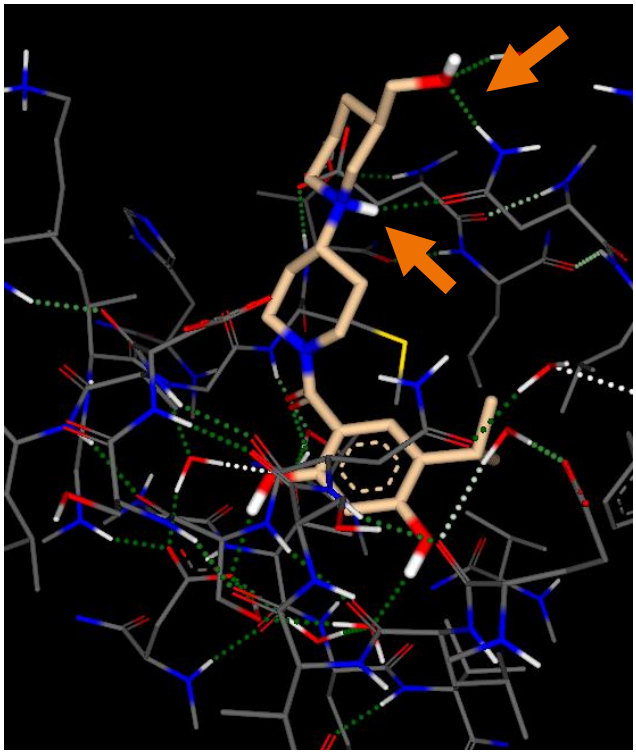


- From the **Display** menu, choose **Display Options**.
- In the **Display Options** dialogue, untick the **Show complexed ligands**.



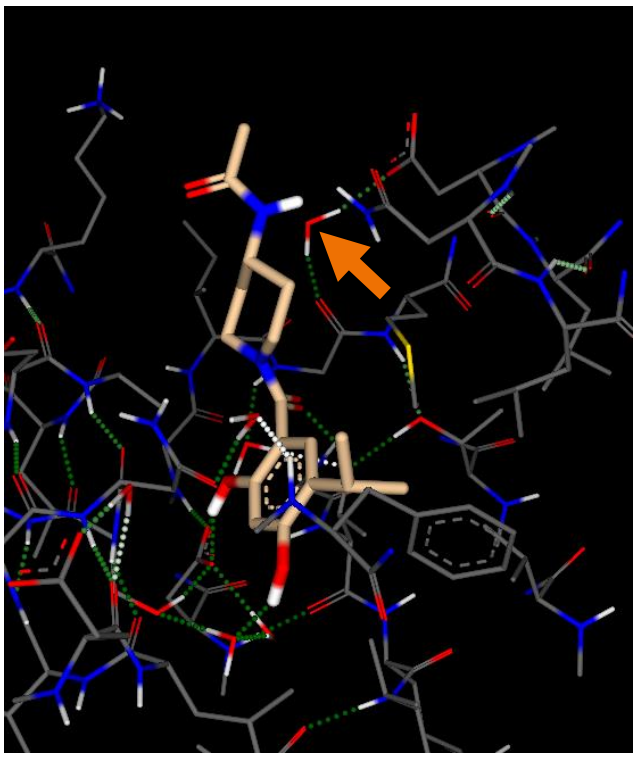
Using the mouse, you can interact with the view of the protein:

- Use the mouse-wheel to zoom in and out
- Use the left mouse button and drag to rotate the view
- Use the right mouse button and drag to pan the view



When you choose the **Show Binding** option, strong hydrogen bond interactions between the ligand and protein are displayed by green dotted lines (white dotted lines indicate weaker hydrogen bond interactions). Focusing, in particular, on the hydroxy-methyl piperidine portion of the ligand, we can see its bi-dentate interaction with HSP90.

- Select the second row in the data set to display the N-acetyl derivative instead.

