

KNIME:

DMPK: Metabolic Hotspots and reactions

Drug Discovery Unit Training Academy
TRAINING

contents

1.0	About this exercise	Page 3
1.1	Installing KNIME	Page 3
2.0	Preparing the Workflow	Page 4
2.1	Opening the Workflow	Page 4
2.2	Adjusting KNIME preferences for Chemistry-enhanced workflows	Page 6
3.0	Using the Workflow	Page 8
3.1	Nodes	Page 8
3.2	Processing the metabolic hotspots table	Page 9
3.3	Highlighting the molecule's functional groups	Page 11
3.4	Let's get chemical!	Page 13
3.5	Does it match reality?	Page 14
4.0	And now...	Page 15
5.0	Answers for in-text questions	Page 16

1.0 About this exercise - DMPK: Metabolic Hotspots and reactions

In this exercise, we will make use of the **KNIME Data Analytics Platform** software to highlight possible metabolic hotspots present in a given drug's structure and which metabolic reactions may be occurring on these hotspots. In addition, the software will perform the reactions occurring at these sites and allow the user to view the products.

1.1 Installing KNIME

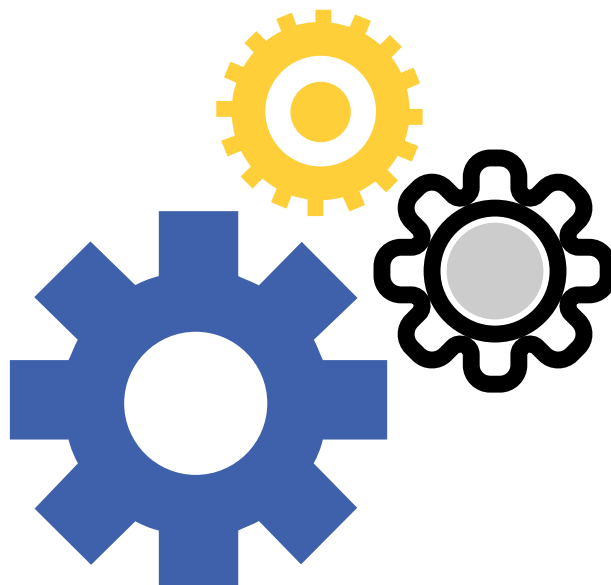
Download KNIME at www.knime.com/downloads. If you don't have administrative privileges on your machine, use the "self-extracting archive" option. If you choose this option, you have to manually extract the files of the downloaded package into a folder you can create and edit.

Note: if you're using the self-extracting setup, you must run KNIME manually by finding the executable file "knime.exe" in the folder you extracted the contents of the downloaded file into.

Also, make sure you have saved the two files required for this exercise to be run in to a folder you have access to:

- A KNIME workflow (.knwf) file, "DDU-TA_DMPK_Chemistry.knwf"

As a particular reminder, this workbook was prepared using KNIME 5.4, the latest version available as of January 2025. KNIME's user interface ("Modern UI") has drastically changed from version 4 and the figures in this workbook will reflect that change.



2.0 Preparing the Workflow

2.1 Opening the Workflow

After opening KNIME, you'll be presented with the software's Home screen. Click on "Local space" and then click "Import workflow" (**Figure 1**).

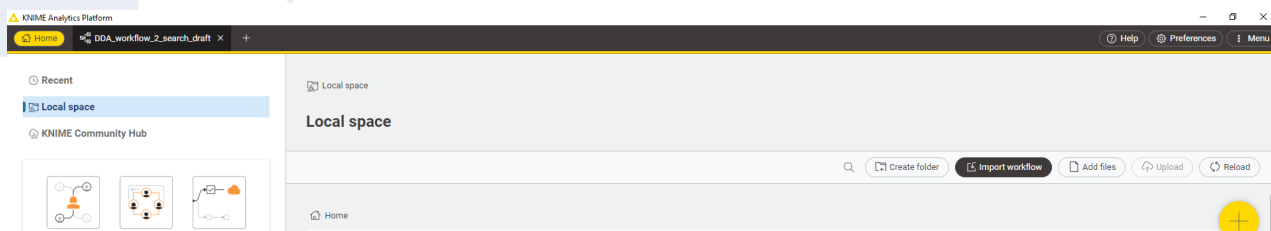


Figure 1 KNIME home screen highlighting the "Import workflow" option.

A dialog box will open. Select the workbook file and press "Open" to open the workflow.

In "Source:", click "Select File", then press the "Browse" button to choose the workflow file (**Figure 2**).

