StarDrop 8

StarDrop™ Worked Example: Inspyra

Contents

Introduction

Anagliptin is a member of the 'gliptin' family of dipeptidyl peptidase-4 (DPP-4) inhibitors. It has been approved in Japan since 2012 for the treatment of type II diabetes mellitus. However, it was withdrawn from other markets due to observed animal toxicity, raising concerns, since disproved, about selectivity over the DPP-2, DPP-8, and DPP-9 receptors.

The lack of differentiation between the DPP-4, DPP-2, DPP-8, and DPP-9 binding pockets suggests that a structure-based approach to optimising selectivity would be challenging; hence, a ligand-based approach using QSAR modelling may provide an alternative.

This example uses Inspyra™ in StarDrop™ to interactively explore optimisation strategies to identify a selective DPP-4 inhibitor with appropriate physicochemical properties. We will start from a data set of compounds representing four chemical series based on pyrrolidine, cyanopyrrolidine, piperazine, and 1,4-diazepane cores. We will explore approaches to expand on the diversity of compounds in the data set and consider design strategies around a specific compound of interest.

Loading data and model files

We will start by loading QSAR models for binding affinity at the four receptor isoforms in this study so that we can evaluate how likely new compounds will be to bind to each receptor. We can use the StarDrop Auto-Modeller™ to build a QSAR models based on our existing data; we have not discussed the Auto-Modeller as part of this exercise, but if you would like more information, please contact support@optibrium.com.

- From StarDrop's **File** menu, choose **Open Data Set**. Go to File name, select 'All Files(*.*)', ,and load QSAR models of DPP-2, DPP-4, DPP-8, and DPP-9 activities by selecting the corresponding model .aim files.

Hint: you can select all four models by holding down the **Ctrl** key on Windows or the **Command key** on macOS.

- Changing to the **Models** area, you can see that these models have been added to the list of **Custom** models under the name of the directory from which they were loaded.
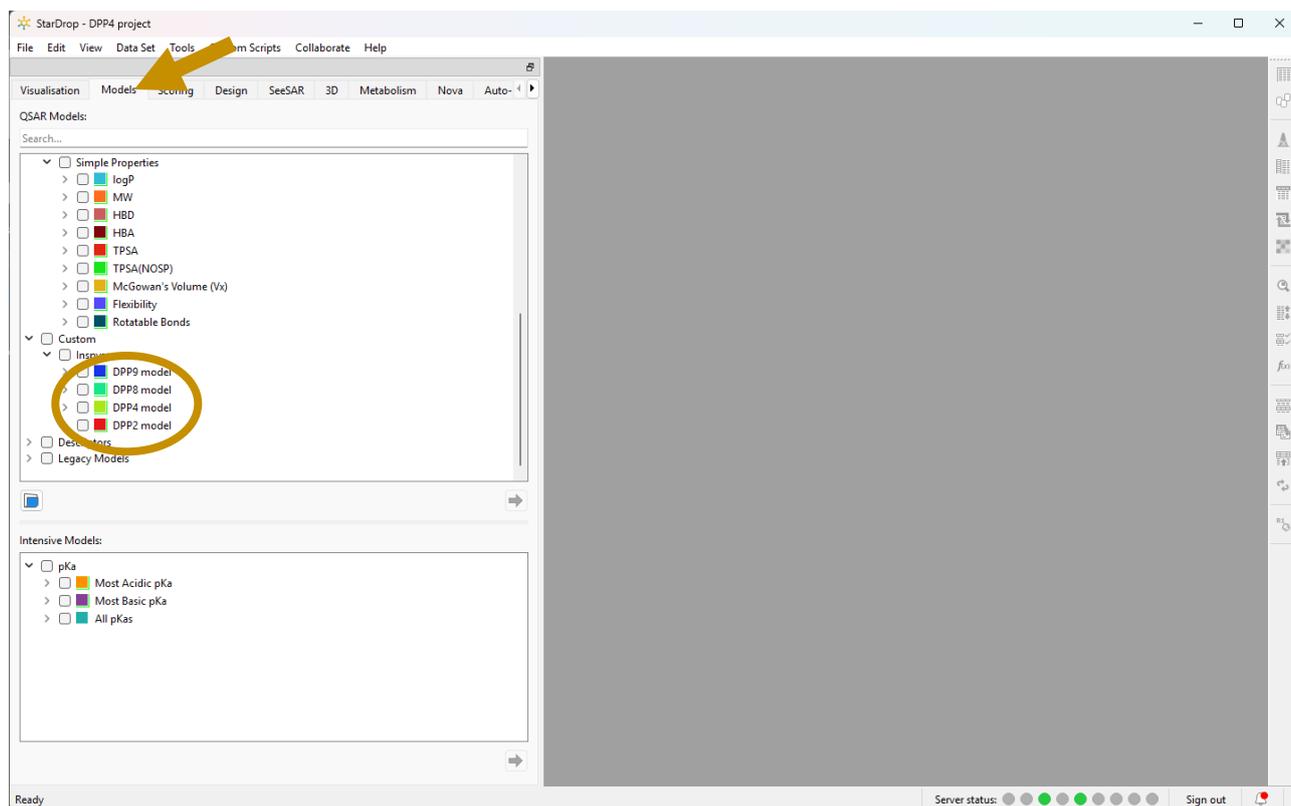
- Now open the StarDrop file **DPP4 project.sdproj** by selecting **Open** from the **File** menu, and navigating to the location where the file was saved.

