

## Searching the PDB for drugs that bind to Christmas Factor

**PDBe** provides a range of services for exploring the binding of small molecules such as drugs, co-factors, co-enzymes, and prosthetic groups to biomacromolecules.

As an introduction to some of the services available, this mini-tutorial uses PDBe tools to explore compounds in the PDB archive that are chemically similar to the **benzothiofene** derivatives bound to **Christmas Factor** in PDB entries **3lc3** and **3lc5**.

The search methods it shows should be of general use to you as you explore small molecules in PDB structures.

Small molecules are dealt with in the PDB archive as formally equivalent to macromolecular residues. Each one is given a unique **three-letter code** but these are often arbitrary and so you can't necessarily guess what the compound might be from the code.

To start searching small molecules using PDBChem, click on the link in the PDBe Services menu on the PDBe home page under **Services > Chemistry > PDB Compounds** (highlighted in red, **screenshot 1**).

### Screenshot 1. The PDB compounds link on the PDBe home page.

From here you get a search menu where you can search using five separate options, each of which has a separate input box (see **screenshot 2**).

The screenshot shows the PDBe home page with the following elements:

- Header:** PDBe PROTEIN DATA BANK EUROPE
- Search Bar:** Enter search term, accession number or PDBe service name [SEARCH]
- Navigation:** Home, Wizard, Education, Resources, Help, About us
- EMBL-EBI's Protein Data Bank in Europe (PDBe)** is the European resource for the collection, organisation and dissemination of data on biological macromolecular structures. [More...](#) [Contact us](#)
- Statistics:** As of 14 December 2011 the PDB contains **77878** entries ([latest entries](#), [latest compounds](#)) and EMDB contains **1207** entries ([latest](#))
- Quick access:**
  - Sequence search
  - PDBe feature
  - RSS feeds
- One-click access to PDB data:** Enter a PDB ID code and click a button below for more information about the PDB entry:
  - Entry summary
  - Download PDB file
  - Download other files
  - Quaternary structure
  - Similar structures
  - Motifs and sites
- Retrieve PDB entries using an external database identifier:** PubMed [Search]
- Find a random PDB entry ...**
  - or one solved by...**
    - X-ray
    - NMR
    - Cryo-EM
    - Hybrid methods
    - Any other method
  - or one containing...**
    - Protein
    - DNA
    - RNA
    - Protein & DNA
    - Protein & RNA
    - Sugar
    - Ligand
    - Intact virus
  - or one released in...**
    - the past week
    - 2010s
    - 2000s
    - 1990s
    - 1980s
    - 1970s
- Left Sidebar (PDBe Tools):**
  - Latest release:** PDB entries, PDB compounds, EMDB entries, PDB entry status
  - Deposit:** PDB (AutoDep), EMDB (EMDep)
  - Search:** Advanced PDB search, Advanced EMDB search, Sequence similarity, Structure similarity
  - Browse:** Enzymes, Folds, Protein families, Taxonomy, Compounds
  - Services:** Quaternary structure, Motifs and sites (Advanced analysis, Quick access, 2D site viewer), Structure comparison, Chemistry, **PDB compounds** (highlighted in red), Other compounds, Data validation

**Screenshot 2.**  
**The Chemical Components**  
**search page**

Currently searching over 13335 ligands.

**PDBeChem : Search By Molecule**

Code	=	⌵	?	<input type="text"/>	
Molecule Name	=	⌵	?	<input type="text"/>	
Formula	formula range	⌵	?	<input type="text"/>	<input type="button" value="edit"/>
Non-Stereo SMILES	(Has Sub-Structure)		?	<input type="text"/>	<input type="button" value="edit"/>
Fragments	(Fragment Expression)		?	<input type="text"/>	<input type="button" value="edit"/>
Combine Criteria With	<input checked="" type="radio"/> AND <input type="radio"/> OR ?				

The first option on the search page is useful if you actually know the **three letter code** of the small molecule you're interested in, and the second if you know the molecule name. In both cases, changing the pull-down menu to the left of the input box from '=' to 'like' means the search results will find hits **containing** the term you've input, rather than **matching it exactly**. So entering **MSE** for example takes you straight to the Chemical Component page for Selenomethionine.

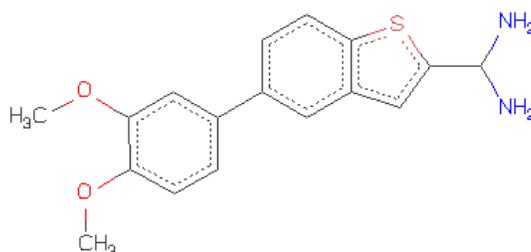
The third option allows you to input the molecules formula into the search box, so searching for '**C6 H12 O6**' will find all the hexose sugars. This search box defaults to **formula range** but you can also explicitly define a range. Inputting '**C6-7 H12 O6**', for example, will find all those molecules containing six or seven carbons, twelve hydrogens, six oxygens.

For our purpose in this tutorial, it is the next search option '**non-stereo SMILES**' which is most useful. In this search box you can input the SMILES string of the molecule, or molecule fragment of interest. SMILES stands for **Simplified Molecular-Input Line-Entry Specification**. For more information about SMILES, see this link: <http://www.daylight.com/smiles/>

More than likely, you won't know the SMILES string of your small molecule, so instead click the '**edit**' button to the right of the option '**non-stereo SMILES**' input box. This will open up another window with the **JME** molecular editor tool (**screenshot 3**). You'll find a link with tips on how to use the editor just below the sketch box.

