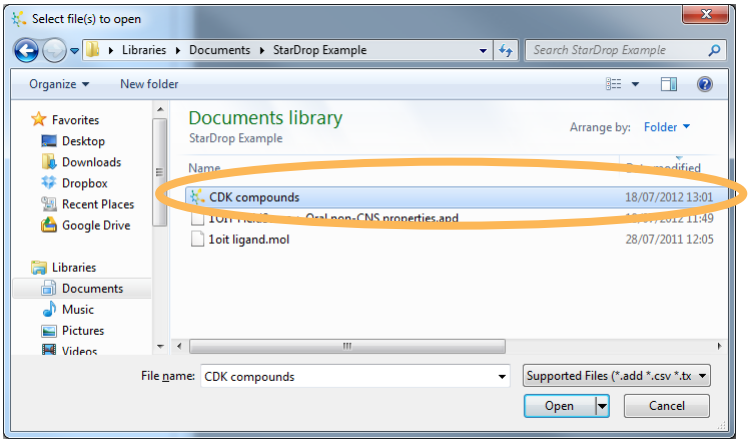
Worked Example:

Multi-Parameter Optimisation of 3D SAR

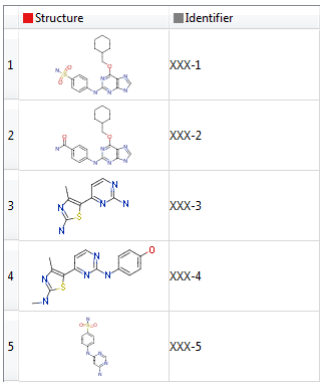
In this example we will explore the multi-parameter optimisation of a series of CDK2 inhibitors, combining a 3D insight into the structure-activity-relationship (SAR) gained from StarDrop's torch3D™ module and predictions of ADME and physicochemical properties, using StarDrop's unique Probabilistic Scoring approach.

Follow the step-by-step guidelines below to design and prioritise new compound ideas:

- Start StarDrop from the Start menu
- Select the **File->Open** menu option and open the data set **CDK compounds**.

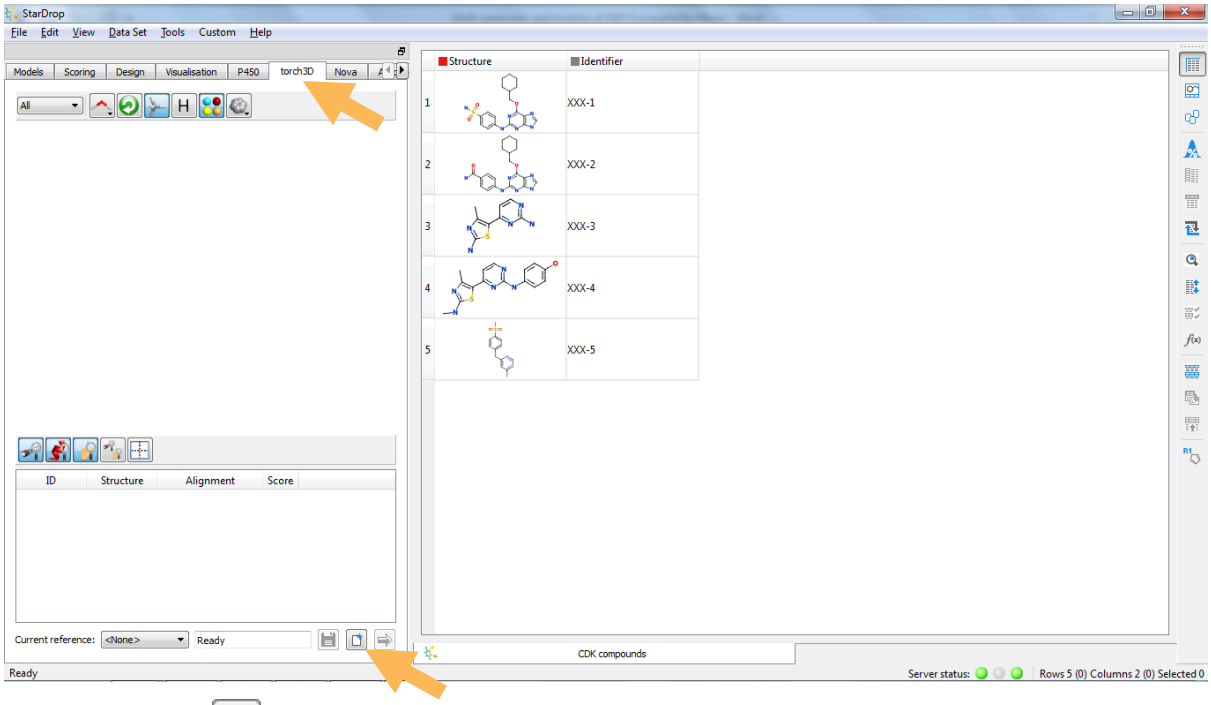


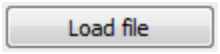
This will open a small data set containing five compounds which are active against the CDK2 target.

