Worked Example:

Addressing Toxicity Risk in Multi-parameter Optimization

In this example we will illustrate how knowledge-based prediction of toxicity can be used within a multi-parameter optimization (MPO) environment to guide the selection and design of compounds with a good balance of properties and reduced risk of toxicity.

We will explore a library of compounds with experimental inhibition data against the cyclooxygenase 2 (COX2) target, with the goal of identifying a high quality lead series, comparable with the best-in-class Celecoxib (Celebrex). However, we should consider that other COX2 inhibitors, including Rofecoxib (Vioxx) and Valdecoxib (Bextra) were withdrawn from the market due to adverse events.

During this example, we will use StarDrop's Probabilistic Scoring approach for multi-parameter optimisation to integrate experimental data, predicted ADME properties from the ADME QSAR module and predictions of toxicity risk from the Derek Nexus module. We will explore the distribution of properties and scores across the chemical diversity of a screening library using StarDrop's chemical space visualisation to quickly target high quality chemistry. Finally, we will use the interactive design tool in StarDrop investigate approaches to modify compounds to reduce the risk of toxicity, guided by the Glowing Molecule visualisation. If you have any questions, please feel free to contact stardrop-support@optibrium.com.

Exercise

Compound Selection

- Start StarDrop from the Start menu.
- Select the **File->Open** menu option, select the data set **COX2 Library** and click **OK**.
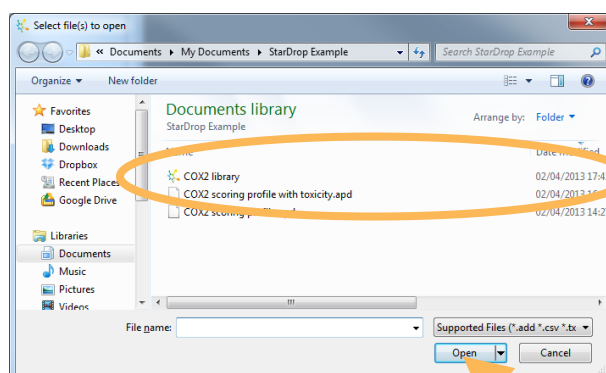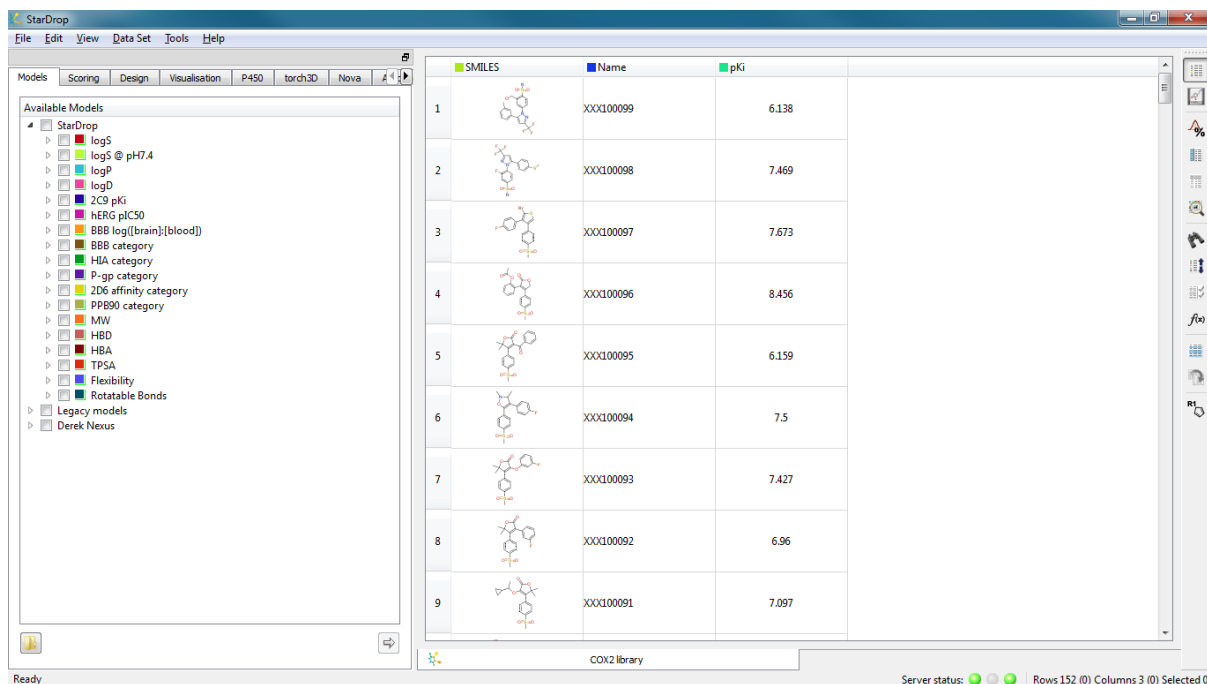
This will open a data set containing 152 compounds with potency data (pK$_i$) against the COX2 target. This represents a small library of compounds which have been screened for potency and from which we would like to identify a high quality lead series.

