Qualifying Online Information Resources for Chemists

NFAIS-CENDI-FLICC 12/8/2008

Antony Williams

Access to Information

- For me…
  - PhD : Libraries primary source of information
  - PostDoc/Academia: Libraries and librarians
  - Eastman Kodak: Software tools and databases
  - Kodak and ACD/Labs: Replaced by the internet
  - Today: The Internet enhanced by a network of collaborators…
  - Librarians have become gurus in using software systems to resource information

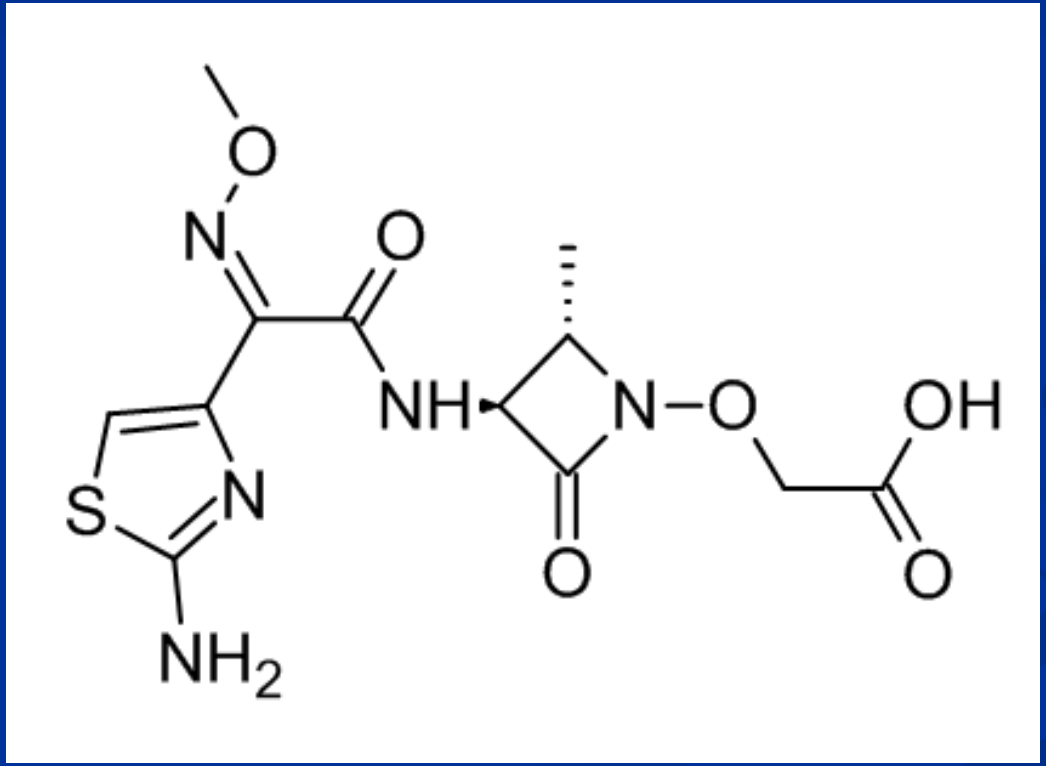
Content is King

- Chemistry "content" is big money – Chemistry publishing and content is worth $100s of millions/year
  - Patent searching
  - Structures and properties
  - Drug databases
  - Literature databases

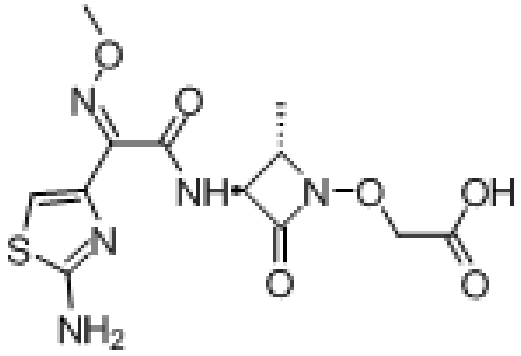
- **Chemical Abstracts Service** (CAS), a division of the ACS is "Gold Standard" in Chemistry related information
  - 101 years of content, $260 million revenue (2006), >40 million substances and 60 million sequences

The Language of Chemistry

- My language….



And its dialects….



90849-080-4 (CAS Number)

Oximonam

{[(2*S*,3*S*)-3-{[(2*Z*)-2-(2-amino-1,3-thiazol-4-yl)-2-(methoxyimino)acetyl]amino}-2-methyl-4-oxoazetidin-1-yl]oxy}acetic acid

Nc1nc(cs1)\C(=N\OC)C(=O)N[C@@H]2C(=O)N(OCC(=O)O)[C@H]2C

InChI=1/C12H15N5O6S/c1-5-8(11(21)17(5)23-3-7(18)19)15-10(20)9(16-22-2)6-4-24-12(13)14-6/h4-5,8H,3H2,1-2H3,(H2,13,14)(H,15,20)(H,18,19)/b16-9-/t5-,8-/m0/s1

FJKOYBHMMTVFHK-TWYJFGHKBO

As a chemist…

- I look for information about chemicals/chemistry
  - What is a particular structure ?
  - What alternative names/identifiers?
  - Reaction synthesis?
  - Physical properties?
  - Analytical data?
  - Purchase?
  - Tell me more?
  - Similar stuff – what other compounds are “like” mine?

Why Journals?

- Journals contain lots of information but are limited – text, charts, graphs and pictures.
- Text-based searches of the internet gets me to articles VERY quickly then articles can disappoint me. **I use what I can afford**. So do others…
  - Google
  - Google Scholar
  - PubMed
- Updating my CV recently was a breeze…the Internet versus other sources

Searching and Reading Articles…

- Searching articles based on chemical structure and substructure is very expensive.. but is changing
- The web IS "tool-ready" so when will publishers deliver?
  - Structures can be shown
  - Spectra can be interactive
  - Graphics don't need to be static
  - Publishers can enhance their articles (Project Prospect from the RSC is an example)

Publications

Reagents and conditions: (**a**) *n*-BuLi, Et_2O, 0 °C to 25 °C. (**b**) Me_3SnCl, 0 °C, 73% over two steps. (**c**) Benzyl chloride, K_2CO_3, DMF, 83%. (**d**) (i) Methyl azidoacetate, sodium methoxide, MeOH. (ii) *o*-Xylene, 160 °C, 47% over two steps. (**e**) (i) NaOH. (ii) $CuSO_4$, Na_2CO_3. (iii) Quinoline, 215 °C, 72% over three steps. (**f**) $POCl_3$, DMF, 87%. (**g**) $MeNO_2$, NH_4OAc, 98%. (**h**) (i) $LiAlH_4$. (ii) Di-*t*-butyldicarbonate, 4-*N*-dimethylaminopyridine, 80% over two steps. (**i**) **Compound 11**, $Pd(PPh_3)_4$, CuCl, LiCl, DMSO, 60 °C, 2 d, 71%. (**j**) Pd black, HCO_2NH_4, 94%. (**k**) (i) Salcomine, O_2, MeCN. (ii) 1% formic acid in H_2O, 77% over two steps, 1:2.5 **Compound 20**:**Compound 21**. DMF, *N,N*-dimethylformamide; salcomine, bis(salicylidene)ethylenediimino cobalt II.

Full size image (54 KB)

Elements

Families

The tryptamine cross-coupling partner **Compound 17** was synthesized beginning with commercially available bromo-salicylaldehyde **Compound 12**. Benzyl protection (yielding **Compound 13**), followed by condensation with methyl azidoacetate and subsequent Hemetsberger indole cyclization, provided 2-carbomethoxyindole **Compound 14**. Subsequent saponification and copper II–mediated decarboxylation in quinoline provided indole **Compound 15** (ref. 17). Formylation, Henry reaction (yielding **Compound 16**), reduction and protection of the nitrogens then gave the tryptamine **Compound 17**.

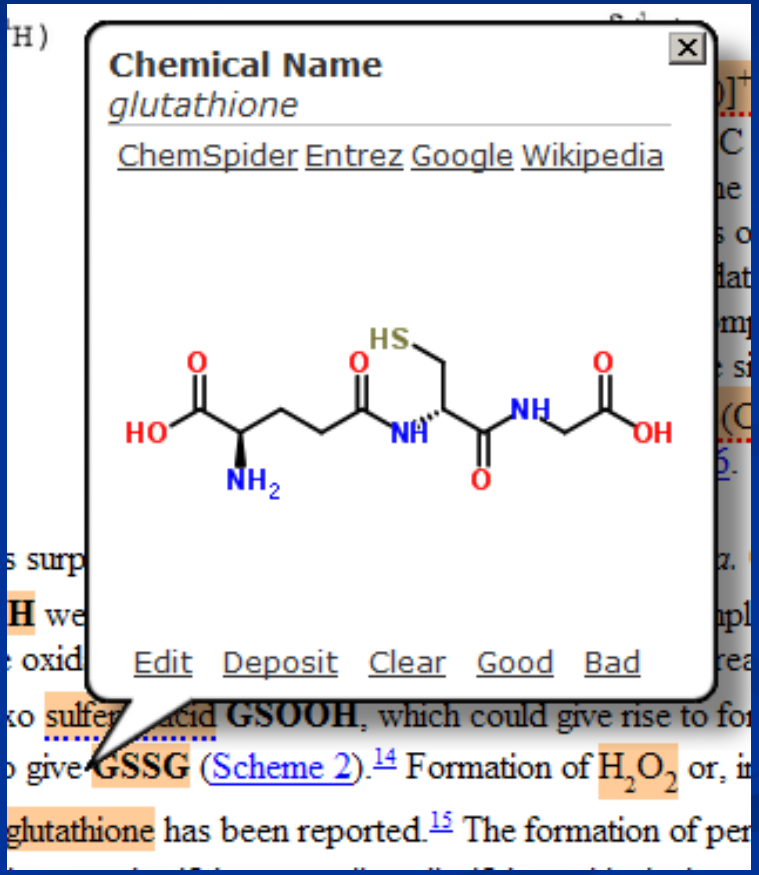
The stage was now set to form the critical C7–C10 bond through Stille cross-coupling. Reaction of **Compound 17** with **Compound 11** in presence of a palladium catalyst did indeed give the protected biaryl **Compound 18** in good overall yield[18]. Debenzylation, followed by subjecting phenol **Compound 19** to salcomine under an oxygen atmosphere, provided a 1:2.5 regioisomeric mixture of *p*- and *o*-quinones **Compound 20** and **Compound 21**, respectively. We were able to separate the undesired regioisomer from the desired *p*-quinone **Compound 20** after selective cleavage of the indole *t*-butylcarbamate (Boc) group under mild conditions (1% formic acid in water).