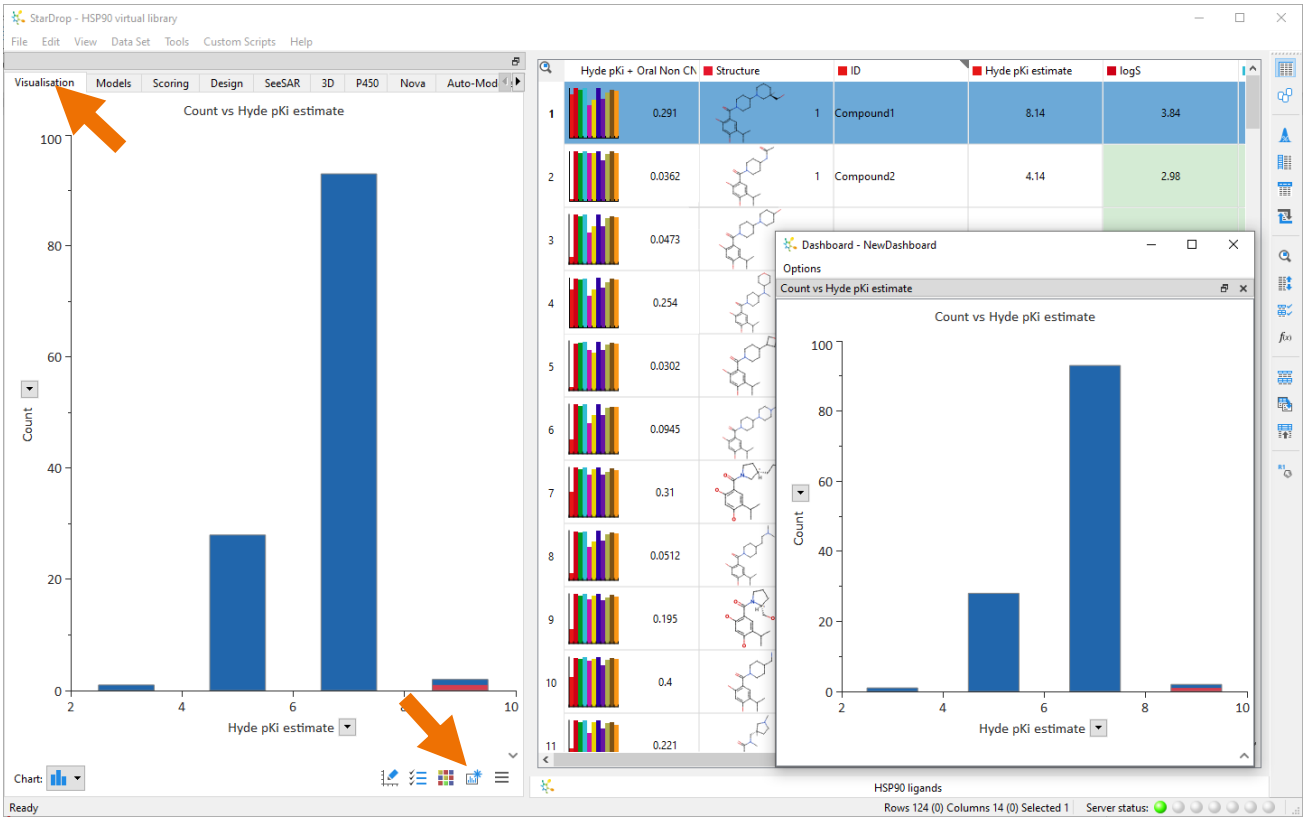


Here we can see that there are no equivalent interactions between this part of the ligand and the protein, which suggests that this may be responsible for the significant reduction in the estimated activity.

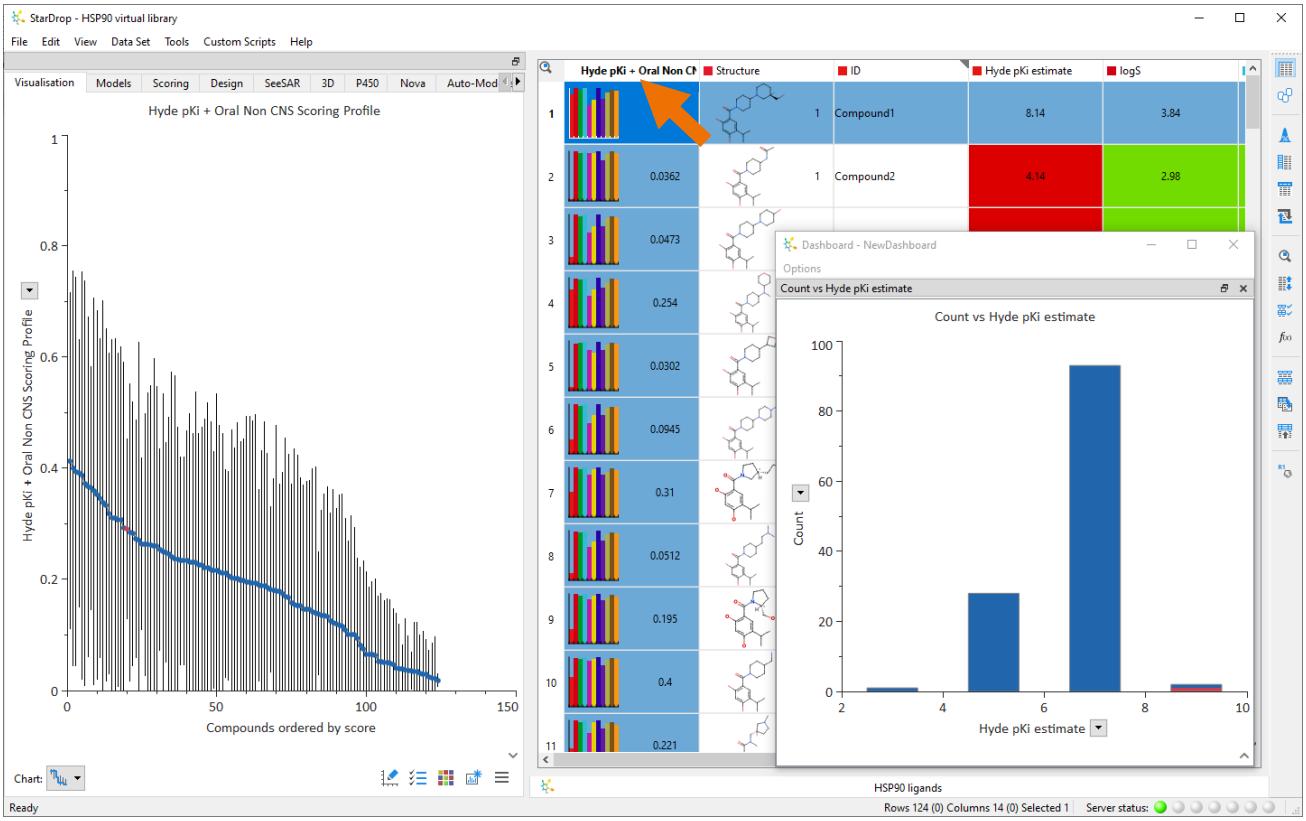
Having considered some of the reasons underlying the estimated activity of the ligands in this virtual library, let's now consider some of their ADME and physicochemical properties.

