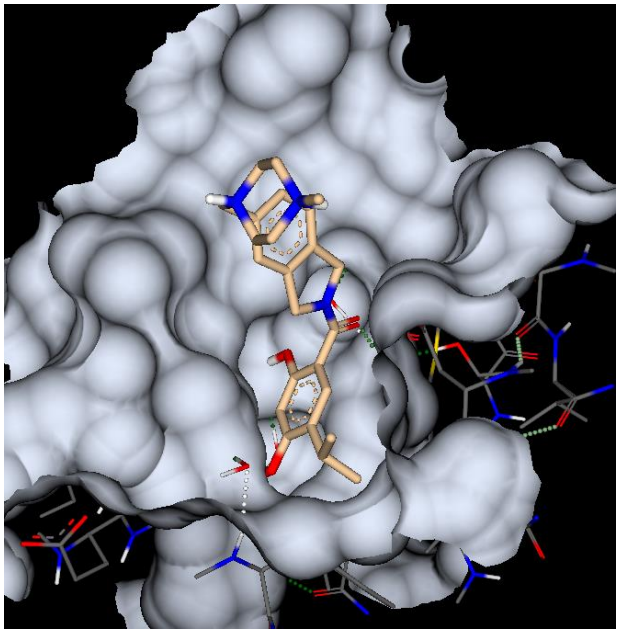
Worked Example:

R-group Clipping of Reagents for Library Enumeration

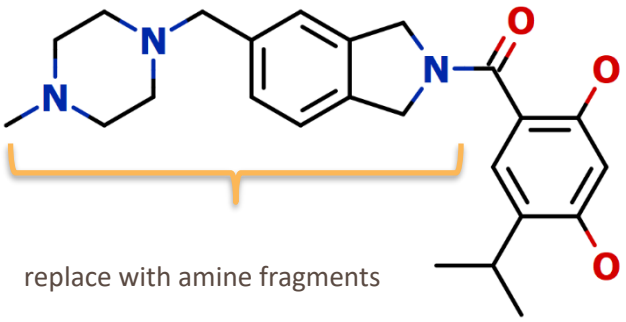
This example explores some of the challenges typically encountered in scaffold-based library design, in particular, the task of creating reagent fragments (clipping) for use in scaffold-based library enumeration. Using StarDrop's R-group clipping tool, we will quickly transform chemical building blocks into their corresponding substituents, ready to enumerate a virtual library in StarDrop's Nova module.



The crystal structure on the right (PDB 2XJX) shows the binding site of Heat Shock Protein 90 (HSP90) with Onalespib as the co-crystallised ligand. Onalespib is a selective, potent HSP90 inhibitor that displays a long duration of anti-tumor activity. The beta resorcinol group forms a tight hydrogen bond network in the binding site, but the 5-(piperazin-1-ylmethyl)-isoindoline does not form any strong interactions with the protein.

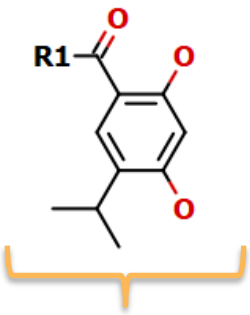
We will prepare a data set of virtual compounds based on an amide coupling reaction with a beta resorcylic acid core and commercially available secondary amines. The resulting amide library will be suitable for further analysis and prioritisation in StarDrop, including multi-parameter optimisation using ADME property calculations and Probabilistic Scoring. The virtual compounds can also be submitted to docking software via StarDrop's Pose Generation Interface to explore binding site interations.

optibrium



replace with amine fragments

Onalespib

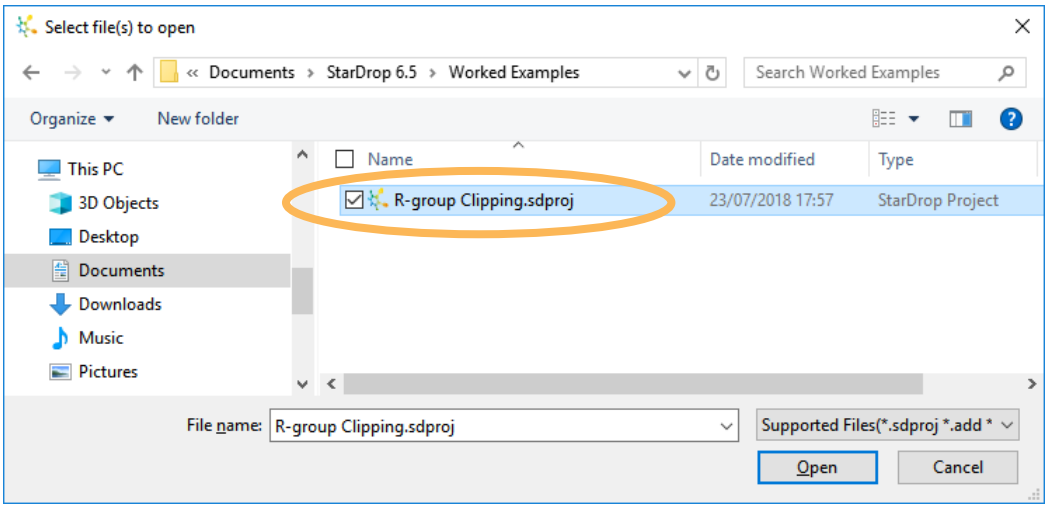


beta resorcylic acid core

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.