Names

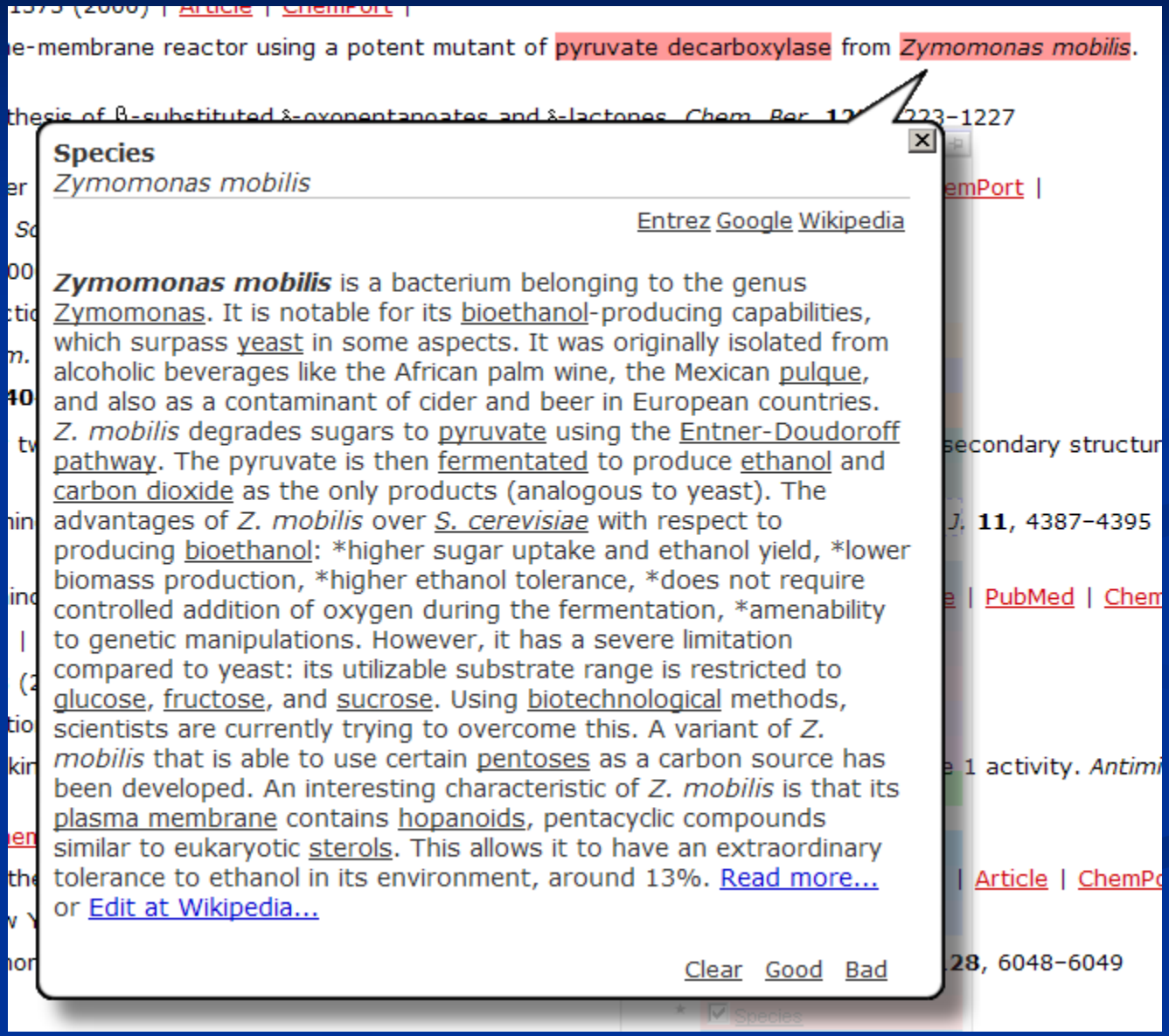
In the final phase of our synthesis, we subjected *p*-quinone **Compound 20** to a large excess of the lithium enolate of *N,N*-dimethylhydantoin. After aqueous workup in air, we isolated *p*-quinone **Compound 22** (Scheme 3). Removal of all remaining protecting groups using boron tribromide then gave crude **Compound 23**, which is a partially reduced version of **5**, the centerpiece of our biosynthetic proposal. Treatment of the crude product with ten equivalents of silver II oxide in methanol and 2% formic acid[19] did indeed give exiguamine A in 46% yield. This single synthetic operation includes two oxidations, the intramolecular nucleophilic attack of a tertiary amine onto an *o*-quinone, tautomerization and, finally, *oxa*-6π-electrocyclization. Silver II oxide proved essential to the reaction cascade, as subjecting crude **Compound 23** to methanol and 2% formic acid under an ambient atmosphere failed to oxidize the catechol moiety, resulting in the recovery of pure **Compound 23**.

Enable Electronic Articles…

- Structures are the language of chemistry
- Show structures to chemists and search/link from there…



Allow Integration…

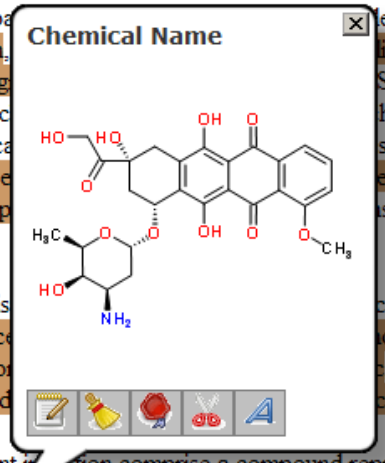


And Extend to Patents…

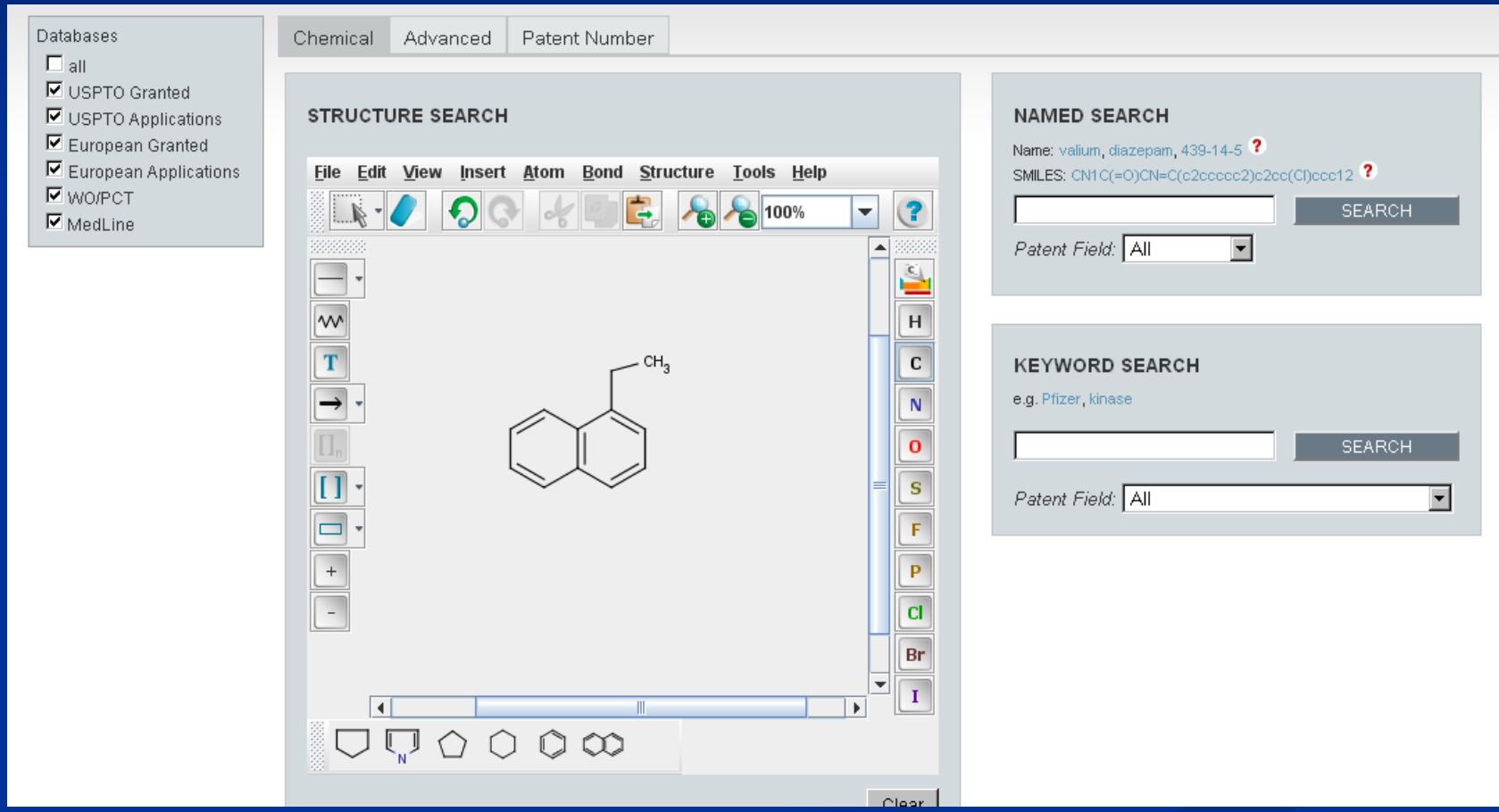
[0179] The term "pharmaceutically acceptable salts" refers to salts prepared from pharmaceutically acceptable non-toxic bases or acids. When the compound of the present invention is acidic, its corresponding salt can be conveniently prepa…le non-toxic bases, including inorganic bases and organic bases. Salts derived from such inorganic bases include aluminum, ammonium, calcium, …lithium, magnesium, manganese (ic and ous), potassium, sodium, zinc and the like salts. Particularly preferred are the ammonium, calcium, mag… Salts derived from pharmaceutically acceptable organic non-toxic bases include salts of primary, secondary, and tertiary amines, as well as cyc…h as naturally occurring and synthesized substituted amines. Other pharmaceutically acceptable organic non-toxic bases from which salts ca…sins such as, for example, arginine, betaine, caffeine, choline, N,N-dibenzylethylenediamine, diethylamine, 2-diethylaminoethanol, 2-dimethylaminoe…e, N-ethylmorpholine, N-ethylpiperidine, glucamine, glucosamine, histidine, hydrabamine, isopropylamine, lysine, methylglucamine, morpholine, p…s, procaine, purines, theobromine, triethylamine, trimethylamine, tripropylamine, tromethamine and the like.

