

Metabolomics Databases

Xiuxia Du, Stephen Barnes

Outline

- Comprehensive metabolomics databases
- Compound databases
- Spectral databases
- Metabolic pathway databases
- Drug databases
- Disease & physiology databases
- Raw data databases

Outline

- **Comprehensive metabolomics databases**
- Compound databases
- Spectral databases
- Metabolic pathway databases
- Drug databases
- Disease & physiology databases
- Raw data databases

Comprehensive databases

- HMDB

HMDB

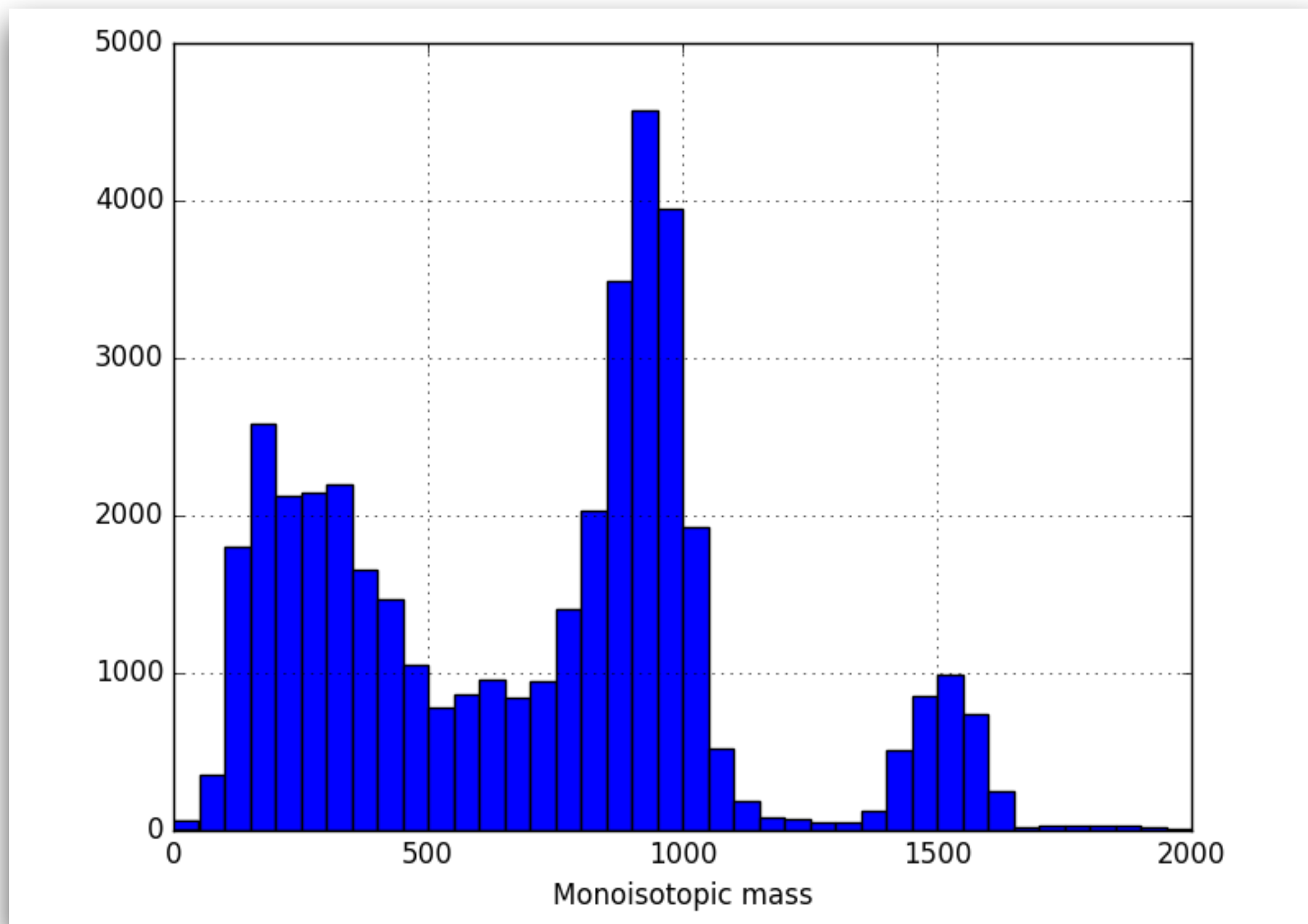
- Overview

Welcome to HMDB Version 3.6

The Human Metabolome Database (HMDB) is a freely available electronic database containing detailed information about small molecule metabolites found in the human body. It is intended to be used for applications in metabolomics, clinical chemistry, biomarker discovery and general education. The database is designed to contain or link three kinds of data: 1) chemical data, 2) clinical data, and 3) molecular biology/biochemistry data. The database contains 41,993 metabolite entries including both water-soluble and lipid soluble metabolites as well as metabolites that would be regarded as either abundant ($> 1 \mu\text{M}$) or relatively rare ($< 1 \text{ nM}$). Additionally, 5,701 protein sequences are linked to these metabolite entries. Each MetaboCard entry contains more than 110 data fields with 2/3 of the information being devoted to chemical/clinical data and the other 1/3 devoted to enzymatic or biochemical data. Many data fields are hyperlinked to other databases ([KEGG](#), [PubChem](#), [MetaCyc](#), [ChEBI](#), [PDB](#), [UniProt](#), and [GenBank](#)) and a variety of structure and pathway viewing applets. The HMDB database supports extensive text, sequence, chemical structure and relational query searches. Four additional databases, [DrugBank](#), [T3DB](#), [SMPDB](#) and [FooDB](#) are also part of the HMDB suite of databases. [DrugBank](#) contains equivalent information on ~1600 drug and drug metabolites, [T3DB](#) contains information on ~3600 common toxins and environmental pollutants, [SMPDB](#) contains pathway diagrams for ~700 human metabolic and disease pathways, while [FooDB](#) contains equivalent information on ~28,000 food components and food additives.

HMDB

- Breakdown by mass



- One metabocard

Showing metabocard for L-Proline (HMDB00162)

Identification Taxonomy Ontology Physical properties Spectra Biological properties Concentrations Links References XML

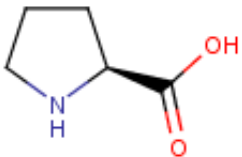
enzymes (29) transporters (1) [Show 30 proteins](#)

[Show Metabolites with Similar Structures](#)

Record Information

| | |
|-----------------------------|-------------------------|
| Version | 3.6 |
| Creation Date | 2005-11-16 15:48:42 UTC |
| Update Date | 2016-06-10 21:23:05 UTC |
| HMDB ID | HMDB00162 |
| Secondary Accession Numbers | None |

Metabolite Identification

| | |
|-------------|--|
| Common Name | L-Proline |
| Description | L-Proline is one of the twenty amino acids used in living organisms as the building blocks of proteins. Proline is sometimes called an imino acid, although the IUPAC definition of an imine requires a carbon-nitrogen double bond. Proline is a non-essential amino acid that is synthesized from glutamic acid. It is an essential component of collagen and is important for proper functioning of joints and tendons. |
| Structure |  <chem>C1CCNC1C(=O)O</chem> |

HMDB

- One metabocard

| | |
|-------------------------------|---|
| Chemical Formula | C ₅ H ₉ NO ₂ |
| Average Molecular Weight | 115.1305 |
| Monoisotopic Molecular Weight | 115.063328537 |
| IUPAC Name | (2S)-pyrrolidine-2-carboxylic acid |
| Traditional Name | L-proline |
| CAS Registry Number | 147-85-3 |
| SMILES | <chem>OC(=O)[C@@H]1CCCN1</chem> |
| InChI Identifier | InChI=1S/C5H9NO2/c7-5(8)4-2-1-3-6-4/h4,6H,1-3H2,(H,7,8)/t4-/m0/s1 |
| InChI Key | InChIKey=ONIBWKKTOPOVIA-BYPYZUCNSA-N |

HMDB

- Searches

The screenshot shows the HMDB website interface. At the top, there is a navigation bar with the HMDB logo, a search bar containing the text 'metabolites', and a 'Search' button. Below the navigation bar, a dropdown menu is open under the 'Search' tab, listing various search methods: ChemQuery Structure Search, Molecular Weight Search, Text Query, Sequence Search, Advanced Search, MS Search, MS/MS Search, GC/MS Search, 1D NMR Search, and 2D NMR Search. To the right of the dropdown, there is a banner for 'The Metabolomics Innovation Centre' with the text 'Specializing in ready to use metabolomics kits.' Below the dropdown, there is a section titled 'Searching HMDB' with a search input field. Below this, there is a text box explaining advanced search capabilities: 'HMDB supports advanced search based on the Lucene query language. HMDB text search supports boolean logic (AND, OR, NOT operations). To match a string exactly, place quotes around your search term (for example "acetic acid" will only match the acetic followed by acid, it will not match acetic or acid alone). You can also search using "wild cards" by inserting a "*" in your search term. For example, searching for "acet*" will match all words starting with "acet". In addition, text search supports parenthetical groupings, and prepended +plus and -minus operators.' Below this text is a table with two columns: 'Example' and 'Description'. The first row shows the example query 'methylhistidine AND poultry' and its description: 'Find all entries containing both methylhistidine and poultry in the metabocard'.

www.hmdb.ca/textquery

HMDB

Browse ▾ Search ▾ Downloads About ▾ Contact Us

Search metabolites Search

The Metabolomics Innovation Centre | Specializing in ready to use metabolomics kits.

ChemQuery Structure Search
Molecular Weight Search
Text Query
Sequence Search
Advanced Search
MS Search
MS/MS Search
GC/MS Search
1D NMR Search
2D NMR Search

Searching HMDB

Search

HMDB supports advanced search based on the [Lucene query language](#). HMDB text search supports boolean logic (AND, OR, NOT operations). To match a string exactly, place quotes around your search term (for example "acetic acid" will only match the acetic followed by acid, it will not match acetic or acid alone). You can also search using "wild cards" by inserting a "*" in your search term. For example, searching for "acet*" will match all words starting with "acet". In addition, text search supports parenthetical groupings, and prepended +plus and -minus operators.

| Example | Description |
|--|---|
| <code>methylhistidine AND poultry</code> | Find all entries containing both methylhistidine and poultry in the metabocard |

HMDB

- Downloads

Protein/Gene Sequences (in FASTA Format)

| Data Set | Released on | Protein Sequences | Gene Sequences |
|-------------------------------------|-------------|--------------------------|--------------------------|
| All Metabolite Metabolizing Enzymes | 2016-07-10 | Download | Download |

Structures (in SDF Format)

| Data Set | Released on | SDF File |
|-----------------------|-------------|--------------------------|
| Metabolite Structures | 2016-07-10 | Download |

Metabolite and Protein Data (in XML format)



























| Data Set | Released on | XML File |
|-----------------|-------------|--------------------------|
| All Metabolites | 2016-07-10 | Download |
| All Proteins | 2016-07-09 | Download |

Spectra

| Data Set | Download Link |
|-------------------------------------|--------------------------|
| Mass Spectra Image Files | Download |
| GC/MS Peak Lists | Download |
| NMR Spectra FIDS Files | Download |
| NMR Spectra Peaklist Files | Download |
| Spectra information (in XML format) | Download |

HMDB

- Download all metabolites

| Name | ^ | Date Modified | Size | Kind |
|--|---|------------------------|----------|------|
|  hmdb_metabolites.xml | | Jul 10, 2016, 2:58 AM | 664.4 MB | XML |
|  HMDB00001.xml | | Apr 13, 2014, 12:50 AM | 60 KB | XML |
|  HMDB00002.xml | | Apr 13, 2014, 12:50 AM | 22 KB | XML |
|  HMDB00005.xml | | Apr 13, 2014, 12:50 AM | 24 KB | XML |
|  HMDB00008.xml | | Apr 13, 2014, 12:50 AM | 22 KB | XML |
|  HMDB00010.xml | | Apr 13, 2014, 12:50 AM | 19 KB | XML |
|  HMDB00011.xml | | Apr 13, 2014, 12:50 AM | 37 KB | XML |
|  HMDB00012.xml | | Apr 13, 2014, 12:50 AM | 21 KB | XML |
|  HMDB00014.xml | | Apr 13, 2014, 12:50 AM | 20 KB | XML |
|  HMDB00015.xml | | Apr 13, 2014, 12:50 AM | 19 KB | XML |
|  HMDB00016.xml | | Apr 13, 2014, 12:50 AM | 21 KB | XML |
|  HMDB00017.xml | | Apr 13, 2014, 12:50 AM | 21 KB | XML |
|  HMDB00019.xml | | Apr 13, 2014, 12:50 AM | 27 KB | XML |
|  HMDB00020.xml | | Apr 13, 2014, 12:50 AM | 55 KB | XML |
|  HMDB00021.xml | | Apr 13, 2014, 12:50 AM | 13 KB | XML |
|  HMDB00022.xml | | Apr 13, 2014, 12:50 AM | 23 KB | XML |
|  HMDB00023.xml | | Apr 13, 2014, 12:50 AM | 29 KB | XML |
|  HMDB00024.xml | | Apr 13, 2014, 12:50 AM | 36 KB | XML |
|  HMDB00026.xml | | Apr 13, 2014, 12:50 AM | 17 KB | XML |
|  HMDB00027.xml | | Apr 13, 2014, 12:50 AM | 27 KB | XML |
|  HMDB00030.xml | | Apr 13, 2014, 12:50 AM | 38 KB | XML |
|  HMDB00031.xml | | Apr 13, 2014, 12:50 AM | 33 KB | XML |
|  HMDB00032.xml | | Apr 13, 2014, 12:50 AM | 20 KB | XML |
|  HMDB00033.xml | | Apr 13, 2014, 12:50 AM | 35 KB | XML |
|  HMDB00034.xml | | Apr 13, 2014, 12:50 AM | 30 KB | XML |
|  HMDB00036.xml | | Apr 13, 2014, 12:50 AM | 29 KB | XML |

Outline

- Comprehensive metabolomics databases
- **Compound databases**
- Spectral databases
- Metabolic pathway databases
- Drug databases
- Disease & physiology databases
- Raw data databases

