StarDrop™ Worked Example:

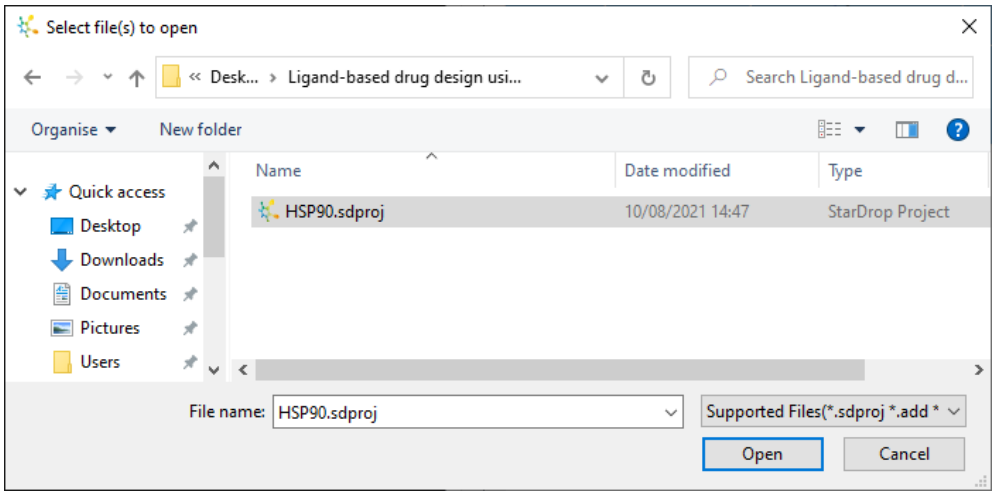
Ligand-based drug design using Surflex eSim 3D™

This worked example uses StarDrop's Surflex eSim3D module to assess a small library of compounds for their similarity to known Heat Shock Protein 90 (HSP90) ligands. The library, created by a *de novo* design process, contains compounds with a beta resorcylic acid core. The objective in this example is to use the Surflex eSim3D module to develop an understanding of the 3D structure-activity relationships (SAR) and then use multi-parameter optimisation to assess the similarity score together with the absorption, distribution, metabolism and excretion (ADME) and physicochemical properties of the library.

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.

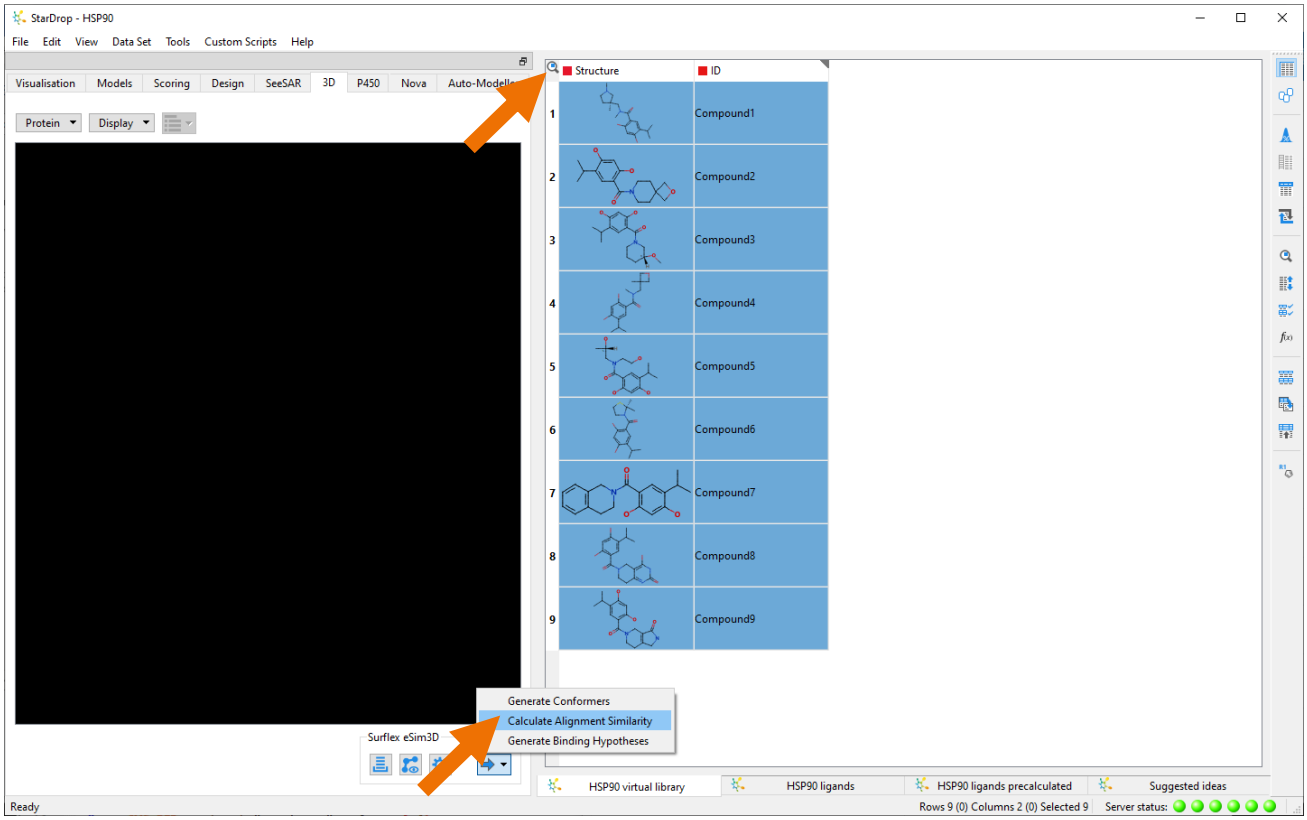
Exercise

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



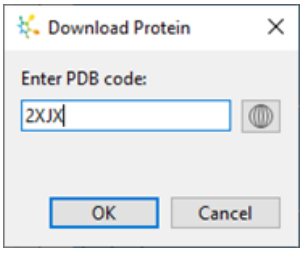
This opens a data set of nine virtual HSP90 inhibitors. The compounds are based around a known HSP90 inhibitor, Onalespib, and share a beta resorcylic acid core. To investigate how similar these compounds are in terms of their shape, electrostatic potential and hydrogen bonding potential to Onalespib, we will do a similarity analysis.

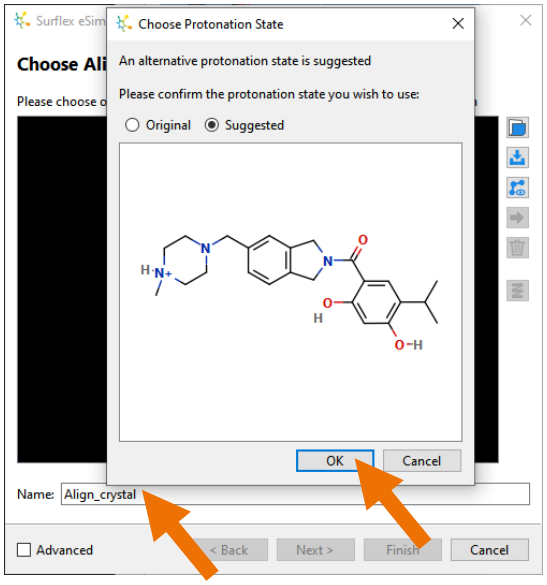
- Select all compounds of the HSP90 virtual library data set by clicking in the top left corner of the data set.
- Click the **Go button menu**  and select **Calculate Alignment Similarity.**



This will prompt you to choose an alignment reference molecule to which to align your compounds. We will use the HSP90 inhibitor Onalespib as a reference, and we will obtain the bioactive conformation of Onalespib from a co-crystal structure.

- Click the download button  to obtain the ligand from a PDB reference.
- Enter the PDB code '2XJX' and click OK.