- Change to the **Visualisation** area.
- Click the **Detach** button to create a new dashboard containing the histogram.

Note: Any other charts you detach will be added to the dashboard, which you can organise and resize by dragging plots and the boundaries between them.

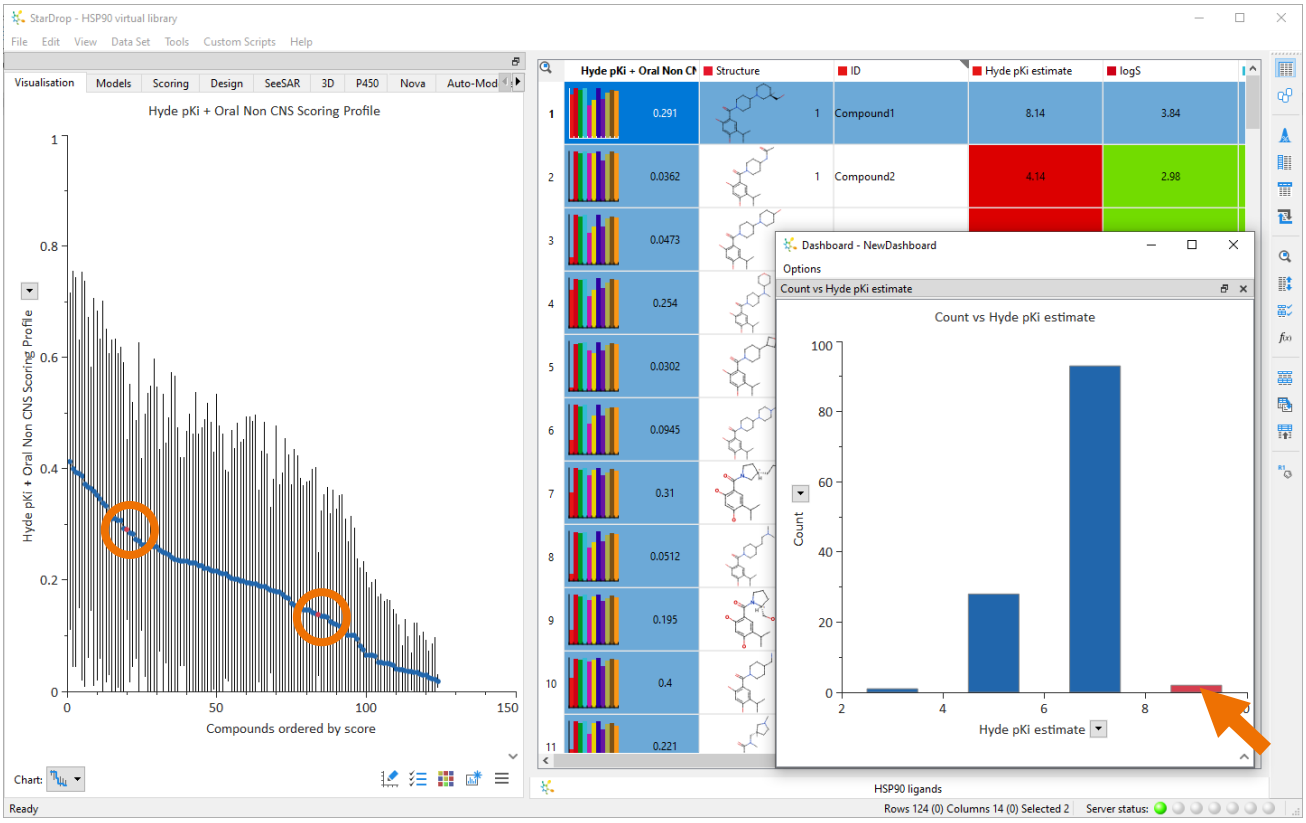


- In the data set, select the score column by clicking on its column header to generate a 'snake' plot.



A ‘snake’ plot shows the scores for a whole data set, ordered from highest to lowest, indicating the actual scores along with the confidence in the scores due to the uncertainty in the underlying data. All of the data in this set have been predicted, which is why there are a lot of large error bars; however, there are still significant differences between the highest and lowest scoring compounds.

- In the histogram, click the bar which contains the compounds with estimated pK_i values above eight.