Compound databases

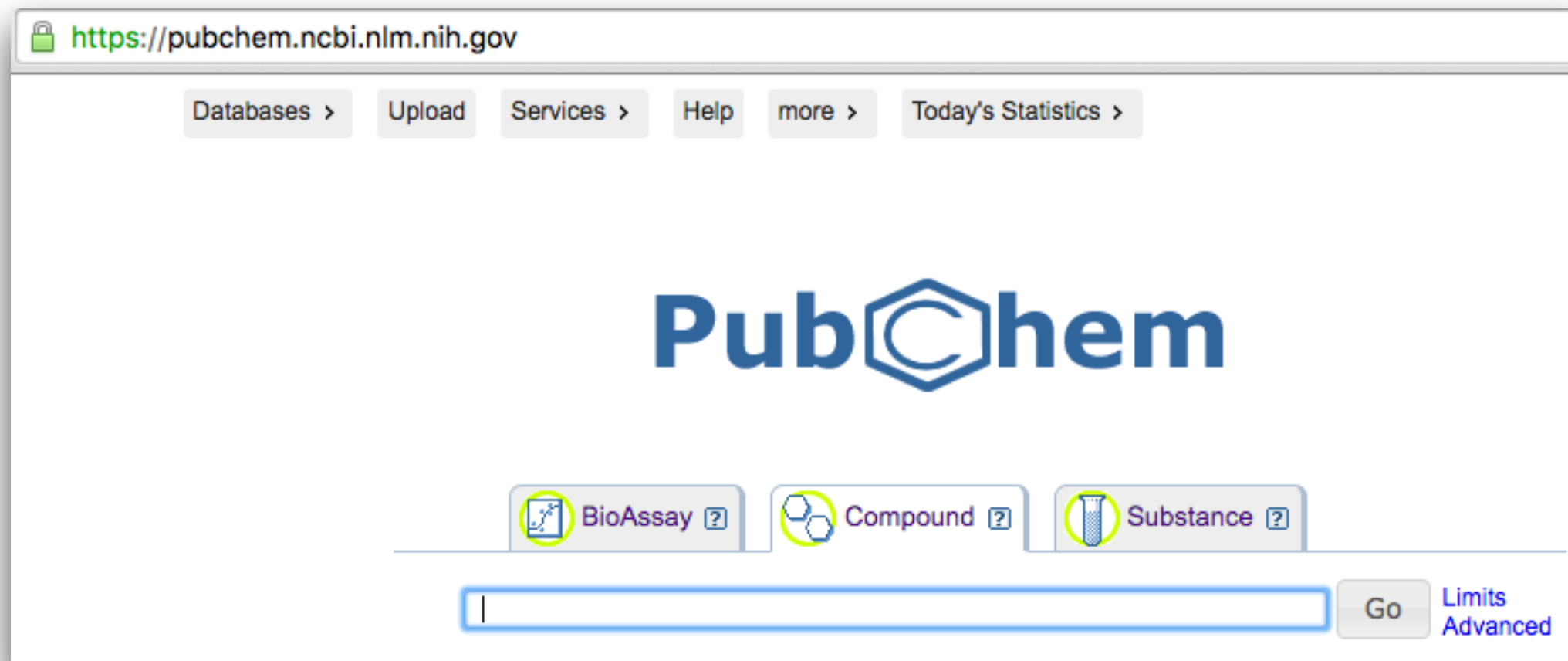
- PubChem
- ChemSpider
- ChEBI
- KEGG Glycan
- IIMDB

Compound databases

- PubChem
- ChemSpider
- ChEBI
- KEGG Glycan
- IIMDB

PubChem

- Website



PubChem

- Statistics (July 12, 2016)

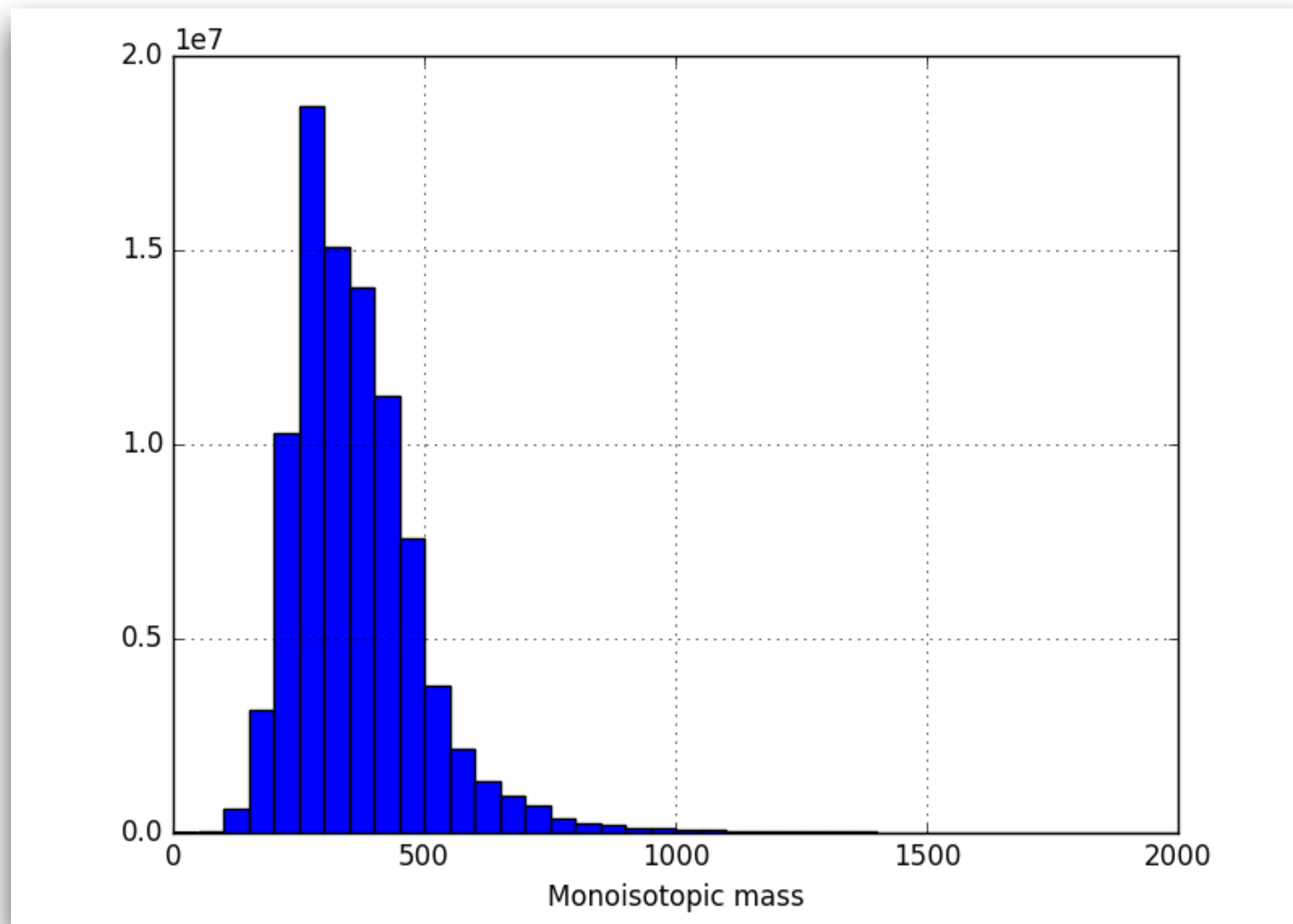
The screenshot shows the PubChem website interface. The address bar displays <https://pubchem.ncbi.nlm.nih.gov/#>. The navigation menu includes 'Databases >', 'Upload', 'Services >', 'Help', 'more >', and 'Today's Statistics >'. A red arrow points to the 'Today's Statistics >' dropdown menu, which is open and displays the following statistics:

| | |
|--------------------|-------------|
| Compounds: | 91,405,981 |
| Substances: | 221,797,218 |
| BioAssays: | 1,218,656 |
| Tested Compounds: | 2,261,888 |
| Tested Substances: | 3,522,315 |
| RNAi BioAssays: | 91 |
| BioActivities: | 230,495,192 |
| Protein Targets: | 10,180 |
| Gene Targets: | 19,778 |

Below the statistics, there are icons for 'BioAssay', 'Compounds', and 'Substances'. At the bottom, there is a search bar with a 'Go' button.

PubChem

- Breakdown by mass

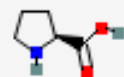


PubChem

- Information on one compound

L-proline

► Cite this Record



| | |
|---------------------------|--|
| PubChem CID: | 145742 |
| Chemical Names: | L-proline; 147-85-3; L-(-)-Proline; Proline; (S)-Pyrrolidine-2-carboxylic acid; (-)-(S)-Proline; More... |
| Molecular Formula: | $C_5H_9NO_2$ |
| Molecular Weight: | 115.13046 g/mol |
| InChI Key: | ONIBWKKTOPOVIA-BYPYZUCNSA-N |
| UNII: | 9DLQ4CIU6V |
| Modify Date: | 2016-07-09 |
| Create Date: | 2004-09-16 |

L-proline is a non-essential amino acid that is synthesized from GLUTAMIC ACID. It is an essential component of COLLAGEN and is important for proper functioning of joints and tendons.

PubChem

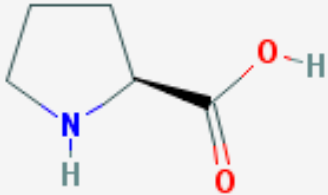
- Information on one compound

Contents

- 1 2D Structure
- 2 3D Conformer
- 3 Biologic Description
- 4 Names and Identifiers
- 5 Chemical and Physical Properties
- 6 Related Records
- 7 Chemical Vendors
- 8 Drug and Medication Information
- 9 Food Additives and Ingredients
- 10 Pharmacology and Biochemistry
- 11 Use and Manufacturing
- 12 Identification
- 13 Safety and Hazards
- 14 Toxicity
- 15 Literature
- 16 Patents
- 17 Biomolecular Interactions and Pathways
- 18 Biological Test Results
- 19 Classification
- 20 Information Sources

1 2D Structure

Search Download Get Image



Magnify

from PubChem

2 3D Conformer

Search Download Get Image

PubChem

- Search and FTP

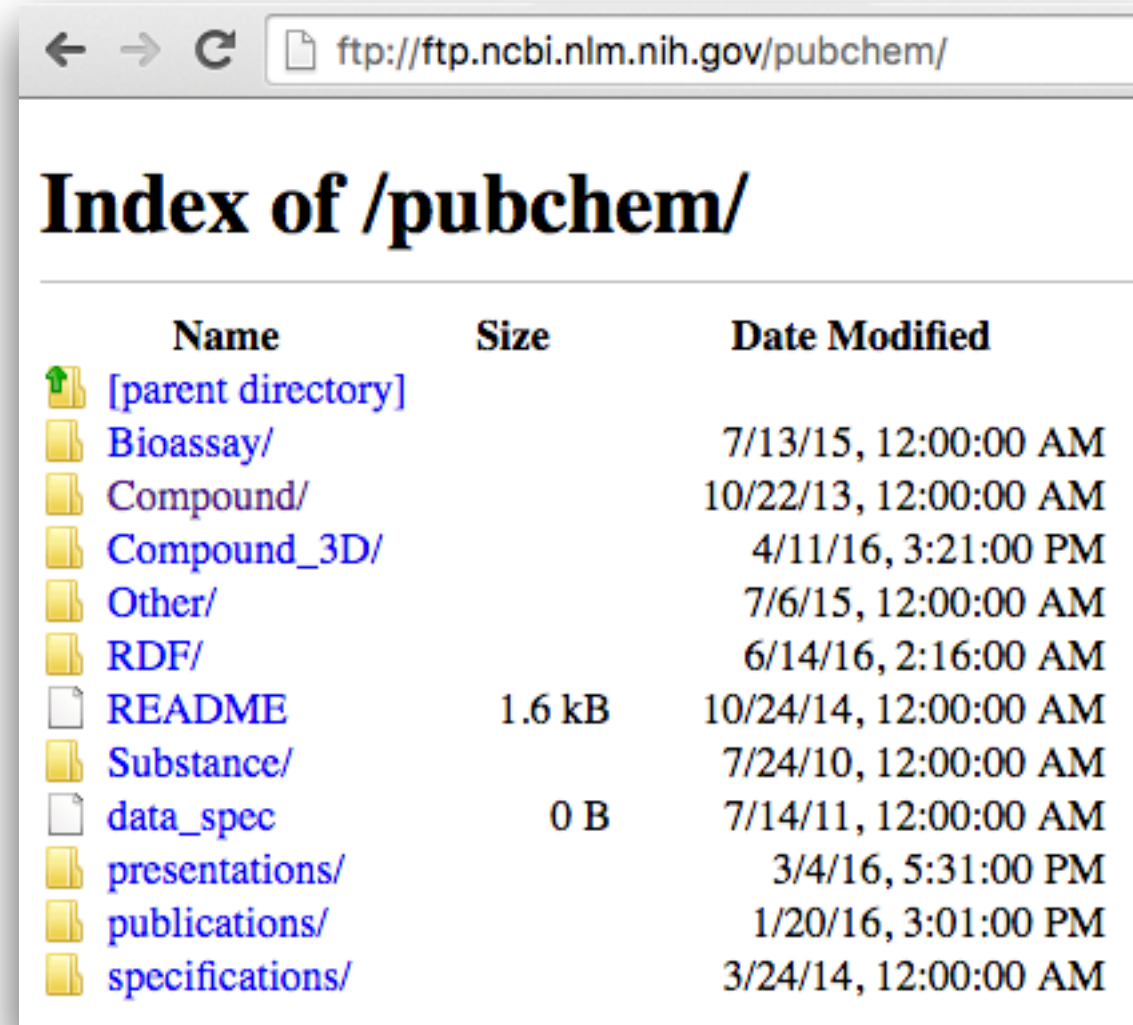
The screenshot displays the PubChem website's search and utility interface. At the top center is the PubChem logo. Below it are three search category buttons: "BioAssay" (with a bioassay icon), "Compound" (with a molecular structure icon), and "Substance" (with a test tube icon). A search input field is positioned below these buttons, followed by a "Go" button and links for "Limits" and "Advanced".

To the right of the search area is a vertical sidebar of utility buttons: "BioAssay Tools", "Structure Search", "3D Conformer Tools", "Structure Clustering", "Classification", "Upload", "Download", and "PubChem FTP". Each button includes a small icon representing its function.

Below the search input field, there is a link to "Try the PubChem Search Beta". A green notification box contains a "New" tag and text stating: "A new article about the PubChem Compound and Substance databases is available. Read more...". At the bottom right of this notification box, there is a "more ..." link and an RSS feed icon.













