



Chemical Database Projects Delivered by RSC eScience

at the FDA Meeting “Development of a Freely Distributable Data System for the Registration of Substances”

Antony Williams

RSC eScience

- What was once just ChemSpider is much more...
 - ChemSpider Reactions
 - Chemicals Validation and Standardization Platform
 - Learn Chemistry Wiki
 - National Chemical Database Service
 - Open PHACTS
 - PharmaSea
 - Global Chemistry Hub

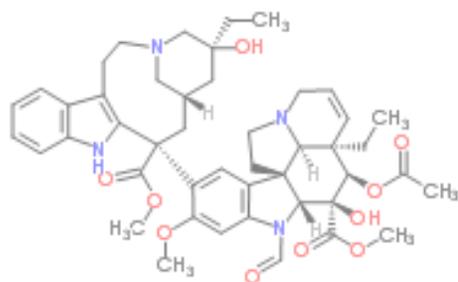
We are known for ChemSpider...

- ***The Free Chemical Database***
- A central hub for chemists to source information
 - >28 million unique chemical records
 - Aggregated from >400 data sources
 - Chemicals, spectra, CIF files, movies, images, podcasts, links to patents, publications, predictions
- A central hub for chemists to deposit & curate data

We Want to Answer Questions

- Questions a chemist might ask...
 - What is the melting point of n-heptanol?
 - What is the chemical structure of Xanax?
 - Chemically, what is phenolphthalein?
 - What are the stereocenters of cholesterol?
 - Where can I find publications about xylene?
 - What are the different trade names for Ketoconazole?
 - What is the NMR spectrum of Aspirin?
 - What are the safety handling issues for Thymol Blue?

I want to know about “Vincristine”



[?](#) 2D 3D Save Zoom

 - 9 of 9 defined stereocentres

Vincristine

ChemSpider ID: **5758**

Molecular Formula: $C_{46}H_{56}N_4O_{10}$

Monoisotopic mass: 824.399644 Da

▼ Systematic name

(2 α ,2' β ,3 α ,4 α ,5 β ,19 β)-22-oxovincal leukoblastine

► SMILES and InChIs

Wikibox

Embed

Deprecate

Watch this record

Manage data slice

Vincristine: Identifiers and Properties

Names and Identifiers

Names and Synonyms [?](#) Database ID(s)

Validated by Experts, Validated by Users, Non-Validated,

(2' β)-22-Oxovincal leukoblastine

200-318-1 [\[EINECS/ELINCS\]](#)

22-Oxovincal leukoblastine

57-22-7 [\[RN\]](#)

vincal leukoblastine, 22-oxo-

vincal leukoblastine, 22-oxo-, (2' β)-

vincal leukoblastine, 22-oxo-, (3 β ,4' β)-

Vincristine [\[Wiki\]](#)

Vincristinum [\[Latin\]](#)

VIN

Properties

Experimental data Predicted - ACD/Labs Predicted - ChemAxon

Data supplied by datasources and users.

- Experimental Physchem Properties

[+](#) Melting Point: 218 - 220 C [?](#) [↗](#)

- Miscellaneous

Appearance: solid [?](#) [↗](#)

Stability: Stable, but may be heat sensitive. Incompatible with strong oxidizing agents. [?](#) [↗](#)

Toxicity: IVN-RAT LD50 1300 mg kg⁻¹, IPR-MUS LD50 5.2 mg kg⁻¹ [?](#) [↗](#)

Safety: Safety glasses, gloves, good ventilation. [?](#) [↗](#)

Vincristine: Vendors and Sources

▼ Chemical Vendors

Data Source	External ID(s)
Pharmten 	S-PTN3183
AvaChem Scientific 	1011, 1011B
AOKChem 	aokchem11421

▼ Data Sources

 Chemical Vendors	Biological Data	Publishers	Metabolism Data	Phys. Properties	Tox/Envir. Data
Natural Products	Data Aggregators	Safety Data	All Data Sources		

Data Source	External ID(s)
ChemBank	NCI60_026703
DiscoveryGate	5978
LeadScope 	LS-228
NIAID	002674
DrugBank	5978, APRD00495
Collaborative Drug Discovery	16374

Vincristine: Articles

▼ Articles



[Links & Reference](#) [RSC Journals](#) [RSC Books](#) [PubMed](#) [MeSH Literature](#) [Google Books](#)

Please [click here](#) to see full list of found articles.

- Liu GT. Optic pathway gliomas. , *J Neuroophthalmol* , Volume 31 , Issue 3 , 2011 Sep
- Soysal T. Thrombotic Thrombocytopenic Purpura After Prophylactic Cefuroxime Axetil Administered in Relation to a Liposuction Procedure. , *Aesthetic Plast Surg* , 2011 Aug 19
- Hoster E. Update on the molecular pathogenesis and clinical treatment of mantle cell lymphoma: report of the 10th annual conference of the European Mantle Cell Lymphoma Network. , *Leuk Lymphoma* , 2011 Aug 18
- North KN. Therapeutics for Childhood Neurofibromatosis Type 1 and Type 2. , *Curr Treat Options Neurol* , 2011 Aug 18
- Witzens-Harig M. Replacement of conventional doxorubicin by pegylated liposomal doxorubicin is a safe and effective alternative in the treatment of non-Hodgkin's lymphoma patients with cardiac risk factors. , *Ann Hematol* , 2011 Aug 18
- Kahl BS. VcR-CVAD induction chemotherapy followed by maintenance rituximab in mantle cell lymphoma: a Wisconsin Oncology Network study. , *Br J Haematol* , 2011 Aug 16
- Geskin L. Clonal T-Cell Receptor {gamma}-Chain Gene Rearrangements in Differential Diagnosis of Lymphomatoid Papulosis From Skin Metastasis of Nodal Anaplastic Large-Cell Lymphoma. , *Arch Dermatol* , Volume 147 , Issue 8 , 2011 Aug
- Suzuki O. [A case of Sjogren syndrome coexistent with MALT lymphoma occurring along the parotid gland and trachea]. , *Nihon Kokyuki Gakkai Zasshi* , Volume 49 , Issue 7 , 2011 Jul
- Goldman S. Efficacy of interstitial continuous vincristine infusion in a bioluminescent rodent intracranial tumor model. , *J Neurooncol* , 2011 Aug 14
- Finkler NJ. Metastatic choriocarcinoma in a viable intrauterine pregnancy treated with EMA-CO in the third trimester: a case report. , *J Reprod Med* , Volume 56 , Issue 7-8 , 2011 Jul-Aug

How did we build it?

- We deal in Molfiles or SDF files – with coordinates
- Deposit anything that has an InChI – we support what InChI can handle, good and bad
- Standardization based on “InChI standardization”
- InChIs aggregate (certain) tautomers

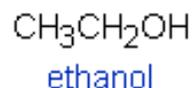
The InChI Identifier

International Chemical Identifier

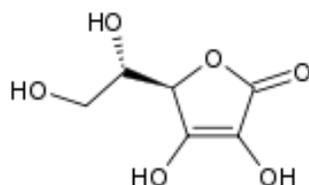
From Wikipedia, the free encyclopedia

(Redirected from [InChI](#))

The **IUPAC International Chemical Identifier (InChI)**, pronounced "INchee") is a textual [identifier](#) for [chemical substances](#), designed to provide a standard and human-readable way to encode molecular information and to facilitate the search for such information in databases and on the web. Developed by [IUPAC](#) and [NIST](#) during 2000-2005, the format and algorithms are non-proprietary and the software is freely available under the [open source LGPL](#) license (though the term "InChI" is a [trademark](#) of IUPAC).^[1]



InChI=1/C2H6O/c1-2-3/h3H,2H2,1H3



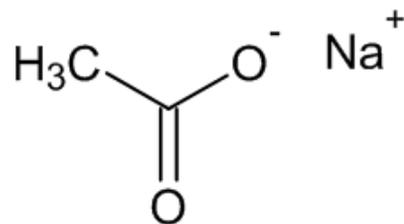
InChI=1/C6H8O6/c7-1-2(8)5-3(9)4(10)6(11)12-5/h2,5,7-10H,1H2/t2-,5+/m0/s1

L-ascorbic acid

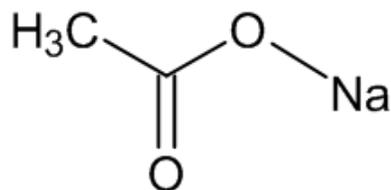
Downsides of InChI

- Good for small molecules – but no polymers, issues with inorganics, organometallics, imperfect stereochemistry. ChemSpider is “small molecules”
- InChI used as the “deduplicator” – **FIRST** version of a compound into the database becomes **THE** structure to deduplicate against...

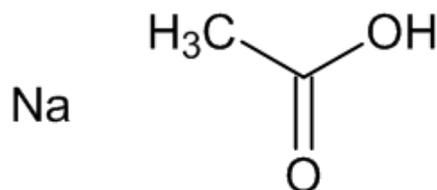
Side Effects of InChI Usage



InChI=1S/C2H4O2.Na/c1-2(3)4;/h1H3,(H,3,4);/q;+1/p-1
VMHLLURERBWHNL-UHFFFAOYSA-M

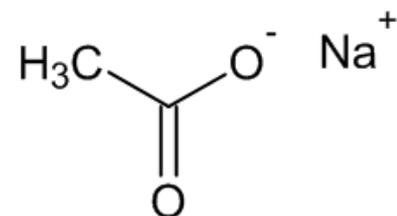


InChI=1S/C2H4O2.Na/c1-2(3)4;/h1H3,(H,3,4);/q;+1/p-1
VMHLLURERBWHNL-UHFFFAOYSA-M

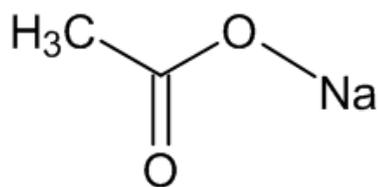


InChI=1S/C2H4O2.Na/c1-2(3)4;/h1H3,(H,3,4);
BDKZHNJTLHOSDW-UHFFFAOYSA-N

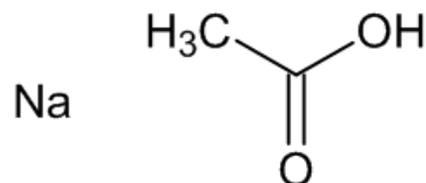
SMILES by comparison...



[Na+].[O-]C(C)=O

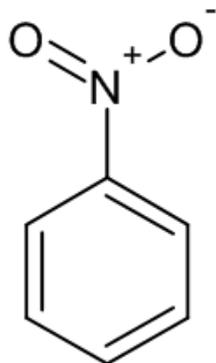


CC(=O)O[Na]



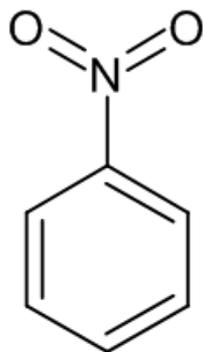
[Na].CC(=O)O

Side Effects of InChI Usage



InChI=1S/C6H5NO2/c8-7(9)6-4-2-1-3-5-6/h1-5H
LQNUZADURLCDLV-UHFFFAOYSA-N

[O-][N+](=O)c1ccccc1



InChI=1S/C6H5NO2/c8-7(9)6-4-2-1-3-5-6/h1-5H
LQNUZADURLCDLV-UHFFFAOYSA-N

O=N(=O)c1ccccc1

Searches: The INTERNET

▼ ChemSpider Searches

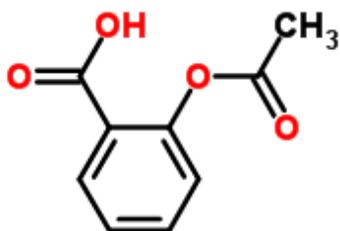
Search ChemSpider for:

- ⚙ Records with the same molecular formula
- ⚙ Compounds with the same skeleton
- ⚙ Use this molecule in a structure search

Search external sites for this structure:

-  Search Google Scholar (by synonym)
-  Search Google for exact structure
-  Search Google for structures with same skeleton

Search by InChI



Cell 2D 3D Save Zoom

Aspirin

ChemSpider ID: **2157**

Molecular Formula: $C_9H_8O_4$

Average mass: 180.157394 Da

Monoisotopic mass: 180.042252 Da

Systematic name

2-Acetoxybenzoic acid

SMILES and InChIs

SMILES:

CC(=O)Oc1ccccc1C(=O)O

Copy

Std. InChI:

InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12

Copy

Std. InChIKey:

BSYNRYMUTXBXSQ-UHFFFAOYSA-N

Copy

About 26,000 results (0.55 seconds)

Aspirin | **C9H8O4** | ChemSpider

www.chemspider.com/Chemical-Structure.2157.html
ChemSpider 2D Image | Aspirin | C9H8O4. Cell; 2D; 3D; Save ...
InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12) Copy
Copied ...