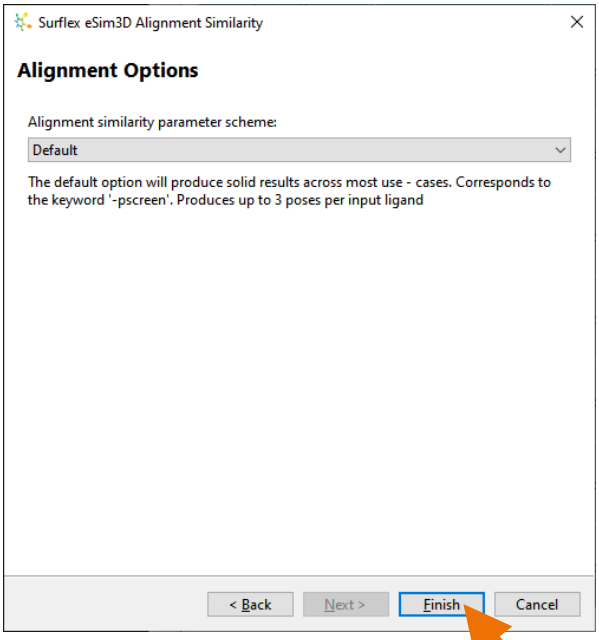


A dialogue box appears, suggesting an alternative protonation state for the nitrogen on the piperazine ring.

- Click **OK** to accept the protonation.
- In the **Name** textbox, enter 'Align_crystal'.

There are further options, such as the ability to add torsional and positional constraints if you select the **Advanced** options. However, we will use the defaults.

- Click **Next**.



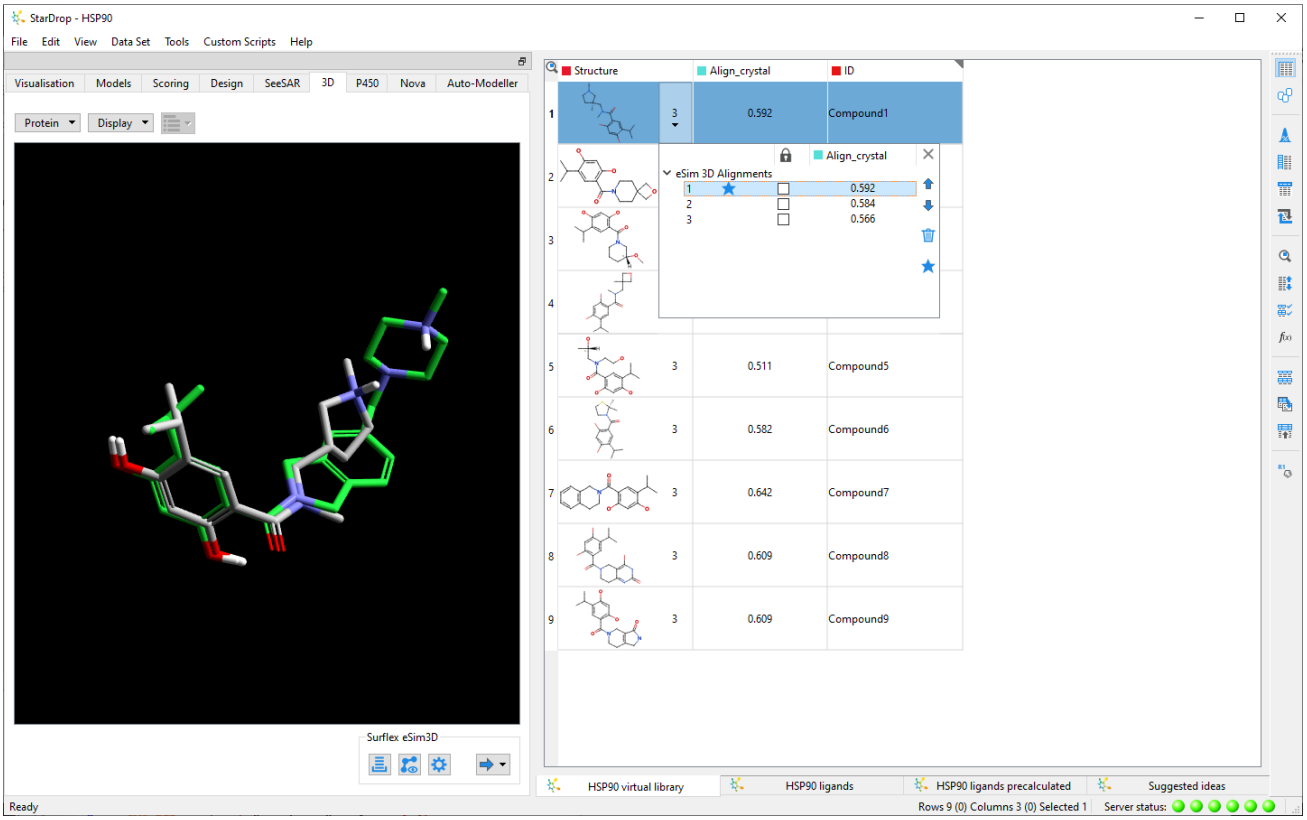
This page gives the option to define the parameter scheme of the alignment similarity analysis. The parameter schemes vary in terms of speed and accuracy.

- Click **Finish** to run the alignment with the default option.

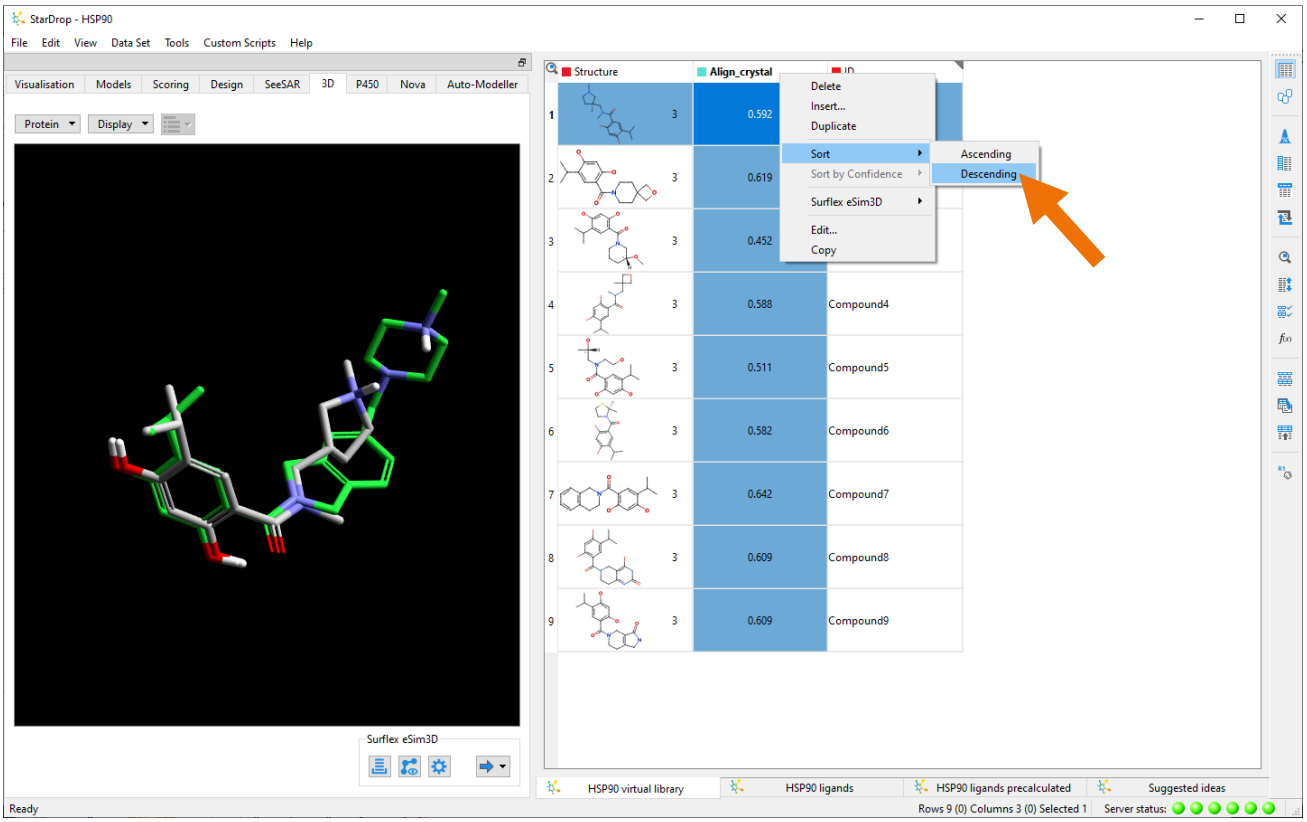
Once the alignment finishes, the calculated similarity value is added to our data set in the new **Align_crystal** column. A number is added to the **Structure** column, which is the number of alignments it has generated. We can view the alignments by clicking the number next to the structure. The best scoring conformation is the primary pose which is denoted by the ★ symbol.