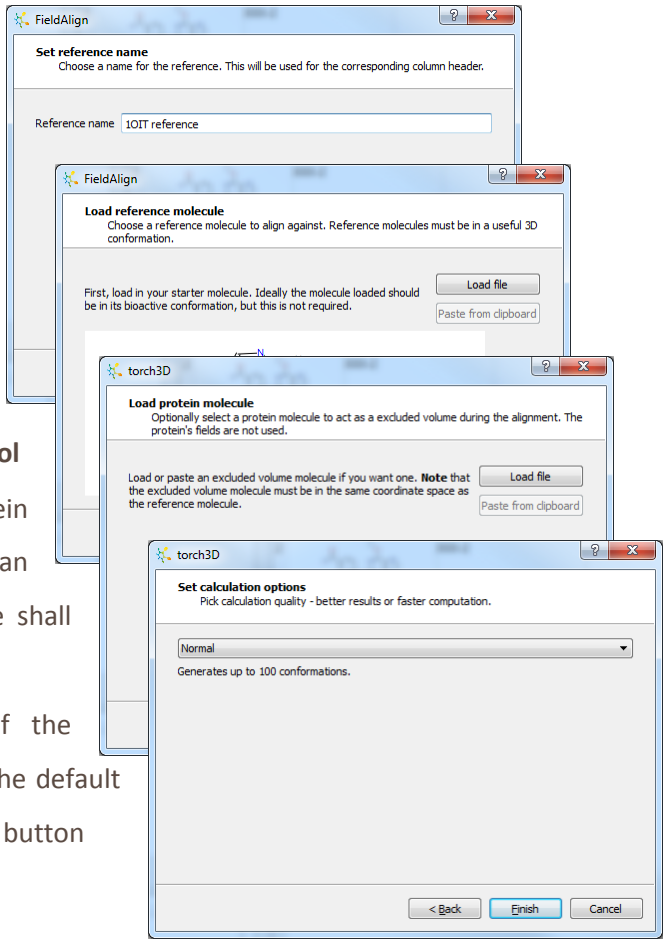


We're going to explore the 3D SAR of these compounds by comparing them to a reference compound with a known bioactive conformation in CDK2, using the torch3D™ tool developed in collaboration with Cresset.

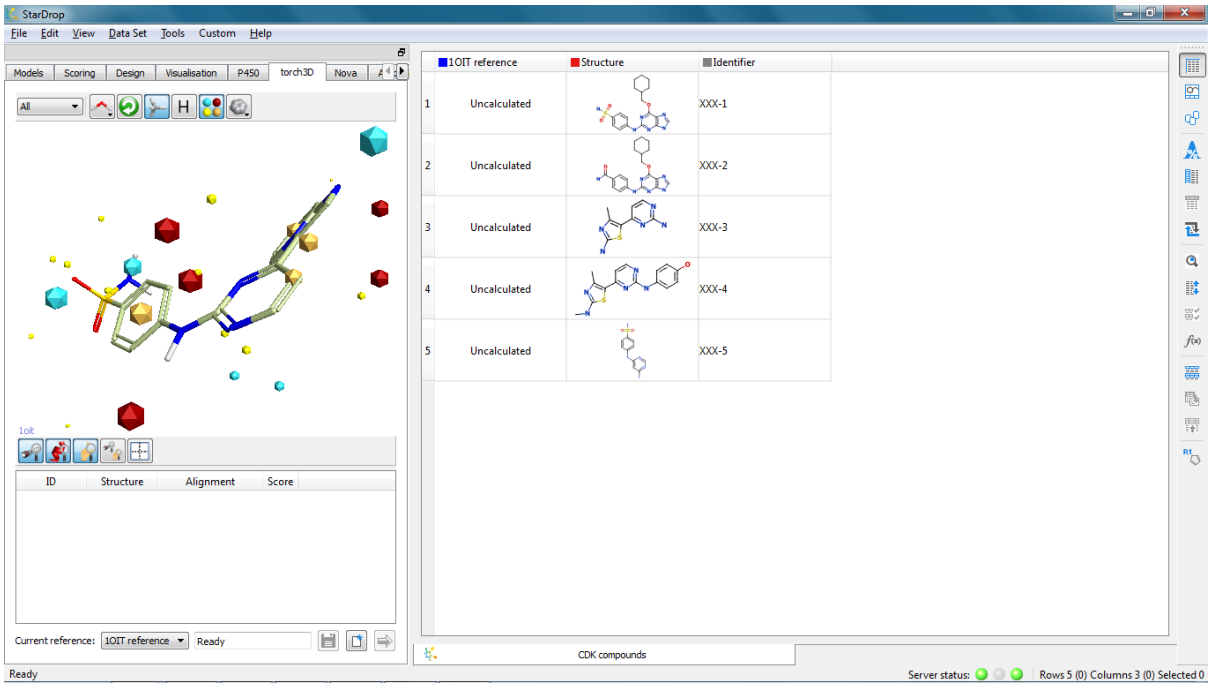
- Change to the **torch3D** tab in StarDrop



- Click the  button to start the torch3D wizard and define a reference against which new compounds will be compared
- On the first page of the Wizard, enter the name "**1OIT reference**" (Please use *exactly* this name because it will be used later)
- Click the **Next** button and then load the reference molecule by clicking the  button and selecting the attached file **1oit ligand.mol**
- Click **Next** to (optionally) load a protein structure that can be used to define an excluded volume - in this example we shall omit this step
- Click **Next** to specify the speed of the calculations – in this case we will use the default **Normal** setting and then click the **Finish** button