This workspace contains a data set called DPP4 Series, with 206 compounds from four series, and experimental data for activities against DPP-4, -2, -8, and -9, where measured. The data set also includes predictions for the pIC50 values for these four isoforms, using the models we have just loaded and some simple physicochemical and ADMET properties. Furthermore, the compounds have been scored using a multi-parameter profile for a selective DPP-4 inhibitor with appropriate ADME and physicochemical properties: good solubility and human intestinal absorption (HIA), lipophilicity (logP) in an appropriate range and good selectivity over inhibition of the hERG ion channel.
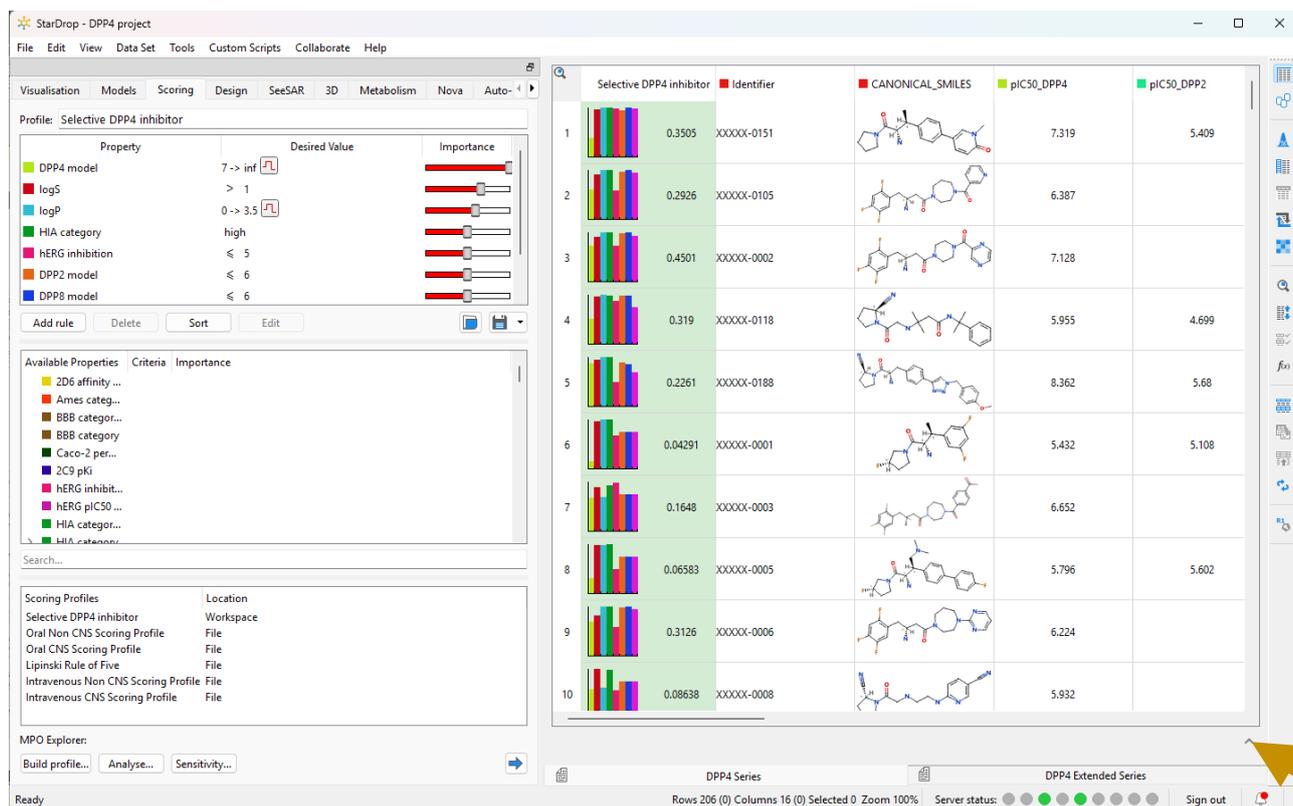
Inspyra Panel