- Select Compound1 to show it in the 3D viewer aligned to Onalespib.
- Click on the number of alignments to view the aligned conformations.

Note: the primary pose can be changed by selecting the conformation and then clicking the star ★ button on the right of the table. The primary pose is the one that is shown in the 3D viewer when the row is selected.

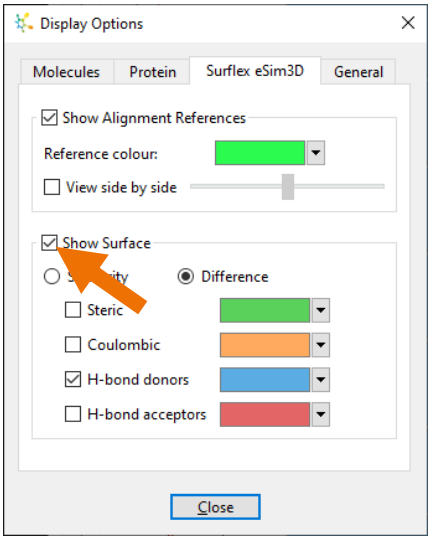


- Right-click on the **Align_crystal** column to bring up the menu and choose **Sort**, then **Descending** to bring the compounds with the highest similarity values to the top.



We can see that the most similar compound to the co-crystallised ligand is Compound7. Select Compound7 in row 1 to see the alignment in the 3D viewer. There is good alignment with the beta resorcylic acid and isoindoline regions of our reference compound, Onalespib. However, Compound7 is much smaller and doesn't have a portion that aligns with the piperazine ring. Therefore despite its alignment, the potency might be much lower.

- Select the  button to bring up the Surflex eSim3D display options.
- Select the **Show Surface** option and explore the similarity and dissimilarity surface maps to see which regions of Compound7 align well.
- Select the **Difference** option and check the box next to **H-bond donors**.

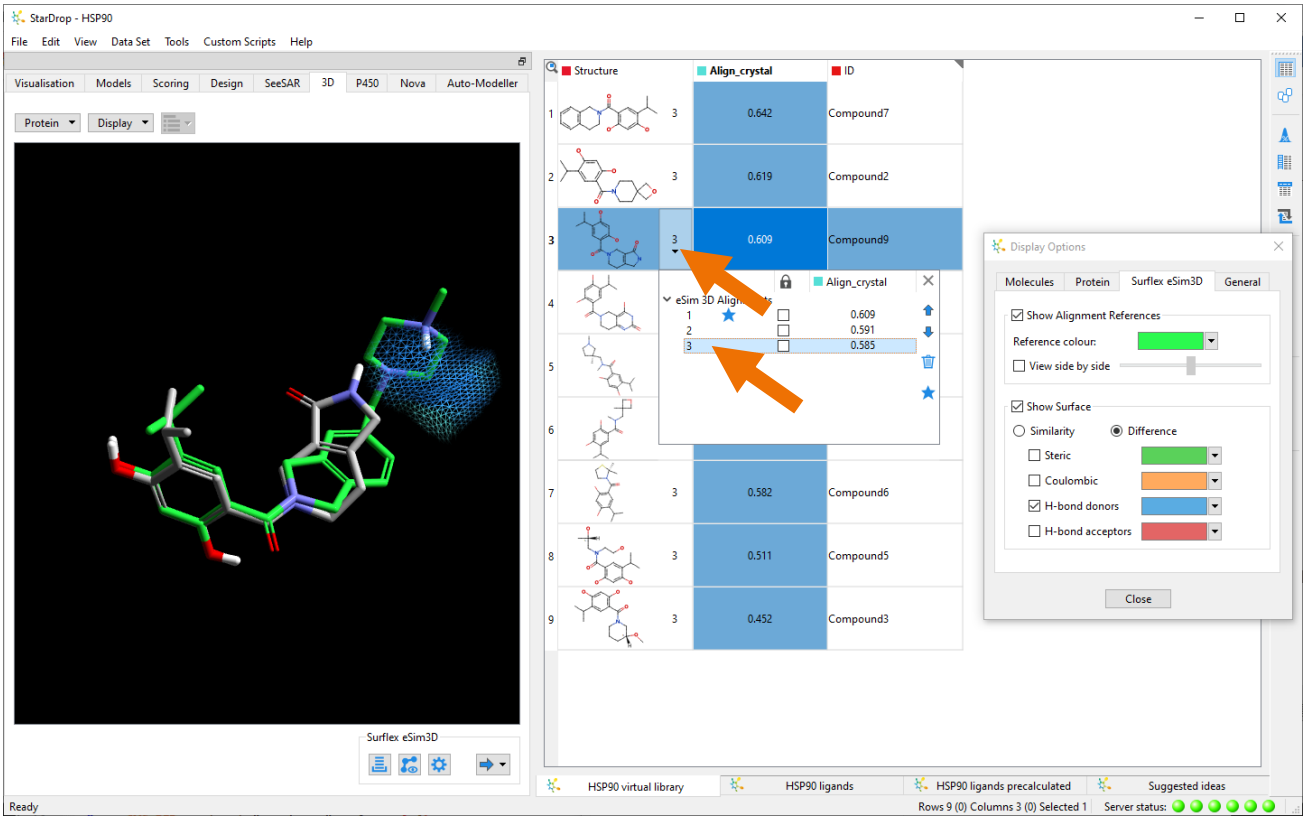


These surfaces highlight regions of similarity and dissimilarity in terms of four properties:

1. Steric – the shape of the molecule
2. Coulombic – the electrostatic potentials of the molecule
3. H-bond donors – where these are positioned
4. H-bond acceptors – where these are positioned

We can see there is a difference highlighted around the distal piperazine. Compound9 is in row 3 and has the opportunity to substitute the amide NH to reach the space occupied by the piperazine.

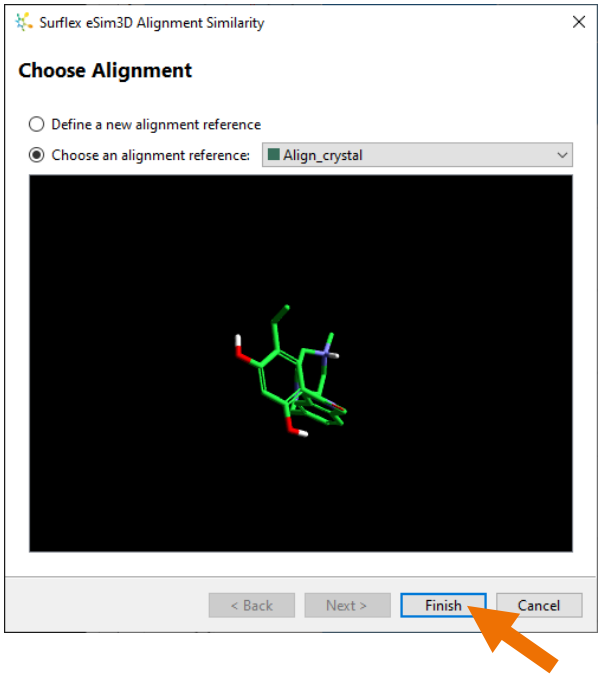
- Select Compound9 in row 3 and explore the alignments for that compound.