PubChem

- FTP



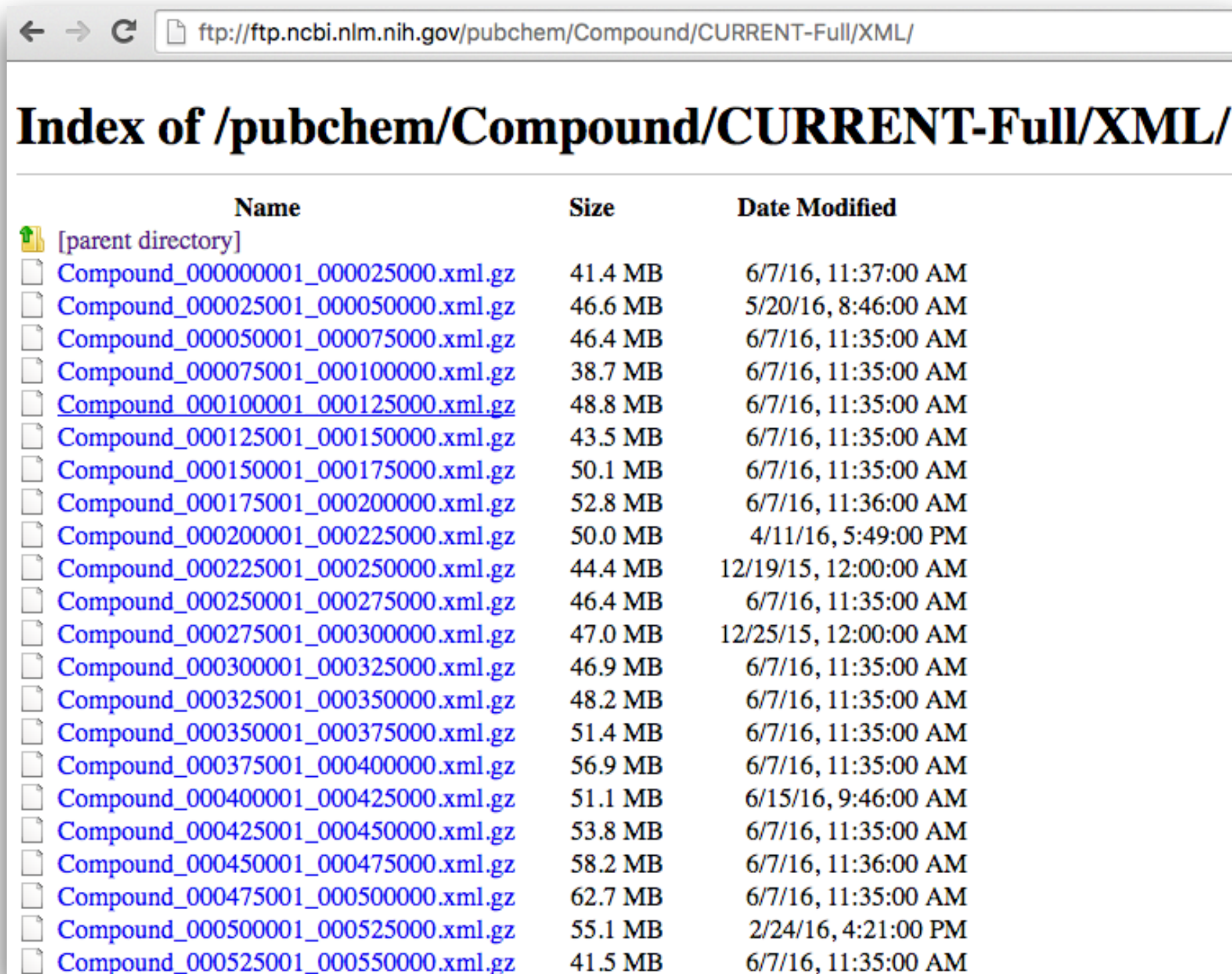
← → ↻ ftp://ftp.ncbi.nlm.nih.gov/pubchem/

Index of /pubchem/


| Name | Size | Date Modified |
|--|--------|-----------------------|
|  [parent directory] | | |
|  Bioassay/ | | 7/13/15, 12:00:00 AM |
|  Compound/ | | 10/22/13, 12:00:00 AM |
|  Compound_3D/ | | 4/11/16, 3:21:00 PM |
|  Other/ | | 7/6/15, 12:00:00 AM |
|  RDF/ | | 6/14/16, 2:16:00 AM |
|  README | 1.6 kB | 10/24/14, 12:00:00 AM |
|  Substance/ | | 7/24/10, 12:00:00 AM |
|  data_spec | 0 B | 7/14/11, 12:00:00 AM |
|  presentations/ | | 3/4/16, 5:31:00 PM |
|  publications/ | | 1/20/16, 3:01:00 PM |
|  specifications/ | | 3/24/14, 12:00:00 AM |

PubChem

- Compound XML files

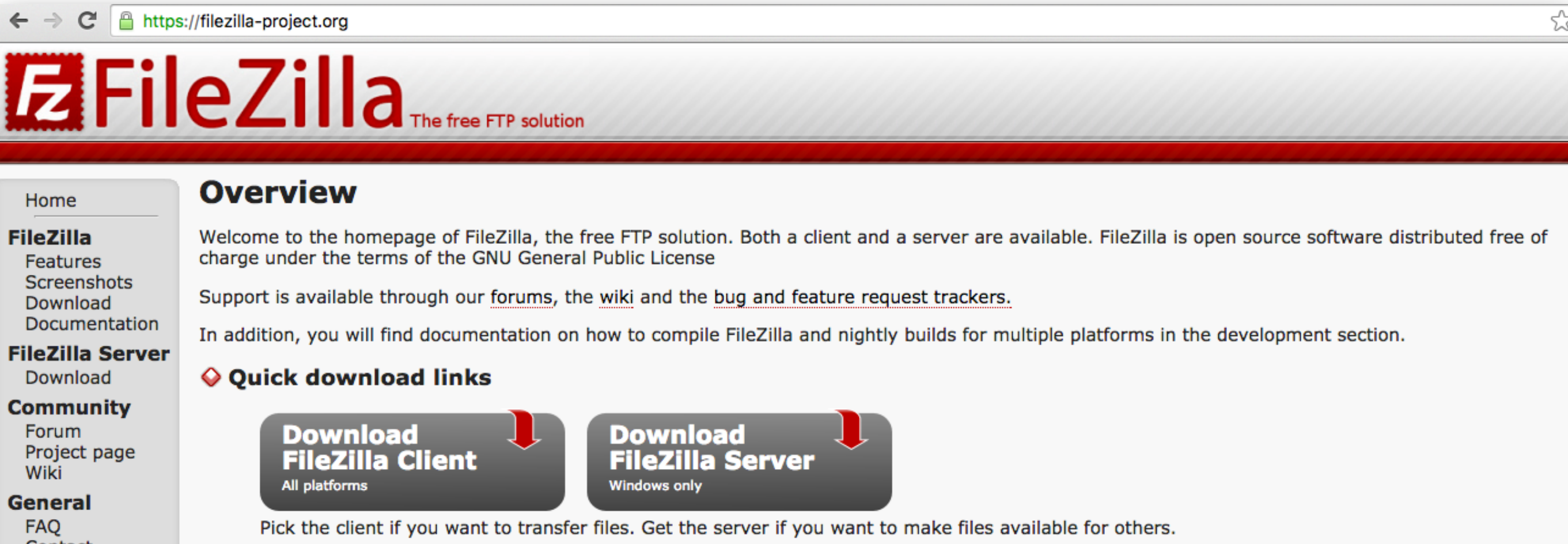


The screenshot shows an FTP directory listing for the path `ftp://ftp.ncbi.nlm.nih.gov/pubchem/Compound/CURRENT-Full/XML/`. The listing is titled "Index of /pubchem/Compound/CURRENT-Full/XML/" and contains a table with three columns: Name, Size, and Date Modified. The files listed are compressed XML files with names following the pattern `Compound_XXXXXX001_YYYYY0000.xml.gz`, where `XXXXXX` represents a 6-digit compound ID and `YYYYY` represents a 5-digit file ID. The sizes range from approximately 38.7 MB to 62.7 MB, and the modification dates range from 2/24/16 to 6/7/16.

| Name | Size | Date Modified |
|--|---------|-----------------------|
|  [parent directory] | | |
| Compound_00000001_000025000.xml.gz | 41.4 MB | 6/7/16, 11:37:00 AM |
| Compound_000025001_000050000.xml.gz | 46.6 MB | 5/20/16, 8:46:00 AM |
| Compound_000050001_000075000.xml.gz | 46.4 MB | 6/7/16, 11:35:00 AM |
| Compound_000075001_000100000.xml.gz | 38.7 MB | 6/7/16, 11:35:00 AM |
| Compound_000100001_000125000.xml.gz | 48.8 MB | 6/7/16, 11:35:00 AM |
| Compound_000125001_000150000.xml.gz | 43.5 MB | 6/7/16, 11:35:00 AM |
| Compound_000150001_000175000.xml.gz | 50.1 MB | 6/7/16, 11:35:00 AM |
| Compound_000175001_000200000.xml.gz | 52.8 MB | 6/7/16, 11:36:00 AM |
| Compound_000200001_000225000.xml.gz | 50.0 MB | 4/11/16, 5:49:00 PM |
| Compound_000225001_000250000.xml.gz | 44.4 MB | 12/19/15, 12:00:00 AM |
| Compound_000250001_000275000.xml.gz | 46.4 MB | 6/7/16, 11:35:00 AM |
| Compound_000275001_000300000.xml.gz | 47.0 MB | 12/25/15, 12:00:00 AM |
| Compound_000300001_000325000.xml.gz | 46.9 MB | 6/7/16, 11:35:00 AM |
| Compound_000325001_000350000.xml.gz | 48.2 MB | 6/7/16, 11:35:00 AM |
| Compound_000350001_000375000.xml.gz | 51.4 MB | 6/7/16, 11:35:00 AM |
| Compound_000375001_000400000.xml.gz | 56.9 MB | 6/7/16, 11:35:00 AM |
| Compound_000400001_000425000.xml.gz | 51.1 MB | 6/15/16, 9:46:00 AM |
| Compound_000425001_000450000.xml.gz | 53.8 MB | 6/7/16, 11:35:00 AM |
| Compound_000450001_000475000.xml.gz | 58.2 MB | 6/7/16, 11:36:00 AM |
| Compound_000475001_000500000.xml.gz | 62.7 MB | 6/7/16, 11:35:00 AM |
| Compound_000500001_000525000.xml.gz | 55.1 MB | 2/24/16, 4:21:00 PM |
| Compound_000525001_000550000.xml.gz | 41.5 MB | 6/7/16, 11:35:00 AM |

PubChem

- Download compound XML files

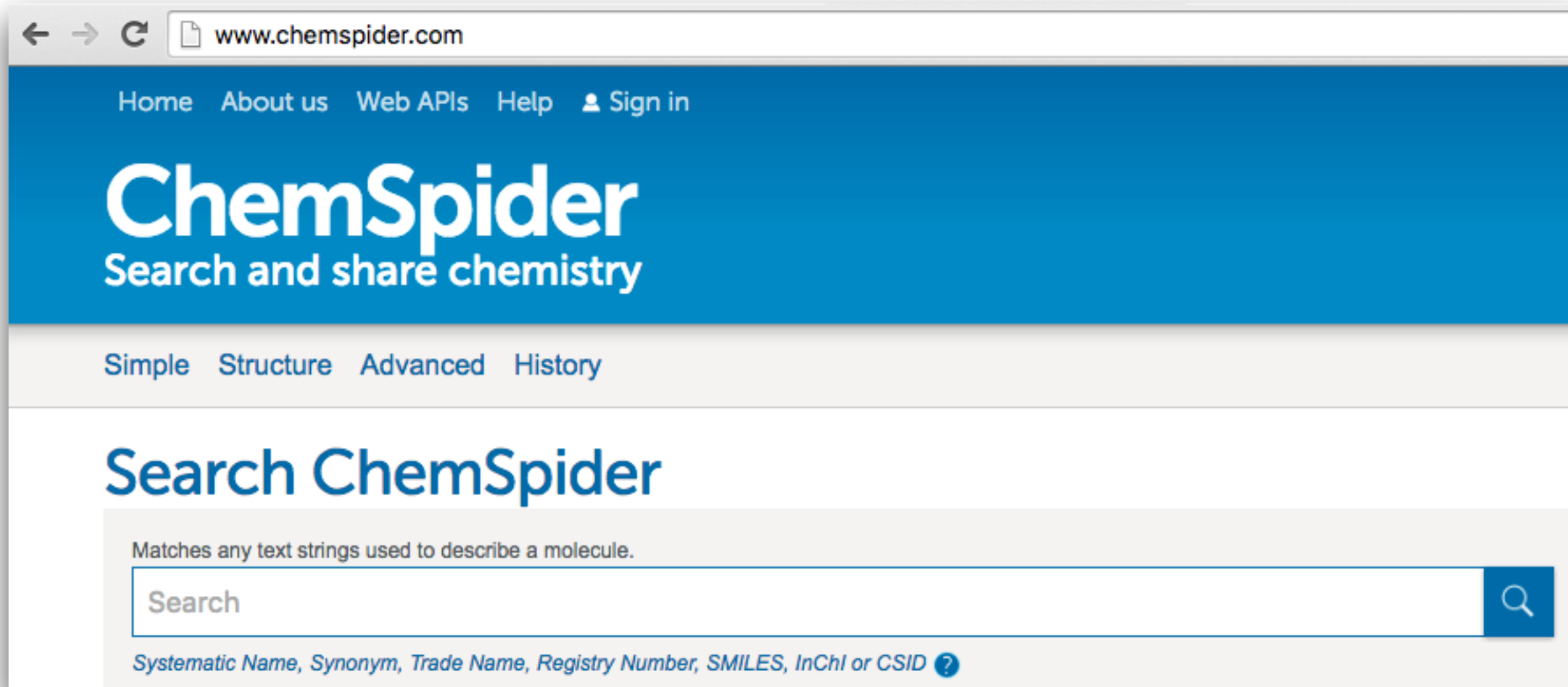


The screenshot shows the FileZilla website homepage. The browser address bar displays <https://filezilla-project.org>. The main header features the FileZilla logo and the tagline "The free FTP solution". A left sidebar contains navigation links for Home, FileZilla (Features, Screenshots, Download, Documentation), FileZilla Server (Download), Community (Forum, Project page, Wiki), and General (FAQ, Contact). The main content area is titled "Overview" and includes a welcome message, support information, and a "Quick download links" section with two buttons: "Download FileZilla Client" (All platforms) and "Download FileZilla Server" (Windows only). A note at the bottom of the download section reads: "Pick the client if you want to transfer files. Get the server if you want to make files available for others."

- Total number of XML files: 4,849
- Number of compounds in each XML file: 25,000

ChemSpider

- Website



ChemSpider

- Information

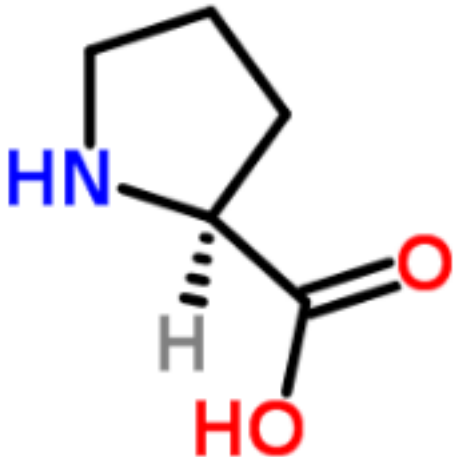
| What is ChemSpider? | Search by chemical names | Search by chemical structure | Find important data |
|---|--|---|--|
| <p><i>ChemSpider</i> is a free chemical structure database providing fast text and structure search access to over 55 million structures from hundreds of data sources.</p> | <ul style="list-style-type: none">• Systematic names• Synonyms• Trade names• Database identifiers | <ul style="list-style-type: none">• Create structure-based queries• Draw structures in the web page• Use structure files from your computer | <ul style="list-style-type: none">• Literature references• Physical properties• Interactive spectra• Chemical suppliers |

57 Million
chemical structures

517
Data sources

ChemSpider


- One compound






The image shows the chemical structure of L-Proline, a five-membered ring containing one nitrogen atom and one carboxylic acid group. The nitrogen atom is labeled 'HN' in blue. The carboxylic acid group is shown with a red 'O' and a red 'HO' group. A hydrogen atom 'H' is shown with a dashed bond, indicating its stereochemistry.

L-Proline

| | |
|-------------------|---|
| Molecular Formula | C ₅ H ₉ NO ₂ |
| Average mass | 115.131 Da |
| Monoisotopic mass | 115.063332 Da |
| ChemSpider ID | 128566 |

 - 1 of 1 defined stereocentres

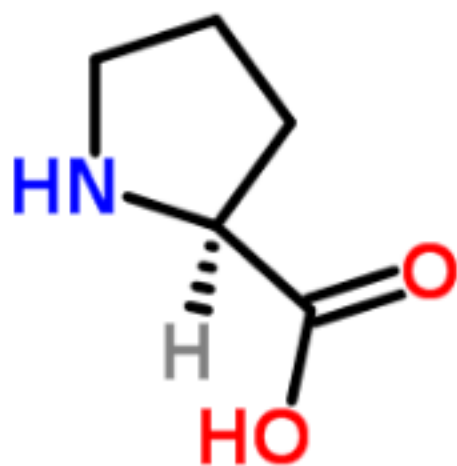
  

▼ **More details:**

This record has not been tagged.