We will use the **Inspyra Panel** to explore potential optimisation strategies based on the compounds in this data set.

- To open the **Inspyra Panel**, click on the ∧ button in the bottom-right of the data set, as indicated above.
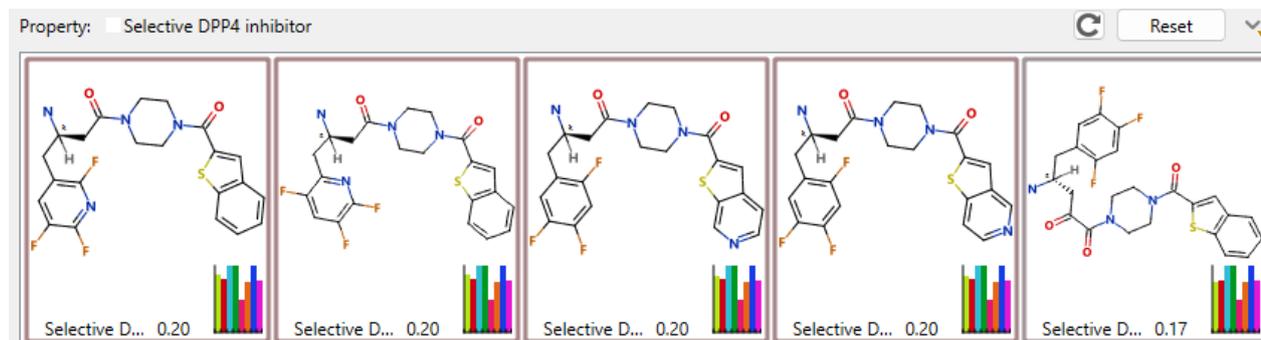
The **Inspyra Panel** shows five ideas generated from the compounds in the data set. By default, the property chosen for optimisation is the **Selective DPP4 inhibitor** score. For each compound idea, the score is shown, along with a histogram indicating each property's impact on the score. This can be changed by selecting a different option from the **Property** menu above the panel.

A **red** border around an idea indicates that the property being optimised (in this case, the score) is higher than the average for the data set; a **blue** border indicates that it is lower than average. The intensity of the colour indicates the magnitude of the increase or decrease.

A new suggestion will replace one idea in the panel every 30 seconds.

You can interact with an idea in the panel to provide feedback.

The **pin** button in the bottom-left will prevent the idea from being replaced so that you can consider it for longer.

In the top left, buttons enable you to provide negative feedback on the idea. The **dislike** button provides strong feedback that you dislike the properties or structure of the compound. The **ignore** button provides feedback that you are not interested in the idea, even if there is not a strong reason to dislike the compound.

The buttons in the top right enable you to provide positive feedback. The **like** button indicates that this is an interesting idea, but not good enough to retain. Finally, the **select** button will add the idea to the data set and provide strong positive feedback.

Clicking on any feedback buttons will replace the idea with a new one. If you have added the idea to your data set, this compound will also be used to generate new ideas.

As you interact with the ideas in the panel, Inspyra will learn from your preferences regarding property ranges and structural diversity and, increasingly, suggest more ideas that match your preferences (assuming that suitable compounds can be generated). For more information about the inference engine and how it responds to our interactions, please see Section 11.3 of the StarDrop Reference Guide, available from the **Reference Guide** option in the **Help** menu.

In this case, we are interested in compound ideas that increase the overall score, indicating a greater likelihood of achieving the balance of properties that we require.

