

Tutorial for browsing and searching the Chemical Functional Ontology (ChemFOnT)

August 15, 2022

Updated September 14, 2022

1. Structure and design of ChemFOnt.

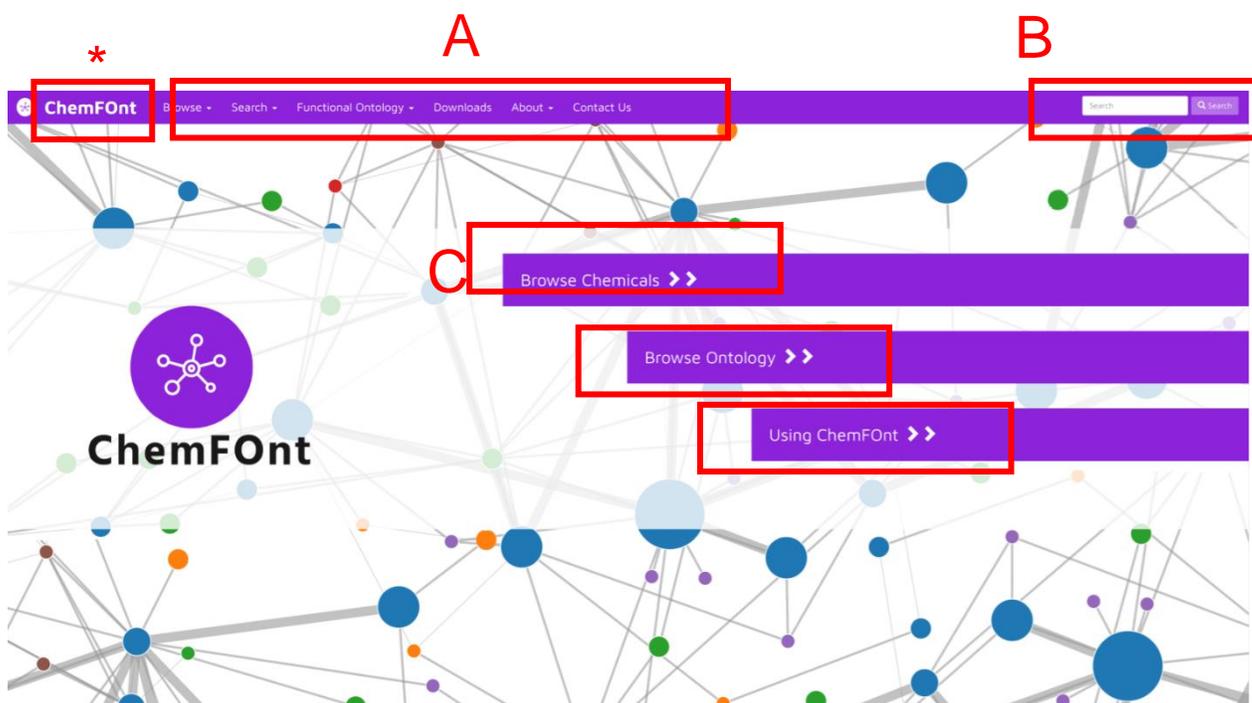
- In simple terms, the **C**hemical **F**unctional **O**ntology or ChemFOnt (<https://chemfont.ca>) is a hierarchical ontology describing the functions and actions of biologically important chemicals.
- ChemFOnt contains information on >319,000 chemicals, including primary metabolites, secondary metabolites, natural products, food chemicals, synthetic food additives, drugs, herbicides, pesticides, and environmental chemicals.
- Within the entry (Molecule Card) of each chemical, a functional ontology describes the roles, disposition, physiological effects, and processes associated with the chemical. Other sections include a chemical taxonomy and physical properties.

The next sections provide details about how to browse and search ChemFOnt.

2. Open your Internet browser and type “https://chemfont.ca” or simply “chemfont.ca” into the search bar and press the “return” (“Enter”) key. Below is an example of having typed in the url (web address) of the ChemFOnt database into the search bar.



- After submitting the search (pressing the Enter key), you will be taken to the ChemFOnt homepage (shown below).
- Next: a few points about the choices available on the homepage.