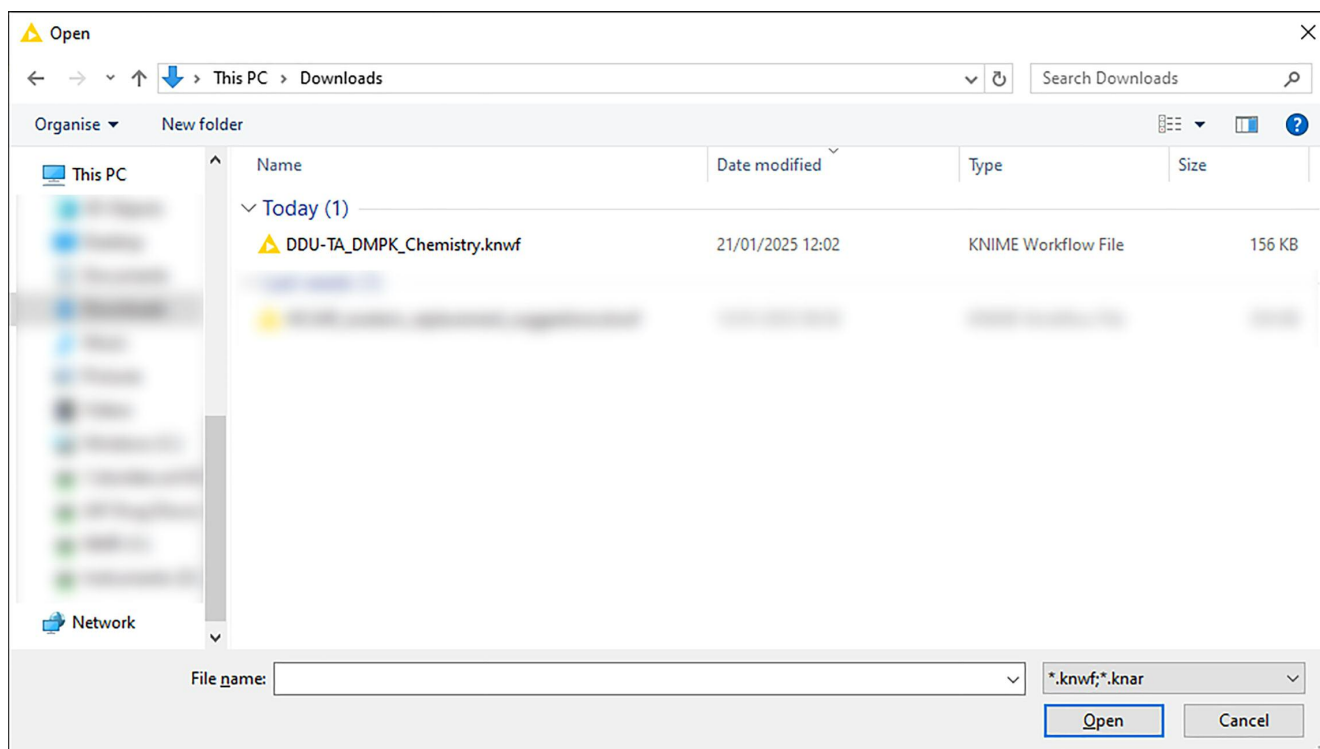


Figure 2 Import dialog box for KNIME.

A new file will appear on the **Local space** (Figure 3). Double click on it to open the workflow.

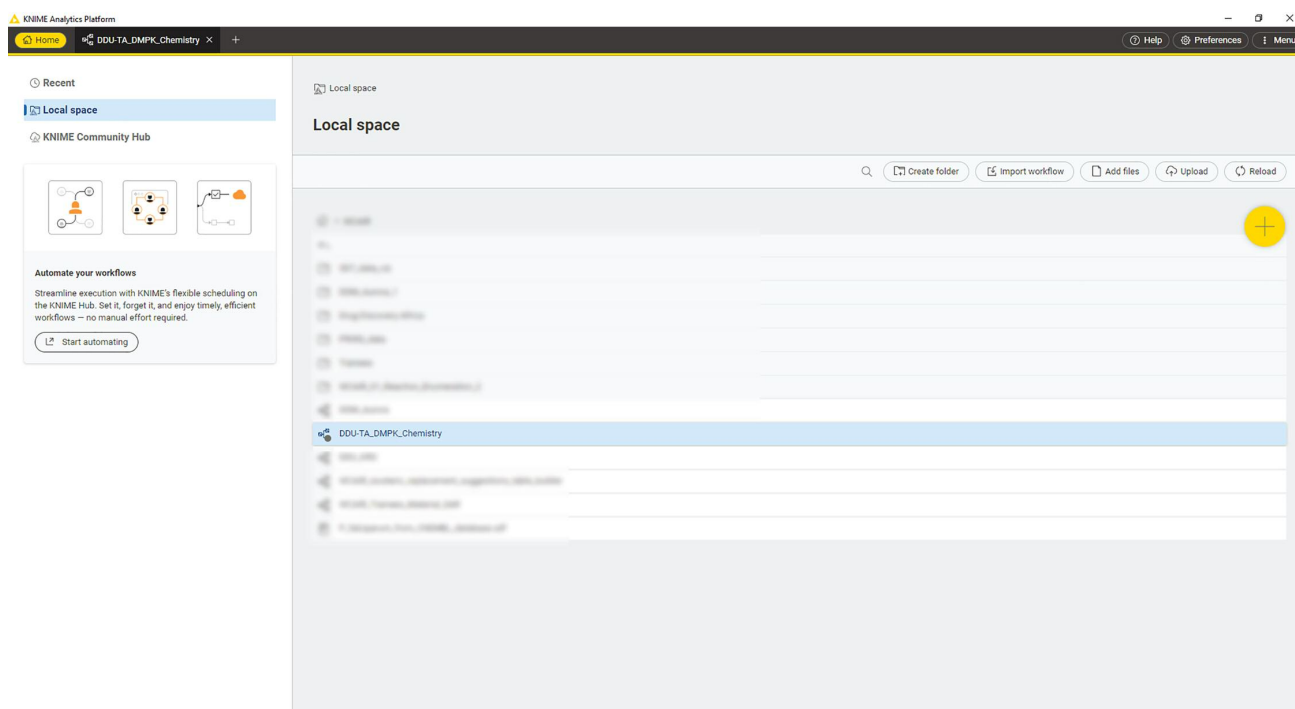


Figure 3 KNIME "Local space" page showing the imported workflow file.

If this is the first time you have used KNIME, KNIME will ask you to install the extensions required to run the workflow (Figure 4).

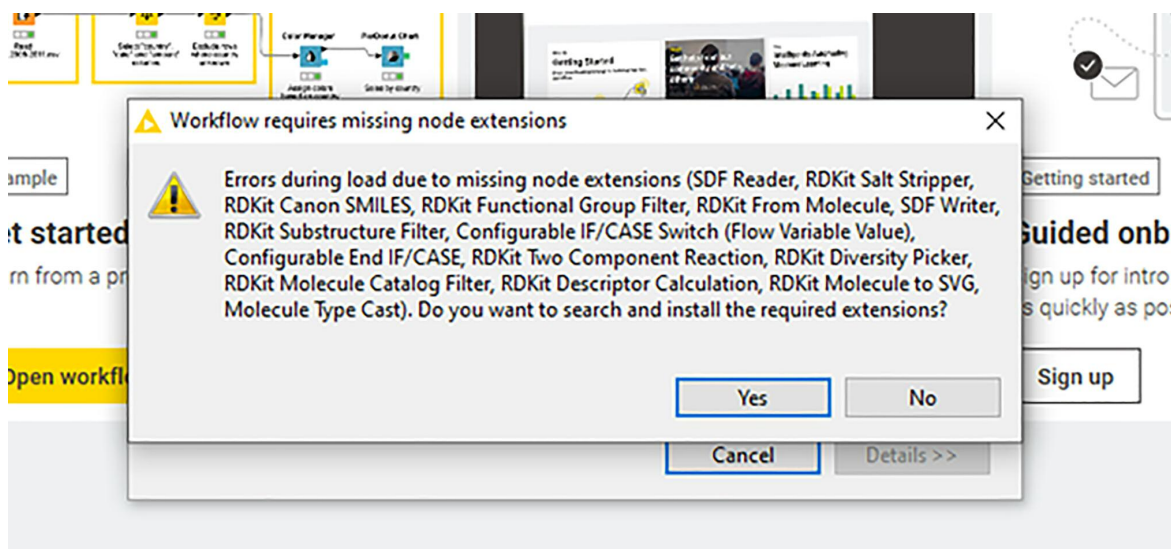


Figure 4 Error message indicating missing nodes that need to be installed to open the workspace. Note: depending on which extensions you may have already installed, the list may be different from the one in the figure.

Click “Yes” and, in the new dialogue box, click “>”.

Ensure that “I accept the terms of the license agreements” is selected prior to clicking “Finish”. The necessary nodes have now been installed. After installing the nodes, KNIME will ask to be restarted. Accept the restart and open the workflow again.