Aspirin

[webbook.nist.gov/.../inchi/InChI%3D1S/C9H8O4/c1-6\(10\)13-8-5-3-2-4-7\(8\)9\(11\)12/h2-5H,1H3,\(H,11,12\);](http://webbook.nist.gov/.../inchi/InChI%3D1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12);)
InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12); Download
the identifier in a file. INChI Trust 2011 Certified Logo ...

acetylsalicylic acid (CHEBI:15365)

www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:15365
Nov 1, 2011 - InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12); InChI=...

About 4,760 results (0.18 seconds)

java - JAX-RS

[stackoverflow.com](http://stackoverflow.com/...)
4 answers - Feb
The InChIName i

12/h2-5H,1H3,(F

Chemical Identifier Search | **BSYNRYMUTXBXSQ-WXRBYKJCCW**

www.chemspider.com/InChIKey=BSYNRYMUTXBXSQ-WXRBYKJ...
Search by Systematic name, Synonym, Trade name, Registry number, SMILES or
InChI. | **BSYNRYMUTXBXSQ-WXRBYKJCCW**.

"**BSYNRYMUTXBXSQ-UHFFFAOYSA-N**"[InChIKey] - PubChem ...

www.ncbi.nlm.nih.gov/pcsubstance?...%22BSYNRYMUTXBXSQ...
Source: chemicalize.org by ChemAxon (92) Compound ID: 2244 SID: 137001131 2.
Source: Cambridge Crystallographic Data Centre (ACSA01) Compound ...

Ligand View: ace

www.brenda-enzym...
Information on enzym

UHFFFAOYSA-N.

Aspirin

webbook.nist.gov/cg...
IUPAC Standard InC

Number: 50-78-2; C

About 4,390 results (0.12 seconds)

"**BSYNRYMUTXBXSQ-UHFFFAOYSA-N**"[InChIKey] - PubChem ...

www.ncbi.nlm.nih.gov/pcsubstance?...%22BSYNRYMUTXBXSQ...
Source: chemicalize.org by ChemAxon (92) Compound ID: 2244 SID: 137001131 2.
Source: Cambridge Crystallographic Data Centre (ACSA01) Compound ...

Aspirin - PubChem

pubchem.ncbi.nlm.nih.gov/... > Chemicals & Bioassays > PubChem
Molecular Formula: C9H8O4 Molecular Weight: 180.15742 InChIKey:
BSYNRYMUTXBXSQ-UHFFFAOYSA-N. The prototypical analgesic used in the
treatment ...

Aspirin - the NIST WebBook - National Institute of Standards and ...

webbook.nist.gov/cgi/cbook.cgi?ID=C50782&Type=THZ...
IUPAC Standard InChIKey: **BSYNRYMUTXBXSQ-UHFFFAOYSA-N**; CAS Registry
Number: 50-78-2; Chemical structure: C9H8O4 This structure is also available ...

Ligand View: acetylsalicylic acid - InchiKey: **BSYNRYMUTXBXSQ ...**

www.brenda-enzymes.org/php/ligand_flatfile.php4?brenda_ligand...
InChIKey: **BSYNRYMUTXBXSQ-UHFFFAOYSA-N** ... 3.1.1.2, induces enzyme
expression in the liver and stimulates enzymatic activity in blood plasma, 693517 ...

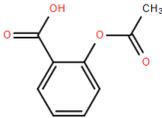
ChemSpider Google Search

<http://www.chemspider.com/google/>

Upload a structure file or image file

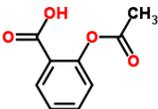
MOL, SDF, CDX, PNG, JPG, GIF

Convert to structure
aspirin
Name, SMILES, InChI or ChemSpider ID



...
A
H
C
N
O
S
F

Search



Aspirin

ChemSpider ID: [2157](#)
Molecular Formula: $C_9H_8O_4$
Average mass: 180.157394 Da
Monoisotopic mass: 180.042252 Da
Systematic name: 2-Acetoxybenzoic acid
Search this compound on [Google Scholar](#)

Web Patents Books Images

[Shelf stable aspirin solutions](#)
US Pat. 4975269 - Jul 31, 1989
There is provided a shelf stable solution of **aspirin**, free of moieties reactive with **aspirin**, suitable for topical application to the skin compounded from **aspirin**, ...

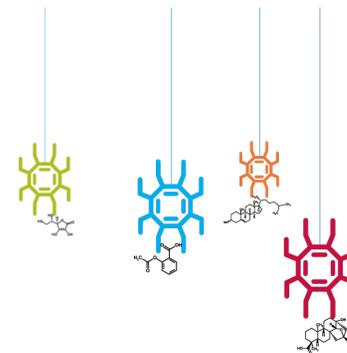
[Suppression of thromboxane levels by percutaneous administration of aspirin](#)
US Pat. 6071896 - Mar 18, 1998 - Gundersen Clinic, Ltd.
... a pharmaceutical composition containing **aspirin**. Articles ...

[Enteric coated aspirin tablets](#)
US Pat. 4857337 - May 24, 1988 - American Home Products Corp. (Del)
Enteric coated **aspirin** tablets rendered shock-insensitive by the provision of a protective coat of hydroxypropyl methylcellulose of at least about 1.5% by weight ...

How did we build it?

- We deal in Molfiles or SDF files – with coordinates
- Deposit anything that has an InChI – we support what InChI can handle, good and bad
- Standardization based on “InChI standardization”
- InChIs aggregate (certain) tautomers
- We deal with “various forms” of data

Crowdsourced “Annotations”



- Users can add
 - Descriptions/Syntheses/Commentaries
 - Links to PubMed articles
 - Links to articles via DOIs
 - Add spectral data
 - Add Crystallographic Information Files
 - Add photos
 - Add MP3 files
 - Add Videos

Description

Description

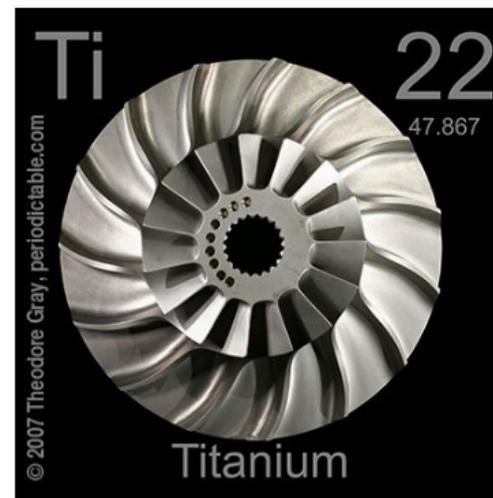
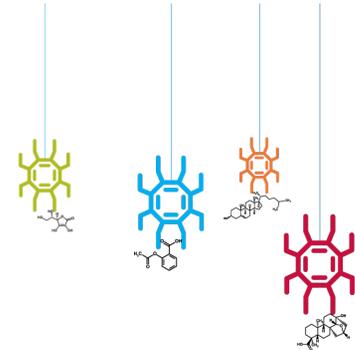


MP3 Download: Simon Cotton from Uppingham School in the UK talks about the cleaning chewing gum and dog mess off of our pavements.

Titanium - Periodic Table of Videos



YouTube



ChemSpider : Spectra Linked

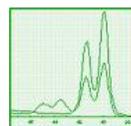
Type: HNMR

Associated Hyperlink: <http://rainier.chem.plu.edu/nutsform.html>

Comments: These data are obtained from the Pacific Lutheran University FTNMR FID Archive

Approved: No

Submitted by: [ChemSpiderman](#)



[OPEN DATA](#)

[Download](#)

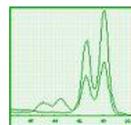
Type: CNMR

Associated Hyperlink: <http://rainier.chem.plu.edu/nutsform.html>

Comments: These data are obtained from the Pacific Lutheran University FTNMR FID Archive

Approved: No

Submitted by: [ChemSpiderman](#)



[OPEN DATA](#)

[Download](#)

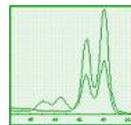
Type: Electron Impact

Associated Hyperlink: <http://webbook.nist.gov/cgi/cbook.cgi?ID=C57885&Units=SI&Mask=200#Mass-Spec>

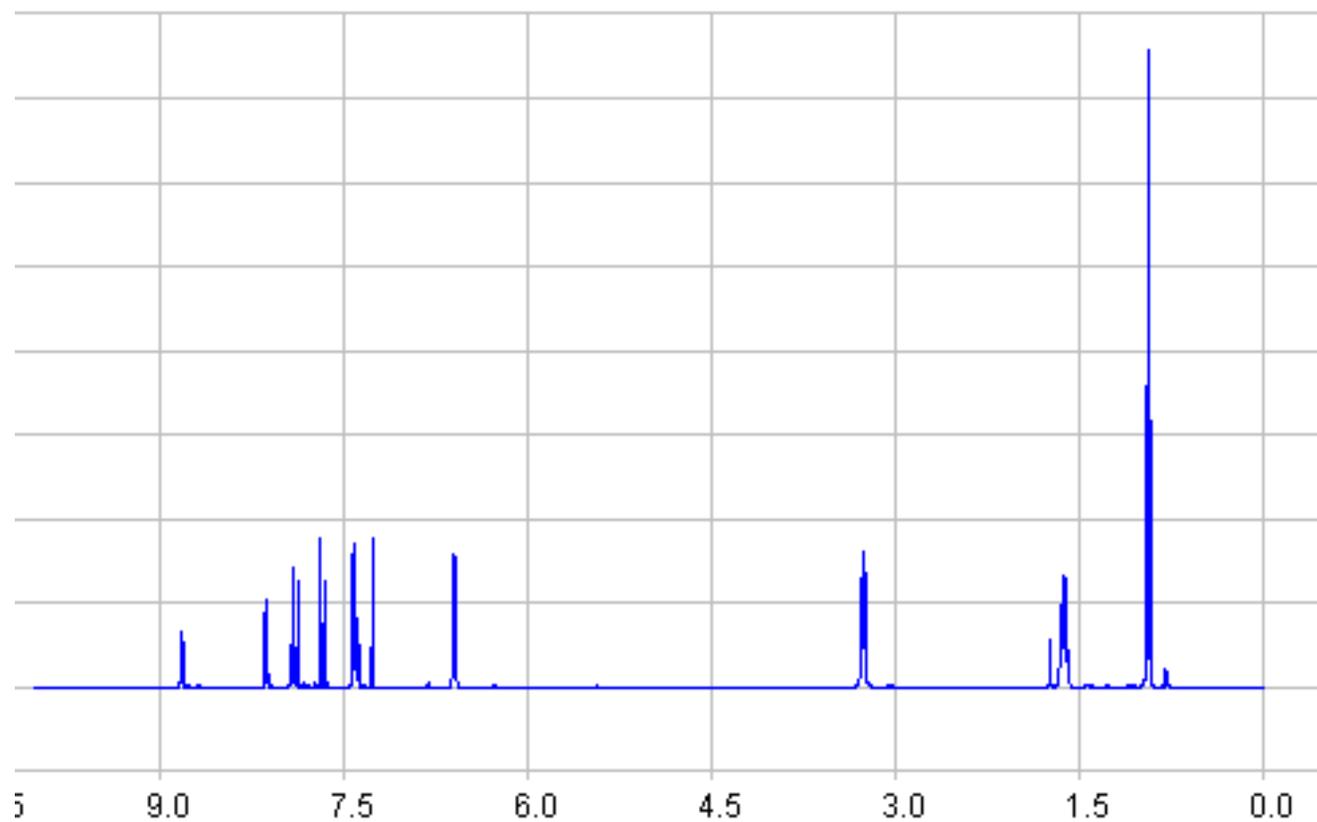
Comments: Provided with permission: March 12 2009. Owner NIST Mass Spectrometry Data Center Collection (C) 2007 copyright by the U.S. Secretary of Commerce on behalf of the United States of America. All rights reserved. Origin T.IIDA NIHON UNIVERSITY, KORIYAMA, FUKUSHIMA-KEN, JAPAN; NIST MS number 67286