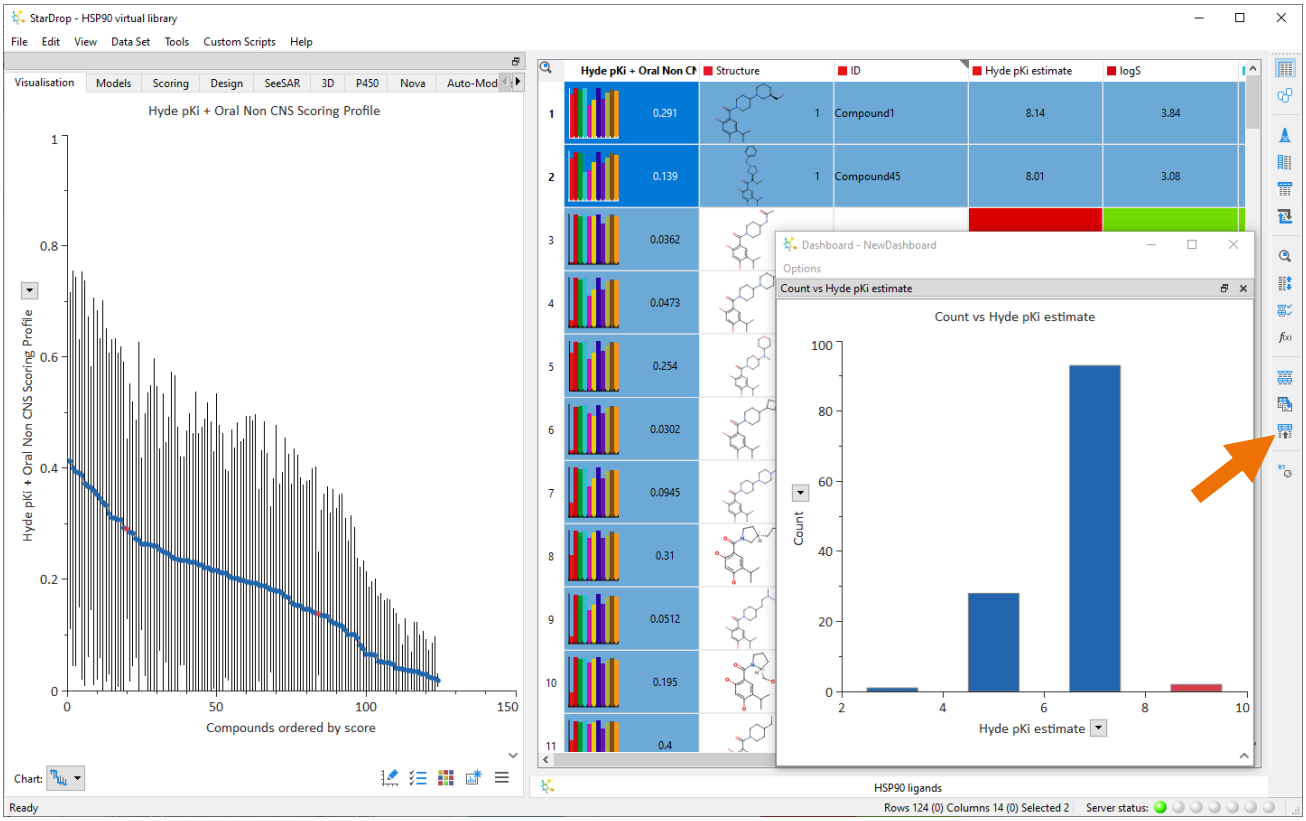
Any selections you make within charts or data sets in StarDrop are reflected everywhere else, so we can quickly see from the snake plot that the two most potent compounds have very different scores.

Let’s take a closer look at the properties and the reason for the low score.