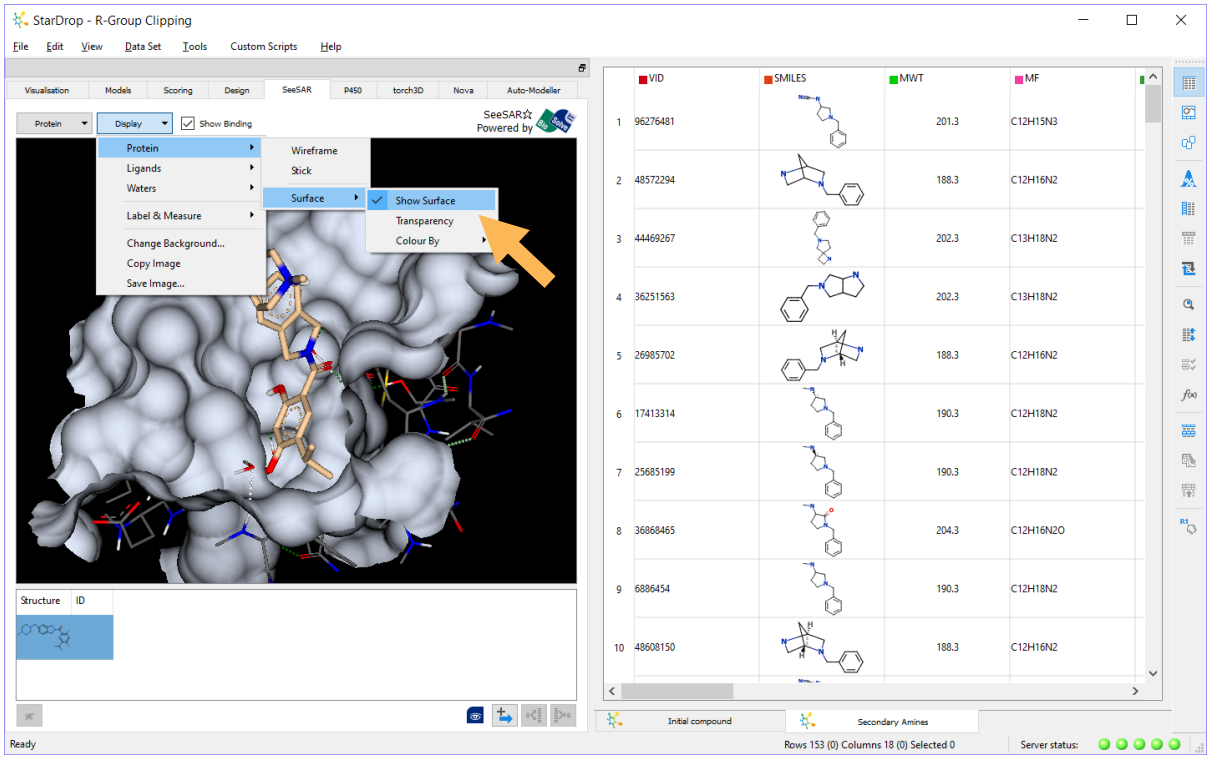
Exercise

- In StarDrop, open the file **R-group Clipping.sdproj** by selecting **Open** from the **File menu**.



On the left, in the SeeSAR area, the protein HSP90 is displayed with its secondary structure and the co-crystallised ligand, Onalespib. To better understand the binding mode of Onalespib we can explore the binding site.

- In the SeeSAR area, click the **Show Binding check box** above the protein to change the view of the protein to focus on the binding pocket.
- To show the binding site surface, select **Protein** from the **Display menu** at the top of the SeeSAR area and then choose **Show Surface** from the **Surface** options**.**



Hint: Using the mouse, you can interact with the view of the protein:

- Use the mouse-wheel to zoom in and out
- Use the left-mouse button and drag to rotate the view
- Use the right-mouse button and drag to pan the view

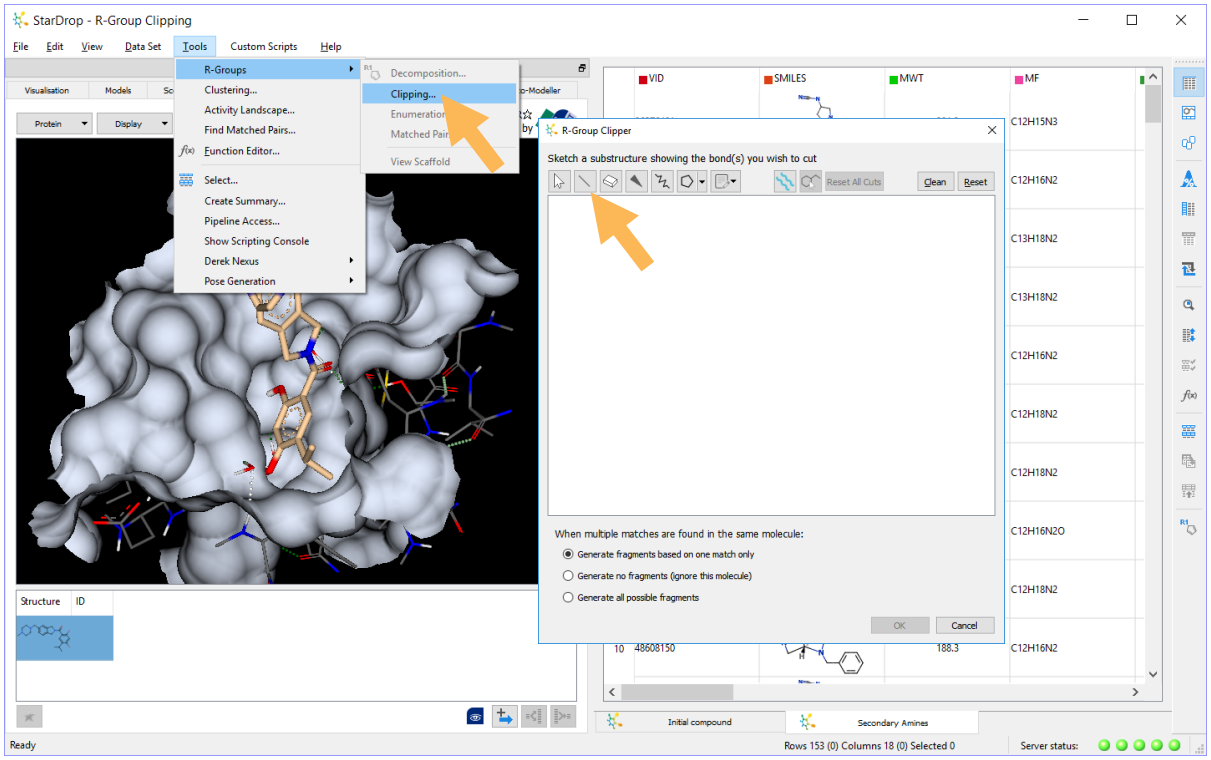
Note: Hydrogen bond interactions are indicated by either green or white dashed lines depending on the strength of the interaction, with green being stronger.