Approved: No

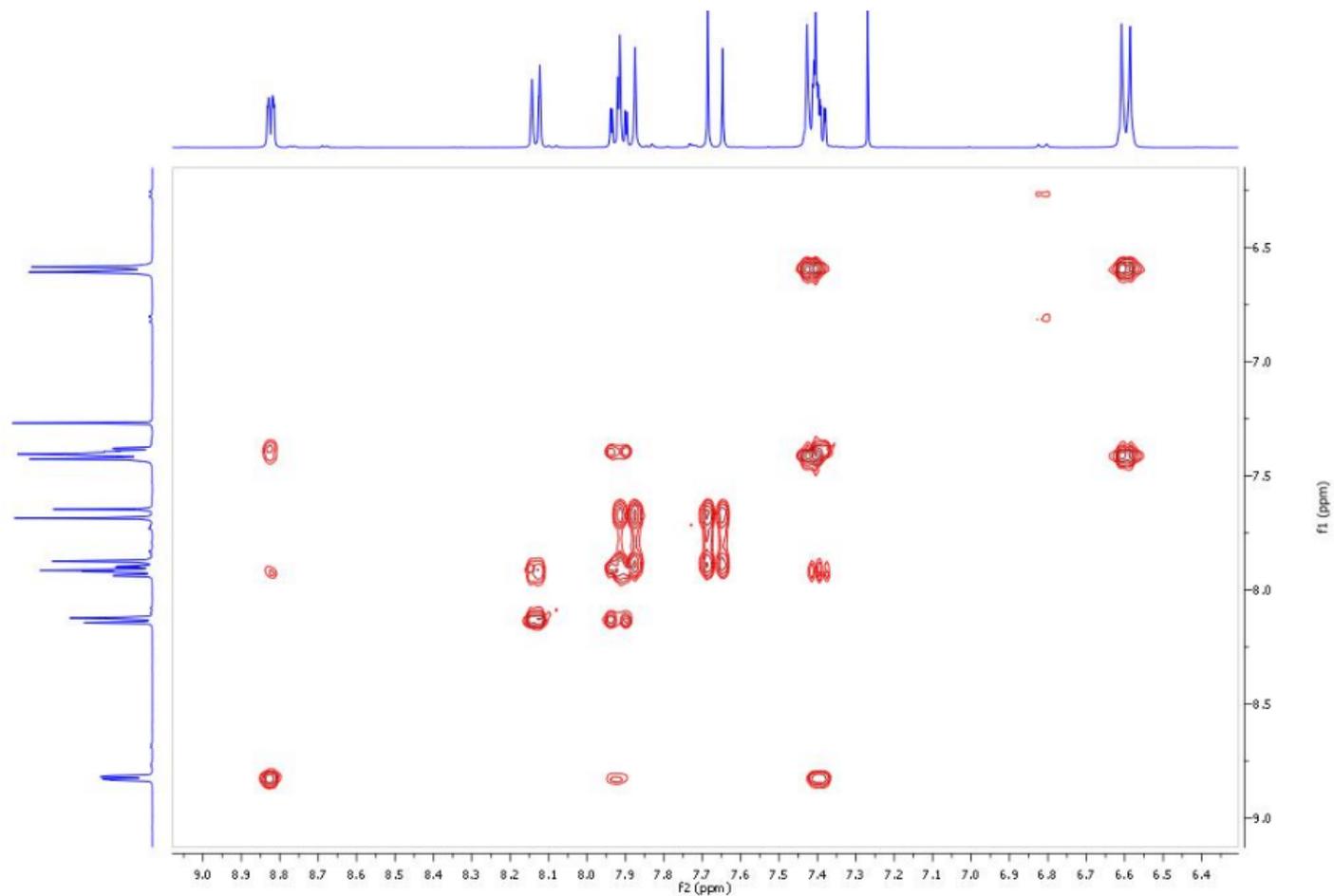
Submitted by: [ChemSpiderman](#)



ChemSpider ID 24528095 H1 NMR



ChemSpider ID 24528095 HHCOSY



How did we build it?

- We deal in Molfiles or SDF files – with coordinates
- Deposit anything that has an InChI – we support what InChI can handle, good and bad
- Standardization based on “InChI standardization”
- InChIs aggregate (certain) tautomers

- We deal with “various forms” of data
- **We are challenged with the complexities of chemical names**

Antony Williams vs Identifiers



Dad, Tony, others

5 email addresses
ChemSpiderman (blog,
Twitter account,
Facebook, Friendfeed)
OpenID

.....



Surname/Nom (1)
WILLIAMS
Given names/Prénoms(2)
ANTHONY JOHN
Nationality/Nationalité (3)

Passport ID



License



SSN



Green Card

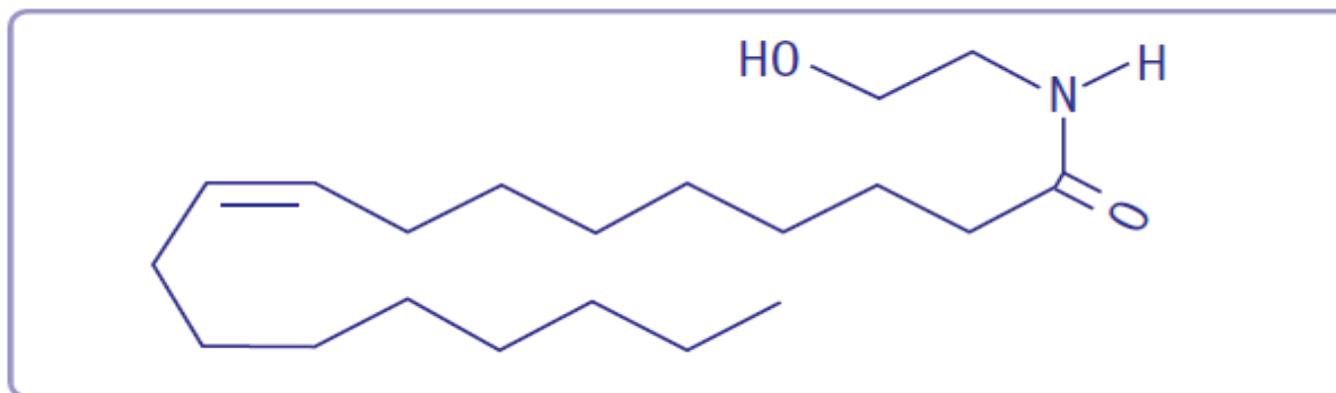
Aspirin names and synonyms

 Depositor-Supplied Synonyms: (Total: 335) 

aspirin 
Acylpyrin 
Colfarit 
Ecotrin 
Acetylsalicylate
ACETYLSALICYLIC ACID 
Enterosarein
Acenterine
Polopiryna 
Micristin 
Acetosal
Acetosalic acid
2-Acetoxybenzoic acid
Enterosarine
Acetophen
Acetosalin
Acetylsal
Aspirdrops
Bialpirina
Clariprin
Entericin
Enterophen

- Text searches depend on correct association
- **>300** suggested identifiers for Aspirin just on PubChem
- **Disambiguation** dictionaries are necessary, not just for authors!

What is a good strategy to search for chemical compounds? We compared various free online databases and search engines for N-oleoylethanolamine (OEA, CAS Registry Number 111-58-0), a very interesting compound and the Nature "lipid of the month" June 2009". OEA has been found to induce satiety and decrease meal frequency, and is therefore a potential therapeutic target for treatment of obesity, diabetes and eating disorders. OEA is also used in the treatment of psoriasis, due to its ceramidase inhibiting effects.

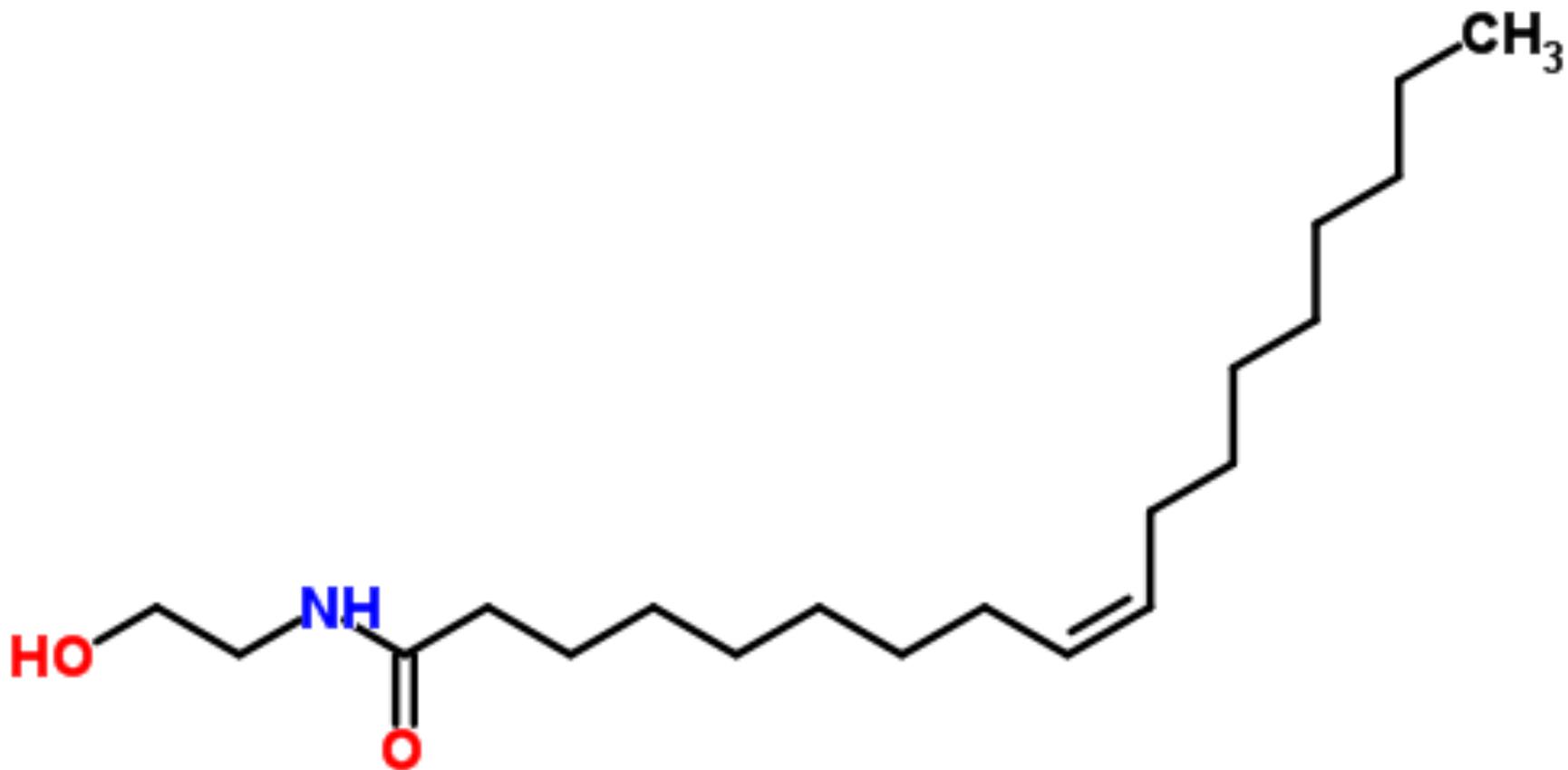


The Final Search Strategy

The final strategy for FreePatentsOnline:

TTL/(Oleylethanolamine OR Oleylethanolamide OR "Oleoyl ethanolamine" OR "N-oleoyl-ethanolamine" OR "oleoyl ethanolamide" OR "N-oleoyl ethanolamide" OR "N-oleoyl ethanolamine" OR "N-Oleoyl-2-aminoethanol" OR "N-Oleylethanolamine" OR Oleylethanolamide OR "Oleyl ethanolamide" OR "oleic acid ethanolamide" OR Oleylethanolamine OR "N-Oleylethanolamine" OR "Oleyl ethanolamine" OR "N-oleylethanolamine" OR "oleic acid ethanolamine") OR **ABST**/(Oleylethanolamine OR Oleylethanolamide OR "Oleoyl ethanolamine" OR "N-oleoyl-ethanolamine" OR "oleoyl ethanolamide" OR "N-oleoyl ethanolamide" OR "N-oleoyl ethanolamine" OR "N-Oleoyl-2-aminoethanol" OR "N-Oleylethanolamine" OR Oleylethanolamide OR "Oleyl ethanolamide" OR "oleic acid ethanolamide" OR Oleylethanolamine OR "N-Oleylethanolamine" OR "Oleyl ethanolamine" OR "N-oleylethanolamine" OR "oleic acid ethanolamine") OR **ACLM**/(Oleylethanolamine OR Oleylethanolamide OR "Oleoyl ethanolamine" OR "N-oleoyl-ethanolamine" OR "oleoyl ethanolamide" OR "N-oleoyl ethanolamide" OR "N-oleoyl ethanolamine" OR "N-Oleoyl-2-aminoethanol" OR "N-Oleylethanolamine" OR Oleylethanolamide OR "Oleyl ethanolamide" OR "oleic acid ethanolamide" OR Oleylethanolamine OR "N-Oleylethanolamine" OR "Oleyl ethanolamine" OR "N-oleylethanolamine" OR "oleic acid ethanolamine")

All Those Names, One Structure



Curated Dictionaries Matter

6,7-Dimethylumazine as a potential ligand for selective recognition of adenine onosite an abasic site in DNA duplexes†

Zhiqiang Ye[†], Burki Rajendar^{†§}, Dai Qing[†], Seiichi Nishizawa^{†¶} and Norio Teramae^{†¶*}

[†]Department of Chemistry, Graduate School of Science, Tohoku University, Aoba-ku, Sendai, 980-8578, Japan. [E-mail: teramae@mail.tains.tohoku.ac.jp](mailto:teramae@mail.tains.tohoku.ac.jp); Fax: +81 22 7956552; Tel: +81 22 7956549

[¶]CREST, Japan Science and Technology Agency (JST), Aoba-ku, Sendai, 980-8578, Japan

Received (in Cambridge, UK) 26th September 2008, Accepted 20th October 2008

First published

6,7-Dimethylumazine
X = AP site (S
the binding affi

Single nucleoti
Thus, simple a
research effort
have recently s
and they succe
other hand, we
fluorescence li
binders or inter
intrahelical nuc
containing DN
amiloride,⁸ and
developing an
1.21 × 10⁶ M⁻¹
adenine. Sever
stabilization of
increase the sta

Manuscript DOI 10.1039/b816876h Compound information for 2-amino-6,7-dimethyl-4-hydroxypteridine ...

http://www.rsc.org/delivert/_articlelinking/cheminfo.asp?MLID=3&compoundtext=2-amino-6,7-dimethyl-4-hydroxypteridine&MSID

RSC Publishing

Compound information '2-amino-6,7-dimethyl-4-hydroxypteridine'

Synonyms:

- 2-amino-6,7-dimethyl-4-hydroxypteridine

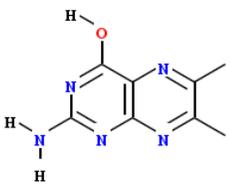
SMILES: OC1=NC(=NC2=NC(=C(N=C12)C)C)N

InChI: InChI=1C8H9N5O/c1-3-4(2)11-6-5(10-3)7(14)13-8(9)12-6/h1-2H3,(H3,9,11,12,13,14)/w14H,9H2

InChIKey: InChIKey=ZKWZUPPXTCQQJL-JPLXFSROCR

CML (Chemical Markup Language) Representation: [Download File](#)

2-D Representation:



Other resources:

- Search for this compound in PubChem
- Search for this compound in SureChem patents

Toolbox

Tools and Resources

- Print this article
- Download PDF
- Email a friend
- Supplementary information
- Prospect View FAQ
- Normal View
- Advanced features
- Find citing articles

Navigation

Hide compounds

Show Gold Book

GX GCA AC-3/3'-AGG TCN CGT TG-5',
1.0 μM; substituted methyl groups enhance
Highlight Terms

enes and detection of genetic mutations.^{1,2}
nacogenomics.³ Consequently, considerable
by fluorescent molecules.⁴ Nakatani *et al.*⁵
sine,^{2b} and guanine-adenine^{5c} mismatches,
plasmon resonance (SPR) assay. On the
have discovered a series of small
typical DNA-drug binding ligands (groove
pseudo-base pairing of ligands with
AP site. In combination with AP site-
⁶ 2-amino-6,7-dimethyl-4-hydroxypteridine,⁷
among these base-selective ligands,
mate A (the 1 : 1 binding constant, K_{11} =
selective detection of SNPs related to
ve the binding affinity of ligands and
s in a DNA strand has been known to

Crowdsourcing ChemSpider

- ChemSpider is crowdsourced
- Community deposition, annotation and curation
- Anyone can “Leave Feedback”
- Registered users can add data

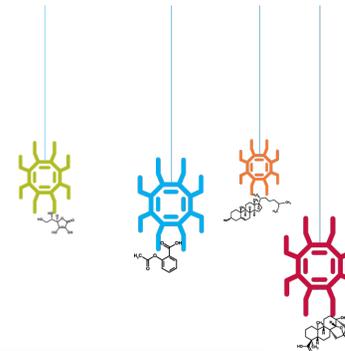
Want to comment
on this record?

Leave Feedback

Add data to this record

Identifier	Description
Image	Spectrum
CIF	Data source
Publication	DOI
PubMed ID	URL

“Curate” Identifiers



▼ Identifiers

Names and Synonyms

[Validated by Experts](#), [Validated by Users](#), [Non-Validated](#), [Removed by Users](#), [Redirected by Users](#), [Redirect Approved by Experts](#)

[Edit](#)

1H-thieno[3,4-d]imidazole-4-pentanoic acid, hexahydro-2-oxo-, (3aR,4S,6aR)-

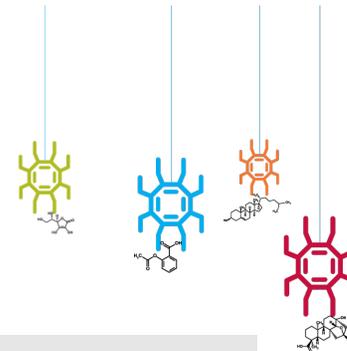
5-[(3aR,4S,6aR)-2-Oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl]pentanoic acid

58-85-5 [RN]

D(+)-Biotin >99.5%

VITAMIN H

“Curate” Identifiers



Names and Synonyms

Select all Deselect all Invert selection

Validated by Experts, Validated by Users, Non-Validated, Removed by Users, Redirected by Users, **Redirect Approved by Experts**

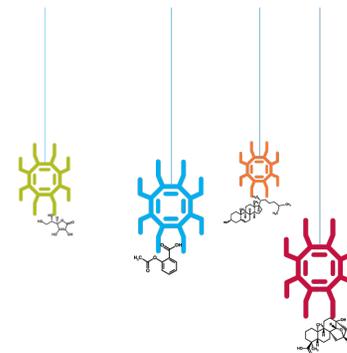
<input type="checkbox"/>	1H-thieno[3,4-d]imidazole-4-pentanoic acid, hexahydro-2-oxo-, (3aR,4S,6aR)-	<input type="button" value="Edit"/>	Antony Williams Antony Williams
<input type="checkbox"/>	5-[(3aR,4S,6aR)-2-Oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl]pentanoic acid	<input type="button" value="Edit"/>	Antony Williams Antony Williams
<input type="checkbox"/>	58-85-5[RN]	<input type="button" value="Edit"/>	Antony Williams
<input checked="" type="checkbox"/>	D(+)-Biotin >99.5%	<input type="button" value="Edit"/>	Antony Williams
<input checked="" type="checkbox"/>	VITAMIN H	<input type="button" value="Edit"/>	Antony Williams

Change Synonym(s) States ✕

Change State To

- Reject**
 - Deleted (approved as wrong)
 - Rejected
- Normal**
 - Normal
- Confirm**
 - Confirmed
- Approve**
 - Approved
 - Common Name
- Redirect**
 - Redirected
 - Redirect Approved

“Curate” Identifiers



Validated by Experts, Validated by Users, Non-Validated

1H-thieno[3,4-d]imidazole-4-pentanoic acid,

5-[(3aR,4S,6aR)-2-Oxohexahydro-1H-thieno[

58-85-5[RN]

~~D(+)-Biotin >99.5%~~

~~VITAMIN H~~

Vincristine: Identifiers and Properties

▼ Names and Identifiers

Names and Synonyms  Database ID(s)

Validated by Experts, Validated by Users, Non-Validated,

(2' β)-22-Oxovincal leukoblastine

200-318-1 [[EINECS/ELINGS](#)]

22-Oxovincal leukoblastine

57-22-7 [[RN](#)]

vincal leukoblastine, 22-oxo-

vincal leukoblastine, 22-oxo-, (2' β)-

vincal leukoblastine, 22-oxo-, (3 β ,4' β)-

Vincristine [[Wiki](#)]

Vincristinum [[Latin](#)]

VIN

Vincristine: Patents

Linked by Name

▼ Patents

Google Patents

USPTO Granted

USPTO Applications

European Granted

European Applications

WO/PCT

Japanese Abstracts

powered by Google™



[Vincristine-containing product](#)

US Pat. 5055449 - Nov 07, 1988 - Akzo N.V.

Vincristine is a mitosis-retarding Vinca alkaloid which is fairly effective in combating various types of cancer, particularly Hodgkins disease and other ...



[Method of preparing vincristine](#)

US Pat. 4375432 - Nov 16, 1981 - Eli Lilly and Company

METHOD OF PREPARING VINCRISTINE CROSS-REFERENCE This application is a continuation-in-part of my copending application Ser. No. 262836 filed May 12, 1981, ...



[Process for the preparation of vincristine](#)

US Pat. 4767855 - Jun 02, 1986 - P. F. Medicament

4767855 25 PROCESS FOR THE PREPARATION OF VINCRISTINE The present invention relates to a particularly efficient process for the preparation of ...



[Method of preparing vincristine](#)

US Pat. 4303584 - Apr 02, 1980 - Eli Lilly and Company

Primary Examiner — Donald G. Daus Assistant Examiner — Diana G. Rivers Attorney, Agent, or Firm — James L. Rowe; Arthur R. Whale [57] **ABSTRACT Vincristine** ...



[Synthetic vinblastine and vincristine derivatives](#)

US Pat. 4144237 - Jun 13, 1977 - The United States of America as represented by the Department of Health, Education and Welfare 13, 1979 [54] **SYNTHETIC VINBLASTINE AND VINCRISTINE DERIVATIVES** [75] Inventor: James P. Kutney, Vancouver, Canada [73] Assignee: The United States of ...

Validated Names for Searching...

Google scholar

"Vincristine" OR "Vincristinum" OR "57-22-7" OR

Search

[Advanced Scholar Search](#)

Scholar

Articles excluding patents

anytime

include citations



[Create email alert](#) Results 1 - 10 of about 196,000. (0.46 s)

Did you mean: ["Vincristine" OR "Vincristin" OR "57-22-7" OR "22-Oxovincalcoblastine" OR "vincalcoblastine, 22-oxo-" OR "\(2'beta\)-22-Oxovincalcoblastine" OR "vincalcoblastine, 22-oxo-, \(2'beta\)-"](#)

[Overcoming of vincristine resistance in P388 leukemia in vivo and in vitro through enhanced cytotoxicity of vincristine and vinblastine by verapamil](#)

[\[PDF\]](#) from [aacrjourn](#)

T Tsuruo, H Iida, S Tsukagoshi... - *Cancer research*, 1981 - AACR

Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, Toshima-ku, Tokyo 170, Japan ... A noncytotoxic dose of verapamil, a coronary vasodilator, enhances the cytotoxicity of Vincristine (VCR) and vinblastine in P388 leukemia and its VCR-resistant subline, ...

[Cited by 994](#) - [Related articles](#) - [All 2 versions](#)

[Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone \(CHOP\) significantly improves the outcome ...](#)

[\[HTML\]](#) from [hematol](#)

W Hiddemann, M Kneba, M Dreyling... - ..., 2005 - [bloodjournal.hematologylibrary.org](#)

From the Department of Internal Medicine III, University of Munich, Klinikum Großhadern; Department of Internal Medicine II, University Hospital Schleswig-Holstein, Campus Kiel; Department of Hematology and Oncology, Hospital St Georg, Hamburg; Department of Internal Medicine III, Klinikum ...