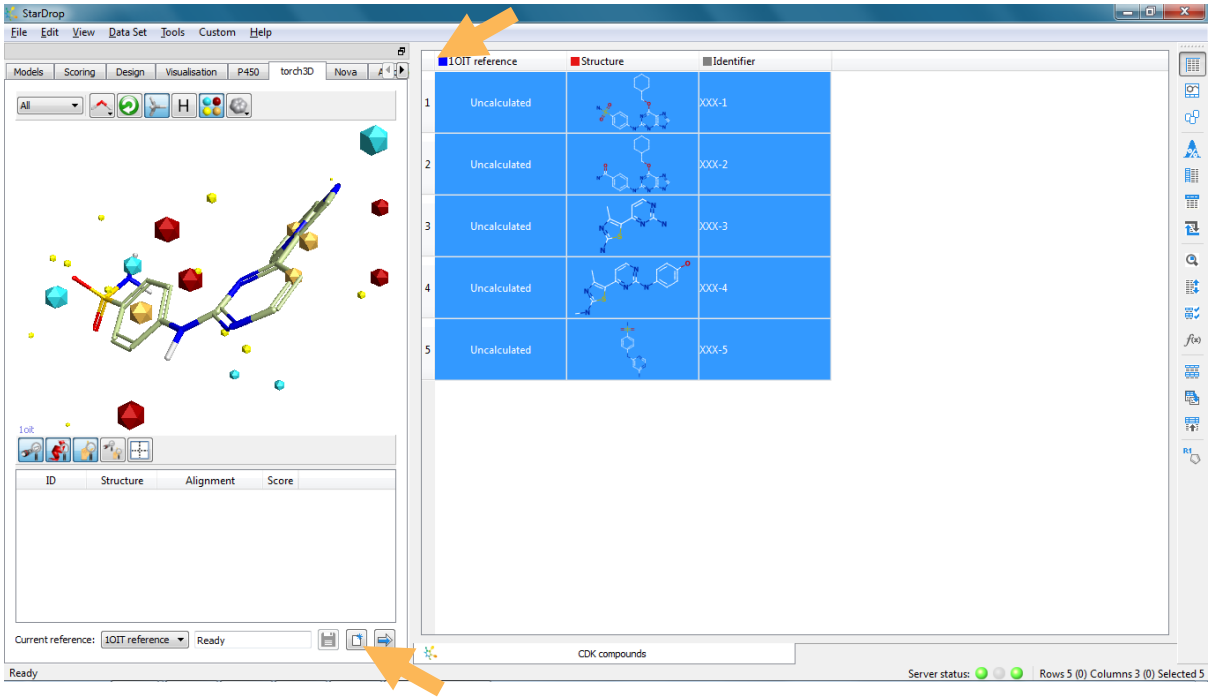


The reference molecule will be displayed and a new column will be added to the data set.

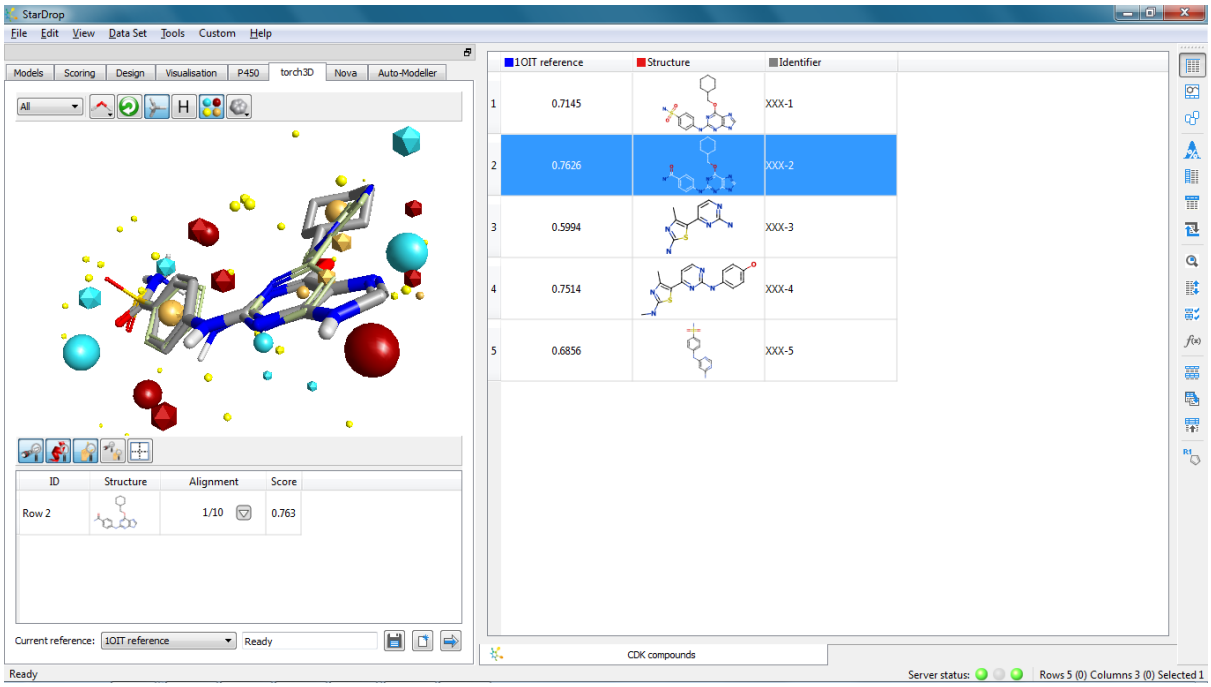


Now the reference has been set up, we can compare the compounds in the data set with the reference, based on the pattern of fields generated by the structures as well as the shape.

- Select all of the rows in the data set (hint: click the top-left corner of the data set) and then click the  button to start the calculations



Once a score has been returned for a compound, select that row to see the compound, and its field points, superimposed over the reference molecule. Using the mouse you can zoom into and rotate the 3D molecules in order to see how their fields compare.



In torch3D larger field points represent stronger points of potential interaction and the field points are coloured as follows:

- Blue: Negative field points (like to interact with positives/H-bond donors on a protein)
- Red: Positive field points (like to interact with negatives/H-bond acceptors on a protein)
- Yellow: van der Waals surface field points (describing possible surface/vdW interactions)
- Gold/Orange: Hydrophobic field points (describe regions with high polarisability/hydrophobicity)