[0180] When the compound of the present invention is …conveniently prepared from pharmaceutically acceptable non-toxic acids, including inorganic and organic acids. Such acids include, for example, ac…orsulfonic, citric, ethanesulfonic, fumaric, gluconic, glutamic, hydrobromic, hydrochloric, isethionic, lactic, maleic, malic, mandelic, methanesulfo…, phosphoric, succinic, sulfuric, tartaric, p-toluenesulfonic acid and the like. Particularly preferred are benzenesulfonic, citric, hydrobromic, hyd…, and tartaric acids.

[0181] The pharmaceutical compositions of the present i…ition comprise a compound represented by Formulas IA, IB, IIA, or IIB (or pharmaceutically acceptable salts thereof) as an active ingredient, a pharmaceutically acceptable carrier and optionally other therapeutic ingredients or adjuvants. Such additional therapeutic ingredients include, for example, cytotoxic agents (alkylators, DNA topoisomerase inhibitors, antimetabolites, tubulin binders); inhibitors of angiogenesis; and other different forms of therapies including kinase inhibitors such as Tarceva, monoclonal antibodies, cancer vaccines, doxorubicin, vincristine, cisplatin, carboplatin, gemcitabine, and taxanes. The compositions include compositions suitable for oral, rectal, topical, and parenteral (including subcutaneous, intramuscular, and intravenous) administration, although the most suitable route in any given case will depend on the particular host



Structure-based Patent Searching SureChem and IBM services



Building a Structure Centric Community for Chemists

What can be done?

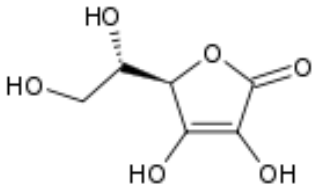
International Chemical Identifier

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]

CH_3CH_2OH
ethanol

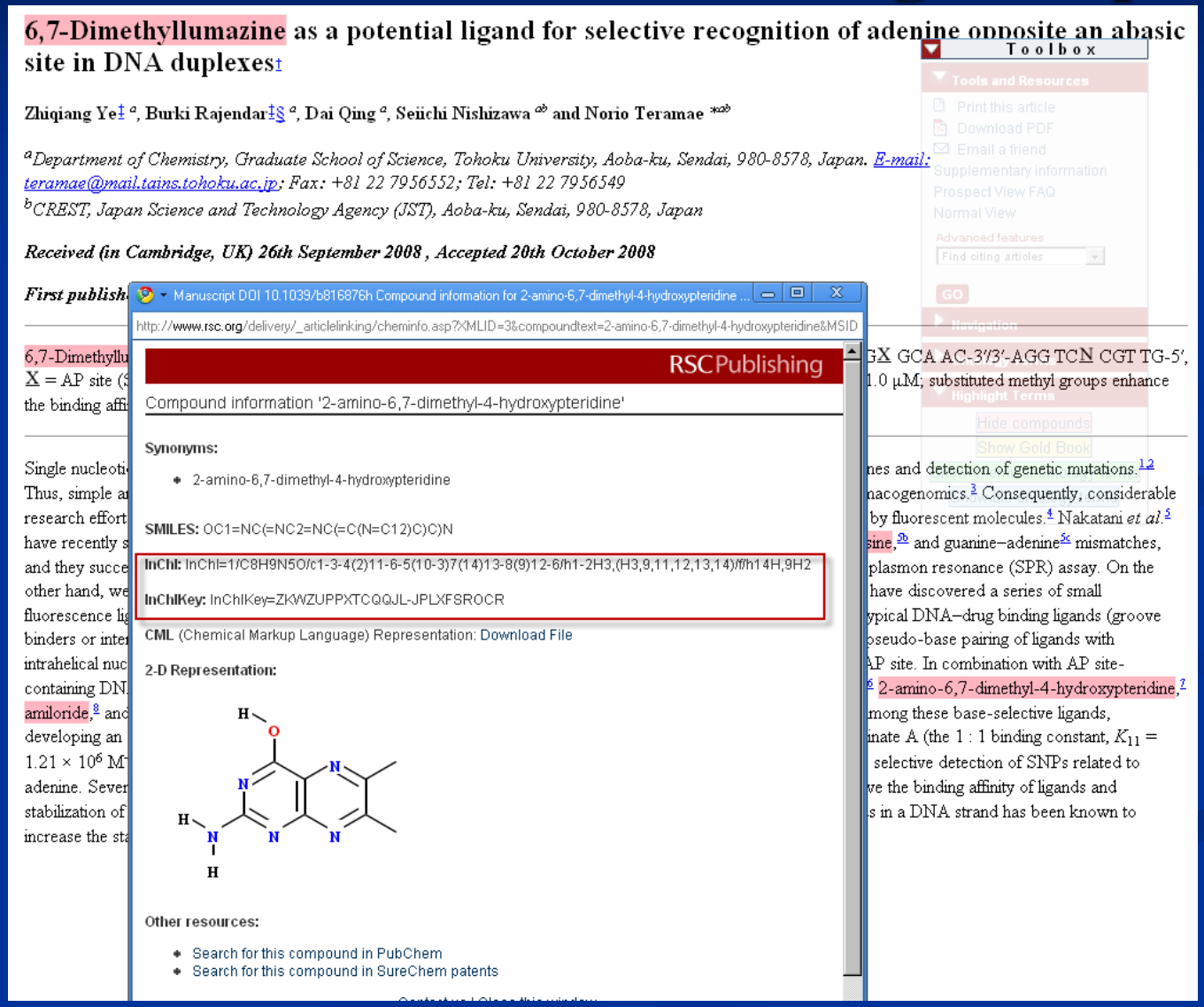
InChI=1/C2H6O/c1-2-3/h3H,2H2,1H3



L-ascorbic acid

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

Publishers should adopt/add InChIs RSC and Nature Publishing Group have!



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

Zhiqiang Ye‡ [a], Burki Rajendar‡§ [a], Dai Qing [a], Seiichi Nishizawa [ab] and Norio Teramae *[ab]

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

[b]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 publish

Toolbox
Tools and Resources
Print this article
Download PDF
Email a friend
Supplementary information
Prospect View FAQ
Normal View
Advanced features
Find citing articles
GO
Navigation
Highlight Terms
Hide compounds
Show Gold Book

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

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

RSCPublishing

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=1/C8H9N5O/c1-3-4(2)11-6-5(10-3)7(14)13-8(9)12-6/h1-2H3,(H3,9,11,12,13,14)/f/h14H,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

Blogs, Wikis, Forums and Collaborative Science

- I have two blogs, one forum and a full blog reader…
  - http://www.chemspider.com/blog
  - http://www.chemspider.com/chemunicating (ChemConnector)

  - http://forum.chemspider.com/
- They are catalytic for collaborations, getting questions answered, garnering comments and feedback
- There are upsides and downsides: http://www.chemspider.com/blog/the-joys-and-frustrations-of-6-months-blogging-in-the-chemistry-community.html

Blogging Experience and Judgments

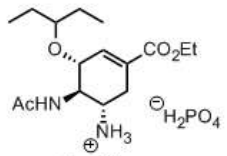
- The blogging community for chemistry small and tight
- Benefits to me
  - Fast feedback – on and offline
  - Extended network, diverse skills
  - Fast way to spread news – a local PressWire
- Most blogs are for information sharing, opinions
- Low in scientific content – the content is "off-blog" – blogs help find it
- Small number of blogs doing real science

TotallySynthetic.com

totallysynthetic*

home about index aup faq

Oseltamivir phosphate (Tamiflu) Pt. 5



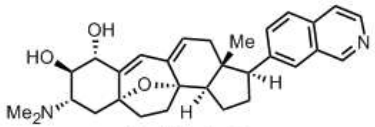
Tamiflu

Shibasaki, Yamatsugu, Yin, Kamijo, Kimura and Kanai. *ACIEE*, **2008**, *EarlyView*. DOI: 10.1002/anie.200804777.

A fifth appearance for my favourite drug 'interloper' in to this natural-product space; alarm bells shouldn't be ringing - just cause I work in pharma doesn't mean I've turned my back on natural products! Tamiflu is of course based on a natural product, shikimic acid - the starting point for the original synthesis. But as natural sources go, it's rather hard to get hold of, and thus damned pricey (£248 for 5g on SA just now). Other routes used involved chemistry that was perhaps a mite 'tetchy' on scale, such as azides and aziridines. A few years

Posted at 2pm on 29/11/08 | 33 comments | Filed Under: Still In The RBF
read on

Cortistatin Pt. III



Cortistatin A

Shair, Lee, and Nieto-Oberhuber, *JACS*, **2008**, *ASAP*. DOI: 10.1021/ja8071918.

A third showing for everybodies favourite androstane, this offering from Matt Shair adds to the quantity of inovative chemistry used in it's contruction. As a quick reminder, first up was Phil Baran, back in May; then came Nicolaou and Chen in August - along with several 'studies towards papers'. However, rather than my going through it all again, have a look at this excellent review by Stefan Bräse which was in ACIEE last month.

If you read it through, you'll notice Nicolaou's use of the Hajos-Parrish ketone (the synthesis of which I discussed in my post on that work); Shair

Posted at 12am on 24/11/08 | 39 comments | Filed Under: Still In The RBF
read on

Latest comments

TB Shikimic acid is made ver..(Go)
Jose milkshake- chemspider mig..(Go)
sjb Re facelift. Whilst I ..(Go)
anon milkshake - we're stuck w..(Go)
milkshake Now that we are in full-b..(Go)
Flower Thank you but in this App..(Go)
cvengo yeah, but then there are ..(Go)
anniechem cvengo - you're looking a..(Go)
cvengo one thing that is annoyin..(Go)

In Other News...

