# PrePeP quick start

1. Run program
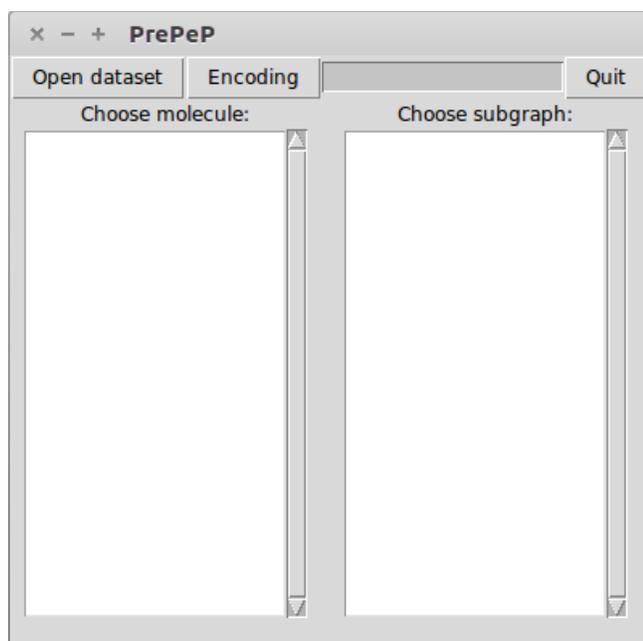
2. Press *Open dataset* and select a set of molecules in .sdf format (**demo.sdf** is provided as an example):
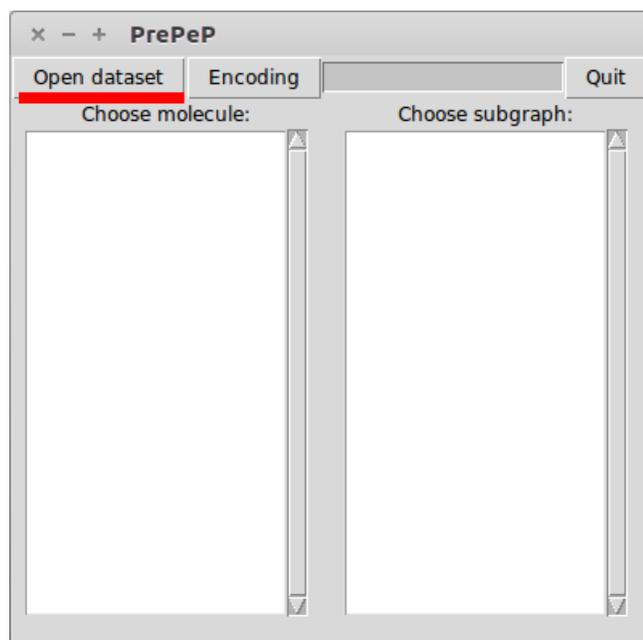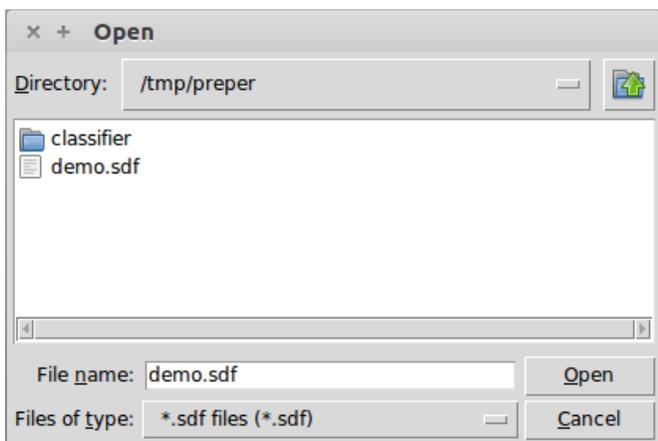
3. Press *Encoding* button. The encoding process will take some time depending on the size of your set. The progress bar will show the whole progress.
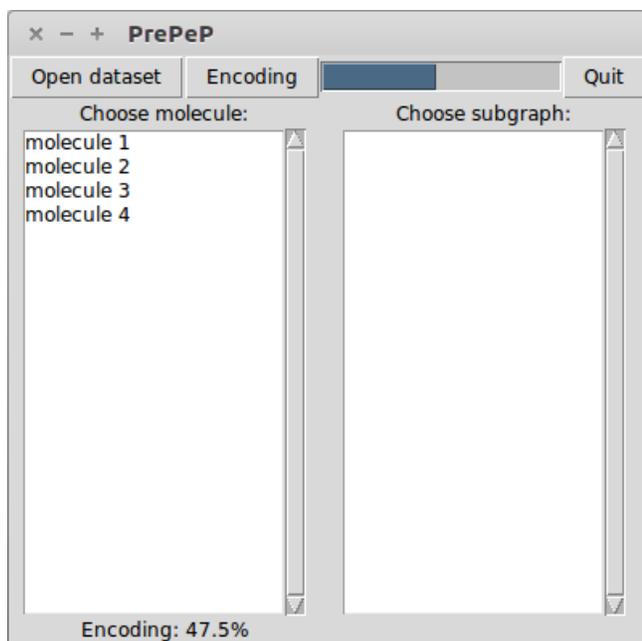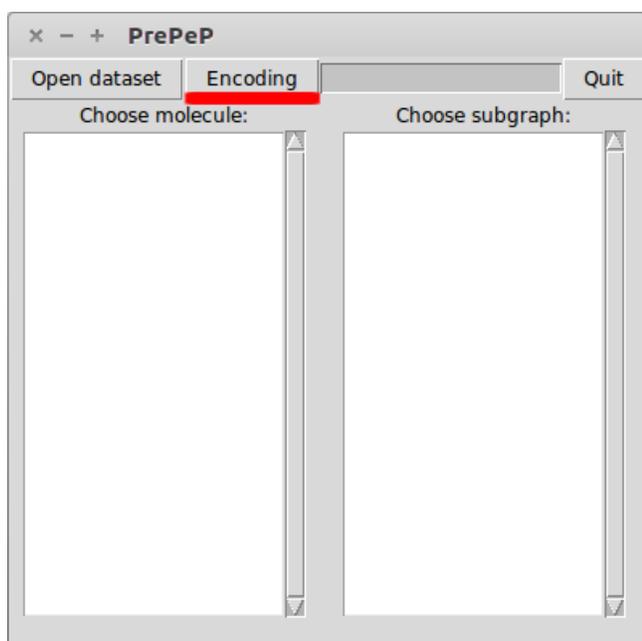
4. When encoding is done left-hand column will show the list of molecules presented in the dataset. To get a result of prediction click on a molecule in this list. The result will appear at the bottom of the left-hand column.
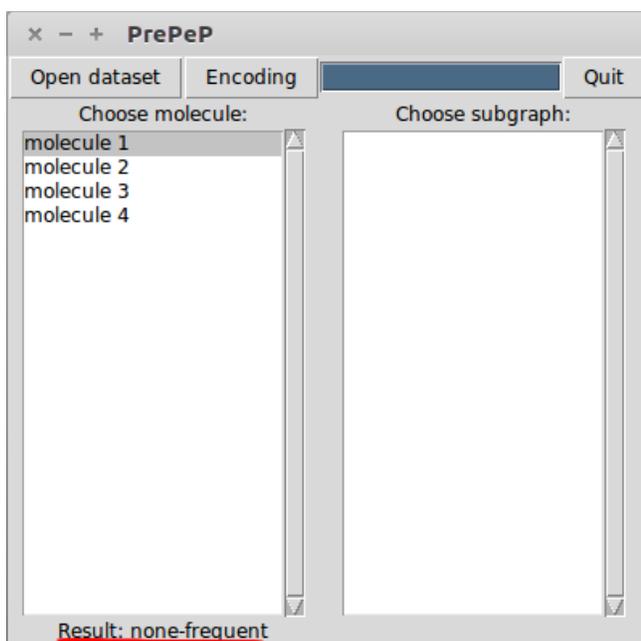
At the same time the molecule will be visualize.
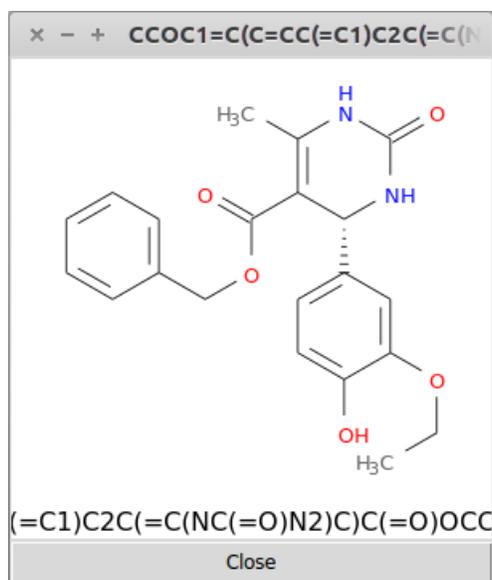
To close the visualization press *Close* button at the bottom.