Names and identifiers | [Properties](#) | [Searches](#) | [Spectra](#) | [Vendors](#) | [Articles](#) | [More](#) ▼

Names and Synonyms | Database ID(s)



L-Proline

| | |
|-------------------|---|
| Molecular Formula | C ₅ H ₉ NO ₂ |
| Average mass | 115.131 Da |
| Monoisotopic mass | 115.063332 Da |
| ChemSpider ID | 128566 |

 - 1 of 1 defined stereocentres



3D



More details:

This record has not been tagged.

Names and identifiers

Properties

Searches

Spectra

Vendors

Articles

More ▾

Search ChemSpider:

 [Compounds with the same molecular formula](#)

 [Compounds with the same skeleton](#)

 [Use this molecule in a structure search](#)

Search Google:

 [Search Google Scholar \(by synonym\)](#)

 [Search Google for exact structure](#)

 [Search Google for structures with same skeleton](#)

ChemSpider


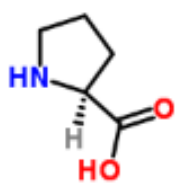

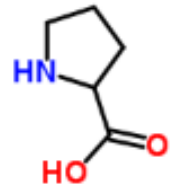
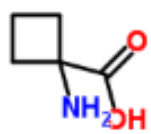

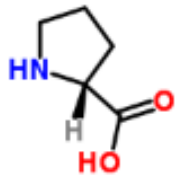
- Compounds with the same molecular formula

Found 501 results

Search term: MF = 'C_{5}H_{9}NO_{2}'

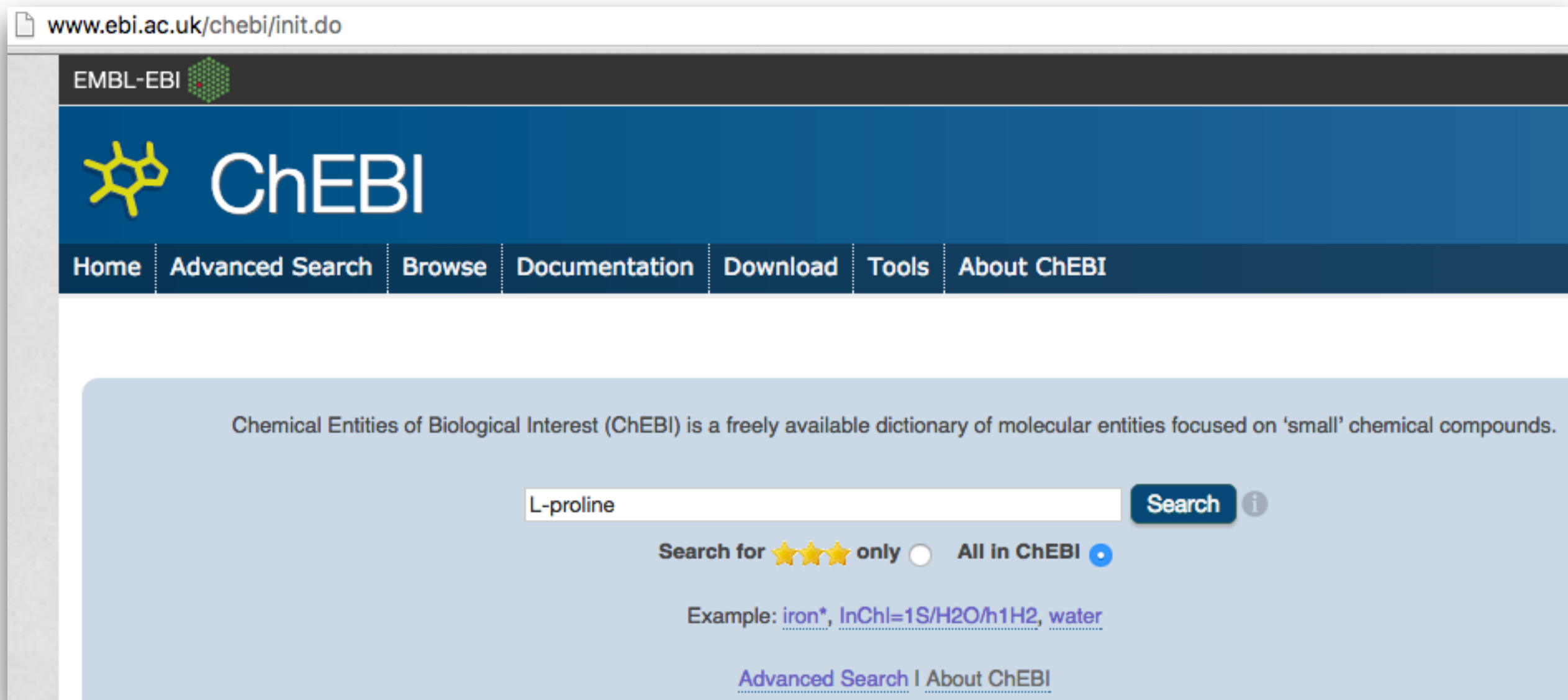


1 2 3 4 5

| ID | Structure | Molecular Formula | Molecular Weight | # of Data Sources | # of References | # of PubMed | # of RSC |
|---|---|---|------------------|-------------------|-----------------|-------------|----------|
| 128566  - 1/1 defined |  | C ₅ H ₉ NO ₂ | 115.1305 | 136 | 11954 | 28480 | 3548 |
| 594  - 0/1 defined |  | C ₅ H ₉ NO ₂ | 115.1305 | 131 | 737 | 28034 | 3319 |
| 80908 |  | C ₅ H ₉ NO ₂ | 115.1305 | 110 | 185 | 8 | 5 |
| 8640  - 1/1 defined |  | C ₅ H ₉ NO ₂ | 115.1305 | 103 | 293 | 701 | 268 |

ChEBI

- Chemical Entities of Biological Interest



The screenshot shows the ChEBI website homepage. At the top, the URL www.ebi.ac.uk/chebi/init.do is visible in the browser's address bar. Below the address bar is a dark blue header with the EMBL-EBI logo and the ChEBI logo (a yellow molecular structure) and the text "ChEBI". A navigation menu below the header includes links for Home, Advanced Search, Browse, Documentation, Download, Tools, and About ChEBI. The main content area has a light blue background and contains the following text: "Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds." Below this text is a search bar containing the text "L-proline" and a "Search" button with an information icon. Under the search bar, there are two radio button options: "Search for ★★★ only" (which is unselected) and "All in ChEBI" (which is selected). Below the radio buttons, there is an example search string: "Example: [iron*](#), [InChI=1S/H2O/h1H2](#), [water](#)". At the bottom of the search area, there are two links: [Advanced Search](#) and [About ChEBI](#).

ChEBI

- About

[ChEBI](#) > About ChEBI

1. Introduction

Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds. The term 'molecular entity' refers to any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer, etc., identifiable as a separately distinguishable entity. The molecular entities in question are either products of nature or synthetic products used to intervene in the processes of living organisms.

ChEBI incorporates an ontological classification, whereby the relationships between molecular entities or classes of entities and their parents and/or children are specified.

ChEBI uses nomenclature, symbolism and terminology endorsed by the following international scientific bodies:

- [International Union of Pure and Applied Chemistry \(IUPAC\)](#)
- Nomenclature Committee of the [International Union of Biochemistry and Molecular Biology \(NC-IUBMB\)](#)

Molecules directly encoded by the genome (e.g. nucleic acids, proteins and peptides derived from proteins by cleavage) are *not* as a rule included in ChEBI.

ChEBI

- One compound

CHEBI:17203 - L-proline

Main

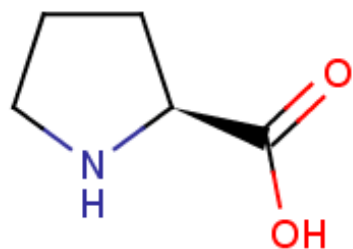
ChEBI Ontology

Automatic Xrefs

Reactions

Pathways

Models



ChEBI Name **L-proline**

ChEBI ID **CHEBI:17203**

ChEBI ASCII Name L-proline

Definition

Pyrrolidine in which the *pro-S* hydrogen at position 2 is substituted by a carboxylic acid group. L-Proline is the only one of the twenty DNA-encoded amino acids which has a secondary amino group α to the carboxyl group. It is an essential component of collagen and is important for proper functioning of joints and tendons. It also helps maintain and strengthen heart muscles.

Stars

☆☆☆ This entity has been manually annotated by the ChEBI Team.

Secondary ChEBI IDs

CHEBI:45159, CHEBI:45100, CHEBI:45040, CHEBI:42067, CHEBI:184637, CHEBI:6286, CHEBI:13154, CHEBI:21373

Supplier Information



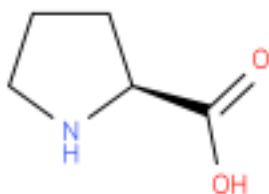
[eMolecules:524642](#), [ZINC00895360](#)

 [Download Molfile](#)

- [Find compounds which contain this structure](#)
- [Find compounds which resemble this structure](#)
- [Take structure to the Advanced Search](#)

Search Results for All in ChEBI

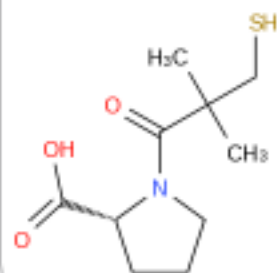
substructure



Edit Search

314 entries found, displaying 1 to 15.

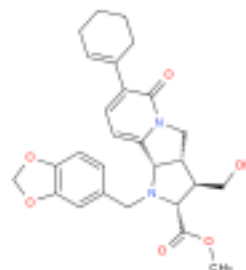
[\(2R\)-1-\(3-mercapto-2,2-dimethyl-1-oxopropyl\)-2-pyrrolidinecarboxylic acid](#)



CHEBI:95254

Stars: ★★☆☆

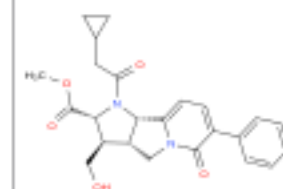
[\(2R,3R,3aS,9bS\)-1-\(1,3-benzodioxol-5-ylmethyl\)-7-\(1-cyclohexenyl\)-3-\(hydroxymethyl\)-6-oxo-3,3a,4,9b-tetrahydro-2H-pyrrolo\[2,3-a\]indolizine-2-carboxylic acid methyl ester](#)



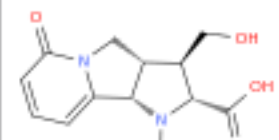
CHEBI:98944

Stars: ★★☆☆

[\(2R,3R,3aS,9bS\)-1-\(1,3-benzodioxol-5-ylmethyl\)-7-\(1-cyclohexenyl\)-3-\(hydroxymethyl\)-6-oxo-3,3a,4,9b-tetrahydro-2H-pyrrolo\[2,3-a\]indolizine-2-carboxylic acid methyl ester](#)



[\(2R,3R,3aS,9bS\)-1-\(cyclopentylcarbamoyl\)-3-\(hydroxymethyl\)-6-oxo-3,3a,4,9b-tetrahydro-2H-pyrrolo\[2,3-a\]indolizine-2-carboxylic acid](#)



CHEBI:98447

Stars: ★★☆☆

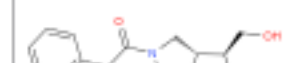
[\(2R,3R,3aS,9bS\)-1-\(cyclopentylmethyl\)-3-\(hydroxymethyl\)-7-\(2-methoxyphenyl\)-6-oxo-3,3a,4,9b-tetrahydro-2H-pyrrolo\[2,3-a\]indolizine-2-carboxylic acid methyl ester](#)



CHEBI:131226

Stars: ★★☆☆

[\(2R,3R,3aS,9bS\)-1-\(cyclopentylmethyl\)-3-\(hydroxymethyl\)-7-\(2-methoxyphenyl\)-6-oxo-3,3a,4,9b-tetrahydro-2H-pyrrolo\[2,3-a\]indolizine-2-carboxylic acid](#)



ChEBI

- One compound

CHEBI:17203 - L-proline

Main

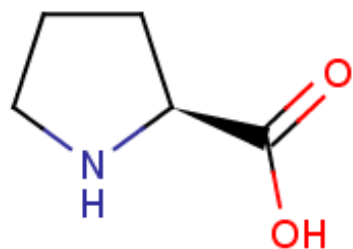
ChEBI Ontology

Automatic Xrefs

Reactions

Pathways

Models



ChEBI Name **L-proline**

ChEBI ID **CHEBI:17203**

ChEBI ASCII Name L-proline

Definition

Pyrrolidine in which the *pro-S* hydrogen at position 2 is substituted by a carboxylic acid group. L-Proline is the only one of the twenty DNA-encoded amino acids which has a secondary amino group α to the carboxyl group. It is an essential component of collagen and is important for proper functioning of joints and tendons. It also helps maintain and strengthen heart muscles.

Stars

☆☆☆ This entity has been manually annotated by the ChEBI Team.


Secondary ChEBI IDs

CHEBI:45159, CHEBI:45100, CHEBI:45040, CHEBI:42067, CHEBI:184637, CHEBI:6286, CHEBI:13154, CHEBI:21373

Supplier Information



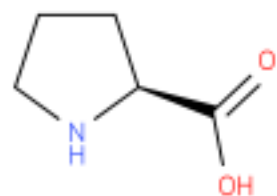
[eMolecules:524642](#), [ZINC00895360](#)

 [Download Molfile](#)

- [Find compounds which contain this structure](#)
- [Find compounds which resemble this structure](#)
- [Take structure to the Advanced Search](#)

Search Results for All in ChEBI

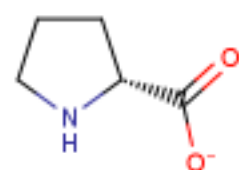
similar structures



Edit Search

125 entries found, displaying 1 to 15.

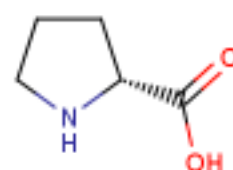
D-prolinate



CHEBI:32867

Stars: ★★★★★

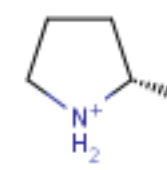
D-proline



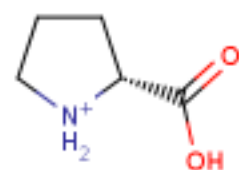
CHEBI:16313

Stars: ★★★★★

D-proline



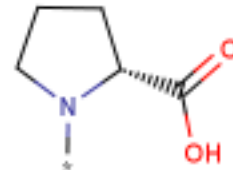
D-prolinium



CHEBI:32868

Stars: ★★★★★

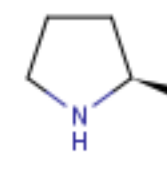
D-prolino group



CHEBI:32870

Stars: ★★★★★

L-prolinate



L-proline



CHEBI:17203

Stars: ★★★★★

L-proline zwitterion



CHEBI:60039

Stars: ★★★★★

L-prolinium



Outline

- Comprehensive metabolomics databases
- Compound databases
- **Spectral databases**
- Metabolic pathway databases
- Drug databases
- Disease & physiology databases
- Raw data databases

Spectral databases

- NIST 14
- METLIN
- MassBank
- MoNA
- Gold Metabolome Database
- Feign GC-MS database
- HMDB
- BMRB
- Madison Metabolomics Consortium Database
- BML-NMR
- mzCloud

Spectral databases

- NIST 14
- METLIN
- MassBank
- MoNA
- Gold Metabolome Database
- Feign GC-MS database
- HMDB
- BMRB
- Madison Metabolomics Consortium Database
- BML-NMR
- mzCloud

NIST 14

- Electron ionization mass spectral library
 - 276,259 spectra of 242,477 unique compounds

- MS/MS library: 234,284 spectra
 - 51,216 ion trap spectra for 42,126 different ions of 8,171 compounds
 - 183,068 collision cell spectra (QTOF and tandem quad) spectra for 14,835 different ions of 7,692 compounds

NIST 14 EI library

NIST MS Search 2.2 - [Name search]

File Search View Tools Options Window Help

MS I m/z ← ?