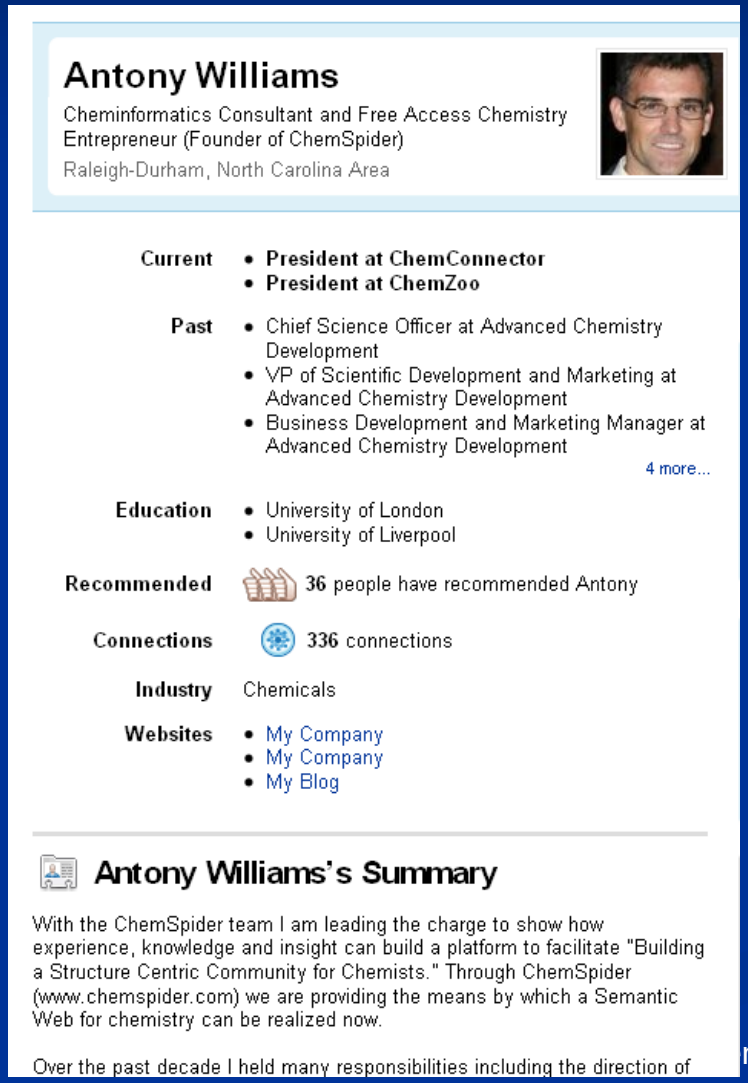
K. A. Woerpel *et al.* Mechanisms for nucleophilic substitutions of cyclic acetals: nucleophile strength versus stereoselectivity. 10.1021/ol8019956

P. A. Evans *et al.* Intermolecular rhodium-catalyzed [3+2+2] cyclization of

Blogs

Carbon-Based Curiosities
Curly Arrow
Dylan's Tenderblog
In The Pipeline
KinasePro
Lamentations on Chemistry
Liquid Carbon
Molecule Of The Day
One In Ten-Thousand

Social Networking for Chemists
Blogs are the start



Antony Williams

Cheminformatics Consultant and Free Access Chemistry Entrepreneur (Founder of ChemSpider)

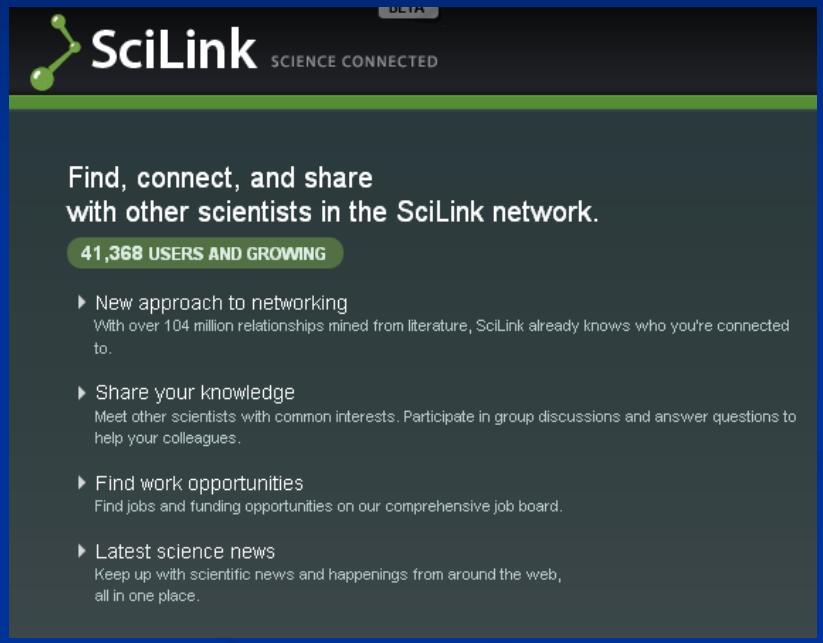
Raleigh-Durham, North Carolina Area

Current	• **President at ChemConnector** • **President at ChemZoo**
Past	• Chief Science Officer at Advanced Chemistry Development • VP of Scientific Development and Marketing at Advanced Chemistry Development • Business Development and Marketing Manager at Advanced Chemistry Development 4 more...
Education	• University of London • University of Liverpool
Recommended	**36** people have recommended Antony
Connections	**336** connections
Industry	Chemicals
Websites	• My Company • My Company • My Blog

Antony Williams's Summary

With the ChemSpider team I am leading the charge to show how experience, knowledge and insight can build a platform to facilitate "Building a Structure Centric Community for Chemists." Through ChemSpider (www.chemspider.com) we are providing the means by which a Semantic Web for chemistry can be realized now.

Over the past decade I held many responsibilities including the direction of



ntric Comm

Collaborative Knowledge Management for Chemists – Wikipedia, Built by a Network

Help us improve Wikipedia by **supporting it financially**.

Log in / create account

article | discussion | edit this page | history

WIKIPEDIA
The Free Encyclopedia

navigation
- Main Page
- Contents
- Featured content
- Current events
- Random article

interaction
- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

search
Go | Search

toolbox
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Paclitaxel

From Wikipedia, the free encyclopedia
(Redirected from Taxol)

Paclitaxel is a mitotic inhibitor used in cancer chemotherapy. It was discovered in a National Cancer Institute program at the Research Triangle Institute in 1967 when Monroe E. Wall and Mansukh C. Wani isolated it from the bark of the Pacific yew tree, *Taxus brevifolia* and named it 'taxol'. When Bristol-Myers Squibb (BMS) licensed the compound for sale they claimed rights to the name as well and people responded by referring to the generic name as 'paclitaxel', but BMS later lost the court case about naming rights. In this formulation paclitaxel is dissolved in Cremophor EL, a polyoxyethylated castor oil, as a delivery agent since paclitaxel is not soluble in water. A newer formulation, in which paclitaxel is bound to albumin as the delivery agent (Protein-bound paclitaxel), is sold commercially by Abraxis BioScience under the trademark Abraxane.[2]

Paclitaxel is now used to treat patients with lung, ovarian, breast cancer, head and neck cancer, and advanced forms of Kaposi's sarcoma. Paclitaxel is also used for the prevention of restenosis.

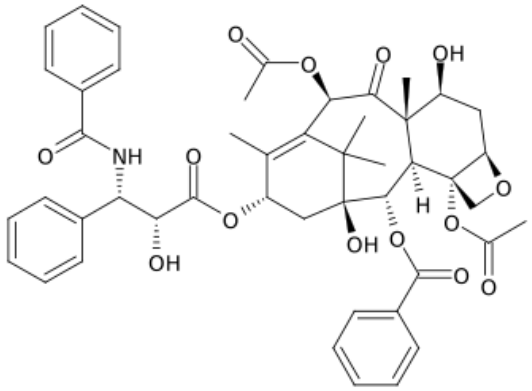
Paclitaxel works by interfering with normal microtubule breakdown during cell division. Together with docetaxel, it forms the drug category of the taxanes. It was the subject of a notable total synthesis by Robert A. Holton.

As well as offering substantial improvement in patient care, paclitaxel has been a relatively controversial drug. There was originally concern because of the environmental impact of its original sourcing, no longer used, from the Pacific yew. The assignment of rights, and even the name itself, to BMS were the subject of public debate and Congressional hearings.