On the right, we have included in this project a data set containing 153 secondary amine structures and their associated meta-data which were retrieved directly from eMolecules. To learn more about querying and retrieving information on eMolecules compounds directly from StarDrop, please visit:

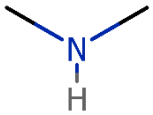
https://www.optibrium.com/community/videos/introduction-to-stardrop-modules-and-features/357-stardropemolecules

The first step in creating the virtual library is to clip the secondary amine reagents into R-group fragments that we can use to enumerate the new library.

- Open the **R-Group Clipper dialogue** by selecting **R-Groups** from the **Tools menu** and choose **Clipping**.

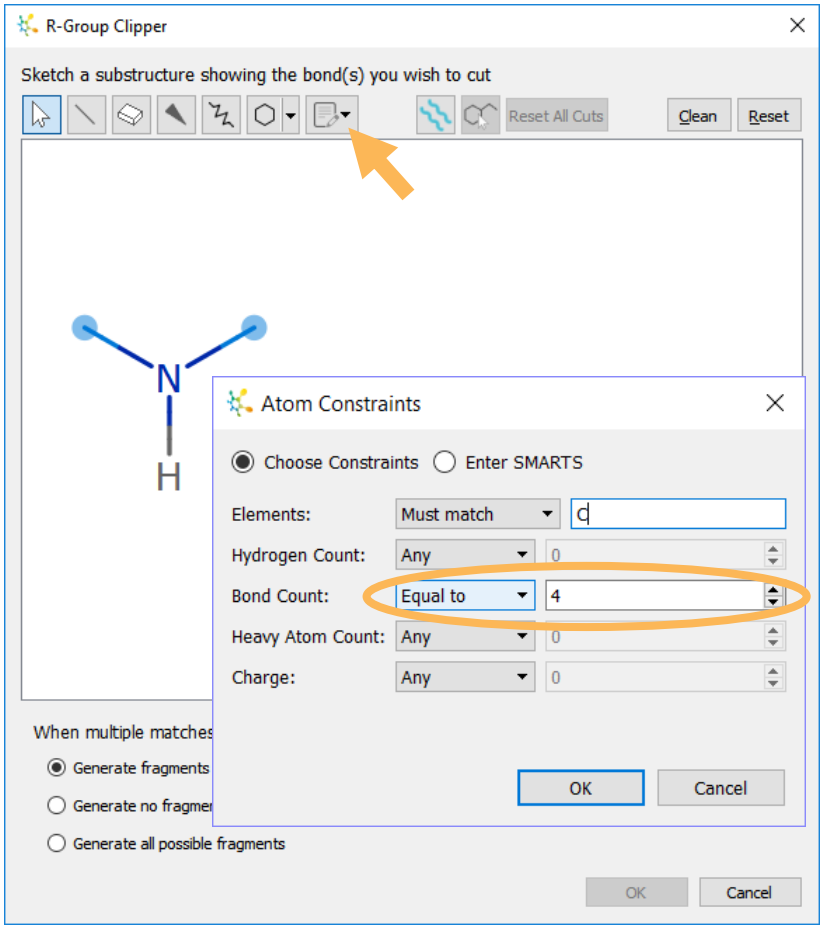


In the R-Group Clipper we can sketch a substructure which defines how compounds in the data set should be clipped. In this case, we will sketch the secondary amine and impose some bond and atom constraints to limit the fragment set to only cyclic, aliphatic, secondary amines.

- In the sketch area use the **Bond tool**  to sketch a simple dimethyl amine molecule. 