5. For a molecule classified as frequent the list of discriminative subgraphs presented in the molecule will appear in the right-hand column.
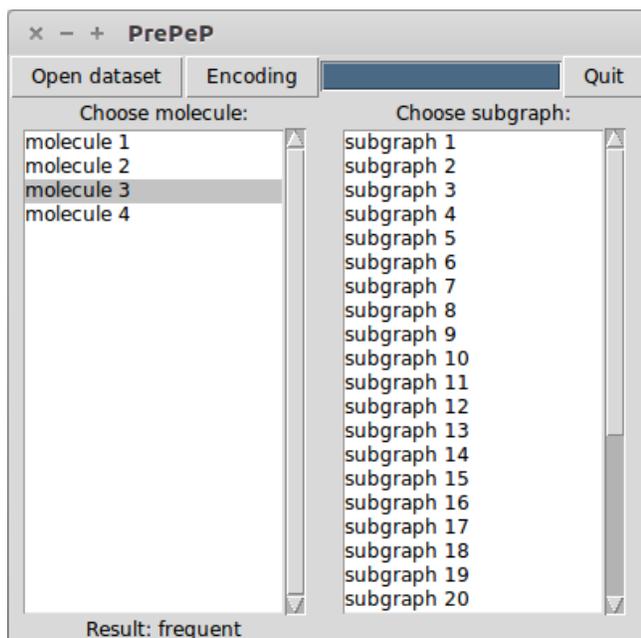
To visualize a discriminative subgraph first close the plot of the molecule, then select a subgraph in the right-hand column. The visualization will appear and corresponding frequency in the data will be shown at the bottom of the right-hand column.