3. **A.** Across the top of the homepage, there are six different menu headings, some with their own pulldown menu: **Browse, Search, Ontology, Downloads, About, and Contact Us.**

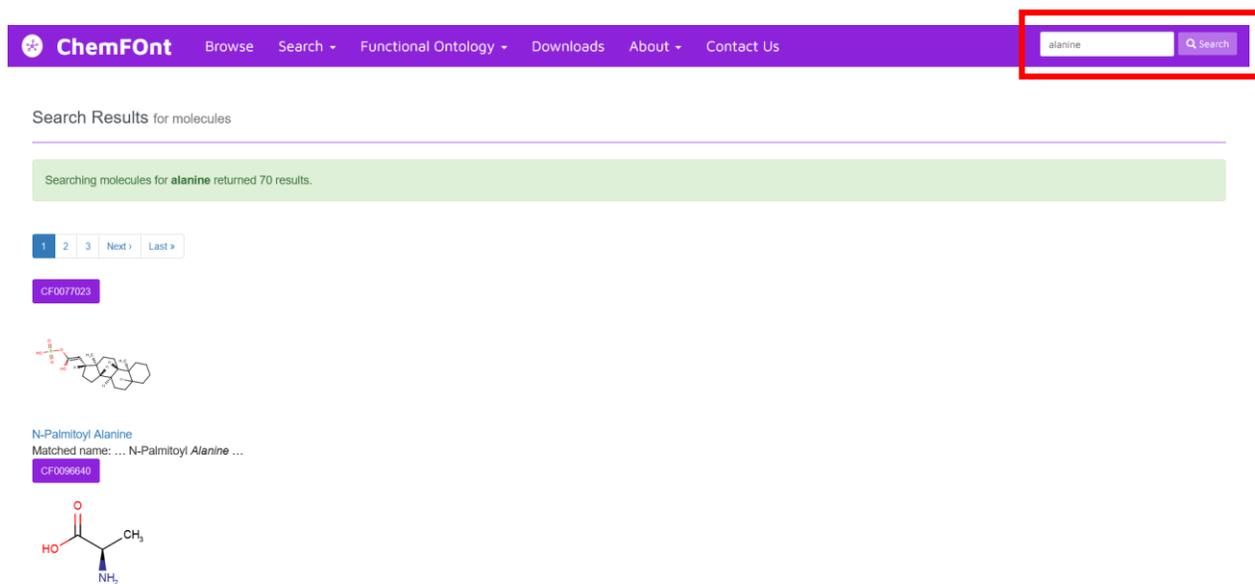
B. There is also a convenient general-text “**Search**” box in the top right corner which enables you to search the entire database. Clicking the purple “Search” button will return a list of results that match closely your query.

C. You can also “Browse Data”, “Search Data”, or learn more about “Using ChemFOnT” by clicking on these hyperlinked bars in the centre of the homepage.

* Clicking on “ChemFOnT” at the top left any time brings you back to this homepage.

➔ Next, we provide examples of general text searches using the “Search” box.

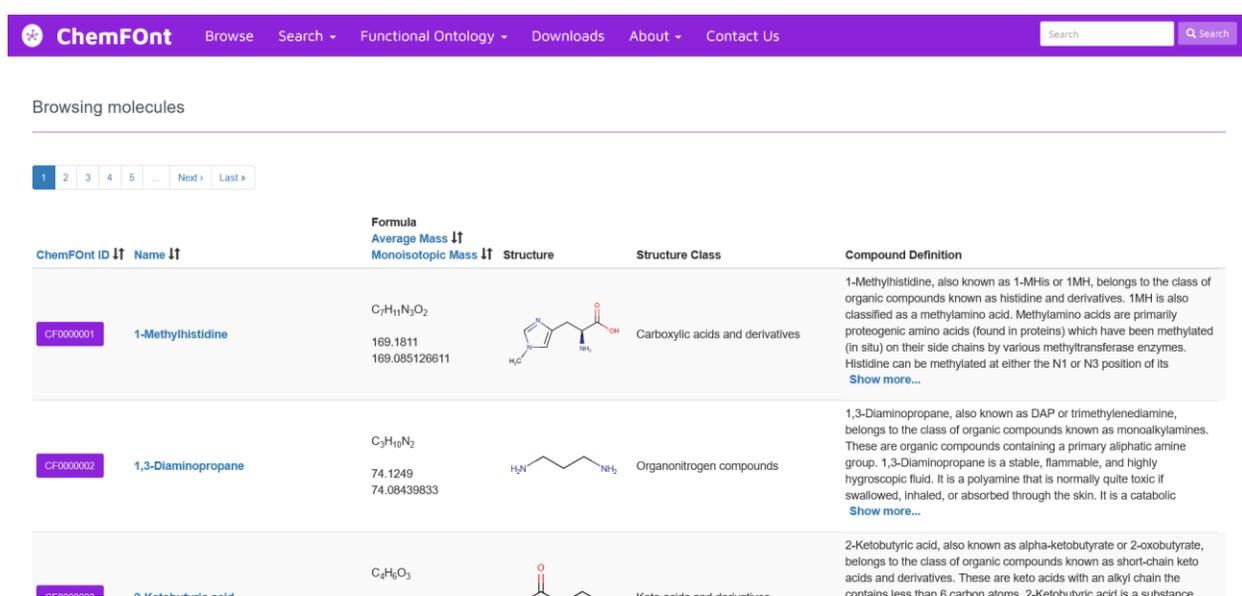
- If you entered “alanine” in the “Search” box, 70 results are returned in which “alanine” is found in the name or definition.