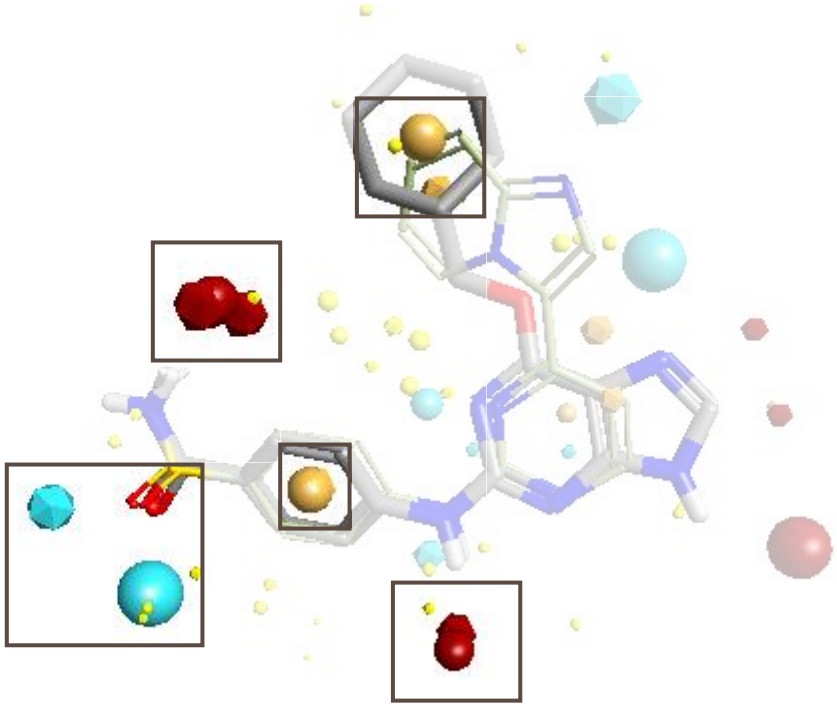
For further details about the interpretation of the field points, please see Section 9.3 of the StarDrop Reference Guide, which can be opened from the **Help->Reference Guide** menu option in StarDrop.

The FieldScore (a value between 0 and 1, where scores closest to 1 are the best match) is calculated by considering how similar the fields around the aligned conformation are to those of the reference compound. The best score is shown, but the scores for the next best conformations (10 in total) are available, enabling you to view other possibilities. Using the up and down arrows in the table below the 3D window you can choose which conformation is displayed.

N.B. The score is an important factor in deciding the validity and potential activity of a particular alignments and molecules. However, it is not the only factor to be considered before embarking on the synthesis of a compound designed in torch3D. The top-scoring result is the one that is the most

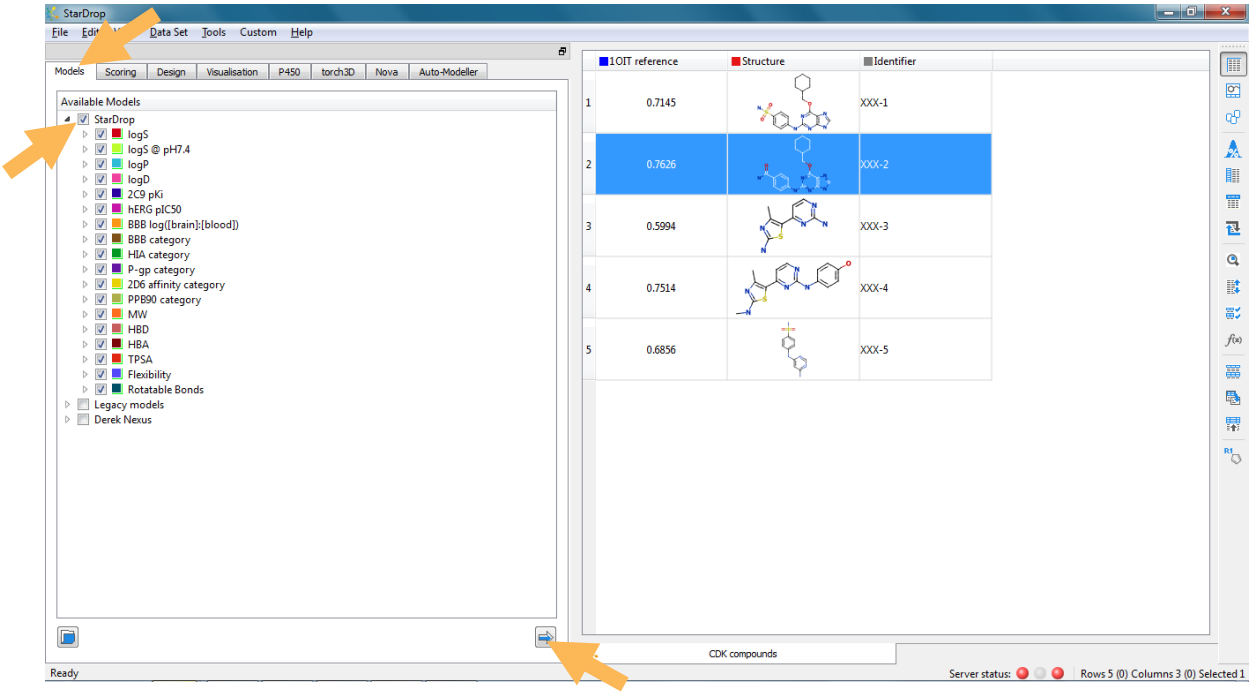
similar to the target molecule in terms of fields and shape. That doesn't necessarily mean that it is the most likely to be active or that it's the one you should make first.

- Explore the different compounds in the data set and note that, despite the difference in the chemical series, there is a strong similarity in the alignment of their fields and their shape, explaining their similar bioactivity. See, for example, compound **XXX-2** below, overlaid on the reference:



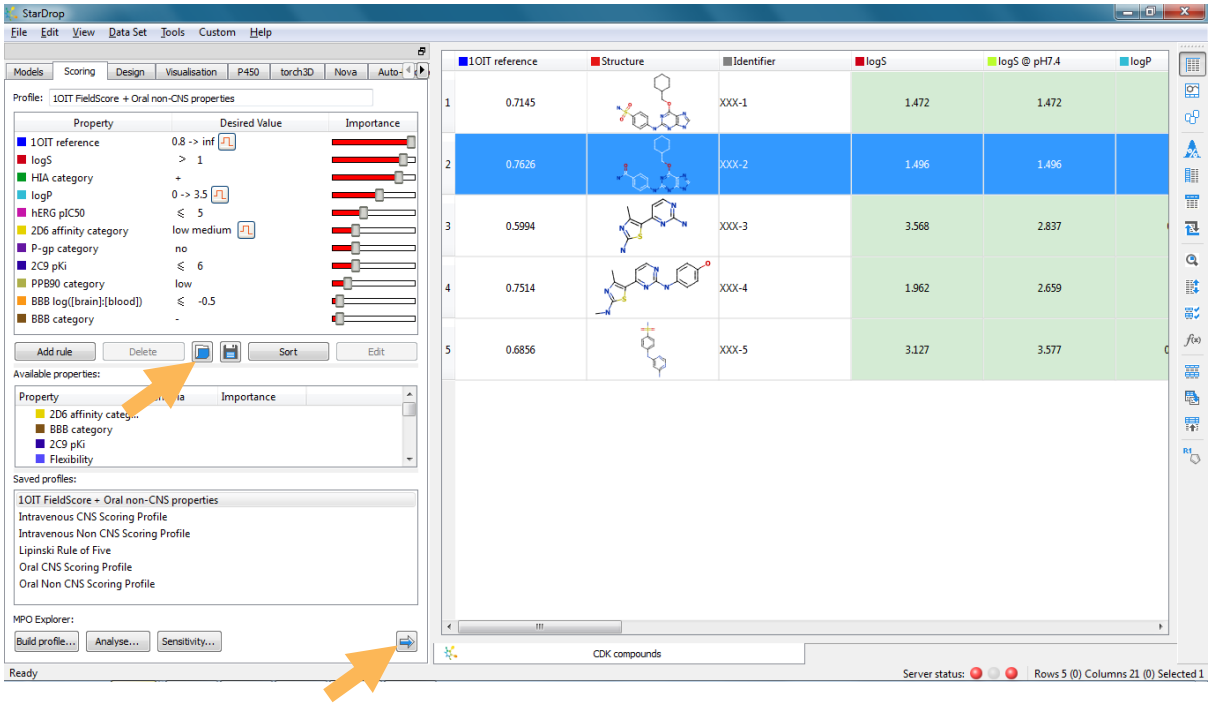
This can provide us with valuable information to guide the design of novel compounds with improved potency. However, potency is not the only factor to consider when optimising compounds, so to help us to design compounds with an optimal *balance* of properties, we will also make predictions of relevant ADME and physicochemical properties using StarDrop's ADME QSAR module.

- Change to the **Models** tab in StarDrop
- In the list of **Available Models**, tick the box next to **StarDrop** to select all of StarDrop's ADME QSAR models and click the  button to calculate the selected properties for all of the compounds in the data set



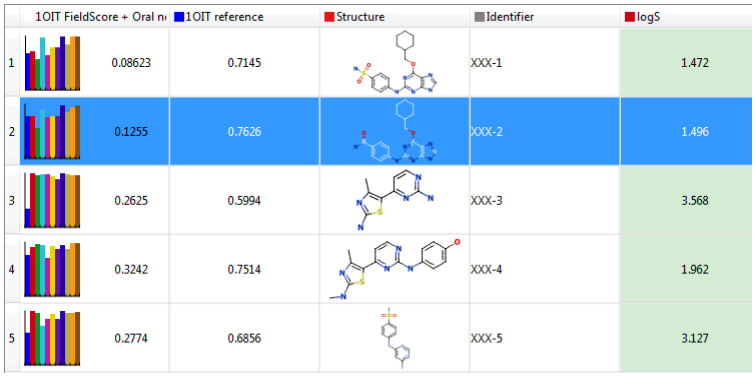
Now we have a lot of data for each of the compounds; too much to easily keep track of all of the properties while designing new compounds. Therefore, we'll use StarDrop's Probabilistic Scoring approach to multi-parameter optimisation to assess each compound's properties against the overall profile required 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.

- Change to the **Scoring** tab in StarDrop
- Click the  button to load a new scoring profile. Load the file **1OIT FieldScore + Oral non-CNS properties**



This shows the profile of property criteria that have been defined by the project team; in this case the team would like a compound with a high FieldScore and appropriate properties for an orally dosed compound against a CNS target. The importance of each criterion has also been specified, allowing acceptable compromises to be defined.

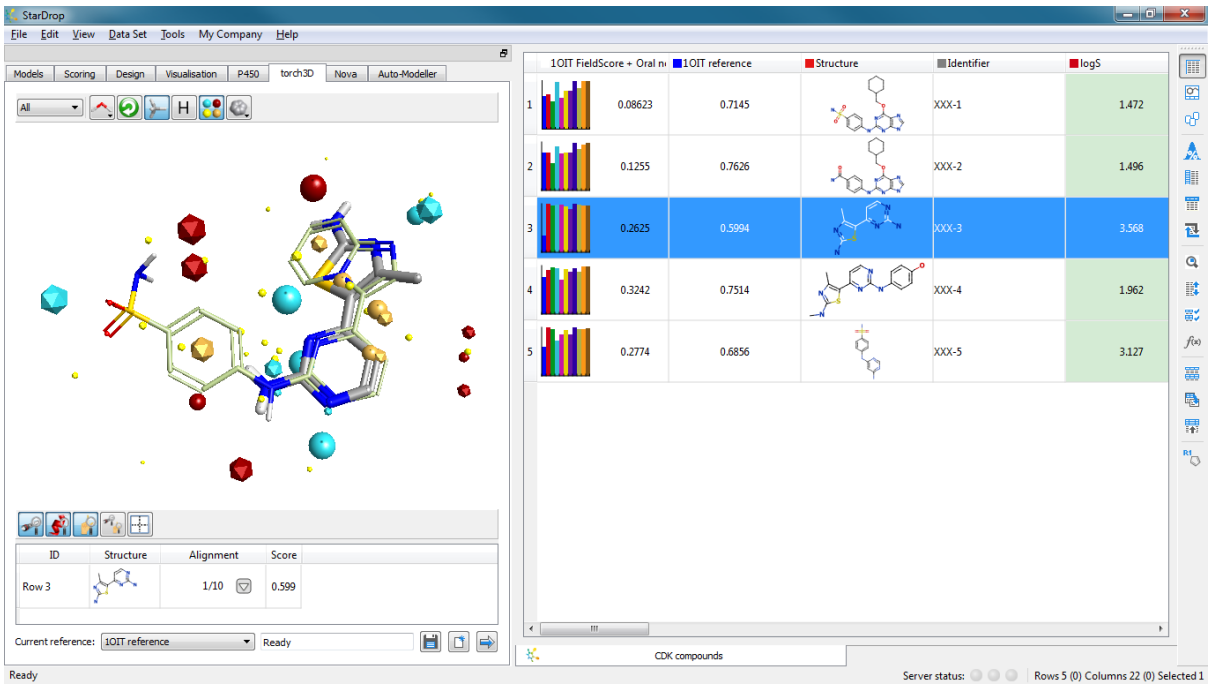
- Click the  button to generate a score for each compound against this profile.