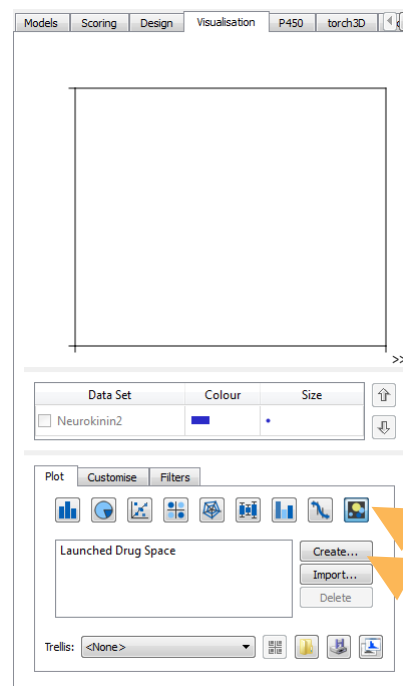- To explore the chemical diversity of this library and the distribution of potency across it, we are going to first take a look at a chemical space plot of this data. Click on the **Visualisation** tab on the left.

- Click on the ▣ button and then click "Create" to create a new chemical space plot.

- In the dialogue box that appears, type in a name for the chemical space plot you're about to create – in the example below we have called it **COX2 Space**.
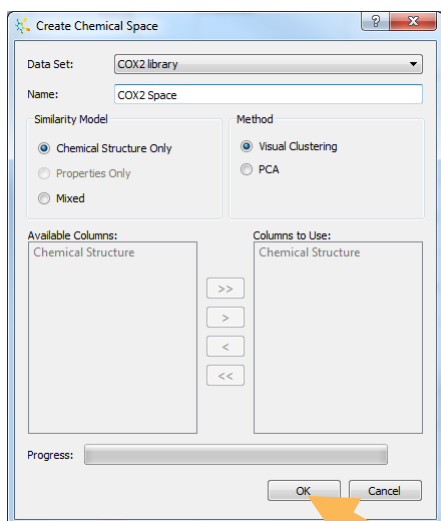
- We can build different types of chemical space plots based chemical structure and/or property data, but in this case we're going to use the defaults, **Chemical Structure Only** and the **Visual Clustering** method, as shown above.

- Click the **OK** button and wait until the chemical space as been created and the dialogue box disappears. When the process is a chemical space will be plotted in the **Visualisation** tab.
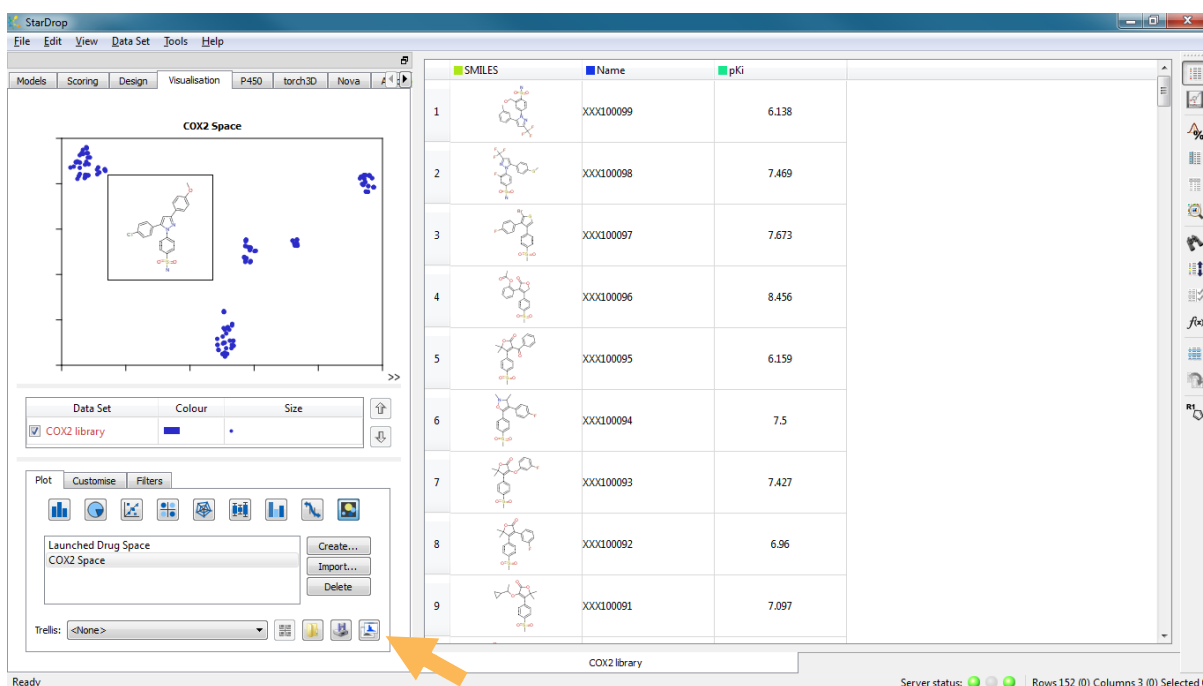
- In the **Visualisation tab**, click (as shown above) to create a separate window containing this chemical space. This will allow us to keep this chemical space plot visible as we explore the library.

