Worked Example:

Addressing Toxicity Risk in Multi-parameter Optimisation

In this example, we will illustrate how knowledge-based prediction of toxicity can be used within a multi-parameter optimisation (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 However, we should consider that other COX2 inhibitors, including Rofecoxib (Vioxx) and Valdecoxib (Bextra), were withdrawn from the market due to adverse events.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 target high-quality chemistry quickly. Finally, we will use the interactive design tool in StarDrop to 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

- In StarDrop, select **Open** from the **File** menu and open the StarDrop project file **COX2.sdproj**.
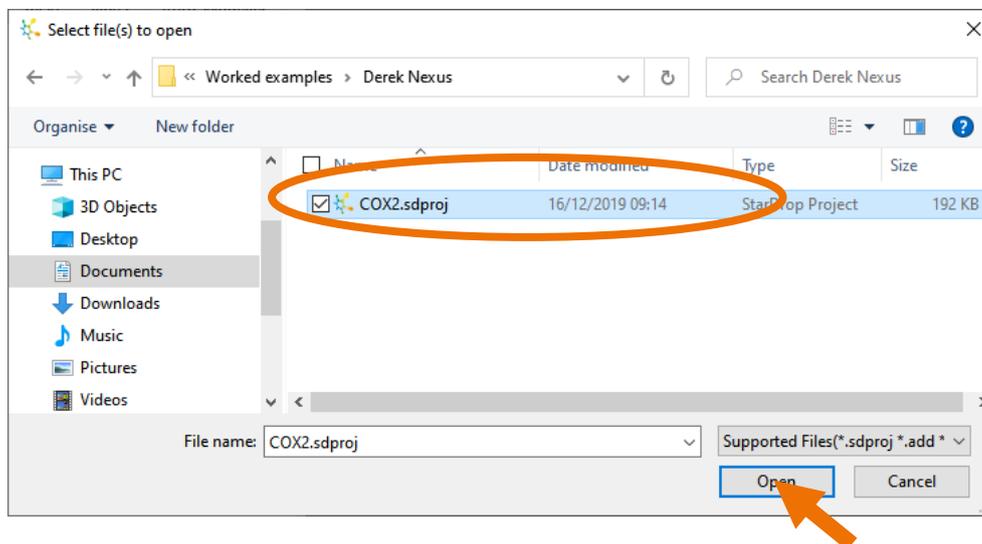
The project contains a data set of 152 compounds with potency data (pK$_i$) against the COX2 target. This represents a small library of compounds screened for potency and from which we would like to identify a high-quality lead series.
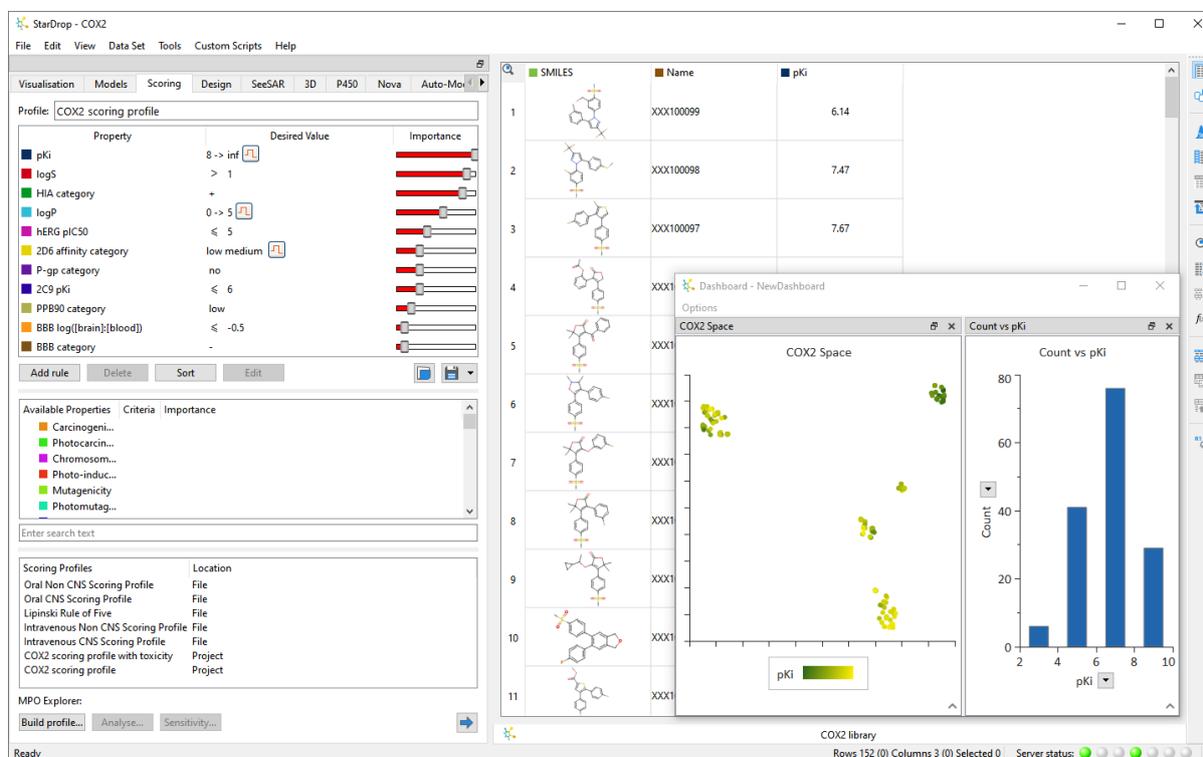
The project also contains a dashboard displaying two charts. On the right is a histogram showing the distribution of pK_i values across the data set. On the left is a chemical space representation of the compounds. Each point in the chemical space represents one compound, and hovering the mouse over a point will display a pop-up of the structure. Any selections made in the data set or the charts will be reflected everywhere else.