The screenshot shows the ChemFOnT website interface. At the top, there is a purple navigation bar with the ChemFOnT logo and menu items: Browse, Search, Functional Ontology, Downloads, About, and Contact Us. A search box in the top right corner contains the text "alanine" and a purple "Search" button. Below the navigation bar, the page displays "Search Results for molecules". A green banner indicates "Searching molecules for alanine returned 70 results." Below this, there are pagination controls showing "1 2 3 Next > Last >". The first search result is for "N-Palmitoyl Alanine" with a matched name "... N-Palmitoyl Alanine ...". The result includes a purple ID tag "CF0098640", a chemical structure of N-palmitoyl-L-alanine, and a smaller chemical structure of L-alanine below it.

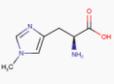
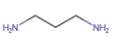
4. “Browse” option. Clicking on the “Browse” option at the top left of the homepage will allow you to browse either the chemicals or the ontology.

a. Clicking “Browse Chemicals will bring you to the Molecule Table, showing chemicals in order of their ChemFOnT ID number. Clicking the “Browse Data” hyperlink bar in the centre of the homepage will bring you to the same Molecule Table.



Browsing molecules

1 2 3 4 5 ... Next > Last >

ChemFOnT ID	Name	Formula	Average Mass	Monoisotopic Mass	Structure	Structure Class	Compound Definition
CF0000001	1-Methylhistidine	C ₇ H ₁₁ N ₃ O ₂	169.1811	169.085126611		Carboxylic acids and derivatives	1-Methylhistidine, also known as 1-MHis or 1MH, belongs to the class of organic compounds known as histidine and derivatives. 1MH is also classified as a methylamino acid. Methylamino acids are primarily proteogenic amino acids (found in proteins) which have been methylated (in situ) on their side chains by various methyltransferase enzymes. Histidine can be methylated at either the N1 or N3 position of its Show more...
CF0000002	1,3-Diaminopropane	C ₃ H ₁₀ N ₂	74.1249	74.08439833		Organonitrogen compounds	1,3-Diaminopropane, also known as DAP or trimethylenediamine, belongs to the class of organic compounds known as monoalkylamines. These are organic compounds containing a primary aliphatic amine group. 1,3-Diaminopropane is a stable, flammable, and highly hygroscopic fluid. It is a polyamine that is normally quite toxic if swallowed, inhaled, or absorbed through the skin. It is a catabolic Show more...
CF0000003	2-Ketobutyric acid	C ₄ H ₆ O ₃				Keto acids and derivatives	2-Ketobutyric acid, also known as alpha-ketobutyrate or 2-oxobutyrate, belongs to the class of organic compounds known as short-chain keto acids and derivatives. These are keto acids with an alkyl chain the contains less than 6 carbon atoms. 2-Ketobutyric acid is a substance Show more...

- The Molecule Table is sortable by selecting a column header in blue text, namely one of the ChemFOnT ID, Name, Average Mass, or Monoisotopic Mass. Other fields shown are the Formula, Structure, Structure Class, and Compound Definition.

➔ Next is an example result from the Molecule Table.

- Clicking on the ChemFOnT ID button or the compound name will take you to the ChemFOnT Molecule Card for that compound. The top portion of the ChemFOnT Molecule Card for 1-methylhistidine (CF0000001) is shown below.
- Each ChemFOnT Molecule Card contains seven major data sections, each with various fields: 1) ChemFOnT Record Information; 2) Molecule Identification; 3) Chemical Taxonomy; 4) Functional Ontology; 5) Physical Properties; 6) External Links; and 7) References. The first four sections are open (expanded) by default.
- For any collapsed section, clicking on the down arrow on the righthand side expands it and reveals further information about that molecule. Clicking on the up arrow for any expanded section collapses that section.