▪ First we will colour the points by the experimental potency, to identify the most potent chemistries. In the detached plot, click on the colour block in the key and, in the **Format by Property** window, select **pKi** from the **Source** menu and click **OK**.
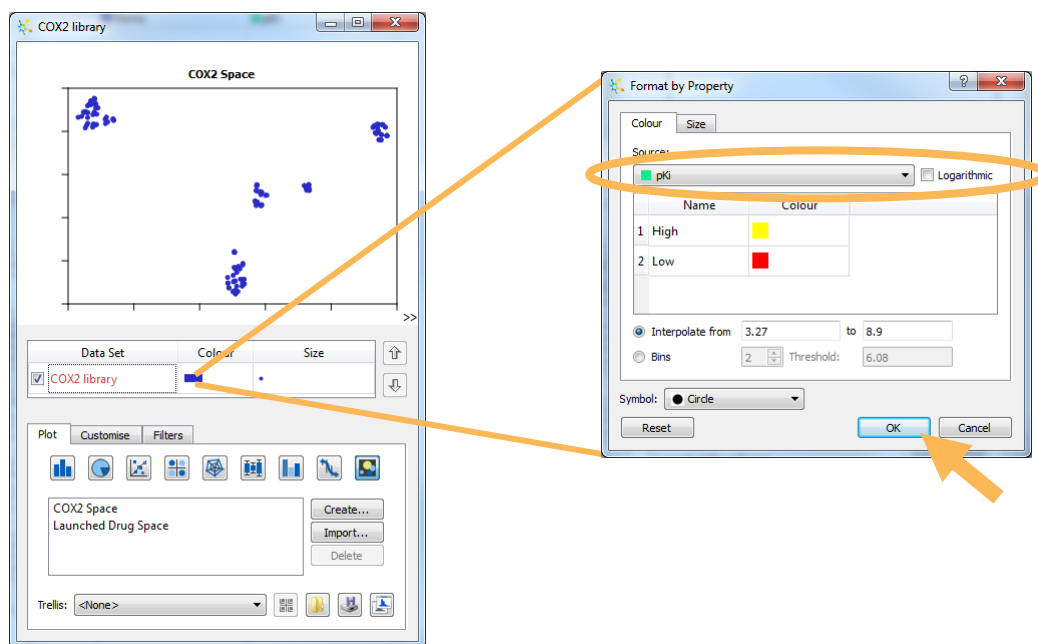
In the resulting plot, we can see the clusters containing the most potent compounds (by default colored in yellow). However, the most potent compounds are not necessarily the best starting point for a high quality lead series; ideally, we would like to identify a potent series that is also likely to have an appropriate balance of absorption, distribution, metabolism and elimination (ADME) properties. StarDrop's Probabilistic Scoring approach to MPO makes it easy to identify the chemistries that are most likely to achieve the required profile.
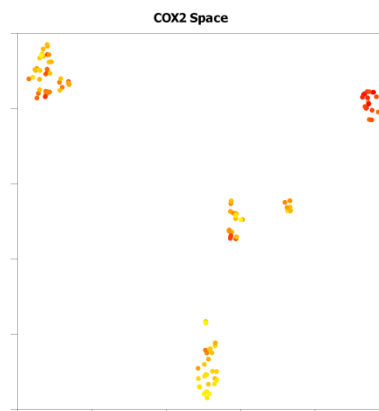
- Change to the **Scoring** tab in StarDrop, and click on the ▢ button to load a new scoring profile. In the **Load Profile** window, select the file **COX2 scoring profile.apd** and click **Open**.