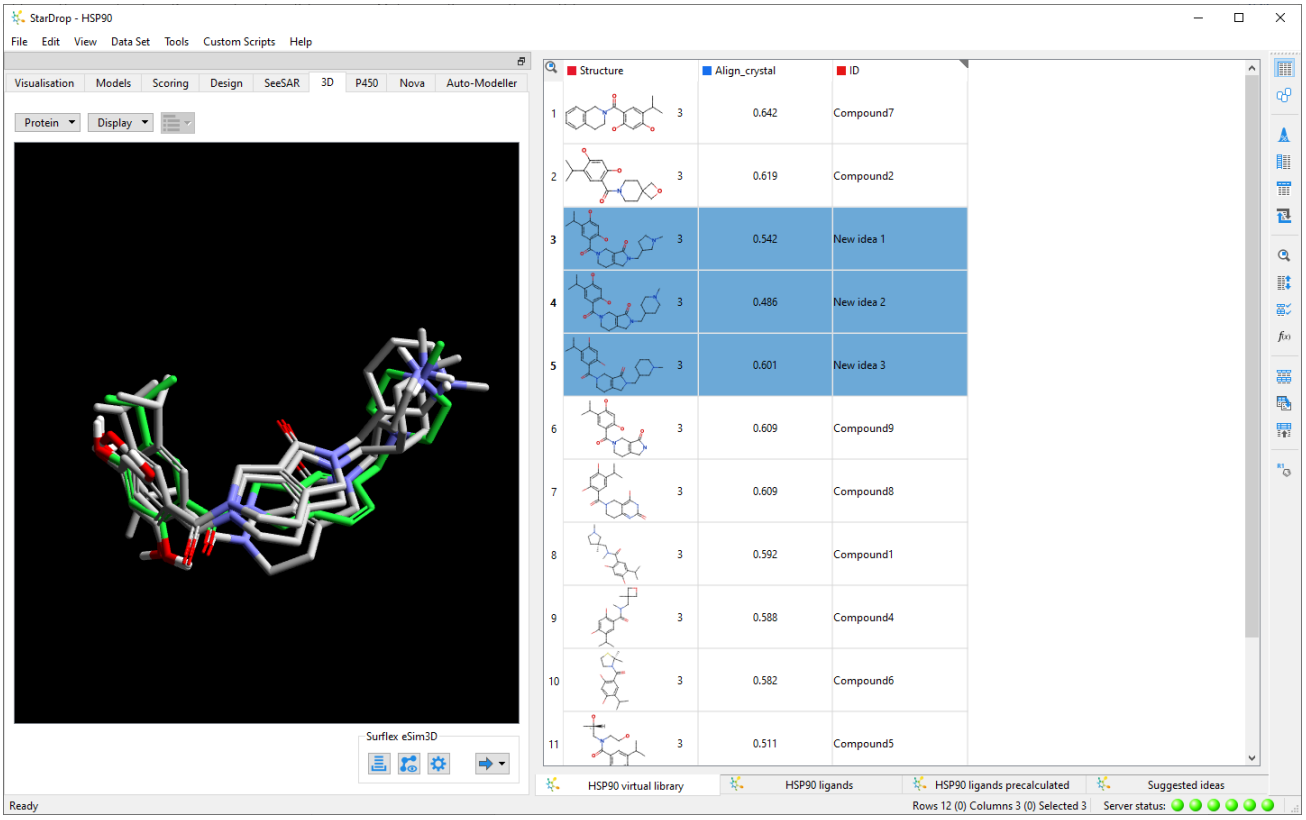
A direct methylene link to a piperazine would not be stable in this compound, but we can instead design a piperidine or pyrrolidine. To save time, we have included suggested ideas based on Compound9 in the fourth data set of this project.

- Click **Close** to close the Display Options.
- Select the **Suggested ideas** data set.
- Select all three compounds by clicking in the top left corner of the data set.
- Go to the **Edit** menu and select **Copy**.
- Go back to the HSP90 virtual library data set, go to the **Edit** menu and select **Paste**.
- Select the three new compounds. Hint- multiple rows can be selected by holding down the Ctrl key.
- Click the **Go button menu**  and select **Calculate Alignment Similarity.**

- The reference we chose in our previous analysis is saved. Click **Finish** to align these new compounds to the reference compound.



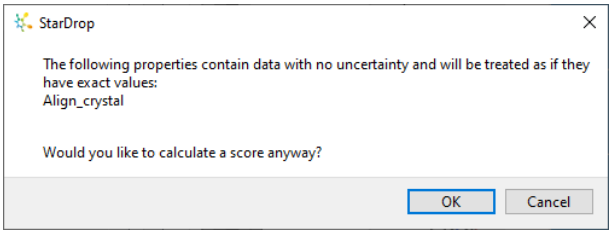
Despite the obvious 2D overlay, these derivatives don't score as well because they disturb the resorcylic acid alignment.