• If you see a compound with a **blue** border, click on the **dislike** button.

• You may choose to **ignore** or **dislike** compounds with a grey border.

• For compounds with a **red** border, click on the **like** button.

• If there is a very good idea with a strong red border, you may choose to **select** the idea and add it to your data set.

You will find that the panel will quickly begin to display more ideas with a strong, red border and hence higher scores.
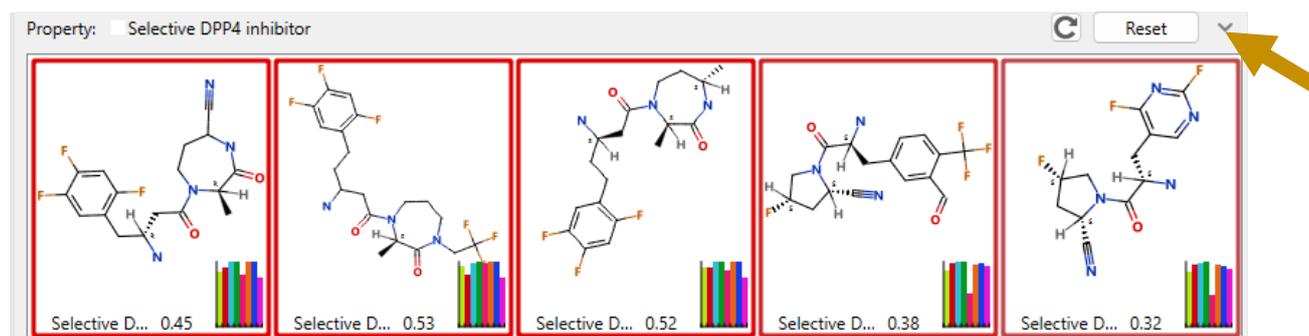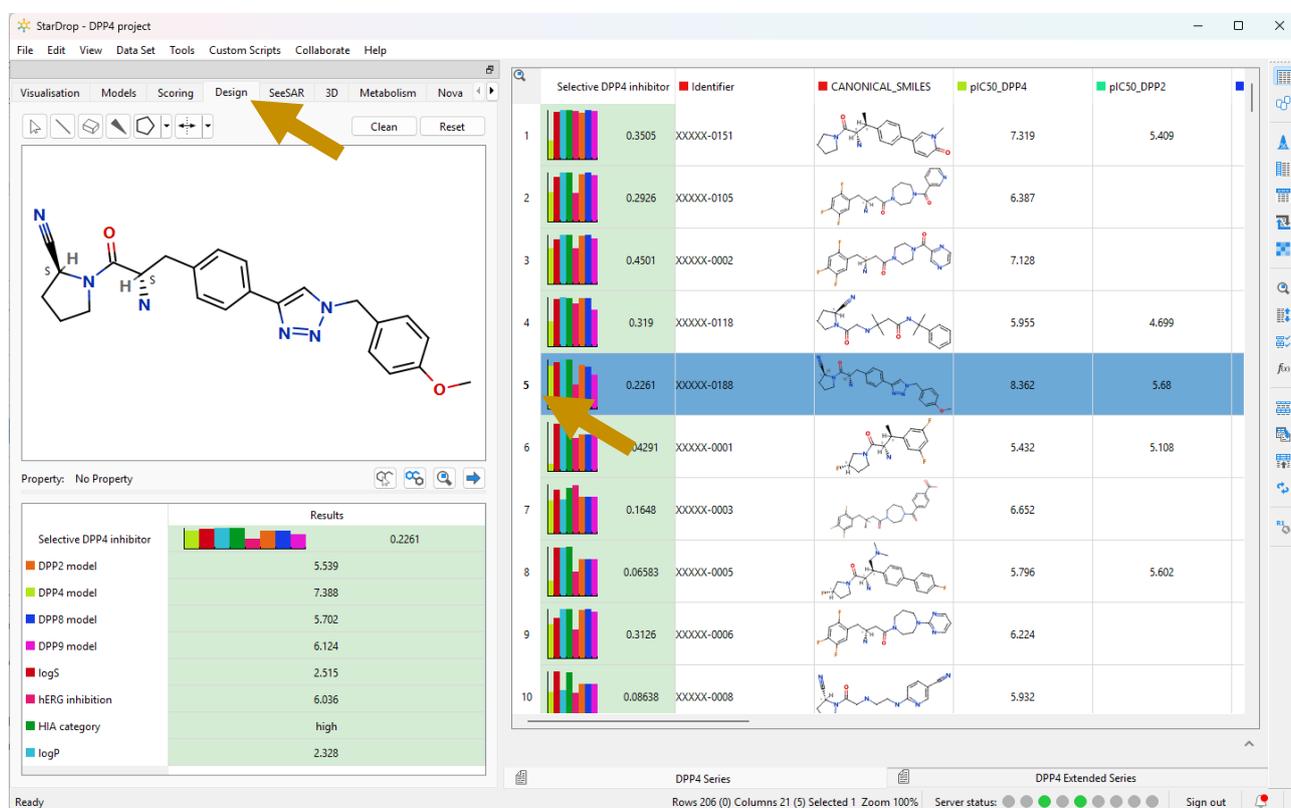
If you preferentially **select** or **like** ideas from one chemotype, you will also find that the ideas appearing in the **Inspyra Panel** are more commonly based on that chemotype.
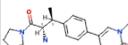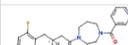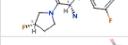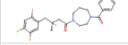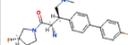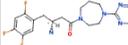
- When you are finished exploring, close the **Inspyra Panel** by clicking on the ⌄ button in the top-right, as shown above.
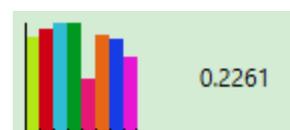
Inspyra Designer