DODECANOICACID Clear a-z mainlib From b

Dodecanedioic acid, diTBDMS
Dodecanedioic acid di-TMS
Dodecanedioic acid, TBDMS de
Dodecanehydrazide, N2-(phenyl)
Dodecanenitrile
Dodecanethioamide, N,N-diethyl
Dodecanethiol-(1)
Dodecanoic acid
Dodecanoic acid, 10-methyl-, me
Dodecanoic acid, 10-oxo-
Dodecanoic acid, 10-undecen-1
dodecanoic acid, 1,10-decanedi
Dodecanoic acid, 1,1',1''-(1,2,3-
Dodecanoic acid, 11-amino-, me
Dodecanoic acid, 1,1'-biphenyl-4
Dodecanoic acid, 1,1'-(dibutylsta
Dodecanoic acid, 1,1-dimethyletl
Dodecanoic acid, 1,1-dimethylpn
Dodecanoic acid, 11-hydroxy-, m
Dodecanoic acid, 11-oxo-, methy
Dodecanoic acid, 1,2,3-propane
Dodecanoic acid, 12-(4-methylp
Dodecanoic acid, 12-amino-
Dodecanoic acid, 12-amino-, lac
Dodecanoic acid, 12-bromo-
Dodecanoic acid, 12-hydroxy-
Dodecanoic acid, 12-hydroxy-, m

(mainlib) Dodecanoic acid

Name: Dodecanoic acid
Formula: C₁₂H₂₄O₂
MW: 200 Exact Mass: 200.17763 CAS#: 143-07-7 NIST#: 221043 ID#
Other DBs: Fine, TSCA, RTECS, EPA, HODOC, NIH, EINECS, IRDB
Contributor: Chemical Concepts
Related CAS#: 8000-62-2; 8045-27-0; 7632-48-6; 203714-07-2
InChIKey: POULHZVOKOAJMA-UHFFFAOYSA-N Non-stereo
Synonyms:
1.n-Dodecanoic acid
2.Neo-fat 12
3.Alphat no. 4
4.ABL

Names Structures Plot/Text Plot

Lib. Search Other Search Names Compare Librarian MSMS

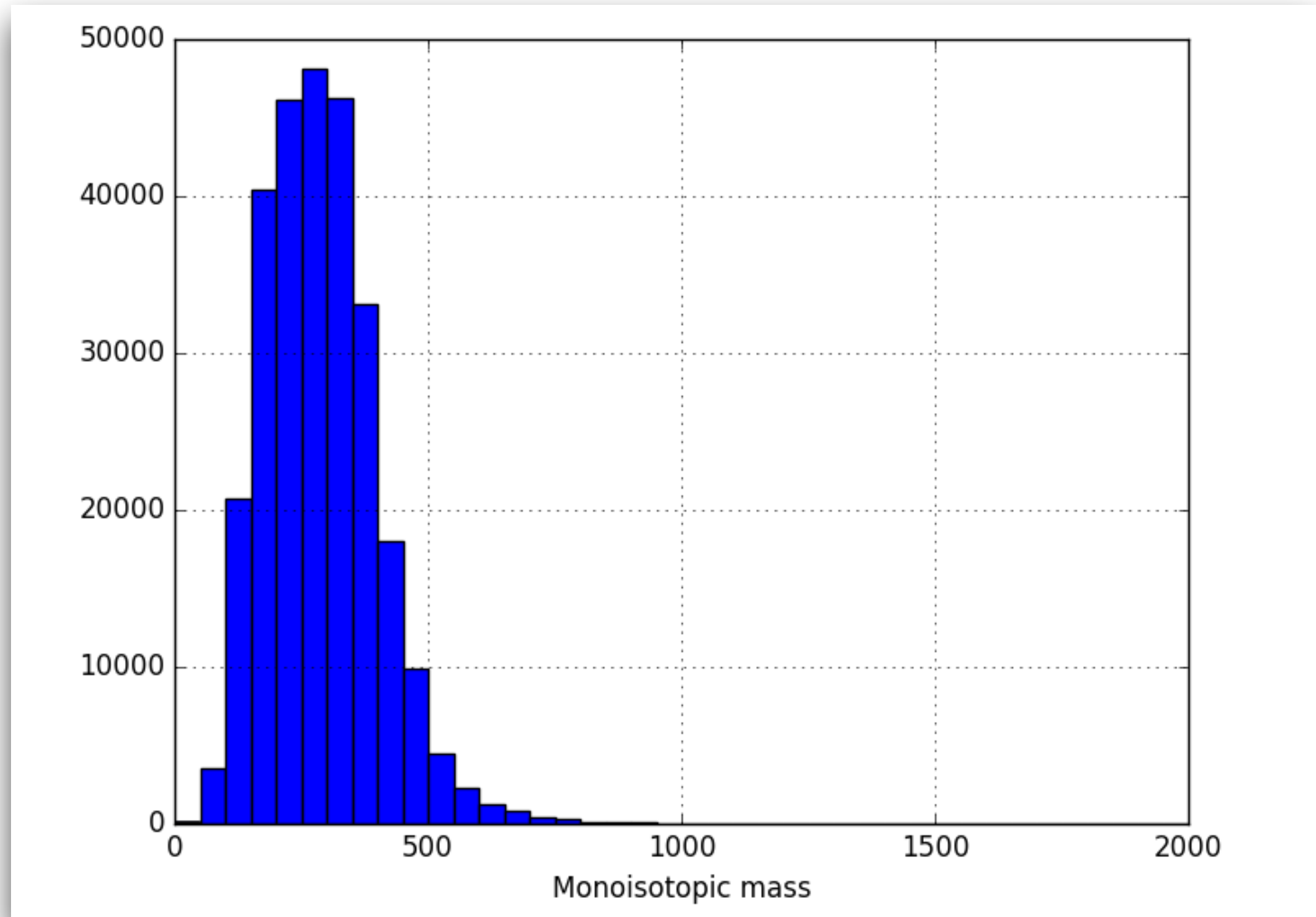
For Help, press F1

NIST 14 EI library

- Focuses on
 - Drugs, metabolites, and poisons
 - Pesticides and fungicides
 - Organics present in soil, water, and air
 - Amino acids, di- and tri-peptides
 - Common sample contaminants
 - Common analytical derivatives of the above

NIST 14 EI library

- Breakdown by mass



NIST 14 MS/MS library

NIST MS Search 2.2 - [MSMS Library]

File Search View Tools Options Window Help

MS I m/z

MW=121
MW=122

- C4H10O2S
 - 2,2'-Thiodiethanol
- C4H10O4
 - meso-Erythritol
 - L-Threitol
- C4H2D4O4
 - Succinic acid-2,2,3,3-d4
- C5H13CIN
- C6H6N2O
- C7H10N2
- C7H6O2
- C8H10O
 - 2-Phenylethanol
 - Benzenemethanol, 2-methyl-
 - Benzenemethanol, .alpha.-n
 - 3-Methylbenzyl alcohol
 - Benzenemethanol, 4-methyl-
 - MS2; P 105.0699; CE 4
 - MS2; P 105.0699; CE 5
 - MS2; P 105.0699; CE 7
 - MS2; P 105.0699; CE 10
 - MS2; P 105.0699; CE 15
 - MS2; P 105.0699; CE 17
 - MS2; P 105.0699; CE 20
 - MS2; P 105.0699; CE 23
 - MS2; P 105.0699; CE 25**
 - MS2; P 105.0699; CE 27

(nist_msms) Benzenemethanol, 4-methyl- [M+H-H2O]+ QTOF 25V

Name: Benzenemethanol, 4-methyl-
Formula: C₈H₁₀O
MW: 122 Exact Mass: 122.073165 CAS#: 589-18-4 NIST#: 1
Other DBs: None
Comment: NIST Mass Spectrometry Data Center
Precursor type: [M+H-H2O]+
Spectrum type: MS2
Precursor m/z: 105.0699
Instrument type: Q-TOF
Instrument: Agilent QTOF 6530
Sample inlet: direct flow injection
Ionization: ESI
In-source voltage: 150
Collision gas: N₂
Collision energy: 25
Ion mode: P
Notes: Consensus spectrum; Acetonitrile/Water/Formic acid
InChIKey: KMTDMTZBNGUNX-UHFFFAOYSA-N Non-ster
Synonyms:
1.4-Methylbenzyl alcohol

Plot/Text Plot

Lib. Search Other Search Names Compare Librarian MSMS

For Help, press F1

METLIN

<https://metlin.scripps.edu/index.php>



Scripps Center for Metabolomics

[MS HOME](#)

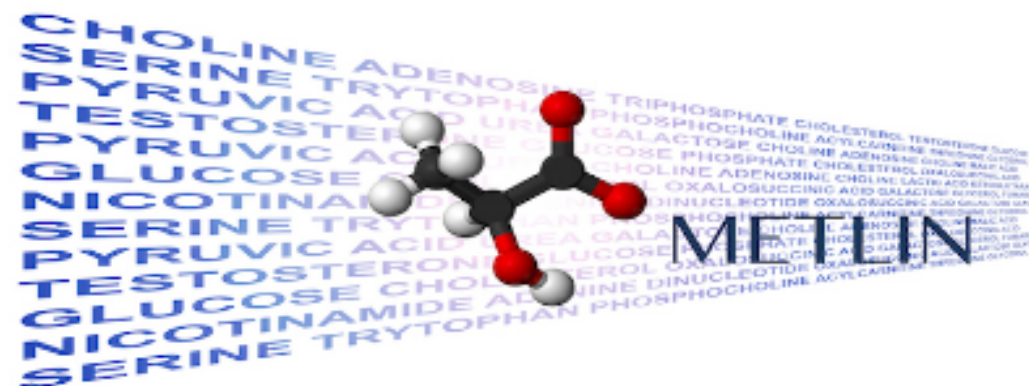
[METLIN](#)

[XCMS Online](#)

[XCMS Institute](#)

[XCMS Public](#)

[Contact](#)



Statistics

- # Metabolites: 242,032
- # High Resolution MS/MS Spectra: 72,268
- # Metabolites w/ High Resolution MS/MS: 14,034

[example](#) | [details...](#)

Functionality

- **Single & Batch**
Precursor Ion (m/z) searching
- **Single & Multiple**
Fragment Ion (m/z) searching
- **Neutral Loss** searching
- **De Novo** Fragment Characterization

METLIN

METLIN: Metabolite Search Advanced

[Simple \(Saved Searches\)](#) | [Advanced](#) | [Batch](#) | [Fragment](#) | [Neutral Loss](#) | [MS/MS Spectrum Match](#) | [Unknowns](#)

MID:

Mass: -

Name:

Formula:

CAS:

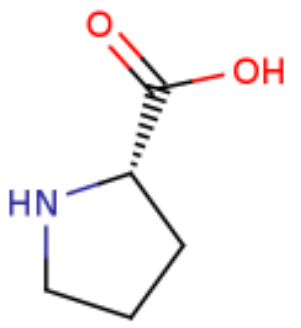
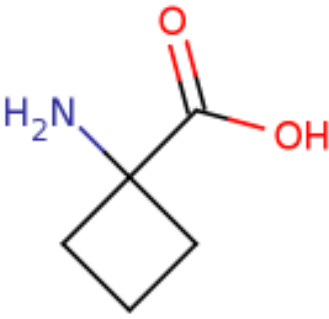
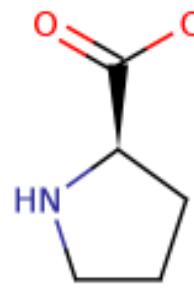
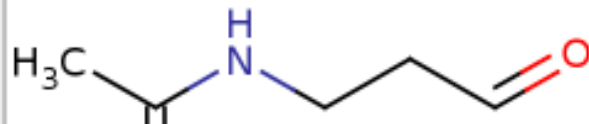
KEGG:

Search only the MS/MS data

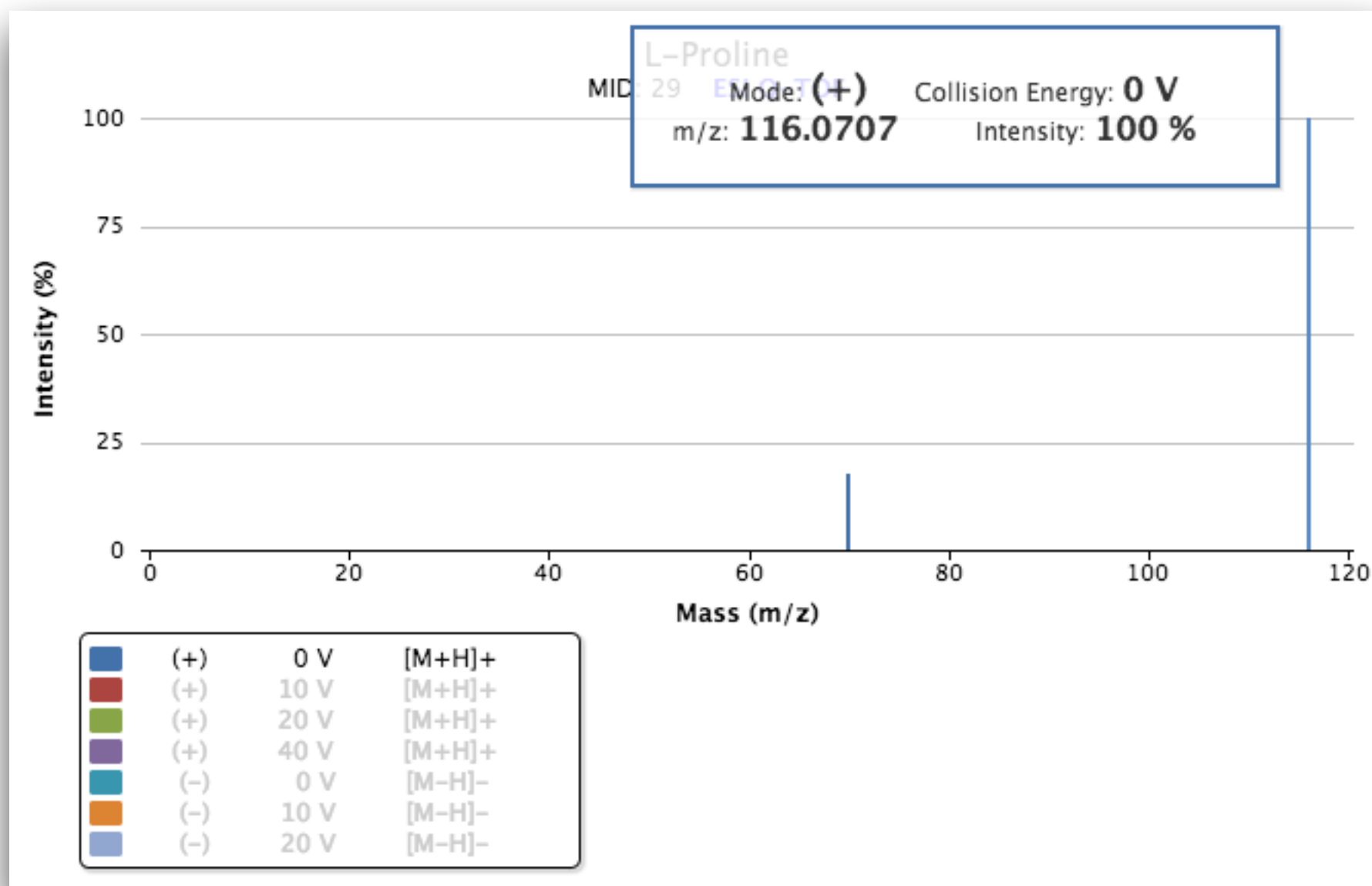
Remove peptides from search

Remove drug from search

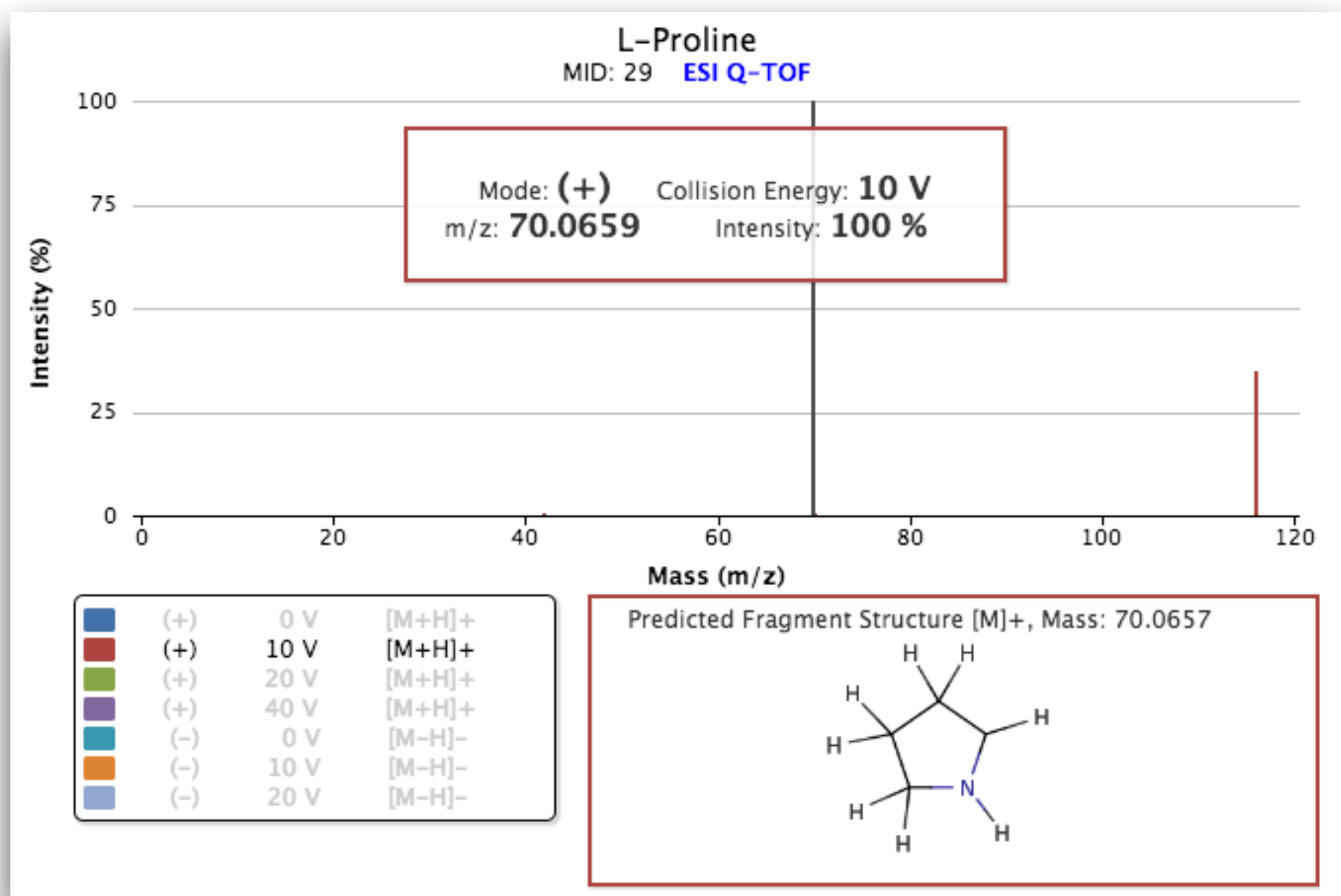
METLIN

| MID | Mass | Name | Formula | CAS | KEGG | MS/MS | Structure |
|-------|----------|------------------------------------|---|------------|--------|----------------------|---|
| 29 | 115.0633 | L-Proline | C ₅ H ₉ NO ₂ | 147-85-3 | C00148 | View |  |
| 44305 | 115.0633 | 1-AMINOCYCLOBUTANE CARBOXYLIC ACID | C ₅ H ₉ NO ₂ | 22264-50-2 | | View |  |
| 58150 | 115.0633 | D-Proline | C ₅ H ₉ NO ₂ | 344-25-2 | C00763 | View |  |
| 72070 | 115.0633 | 3-Acetamidopropanal | C ₅ H ₉ NO ₂ | | C18170 | NO |  |

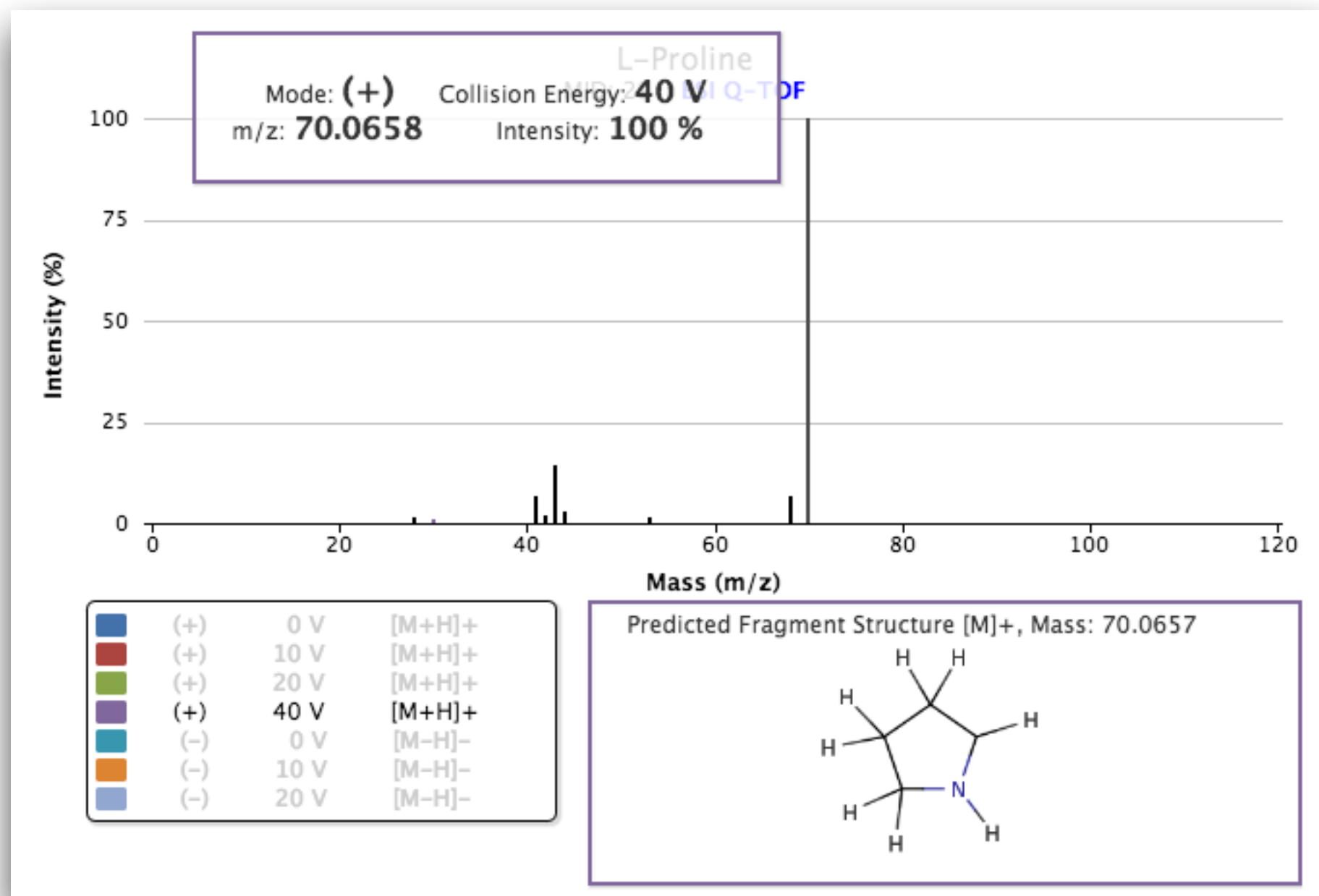
METLIN



METLIN



METLIN



MassBank

- About

www.massbank.jp/en/about.html



High Quality Mass Spectral Database

 Database Service

 Statistics

 Publications

 Download

 Manuals

 About MassBank

 Contact

 Consortium Members

About MassBank

What is MassBank?

MassBank is the first public repository of mass spectral data for sharing them among scientific research community. MassBank data are useful for the chemical identification and structure elucidation of chemical compounds detected by mass spectrometry.

Features

- Distributed database
- High precision and accurate mass spectra of biologically endogenous and exogenous substances
- Mass spectral search by exact m/z and browsing interface
- Merged spectra as the reference data independent to analytical methods

A merged spectrum is generated from spectra of the identical compound measured in different CID conditions.

The MassBank Record for each merged spectrum includes information of its original spectra.

MassBank

- Database services

www.massbank.jp/en/database.html



High Quality Mass Spectral Database

Database Service

Spectrum Search

Quick Search

Peak Search

Substructure Search

Metabolome Prediction

Spectral Browser

Batch Service

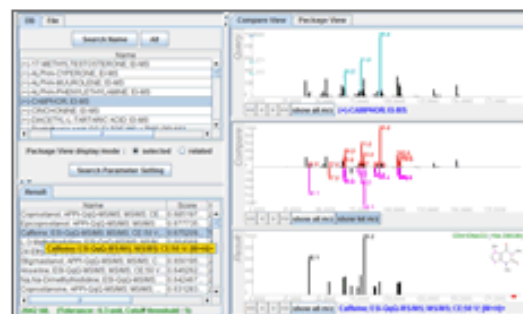
Browse Page

Record Index

Statistics

Database Service

Spectrum Search



Search similar spectra on a peak-to-peak basis

Retrieves spectra similar to user's spectrum in terms of the m/z value. This search is helpful to identify chemical compound by comparing similar spectra on a 3D-display.

[Spectrum Search](#)

[User's Manual](#)

Quick Search

The screenshot shows the Quick Search interface. It features a search form with two tabs: 'Search by Keyword' and 'Search by Peak'. The 'Search by Keyword' tab is active, showing a text input field for 'Compound Name' and a dropdown menu for 'Instrument Type'.

Keyword search of chemical compounds

Retrieves the chemical compound(s) specified by chemical name or

MassBank

← → ↻

Peak Search

[Home](#) | [Spectrum](#) | [Quick](#) | [Peak](#) | [Substructure](#) | [Prediction](#) | [Browser](#) | [Batch](#) | [Browse](#) | [Index](#) | MassBank ID:

Search of **Peaks** **Peak Differences**
Search by **m/z-Value** **Molecular Formula**

| | m/z | Formula |
|--------------------------------------|---------------------------------------|----------------------|
| AND <input type="button" value="v"/> | <input type="text" value="116.0707"/> | <input type="text"/> |
| AND | <input type="text"/> | <input type="text"/> |
| AND | <input type="text"/> | <input type="text"/> |
| AND | <input type="text"/> | <input type="text"/> |
| AND | <input type="text"/> | <input type="text"/> |
| AND | <input type="text"/> | <input type="text"/> |

Rel.Intensity Tolerance

Instrument Type

EI EI-B
 EI-EBEB
 GC-EI-QQ
 GC-EI-TOF

ESI CE-ESI-TOF
 ESI-ITFT
 ESI-ITTOF
 ESI-QTOF
 LC-ESI-IT

MS Type

All MS MS1 MS2 MS3 MS4

Ion Mode

Positive Negative Both

MassBank

Peak Search Results (Peaks by *m/z* value)

[Home](#) | [Spectrum](#) | [Quick](#) | [Peak](#) | [Substructure](#) | [Prediction](#) | [Browser](#) | [Batch](#) | [Browse](#) | [Index](#) | MassBank ID:

Search Parameters :

m/z: 116.0707 Rel.Int: 100 Tol.(unit): 0.3

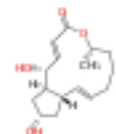
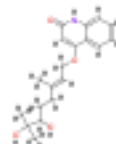
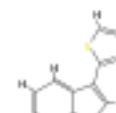
Instrument
Type: **LC-ESI-QTOF**
MS Type: **MS2**
Ion Mode: **Positive**

[Edit / Resubmit Query](#)

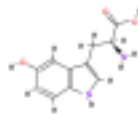
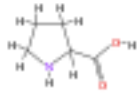
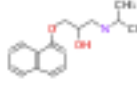
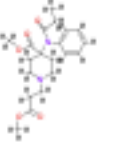

Results : **113 Hit.** (41 - 81 Displayed)

[First](#) [Prev](#) [1](#) [2](#) [3](#) [Next](#) [Last](#) (Total 3 Page)

▼ Results End

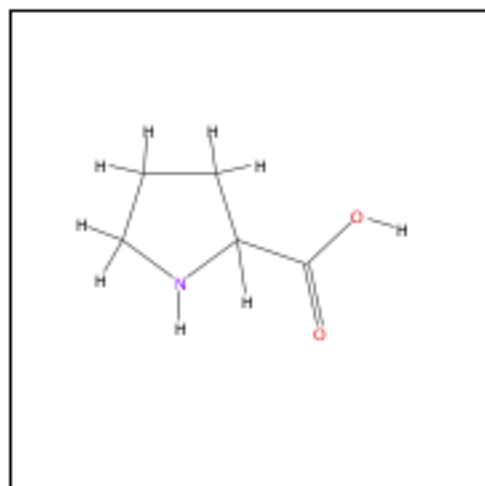
| <input type="checkbox"/> | Name ▲ | Formula / Structure | ExactMass | ID |
|--------------------------|--|---|------------------|----|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Brefeldin-A 1 spectrum | C16H24O4  | 280.16746 | |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Bucharaine 1 spectrum | C19H25NO4  | 331.17836 | |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Camalexin 2 spectra | C11H8N2S  | 200.04082 | |

MassBank

| | | | |
|---|---|-------------------------|------------------------------|
| <p><input checked="" type="checkbox"/> Oxitriptan 1</p> <p>1 spectrum</p> | <p>C₁₁H₁₂N₂O₃</p>  | <p>220.08479</p> | |
| <p><input type="checkbox"/> Proline</p> <p>2 spectra</p> <p> LC-ESI-QTOF; MS2; CE:10 eV; [M+H]⁺ LC-ESI-QTOF; MS2; CE:15 eV; [M+H]⁺ </p> | <p>C₅H₉NO₂</p>  | <p>115.06333</p> | <p>PB000449 PB000450</p> |
| <p><input checked="" type="checkbox"/> Propranolol</p> <p>1 spectrum</p> | <p>C₁₆H₂₁NO₂</p>  | <p>259.15720</p> | |
| <p><input checked="" type="checkbox"/> Remifentanil</p> <p>4 spectra</p> | <p>C₂₀H₂₈N₂O₅</p>  | <p>376.19982</p> | |
| <p><input checked="" type="checkbox"/> S-Lactoylglutathione</p> <p>1 spectrum</p> | <p>C₁₃H₂₁N₃O₈S</p>  | <p>379.10494</p> | |