- Select the right-hand histogram bar to select the compounds with the highest pK_i values.
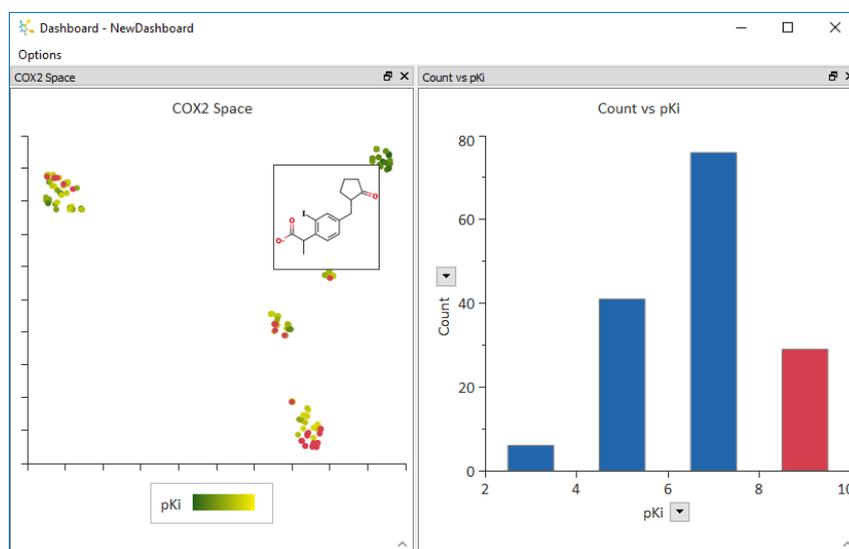
In the chemical space, we can see the clusters containing the most potent compounds (by default coloured in yellow). However, the most potent compounds are not necessarily the best starting point for a high-quality lead series; ideally, we would like to identify a potent series that is also likely to 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.