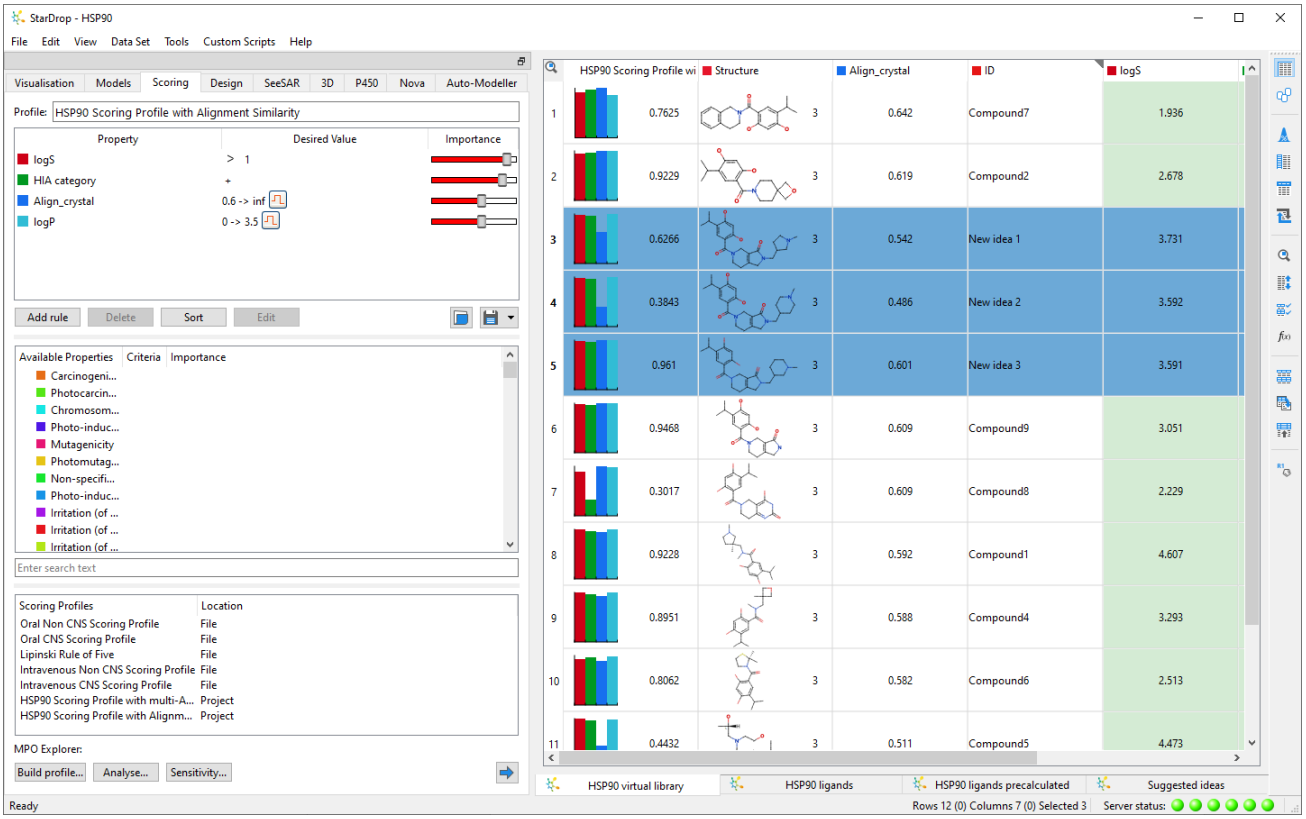


The alignment score isn't the only important consideration, and these derivatives may be superior in other properties. Therefore, we'll use StarDrop's Probabilistic Scoring approach to multi-parameter optimisation to assess each compound's properties against the overall profile required by the project. We can then use this score to track our progress as we attempt to design compounds with an improved *balance* of properties.

- Select the **Scoring tab** and run the 'HSP90 Scoring Profile with Alignment Siliarity' scoring profile by clicking the go button  .

A message will appear stating that the Align_crystal column doesn't have an uncertainty associated with the data. Click **OK** to continue.





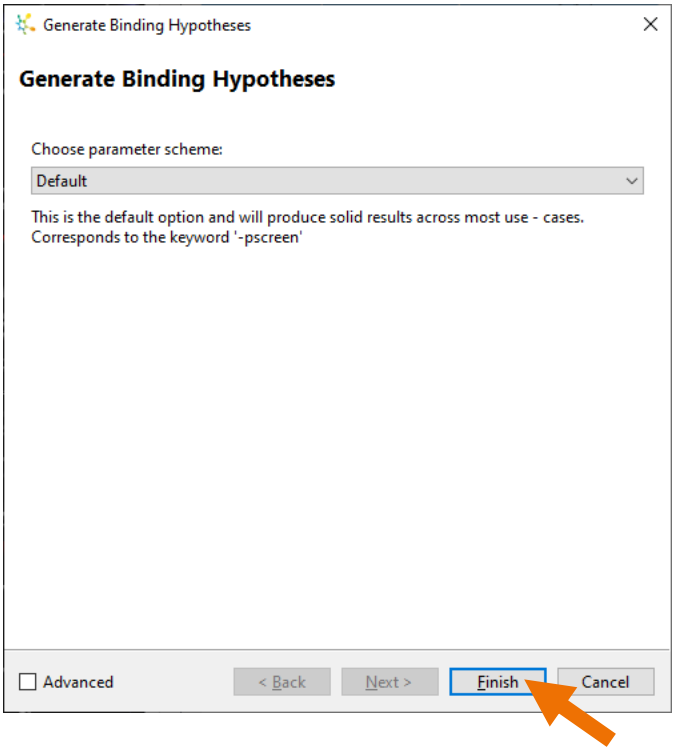
The score is in the range of 0 to 1 (the higher, the better) and represents the likelihood of success of each compound against the overall profile of property requirements, taking into account not only the property values but also the uncertainty in each prediction or measurement. The histogram shows the impact of each individual property on the overall score; the colours correspond to the key in the scoring profile. By hovering the cursor over the histogram, we can see that there is a lower bar for the Align_crystal similarity score for two of the new ideas.

Just as designing against one property could lead to bias, and missed opportunities, comparing the alignment similarity to only one ligand might introduce bias. There are several ligands that bind to HSP90, and five are included in the 'HSP90 ligands' tab. We shall use all of these ligands to generate a binding hypothesis to gain a better understanding of the potential shape and interactions of an ideal ligand.

- Select the second data set of the project called ‘HSP90 ligands’.

This data set shows Onalespib and four other compounds with their HSP90-bound conformation determined. These compounds all contain the beta resorcylic acid core.

- Select all compounds of this data set by clicking in the top left corner of the data set.
- Go back to the 3D area and click the **Go button menu**  and select **Generate Binding Hypotheses**.

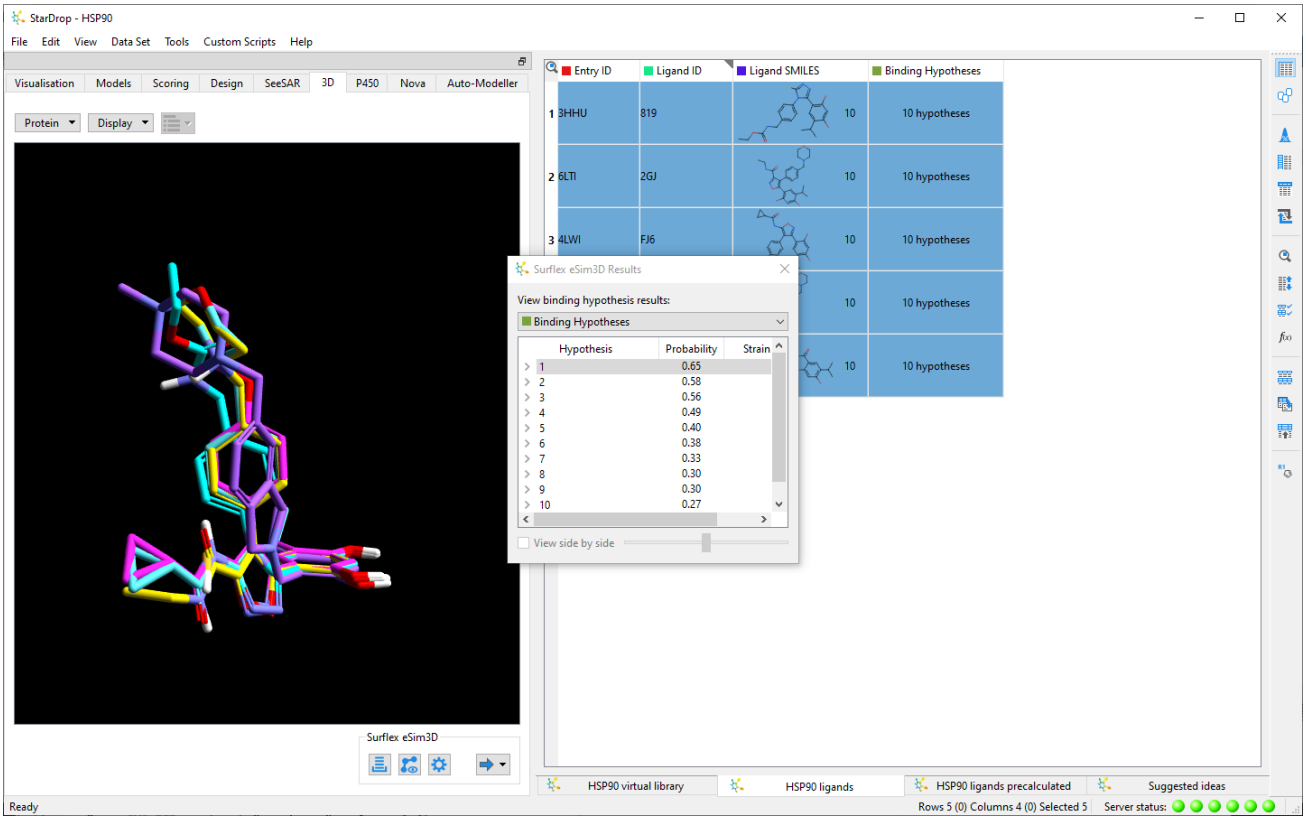


The parameter scheme specifies how intensive a search to conduct for binding hypotheses. Choosing an intensive calculation will lead to more accurate results but will take longer to compute. You can set additional configuration options by selecting **Advanced**, but in this case, we will use the default setting.