Contents [hide]

1 History
- 1.1 The plant screening program, isolation, and preclinical trials
- 1.2 Early clinical trials, supply and the transfer to BMS

2 Production



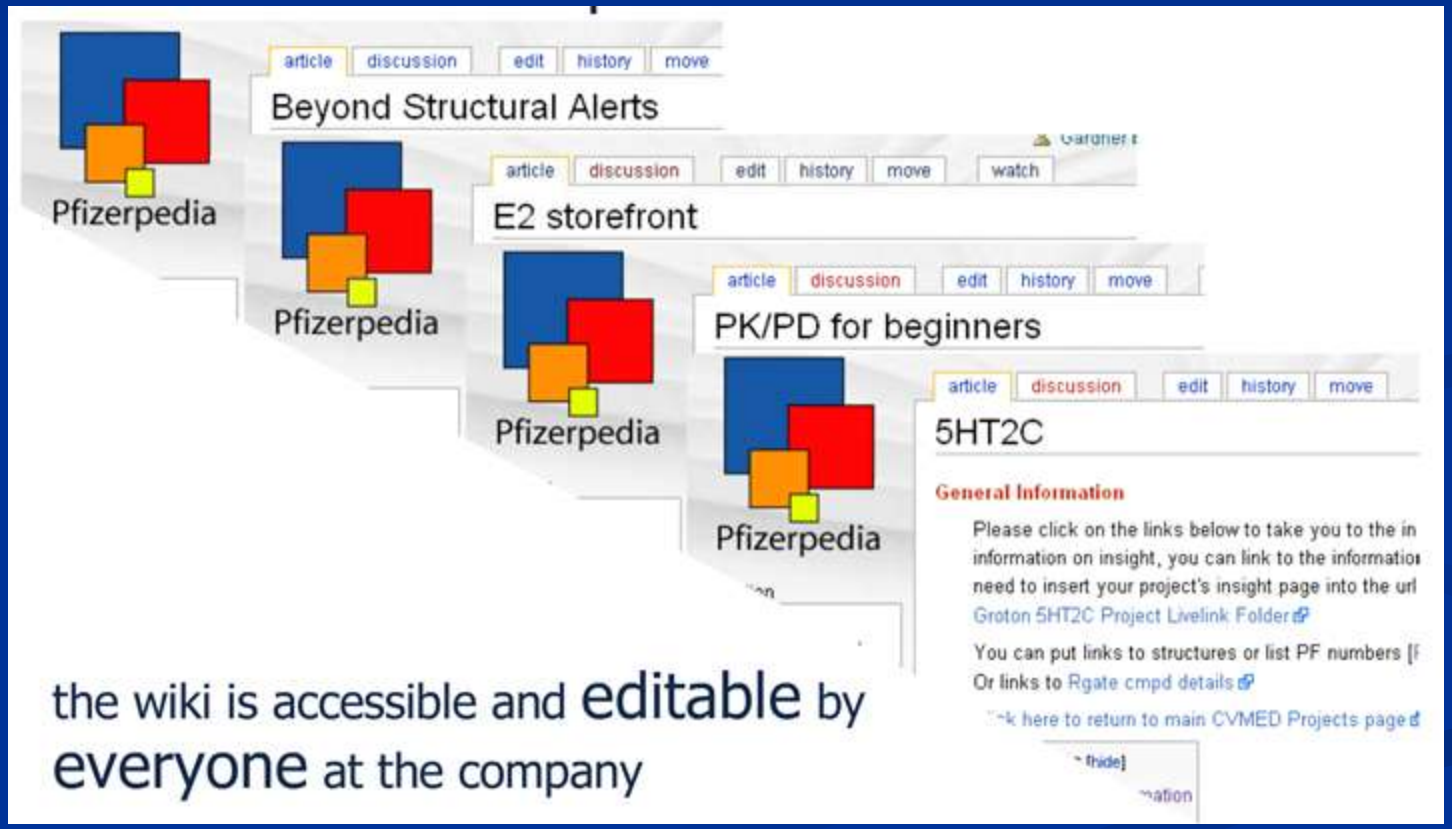
Paclitaxel

Systematic (IUPAC) name
(1*S*,2*S*,3*R*,4*S*,7*R*,9*S*,10*S*,12*R*,15*S*) -4,12-Diacetoxy-15-{[(2*R*,3*S*)-3-(benzoylamino)-2-hydroxy-3- phenylpropanoyl]oxy}-1,9- dihydroxy-10,14,17,17-tetramethyl -11-oxo-6-oxatetracyclo [11.3.1.0~3,10~.0~4,7~] heptadec-13-en-2-yl benzoate
Identifiers
CAS number 33069-62-4
ATC code L01CD01

Building a Structure Centric Community for Chemists

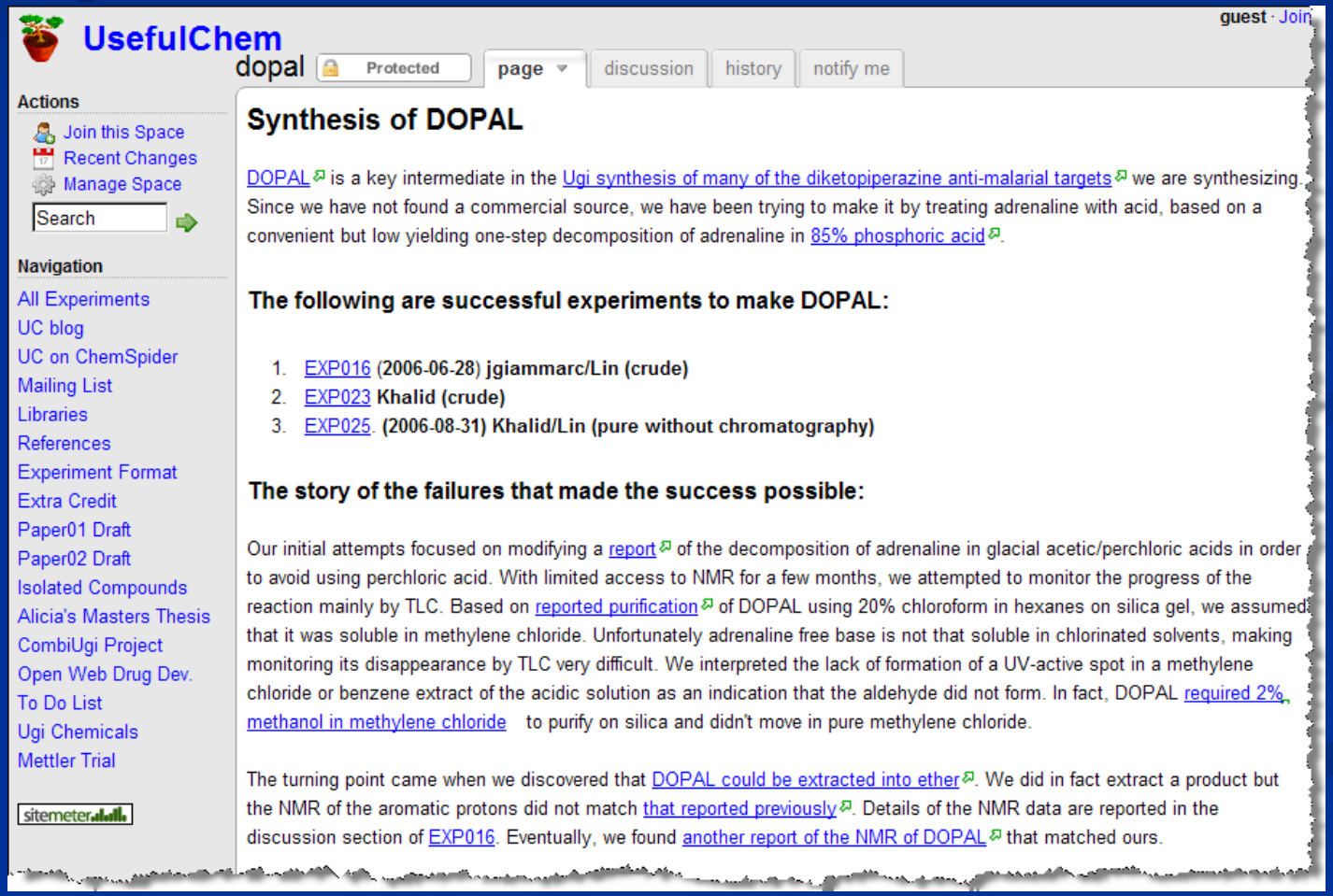
Collaborative Authoring for Drug Discovery

- Pfizerpedia



Collaborative Authoring in Academia

- Group level collaboration via Wikis



UsefulChem

guest · Join

dopal | Protected | page | discussion | history | notify me

Actions

- Join this Space
- Recent Changes
- Manage Space

Search

Navigation

- All Experiments
- UC blog
- UC on ChemSpider
- Mailing List
- Libraries
- References
- Experiment Format
- Extra Credit
- Paper01 Draft
- Paper02 Draft
- Isolated Compounds
- Alicia's Masters Thesis
- CombiUgi Project
- Open Web Drug Dev.
- To Do List
- Ugi Chemicals
- Mettler Trial

sitemeter

Synthesis of DOPAL

DOPAL is a key intermediate in the Ugi synthesis of many of the diketopiperazine anti-malarial targets we are synthesizing. Since we have not found a commercial source, we have been trying to make it by treating adrenaline with acid, based on a convenient but low yielding one-step decomposition of adrenaline in 85% phosphoric acid.

The following are successful experiments to make DOPAL:

1. EXP016 (2006-06-28) jgiammarc/Lin (crude)
2. EXP023 Khalid (crude)
3. EXP025. (2006-08-31) Khalid/Lin (pure without chromatography)

The story of the failures that made the success possible:

Our initial attempts focused on modifying a report of the decomposition of adrenaline in glacial acetic/perchloric acids in order to avoid using perchloric acid. With limited access to NMR for a few months, we attempted to monitor the progress of the reaction mainly by TLC. Based on reported purification of DOPAL using 20% chloroform in hexanes on silica gel, we assumed that it was soluble in methylene chloride. Unfortunately adrenaline free base is not that soluble in chlorinated solvents, making monitoring its disappearance by TLC very difficult. We interpreted the lack of formation of a UV-active spot in a methylene chloride or benzene extract of the acidic solution as an indication that the aldehyde did not form. In fact, DOPAL required 2% methanol in methylene chloride to purify on silica and didn't move in pure methylene chloride.

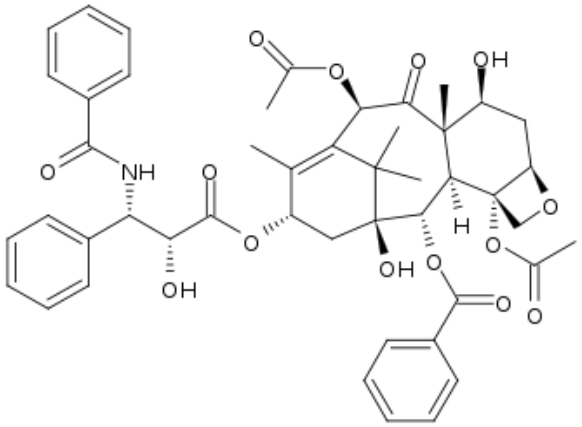
The turning point came when we discovered that DOPAL could be extracted into ether. We did in fact extract a product but the NMR of the aromatic protons did not match that reported previously. Details of the NMR data are reported in the discussion section of EXP016. Eventually, we found another report of the NMR of DOPAL that matched ours.

Wikis for Science

- Who in the room hasn't used Wikipedia?
- Is it trustworthy?
- What are the advantages and disadvantages of the Wiki environment?
- How suitable is it for Chemistry?