The **Inspyra Designer** helps you explore optimisation strategies around a single compound, generating ideas for new substitutions or core and linker replacements.

- Click on the **Design** tab to change to the interactive design area.

- Select **Row 5** from the data set (compound ID XXXXX-0188).

This compound has a predicted hERG liability, as indicated by the lower pink bar in the scoring histogram.

The structure of compound ID XXXXX-0188 will be displayed in the design area. We can manually edit the structure in the designer to consider potential modifications to improve the predicted hERG inhibition. However, the **Inspyra Designer** can propose replacements for regions of the compound to achieve this goal.

- Choose **hERG inhibition** from the **Property** menu in the bottom-left of the designer.

- Click on the button below the designer to activate the **Inspyra Designer**.
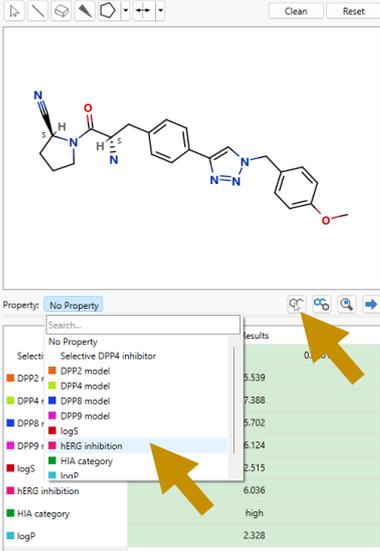
- Left click and draw a lasso around the anisole group at the right-hand side of the structure.

The message "Generating suggestions..." will appear in the bottom-left of the designer, and after a brief period, three suggested replacements for the anisole will appear in circles adjacent to the anisole group.
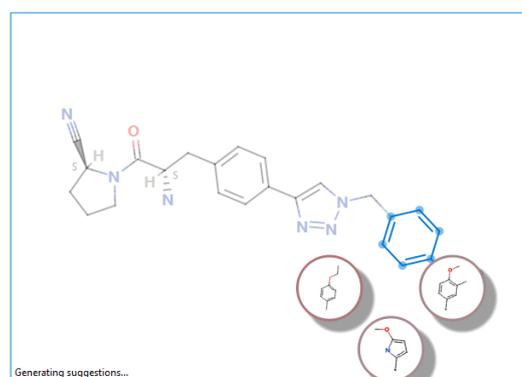

- Hover over a suggestion and scroll with the **mouse wheel** to reveal additional suggestions.

As you wait, more ideas will be generated, and you can scroll further through the list.
The replacements have borders that are coloured in the same way as the ideas in the **Inspyra Panel**; a **red** border indicates a substitution that is predicted to increase the hERG inhibition, and a **blue** border indicates a substitution that decreases the predicted hERG inhibition.