[Cited by 528](#) - [Related articles](#) - [BL Direct](#) - [All 8 versions](#)

[Increased accumulation of vincristine and adriamycin in drug-resistant P388 tumor cells following incubation with calcium antagonists and calmodulin inhibitors](#)

[\[PDF\]](#) from [aacrjourn](#)

T Tsuruo, H Iida, S Tsukagoshi... - *Cancer Research*, 1982 - AACR

Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, Toshima-ku, Tokyo 170, Japan ... Some calcium antagonists and calmodulin inhibitors enhance the intracellular levels of vincristine and Adriamycin in vincristine- and Adriamycin-resistant P388 ...

[Cited by 459](#) - [Related articles](#) - [All 2 versions](#)

And yes..there are challenges

<http://www.cas.org/legal/infopolicy>

Egon asks: What is dapagliflozin? Check in ChemSpider.... »

31
08
2007

How Many Electronic Databases Have More Than 10000 CAS Numbers?

Posted by: Antony Williams in [ChemSpider Chemistry](#)

Copyright©2007 Antony Williams

I have been invited to write an article regarding Open Access Chemistry Databases and am in the process of gathering information. During one of my google searches I happened across a statement I was aware of but had forgotten until recently. It relates to the ability to use CAS numbers on a website. Specifically, from the [CAS Information Use Policies](#) of 2005 it says, quote:

- i. A User or Organization may include, without a license and without paying a fee, up to 10,000 CAS Registry Numbers or CASRNs in a catalog, website, or other product for which there is no charge. **The following attribution should be referenced or appear with the use of each CASRN: CAS Registry Number is a Registered Trademark of the American Chemical Society.** CAS recommends the verification of the CASRNs through CAS Client ServicesSM.

Licensing Data is Tough...

PLOS | COMPUTATIONAL BIOLOGY

Browse | For Authors | About Us

Search  advanced search

 OPEN ACCESS

PERSPECTIVE

3,181	1	26	103
VIEWS	CITATION	ACADEMIC BOOKMARKS	SOCIAL SHARES

Why Open Drug Discovery Needs Four Simple Rules for Licensing Data and Models

Data Licensing

Access to the web interface of ChEMBL is made under the EBI's [Terms of Use](#). The ChEMBL data is made available on a [Creative Commons Attribution-Share Alike 3.0 Unported License](#).

DrugBank is offered to the public as a freely available resource. Use and re-distribution of the data, in whole or in part, for commercial purposes requires explicit permission of the authors and explicit acknowledgment of the source material (DrugBank) and the original publication (see below). We ask that users who download significant portions of the database cite the DrugBank paper in any resulting publications.

Data Licensing, Open Data

- The use of CAS data in third party Data Mining Tools is permitted as long as CAS Records are downloaded via STN[®] AnaVist[™]. All of these new "**freedoms**" are aimed at further enabling the dissemination of scientific information and the advancement of scientific research.
- **CAS does not permit the building of Databases that have wide and general availability** and no longer fulfill the purpose of individual or team research that CAS permits but instead serve as a substitute for the use of CAS Databases.

A Comment on Quality

- For >28 million chemical compounds there are *some* errors:
 - “Incorrect” structure representations
 - Mismatched name-structure relationships
 - Experimental properties (the values, the units)
 - Real vs. virtual compounds – text-mining and conversion
- We have deprecated a LOT of data...

Identifier Dictionaries

- Reciprocal curation processes...share curation with each other.
- If a database has a compound already then use InChiKeys to match “suggested” validation against the compound.
- A series of “added” and “removed” synonyms against InChiKeys for matching.

Federated Data Curation Sharing

Who wants to work with us?

```
<?xml version="1.0"?>
<curations since="2011-02-01T00:00:00">
  <curation status="Added" synonym="229614-55-5" InChIKey="XRQDFNLINLXZLB-GKWMMFDUSA-N"
    date="2011-02-18T10:03:31"/>
  <curation status="Removed" synonym="Cobalt chloride (CoCl2)" InChIKey="GVPFVAHMJGGAJG-
    UHFFFAOYSA-L" date="2011-03-22T09:53:36"/>
  <curation status="Removed" synonym="Additive Screening Solution 24/Fluka kit no 78374"
    InChIKey="GVPFVAHMJGGAJG-UHFFFAOYSA-L" date="2011-03-22T09:53:36"/>
  <curation status="Removed" synonym="Cobalt chloride 0.1 M solution" InChIKey="GVPFVAHMJGGAJG-
    UHFFFAOYSA-L" date="2011-03-22T09:53:36"/>
  <curation status="Removed" synonym="junk" InChIKey="GVPFVAHMJGGAJG-UHFFFAOYSA-L"
    date="2011-03-22T09:53:36"/>
  <curation status="Added" synonym="Tin Oxide" InChIKey="XOLBLPGZBRYERU-UHFFFAOYSA-N"
    date="2011-03-22T10:33:03"/>
  <curation status="Removed" synonym="Not Found" InChIKey="OKBVMLGZPNDWJK-UHFFFAOYSA-N"
    date="2011-03-23T06:36:44"/>
  <curation status="Added" synonym="1,4-Diaminonaphthalene" InChIKey="OKBVMLGZPNDWJK-
    UHFFFAOYSA-N" date="2011-03-23T06:37:33"/>
  <curation status="Added" synonym="1,4-Naphthalenediamine" InChIKey="OKBVMLGZPNDWJK-
    UHFFFAOYSA-N" date="2011-03-23T06:37:33"/>
```

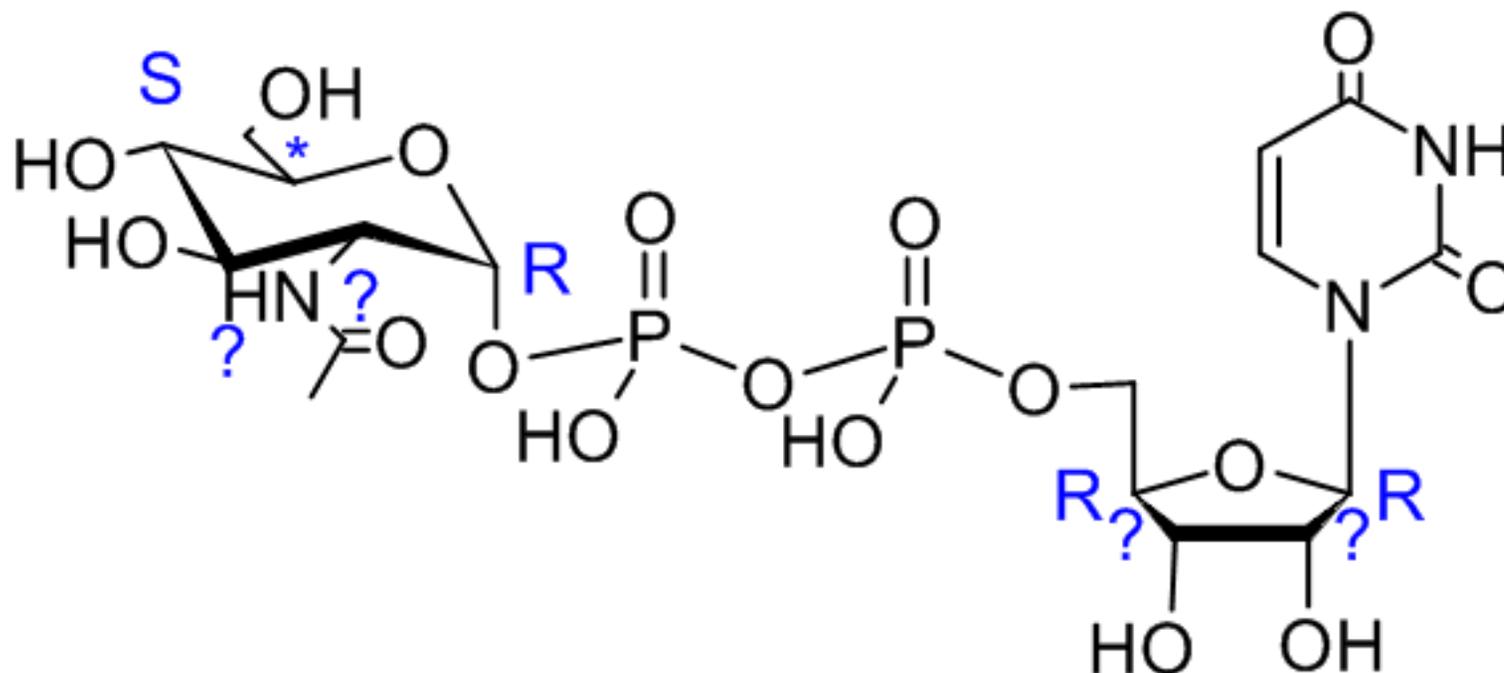
Structure Validation using feed

- Look for approved synonyms
- Compare feed InChIKey with database InChIKey
- If different, flag for inspection

Many Problems Can be Solved...

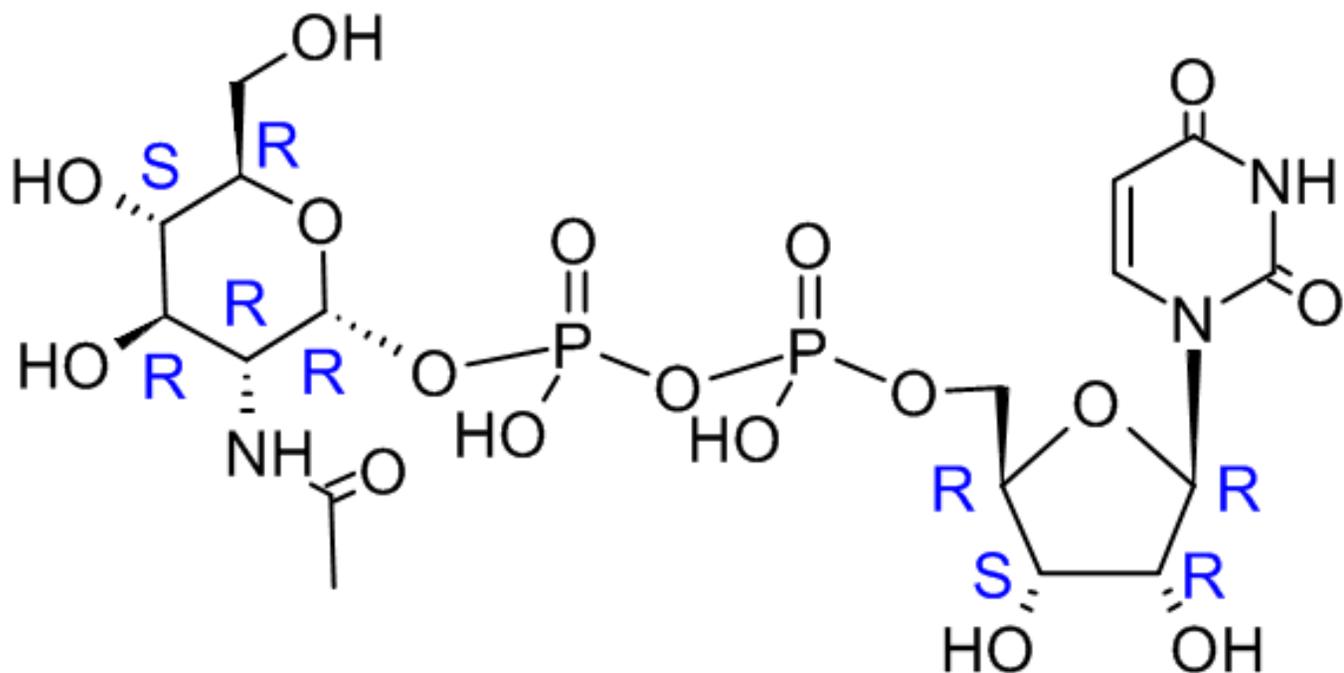
- Clean up databases – structure validation, structure standardization
- Warn about
 - Valency, charge balance, depiction issues, bond types, absent stereo, and another 100 rules (or so...)
- Standardize
 - Agree community rules to “Standardize”

Structure Validation



LFTYTUAZOPRMMI-OIYKPMIKSA-N

Structure Validation - Fixed



LFTYTUAZOPRMMI-CFRASDGPSA-N

What needs to happen?