In the **Scoring** area, the project contains a **COX2 scoring profile** that we have already created.

- Click on the button at the bottom of the **Scoring** area 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 **Descending** from the **Sort** menu. The highest scoring compounds will move to the top of the data set.

We can colour the points in the chemical space by the overall score.

- Click somewhere in the chemical space where there are no compounds to clear the selection.

- Click the arrow in the bottom-right corner of the chemical space to display the controls and then click on the **Format** button .

- In the **Format Chart Data** dialogue, choose to **Colour by** the **COX2 scoring profile** and then click the **Close** button.
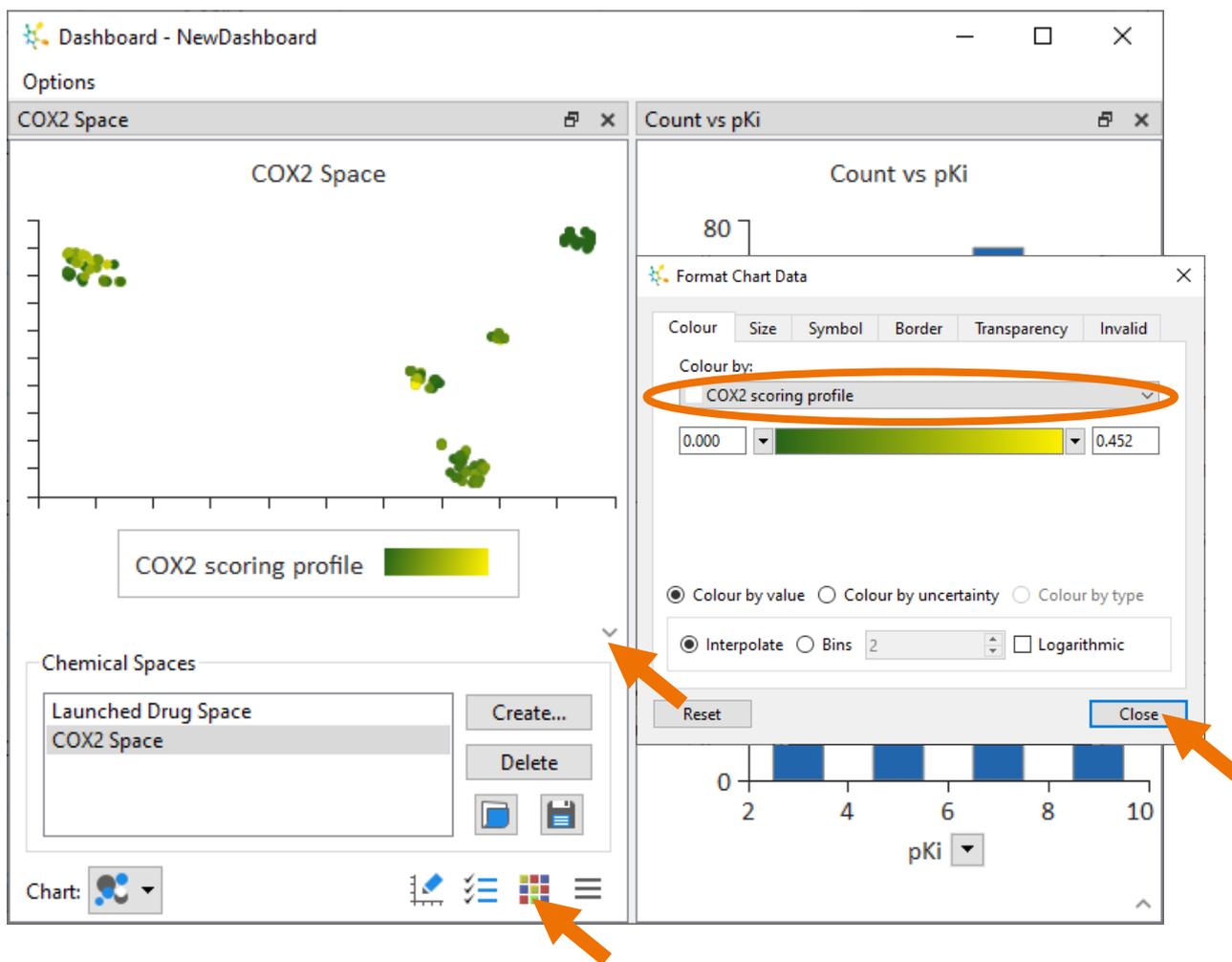
- Click the small arrow below the chemical space again to hide the controls.