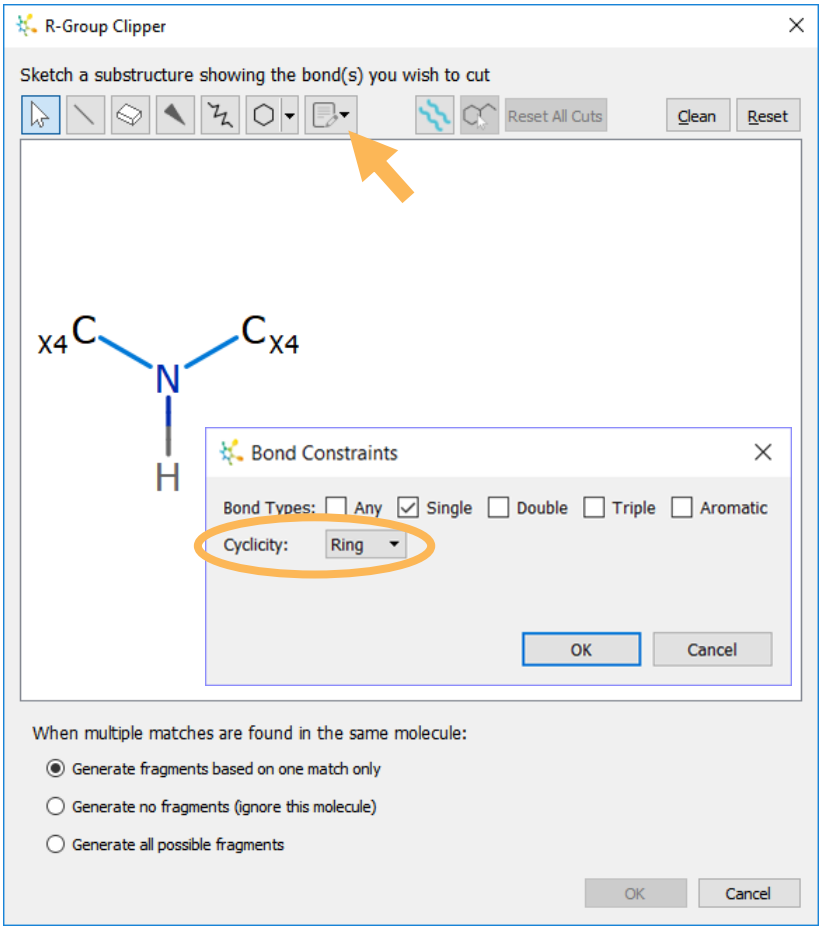
Hint: To specify an element, hover over an atom and type the element symbol, in this case “N” and “H”.

- To add some atom constraints to the two carbon atoms, first select them both by pressing the **CTRL key** while using the **Selection tool**  .

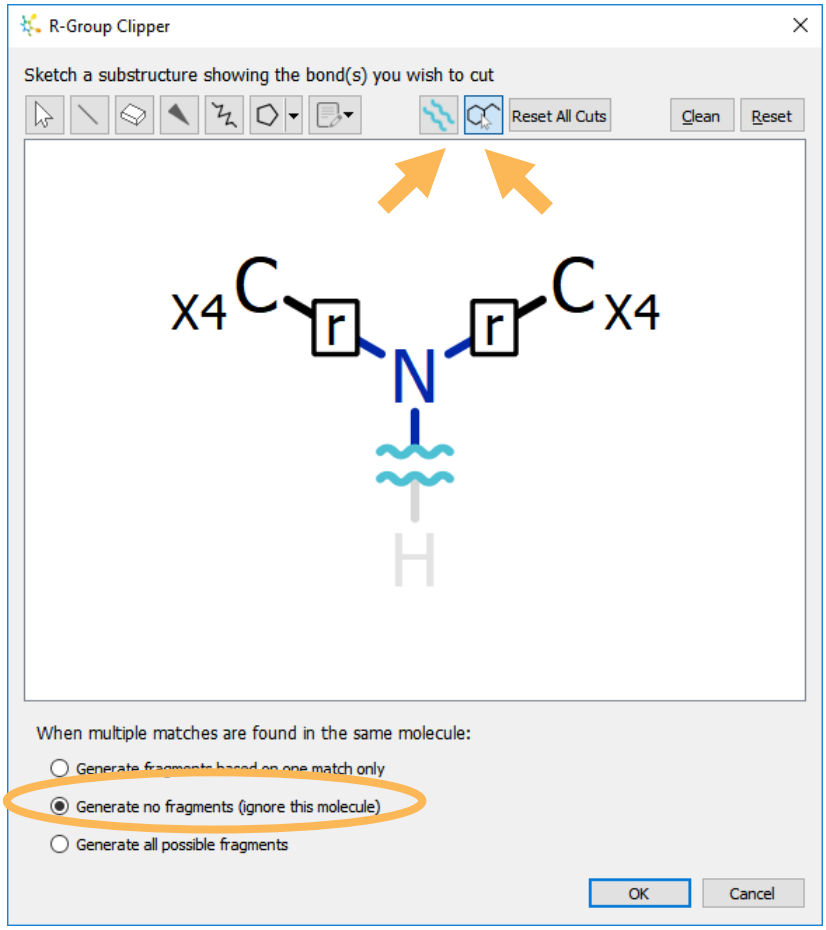


- Click on the **Constraints menu**  and choose **Edit Atom Constraints** to display the **Atom Constraints dialogue**.
- Specify that the bond count for each carbon atom should be **Equal to 4** and click **OK.**
- To add some bond constraints, first select the two N-C bonds by pressing the **CTRL key** while using the **Selection tool**  .

- Choose **Edit Bond Constraints** from the **Constraints menu**  to display the **Bond Constraints dialogue**.
- Select **Ring** from the **Cyclicity options** to specify that these bonds must be single bonds that are part of a ring.
- Click **OK.**



Now we need to specify where we would like the amines to be clipped and define the excluded fragment.