2.2 Adjusting KNIME preferences for Chemistry-enhanced workflows

To view the chemical structures, we need to set the proper renderer for KNIME.

Click on the Preferences button on the top right corner of the KNIME interface.

A dialog box will open (**Figure 5**).

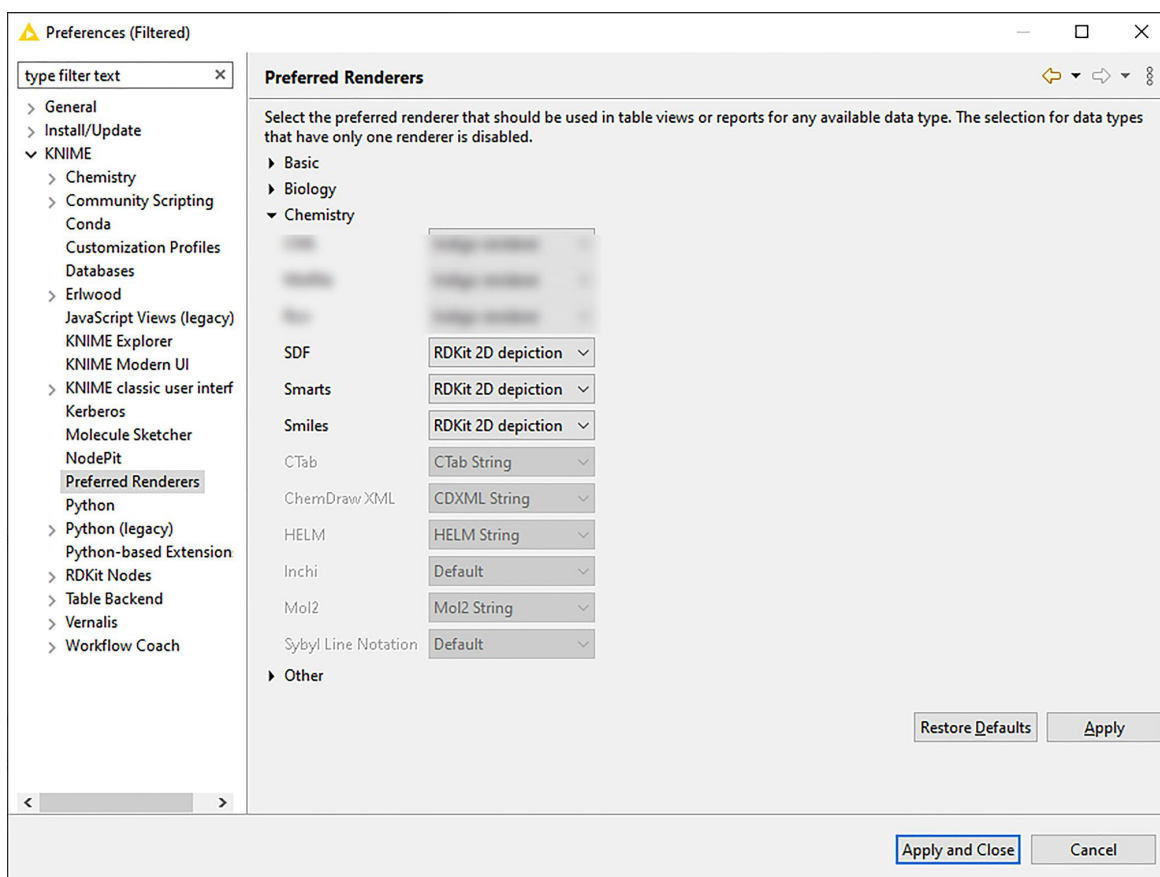


Figure 5 KNIME Preferences dialog box with the 'Preferred Renderers' pane open.

On the left pane, expand KNIME > “Preferred Renderers” > “Chemistry”.

Make sure that the renderer for “SDF”, “SMARTS” and “SMILES” is the “RDKit 2D depiction” *

* **RDKit** was from one of the extensions you installed in the previous step, and this will tell KNIME to use RDKit to recognize, interpret, and render (i.e. draw) chemical structures.

On the left-hand pane, click “KNIME Modern UI” (**Figure 6**) and make sure that the option “All nodes” is selected instead of “Starter nodes”.

