

Chemical Structure/Substructure Searching of Patents

A Mash-up of ChemSpider and the SureChem Patent Database

Introduction

ChemSpider offers the integration to a multitude of databases and recently has been integrated to the SureChem patent database. SureChem offers users the ability to perform online searches of the US, European and WO full text patents collections via text and chemical structure. ChemSpider has indexed the chemical structures associated with all patents on the SureChem patent database and provides a direct link between the services. This technical note will detail how to perform searches on ChemSpider and access related patent data. Users are recommended to read the [Manual](#) for deeper explanations of each of the points described here.

Accessing Searches

To access ChemSpider enter www.chemspider.com into the navigation entry box of an Internet Browser. The Home Page offers the route by which to navigate across the majority of the website. The Navigation interface is illustrated below with descriptive annotations.



Click on the Search button to access searching capabilities of the ChemSpider database

Performing Text-Based Searches

The most basic search screen is as shown below:

Search a database of >16 million chemical structures
Provide Feedback About How You Use ChemSpider at Our Blog

[View Manual](#) [ChemSpider Search Plugin](#) [Google Toolbar Button](#)

Systematic Name, Synonym, Trade Name,
Registry Number, SMILES or InChi 



Structures can be searched by inputting a chemical name, systematic name, SMILES string, InChI string or registry number. The identifiers are described below.

Identifier Type
Systematic Name – this includes two primary types of systematic nomenclature – IUPAC names and Index (CAS-type) names.
Trivial Name – this allows searching by one or more trivial or trade names associated with a chemical structure.
Registry Numbers – search on various registry numbers associated with chemical structures
SMILES strings – the SMILES strings listed in the database are those generated using OpenBabel software
InChI strings – InChI strings are canonicalized by default. Since all InChI strings utilize the InChI DLLs provided by IUPAC searching should invoke a simple copy and paste into the search box.

Some example identifiers for the same molecule are shown below:

Systematic names: **8-chloro-1-methyl-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepine**

Trivial names: **Xanax**

Registry numbers: **249-349-2 (EINECS Number), 28981-97-7**

SMILES codes: **C1c2cc3/C(=N\Cc1nnc(C)n1c3cc2)c4ccccc4**

InChI strings: **InChI=1/C17H13ClN4/c1-11-20-21-16-10-19-17(12-5-3-2-4-6-12)14-9-13(18)7-8-15(14)22(11)16/h2-9H,10H2,1H3**

Typing in any of these identifiers will provide a results screen. An example results screen is shown below.

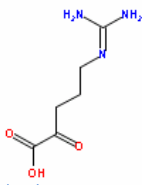
Structure	Molecular Formula	Molecular Weight	Monoisotopic Mass, Da	ACD/LogP	ACD/LogD (pH 5.5)	ACD/LogD (pH 7.4)	#H bond acceptors	#H bond donors	#Freely Rotating Bonds	Polar Surface Area	Molar Volume, cm3	
	C6H4Cl2	181.45	179.93	3.818	3.82	3.82	0	0	0	0	125.2	
	C19H30O2	290.44	290.2246	3.751	3.75	3.75	0	2	1	1	26.3	267.6
	C6H5ClO2	181.02	179.9745	3.761	3.76	3.76	0	2	2	2	18.46	115.2
	C19H28O2	288.42	288.2089	3.421	3.42	3.42	0	2	1	1	26.3	256.9
	C20H32O4	336.47	336.2301	3.895	2.99	1.19	0	4	3	16	44.76	323.5



Technical Note

Click on any of the structures of interest to expand to the screen of interest to provide a single record as shown below.

INHERENT PROPERTIES, IDENTIFIERS AND REFERENCES



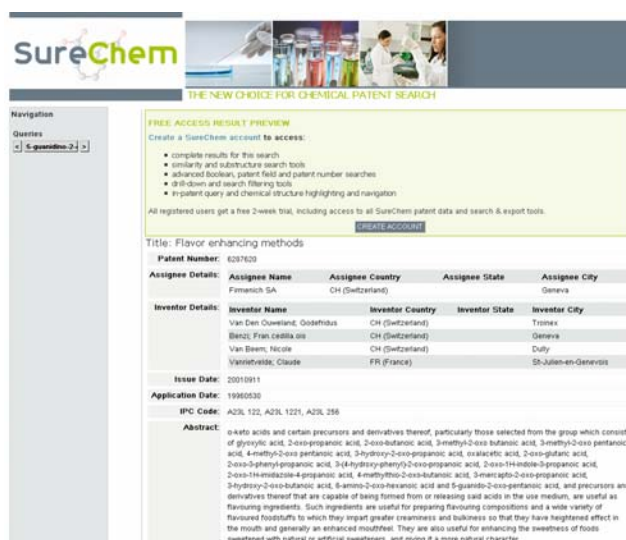
ChemSpider ID: 542
Empirical Formula: C₆H₁₁N₃O₃
Molecular Weight: 173.1698
Nominal Mass: 173 Da
Average Mass: 173.1698 Da
Monoisotopic Mass: 173.080041 Da

SMILES: O=C(C(=O)O)CCC/N=C(\N)N
InChI: InChI=1/C6H11N3O3/c7-6(8)9-3-1-2-4(10)5(11)12/h1-3H2,(H,11,12)(H4,7,8,9)/f/h11H,7-8H2
Data Source(s): [CambridgeSoft Corporation \(1\)](#), [ChEBI \(1\)](#), [ChemDB \(1\)](#), [ChemIDplus \(1\)](#), [KEGG \(2\)](#), [NIH/NLM \(1\)](#), [SureChem \(1\)](#), [Thomson Pharma \(1\)](#)

If there are patents available these will be shown in the Data Sources by the name SureChem. The number in the parentheses indicates the number of patents containing that structure in the SureChem database. In the example above there is only one patent. Click on the SureChem name (or any of the other links in the Data Sources) and the links to data sources will be shown as indicated below.