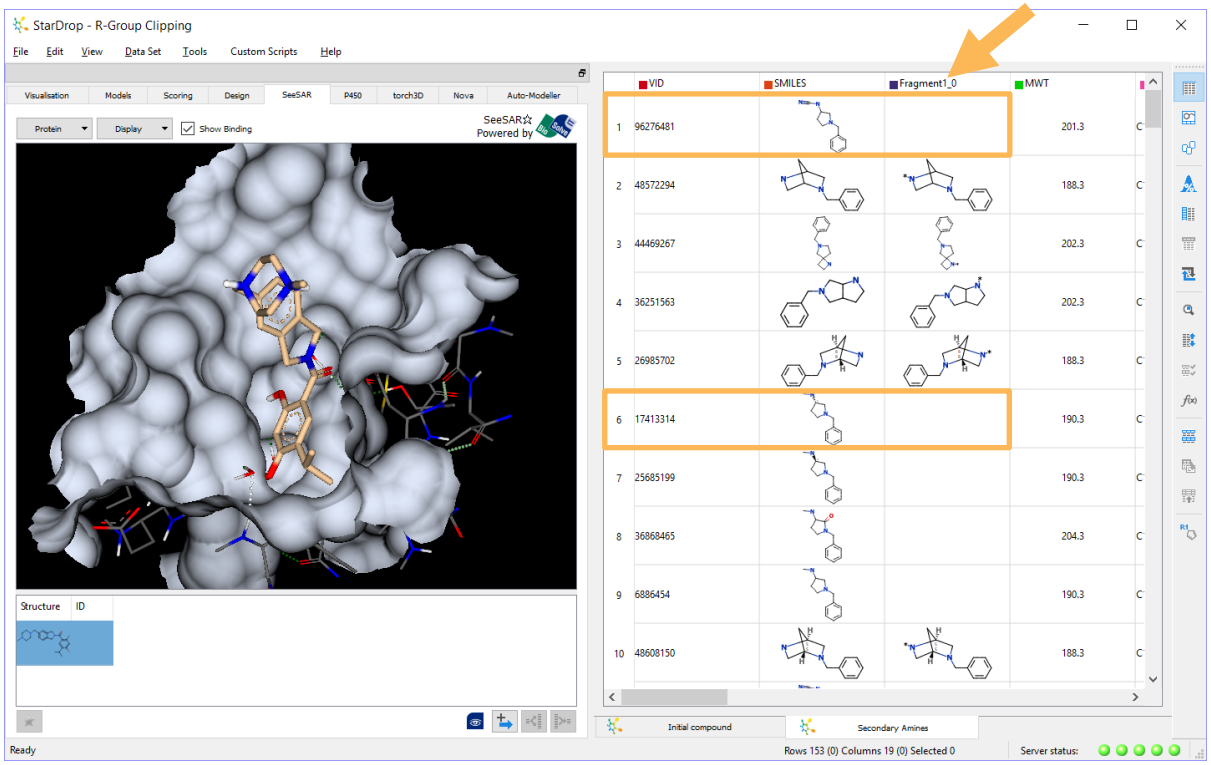
- Select the **Cut button**  and click on the N-H bond to clip this bond.
- Select the **Choose button**  and then click on the Hydrogen to exclude it from the generated fragments.

When comparing molecules in the data set with the specified substructure to determine where to clip them, it is possible that multiple matches might be found within the same molecule. At the bottom of the **R-Group Clipper dialogue** you can specify what should happen when this occurs and, in this example, we will with ignore these secondary amines and generate no fragments when multiple fragments are possible.

- Select the second option to **Generate no fragments (ignore this molecule)** and click **OK.**

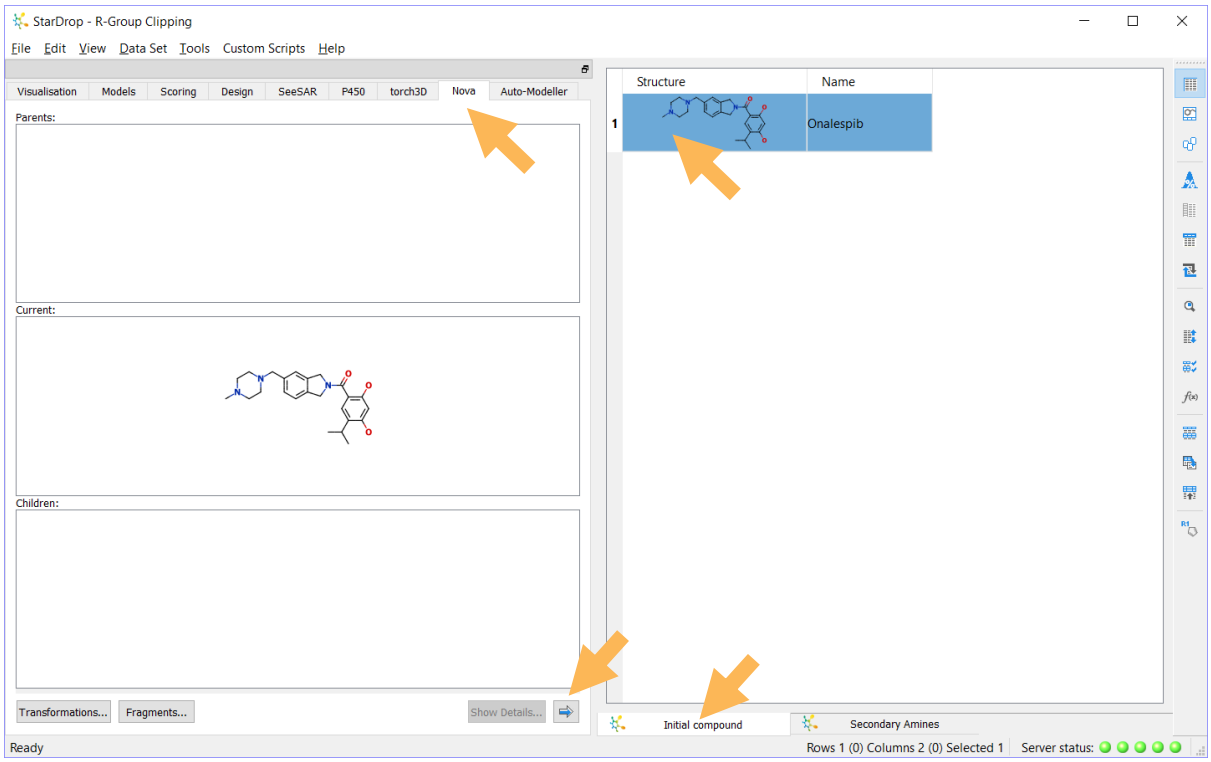
The fragments will be generated in a new column called **Fragment1_0** as shown in the screenshot below with a * indicating the open valence.