**Screenshot 3.**  
**The Molecular Editor for**  
**Chemical Substructure**  
**Search**

PDB entry 3lc3 contains the inhibitor 1-[5-(3,4-dimethoxyphenyl)-1-benzothiophen-2-yl]methanediamine (shown **graphically** below).



Let's sketch just the double-ringed **benzothiophene** moiety in the editor.

Firstly, select the **benzene ring** icon and drag it into the editor (number **1** in **screenshot 4** below).

Next, select the 5 membered ring (number **2** in **screenshot 4**) and while hovering over a bond in the benzene ring, click the mouse to fuse this second ring to the benzene ring. Now select 'S' from the left-hand side menu (number **3** in **screenshot 4**) and click the mouse over the right place in the 5-membered ring to swap the carbon for a sulphur.

Finally, select the the double bond icon and add a double bond into the 5 membered ring to form benzothiophene.

**Screenshot 4.**  
**Building the benzothiophene rings**

Now hit '**submit**' to take you back to the search page. You will see that the **SMILES** string for what you've drawn has been automatically filled in to the input box.

There are other useful options in the editor tool. For example you can upload a molecule to alter in the editor, which might make it more easy if you want to search for a complex fragment. You can also decide whether you want to search for fragments as **part** of a ring or optionally to **disregard** bond order.

Once the SMILES string has been automatically generated on the **PDBChem** search page press '**Search**' to find all those small molecules in the **PDB chemical component** dictionary that contain this **benzothiophene** group.

You will see a results page similar to that below (**screenshot 5**). This show that there are over 50 molecules containing this group, listed with the **three letter code**, the molecule's name, and an image of the structure. Hovering over the image will give you a larger view.

**Screenshot 5.**  
**The PDBeChem search results.**

**PDBeChem : Search Results**  
Total number of results: 54

Code	Name	Formula	Structure	Superseded By
03T	1-benzothiazophene-2-sulfonamide	C8 H7 N O2 S2		
03S	3-CHLORO-6-FLUORO-N-[2-[4-[5-PROPAN-2-YL-1,3,4-THIAZOL-2-YL]SULFAMOYL]PHENYL]ETHYL]-1-BENZOTHIAZOPHENE-2-CARBOXAMIDE	C22 H20 Cl F N4 O3 S3		
03B	4-[3-[4-methyl-1-benzothiazophen-3-yl]methyl]-2-oxo-2,3-dihydro-1H-benzimidazol-1-yl]butanoic acid	C21 H20 N2 O3 S		
03U	1-[3-[(4-(5S)-3,3-dimethyl-1-oxo-2-oxa-7-azaspiro[4.5]dec-7-yl)pyridin-1-yl]carbonyl]-1-benzothiazophen-2-yl]-3-ethylurea	C27 H36 N4 O4 S		

From the search results, you can click the **three letter code** to take you to a descriptive page for each ligand (for example, **03T**, whose page is shown in **screenshot 6**) and from here you can see useful information about each molecule.

**Screenshot 6.**  
**Chemical Component Summary page for 03T**

**03T : Summary**

Code: 03T  
One letter code: X  
Molecule name: 1-benzothiazophene-2-sulfonamide

Program	Version	Name
ACDLabs	12.01	1-benzothiazophene-2-sulfonamide
OpenEye OEToolkits	1.7.2	1-benzothiazophene-2-sulfonamide

Formula: C8 H7 N O2 S2  
Formal charge: 0  
Molecular weight: 213.277 Da

Type	Program	Version	Descriptor
SMILES	ACDLabs	12.01	O=S(=O)(c1ccc2c(c1)ncn2)
SMILES	CACTVS	3.3.70	N[S](=O)(=O)c1ccc2c(c1)ncn2
SMILES	OpenEye OEToolkits	1.7.2	c1ccc2c(c1)ncn2S(=O)(=O)N
Canonical SMILES	CACTVS	3.3.70	N[S](=O)(=O)c1ccc2c(c1)ncn2
Canonical SMILES	OpenEye OEToolkits	1.7.2	c1ccc2c(c1)ncn2S(=O)(=O)N

IUPAC InChI: NC(=O)S1=NC2=CC=CC=C2S1  
IUPAC InChI key: UZMQSZBTPQHLAH-UHFFFAOYSA-N

**Supplementary Information**

Atom count: 20 (13 without Hydrogen)  
Polymer type: Bound ligand  
Type description: NON-POLYMER  
Type code: HETAIN  
Is modified: No  
Standard parent: Not Assigned  
Defined at: 2011-10-05  
Last modified at: 2011-10-14  
Status: Released  
Obsoleted: Not Assigned

The **Chemical Component Summary page** contains the names and synonyms of the molecule as well as several different chemical descriptors, formula and molecular weight. The links to the left allow you to explore the atom names and bonds in the molecule, and to see a list of PDB entries which contain the molecule.

The **'Download Links'** section is where you can download the coordinates of the molecule, and its dictionary should you want to use this molecule in your own structure determination.