Data Integration and Data Mining with ChemProt

In this exercise you will get hands-on training on our new web-server, ChemProt. ChemProt is a resource of annotated and predicted chemical-protein interactions that have been compiled from over 1 100 000 unique chemicals with biological activity for more than 15000 proteins.

ChemProt can assist in the in silico evaluation of small molecules (drugs, environmental chemicals and natural products) with the integration of molecular, cellular and disease-associated proteins complexes. It suggests diseases and side effects associated with protein targets, and provides biological information for a protein when it is part of a disease complex.

In this exercise you will familiarize yourselves with different features of ChemProt.

Go first to the ChemProt server:

http://www.cbs.dtu.dk/services/ChemProt-2.0/

and read the “Instructions” section. Then:

1) Search for compounds bioactivity

- Write the name Citalopram in “Type a Compound name” and click on “Submit”

On the heatmap that appears, how many proteins are annotated to this compound and how many similar compounds have you retrieved?

- Set the Ki threshold between 0.48 and 1 and tick on “Min” under “Display Settings: Values”
Which proteins are now associated with this compound? (Cycles on the heatmap ranging from red (min) to blue (max).) Which disease categories are associated with these proteins?

- Click on the “Disease” field for one of the selected proteins (for example, P31645).

How many proteins interact with this protein? What are the significant biological outcomes associated to it?

- Save the results by clicking on “Download results”.
- Return to the starting page and conduct a similar search for a compound called Fluoxetine.

How many compounds and proteins do you retrieve? One of the targeted proteins is P28223. Which are the disease categories associated with it?

- Click on the “Disease” field

How many proteins interact with Fluoxetine and what are the significant biological outcomes?

- Search in the literature for the therapeutically action of Citalopram and Fluoxetine. Is there agreement with your findings?
- Save the results by clicking on ”Download results”.

2) Search by Disease name

- Write the name Schizophrenia in “Search by Disease name“ and click on ”Submit“
How many proteins are associated with this disease? Is there any overlap with the list of proteins you found for the two compounds above? (Hint: Open the downloaded result files in e.g. excel and search for the proteins)

3) Search by Side effect

• Write the name amnesia in “Search by Side effect” and click on “Submit”

How many proteins are associated with the side effect? Is there any overlap with the list of proteins you found for the two compounds above?

Based on your findings, what could you conclude about Citalopram and Fluoxetine?