Wikipedia Chemistry Curation project

- Only ca. 5000 organic structures
- A year of work for a team of 6 people
- Many errors removed in the process.
- Slow and torturous process
- CAS collaborating in the process



Paclitaxel

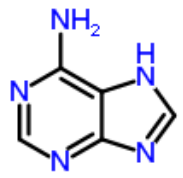
Systematic (IUPAC) name	
(1*S*,2*S*,3*R*,4*S*,7*R*,9*S*,10*S*,12*R*,15*S*) -4,12-Diacetoxy-15-{[(2*R*,3*S*)-3-(benzoylamino)-2-hydroxy-3- phenylpropanoyl]oxy}-1,9- dihydroxy-10,14,17,17-tetramethyl -11-oxo-6-oxatetracyclo [11.3.1.0~3,10~.0~4,7~] heptadec-13-en-2-yl benzoate	
Identifiers	
CAS number	33069-62-4
ATC code	L01CD01
PubChem	36314
DrugBank	APRD00259
ChemSpider	10368587
Chemical data	
Formula	$C_{47}H_{51}NO_{14}$
Mol. mass	853.906 g/mol

Wikipedia via ChemSpider…

Inherent Properties, Identifiers and References

2D | 3D

Quick Links: Permalink Similar Isomers



load | save | zoom

ChemSpider ID:	185
Empirical Formula:	$C_5H_5N_5$
Molecular Weight:	135.1267
Nominal Mass:	135 Da
Average Mass:	135.1267 Da
Monoisotopic Mass:	135.054495 Da

Systematic Name:	7H-purin-6-amine
SMILES:	n1c(c2c(nc1)ncn2)N
InChI:	InChI=1/C5H5N5/c6-4-3-5(9-1-7-3)10-2-8-4/h1-2H,(H3,6,7,8,9,10)
InChIKey:	GFFGJBXGBJISGV-UHFFFAOYAT

Wikipedia Article(s)

License

Adenine is a purine with a variety of roles in biochemistry including cellular respiration, in the form of both the energy-rich adenosine triphosphate (ATP) and the cofactors nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD), and protein synthesis, as a chemical component of DNA and RNA. the shape of adenine is complementary to either thymine or uracil. Read more... or Edit at Wikipedia...

Associated Data Sources and Commercial Suppliers

Filter

Supplemental Information

Disclaimer

Links & References

Eust quio et al.. Discovery and characterization of a marine bacterial SAM-dependent chlorinase, *Nature Chemical Biology, doi: 10.1038/NChemBio.2007.56, published online 2 December 2007.*
[DOI: 10.1038/NChemBio.2007.56]

E. L. Willighagen, H. M. G. W. Denissen, R. Wehrens, and L. M. C. Buydens. On the Use of 1H and 13C 1D NMR Spectra as QSPR Descriptors, *J. Chem. Inf. Model., 46 (2), 487-494, 2006*
[PubMed: 16562976] [DOI: 10.1021/ci050282s]

Collaborative Drug Discovery, Inc.
$1.9M from the Gates Foundation

CDD COLLABORATIVE DRUG DISCOVERY

Username

Home | Product Overview | Who Uses CDD? | Academics | Industry | Foundations | Privacy & Security | Pul

Home /

Public Access

CDD hosts Public Access Data relevant to drug discovery from leading research groups around the world.

Scientists who wish to view or mine CDD's repository of Public Access Data can register for a free account.

Scientists who wish to contribute Public Access Data should contact us.

	Shared Publicly	Description	Group	Molecules
☐	**FDA/Orphan Drugs** PI: Christopher Lipinski Published: 10/26/2007	FDA approved drugs with designated indications, sponsor name and chemical structures (when available)	Known drugs	1721
☐	**TB: TAACF Assay Results** PI: Bernard Munos Published: 2/12/2008	Antibacterial activity of a publicly available library of 812 compounds against Mycobacterium tuberculosis (H37Rv) in	TB Early Phase Drug Discovery Program	812

The Quality of Data Online…

- Content is king – quality costs. Curation is expensive!
- Data online are "filthy".
  - Gathering data is the "easy part"
  - Structures are COMMONLY incorrect
- Informatics tools exist already
  - Hold millions of structures and associated data
  - Structure/substructure/text searching
  - Data downloads, data uploads, editing, annotation

Rich Online Data Resources for Chemists and the Life Sciences

- PubChem
- Pubmed
- Wikipedia
- ChemSpider
- Drugbank
- ChEBI
- ChemIDPlus
- DailyMed
- And many more…

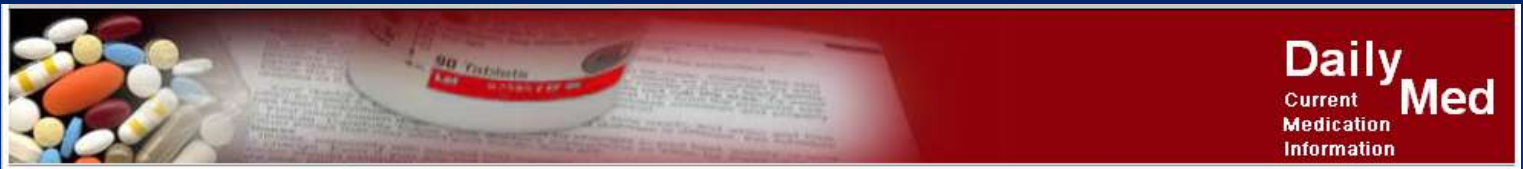
PubChem

PubChem

From Wikipedia, the free encyclopedia

PubChem is a database of chemical molecules. The system is maintained by the National Center for Biotechnology Information (NCBI), a component of the National Library of Medicine, which is part of the United States National Institutes of Health (NIH). PubChem can be accessed for free through a web user interface. Millions of compound structures and descriptive datasets can be freely downloaded via FTP. PubChem contains substance descriptions and small molecules with fewer than 1000 atoms and 1000 bonds. The American Chemical Society tried to get the U.S. Congress to restrict the operation of PubChem, because they claim it competes with their Chemical Abstracts Service.[1]. More than 80 database vendors contribute to the growing PubChem database.[2]

DailyMed



Options

- Home
- E-mail Label Information
- Downloads
- SPL History
- Print this Label
- Notify of Updates
- Contact Us

Additional Resources

- Report Adverse Event
- MedlinePlus Information
- Find Clinical Trials
- Biochemical Data Summary
- Search PubMed Articles
- Presence in Breast Milk

Download the FDA official PDF of this label

Search By Drug Name: GO

RxNorm Names

Review RxNorm Normal Forms

Tolinase (tolazamide) Tablet
[Pharmacia and Upjohn Company]

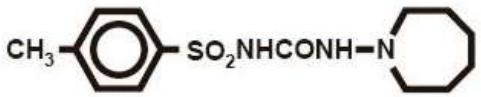
Drug Label Sections

Description | Clinical Pharmacology | Indications & Usage | Contraindications | Warnings | Precautions
Adverse Reactions | Overdosage | Dosage & Administration | How Supplied | Patient Counseling Information
Supplemental Patient Material | Boxed Warning | Patient Package Insert | Highlights | Full Table of Contents

DESCRIPTION

TOLINASE Tablets contain tolazamide, an oral blood glucose lowering drug of the sulfonylurea class. Tolazamide is a white or creamy-white powder with a melting point of 165° to 173° C. The solubility of tolazamide at pH 6.0 (mean urinary pH) is 27.8 mg per 100 mL.

The chemical names for tolazamide are (1) Benzenesulfonamide, *N*-[[(hexahydro-1*H*-azepin-1-yl) amino] carbonyl]-4-methyl-; (2) 1-(Hexahydro-1*H*-azepin-1-yl)-3-(*p*-tolylsulfonyl)urea and its molecular weight is 311.40. The structural formula is represented below:



TOLINASE Tablets for oral administration are available as scored, white tablets containing 100 mg, 250 mg or 500 mg tolazamide. Inactive ingredients: calcium sulfate, docusate sodium, magnesium stearate, methylcellulose, sodium alginate.

Caution! Question Everything!

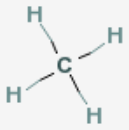
ASN.1 XML SDF

Charcoal - Compound Summary (CID: 297)

An amorphous form of carbon prepared from the incomplete combustion of animal or vegetable matter, e.g., wood. The activated form of charcoal is used in the treatment of poisoning. (Grant & Hackh's Chemical Dictionary, 5th ed)

Table of Contents

- Drug and Chemical Information
  - Pharmacological Action
  - Pharmacological Classification
  - Chemical Classification
  - Safety and Toxicology
  - Literature Links
  - Literature Mining
- Synonyms
- Properties
- Descriptors
- Compound Information
- Substance Information
  - Category
- Exports



Compound ID	**297**
Molecular Weight	16.04246 [g/mol]
Molecular Formula	CH_4
H-Bond Donor	0
H-Bond Acceptor	0

Plumbago (graphite)
Carbon-12
Philblack N 550
Philblack N 765
DIAMOND
Monarch 700
Witcarb 940
Graphite (synthetic)
Irgalite 1104

METHANAL, OXOMETHANE, OXYMETHYLENE, METHYLENE OXIDE, FORMIC ALDEHYDE, METHYL ALDEHYDE

1-Chlorobenzylethyl-3,5,7,9,11,13,15-heptaisobutylpentacyclo[9.5.1.1(3,9).1(5,15).1(7,13)]octasiloxane, mixture of isomers