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. The histogram shows the impact of each individual property on the overall score; the colours correspond to the key in the scoring profile.

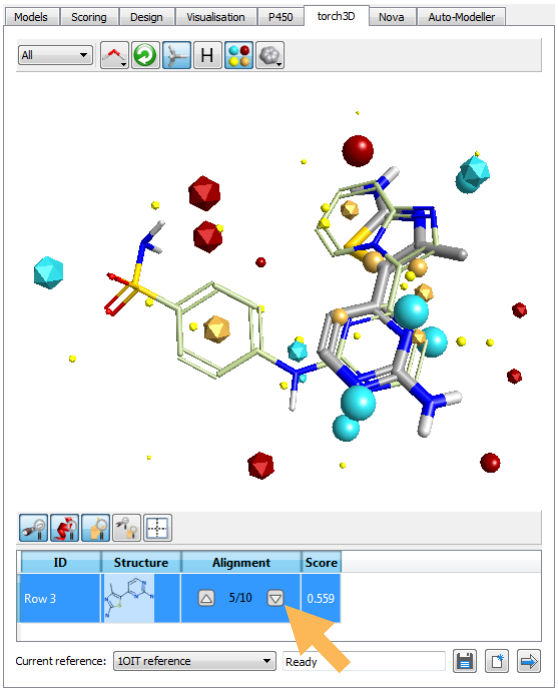
From the scores, we can see that the highest-scoring compounds achieve a score of approximately 0.3. Given the uncertainty in the predicted data, this is quite good, but ideally we would like to try to find even better compounds to synthesize and test, so let's explore the multi-dimensional optimisation of these compounds.

- Return to the **torch3D** tab in StarDrop and select row three in the data set which contains the compound **XXX-3**



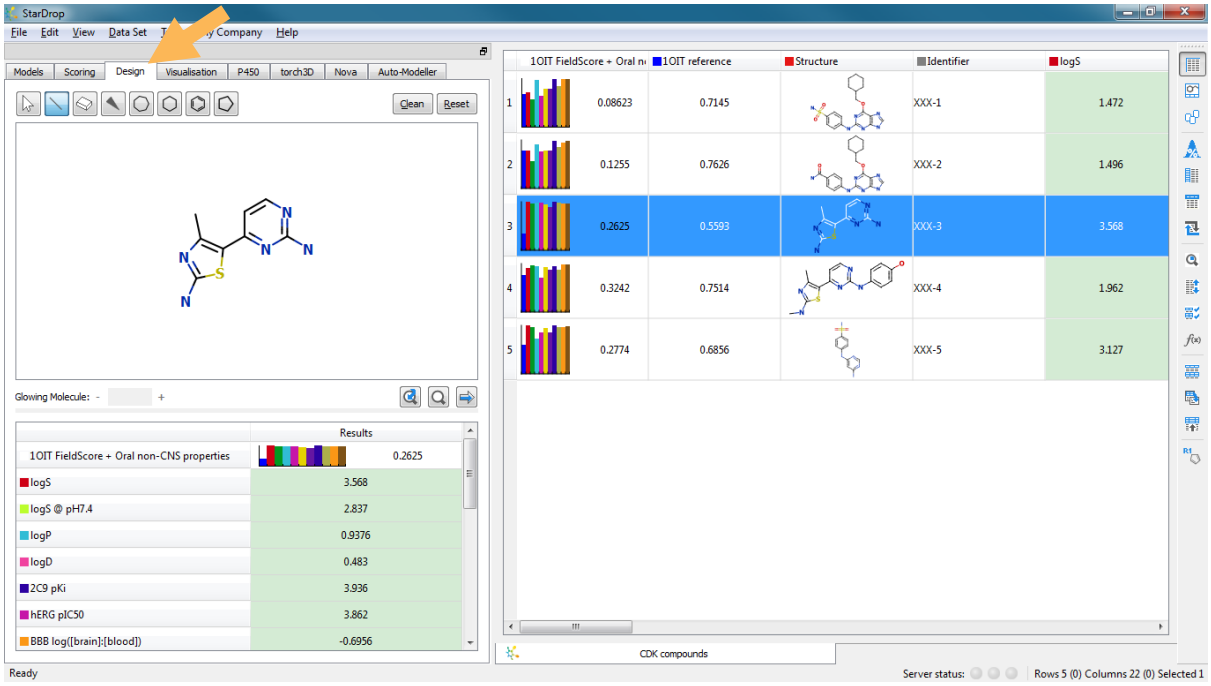
This small fragment aligns well with the reference compound and, despite its small size, achieves a reasonably good FieldScore. The best alignment identified by torch3D is shown by default, but torch3D compares many, energy-minimised conformations with the reference and different alignments with good FieldScores may suggest possible alternative binding modes and indicate opportunities for further optimisation. As an example, we will explore alternative alignments of compound **XXX-3**.

- Explore alternative alignments of compound XXX-3 by clicking the  or  buttons in the **Alignment** column in the table under the 3D visualisation. It is notable that one alternative conformation generated by torch3D is very similar to the highest scoring, except that the piperidine ring is rotated by 180°, giving rise to a different orientation of the amine. This has a similar score to the best alignment with the potential to extend this fragment and form additional interactions.

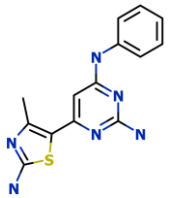


To explore the effect of different substitutions on the piperidine, we're going to use StarDrop's interactive designer.