ChemFOnT Browse Search Functional Ontology Downloads About Contact Us Search

Showing molecule card for 1-Methylhistidine (CF0000001)

Record Information	
Version	1.0
Creation Date	2022-06-27 23:05:00 UTC
Update Date	2022-06-30 20:34:58 UTC
Chemfont ID	CF0000001

Molecule Identification	
Common Name	1-Methylhistidine
Definition	1-Methylhistidine, also known as 1-MHIs or 1MH, belongs to the class of organic compounds known as histidine and derivatives. 1MH is also classified as a methylamino acid. Methylamino acids are primarily proteogenic amino acids (found in proteins) which have been methylated (in situ) on their side chains by various methyltransferase enzymes. Histidine can be methylated at either the N1 or N3 position of its imidazole ring, yielding the isomers 1-methylhistidine (1MH; also referred to as pi-methylhistidine) or 3-methylhistidine (3MH; tau-methylhistidine), respectively. There is considerable confusion with regard to the nomenclature of the methylated nitrogen atoms on the imidazole ring of histidine and other histidine-containing peptides such as aserine. In particular, older literature (mostly prior to the year 2000) designated aserine (Npi methylated) as beta-alanyl-N1-methyl-histidine, whereas according to standard IUPAC nomenclature, aserine is correctly named as beta-alanyl-N3-methyl-histidine. As a result, many papers published prior to the year 2000 incorrectly identified 1MH as a specific marker for dietary consumption or various pathophysiological effects when they really were referring to 3MH (PMID: 24137022). Recent discoveries have shown that 1MH is produced in essentially all mammals (and other vertebrates) via the enzyme known as METTL9 (PMID: 33563959). METTL9 is a broad-specificity methyltransferase that mediates the formation of the majority of 1MH present in mammalian proteomes. METTL9-catalyzed methylation requires a His-x-His (HxH) motif, where "x" is a small amino acid. This HxH motif is found in a number of abundant mammalian proteins such as ARMC6, S100A9, and NDUFB3 (PMID: 33563959). Because of its abundance in many muscle-related proteins, 1MH has been found to be a good biomarker for the consumption of meat (PMID: 21527577). Dietary studies have shown that poultry consumption (p-trend = 0.0006) and chicken consumption (p-trend = 0.0003) are associated with increased levels of 1MH in human plasma (PMID: 30018457). The consumption of fish, especially salmon and cod, has also been shown to increase the levels of 1MH in serum and urine (PMID: 31401679). As a general rule, urinary 1MH is associated with white meat intake (p < 0.001), whereas urinary 3MH is associated with red meat intake (p < 0.001) (PMID: 34091671).
Structure	

4.

b. Clicking “Browse Ontology” will bring you to the Ontology Viewer, showing the ontology in a hierarchical format. Clicking the “Browse Ontology” hyperlink bar in the centre of the homepage will bring you to the same Ontology Viewer.

The screenshot displays the ChemFont Ontology Viewer interface. On the left, a hierarchical tree of ontology terms is shown, with 'Chemical reaction' highlighted in blue and enclosed in a red box. The main content area on the right is titled 'Chemical reaction' and provides a definition: 'A process in which one or more substances, the reactants, are converted to another or several other different substances, the products (Encyclopedia Britannica)'. Below this, a table lists four chemical reactions with their ChemFont IDs, names, structures, and structure classes.

ChemFont ID	Name	Structure	Structure Class
CF-00002309	Cer(d18:1/16:0)		Sphingolipids
CF-00002310	Cer(d18:1/18:0)		Sphingolipids
CF-00001646	2,4,7-Decatrienoic acid		Fatty Acyls
CF-00002311	Cer(d18:1/20:0)		Sphingolipids

- The ontology can be browsed by clicking on any of the terms in the top left of the page. Clicking on one of these terms will show the child terms, as well as any chemicals belonging to the term.
- Clicking on a leaf node of the ontology browse will allow you to see any chemicals that belong to this term.

