

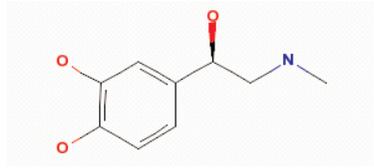


PubChem (PC) is a publicly available collection of over 10 million unique small molecules and their biological activities. PubChem is a component of the NIH Molecular Libraries Roadmap Initiative and is implemented in the NCBI Entrez system as a set of three databases:

PC Substance, PC Compound, and PC BioAssay
Access PubChem at pubchem.ncbi.nlm.nih.gov

PC Substance

Substance Summary:
Compound Displayed



Source: NIAID (122802)
Each Substance contains a link to the original depositor.

PC Substance — contains original data from depositors. Substances may or may not represent well-characterized chemicals.

PC Compound — contains standardized versions of only those Substance records that represent well-defined chemicals.

PC Compound

Compound Summary:



SID: 219191

CID: 5816

Related Substances:

- Same: 13 Links
- Same, Connectivity: 35 Links
- Same, Stereochemistry: 14 Links
- Same, Isotopes: 33 Links

Similar Substances: 613 Links

Structure Search

A Substance is assigned an SID and a link to the CID of the Compound that represents it (when one exists).

Each Compound is linked to the BioActivity data for all of the identical Substances that it represents.

Compounds may contain links to other data, including Protein Sequences, Toxicology information and protein Structures.

Each Compound is linked to the set of identical Substances that it represents and to pre-computed sets of structurally Related Compounds.

CID: 5816
A Compound is assigned a CID.

BioActivity: Summary

All: 28 Links
Active: 1 Link
Inactive: 25 Links
Inconclusive: 1 Link

Protein Sequences: 93 Links

NLM Toxicology: Link

Substances:

All: 46 Links
Same: 13 Links
Mixture: 33 Links

Related Compounds:

Same, Connectivity: 5 Links
Same, Stereochemistry: 2 Links
Same, Isotopes: 3 Links

Similar Compounds: 217 Links

Structure Search

PC BioAssay

PubChem BioAssay — contains links to active, inactive and inconclusive Compounds and Substances, along with links to the target Protein and Gene (where applicable).

AID: 410 *A BioAssay is assigned an AID and is linked to the original depositor.*

Name: p450-cyp1a2
Data Source: NCGC

Displays/downloads the entire data set *Displays/downloads a defined portion of the data set*

Test Results:

Links:

Compounds: All: 8342 Active: 4173 Inactive: 3669 Inconclusive: 713

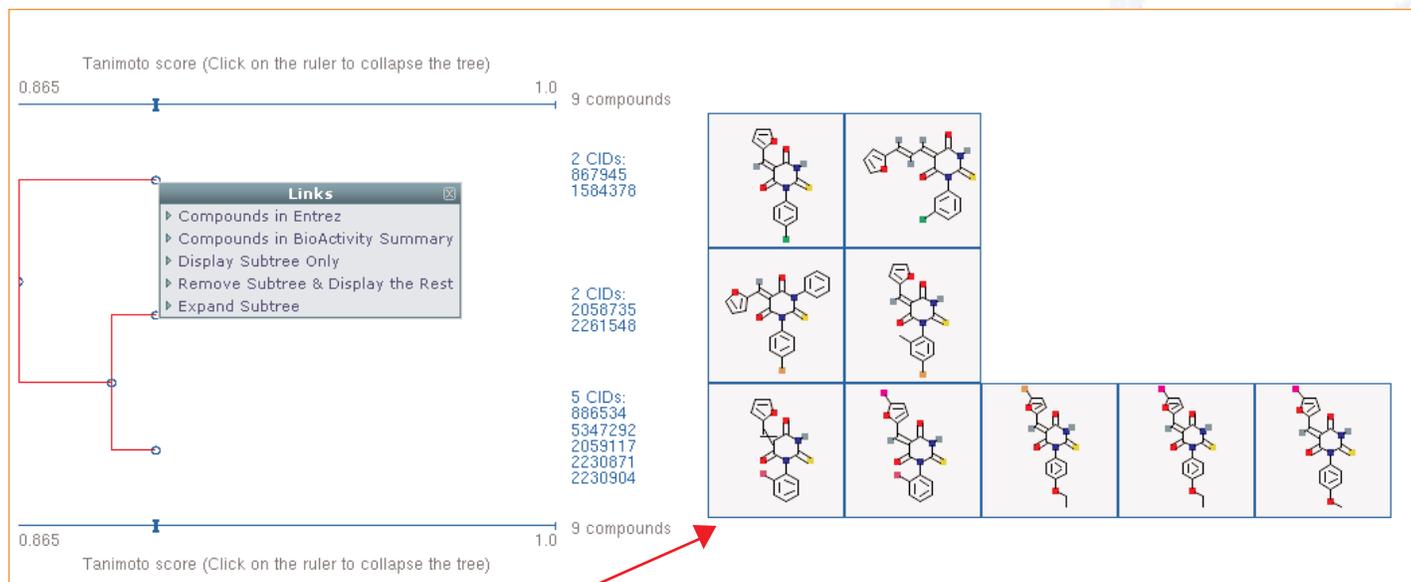
Substances: All: 9198 Active: 4446 Inactive: 4019 Inconclusive: 733