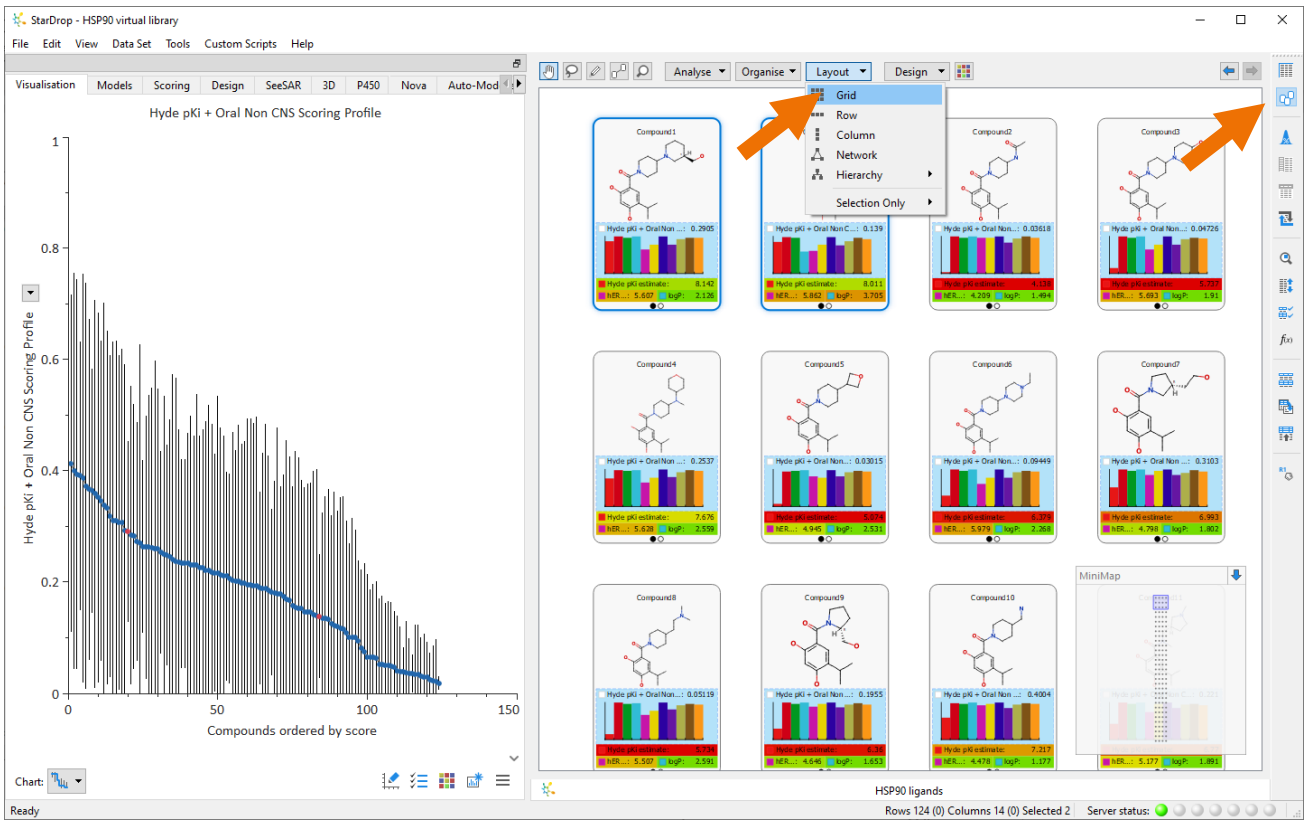
- Close the dashboard (you can ignore any prompt to save it)

One of the selected rows is at the top of the data set, but we can move the other selected row to be alongside it at the top of the data set.

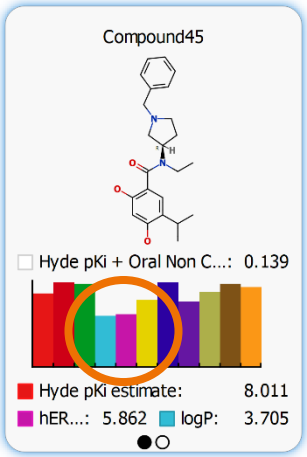
- Click the **Move Selected Items to Top** button on the right-hand toolbar.



- Change into Card View by clicking the **Card View** button on the right-hand toolbar.
- From the **Layout** menu at the top of Card View, choose **Grid**.

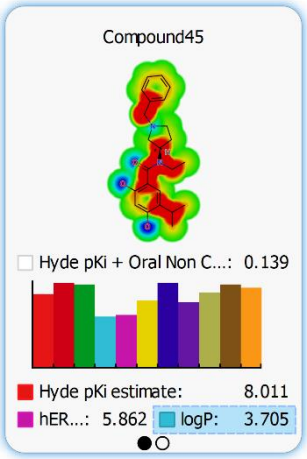


- In Card View, click on the desktop to clear the selection
- Zoom in to look at the properties of the second compound, called **Compound45** (use the middle mouse wheel or in Windows use the **Ctrl** and **-/=** keys to zoom in and out, on macOS, use **Command** and **-/=**. You can also use the touchpad to zoom in and out. The zoom will be centred on the mouse position.)



In the scoring histogram, we can see that the two lowest bars, which indicate potential problems, are the cyan and the pink ones. These correspond to logP and hERG pIC_{50}, respectively, and the values of these two properties are shown below. When scoring the compounds, we indicated that we would ideally like a logP value between 0 and 3.5. We also indicated that an ideal hERG pIC_{50} value would be below 5.

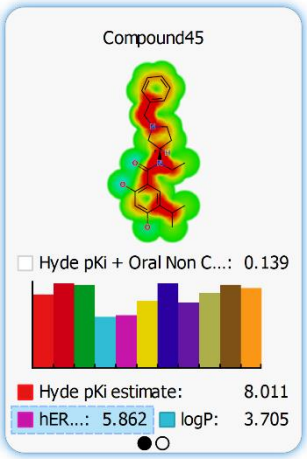
StarDrop's Glowing Molecule™ highlights regions of the molecule that are having a significant effect upon a predicted property. This can help to guide the design of new compounds by indicating regions where a small change in structure may have a significant impact on the predicted property.



- Click on the logP prediction on the card to activate the Glowing Molecule visualisation for logP.

Regions glowing red tend to increase the predicted value, whereas regions glowing blue tend to decrease the predicted value.