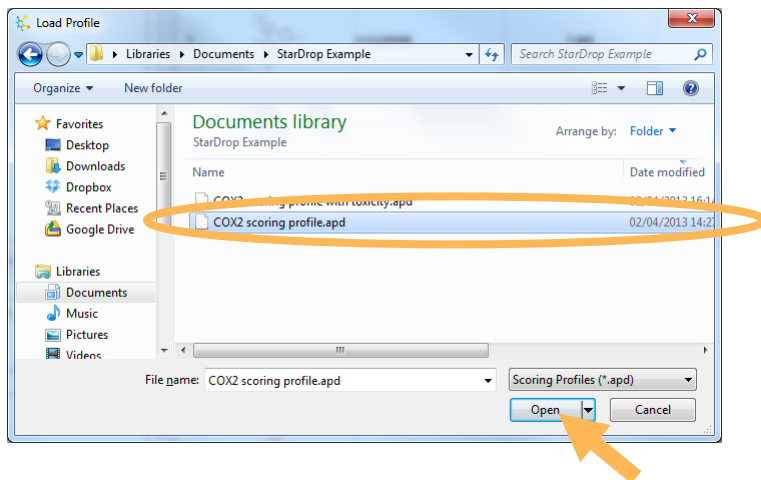
- The scoring profile will be displayed in the Scoring tab, as shown below. Click on the ⇨ button to calculate the score for each compound based on the experimental potency and predicted ADME properties. The models required to predict the ADME properties in the profile will be run automatically and the results will appear as columns in the data set.
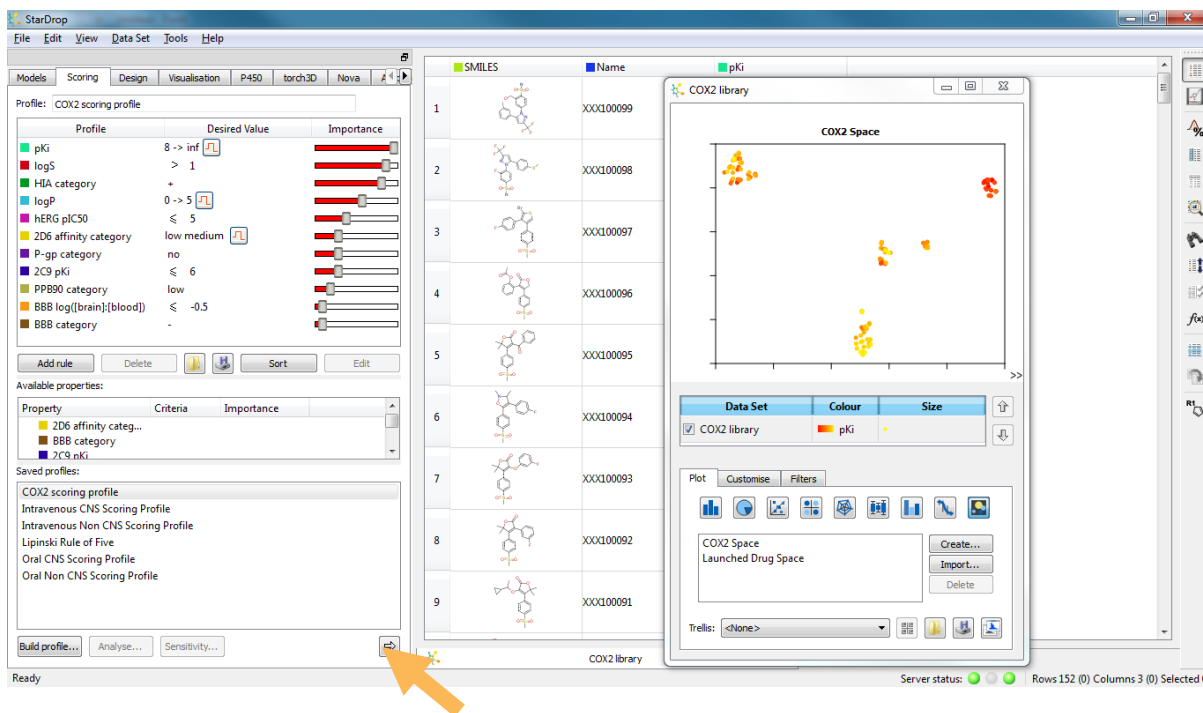
- Right-click on the COX2 scoring profile column header in the data set and select the **Sort->Descending** menu option. The highest scoring compounds will come to the top of the data set.
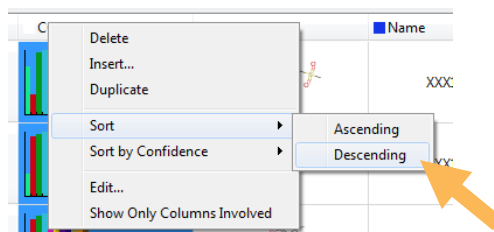
- We can color the points in the chemical space by the overall score. To do this, click on the color block in the key of the detached plot, select **COX2 scoring profile** from the **Source** menu in the **Format by Property** window and click **OK**, as shown below.
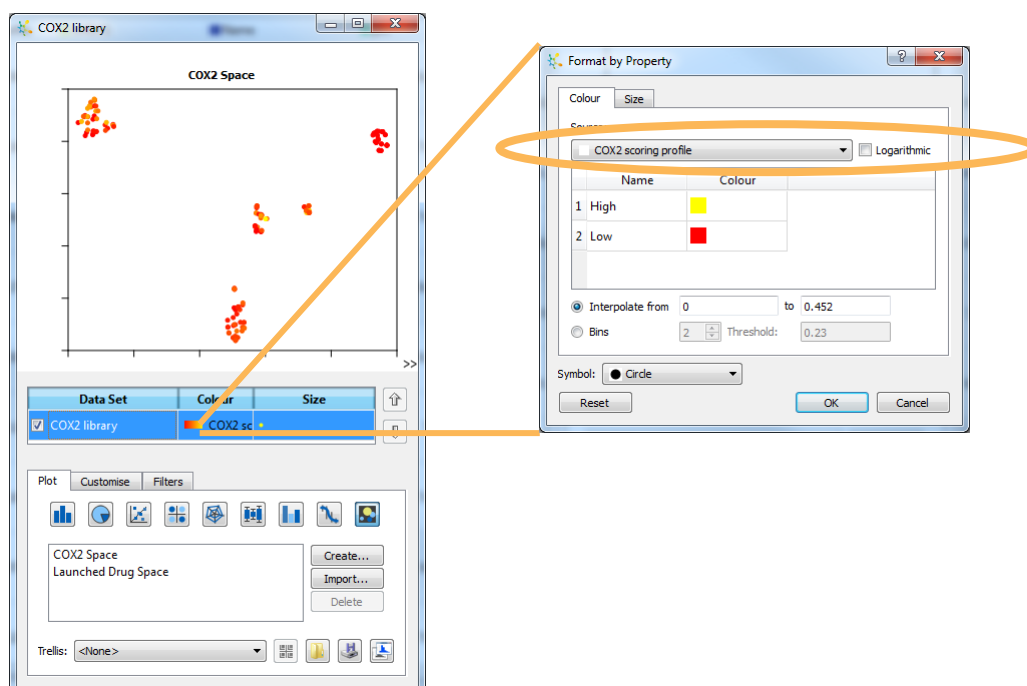