- If we could validate
 - Catch errors in databases (and clean)
 - Proactively catch errors in publications/patents
 - Reduce junk in the ether – improve QUALITY!
- If we standardized
 - Interlinking should improve

CVSP: result of processing

[Home](#) [Upload](#) [Submissions](#) [Profile](#) [Admin](#) [Help](#)

Status: Processed
File: Kinase_dose_response_IC90_and_selective_SI.sdf (1248 records)
Standardization Type: Validate and Standardize
Validation errors: 0 records
Validation warnings: 175 records
Submission Actions: **Reprocess** **Delete**

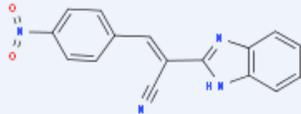
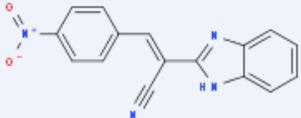
Record Actions: **Download using filter settings**
Download standardized
Download selected

Enable Auto Refresh
 [Mark File as Demo](#)

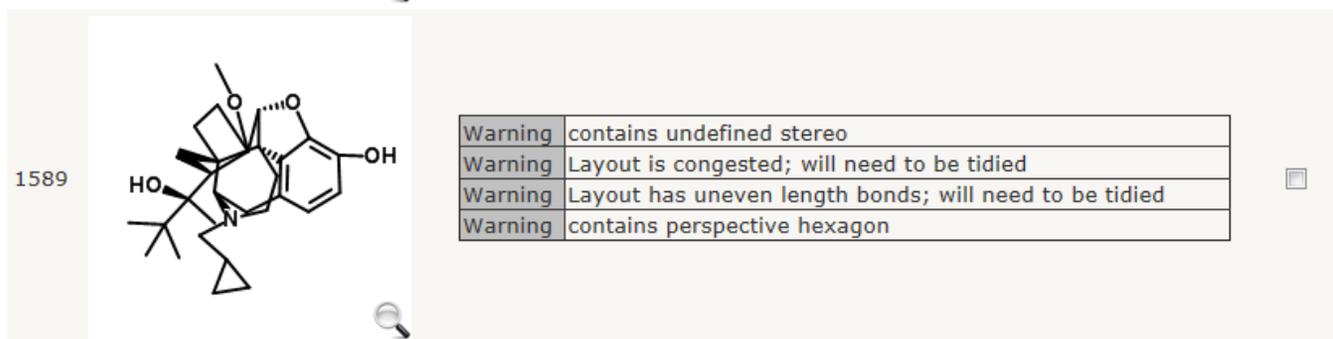
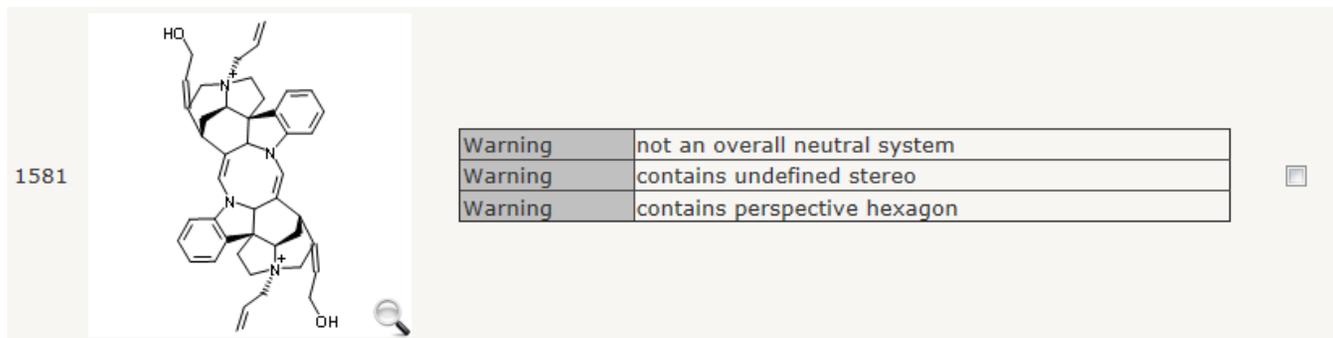
Filter records by issue type: AND Show standardized records only

All Records
All Records
Warning
Info

1 2 3 4 5 6 7 8 9

#ID	ChemSpider ID	Original	Validation Messages	Standardized				
1	   4576023		<table border="1"><tr><td>Warn</td><td>contains pentavalent nitro nitrogen</td></tr><tr><td>Info</td><td>Contains 3d coordinates, may need to be tidied</td></tr></table>	Warn	contains pentavalent nitro nitrogen	Info	Contains 3d coordinates, may need to be tidied	
Warn	contains pentavalent nitro nitrogen							
Info	Contains 3d coordinates, may need to be tidied							

NCATS Dataset



Status: Processed
 File: DrugBank_Total.sdf.zip (6516 records)
 Standardization Type: Validate and Standardize
 Validation errors: 73 records
 Validation warnings: 1079 records
 Submission Actions: **Reprocess** **Delete**

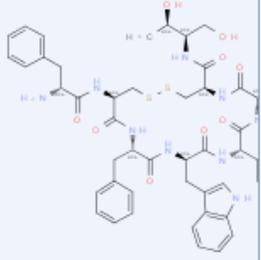
Record Actions: **Download using filter settings**
Download standardized
Download selected

- Enable Auto Refresh
- [Mark File as Demo](#)

Filter records by issue type: Warning AND by Issue

- All
- All
 - contains non-metal-transition metal bond
 - contains aluminium-non-metal bond
 - contains pentavalent nitro nitrogen
 - contains covalent metal-nitrogen bond
 - contains covalent metal-oxygen bond
 - nitrogenous base in acid form
 - contains ethane molecule(s)
 - not an overall neutral system
 - consists of more than one neutral molecule
 - contains unknown stereo bond
 - completely undefined stereo - enantiomers
 - completely undefined stereo - mixtures
 - partially undefined stereo - epimers
 - partially undefined stereo - mixtures
 - contains stereobond in six-membered ring
 - contains L-pyranose: intentional?
 - contains enol function
 - contains N=C-OH tautomer of a carbonyl compound
 - contains nitroso form of oxime

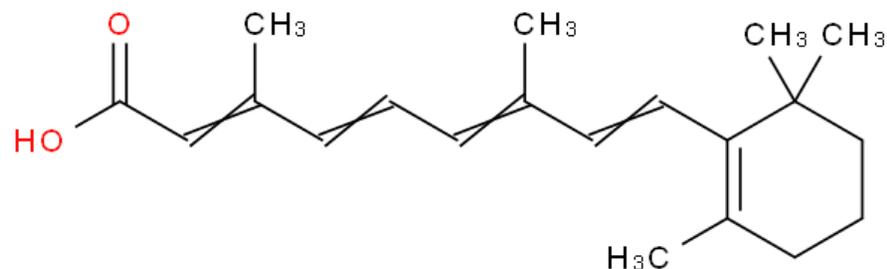
AND Show standardized records only

#ID	ChemSpider ID	Original	Standardized
4	   10482007	 Download	 No change

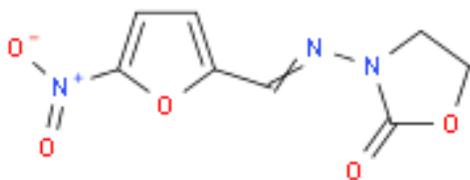
DrugBank dataset (6516 records)

- Marked as Errors (arbitrary)
 - 2 records with query bonds
 - 3 records with invalid atoms (asterisk in polymers)
 - Unusual valence: ~70 (oxygen 3, sulfur 3 and 5, Mg 4, B 5, etc.)
- Warnings
 - INCHI not matching structure (100+)
 - SMILES not matching structure (100+)

- DrugBank ID: DB00755
- InChI=1S/C20H28O2/c1-15(8-6-9-16(2)14-19(21)22)11-12-18-17(3)10-7-13-20(18,4)5/h6,8-9,11-12,14H,7,10,13H2,1-5H3,(H,21,22)/b9-6+,12-11+,15-8+,16-14+



- DrugBank ID: DB00614



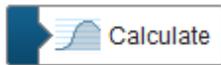
Warn	contains unknown stereo bond
Warn	depositor-specified name(s) do not match the structure : furazolidone
Warn	depositor-specified InChIs do not match the structure : InChI=1S/C8H7N3O5/c12-8-10(3-4-15-8)9-5-6-1-2-7(16-6)11(13)14/h1-2,5H,3-4H2/b9-5+

Connecting Chemistry across the web

- So much of what is seen on ChemSpider is retrieved in real time using services

▼ Properties

Predicted - ACD/Labs Predicted - ChemAxon

 Calculate

Click to predict properties on the Chemicalize site

▼ Articles

Links & Reference RSC Journals RSC Books PubMed MeSH Literature Google Books

Results 1 - 10 of 1113 (Click [here](#) to explore results)

Metal Phosphonate Chemistry: From Synthesis to Applications, Chapter 14 Structural Diversity in Metal Phosphonates
Konstantinos D. Demadis, Nikoleta Stavgianoudaki, 2012, Pages 438-492

Amino Acids, Peptides, and Proteins, Metal complexes of amino acids and peptide. J. S. Davies, K. B. Nolar
282-332

Organophosphorus Chemistry, Physical methods. D. W. Allen, J. C. Tebby, R. N. Slinn, M. C. Salt, 2000, Vol. 1, Pages 185-208

Organophosphorus Chemistry, Phosphates and phosphonates of biochemical interest. D. W. Hutchinson, O. S. Schmechel, 1997, Pages 185-208

▼ Articles

Links & Reference RSC Journals RSC Books PubMed MeSH Literature Google Books

 [Glyphosate Resistance in Crops and Weeds: History, Development, and Management](#)
by unknown
2010 - 321 pages
[books.google.com](#)

 [Glyphosate: a unique global herbicide](#)
by John E. Franz, Michael K. Mao, James A. Sikorski
1997 - 653 pages
[books.google.com](#)

Connecting Chemistry across the web

▼ **Patents**

[Google Patents](#) [USPTO Granted \(9099\)](#) [USPTO Applications \(9505\)](#) [European Granted \(1637\)](#) [European Applications \(1923\)](#)
[Japanese Abstracts \(178\)](#)

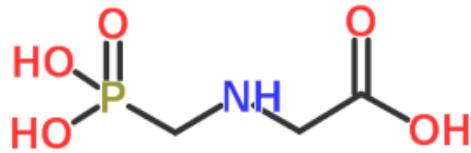
178 patents found in Japanese Abstracts

Patent No.	Title
2000112761	CIRCUIT FOR DIVIDING DIVIDEND BY 2^n THROUGH THE USE OF ROUND-UP OR ROUND-DOWN AP
08134088	ISOLATION OF N-PHOSPHONOMETHYLGLYCINE
09249509	GLYPHOSATE SOLUTION PREVENTING COLORATION

Online Predictions

Identifier Conversion and Properties Predictions ([Read Manual](#))

Click the image below to enter a chemical structure:



Identifiers

SMILES: C(C(=O)O)NCP(=O)(O)O

InChI: InChI=1/C3H8NO5P/c5-3(6)1-4-2-10(7,8)9/h4H,1-2H2,(H,5,6)(H2,

InChIKey: XDDAORKBJWWYJS-UHFFFAOYAE

Std. InChI: InChI=1S/C3H8NO5P/c5-3(6)1-4-2-10(7,8)9/h4H,1-2H2,(H,5,6)(H2,

Std. InChIKey: XDDAORKBJWWYJS-UHFFFAOYSA-N

Inherent Properties

Molecular Formula: C₃H₈NO₅P

Molecular Weight: 189.073

Nominal Mass: 189.000

Average Mass: 189.073

Monoisotopic Mass: 189.014

Calculated Properties [?](#)

ACD/LogP: 0.272±0.588 ± 0.8

of Rule of 5 Violations: 0

Number of Hydrogen Bond Acceptors: 6

Number of Hydrogen Bond Donors: 4

Number of Freely Rotatable Bonds: 4

Polar Surface Area: 116.670

Boiling Point: 465.8±55.0 ± 55.0 Celsius

Flash Point: 235.528 °C ± 31.5 Celsius

Enthalpy of Vaporization: 79.713 ± 6.0 kJ/mol

Molar Volume: 100.656 ± 3.0 cm³

Web Services Open Up Collaboration

- Agilent, Bruker, Waters and Thermo all use our web-based services for compound lookup
- Many academic sites integrating directly – metabonomics, name lookup, semantic markup
- Mobile app integration
- Commercial structure drawing packages

Web Services

Search

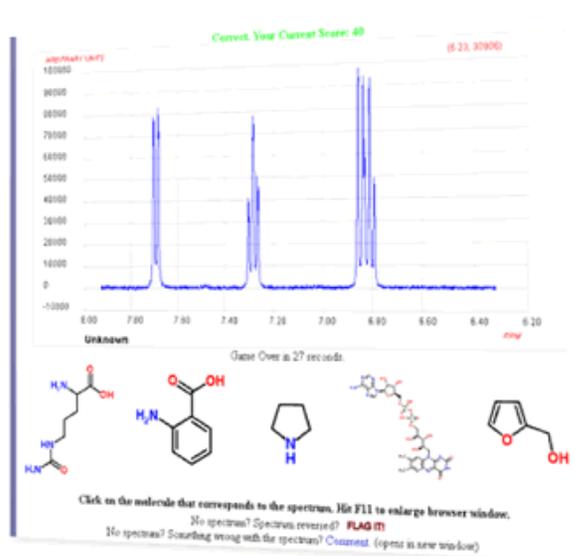
Please send all feedback to development-at-chemspider-dot-com

NOTE: Some operations require a "Security Token". To obtain a token please complete the [registration](#) process. Security Token is listed at [Profile](#) page.