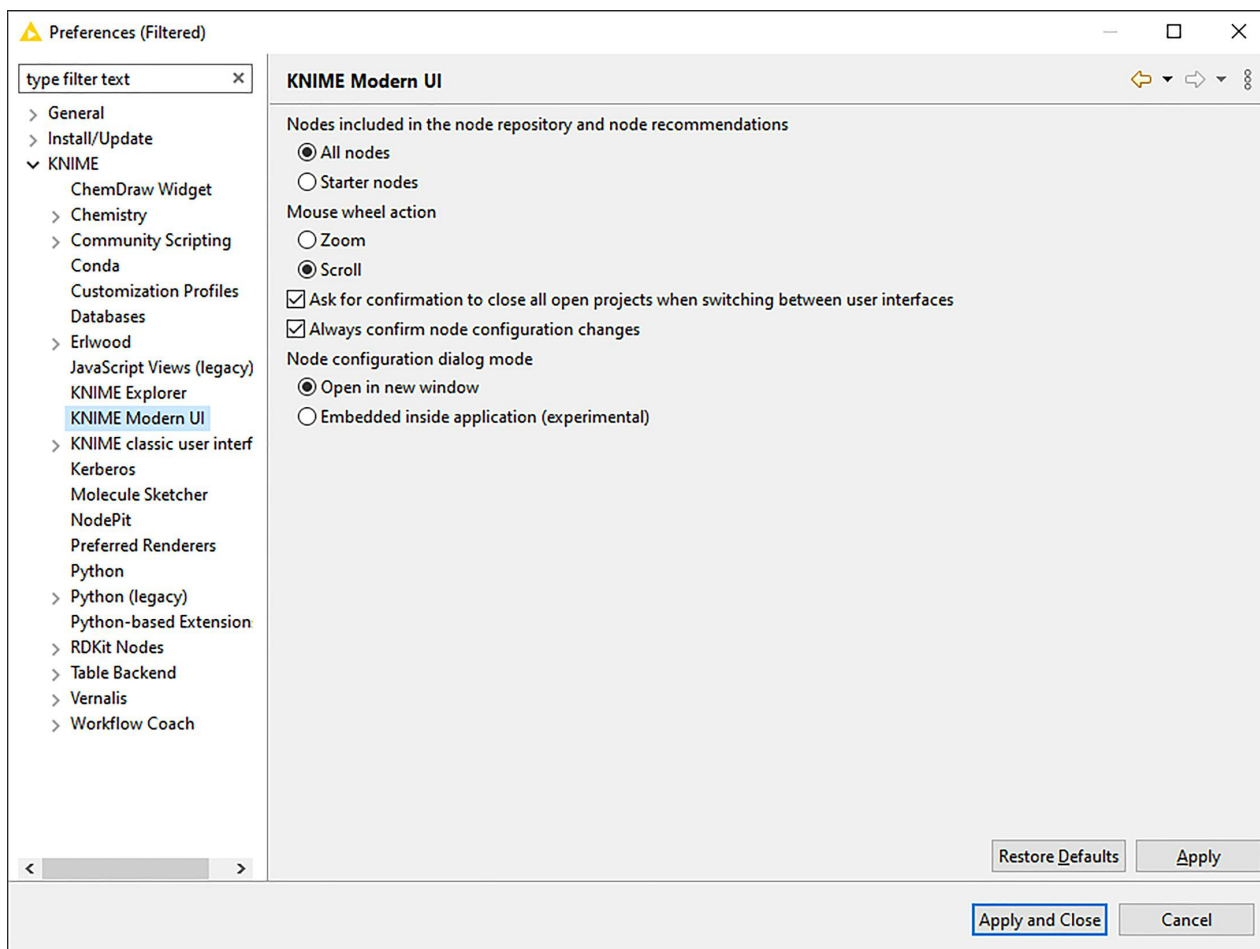


Figure 6 KNIME Preferences dialog box showing the “KNIME Modern UI” options.

Press “Apply and Close” when done.

3.0 Using the workflow

With the workflow open, KNIME interface will show the file as a graphical workflow (**Figure 7**).

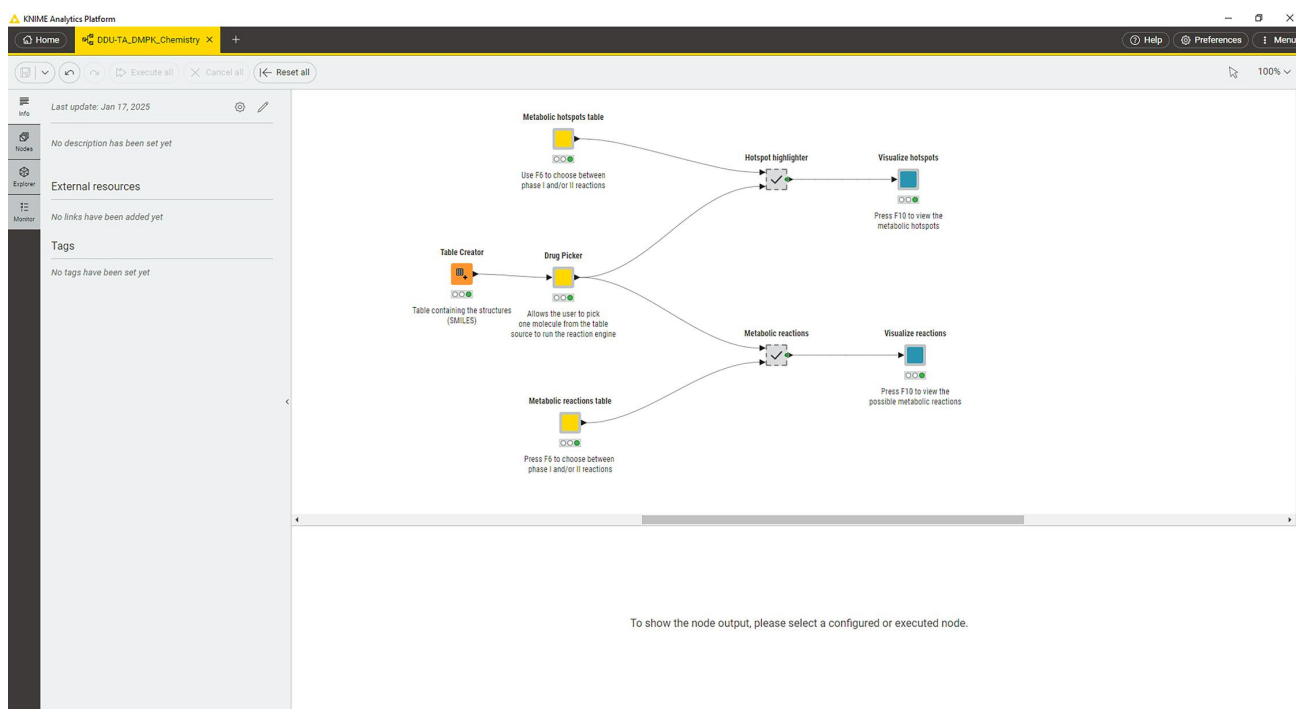


Figure 7 KNIME with the workflow loaded.

3.1 Nodes

With the workflow loaded, you will see several squares connected by arrows.

Each square, connected by arrows, is called a “node” (**Figure 8**). A node indicates a data transformation. KNIME compartmentalises the transformation and allows the user to see what is happening after each node execution.

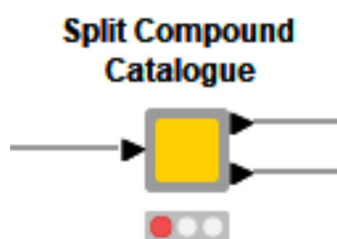
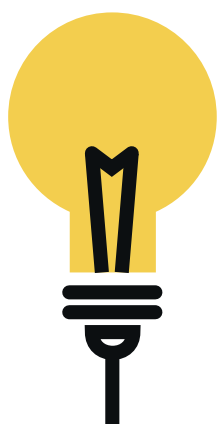


Figure 8 A KNIME node.



An arrow to the left of the node is the “input port”, i.e: the data table that will be transformed by the node. One or more arrows to the right are the “output ports”, or the data tables that are the result of the transformation done by the node.


The traffic lights at the bottom of each node denotes the node state.

- A **red light** means that the node has not been executed yet or is not ready to be run (requiring configuration)
- A **yellow light** means that the node is configured and ready to be executed
- A **green light** means that the node was successfully executed.


You will notice that all the nodes are marked as yellow or red because they are not ready to be executed. You need to configure the nodes so they can be properly executed.

3.2 Processing the metabolic hotspots table

We will begin from the top, with the “Metabolic hotspots table” component.

Select the node > press  (execute) or F7 (**Figure 9**) to run the component.

The component has been preset, for this exercise, to filter only phase I metabolic hotspots.

To change the preset, press  or F6 to configure with a different filter at your convenience (**Figure 9**).