From this, we can see that the most potent cluster is no longer yellow, indicating that it is unlikely to have acceptable ADME properties. Alternative regions have more yellow points and are more likely to have good potency and good ADME properties. Note that Celecoxib is one of the top-scoring compounds, giving us confidence that the scoring profile prioritises high-quality compounds for COX2 inhibitors.

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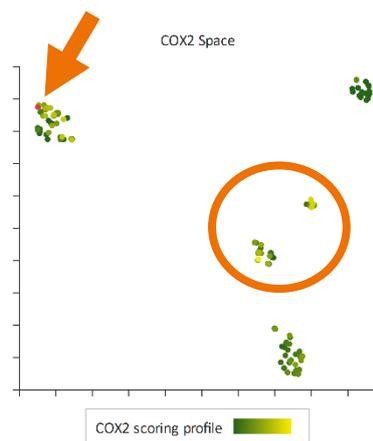
So far, we have not considered the potential for toxicity, and for this, we will use the **Derek Nexus** module for StarDrop.

- Click on the **Models** tab.

At the bottom of the list of models, you will see the **Derek Nexus** branch of models.

- From the list, select the **Carcinogenicity**, **Genotoxicity** and **Hepatotoxicity** endpoints by checking the adjacent boxes, as shown below (**Note:** Hepatotoxicity is in the **Organ toxicity** branch under **Derek Nexus**).

- Click the ⮕ button to run these predictions. This will take a few minutes, and although the columns will be created immediately, the results will be filled in as they are returned from the server.
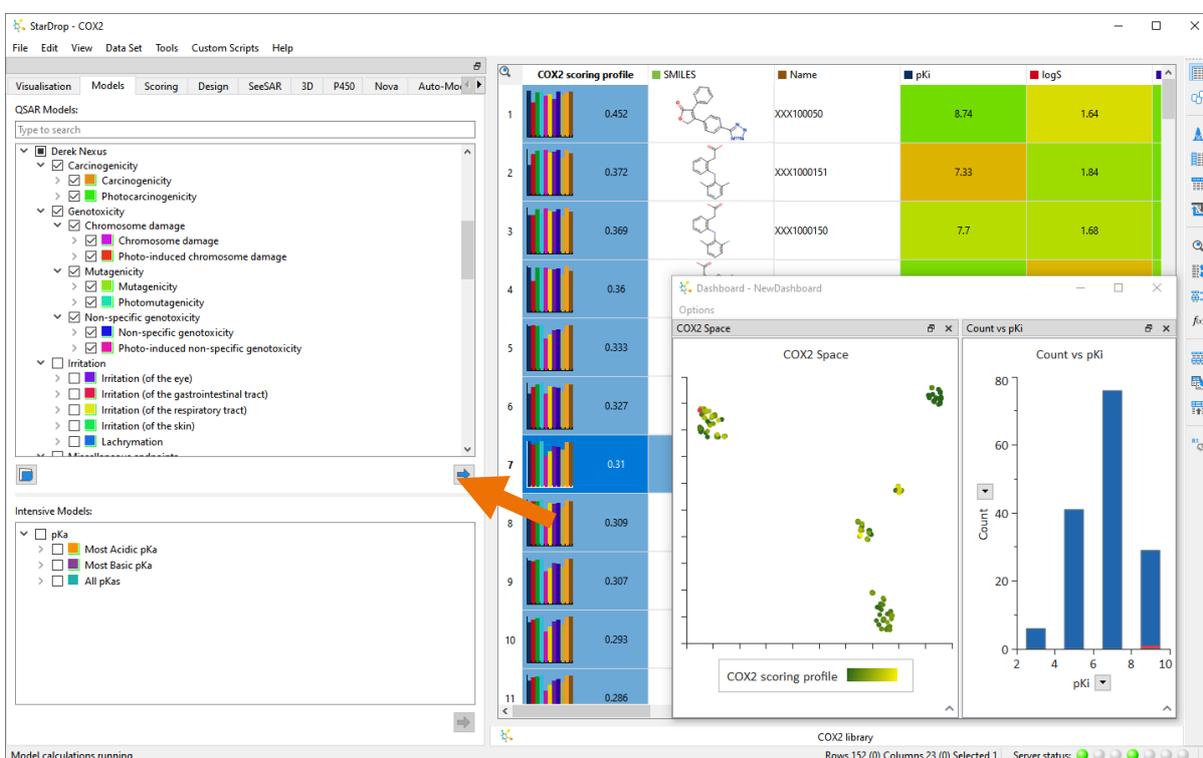
- Minimise the dashboard and scroll across the data set to see the predictions for the individual endpoints.