Note: Some rows will not contain a fragment due to the exclusion criteria we specified. Examples are highlighted in the screenshot below.



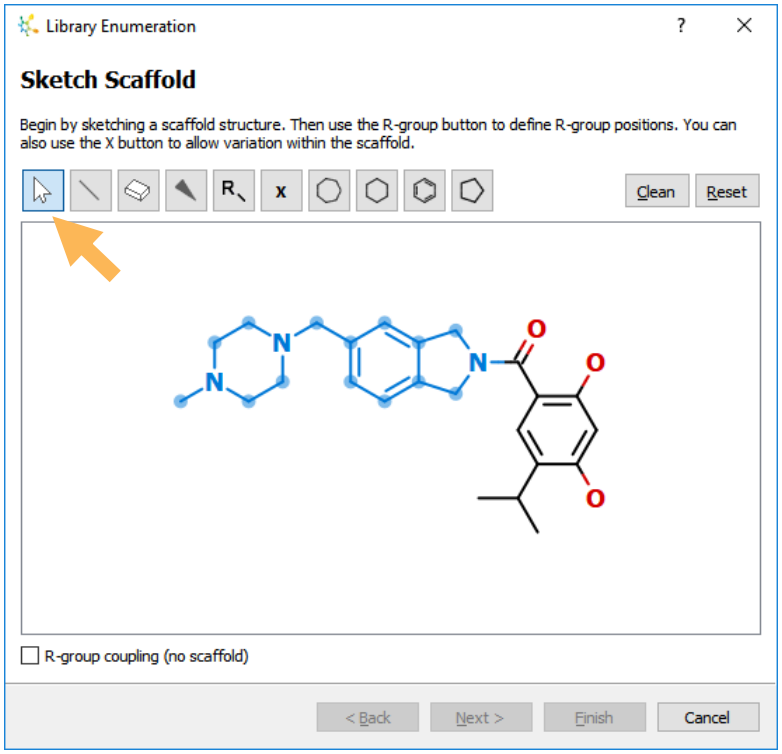
Using this set of fragments, we can now enumerate an amide library using a resorcylic acid scaffold derived from Onalespib. The structure of Onalespib is available in the **Initial Compound** data set which is also part of this project.

- Click on the **Initial Compound** data set tab and select the row containing Onalespib.
- Click on the **Nova tab**.

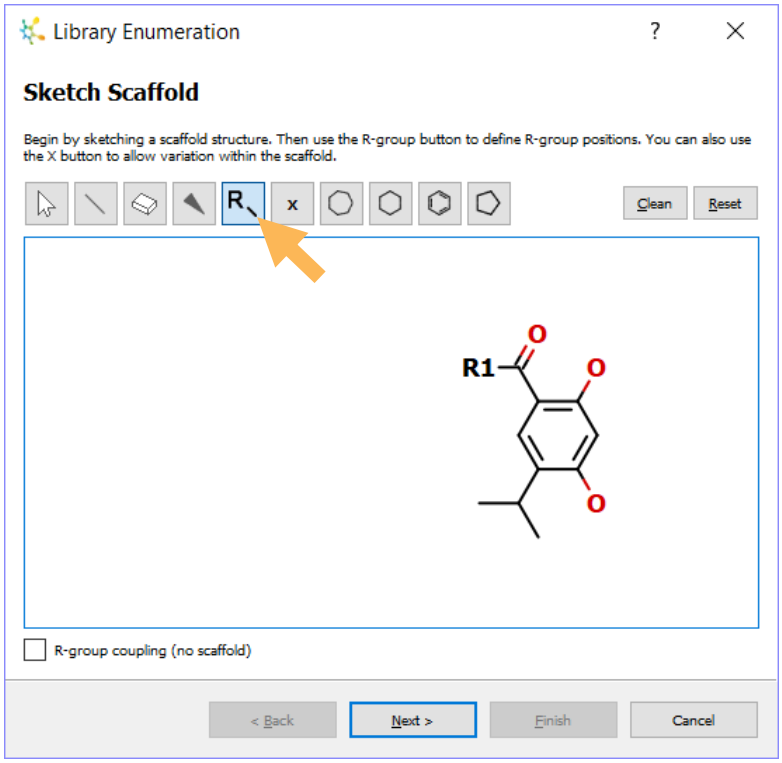


- Click  at the bottom of the Nova area to start the enumeration. In the wizard that appears, select **Library Enumeration** and click **Next**.

The **Sketch Scaffold page** will be shown containing Onalespib. If desired, we could sketch a new scaffold by clicking the **Reset button**, but in this case, we'll edit the displayed compound to create the scaffold for our new library.