- Click **Finish** to start the generation of binding hypotheses.

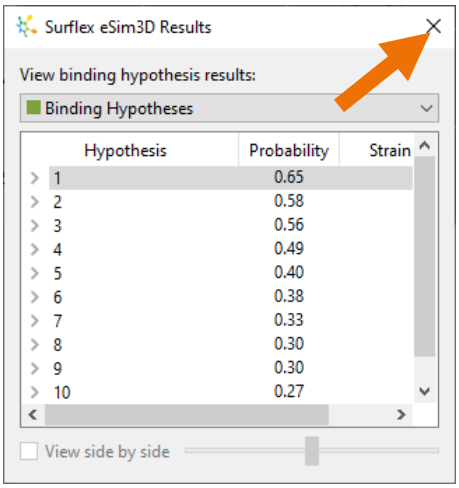
This will take a few minutes to complete. If you would not like to wait, you can view the results in the third data set of the project, called ‘HSP90 ligands precalculated’, by clicking this button  under the 3D viewer.

There are ten binding hypotheses generated, and the probability of each of the hypotheses being accurate is shown.



- Select the  button to bring up the Surflex eSim3D display options and explore the similarities and dissimilarities between these compounds.

We can see good alignment around the beta resorcylic acid region. Regions showing similarity in the shape or potential interactions between the ligands indicate likely steric constraints or conserved interactions. Areas of steric dissimilarity indicate regions where there is more freedom within the binding site. We now have a better reference for the virtual compound alignment as it is made up of several examples of HSP90 ligands. We will now compare our virtual compounds to the top binding hypothesis.

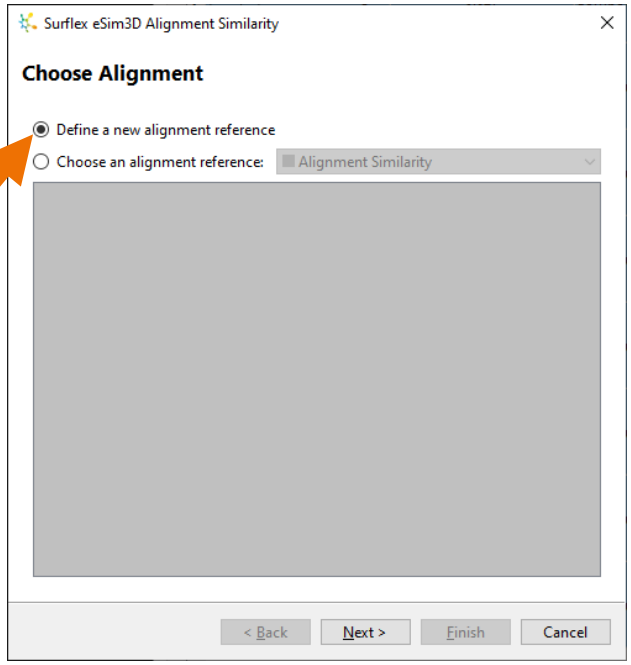


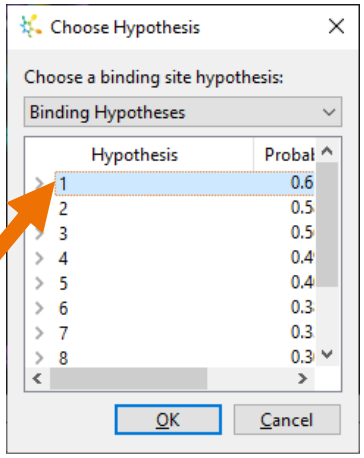
- Click on the cross at the top right of the binding hypothesis results to close them.

Note: the results can be accessed again by clicking this button  under the 3D viewer.

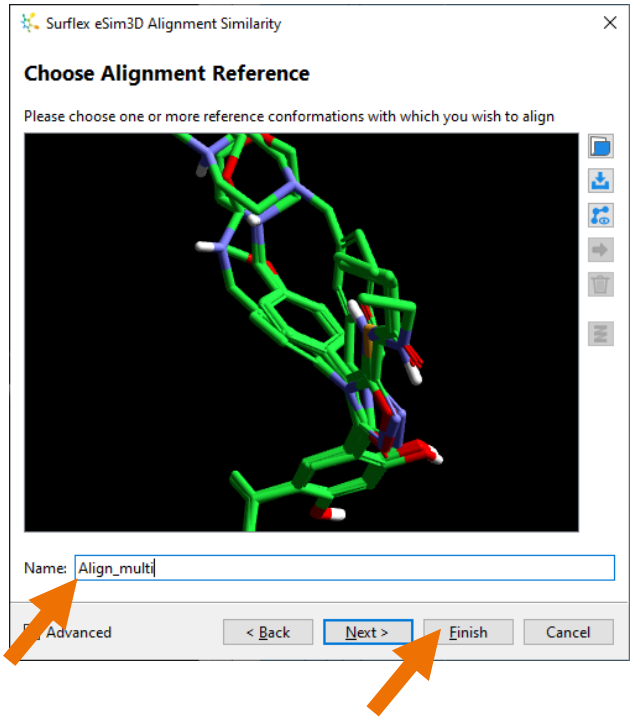
- Go back to the first data set, 'HSP90 virtual library'.
- Select all compounds by clicking in the top left of the data set.
- Click the **Go button menu**  and select **Calculate Alignment Similarity**.

- Select **Define a new alignment reference** and click **Next.**
- Click the  button to select from the binding hypotheses for a reference to align against.
- Choose the first binding hypothesis and click **OK**.