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. In ascending order of likelihood, other possible results are: **Inactive**, **Equivocal**, **Plausible** and **Probable**. Alerts are highlighted in orange, 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 visualise the results in the library's chemical space.

- Restore the dashboard.

- In the chemical space plot window, open the
  controls again and change the formatting to
  colour by the **Hepatotoxicity**.

COX2 Space

In the resulting chemical space (shown right), we can
see that some regions of the chemical space are
highlighted as having an 'equivocal' or 'plausible'
likelihood of hepatotoxicity, including one of the high
scoring regions.

We can include the toxicity predictions in a scoring
profile to prioritise potent compounds with a good
balance of ADME properties and a reduced risk of toxicity.

- Change to the **Scoring** area again and select the file **COX2 scoring profile with toxicity** from
  the list at the bottom. This is another scoring profile that we have already saved as part of the
  project.

- Click on the [→] button to calculate the score for each compound based on the experimental
  potency and predicted ADME properties and toxicities.

Note: You will see a warning about some of the data having zero uncertainty but click the **OK** button.
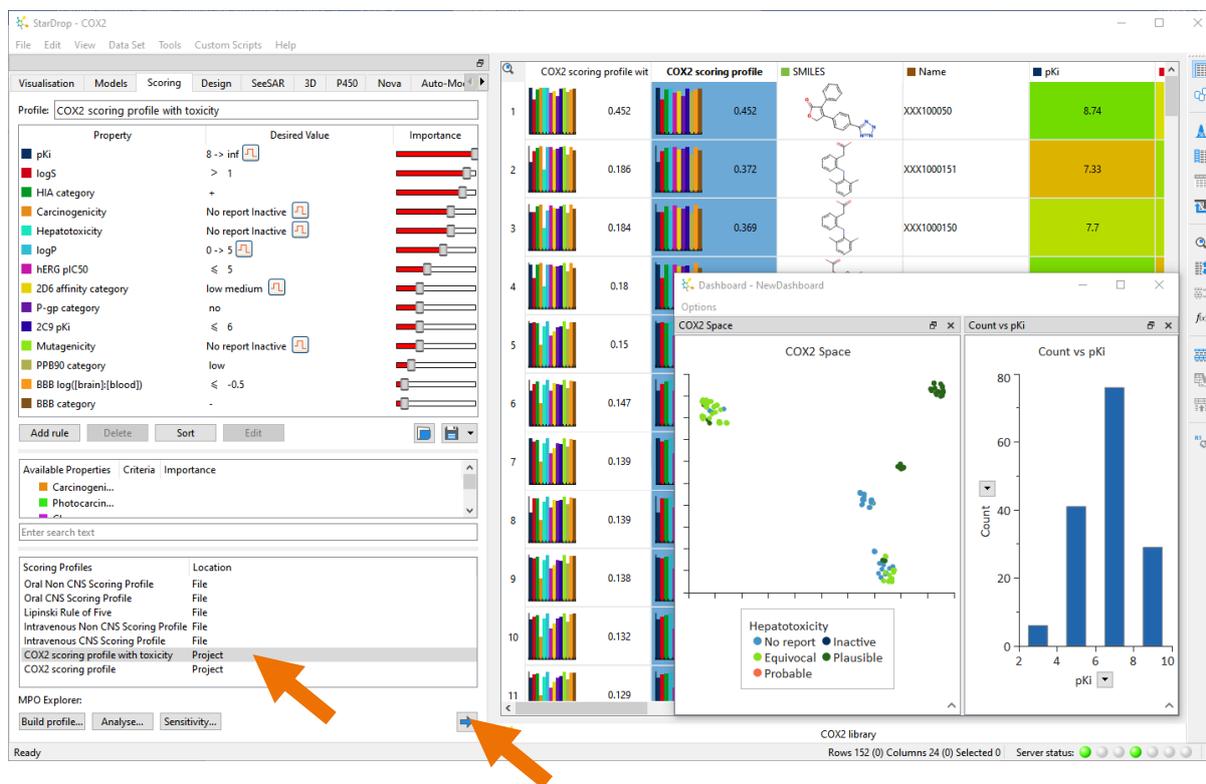
- Right-click on the header of the new score column and sort the data set in descending order again.
- Colour the points in the chemical space using the new score calculated with the toxicity endpoints in the same way as we did before.