Proline; LC-ESI-QTOF; MS2; CE:15 eV; [M+H]⁺

Mass Spectrum Chemical Structure



ACCESSION: PB000450
RECORD_TITLE: Proline; LC-ESI-QTOF; MS2; CE:15 eV; [M+H]⁺
DATE: 2016.01.19 (Created 2008.01.02, modified 2013.06.04)
AUTHORS: Boettcher C, Institute of Plant Biochemistry, Halle, Germany
LICENSE: [CC BY-SA](#)
COMMENT: IPB_RECORD: 261
COMMENT: CONFIDENCE: confident structure

CH\$NAME: Proline
CH\$NAME: pyrrolidine-2-carboxylic acid
CH\$COMPOUND_CLASS: Natural Product; amino acid
CH\$FORMULA: [C5H9NO2](#)
CH\$EXACT_MASS: 115.06333
CH\$SMILES: C1CC(NC1)C(=O)O
CH\$IUPAC: InChI=1S/C5H9NO2/c7-5(8)4-2-1-3-6-4/h4,6H,1-3H2,(H,7,8)
CH\$LINK: INCHIKEY [ONIBWKKTOPOVIA-UHFFFAOYSA-N](#)
CH\$LINK: KEGG [C00148](#)
CH\$LINK: PUBCHEM CID:[614](#)

AC\$INSTRUMENT: API QSTAR Pulsar i
AC\$INSTRUMENT_TYPE: LC-ESI-QTOF
AC\$MASS_SPECTROMETRY: MS_TYPE MS2
AC\$MASS_SPECTROMETRY: ION_MODE POSITIVE
AC\$MASS_SPECTROMETRY: COLLISION_ENERGY 15 eV
AC\$MASS_SPECTROMETRY: IONIZATION ESI

MS\$FOCUSED_ION: PRECURSOR_TYPE [M+H]⁺

PK\$SPLASH: [splash10-01b9-8900000000-000cb5f6ae87b559450a](#)

MassBank

mona.fiehnlab.ucdavis.edu/spectra/splash/splash10-01b9-8900000000-000cb5f6ae87b559450a

MoNA - MassBank of North America

Browse ▾

Statistics ▾

Resources ▾

Download

Upload

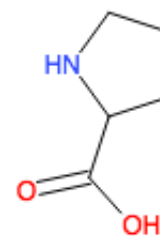
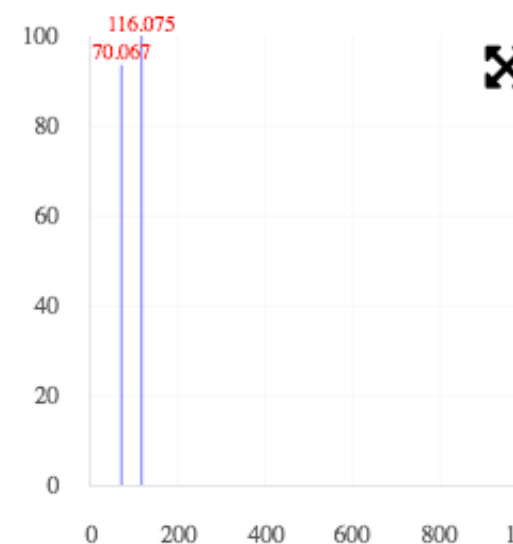
Search...

display generated query



Proline

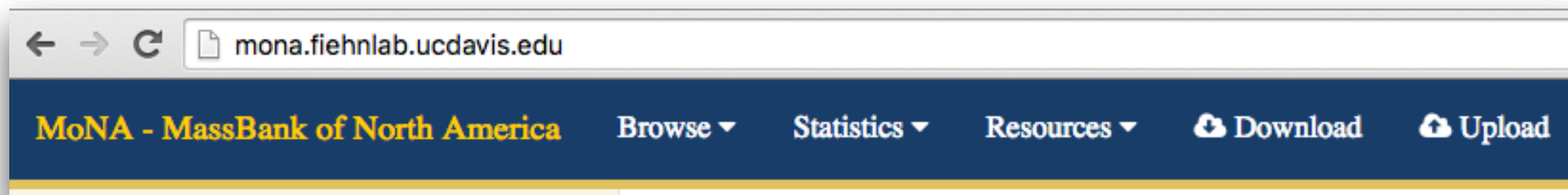
Score: ★★★★★★★★★★☆☆



| | |
|--------------------|------------------------------|
| Q accession | PB000450 |
| Q authors | Boettcher C, Institute of... |
| Q collision energy | 15 eV |
| Q compound class | Natural Product; amino ac... |
| Q exact mass | 115.0633 |
| Q instrument | API QSTAR Pulsar i |
| Q instrument type | LC-ESI-QTOF |
| Q ion mode | positive |
| Q ionization | ESI |
| Q license | CC BY-SA |
| Q mass accuracy | 8754.0881 |
| Q mass error | -1007.2745 |
| Q ms level | MS2 |
| Q origin | PB000450.txt |
| Q precursor type | [M+H] ⁺ |

Originally submitted to the MassBank Spectral Database as [PB000450](#)

MoNA



Welcome to MoNA

MassBank of America (MoNA), is an auto curating repository for storing, comparing and querying mass spectra of chemical compounds. It is metadata centric and it was designed to allow easy integration into other tools by utilize its REST based application programming interface. At the current time it contains over 200k predicted and 40k unique experimental mass spectra and their associated metadata. The predicted spectra were obtained by utilizing the lipid blast library and the experimental spectra were acquired from 30 different facilities all over the world, including several major research facilities in the United States and Japan. MoNA is utilizing the InChI Key as unique identifier for chemicals and is designed for easy scalability and expandability. This is realized by utilizing common applications like nginx, grails, AngularJS, postgresSQL and tomcat. MoNA is currently integrated in applications like MSDial, BinBase, MZMine and the statistics package R. This was accomplished by utilizing its REST based API, which is also utilized by its main AngularJS based web interface. We consider MoNA to be highly useful for crosslinking mass spectra in publications, identification of unknowns and integration in data acquisition software.

Metabolic pathway databases

- KEGG
- MetaCyc
- HumanCyc
- BioCyc
- Reactome
- WikiPathways

Drug databases

- DrugBank
- Therapeutic target databases
- PharmGKB
- STITCH
- SuperTarget

Disease & physiology databases

- OMIM
- METAGENE
- OMMBID

Raw data databases

- Metabolomics Workbench
 - Funded by the NIH Common Fund Metabolomics Program
 - Serve as a national and international repository for metabolomics data and metadata
 - Provide access to raw data, metabolite standards, protocols


Raw data databases

- Metabolomics Workbench
- MetaboLights

Metabolomics workbench

- Website

www.metabolomicsworkbench.org



METABOLOMICS WORKBENCH

Log in / Register

[Home](#) | [Metabolomics Update](#) | [Data](#) | [Protocols](#) | [Standards](#) | [Resources](#) | [NIH Metabolomics](#) | [Training](#) | [About](#)

Personnel (Restricted access)

Welcome to the UCSD Metabolomics Workbench, a resource sponsored by the Common Fund of the National Institutes of Health.

Metabolomics News

06-06-2016 - **Human Metabolome Gene/Protein Database (MGP)** - The Human Metabolome Gene/Protein Database (MGP) of metabolome-related genes and proteins is now available via the Metabolomics Workbench; MGP contains data for over 7300 genes and over 15,500 proteins.

▶ [news archive](#)

Events Calendar

July 17-21, 2016
Birmingham, AL, USA

[2016 International Summer Sessions in](#)

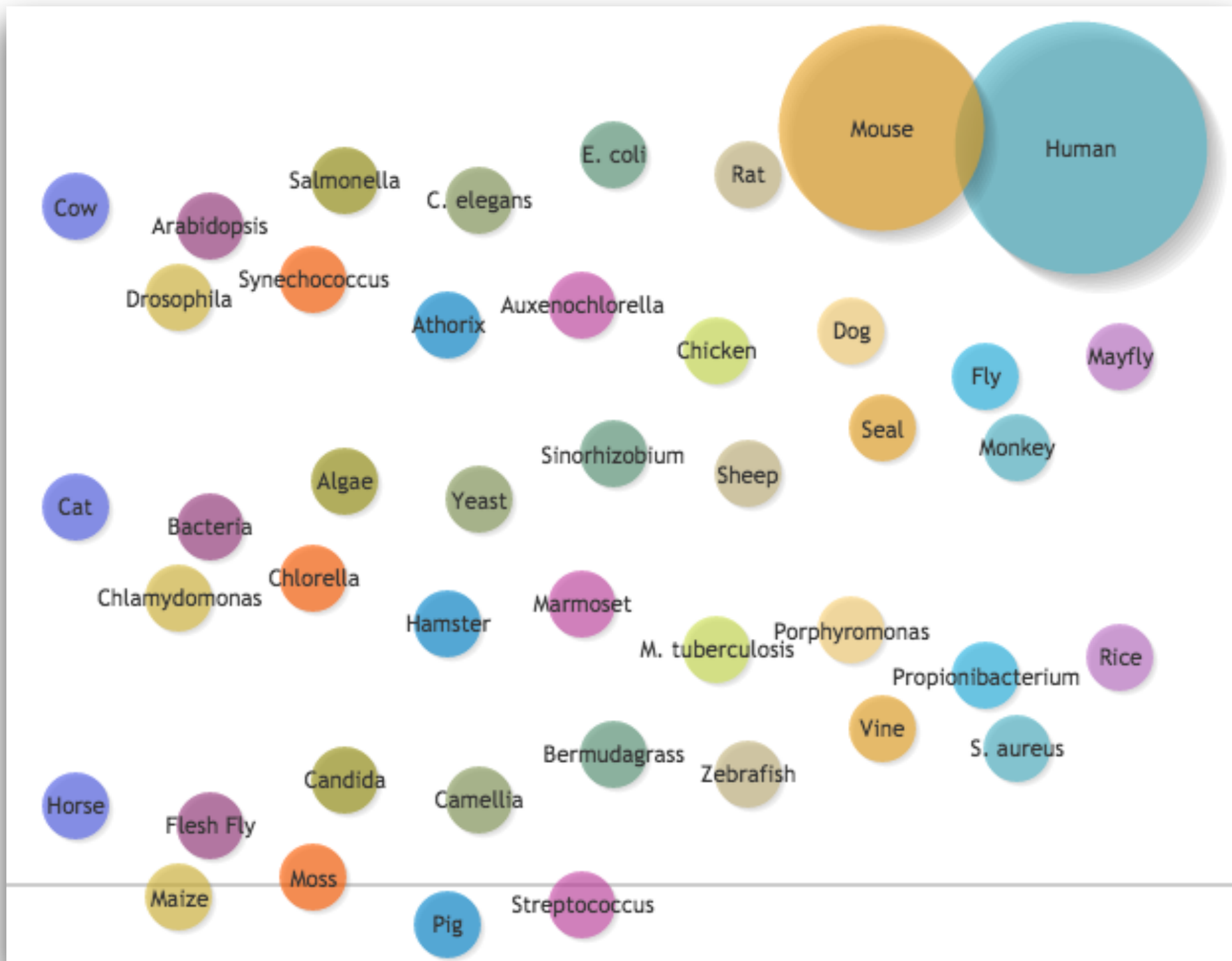
Raw data databases

- Metabolomics Workbench: **summary of all studies**

| Study ID ↑↓ | Study Title ↑↓ | Species ↑↓ | Institute ↑↓ | Analysis ↑↓ | Release Date | Version | Samples | Download (* : Contains raw data) |
|----------------|---|----------------------|---------------------------------|----------------|--------------|---------|---------|-------------------------------------|
| ST000001 | Fatb Induction Experiment (FatBIE) | Arabidopsis thaliana | University of California, Davis | MS | 2013-02-14 | 1 | 24 | Uploaded data (476K) |
| ST000002 | Intestinal Samples II pre/post transplantation | Homo sapiens | University of California, Davis | MS | 2013-02-22 | 1 | 12 | Uploaded data (664K) |
| ST000003 | Metabolomic analysis of mouse embryonic fibroblasts, embryonic stem cells, and induced pluripotent stem cells | Mus musculus | University of California, Davis | MS | 2013-02-15 | 1 | 18 | Uploaded data (5.3G)* |
| ST000004 | Lipidomics studies on NIDDK / NIST human plasma samples | Homo sapiens | LIPID MAPS | MS | 2013-03-17 | 1 | 8 | Uploaded data (48K) |
| ST000005 | Timecourse on RAW 264.7 cells treated with Kdo2-Lipid A and compactin | Mus musculus | LIPID MAPS | MS | 2013-03-22 | 1 | 696 | Uploaded data (56K) |
| ST000006 | White Wine Study | Vitis vinifera | University of California, Davis | MS | 2013-03-23 | 1 | 102 | Uploaded data (532K) |
| ST000007 | Rice Infection Study | Oryza sativa | University of California, Davis | MS | 2013-03-24 | 1 | 60 | Uploaded data (1.7M) |