If you point at a suggested substitution, you can interact with it in a similar manner to the ideas in the **Inspyra Panel** to **dislike**, **ignore**, or **like** a suggestion. You can **select** a suggested substitution by clicking on it.
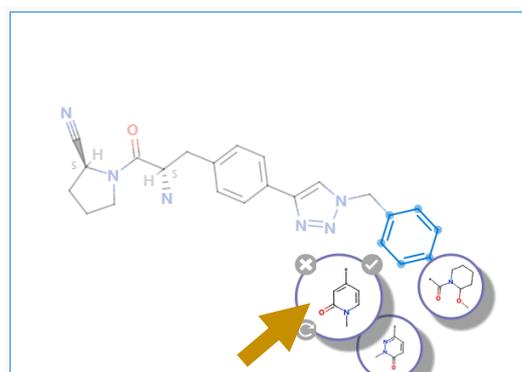
- Click on a suggestion with a blue border that is predicted to decrease the hERG inhibition.
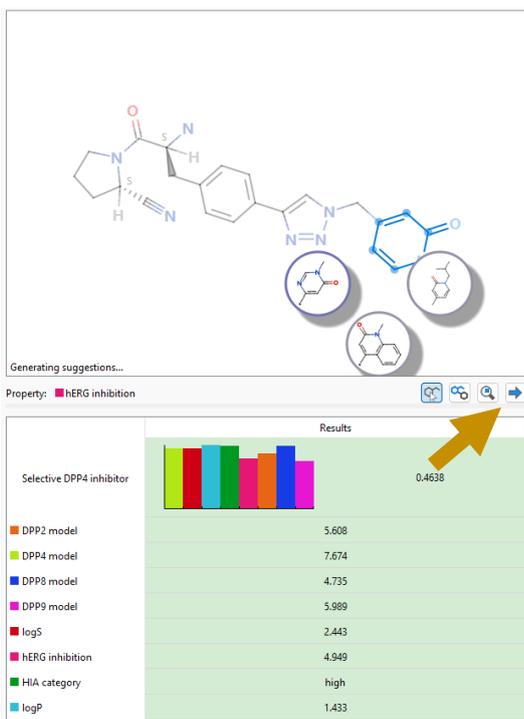
The selected substitution will be made on the compound in the designer to generate a new structure, and the properties of this new compound will be calculated and displayed below. Note that the predicted hERG inhibition of this new compound is lower than the original (4.949 versus 6.036 in the example shown to the left).

Suggestions for further substitutions based on this new compound will be generated and displayed.

• You can click the ➡ button to add the new compound idea to the data set and store it permanently.

Selecting other regions of the compound will enable you to explore other potential modifications.
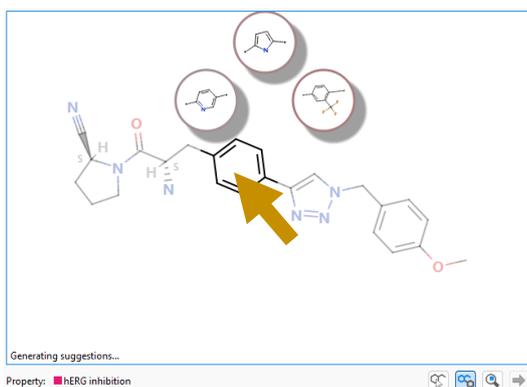We'll now explore the **Fragmentation Mode** of the **Inspyra Designer**.

• Select compound ID XXXXX-0188 again to replace the compound in the designer.

• Click on the 🔗 button below the designer to enter **Fragmentation Mode**.
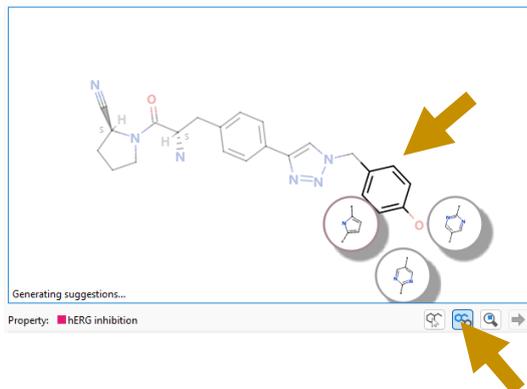
The compound will be automatically fragmented into rings, linkers, and substituents, with replacements being suggested for each fragment. Over time, more suggestions will be generated for each fragment. Initially, one region will be automatically selected; in this case, the phenyl ring of the anisole group.