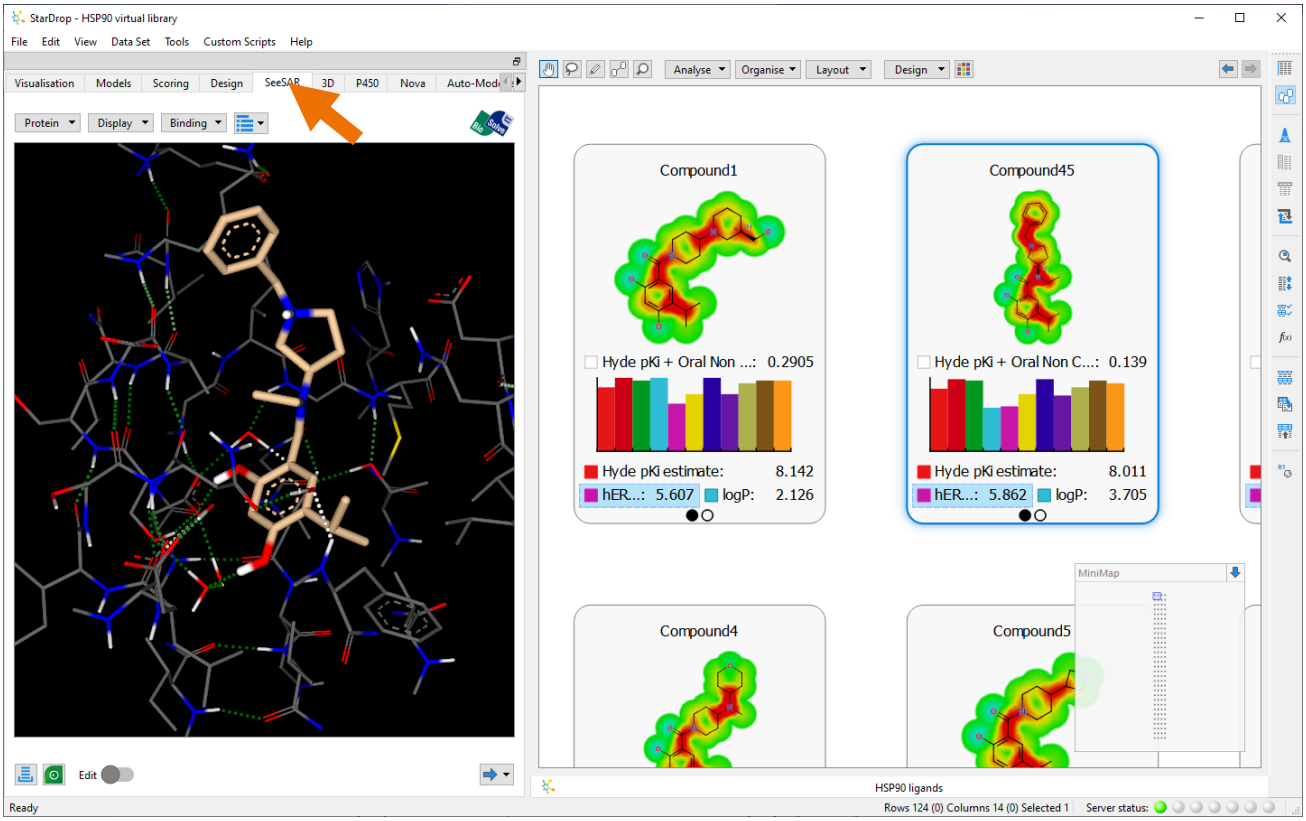
- Now select the hERG pIC50 property to view the Glowing Molecule representation for this prediction.



For both the logP and hERG pIC_{50}, we can see that regions around the beta resorcylic, phenyl and pyrrolidine rings are red, indicating that these parts of the molecule are contributing to an increase in these properties.

We can now easily compare the 2D SAR indicated by the predictive models with the 3D SAR indicated by the docking pose.