From this, we can see that the most potent cluster is now quite red, indicating that it is unlikely to have acceptable ADME properties. Alternative regions are highlighted in yellow and are more likely to have both good potency and ADME properties. Note that Celecoxib is one of the top-scoring compounds, giving us confidence that the scoring profile is prioritizing high quality compounds for COX2 inhibitors.

- Select row 7, containing Celecoxib, to highlight this compound in blue in the chemical space. Note that two other regions, circled to the right, are indicated as potential regions of high quality chemistry.
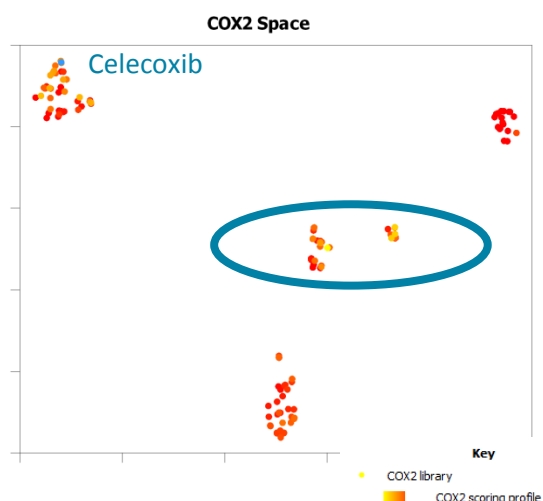
So-far we have not considered the potential for toxicity and for this we will use the **Derek Nexus** module for StarDrop.

- Change to the **Models** tab of StarDrop. At the bottom of the list of models you will see the **Derek Nexus** branch, which you can expand by clicking on the arrow (▷) adjacent to the name to see the full list of toxicity predictors provided by the Derek Nexus module. Select the **Carcinogenicity**, **Genotoxicity**, **Hepatotoxicity** and **Mutagenicity** endpoints, by checking the adjacent boxes, as shown below. Click the ⇨ button to run these predictions. This will take a few minutes...
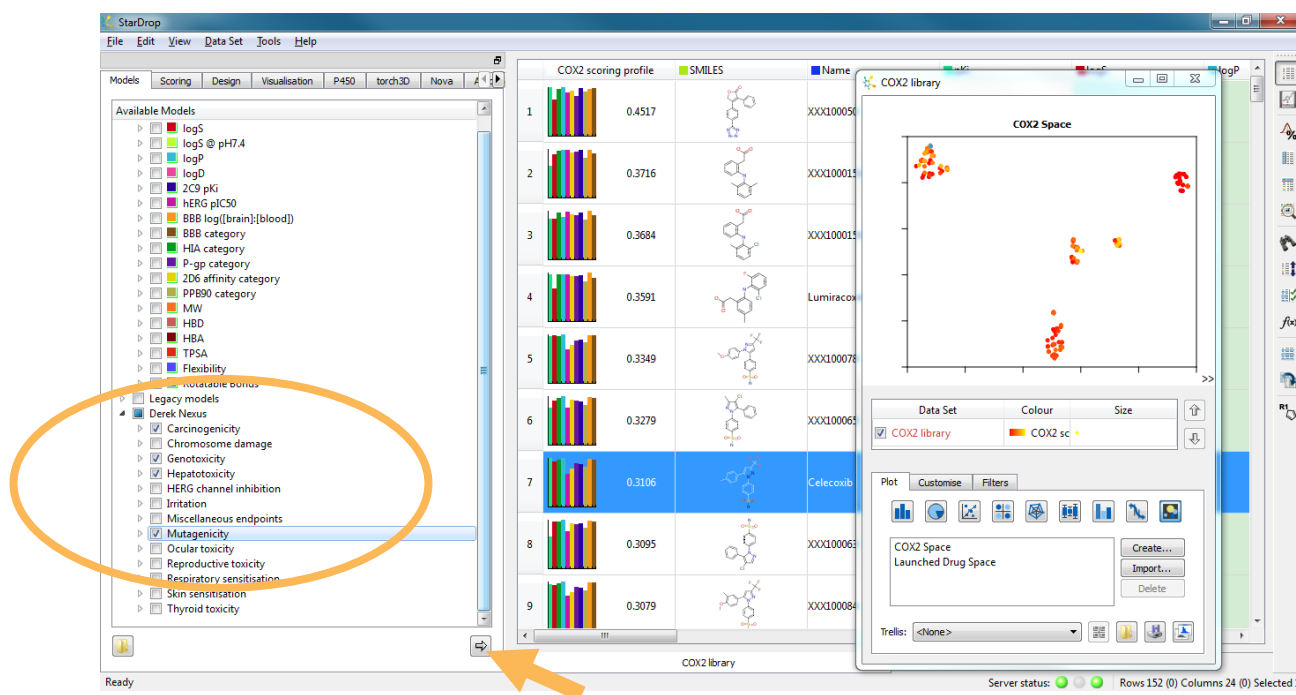
Scrolling the data set to the right will allow you to see the predictions for the individual endpoints in additional columns. These correspond to structural alerts with levels of likelihood based on precedence from experimental data. A **No report** outcome for an endpoint indicates that the query compound does not contain any structural alerts associated with that endpoint and there are no reasons based on the physical properties of the compound to predict either activity or inactivity. Other possible results, in ascending order of likelihood, are: **Equivocal**, **Plausible** and **Probable**. Alerts are highlighted in red, as shown below. From this we can see that there is an indication of the potential for toxicity of some of the top-scoring compounds.