- Change to the **Design** tab in StarDrop. The selected compound **XXX-3** will appear in the designer, ready to modify



- Modify the structure to be the one shown below using the chemical structure editor

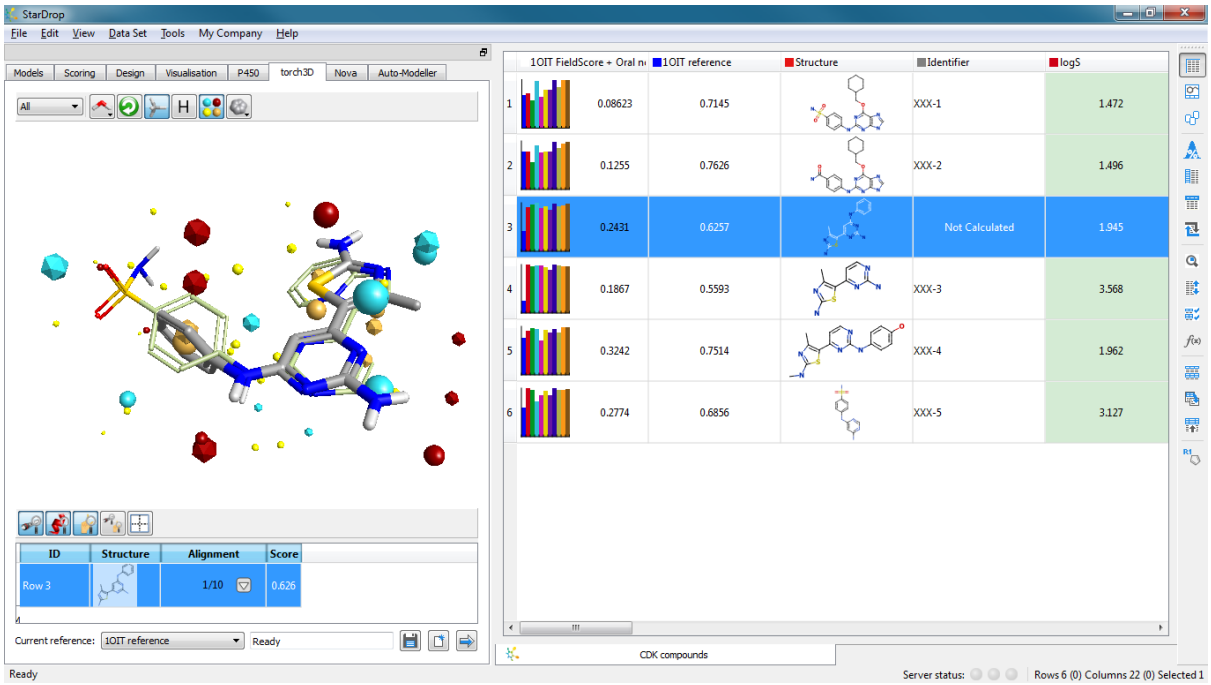


Hints:

- The  tool allows you to add atoms and bonds.
- To change the element of an atom, simply point at the atom and type the symbol. To change from one heteroatom to another, point at the atom, delete the symbol and type the new symbol.
- The  tool allows you to add a benzene ring. To add a ring, position the ring over an atom and click to fuse it onto the molecule (a blue circle will appear).

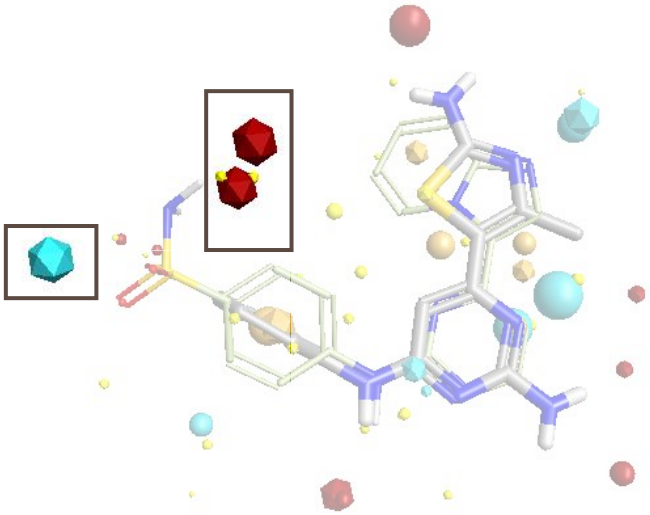
You will notice that the results for the QSAR models below the editor update instantly, as you edit the structure. However, the FieldScore calculation takes too long to provide an instant update and so isn't displayed in the list. To calculate the FieldScore, we'll add the new compound to the data set.

- Click the  button on the **Design** tab to add the structure to the data set
- Return to the **torch3D** tab and click the  button to calculate the FieldScore for this new compound



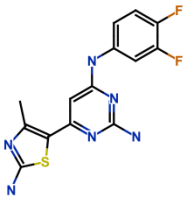
This has resulted in an increase in the FieldScore and this compound has the highest probabilistic score so-far, because the ADME and physicochemical properties for this compound are also predicted to be good (you may need to return to the **Scoring** tab and re-run the scoring calculation by clicking  to update the scores).

However, there may be an opportunity for further optimisation by substitution on the benzene; note the strong positive and negative field points in the equivalent region of the reference, highlighted below:

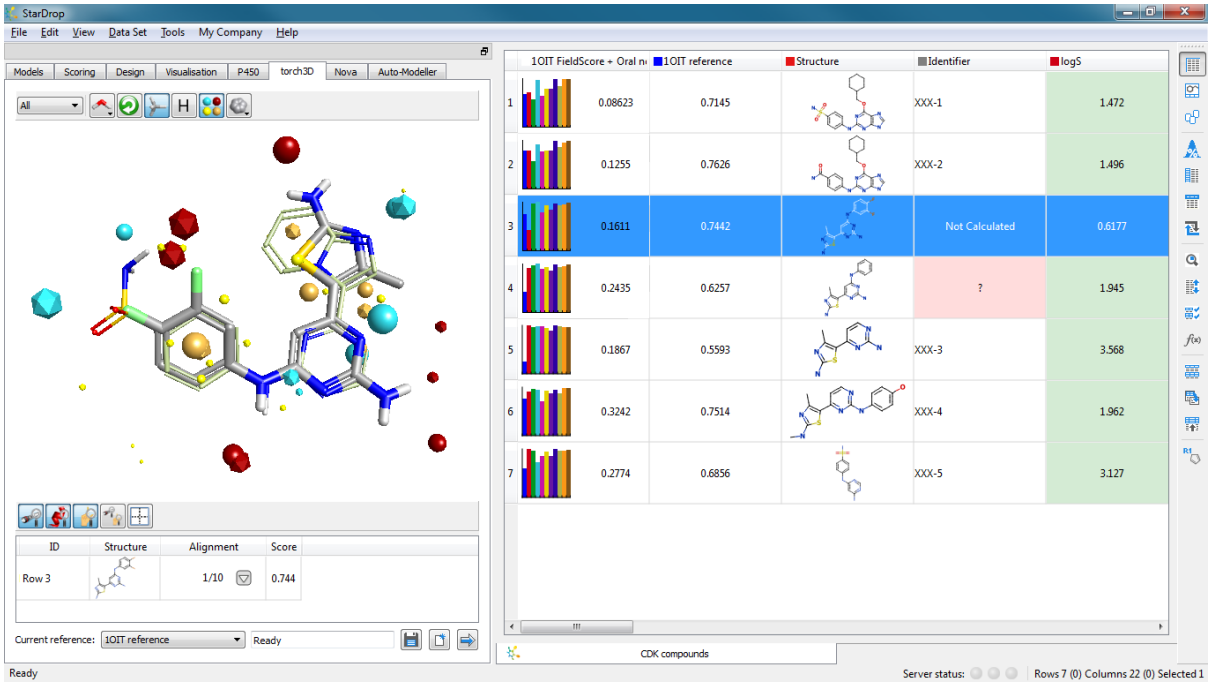


One approach to reproducing this field pattern would be by substituting a difluoro benzene. Let's try that:

- Return to the **Design** tab and add the structure below to the data set



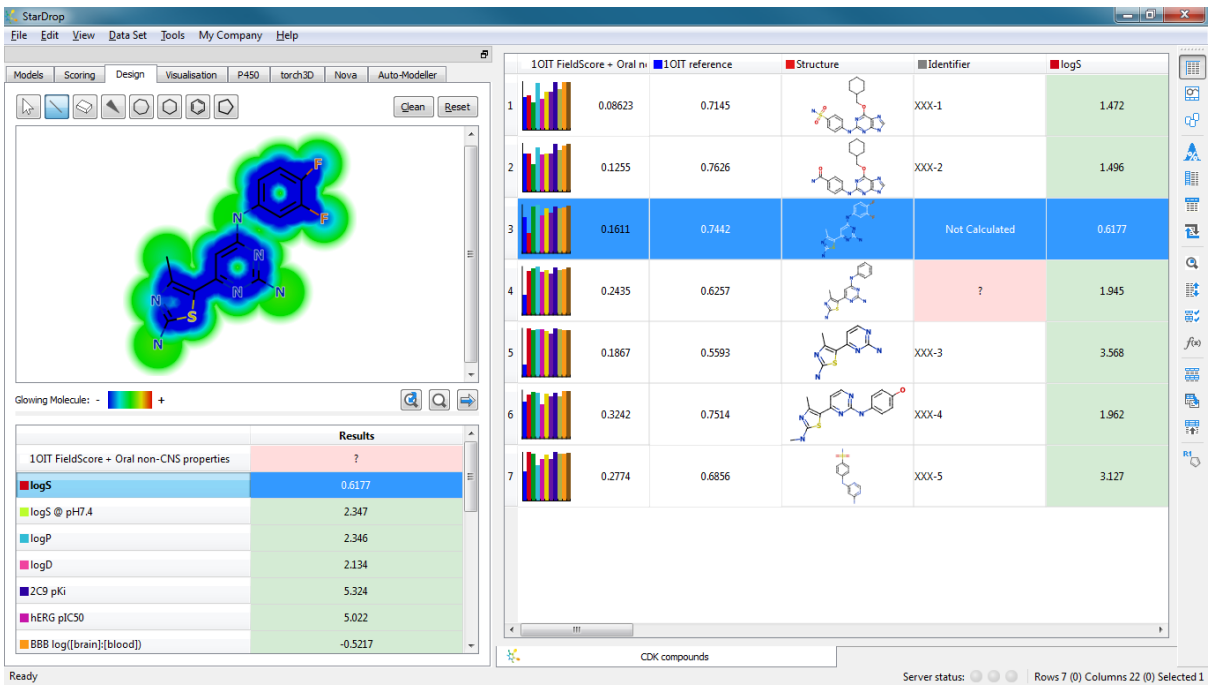
- In the **torch3D** tab, calculate the FieldScore for this compound and the overall probabilistic score will be updated (again, you may need to return to the **Scoring** tab and re-run the scoring calculation by clicking  to update the scores)



Here you can see that the substitution has had the desired effect on the FieldScore. However, this has been at the cost of significantly reducing the solubility resulting in a decrease in the overall score, suggesting that this is unlikely to yield a high quality compound.

You can see the impact of the Fluorine substitutions on the predicted solubility using StarDrop's Glowing Molecule.

- Change to the Design tab and click on the logS value below the editor to see the Glowing Molecule for this compound



The blue colour around the Fluorines indicates their significant negative effect on the solubility.

- Using torch3D, the ADME QSAR modules and probabilistic scoring, explore different substitutions to see how high a score you can achieve by simultaneously improving the FieldScore, ADME and physicochemical properties. Looking at the other compounds in the data set may provide some useful suggestions to start with...

This example has used some of StarDrop's capabilities to explore the multi-parameter optimisation of a series of compounds; in particular probabilistic scoring, interactive design and Glowing Molecule along with the optional torch3D and ADME QSAR modules.