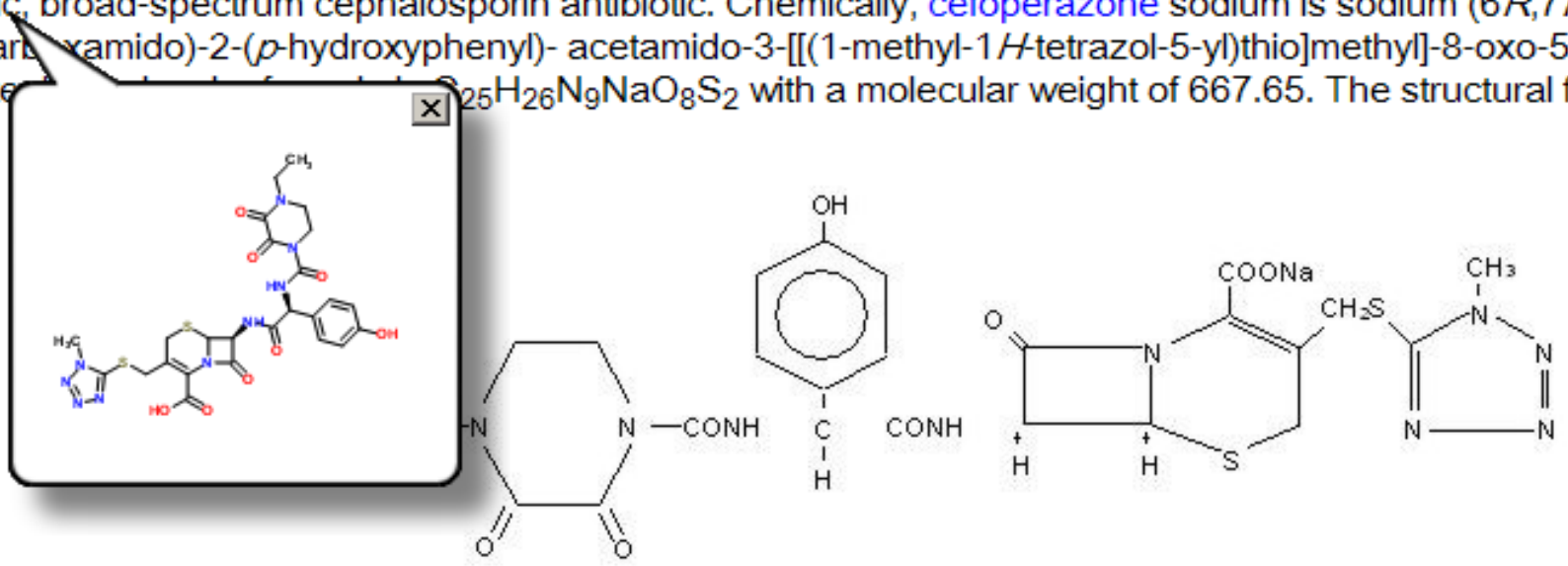
Question Everything
www.dhmo.org

Quality of Structures!!!

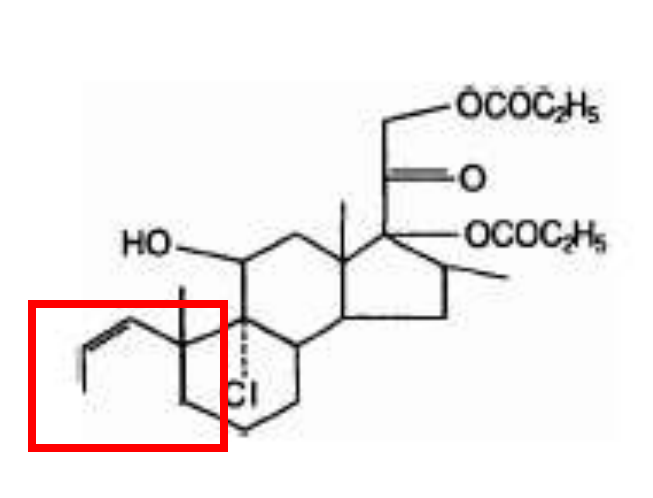
Cefobid (cefoperazone) Powder, For Solution
[Roerig]

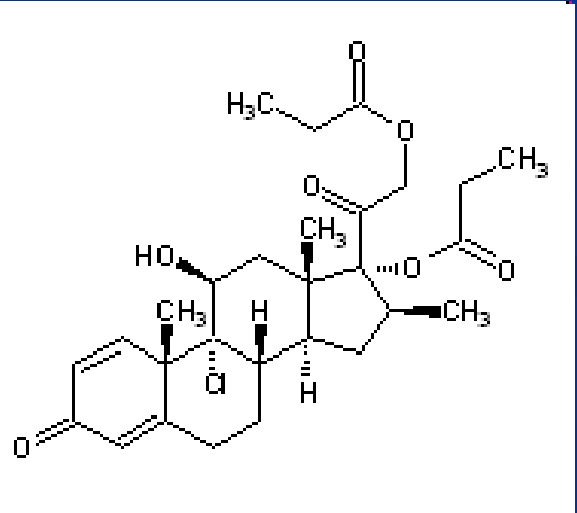
DESCRIPTION

CEFOBID® (sterile cefoperazone), formerly known as sterile cefoperazone sodium, contains cefoperazone a
semisynthetic, broad-spectrum cephalosporin antibiotic. Chemically, cefoperazone sodium is sodium (6*R*,7*R*)
piperazinecarb xamido)-2-(*p*-hydroxyphenyl)- acetamido-3-[[(1-methyl-1*H*-tetrazol-5-yl)thio]methyl]-8-oxo-5-th
2-carboxylate $C_{25}H_{26}N_9NaO_8S_2$ with a molecular weight of 667.65. The structural for



Quality of Structures





Crowdsourcing

- Chemistry databases enhanced by crowdsourcing
- Chemistry databases can be connected to articles, vendors, properties, spectra, etc.
- A platform for deposition, curation and distribution ?
- This is the future… existing business models are at risk

Wendy Warr

STM on the advance

Wendy Warr scours the scientific, technical and medical sector in search of innovation

By Wendy Warr 05 Dec 2008

- "…some publishers are responding vigorously to market forces, but **the steady growth of free information resources is a real threat to them**."

http://www.iwr.co.uk/information-world-review/features/2232039/stm-advance

Trademark Infringement But Real Competition…

ACS Takes Legal Action Against Google

Google search service is said to infringe SciFinder Scholar trademark

AALOK MEHTA

The American Chemical Society filed a complaint on Dec. 9 against Google Inc. in U.S. District Court for the District of Columbia. The complaint contends that Google's use of the trademark "Scholar" for its Google Scholar literature-search engine constitutes trademark infringement and unfair competition.

A beta version of Google Scholar (http://scholar.google.com) debuted in mid-November. The search service allows users, at no cost, to "search specifically for scholarly literature, including peer-reviewed papers, theses, books, preprints, abstracts, and technical reports from all broad areas of research," according to a Google website.

http://publicaccess.nih.gov/

National Institutes of Health Public Access

The Public Access Policy ensures that the public has access to the published results of NIH funded research to help advance science and improve human health.

Address Copyright

Before you sign a publication agreement or similar copyright transfer agreement, **make sure that the agreement allows the paper to be submitted to NIH** in accordance with the Public Access Policy.

Publishers and Open Access

NIH "open access" policy causing publishing companies angst

By John Timmer | Published: July 28, 2008 - 11:16PM CT

"It's clear that the academic publishing world is in a state of flux. Nobody's quite figured out **how to make an open access business model work**, but even most publishers recognize that the public and scientific community benefit from having access to the research they've paid for. "

Chemistry Publishing and "Structures"???

PubChem

PubChem

From Wikipedia, the free encyclopedia

PubChem is a database of chemical molecules. The system is maintained by the National Center for Biotechnology Information (NCBI), a component of the National Library of Medicine, which is part of the United States National Institutes of Health (NIH). PubChem can be accessed for free through a web user interface. Millions of compound structures and descriptive datasets can be freely downloaded via FTP. PubChem contains substance descriptions and small molecules with fewer than 1000 atoms and 1000 bonds. The American Chemical Society tried to get the U.S. Congress to restrict the operation of PubChem, because they claim it competes with their Chemical Abstracts Service.[1]. More than 80 database vendors contribute to the growing PubChem database.[2]

InChIs
Structure but NOT substructure

Google InChI=1/C8H9N5O/c1-3-4(2)11-6-5(10-3)7(14)13-8(9)12-6/h1-2 Search Advanced Search Preferences

Web Results 1 - 10 of about 64 for InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6). (0.72 seconds)

(2-aminoethyl)phosphonic acid (CHEBI:15573)
InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H. InChI=1/C2H8NO3P/c3-1 -2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H ...
www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:15573 - 37k - Cached - Similar pages -

339 -OEChem-12040710492D 15 14 0 0 0 0 0 0 0999 V2000 3.4030 ...
... 0 1 4 2 0 0 0 0 1 6 1 0 0 0 0 2 14 1 0 0 0 0 3 15 1 0 0 0 0 5 7 1 0 0 0 0 5 12 1 0 0 0 0 5 ...
InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H ...
www.bmrb.wisc.edu/metabolomics/standards/2_Aminoethylphosphonic_acid/lit/SID_8143201.sdf - 4k - Cached - Similar pages -

data_bmse000309 save_entry_information _Entry.Sf_category ...
InChi_code ; **InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H** ;
_Chem_comp.Mon_nstd_flag ? _Chem_comp.Std_deriv_one_letter_code ? _Chem_comp. ...
www.bmrb.wisc.edu/metabolomics/standards/2_Aminoethylphosphonic_acid/nmr/bmse000309/bmse000309.str - 33k - Cached - Similar pages -
More results from www.bmrb.wisc.edu »

Metabolomics At NMRFAM: records
InChI: **InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)**. Experimental Water Solubility:. Predicted Water Solubility: 35.25 mg/ml [Predicted by ALOGPS] ...
mmcd.nmrfam.wisc.edu/test/cqsearch.py?cqid=cq_02114 - 12k - Cached - Similar pages -

Bacillus anthracis 2-aminoethylphosphonate
InChI: **InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)**. Unification Links: LIGAND:C03557 , PubChem:339. In Pathway Reactions as a Reactant: ...
pathways.jcvi.org/GB21/NEW-IMAGE?object=CPD-1106 - 8k - Cached - Similar pages -

Ciliatine - PubChem Public Chemical Database
IUPAC Name: 2-aminoethylphosphonic acid. Canonical SMILES: C(CP(=O)(O)O)N InChI: **InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H**. Compound Info: ...
pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=339 - Similar pages -

2-aminoethylphosphonic acid
InChI: **InChI=1/C2H8NO3P/c3-1-2-7(4,5)6/h1-3H2,(H2,4,5,6)/f/h4-5H**. InChIKey: InChIKey=QQVDJLLNRSOCEL-NUMVZRSTCW SMILES: C(CP(=O)(O)O)N ...
sci-toys.com/scichem/jqp006/339.html - 13k - Cached - Similar pages -

The InChI Resolver

RSC and ChemSpider develop InChI Resolver

01 December 2008

An InChI Resolver, a unique free service for scientists to share chemical structures and data, will be developed by a collaboration between ChemZoo Inc., host of ChemSpider, and the Royal Society of Chemistry.