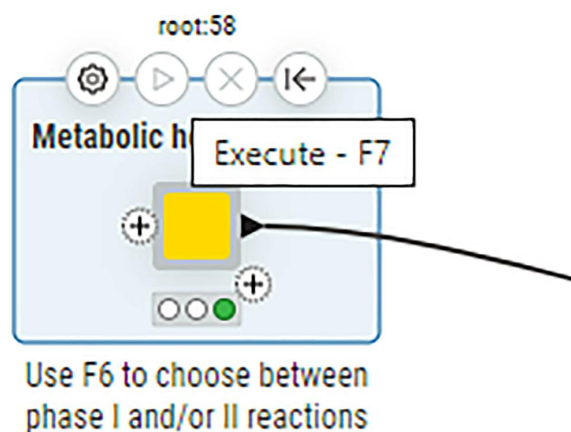


Figure 9 Metabolic hotspots table component highlighting the icon options.

You can inspect the component output using your right click button > “Open Output Port” > “Table” (**Figure 10**).

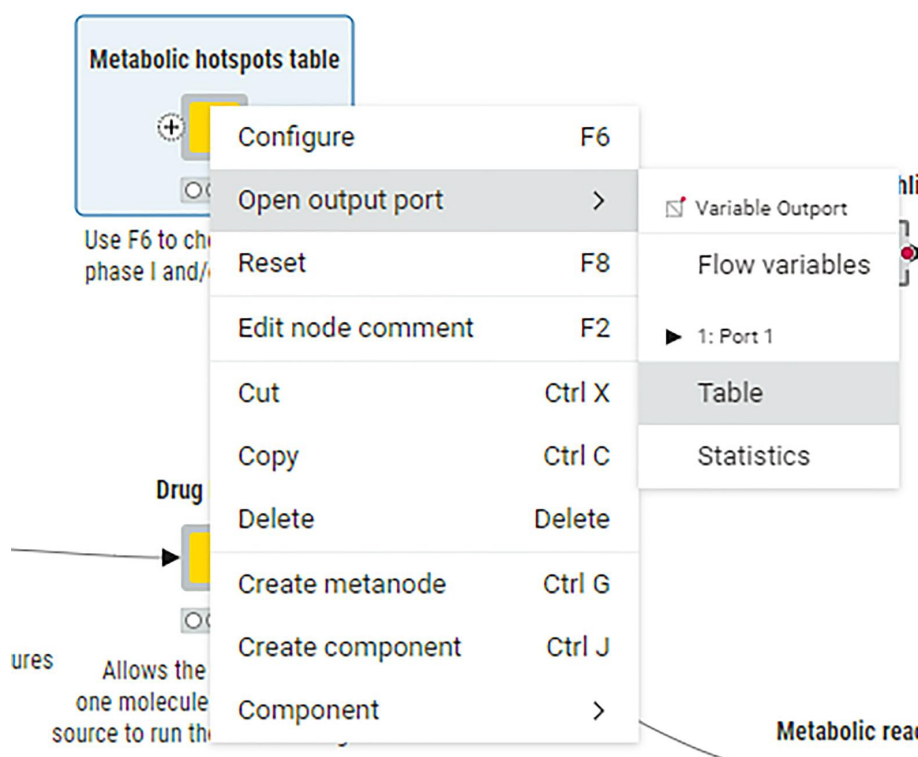


Figure 10 Opening the output port of a KNIME component or node.

How many phase I reactions are present in the Table?

Now for KNIME to read the “Drug Picker” component.

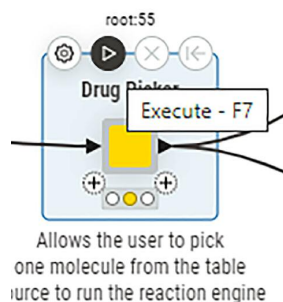
Click on the “Drug Picker” component >  or press F7 to execute the node (**Figure 11**).

The node has been preset to choose the anti-retroviral drug Delavirdine for this exercise.

*You can add other drugs to the “Table Creator” node and configure it to choose other substrates. We will leave it unchanged for this exercise.

Figure 11

“Drug Picker” node highlighting the “execute” option.

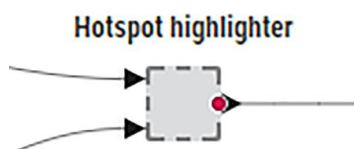


With both the “Metabolic hotspots” and the “Drug Picker” correctly executed, we can run the “Hotspot highlighter”, which will show you where in Delavirdine’s structure there are metabolic hotspots (if any).

3.3 Highlighting the molecule's functional groups

The next step in this exercise is to find and highlight the drug's metabolic hotspots. We will use the "Hotspot highlighter" metanode to do this (Figure 12).

Figure 12 "Hotspot highlighter" metanode.



Briefly, this metanode looks for all the hotspots defined in the "Metabolic hotspots" component and search for them in the picked drug's structure. Any functional group found in the drug's chemical structure will be highlighted and prepared for visualization.

Click on "Hotspot highlighter" >  or press F7 to execute the node.

Examine the metanode output. How many rows does this table have?

A closer look at the metanode output can provide you with interesting information. Figure 13 shows an example of the expected output. Some columns are detailed below.

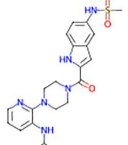
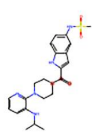
RowID	Drug String	Structure Smiles	Ch... String	C... St...	M... List	Hotspot description String	Hotspot (Highli... SVG image	Site for String	Metabolic Phase String
Row2	Delavirdine		ChEMBL: 136817	[14,15;	Ester or amide			Ester or Amide hydrolysis	1

Figure 13 Details of the "Hotspot highlighter" metanode output table.

- **Hotspot description:** description of the functional group that can be metabolized
- **Hotspot (Highlighting):** the chemical structure of the drug with the hotspot highlighted (in red)
- **Site for:** which reaction(s) this functional group can undergo while it is metabolized

Apart from the reaction on Figure 13, which other reactions can you see in the result table? Are they familiar to you?

The same information in the table (Figure 13) can be seen in a more user-friendly way by using the “Visualize hotspots” component.



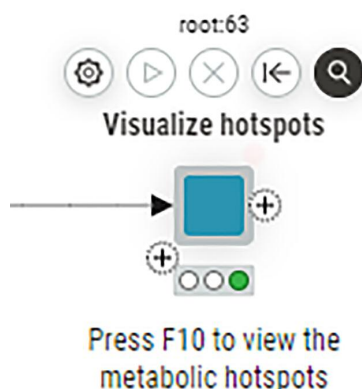
Click on the “Visualize hotspots” component > “Execute and Open View” or  followed by  or press F10 (Figure 14).

Figure 14 The “Visualize hotspots” component.



A new window will open (Figure 15) displaying the results from the “Hotspot highlighter” in detail, presenting every metabolic hotspot that falls under a certain criterion in a different card.