The following operations are supported. For a formal definition, please review the [Service Description](#).

- [AsyncSimpleSearch](#)
Search by Name, SMILES, InChI, InChIKey, etc. Returns transaction ID which can be used to access search status and result. Security token is required.
- [CSID2ExtRefs](#)
Return a list of external references (data sources). Security token with Service Subscriber role is required.
- [GetAsyncSearchResult](#)
Returns the list of CSIDs found by AsyncSearch operation. Security token is required.
- [GetAsyncSearchResultPart](#)
Return a slice of the list of CSIDs found by AsyncSearch operation. Returns full list if start = 0 and count = -1. If (start + count) > (# of results) all results starting at start position are returned. Security token is required.
- [GetAsyncSearchStatus](#)
Query asynchronous operation status. Requires transaction ID returned by AsyncSearch operation. Security token is required.
- [GetCompoundInfo](#)
Get record details (CSID, InChIKey, InChI, SMILES) by CSID. Security token is required.

ChemSpider Everywhere: Spectral Game

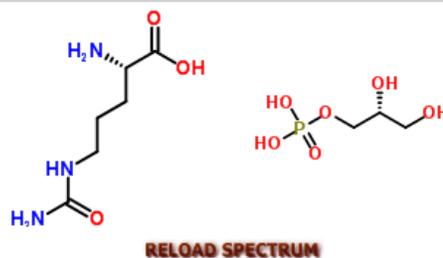
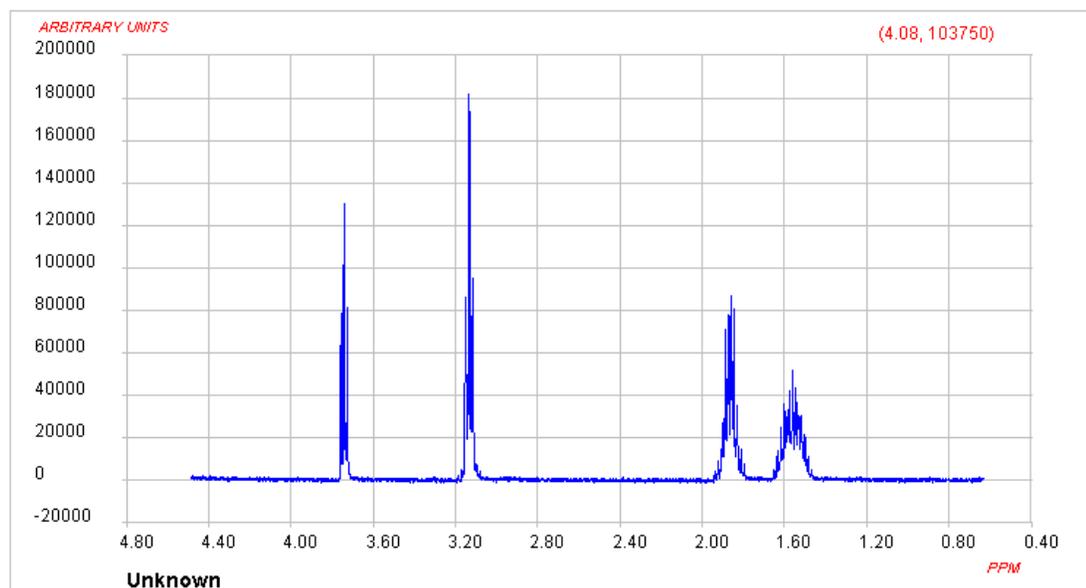


How to play

Each round you'll be presented with a new spectrum. You have to select the molecule that matches the spectrum. For each molecule you identify correctly you'll get one point. The game continues until you get one wrong. Enter your name, choose a group (optional), select the type of spectrum you want and click play.

ChemSpider Everywhere

Crowdsourced Curation of Spectra



Click on the molecule that corresponds to the spectrum. Hit F11 to enlarge browser window.

No spectrum? Spectrum reversed? **FLAG IT!**

No spectrum? Something wrong with the spectrum? [Comment](#). (opens in new window)

Web Services Integrate INTERNAL Projects

- Integration between ChemSpider and...
 - Our publishing platform for structure display
 - ChemSpider SyntheticPages
 - LearnChemistry Wiki
 - National Chemical Database Service
 - And.....a growing list.....

What ChemSpider Does Not Handle

- Polymers
- Markush structures
- Organometallics
- Many Inorganics
- Materials
- Reactions...but....

ChemSpider Reactions

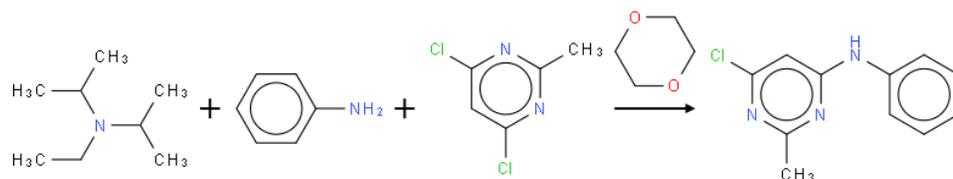
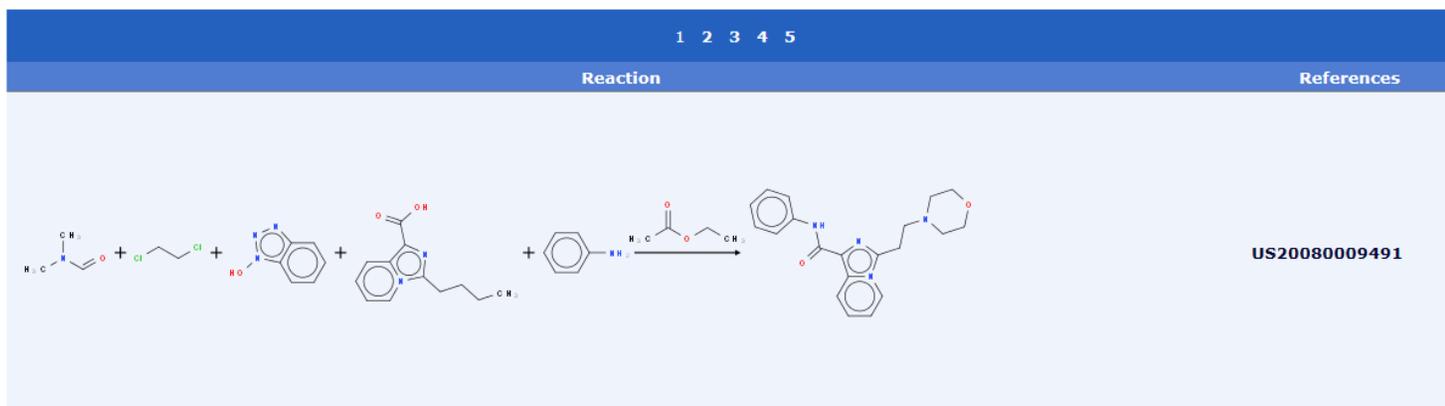


ChemSpider Reactions (Alpha)

Exact Component Any Component USPTO

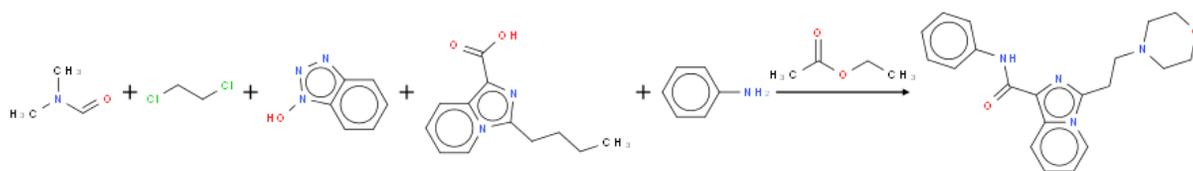
Query exactly matches one of components. Example: c1cccc1

Total: 100

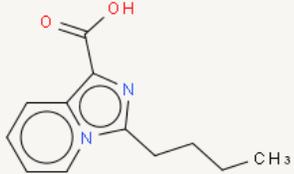
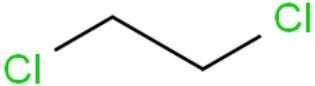
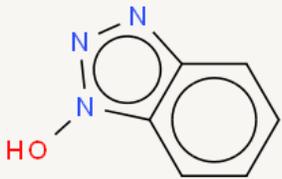


US20080014189

http://csr.dev.rsc-us.org/Reaction.aspx?rec_id=579



Find Sub Reactions Find Similar Reactions

Molecule	Role	CSID	Links																
	Reactant		<table border="1"> <thead> <tr> <th>Reactant</th> <th>Agent</th> <th>Product</th> <th>Any</th> </tr> </thead> <tbody> <tr> <td>Exact</td> <td>Exact</td> <td>Exact</td> <td>Exact</td> </tr> <tr> <td>Super</td> <td>Super</td> <td>Super</td> <td>Super</td> </tr> <tr> <td>Similar</td> <td>Similar</td> <td>Similar</td> <td>Similar</td> </tr> </tbody> </table>	Reactant	Agent	Product	Any	Exact	Exact	Exact	Exact	Super	Super	Super	Super	Similar	Similar	Similar	Similar
Reactant	Agent	Product	Any																
Exact	Exact	Exact	Exact																
Super	Super	Super	Super																
Similar	Similar	Similar	Similar																
	Reactant	13837650	<table border="1"> <thead> <tr> <th>Reactant</th> <th>Agent</th> <th>Product</th> <th>Any</th> </tr> </thead> <tbody> <tr> <td>Exact</td> <td>Exact</td> <td>Exact</td> <td>Exact</td> </tr> <tr> <td>Super</td> <td>Super</td> <td>Super</td> <td>Super</td> </tr> <tr> <td>Similar</td> <td>Similar</td> <td>Similar</td> <td>Similar</td> </tr> </tbody> </table>	Reactant	Agent	Product	Any	Exact	Exact	Exact	Exact	Super	Super	Super	Super	Similar	Similar	Similar	Similar
Reactant	Agent	Product	Any																
Exact	Exact	Exact	Exact																
Super	Super	Super	Super																
Similar	Similar	Similar	Similar																
	Reactant	68282	<table border="1"> <thead> <tr> <th>Reactant</th> <th>Agent</th> <th>Product</th> <th>Any</th> </tr> </thead> <tbody> <tr> <td>Exact</td> <td>Exact</td> <td>Exact</td> <td>Exact</td> </tr> <tr> <td>Super</td> <td>Super</td> <td>Super</td> <td>Super</td> </tr> <tr> <td>Similar</td> <td>Similar</td> <td>Similar</td> <td>Similar</td> </tr> </tbody> </table>	Reactant	Agent	Product	Any	Exact	Exact	Exact	Exact	Super	Super	Super	Super	Similar	Similar	Similar	Similar
Reactant	Agent	Product	Any																
Exact	Exact	Exact	Exact																
Super	Super	Super	Super																
Similar	Similar	Similar	Similar																

DERA

- Digitally Enabling the RSC Archive...back to 1841
 - Extracting data and making available via appropriate platform
 - Chemicals
 - Reactions
 - Analytical Data
 - Figures

[Home](#) > [Student Corner](#) > [student-corner-news-updates](#) >

CSIR-OSDD Signs MoU with Royal Society of Chemistry

posted Jan 28, 2013, 8:59 PM by Anshu Bhardwaj [updated Jan 30, 2013, 12:18 AM]



In pursuit of common aims and to raise awareness of the importance of cheminformatics to accelerate the discovery of novel therapies for neglected diseases like TB and Malaria, CSIR-OSDD signed a Memorandum of Understanding (MoU) with Royal Society of Chemistry (RSC), the largest European organisation for advancing chemical sciences on 29th Jan 2013. The MoU

signing ceremony, conducted at the Council of Scientific & Industrial Research, Anusandhan Bhawan was graced by the presence of Hon'ble Science & Technology Minister, Shri Jaipal Reddy.