• As before, you can scroll through these ideas using the mouse wheel and **ignore**, **dislike** and **like** suggestions.

If you **like** a suggestion by clicking on the ✅, this will be remembered, and you can **like** multiple replacements for each region. Choose some suggested replacements for replacement of the phenyl ring by **liking** them.
• Now, click on the phenyl ring in the middle of the molecule to view the suggestions for this region.
• Again, choose some suggested replacements by clicking on the ✅.
• Feel free to explore other regions of the compound and choose some replacements.

- Finally, click the ➡ button, and this will add compounds to the data set corresponding to every combination of the replacements you have liked for each region of the compound.

The assumption underlying **Fragmentation Mode** is that if replacements for different regions of a compound are interesting, it is likely that the combinations of these replacements will also be worthy of consideration.

The first part of this worked example has illustrated how you can interact with Inspyra in different ways to generate suggestions for new compound ideas and substitutions around individual compounds. In the next section, we will illustrate how you can use the Inspyra Panel in combination with Matched Series Analysis.

Inspyra Panel with Matched Series Analysis

In this section, we will explore ways to use the **Inspyra Panel** combined with Matched Series Analysis (MSA). MSA goes beyond Matched Pair Analysis and uses a longer series of matched compounds to propose new substitutions that are likely to improve activity (or other properties if using additional MSA knowledge bases). See Section 10.3 of the StarDrop Reference Guide for more information on MSA, available from the **Reference Guide** option in the **Help** menu.
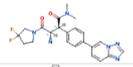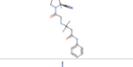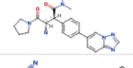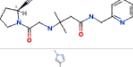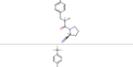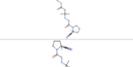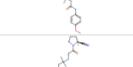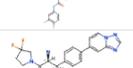
The **Inspyra Panel** provides a way to interactively explore suggestions for new compounds generated with MSA that are likely to improve activity, even when a corresponding predictive model is unavailable.

MSA requires experimental data on a sufficient number of compounds with which to analyse the structure-activity relationships within the chemical series of interest. As the DPP-4 project progressed, more compounds were synthesised and tested, resulting in a larger data set, making it suitable for MSA.

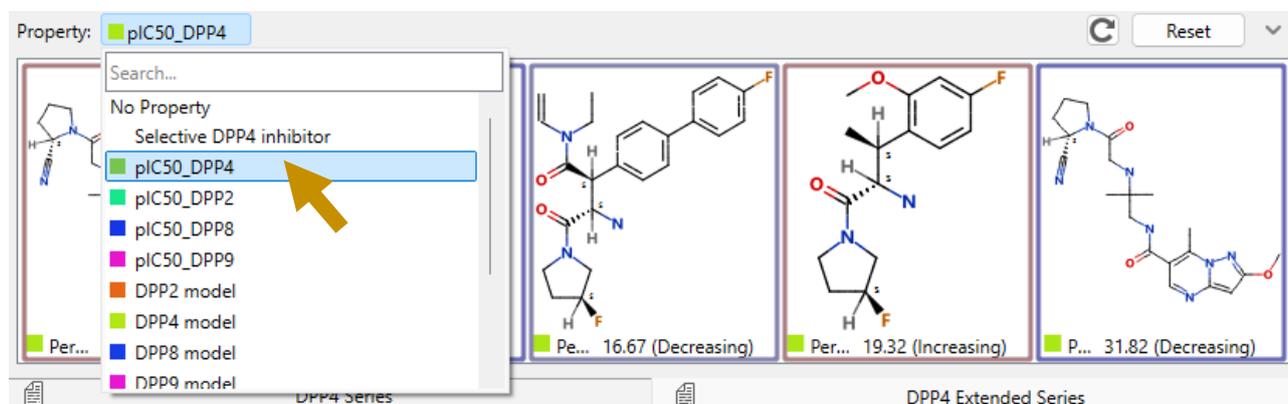- Change to the **DPP-4 Extended Series** data set by clicking on the tab at the bottom.

This data set contains 293 compounds that were generated by expanding the same series as in the previous data set.

- Open the **Inspyra Panel** again by clicking on the ⌃ button.