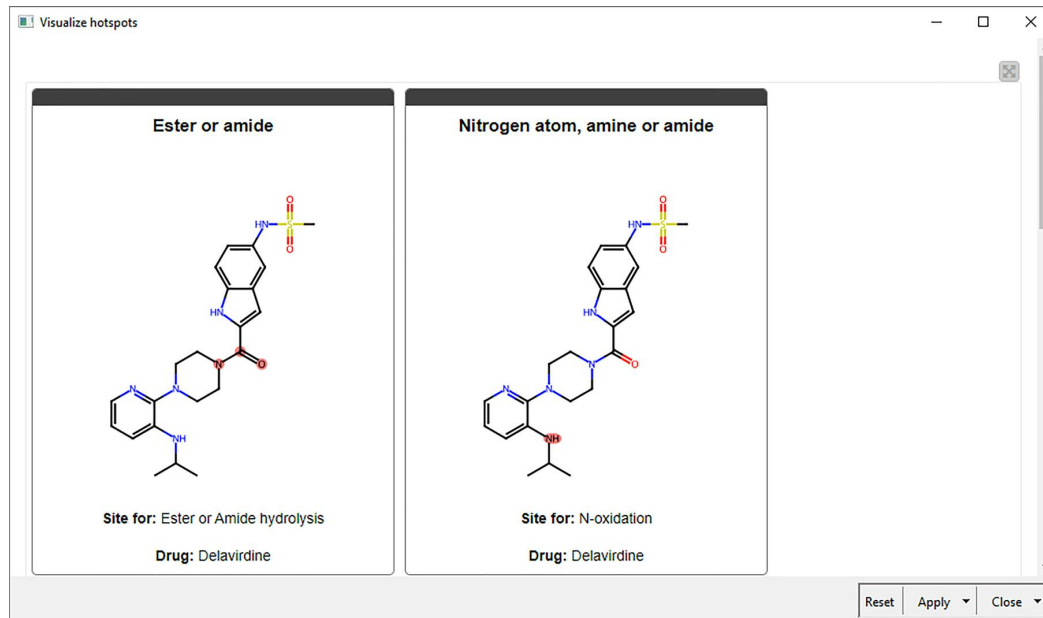


Figure 15 The interactive view from the “Visualize hotspots” component.

Looking at the different “cards” showing the metabolic hotspots, is there anything that seems strange to you?

When you close this window, you can safely “discard changes” when closing it. This component is not passing information further on, so nothing will be stored in the workflow.

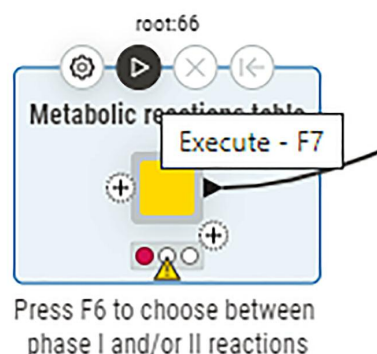
Now that we have seen the metabolic hotspots, what about the reactions themselves?

3.4 Let's get chemical!

To run the metabolic reactions on Delavirdine (our example drug), we need to execute the “Metabolic reactions table” component. This component will generate a table with the possible metabolic reactions that can happen for our example drug.

Click on the component >  or F7 (**Figure 16**).

Figure 16 The “Metabolic reactions table” component.



- * As previous, this component is also preset to run only phase I metabolic reactions. You may change this later by using the component's “configure” option or pressing F6. For this exercise, we will be only using phase I reactions.

Examine the component's output. How many reactions are there?

At first, you may feel that there's something wrong with the component, as the reactions have a “2D depiction failed” message. Don't worry about this. The queries are complex (we are depicting chemical reactions in generic terms using an extension of SMILES named **SMIRKS**) and RDKit sometimes cannot render them properly.

With that done, the next step is to run the two remaining components, “Metabolic reactions” and “Visualize reactions”, which we'll use to inspect the reactions and their results (products). Since the two nodes are connected, all that you need to do is to go to the last node – “Visualize reactions” – select it and choose “Execute and Open View” (**Figure 17**).

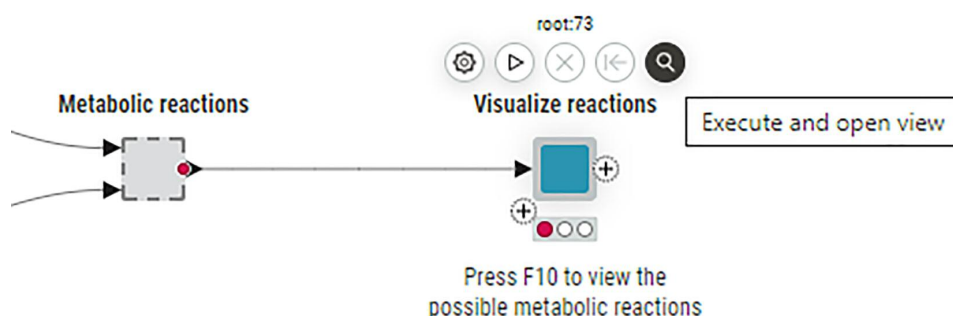
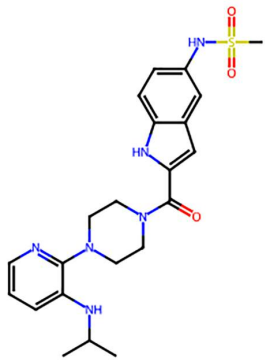
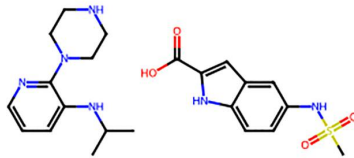


Figure 17 “Metabolic reactions” and “Visualize reactions” components.

A new window with an interactive table will open (**Figure 18**) showing, in each row, the chemical structure of the drug substrate, a description of the reaction being performed, and the structure(s) of the product(s).

The screenshot shows a window titled "Visualize reactions" with a "Show 10 entries" dropdown. It contains a table with three columns: "Substrate (SVG)", "Description", and "Product (SVG)".

Substrate (SVG)	Description	Product (SVG)
	Ester or Amide hydrolysis	

At the bottom right of the window are buttons for "Reset", "Apply", and "Close".

Figure 18 Interactive view of the "Visualize reactions" component.

3.5 Does it match reality?

The purpose of this exercise is to show which metabolism reactions can happen* with a given drug structure (in our case, the anti-retroviral drug Delavirdine).

*This exercise is not intended to give any kind of prediction regarding what will happen when a particular drug is metabolized. Our xenobiotic metabolic apparatus is comprised of several different enzymes and a large group of them are isoforms of the CYP450 monooxidase. Different isoforms have different selectivity for its substrates (some prefer planar molecules, some prefer basic compounds, other prefer neutral or acid molecules; most of them will metabolize more lipophilic compounds, etc.) meaning that what we see, in real-world case of a drug's metabolic pathway, is a complex chain of these reactions, as the drug's primary metabolites can be further metabolized.