We can visualize the results in the library's chemical space.

▪ In the chemical space plot window, click on the color block in the key (as we did previously) and select **Hepatotoxicity** from the **Source:** menu in the **Format by Property** window. In the resulting chemical space (shown right), we can see that some regions of the chemical space are highlighted as having a 'plausible' likelihood of hepatotoxicity, including one of the high scoring regions.
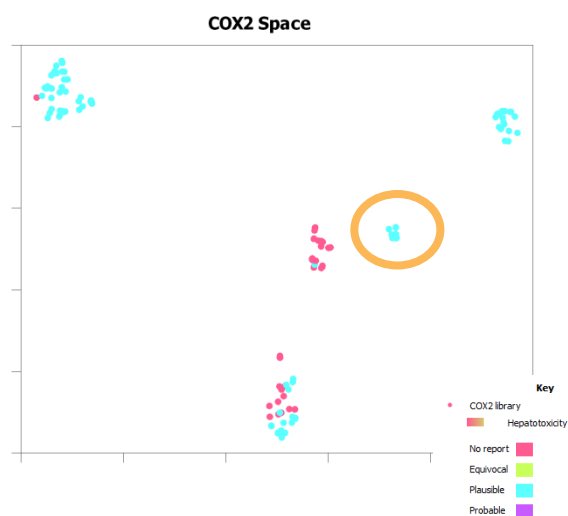
We can include the toxicity predictions in a scoring profile to prioritize potent compounds with a good balance of ADME properties and a reduced risk of toxicity.