Chemical Database Service

National Chemical Database Service

The RSC will be operating the EPSRC National Chemical Database Service from 2013-2017

What is the RSC's vision for the Service?

We intend to build the Service for the future - to develop a chemistry data repository for UK academia, and to build tools, models and services on this data store to increase the value and impact of researchers' funded work. We will continue to develop this data store through the lifetime of the contract period and look forward to working with the community to make this a world-leading exemplar of the value of research data availability.

The Service will also offer access to a suite of commercial databases and services. While there will be some overlap with currently provided databases popular with the user community we will deliver new data and services and optimize the offering based on user feedback.

When will the Service be available?

The Service will start on 2nd January 2013, and will be available at cds.rsc.org

The database services we are working to have available at launch are the Cambridge Structural Database, ACD/ILab and Accelrys' Available Chemicals Directory. The Service will also include integrated access to the [RSC's award winning ChemSpider database](#). As 'live' dates for other services become clear, they will appear here.

How will this affect existing users?

While the CSD and ACD/ILab are continuing, and additional new data sources will be added, access to several of the other existing databases will be ended. Interfaces and access mechanisms will change as we deliver the services differently.

We realise that existing users will be affected by the change in access to

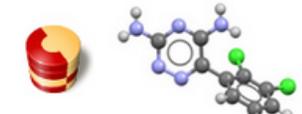
Available databases

Click boxes for further information and login.

ACD/ILab2



CSD



MOS, NPU, CCR



SPRESIweb



Available Chemicals Directory



ICSD



ChemSpider



In partnership with the EPSRC



Data for life sciences

LETTERS

NATURE CHEMICAL BIOLOGY VOLUME 3 NUMBER 5 MAY 2007

nature
chemical biology

Chemical genetics reveals a complex functional ground state of neural stem cells

Phedias Diamandis¹⁻⁴, Jan Wildenhain⁴, Ian D Clarke^{1,2}, Adrian G Sacher^{1,2}, Jeremy Graham^{1,2}, David S Bellows³, Erick K M Ling^{1,2,5}, Ryan J Ward^{1,2,5}, Leanne G Jamieson^{1,2,5}, Mike Tyers^{3,4} & Peter B Dirks^{1,2,5,6}

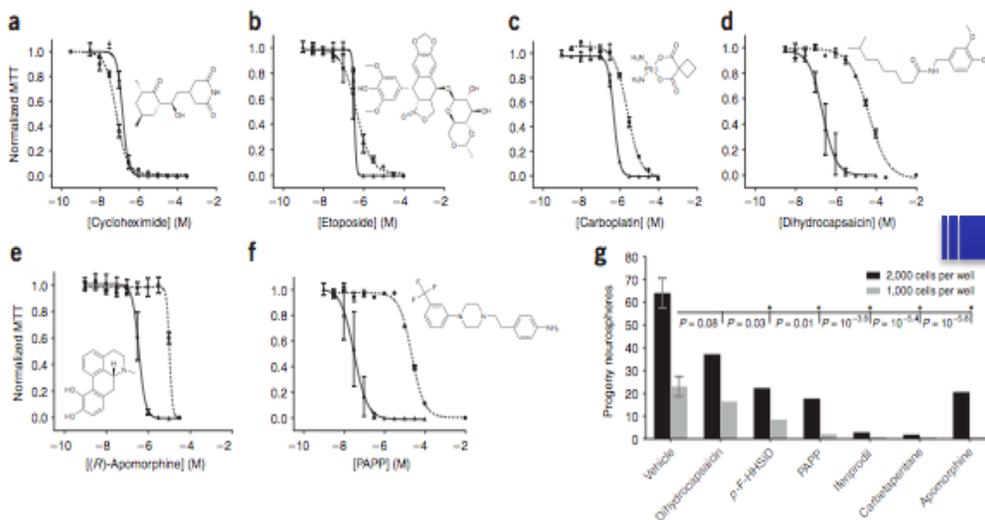


Figure 2 Identification of potent NPC-specific compounds. (a-f) Dose-response curves and chemical structures of controls: cycloheximide (a), etoposide (b) and carboplatin (c), and of selected newly identified compounds: dihydrocapsaicin (d), apomorphine (e) and PAPP (f). Each plot shows the fitted sigmoidal logistic curve to MTT proliferation assay readings of both astrocytes (-●-) and neurosphere cultures (-▲-). Values represent the mean and



© 2007 Nature Publishing Group <http://www.nature.com/naturechemicalbiology>

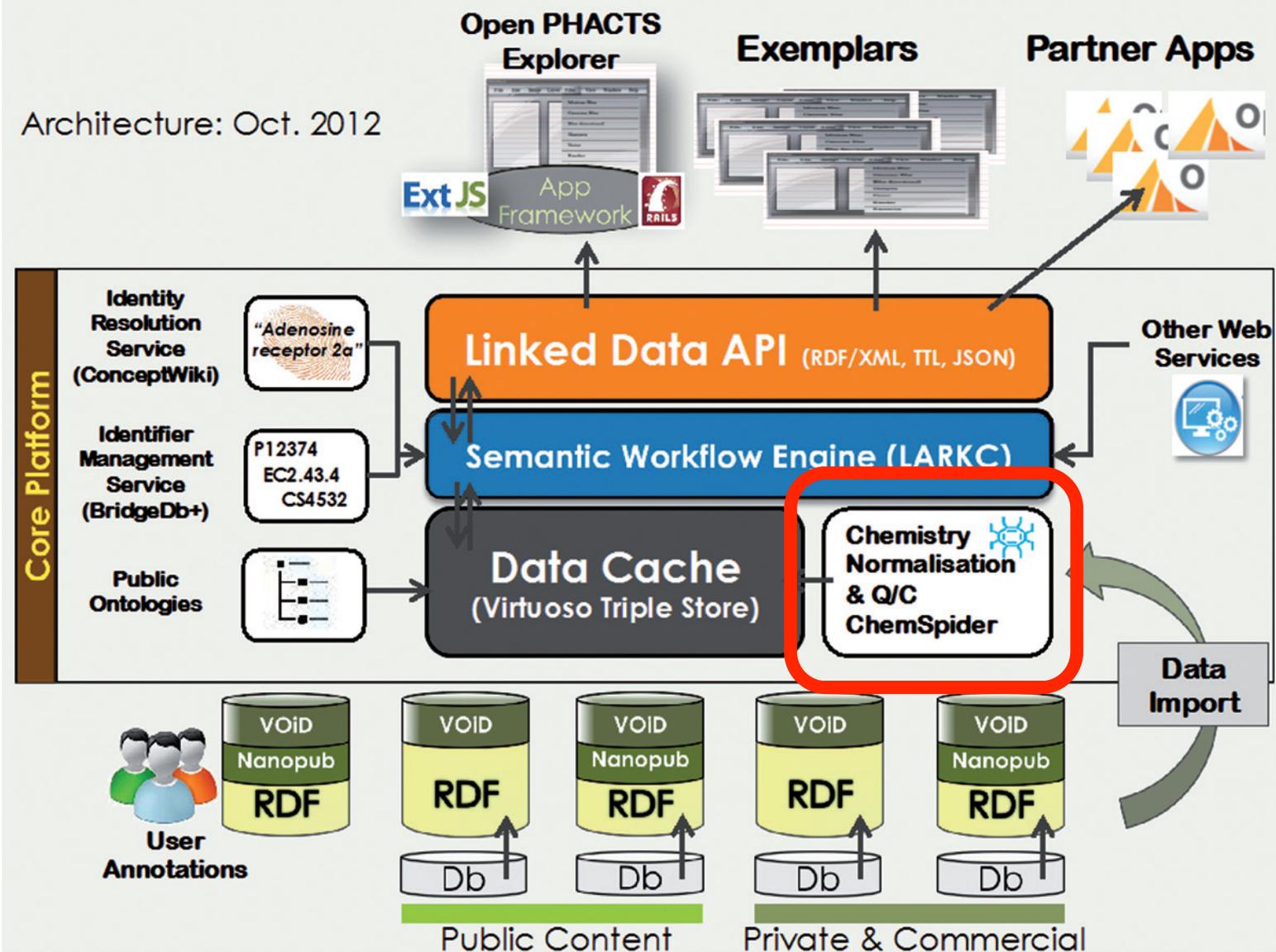
OpenPHACTS

- Open PHACTS is an Innovative Medicines Initiative (IMI) – 3 years project
- To reduce the barriers to drug discovery in industry, academia and for small businesses
- To build an open platform, integrating chemistry and biology data from public domain resources
- Semantic web platform
- Open Standards, Open Data and Open Source



- **Crowdsourcing across drug discovery**
- Open PHACTS : partnership between European Community and European Pharma Companies
- 22 partners, 8 pharmaceutical companies, 3 biotechs working together for 3 years
- Freely accessible for knowledge discovery and verification.
 - Data on chemistry and biology
 - Pharmacological profiles
 - Proprietary and public data sources.

Architecture: Oct. 2012

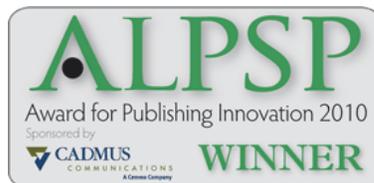


PharmaSea

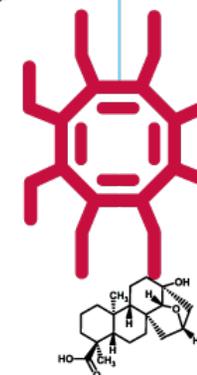
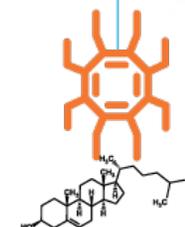
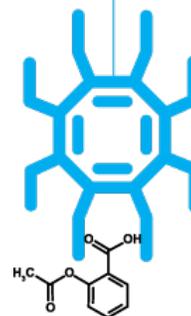
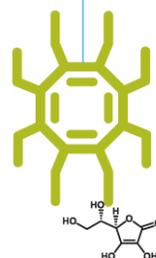
- FP7 Initiative. PharmaSea: increasing value and flow in the marine biodiscovery pipeline (2012-2017)
- Improve the quality, volume and value of active agents discovered in the marine environment and increase the speed at which they can be delivered
- RSC: Providing dereplication via ChemSpider, analytical data algorithms, integration with computer-assisted structure elucidation algorithms

Conclusions

- RSC eScience supporting increasing number of grant-based projects
- ChemSpider grows daily – community depositions and data from RSC Content with a focus on expanding data while improving quality
- ChemSpider is an integration platform for MANY projects through web services
- CVSP processing is available to use and provide feedback – will be available as a service also
- We believe in curation **sharing** - who wants to collaborate?



Thank you



Email: williamsa@rsc.org

Twitter: ChemConnector

Blog: www.chemspider.com/blog

Personal Blog: www.chemconnector.com

SLIDES: www.slideshare.net/AntonyWilliams

RSC | Advancing the
Chemical Sciences



ChemSpider

Building community for chemists