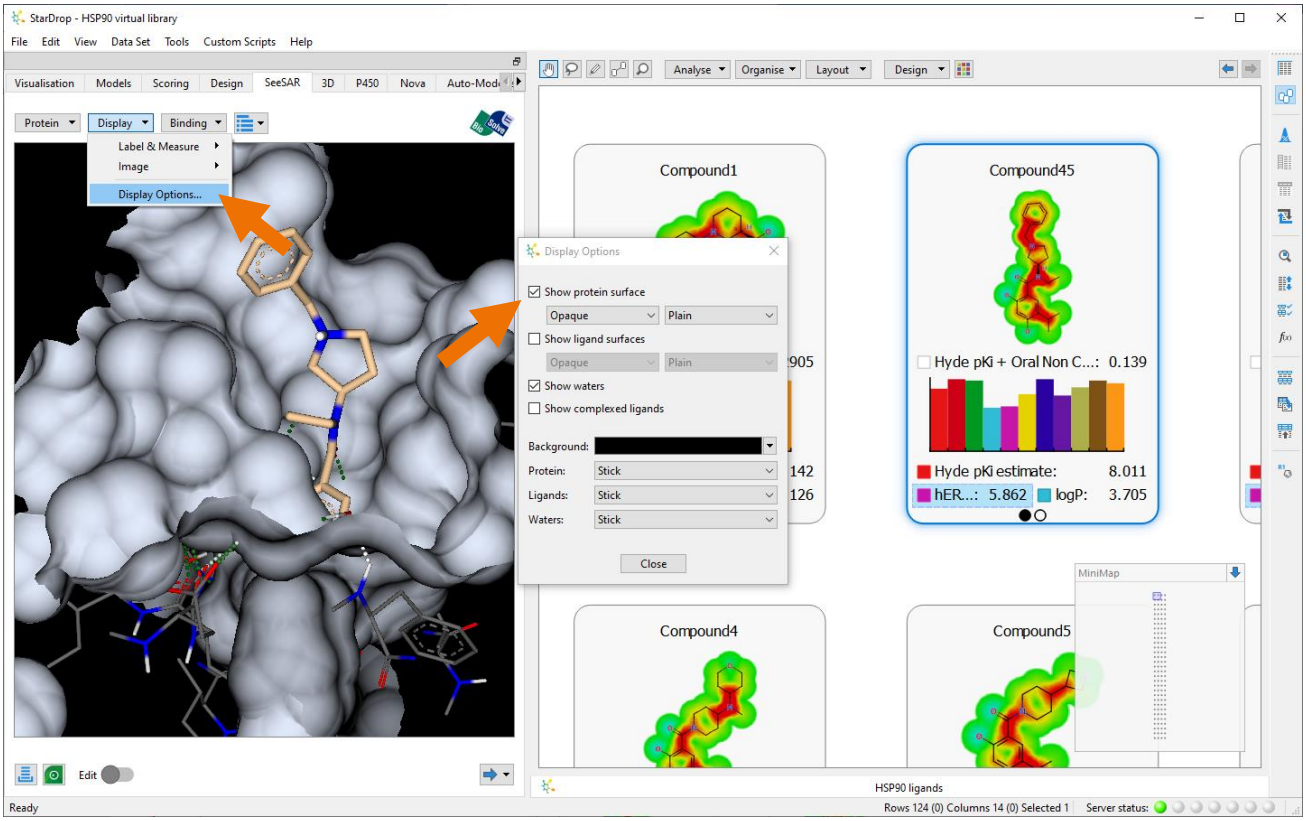
- Change back to the **SeeSAR** area, and you will see this compound shown in SeeSAR.



- From the **Display** menu at the top of the SeeSAR area, choose **Display Options**.

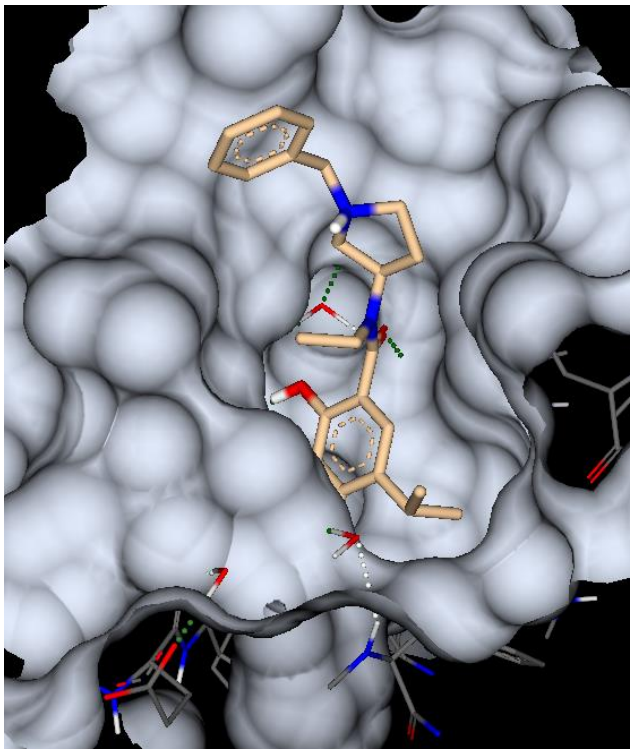
Here we can choose from a number of display options, and the SeeSAR area will update to show each effect.

- In the **Display Options** dialogue, tick the box next to **Show protein surface**.