- Change to the **Scoring** tab in StarDrop, and click on the 🗂 button to load a new scoring profile. In the **Load Profile** window, select the file **COX2 scoring profile.apd** and click **Open**.
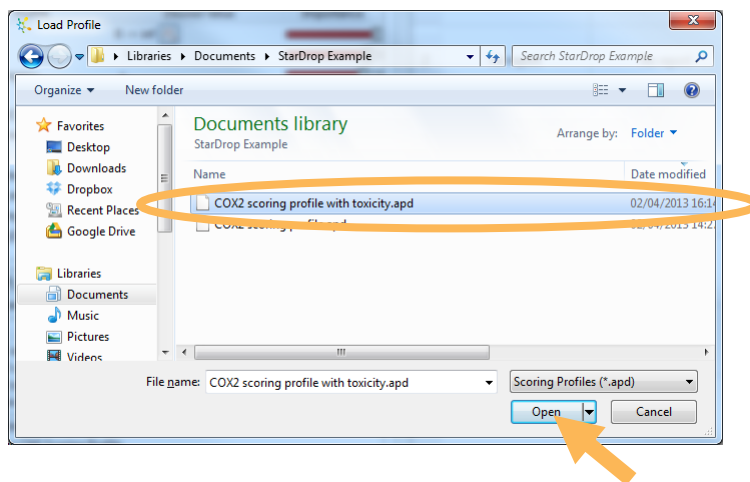
- The scoring profile will be displayed in the Scoring tab, as shown below. Click on the ➡ button to calculate the score for each compound based on the experimental potency and predicted ADME properties.
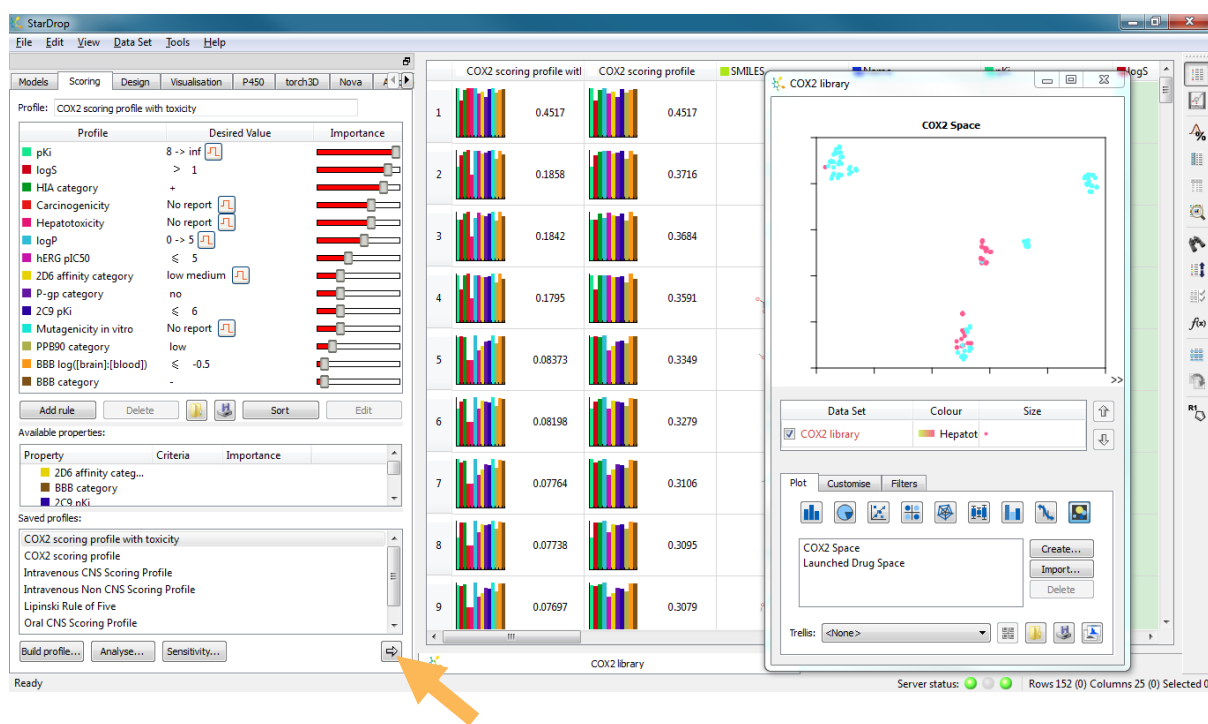
- By colouring the points in the chemical space using the new score calculated with the toxicity endpoints, in the same way as we did before, we can see that the region of the chemical space circled on the 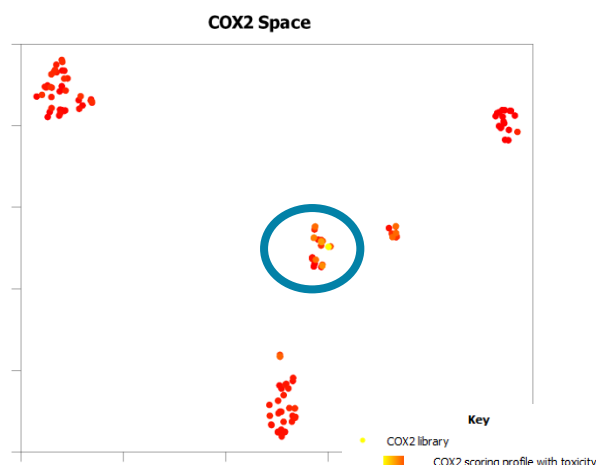right contains the majority of the top-scoring compounds. This indicates that this chemistry has the highest chance of providing a high quality series.

COX2 Space

Key
COX2 library
COX2 scoring profile with toxicity

You can find the corresponding points in the data set by drawing around the points in the chemical space while holding down the left mouse button.

Based on these results, an appropriate strategy would be to focus experimental and synthetic resources on the highest scoring compounds, identified above. However, given the potency and good properties of some alternative areas of chemistry, these should not be rejected outright. As reflected by their scores, predictions of toxicity and other properties are not definitive and therefore it would be appropriate to sample a small number of additional compounds for further tests to confirm the predicted hypothesis and avoid missing valuable opportunities due to uncertain data.
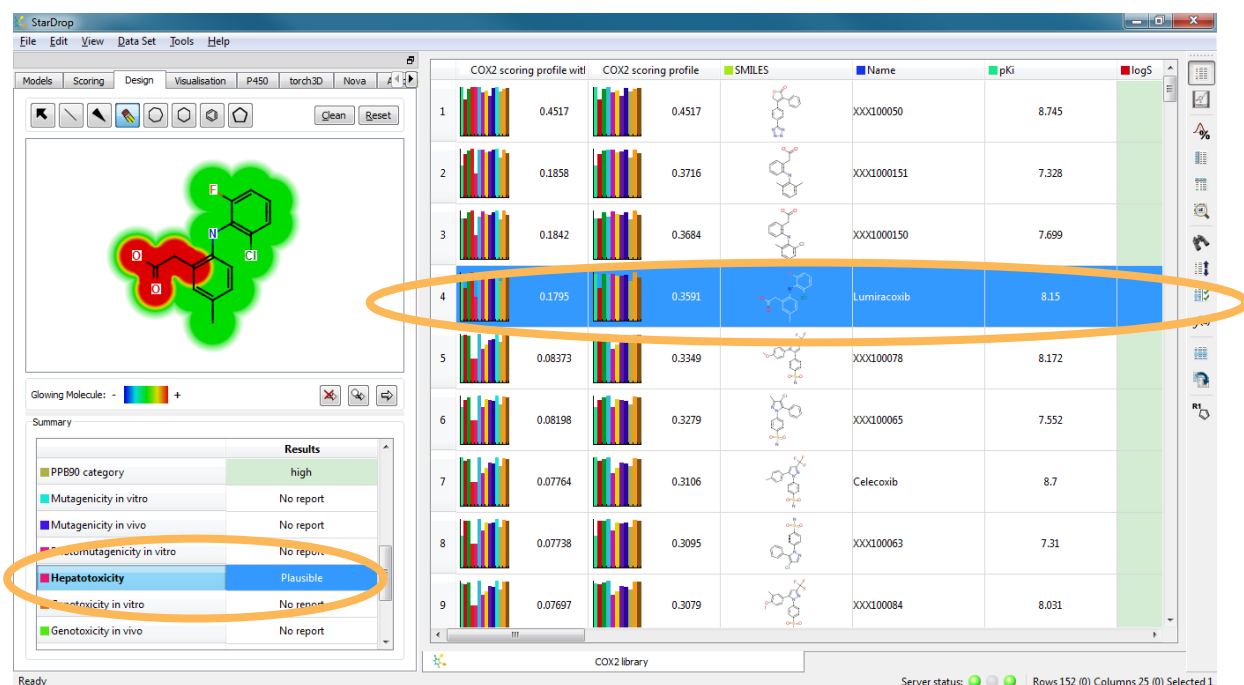
Compound Design