Protein: 1 Link

Gene: 1 Link

PubChem Structure Search and Clustering

Compounds may be retrieved based on identity, similarity (by Tanimoto score), substructure, superstructure, or molecular formula. The search results can be clustered on a tree or uploaded to Entrez for further analysis.



BioActivity summaries are linked from the summary pages of all tested Substances and Compounds.

- Structure Activity Analysis ?
- Structure Clustering ?
- Selected BioAssays to Entrez ?

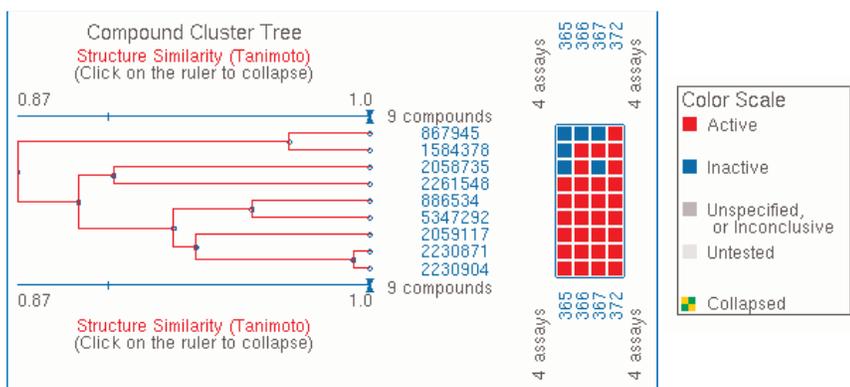
PubChem BioActivity Summary

Shows all BioAssays in which the input set of Substances/Compounds have been tested

#	<input type="checkbox"/>	AID	Active	Inactive	Total Tested	Data	Name
1	<input type="checkbox"/>	372	9		9	Data	HIV-1 RNase H Inhibition
2	<input type="checkbox"/>	366	8	1	9	Data	human RNase H Inhibition
3	<input type="checkbox"/>	367	7	2	9	Data	HIV-2 RNase H Inhibition
4	<input type="checkbox"/>	365	6	3	9	Data	E. coli RNase H Inhibition

PubChem Structure Activity Analysis

The PubChem SAR tool displays bioactivity data in a heatmap where each row represents a CID and each column an AID. The Compounds can be clustered either by structural or activity similarity, and the BioAssays by activity or protein target similarity.



PubChem Power User Gateway (PUG): <ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem.pug.pdf>

PubChem Help: pubchem.ncbi.nlm.nih.gov/help.html