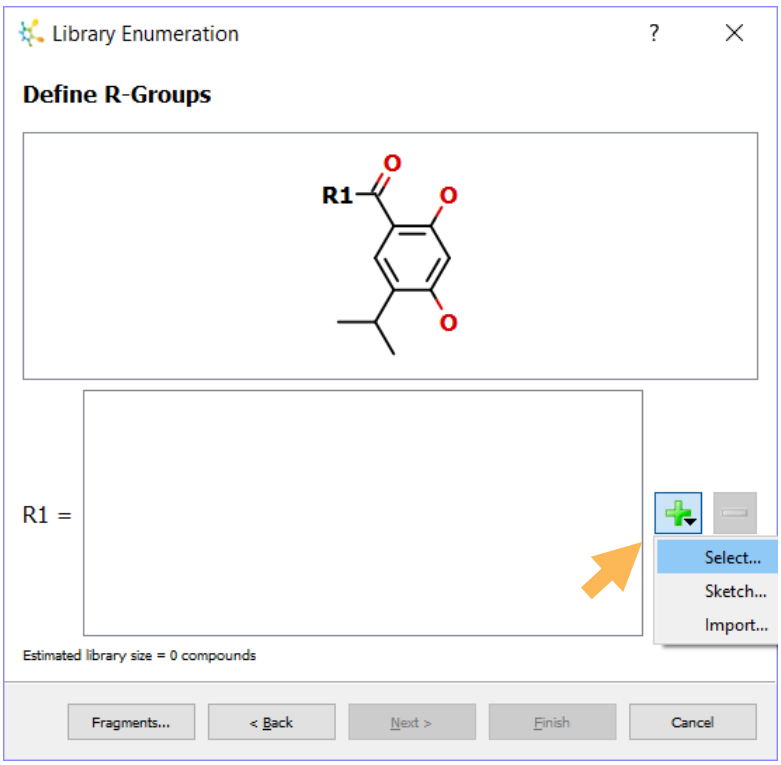


- Use the **Select tool**  to lasso the amine portion of the molecule.
- Click the **Delete key** to delete the highlighted atoms.
- Use the **R-group tool**  to add an R-group by clicking on the atom to which it should be connected.
- Click **Next**.

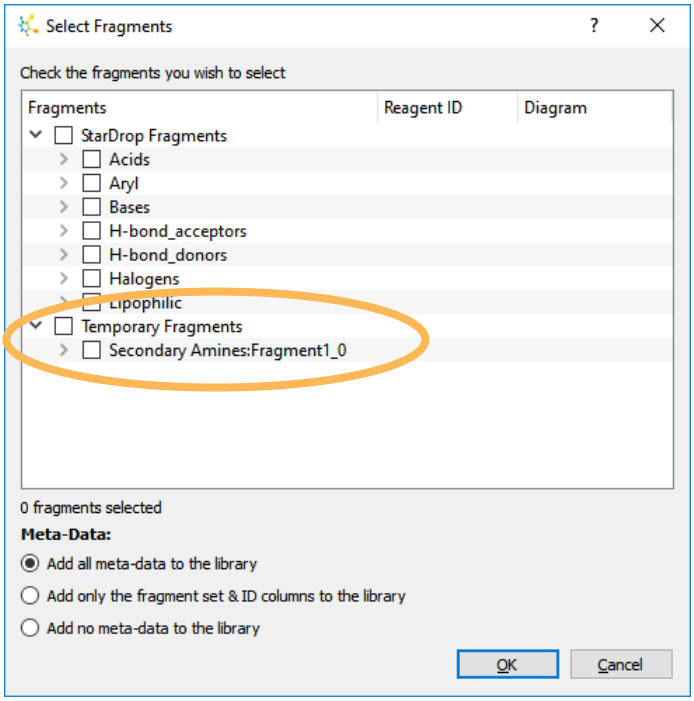


The **Define R-groups page** is displayed. Here we will define the list of secondary amine fragments to use in the enumeration.

- Click the **Add button**  next to **R1** and choose **Select** to open the library of predefined substituent groups.

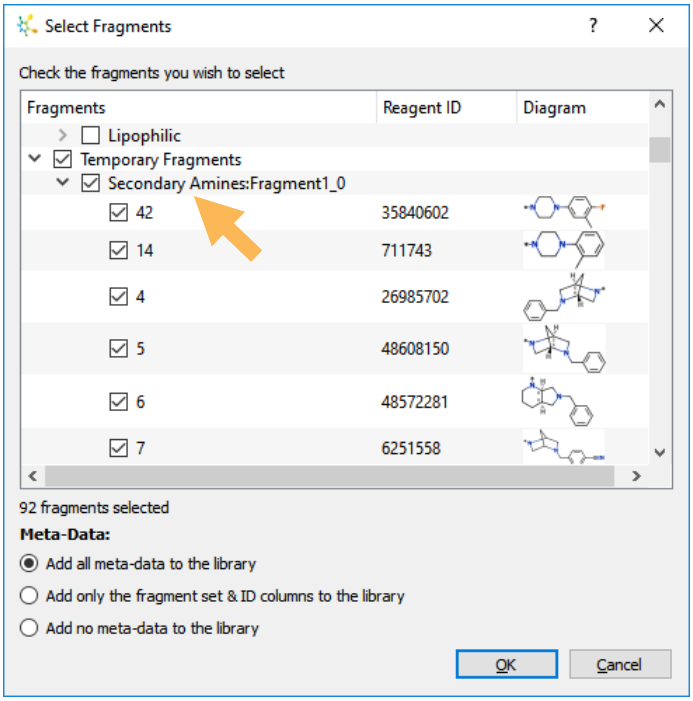


In the fragment library you will see all the fragments that have been previously saved. The fragments derived from the R-group clipping of the amine library are available at the bottom of the list. They are listed as "Temporary Fragments" because they are from one of the project data sets and have not explicitly been added to the library for future use in other StarDrop projects.



Hint: To add a set of fragments to the library permanently, right-click on the fragment column header in the data set and choose Add Data Set to Fragment Library from the menu.

- Tick the box next to **Secondary Amines:Fragment1_0** to select these fragments.



The **Meta-Data** options enable you to specify what, if any, data from the fragment library are added to the new series data set.