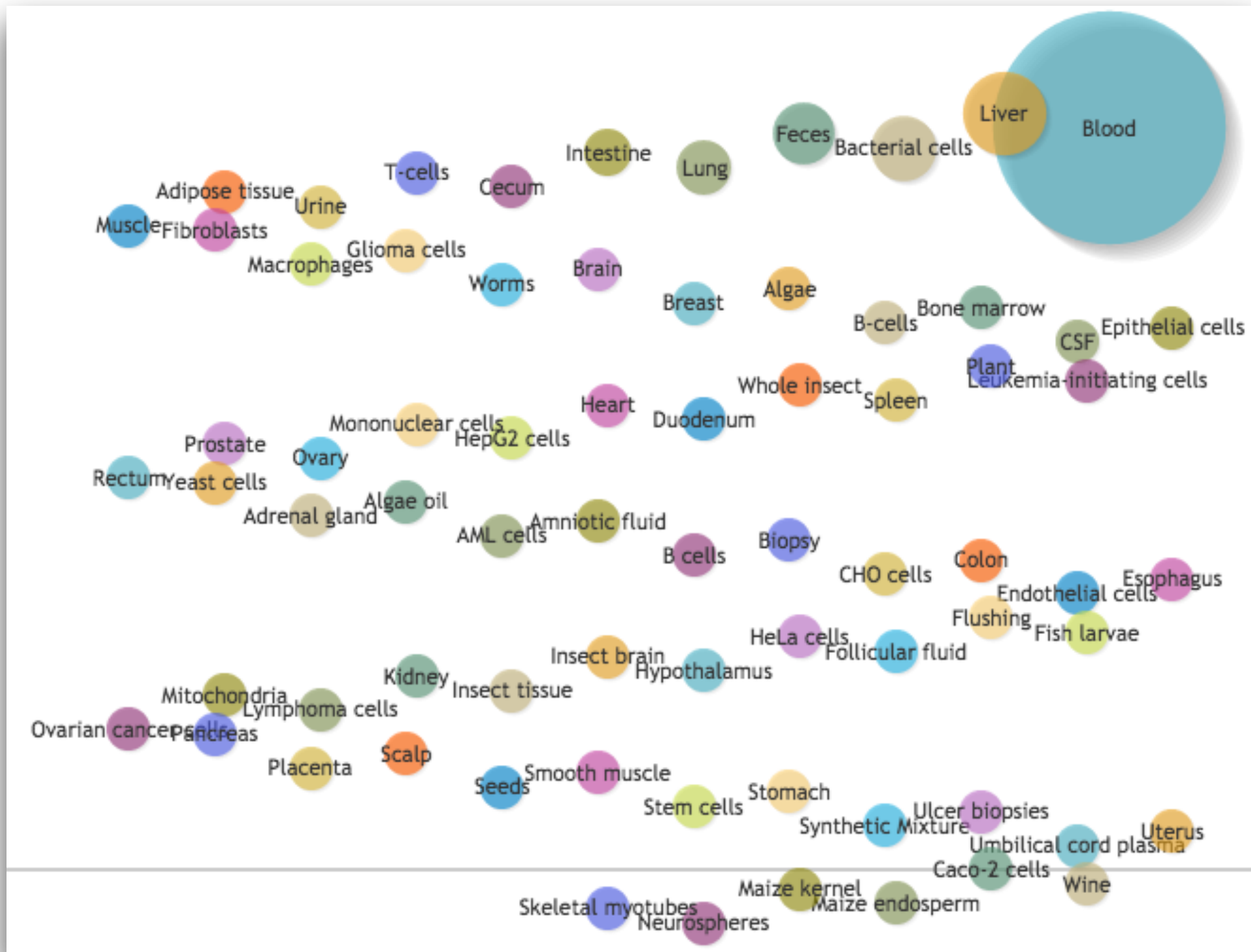
Raw data databases

- Metabolomics Workbench: **species**



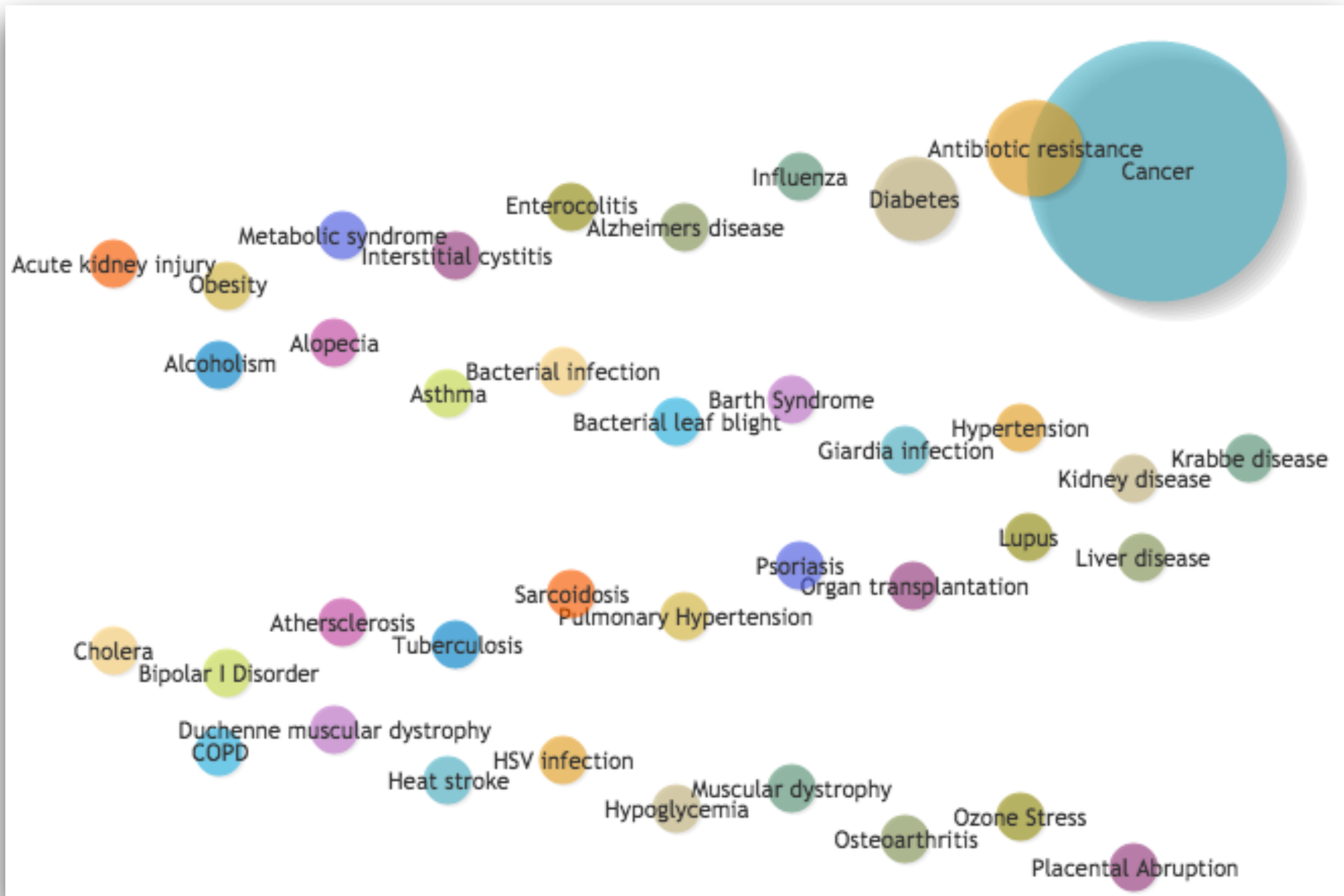
Raw data databases

- Metabolomics Workbench: **sample sources**



Raw data databases

- Metabolomics Workbench: **diseases**



Raw data databases

- MetaboLights

The screenshot shows the MetaboLights website homepage. At the top, there is a browser address bar with the URL www.ebi.ac.uk/metabolights/. Below the address bar is a cookie consent banner. The main header features the EMBL-EBI logo on the left and navigation links for Services, Research, Training, and About us on the right. The MetaboLights logo is prominently displayed in the center, accompanied by a search bar with a search button and example search terms: alanine, Homo sapiens, urine, MTBLS1. A secondary navigation bar contains links for Home, Browse Studies, Browse Compounds, Browse Species, Analysis, Download, Help, Give us feedback, and About, along with Submit Study and Login buttons. The main content area is divided into three columns: 'MetaboLights' (introduction), 'Download' (pre-packaged ISAcreeator and experiments), and 'Tweets by @MetaboLights' (social media updates).

EMBL-EBI

This website uses cookies. By continuing to browse this site, you are agreeing to the use of our site cookies. To find out more, see our [Terms of Use](#). OK

Services Research Training About us

MetaboLights

Search

Examples: [alanine](#), [Homo sapiens](#), [urine](#), [MTBLS1](#)


Home Browse Studies Browse Compounds Browse Species Analysis Download Help Give us feedback About Submit Study Login


MetaboLights

MetaboLights is a database for Metabolomics experiments and derived information.


The database is cross-species, cross-technique and covers metabolite structures and their reference spectra as well as their biological roles, locations and concentrations, and experimental data from metabolic experiments. [About MetaboLights](#).


Download

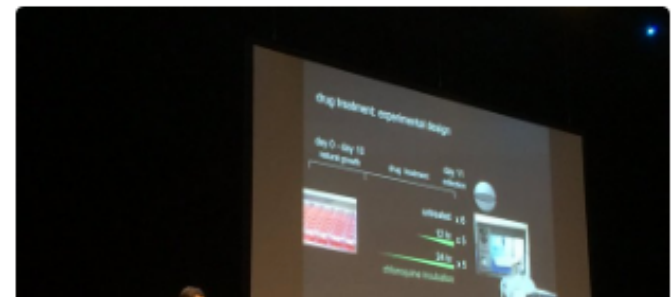
 **Pre-packaged ISAcreeator download.** To make it easy for new users, please download and just unzip our pre-packaged ISAcreeator with plugin and configurations.

 **Experiments.** All public MetaboLights experiments can be downloaded from our public [ftp archive](#). Please find zip archives under the "studies" folder. Each public study can be found in the corresponding MTBLS-id folder. Complete experiments can be opened with [ISAcreeator](#) or you can extract the archives using your normal unzip program.

Tweets by @MetaboLights

 MetaboLights Retweeted

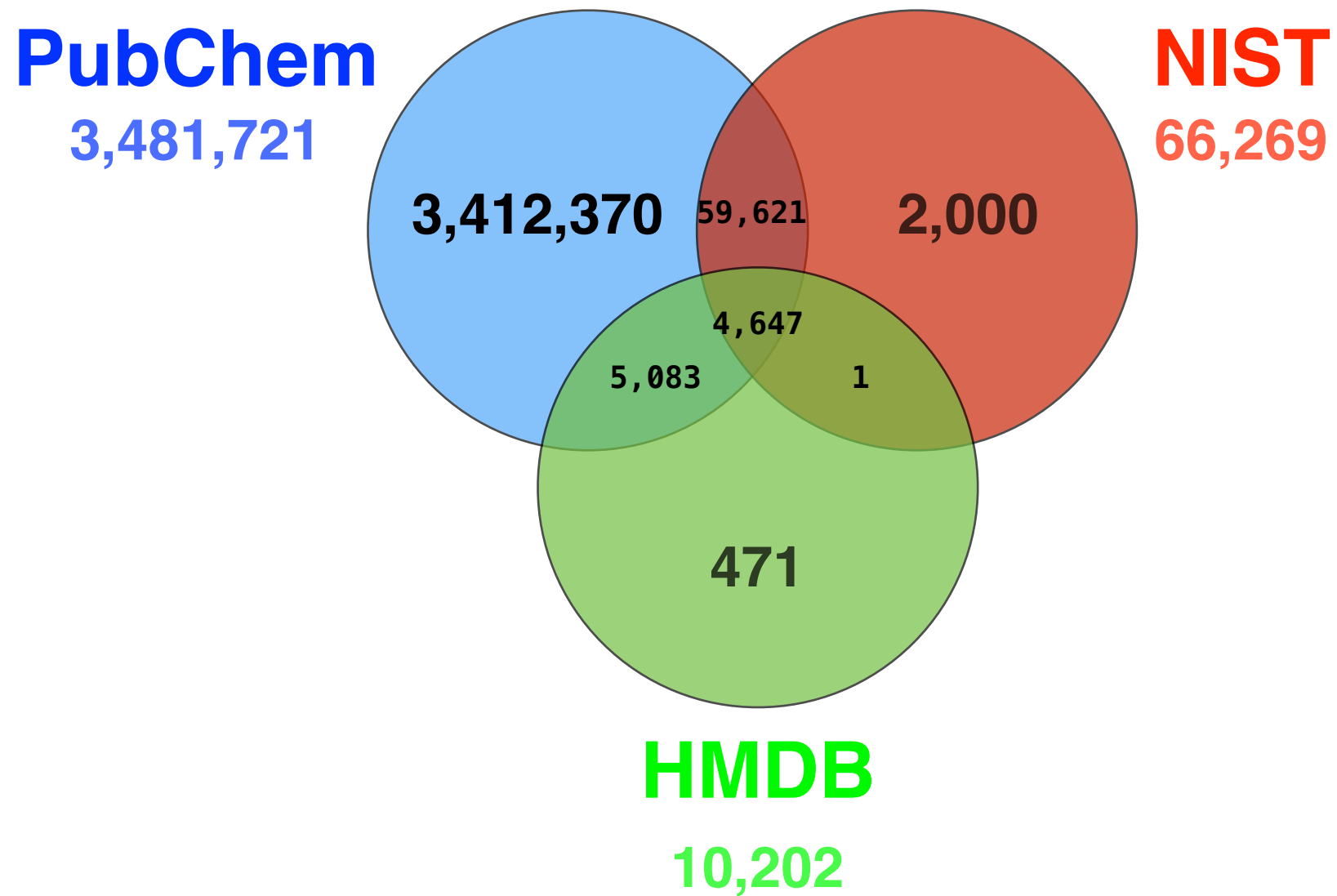
 **Reza Salek** @metaboknight
Andrew Palmer; on Spatio-temporal [#metabolomics](#) of tumours using [@chebit](#), BiNChE ebi.ac.uk/chebi/tools/bi...



[Embed](#) [View on Twitter](#)

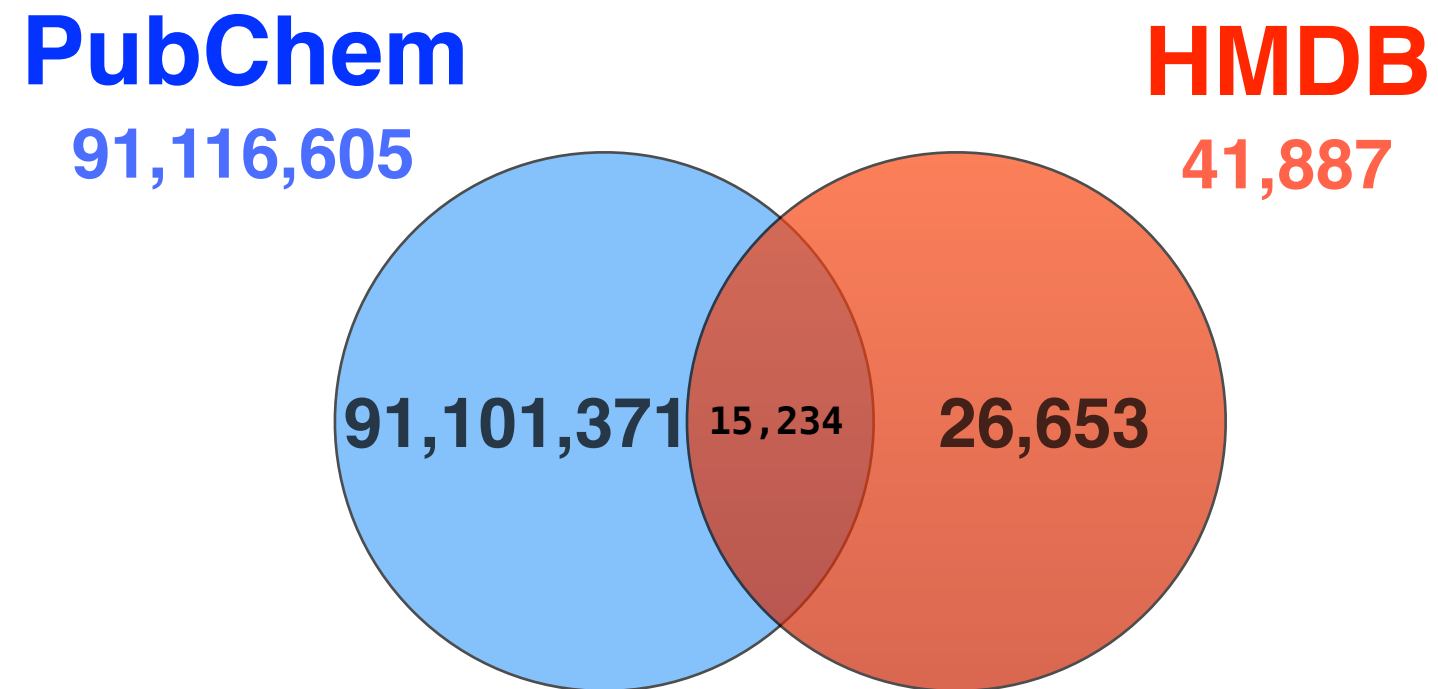
PubChem, NIST, and HMDB, again

- In terms of unique molecular formula



PubChem, NIST, and HMDB, again

- In terms of unique InChi Key



Acknowledgement

- Aleksandr Smirnov



Thank you!