N=C(/C=C/C=C)C

NH

CH3

H2C
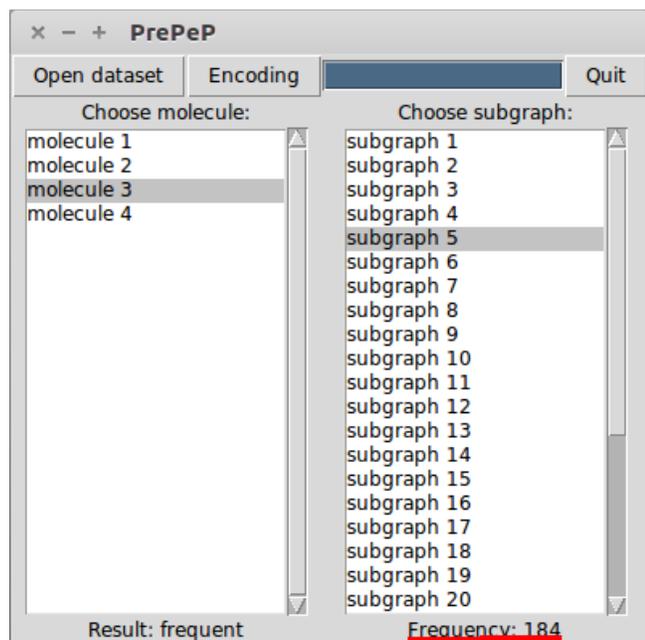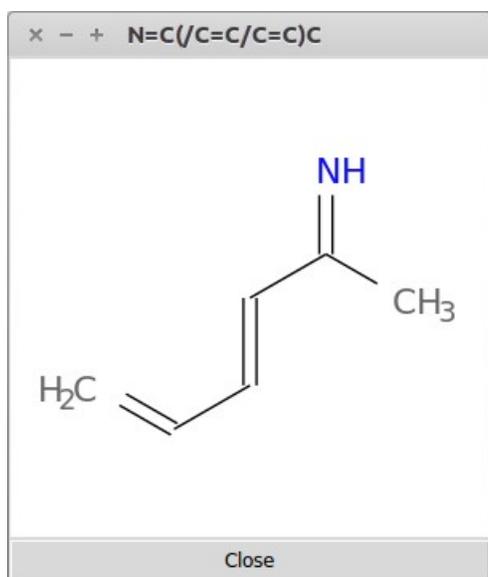
Close

6. To quit the program, first, press *Close* button at the end of the plot (if any), second, press *Quit* button in the main window.

PrePeP

Open dataset   Encoding   Quit

Choose molecule: Choose subgraph:

molecule 1   subgraph 1
molecule 2   subgraph 2
molecule 3   subgraph 3
molecule 4   subgraph 4
                    subgraph 5
                    subgraph 6
                    subgraph 7
                    subgraph 8
                    subgraph 9
                    subgraph 10
                    subgraph 11
                    subgraph 12
                    subgraph 13
                    subgraph 14
                    subgraph 15
                    subgraph 16
                    subgraph 17
                    subgraph 18
                    subgraph 19
                    subgraph 20

Result: frequent   Frequency: 184