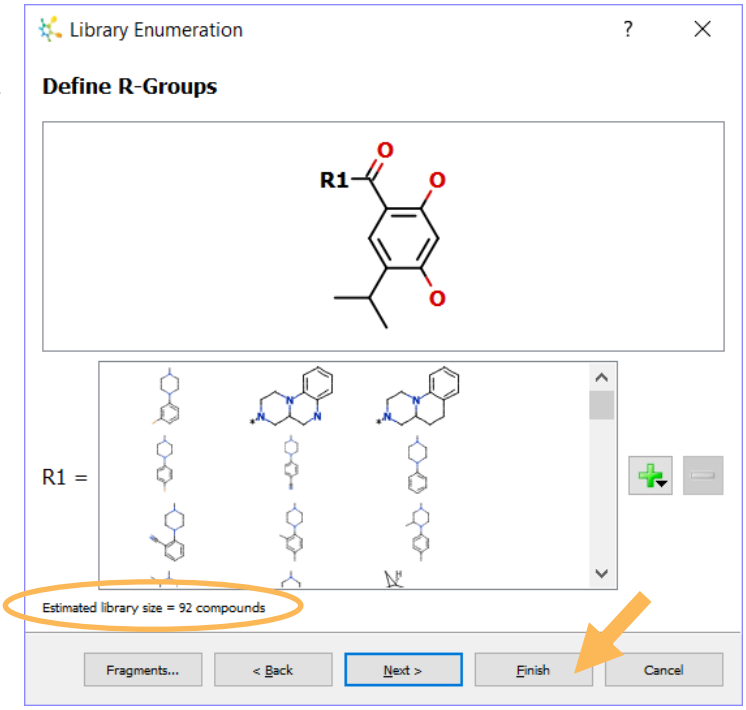
- Select the **Add all meta-data to the library** option and click **OK.**

Note that with this selection, the columns of information imported from eMolecules will be added to the enumerated library of structures making it easy to see which reagents are required for each of the virtual compounds.

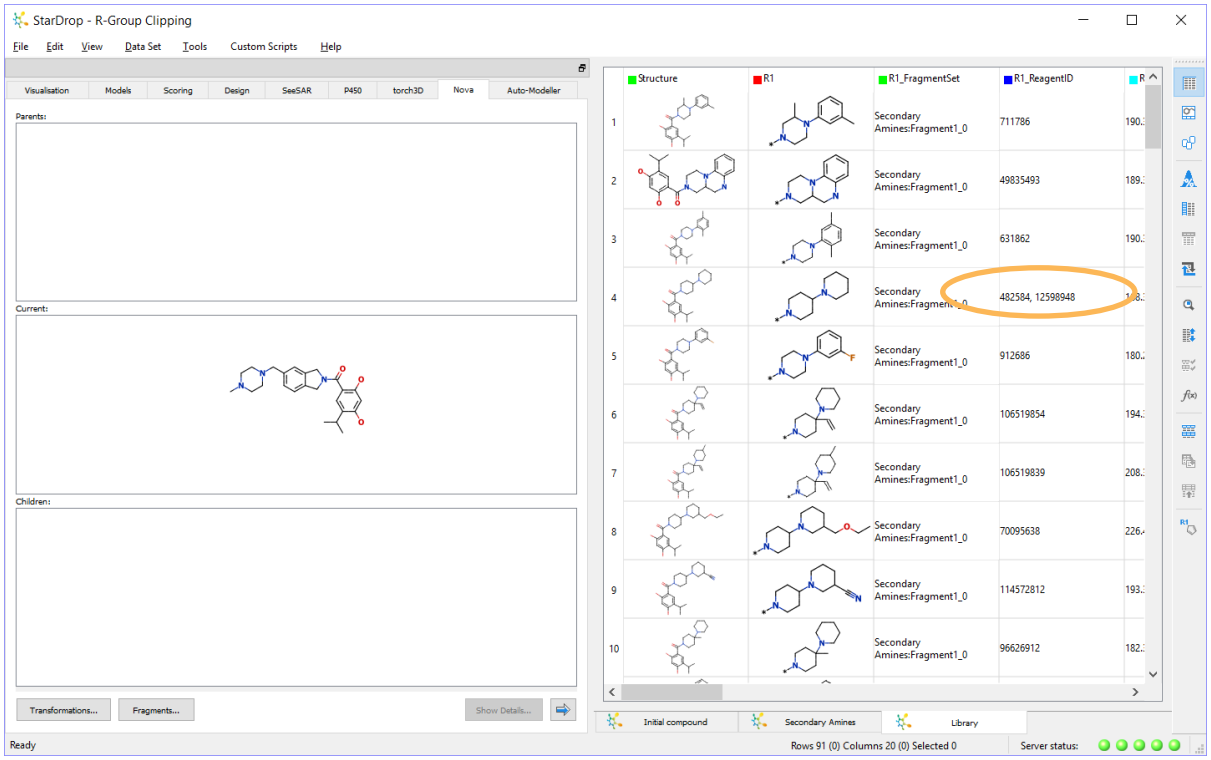
The fragments selected will be shown next to R1. If we wish to add further fragments we can do so by clicking the **Add button**  again, but in this case, we will just use the fragments we already have.

The estimated library size is 92 compounds.

- Click the **Finish** button.



A new data set will be added to the project called **Library**. It contains 91 structures along with all the reagent meta-data from eMolecules.



Note: Two of the fragments generated during the clipping process were duplicates, originating from different starting amines resulting in there being 91 unique enumerated compounds. The meta-data for both fragments have been preserved for this compound, as highlighted above.

Hint: Scroll to the right to see additional meta-data from eMolecules including the hyperlinks to the reagents in the eMolecules web site for convenient reagent ordering.

The data set can now be used with all StarDrop's capabilities for optimising and selecting compounds. The library can also be evaluated by docking in the HSP90 binding site using StarDrop's Pose Generation Interface to provide seamless integration with docking models from third-party platforms. Whilst this is beyond the scope of this exercise, if you would like to learn more about the Pose Generation Interface, please visit the following link in our online community videos:

https://www.optibrium.com/community/videos/introduction-to-stardrop-modules-and-features/375-pgi