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

COX2 scoring profile with toxicity
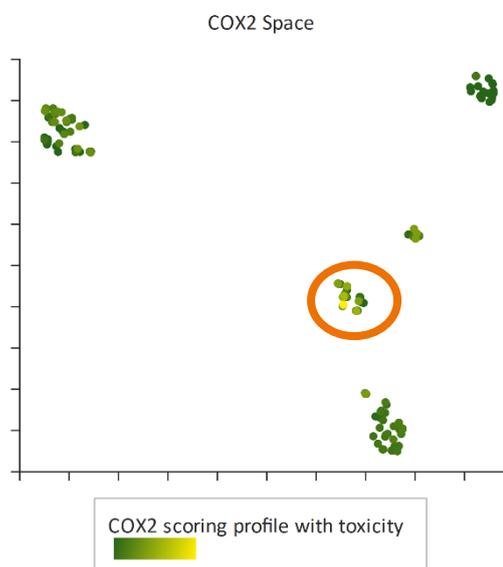
You can find the corresponding points in the data set by drawing around the points in the chemical space with the mouse.

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 compound selection, 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.

COX2 Space

COX2 scoring profile with toxicity

- Draw around the cluster of compounds shown to the right to select the corresponding compounds in the data set.

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 and the Glowing Molecule™ visualisation to identify the substructures causing alerts and guide efforts to redesign compounds with reduced toxicity risk.

- Minimise the dashboard.
- Change to the **Design** area and select the row containing Lumiracoxib (row 10).

The structure will be copied into the editor in the Design area.

- In the list of predicted properties below the editor, select **Hepatotoxicity** to highlight the structural alert for this endpoint.

We can use the editor to modify the structure of the carboxymethyl.

- Use the **eraser** tool to delete the methylene and then use the **bond** tool to join the carboxyl group directly to the phenyl before clicking the **Clean** button, as shown in the sequence below.

With each edit, all the predicted properties in the table below will update instantly, allowing all 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 visualisation, 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.