Figure 19 shows the observed and identified metabolites for Delavirdine in rats (this was adapted from Chang, M., et al. *Drug Metab. Dispos.* **1997**, 25, 228-242) [[PUBMED link](#)].

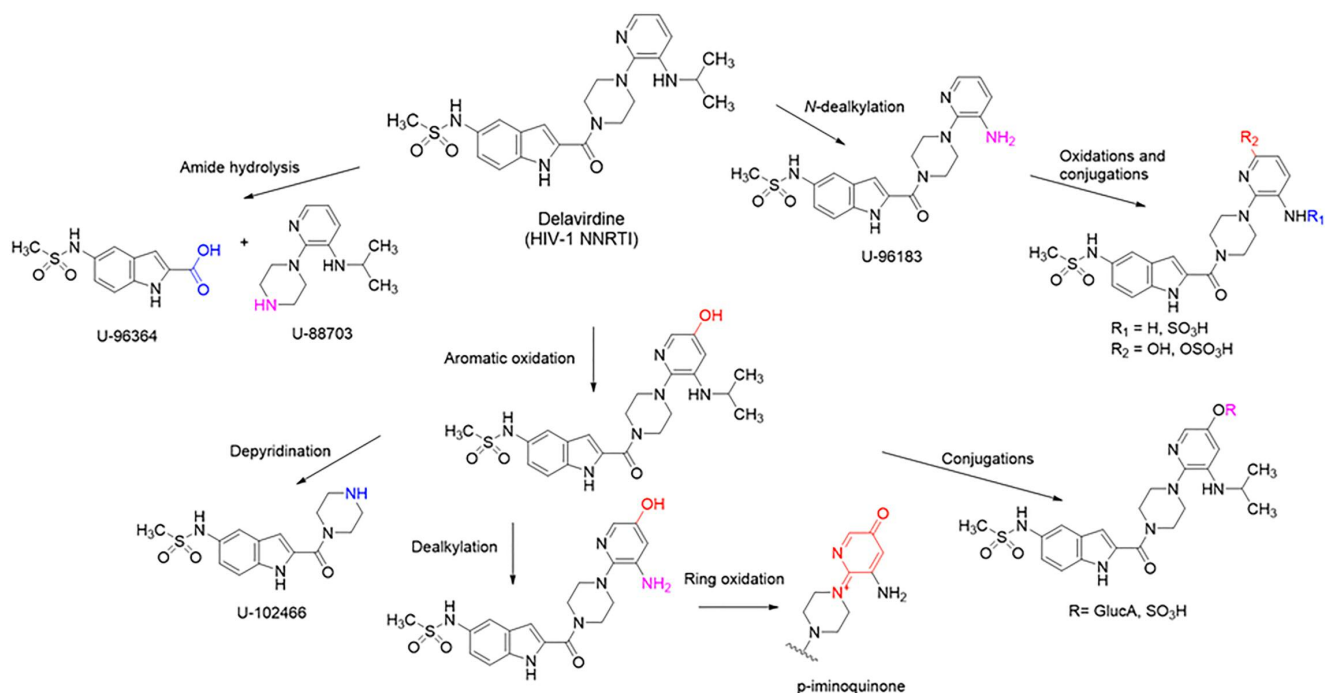


Figure 19 Metabolic pathway for the anti-retroviral drug Delavirdine. Adapted from Chang, M., et al.

Comparing your results from the “Visualize reactions” component and what it is shown on **Figure 19**, what is different and what it is not? What could be the cause for these perceived differences?

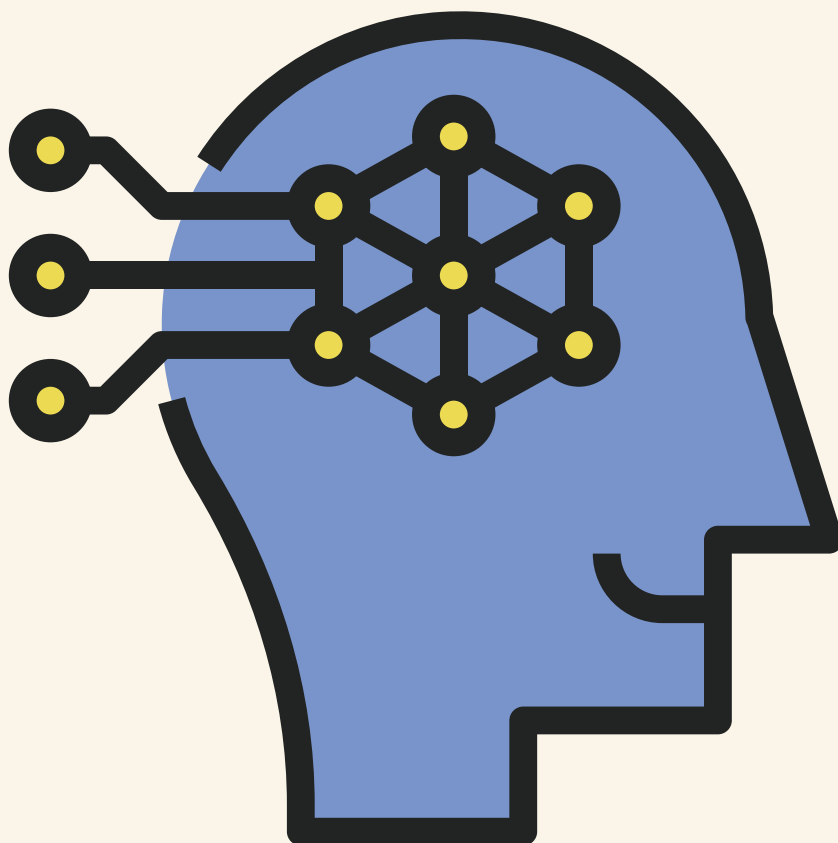
4.0 And now...

You know now how to use KNIME and its advanced data mining features to define metabolic hotspots and metabolic reactions in a programmable manner, how to apply these reactions on a drug substrate, and inspect the results in a user-friendly way.

And, as stated before, this is just the tip of the iceberg. KNIME can do a lot more for chemistry and life sciences-based applications. You can read more about it on the [KNIME website](#).

As a follow-up, the exercise has another drug that can be picked to run the workflow from the beginning (Propranolol). Why don't you give it a try? How would the exercise's result compare to the [list of known metabolites of this drug](#)?

5.0 Answers for in-text questions



5.0 Answers for in-text questions

3.2 Processing the metabolic hotspots table

How many rows does this table have?

There are 28 rows (reactions) on the “Metabolic hotspots table” output.