Data Source	Number of References	External ID(s)
CambridgeSoft Corporation	1	5451
ChEBI	1	CHEBI:18253
ChemDB	1	6916250
ChemIDplus	1	003715104
KEGG	2	C03771, C05935
NIH/NLM	1	558
SureChem	1	6287620
Thomson Pharma	1	00417081

In order to view the patent simply click on the External Link ID, in this case the number [6287620](#). The patent is shown below. It may be necessary to create a trial account for free access.



SureChem
THE NEW CHOICE FOR CHEMICAL PATENT SEARCH

Navigation
 Queries
 1: 6287620

FREE ACCESS RESUME PREVIEW
 Create a SureChem account to access:

- complete results for this search
- similarity and substructure search tools
- advanced boolean, patent field and patent number searches
- diff-chem and search filtering tools
- in-patent query and chemical structure highlighting and navigation

All registered users get a free 2-week trial, including access to all SureChem patent data and search & export tools.

[CREATE ACCOUNT](#)

Title: Flavor enhancing methods
Patent Number: 6287620

Assignee Details:	Assignee Name	Assignee Country	Assignee State	Assignee City
	Firmenich SA	CH (Switzerland)		Geneva

Inventor Details:	Inventor Name	Inventor Country	Inventor State	Inventor City
	Van Den Ouweland, Godefridus	CH (Switzerland)		Trossen
	Benz, Fran Cédric	CH (Switzerland)		Geneva
	Van Beem, Nicole	CH (Switzerland)		Dully
	Vignérat, Claude	FR (France)		St-Julien-en-Genevois

Issue Date: 2010011
Application Date: 19980530
IPC Code: A23L 122, A23L 1221, A23L 258

Abstract: keto acids and certain precursors and derivatives thereof, particularly those selected from the group which consists of glyoxylic acid, 2-oxo-propanoic acid, 2-oxo-butanoic acid, 3-methyl-2-oxo-butanoic acid, 3-methyl-2-oxo-pentanoic acid, 4-methyl-2-oxo-pentanoic acid, 3-hydroxy-2-oxo-propanoic acid, oxalacetic acid, 2-oxo-glutaric acid, 2-oxo-5-phenyl-propanoic acid, 3-(4-hydroxy-phenyl)-2-oxo-propanoic acid, 2-oxo-1H-indole-3-propanoic acid, 2-oxo-1H-imidazole-4-propanoic acid, 4-methylthio-2-oxo-butanoic acid, 3-methylthio-2-oxo-propanoic acid, 3-hydroxy-2-oxo-butanoic acid, 6-amino-2-oxo-hexanoic acid and 5-guanido-2-oxo-pentanoic acid, and precursors and derivatives thereof that are capable of being formed from or releasing said acids in the use medium, are useful as flavouring ingredients. Such ingredients are useful for preparing flavouring compositions and a wide variety of flavoured foodstuffs to which they impart greater creaminess and bulkiness so that they have heightened effect in the mouth and generally an enhanced mouthfeel. They are also useful for enhancing the sweetness of foods sweetened with natural or artificial sweeteners, and giving it a more natural character.



To highlight the name searched in the patent article simply click on the Query field as shown below.



and the text will be highlighted as shown:

Claim(s):	What is claimed is: 1. A method to improve, enhance or modify gustative impressions related to texture, volume or fullness, as perceived upon tasting, of a flavoring composition or a food product selected from the group consisting of light products having a low fat or sugar content, meat or meat-flavored products, bouillons, vegetable or vegetable-flavored products, fruit-juices, lemonades, and orange-flavored products, which method comprises adding to said composition or product a flavor and/or taste perceptible amount of one or more active ingredients selected from the group consisting of (a) the α -keto acids glyoxylic, 3-methyl-2-oxo-butanoic, 3-methyl-2-oxo-pentanoic, 4-methyl-2-oxo-pentanoic, 3-hydroxy-2-oxo-propanoic, oxalacetic, 2-oxo-glutaric, 2-oxo-3-phenyl-propanoic, 3-(4-hydroxyphenyl)-2-oxo-propanoic, 2-oxo-1H-indole-3-propanoic, 2-oxo-1H-imidazole-4-propanoic, 4-methylthio-2-oxo-butanoic, 3-mercapto-2-oxo-propanoic, 3-hydroxy-2-oxo-butanoic, 6-amino-2-oxo-hexanoic and 5-guanidino-2-oxo-pentanoic acid , and (b) the edible salts of the said α -keto acids, or a mixture of at least one of the compounds specified under (a) and (b) with the corresponding amino acid or acids.
------------------	---

To see other chemical compounds in the patent that have been identified by SureChem, select one of the two 'Markup' options at the top of the document. "All Compounds" includes every chemical name identified by SureChem while "Compounds with Structures" only shows those chemical names that SureChem has identified and been able to convert to a chemical structures.

Search filters and data export options are available on the right side of the screen, allowing you to narrow search results and then export the related chemical structures to your desktop. For more information about SureChem, visit www.surechem.org.

Performing Structure-Based Searches

The details of structure and substructure searching are detailed in Section 2.3.3 of the ChemSpider Manual available at <http://www.chemspider.com/ChemSpiderManual.aspx>. Please review the manual for details about how to perform these searches and the linking out to patents that may be examined from the results list.

Conclusion

ChemSpider and SureChem integration provides online access to text and structure/substructure-based searching of both American and International Patent Databases. This is an ideal mash-up to further enable ChemSpider users to derive information regarding molecules of interest.