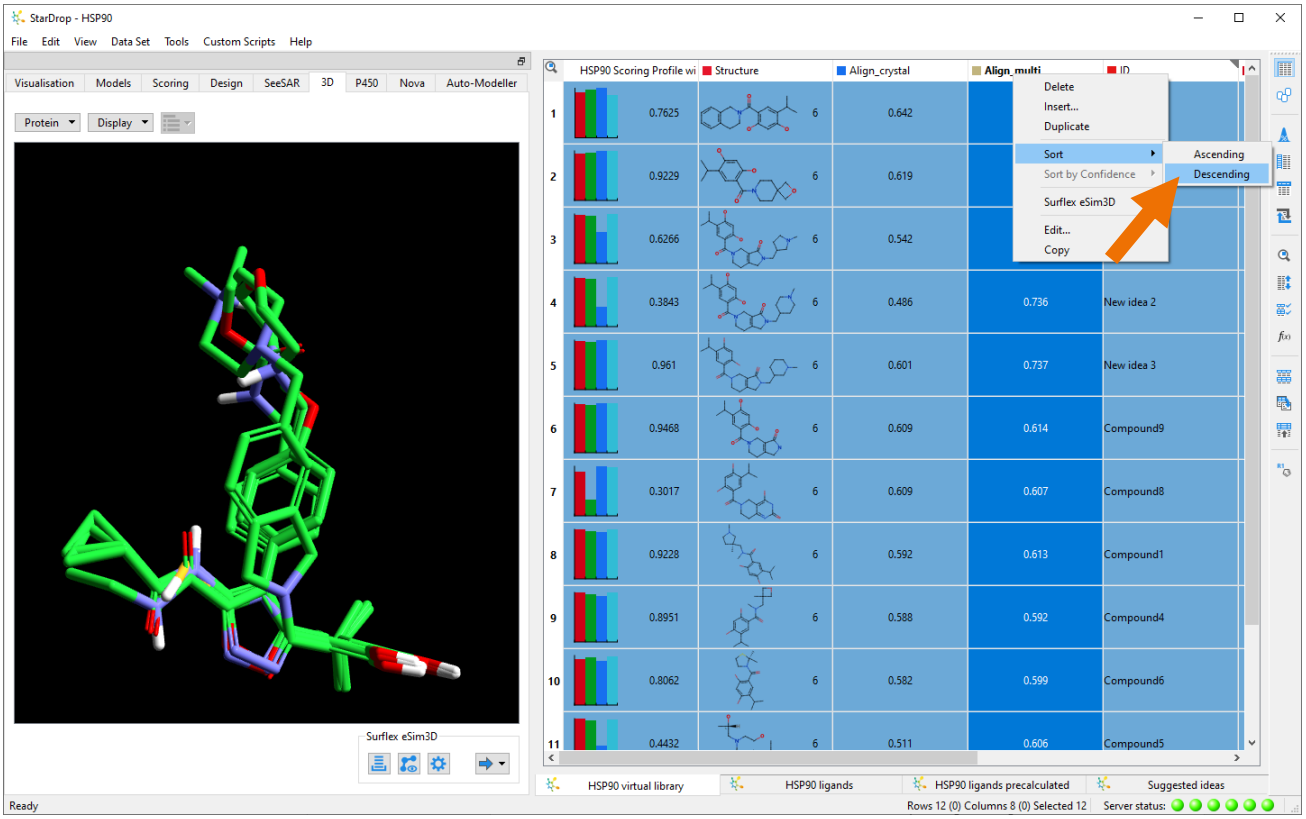


- Give the alignment the name ‘Align_multi’ and then click **Finish** in the main dialogue box.



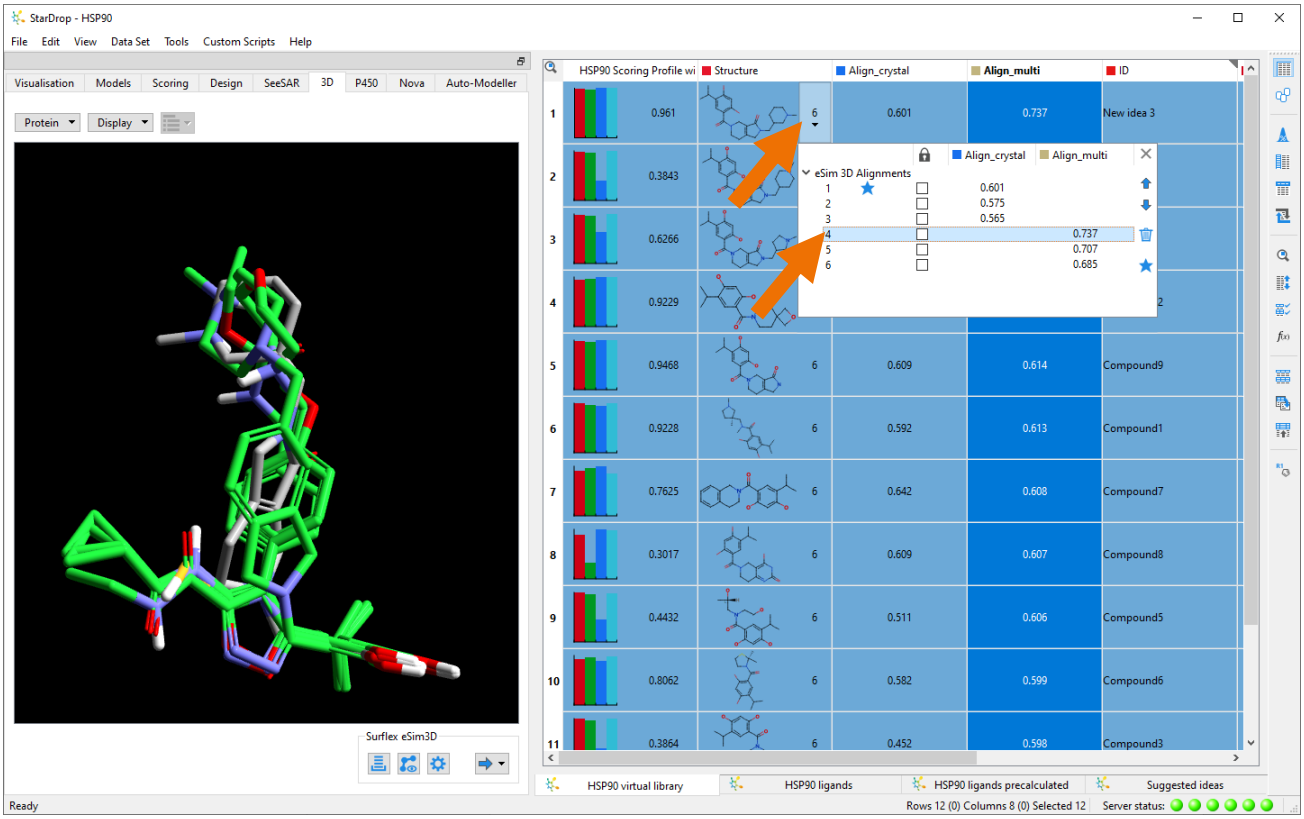
The calculated similarity is again shown in a new column, this time called ‘Align_multi’.

- Right-click on the **Align_multi** column to bring up the menu and choose **Sort**, then **Descending** to bring the compounds with the highest similarity values to the top.



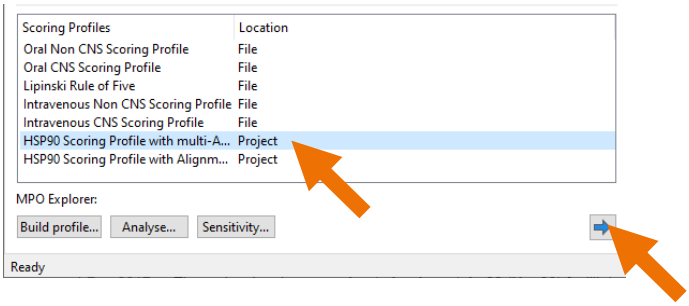
The ordering of the compounds has changed, and the top compounds are now the new ideas. We can see that the similarity score has increased considerably, showing that we are aligning well with the target containing multiple HSP90 ligands. Simultaneously considering different ligands enables us to consider a range of conformations and interactions that could be accommodated by the target binding site.