3.3 Highlighting the molecule’s functional groups

Examine the metanode output. How many rows does this table have?

The node can find 5 hotspots in Delavirdine’s chemical structure. Each one of them is represented as a row in the output table.

Apart from the reaction on **Figure 13**, which other reactions can you see in the result table? Are they familiar to you?

The reactions listed are:

- Ester or amide hydrolysis
- *N*-oxidation
- Aromatic ring C-H oxidation
- Oxidative dealkylation
- Pyridine ring *N*-oxidation

The answer for the second question may vary depending on how familiar you are with metabolic reactions. Normally, one reaction that is seldomly mentioned is the pyridine ring *N*-oxidation leading to a pyridine *N*-oxide as metabolite.

Looking at the different “cards” showing the metabolic hotspots, is there anything that seems strange to you?

There may be something strange around the aromatic ring *para* C-H bond and heteroatom aliphatic *alpha* C-H bond cards. It is worth noting that the aromatic ring being shown in the card has two substituents, and, therefore, there are two distinct *para* positions in that ring, one relative to each of the substituents in the ring. Something similar can be observed for the aliphatic *alpha* C-H bonds. Some substrates may have more than one valid C-H bond that falls into this category, and the component is able to locate and display all that matches in a single card.

One thing that is important to mention is that not all *para* C-H bonds have the same reactivity, though. **Figure 19** shows that Delavirdine’s pyridine ring is oxidized preferentially at the *para* position regarding the piperazine ring. U-96183 (the des-isopropyl metabolite) is oxidized at the *para* position regarding the free amino group.

3.4 Let's get chemical!

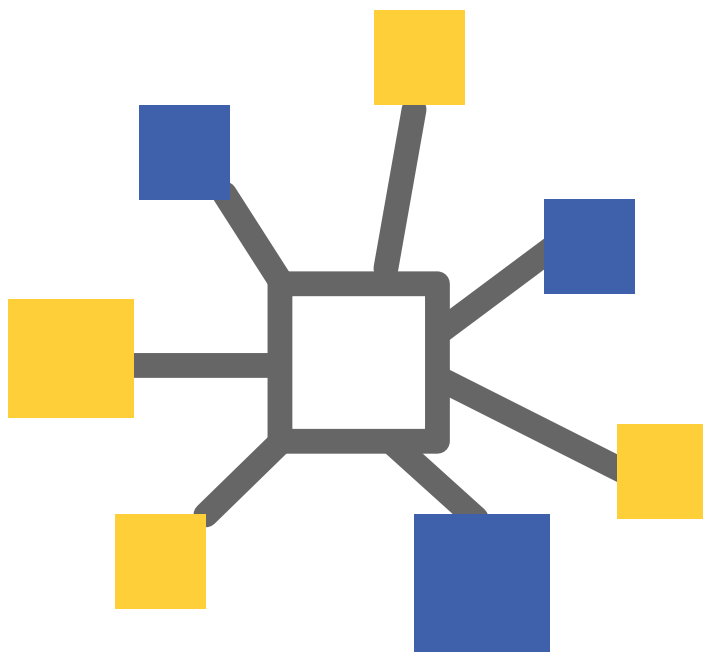
Examine the component’s output. How many reactions are there?

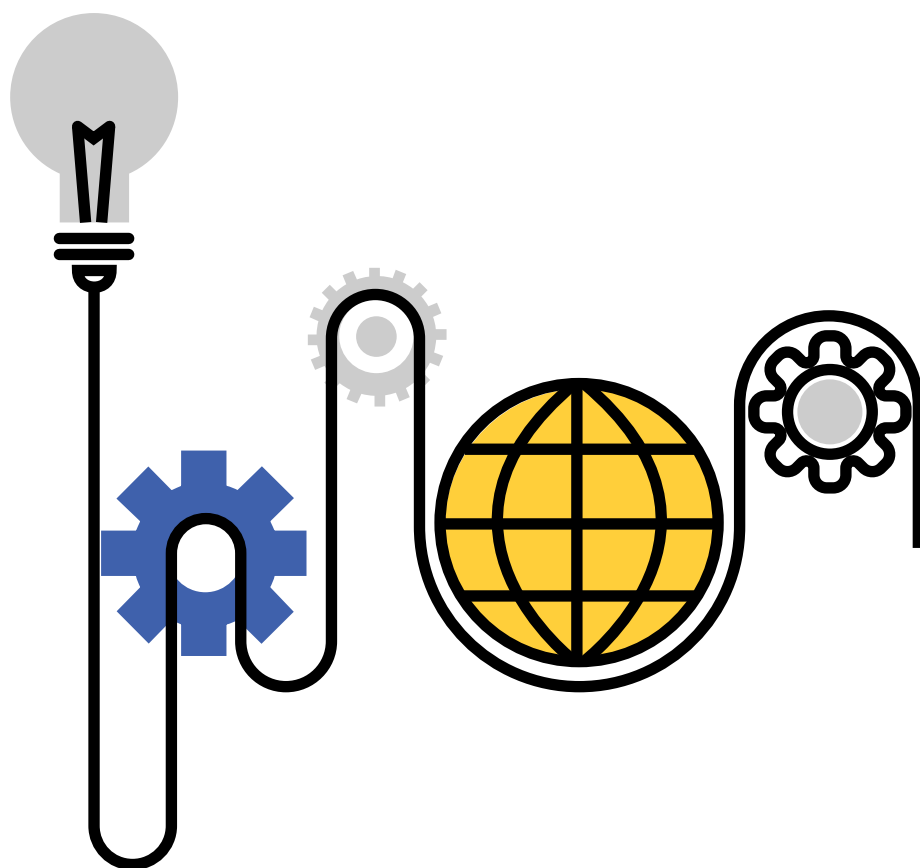
There are 33 distinct reactions being represented in the table.

3.5 Does it match reality?

Comparing your results from the “Visualize reactions” component and what it is shown on Figure 19, what is different and what it is not? What could be the cause for these perceived differences?

Some reactions are present in both (amide hydrolysis, oxidative N-dealkylations, aromatic C-H oxidations). What is missing in the results, at first, are the phase II conjugations (which we left out as part of the exercise), and the fact that the primary metabolites are also further metabolized into smaller, more polar structures. Also, one reaction that is worth noting is a very rare one - *depyridination* leading to the metabolite identified as U-102466 (Figure 19) - that happens **after** the pyridine ring is oxidized. Since this workflow is only running the metabolic reactions once, using Delavirdine as starting point, this exercise is not suitable to investigate what happens with the primary metabolites.





Drug Discovery Unit Training Academy

School of Life Sciences

University of Dundee

Dow Street

Dundee DD1 5EH

@ ContactDDU@dundee.ac.uk

<https://drugdiscovery.dundee.ac.uk/>



University
of Dundee