Chemical reaction

ChemFont ID: CF000000111
Definition: A process in which one or more substances, the reactants, are converted to another or several other different substances, the products (Encyclopedia Britannica)
Definition ID: CF000001211

ChemFont ID	Name	Structure	Structure Class
CF-00002209	Cer(d18:1/16:0)		Sphingolipids
CF-00002210	Cer(d18:1/18:0)		Sphingolipids
CF-000016246	2,4,7-Decatrienoic acid		Fatty Acyls
CF-00002211	Cer(d18:1/20:0)		Sphingolipids

Showing 1 to 4 of 4 entries

- Term information, including ChemFont ID, Definition and Definition ID can be found at the top of the page for any ChemFont term. Additionally, by clicking on either the ChemFont ID box or the Name, you can view the ChemFont Molecule Card.

Chemical reaction

ChemFont ID: CF000000111
Definition: A process in which one or more substances, the reactants, are converted to another or several other different substances, the products (Encyclopedia Britannica)
Definition ID: CF000001211

ChemFont ID	Name	Structure	Structure Class
CF-00002209	Cer(d18:1/16:0)		Sphingolipids
CF-00002210	Cer(d18:1/18:0)		Sphingolipids
CF-000016246	2,4,7-Decatrienoic acid		Fatty Acyls
CF-00002211	Cer(d18:1/20:0)		Sphingolipids

Showing 1 to 4 of 4 entries

- As with Compounds, ChemFont terms can be searched for via the top right search box.

5. “Search” options. Clicking on the “Search” option at the top left of the homepage reveals a pulldown menu with three options including:

- Advanced Search (advanced search based on the textual fields of a Molecule Card);
- ChemQuery Structure Search (to search for a molecule by its chemical structure, whether by using its actual drawn structure or inserting its SMILES or InChI); and
- Text Query (the same search performed as using the search box in the top right of the homepage or the “Search Data” hyperlink bar in the centre of the homepage).



The website has provided online instructions for the Advanced Search and Text Query that are easy to follow.

➔ As the “ChemQuery Structure Search” is not straightforward, the following is an example to demonstrate how it is used.

For ChemQuery searches: The MarvinView applet from ChemAxon allows users to interactively draw structures or paste InChI or SMILES strings onto a palette. Click on the purple “Search” button and the MarvinView applet will be displayed.

ChemQuery Search by structure

Structure Search

Search Options

Similarity
 Substructure
 Exact

Similarity threshold: 0.7

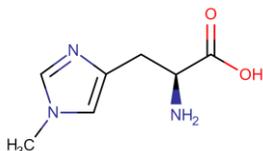
Molecular Weight Filter: e.g. 100 to e.g. 250

Maximum Results: 100

Search

Marvin JS
by Chemaxon

- If you were searching for the compound 1-methylhistidine, using the interactive palette,
- you could draw its chemical structure as below;



- insert its InChI (InChI=1S/C7H11N3O2/c1-10-3-5(9-4-10)2-6(8)7(11)12/h3-4,6H,2,8H2,1H3,(H,11,12)/t6-/m0/s1); or
- insert one of its SMILES strings (e.g., CN1C=NC(C[C@H](N)C(O)=O)=C1).
- Inserting the SMILES string “CN1C=NC(C[C@H](N)C(O)=O)=C1” (without the quotation marks) generates the below chemical structure as shown within the applet.

ChemFont Browse Search Functional Ontology Downloads About Contact Us

Structure Search

Search Options

Similarity Substructure Exact

Similarity threshold: 0.7

Molecular Weight Filter: e.g. 100 to e.g. 250

Maximum Results: 100

Search

MarvinJS Tutorials

- Selections for a narrowed search are also provided. Selecting the “Similarity” match option, then clicking the “Search” button returns these results.

Results 1 — 8 of approximately 8 results

← Previous

CF000000015 Score: 1.0		<p>Formula: C₇H₁₁N₃O₂</p> <p>Monoisotopic mass: 169.0851</p> <p>Molecular weight: 169.1811</p>
CF000015370 Score: 0.849		<p>Formula: C₆H₉N₃O₂</p> <p>Monoisotopic mass: 155.0695</p> <p>Molecular weight: 155.1546</p>
CF000005855 Score: 0.628		<p>Formula: C₇H₁₁N₃O₂</p> <p>Monoisotopic mass: 169.0851</p> <p>Molecular weight: 169.1811</p>
CF000081822 Score: 0.774		<p>Formula: C₉H₁₃N₃O₃</p> <p>Monoisotopic mass: 211.0957</p> <p>Molecular weight: 211.2210</p>