So-far we have been considering strategies for selection of compounds, taking into account the risk of toxicity. Now we will illustrate how we can explore strategies to redesign compounds and reduce the risk of toxicity.

- Draw around the cluster of compounds shown to the right, holding down the left mouse button, to select the corresponding compounds in the data set.

COX2 Space

Key
COX2 library
COX2 scoring profile with toxicity
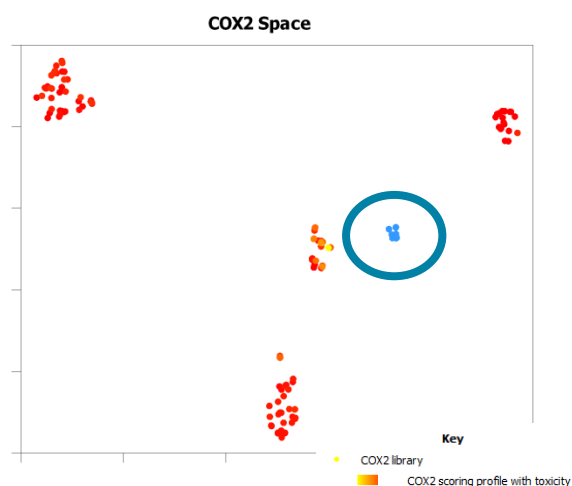
Note that one of the selected compounds is Lumiracoxib. This compound has been withdrawn from the market in several countries, mostly due to hepatotoxicity concerns, and has never been approved for use in the United States. We can use the interactive designer in StarDrop and the Glowing Molecule™ visualization, to identify the substructures causing alerts and guide efforts to redesign compounds with reduced risk of toxicity.

- Change to the **Design** tab in StarDrop and select the row containing Lumiracoxib, as shown below. The structure will be copied into the editor in the design tab. In the list of predicted properties below the editor, select **Hepatotoxicity** to highlight the structural alert for this endpoint.
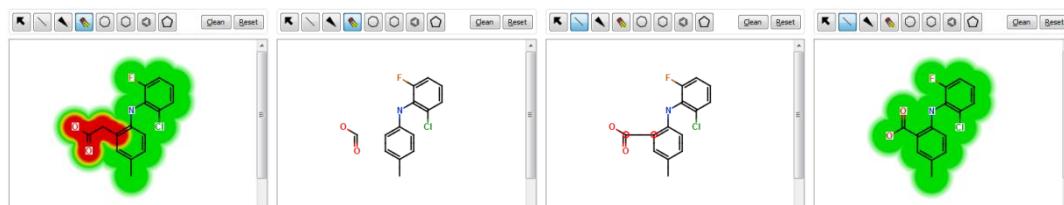

- Use the editor to modify the structure of the carboxymethyl. For example, use the eraser tool ( ) to delete the methylene and the bond tool ( ) to join the carboxyl group directly to the phenyl, as shown in the sequence below.

With each edit, all of the predicted properties in the table below will update instantly, allowing all of the properties to be monitored simultaneously.

- Finally, find compound XX100078 and use the designer to investigate strategies to reduce the risk of hepatotoxicity and carcinogenicity, in the same way as above.

Conclusion

This example has used some of StarDrop's capabilities to illustrate how knowledge-based prediction of toxicity, provided by the Derek Nexus module for StarDrop, can be used in early drug discovery to reduce the risk of encountering toxicity issues downstream. We have used chemical space visualization, probabilistic scoring and interactive design with Glowing Molecule™, supported by StarDrop's optional ADME QSAR module to predict key ADME properties.

If you have any questions, please feel free to contact stardrop-support@optibrium.com.