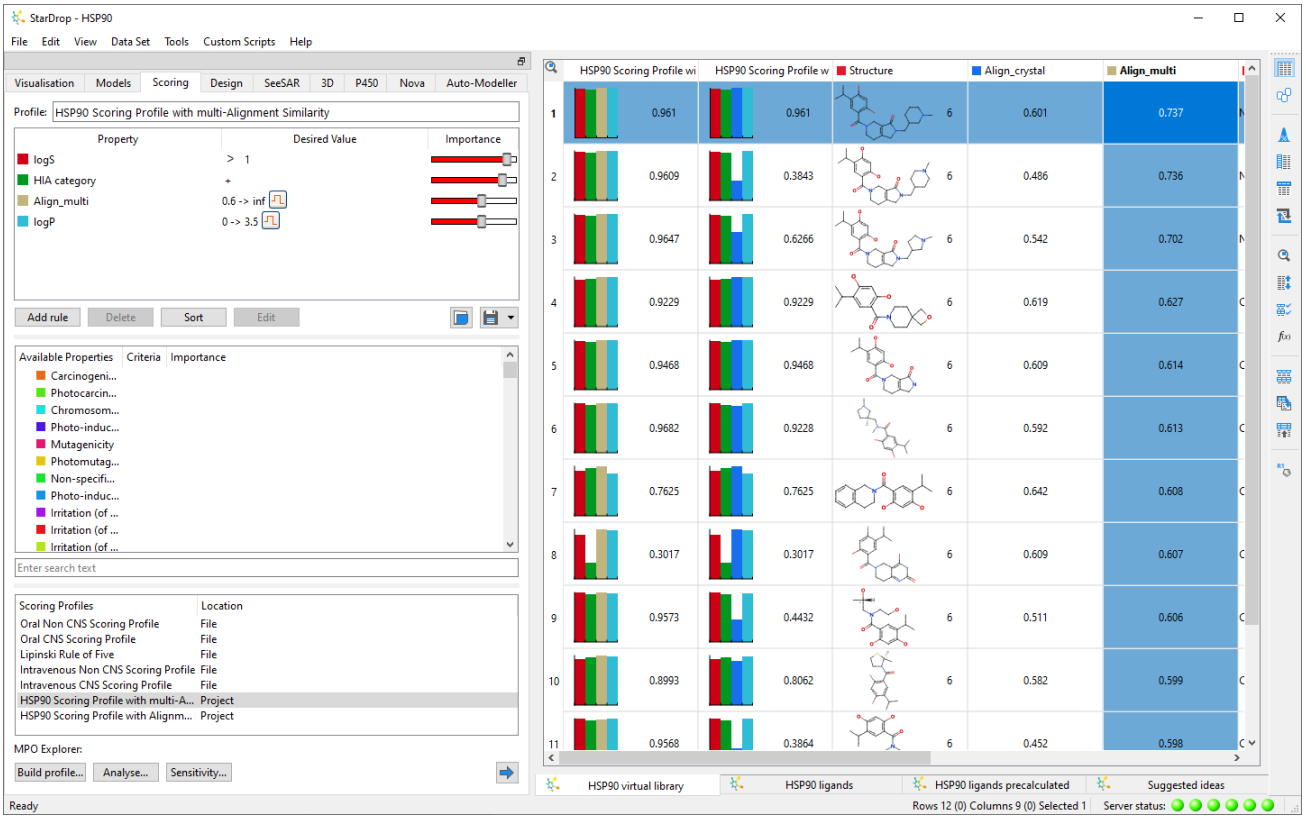
- Click on the compound in row 1 to show it in the 3D viewer.
- Click on the number of conformations next to the structure to display the drop-down menu showing the available alignments.
- Select the fourth conformation (the top-scoring alignment with the multi-ligand target).



The new compounds are likely to have improved potency over the original compounds due to their similarity to known HSP90 ligands. However, potency is not the only factor to consider when optimising compounds, so we will again use a scoring profile to look at the balance of properties.

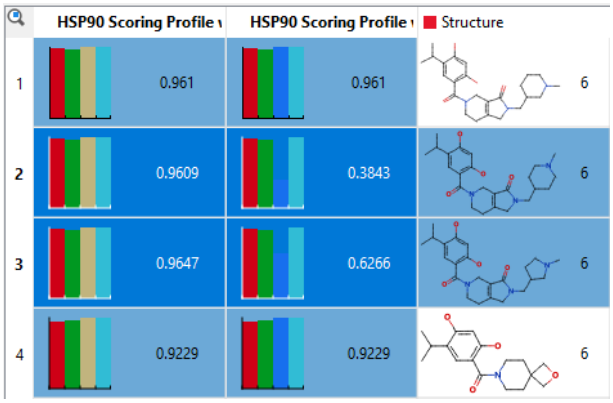
- Change to the **Scoring** area.
- Select the **HSP90 Scoring Profile with multi-Alignment Similarity** scoring profile in the list of available scoring profiles.
- Run it by clicking the Go button  .

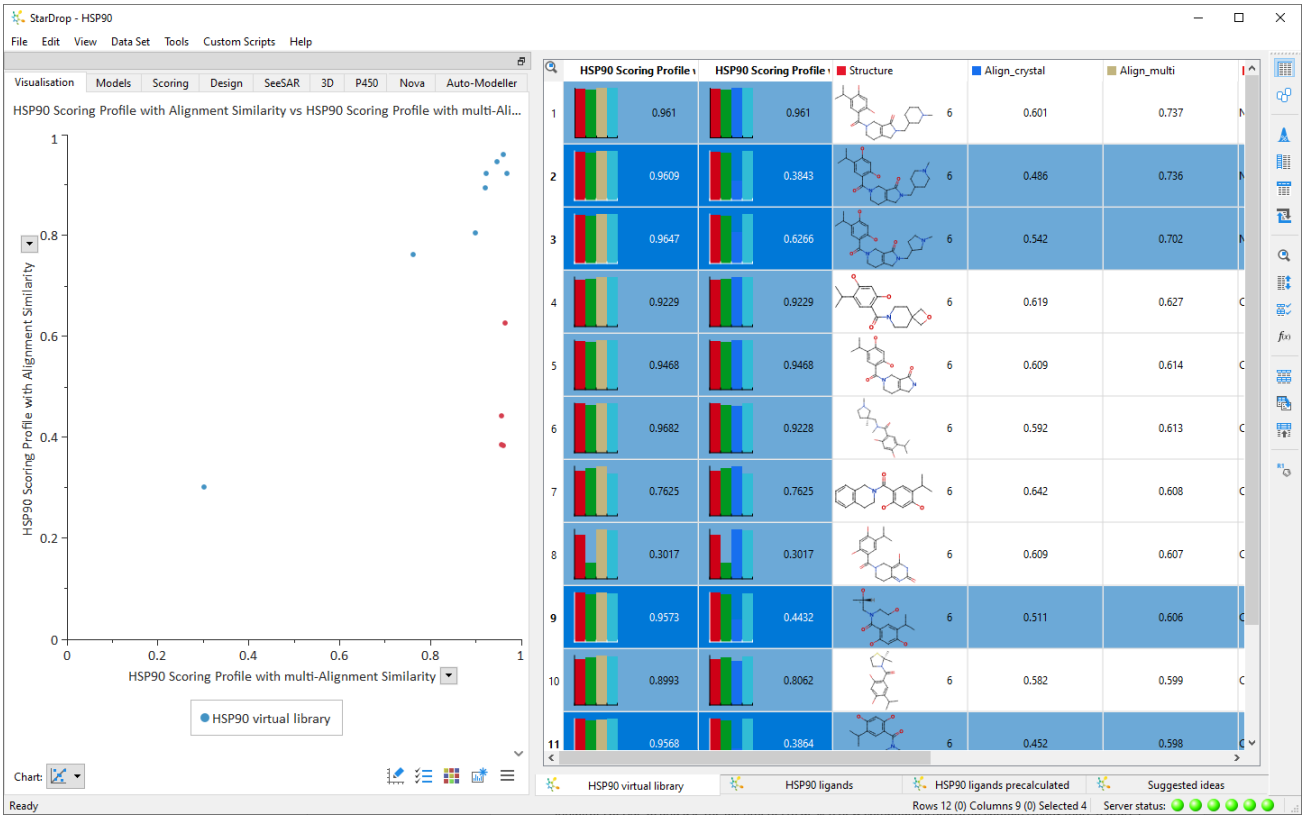




We can see that the scores for some compounds have changed considerably due to the better alignment with the multi-ligand binding hypothesis. We can compare these two multi-parameter scores in the Visualisation area.

- Change to the **Visualisation** area.
- Select the first scoring profile column by clicking on the column header.
- Hold down the Ctrl key if you are using Windows or the Command key if using a mac and select the second scoring profile column.





We can see that the suggested ideas compounds now score much better than with the single co-crystallised ligand of HSP90. We can draw around data points that differ considerably between the two scores. The alignment scores for Compound3, Compound5 and the three new designs improved, and this translates to better multi-parameter scores.

This example illustrates how we can use Surflex eSim3D to compare virtual compounds against known inhibitors of HSP90 and use the alignments to design new compounds. 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.