- Select **pIC50_DPP4** from the **Property** menu in the top left of the **Inspyra Panel**.

This will display the message "Running Matched Series Analysis…". When complete, the ideas that appear will display the percentage chance of improving the activity and whether the new idea will likely increase or decrease the pIC50 for DPP-4. A **red** border around an idea indicates an idea that is likely to increase the activity relative to other compounds in the same matched series. A **blue** border indicates an idea that is likely to decrease the activity. The intensity of the colour indicates the probability of this change.
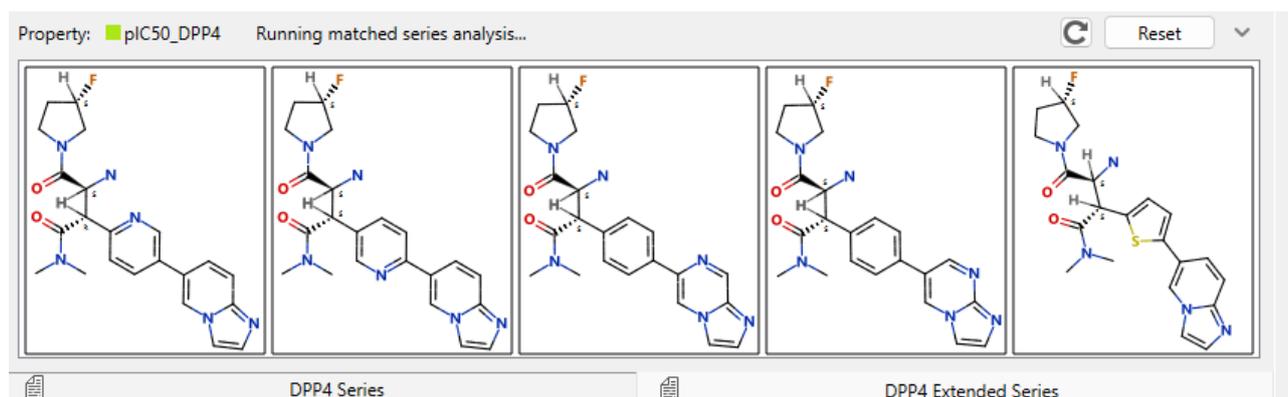
In this case, we wish to increase the pIC$_{50}$ for our primary target, so **like** ✅ or **select** ➕ ideas with red borders and **dislike** ❌ or **ignore** 🔄 ideas with blue borders. You will find that the ideas presented will quickly be dominated by ideas that are more likely to increase the pIC50.
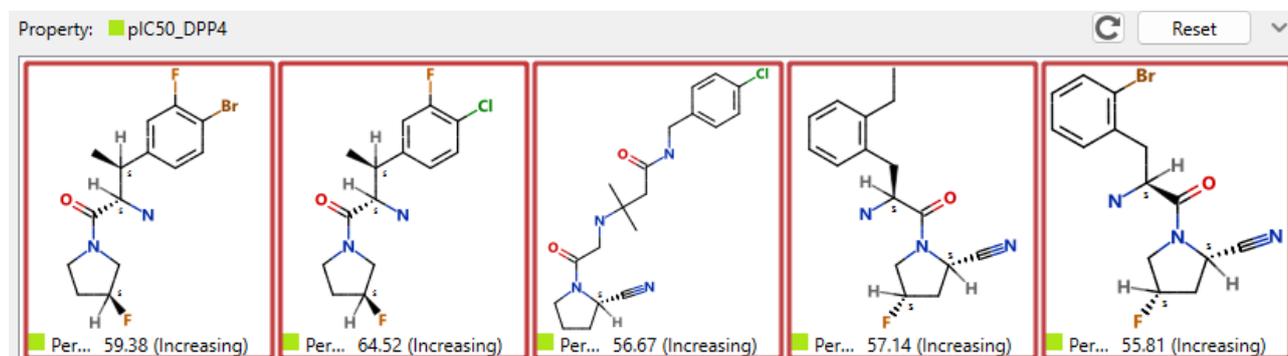
Conclusion

This worked example has explored a variety of ways to work with Inspyra in StarDrop. If you have any questions about this example or would like help applying these techniques in your work, please contact support@optibrium.com. Feel free to explore our online community at https://optibrium.com/resources/ for additional tutorials and videos.