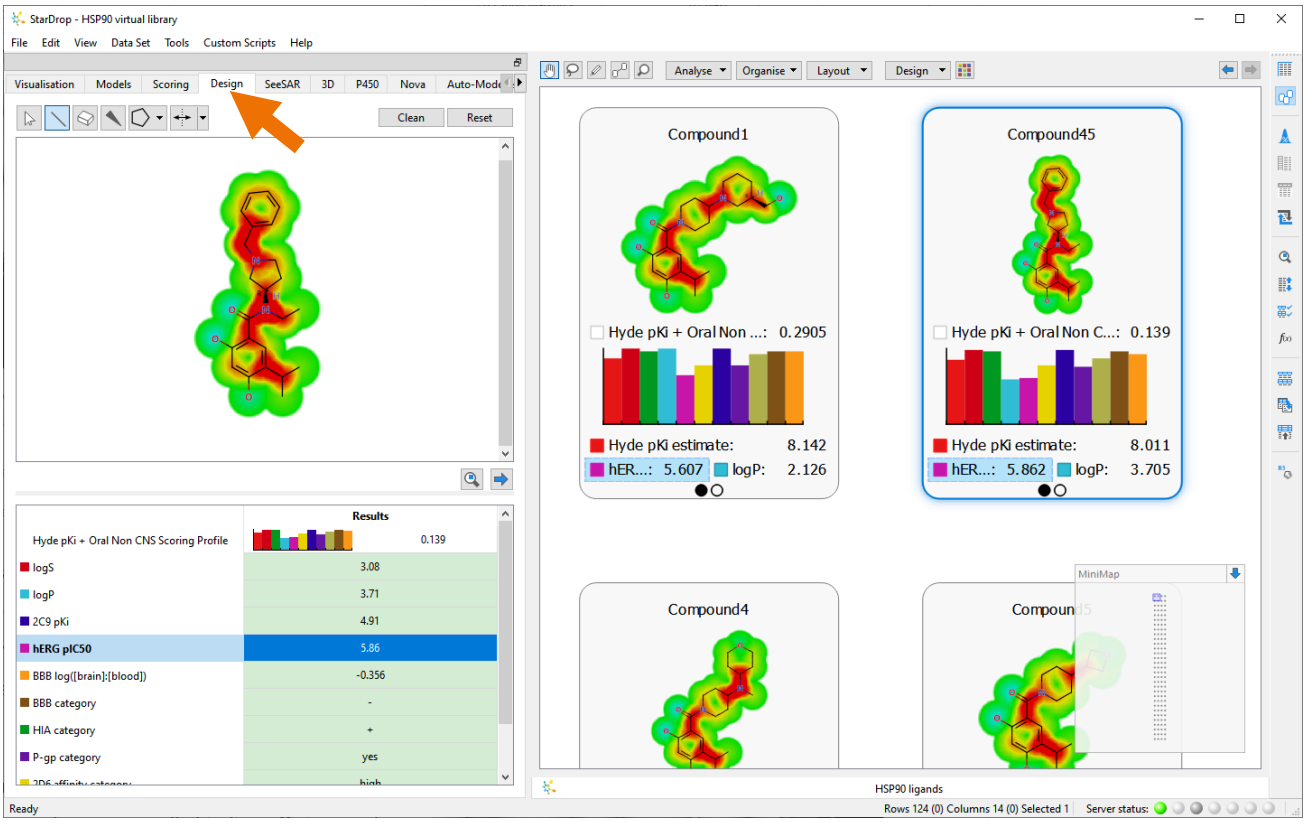
- Click the **Close** button to close the Display Options dialogue.
- Rotate the structure within the SeeSAR area to see how the ligand fits into the binding site.

Whereas the resorcylic acid core sits deep within the pocket and forms a strong hydrogen-bonding network with the protein, the phenyl and pyrrolidine rings sit outside the pocket, showing no obvious strong interactions.



This suggests that these regions may not contribute strongly to the compound's binding affinity, so these rings may well be a good place to consider making modifications in order to mitigate the risks of high logP and hERG inhibition without having a large effect on activity. We can use StarDrop's interactive design tool to explore the impact of possible changes.

- Change to the **Design** area.



Here we can edit and design new compounds, and the selected molecule, Compound45, is already shown. Below the molecule editor, you can also see predicted properties which will update as we modify the compound structure.

- Click on the **Bond** tool and add a carboxylic acid at the meta position on the phenyl ring.

Note that to draw a bond, you simply hover the bond tool over the attachment point and then click and drag. To draw a heteroatom, hover the bond tool over the atom you wish to modify and type the element symbol. To change to a double bond, click the bond tool on the bond you wish to turn into a double bond. To undo any changes type **Ctrl-z** on Windows or **Command-z** on macOS. You can use the mouse-wheel to zoom in and out of the editing area if necessary (or use the **Ctrl** and **-/=** keys on Windows or **Command** and **-/=** keys on macOS).

While editing the molecule, you will see that the properties below the design area will update to show new predictions for the displayed molecule. If a property in the list is selected, then you can also see the Glowing Molecule for that property update while editing. The carboxylic acid derivative has a logP of 3.26 and a hERG pIC_{50} of 4.66, and so this compound would appear to have a better balance of ADME properties (you can see all the other properties as well to confirm that none of these has become significantly worse).

- Add this new molecule to the data set by clicking the button.

We can continue to investigate other alternatives, and to find out the effect any changes have had on the predicted activity we could use the SeeSAR Pose and SeeSAR Affinity modules. Alternatively, it is also possible to use StarDrop's Pose Generation Interface to link StarDrop with your preferred docking program to generate poses and score the resulting fit to the binding pocket.

To learn more about the SeeSAR Pose and SeeSAR Affinity modules, go to:

SeeSAR Pose module - https://www.optibrium.com/videos/stardrops-seesar-pose-module/

SeeSAR Affinity module - https://www.optibrium.com/videos/stardrops-seesar-affinity-module/

For an example of the Pose Generation Interface, go to: https://www.optibrium.com/videos/pose-generation-interface/

This example illustrated how we are able to combine our understanding of the 2D and 3D SAR to suggest improvements to compounds from a virtual library designed as inhibitors for HSP90. StarDrop contains many more features for selecting and designing compounds, and further worked examples are available from the Optibrium community at http://www.optibrium.com/community/tutorials. If you have any questions or feedback, please contact stardrop-support@optibrium.com, and we will be happy to help.