EMBL-EBI Tutorial

Louisa Bellis (April 2012)

Patent Searching in Chemical Databases

You will need access to **ChEMBL** (www.ebi.ac.uk/chembldb) and **ChEBI** (www.ebi.ac.uk/chebi)

Examples:

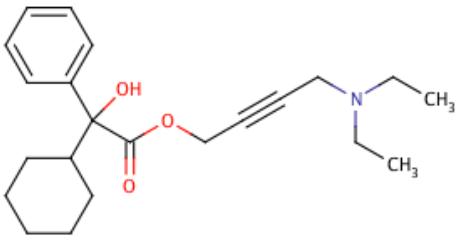
You have been asked to search for more information about the following compound(s). You can either use their systematic name or draw the structures.

Things to consider:

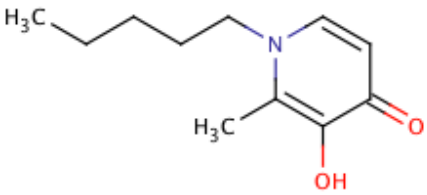
- What, if any, patents are available for these compounds?
- What literature/patents are available?
- Does ChemSpider have any patents?
-

HINT 1: ChEBI contains patent links, so their website is a good place to start.

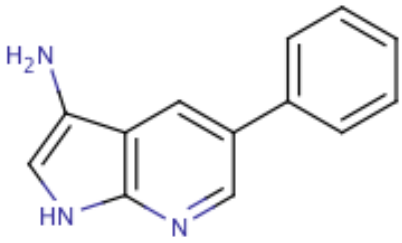
1. Oxybutynin: synonyms: Oxytrol, Ditropan



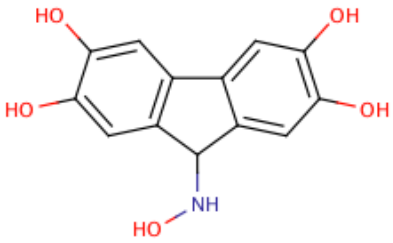
2. 3-Hydroxy-2-methyl-1-pentyl-1H-pyridin-4-one



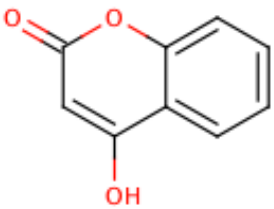
3. 5-Phenyl-1H-pyrrolo[2,3-b]pyridin-3-ylamine



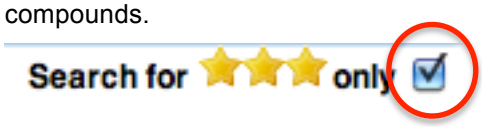
4. 9-Hydroxyamino-9H-fluorene-2,3,6,7-tetraol



5. 4-Hydroxycoumarin-based compounds, including dicoumarol



HINT 2: Untick the check box when running ChEBI searches to include the ChEMBL compounds.



Extra Questions

1. Search for **Dopamine** in using the Search Field at the top of the home page and clicking on "**Compounds**":

   a. How many compounds are brought back?
   b. What are some synonyms for Dopamine?
   c. How many **Functional Assays** are there for Dopamine?

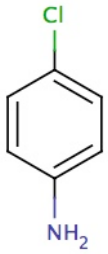
2. Using the "**Targets**" from the Search Field on the homepage:
   a. How many **targets** are brought back if you search for "**Breast Cancer**"?
   b. What is the **Uniprot Accession ID** for the target with the most bioactivity data?
   c. What are the different bioactivity types for this target (Hint: Look at the Bioactivity Summary pie chart in the Target Report Card)

3. Under the **Browse Targets** tab on the homepage, use the mini arrow to access the lower levels:
   a. How many targets are classed as Cytochrome P450?
   b. How many are classed as 7TM2?

4. Search for the target **Maltase-glucoamylase**.
   a. How many compounds and bioactivities are linked to this target?
   b. What are some of the synonyms for this target?
   c. Filter on the activity types of Ki and IC50 and download the bioactivities as an XLS spreadsheet

5. Search for the target **Cannabinoid CB2 Receptor**
   a. What are the Uniprot Accession IDs for these targets?
   b. Copy the FASTA file for **P34972** (www.uniprot.org) and run a Target Blast search. What are the % identity scores of CB1 and the CB2 receptors?

6. Search for the following compound (note: you will have to draw it in using the compound search tab)



   a. How many resulting compounds are there with a substructure search?
   b. How many resulting compounds are there with a >=90% similarity search?

7. Search for **GABA transporter 1**
   a. Filter the bioactivities on **IC50 < 1000nM** for all the targets and download as an XLS spreadsheet
   b. Copy the column list of **ChEMBL_IDs** for the compounds related to these targets and paste them into the Compound Search box
   c. Filter the bioactivities again on IC50 <1000nM and download as an XLS spreadsheet
   d. Are there any other targets, other than GABA 1, that are in the spreadsheet?

8. Go to **Browse Drugs**
   a. Using the magnifying glass icon in the lower left hand corner, search the known drugs for all of the compounds that have **Topical Delivery**.
   b. Download the data as an SDFile (benzene icon) and XLS spreadsheet (X icon).
   c. Copy the column of **Salt_CHEMBL_IDs** and use these to search the main ChEMBL database and download the associated bioactivities as a spreadsheet, after filtering on various activity endpoints of your choice.

9. Search for **liver microsomes** in Assays
   a. What assay has the highest number of bioactivities?
   b. How many compounds have been assayed against this target?
   c. What is the reference for this assay and who is the first author on this?