Using the InChI - an IUPAC standard identifier for compounds - scientists can share and contribute their own molecular data and search millions of others from many web sources. The RSC/ChemSpider InChI Resolver will give researchers the tools to create standard InChI data for their own compounds, create and use search engine-friendly InChIKeys to search for compounds, and deposit their data for others to use in the future.

The future of publishing

'The wider adoption and unambiguous use of the InChI standard will be an important development in the way chemistry is published in the future, and the further development of the semantic web,' comments Robert Parker, Managing Director of RSC Publishing.

Peer Review and Wikis
Peter Frishauf, founder of Medscape

- "Andrew Grove, … Intel Corporation, **likens traditional peer-review systems to Middle Ages guilds**. He calls for "cultural revolution" in publishing to reinvent peer review."

- "That revolution will emerge as **a variant of Wikipedia**. Medical publishing, peer review, research, patient care, and commerce will be transformed. And for the better."

Conclusions

- The internet enables chemistry – and at a reduced cost
- Web 2.0 is here and improving quality – to benefit 3.0
- Question Quality!
- Crowdsourcing for expansion, curation and integration
- Classical models may die quite quickly – business models must change soon or fail
- Publishers – heed the profileration of InChIs for Chemistry

The End of Traditional Publishing

- Peter Frishauf, **The Medscape Journal of Medicine** makes two predictions
  - Within 5 years, most medical journals will be **open-access**. […] provide access to trusted articles and data at no cost.
  - Peer review as we know it will disappear. Rather than the secretive prepublication review process followed by most publishers today, *including Medscape*, **most peer review will occur transparently, and *after publication***.

The ChemSpider Journal – 12/2008
www.chemspider.com

Research In Progress: Predicting Potential Endogenous/Exogenous P-gp Substrates.

Sean Ekins [‡, §, *]

[‡]Collaborations in Chemistry, Jenkintown, PA 19046, USA; §Department of Pharmaceutical Sciences, University of Maryland, Baltimore, MD 21202, USA, [*] Department of Pharmacology, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA.

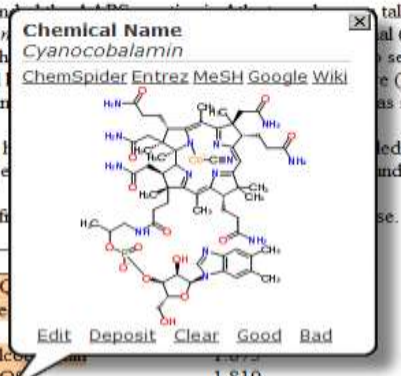
Research

I recently atten talk in a session chaired by Dr. James Polli on *In vitro oral Drug transport data: approaches to data analysis and* al (Q)SAR Chemists Perspective". At the end of which I was asked a question regarding whether I had used the various publish search for endogenous substrates and inhibitors. I briefly mentioned earlier in July I had used the previously published and (Ekins et al. 2002; Chang et al. 2006) to search a multiconformer version of the human metabolome database (http://www.h as searched and created in Catalyst Discovery Studio 2.0 and the results may be of interest to a wider audience.

The following led in Table 1. Including the well known P-gp substrate digoxin (actually used in deriving the model (Ekins et al. 2002; Chang e nd was however Coenzyme Q10 (Figure 1).

Table 1. Hits f se.

Name	
Coenzyme	
Lacto-N-de	
Digoxin	
Glutathionylc	
Coenzyme Q	1.819
Ubiquinol	1.809
Cyanocobalamin	0.855
Ubiquinone Q4	0.242



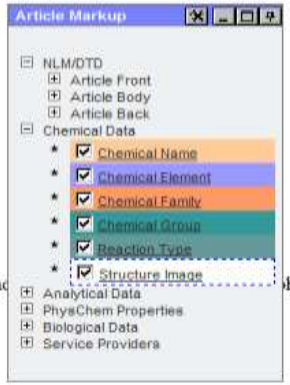
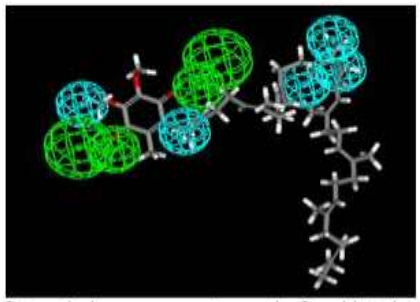


Figure 1. P-gp substrate pharmacophore - showing the mapping of Coenzyme Q10. Pharma hobes (cyan) and hydrogen bond acceptors (green).



Interestingly a very recent paper by Itagaki et al., suggests the top scoring Coenzyme Q10 interacts with P-gp (see PubMed abstract below).

The Story of NAPE

Hi:
I find that newsworthy molecules can make good test cases for searching. Over the years I find gradually better results, yet the failures continue to appear and expose the difficulties in chem nomenclature that I know you are working hard on. I think you'd like to know about this one.

So a news piece comes out about NAPE: N-acylphosphatidylethanolamine
http://www.hhmi.org/news/shulman20081126.html

so I search chemspider and get nothing, and in pubchem get 17 million - gee, which is worse :)

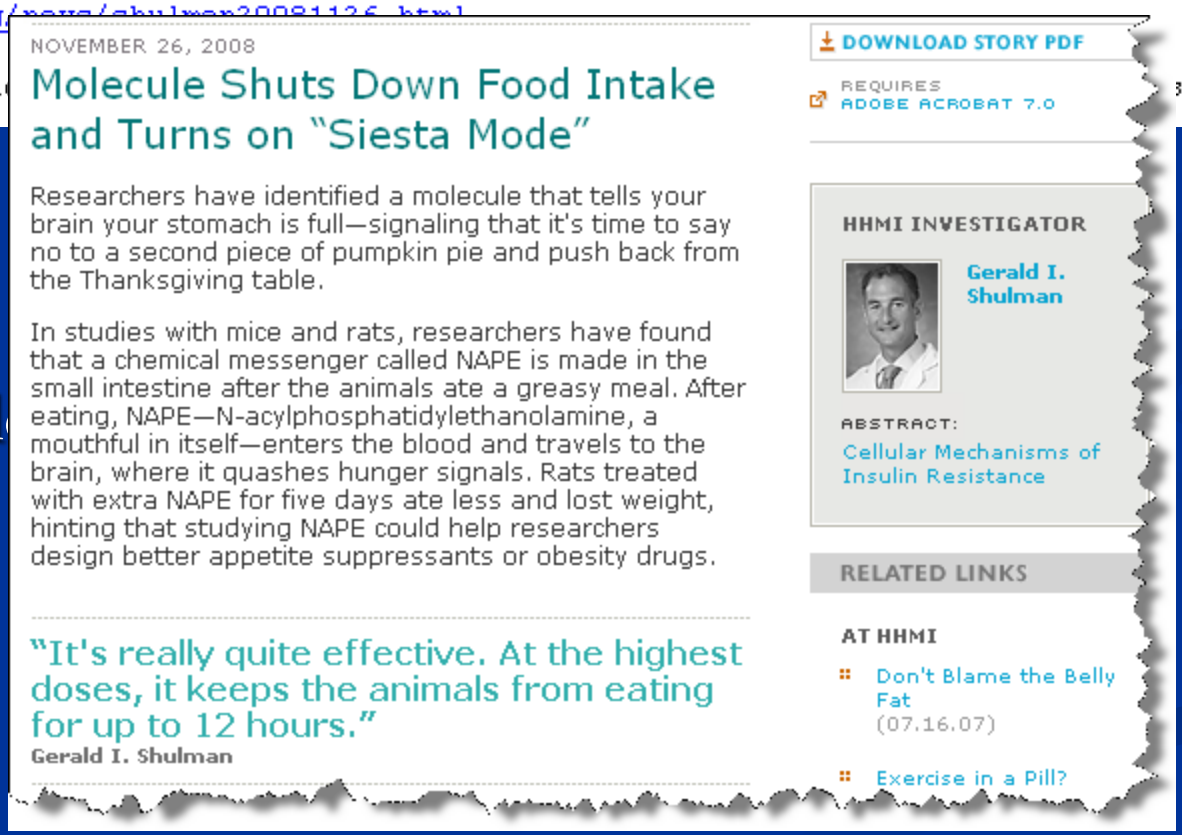
NAPE

Hi:
I find that newsworthy molecules can make good test cases for searching. Over the years I find gradually better results, yet the failures continue to appear and expose the difficulties in chem nomenclature that I know you are working hard on. I think you'd like to know about this one.

So a news piece comes out about NAPE: N-acylphosphatidylethanolamine
http://www.hhmi.org/news/shulman20081126.html

so I search chemspi worse :)

- Googl



NOVEMBER 26, 2008

Molecule Shuts Down Food Intake and Turns on "Siesta Mode"

Researchers have identified a molecule that tells your brain your stomach is full—signaling that it's time to say no to a second piece of pumpkin pie and push back from the Thanksgiving table.

In studies with mice and rats, researchers have found that a chemical messenger called NAPE is made in the small intestine after the animals ate a greasy meal. After eating, NAPE—N-acylphosphatidylethanolamine, a mouthful in itself—enters the blood and travels to the brain, where it quashes hunger signals. Rats treated with extra NAPE for five days ate less and lost weight, hinting that studying NAPE could help researchers design better appetite suppressants or obesity drugs.

"It's really quite effective. At the highest doses, it keeps the animals from eating for up to 12 hours."
Gerald I. Shulman

DOWNLOAD STORY PDF
REQUIRES ADOBE ACROBAT 7.0

HHMI INVESTIGATOR

Gerald I. Shulman

ABSTRACT:
Cellular Mechanisms of Insulin Resistance

RELATED LINKS

AT HHMI

- Don't Blame the Belly Fat (07.16.07)
- Exercise in a Pill?

NAPE

Google Search of : “chemical structure of N-acylphosphatidylethanolamine”

Did you mean: chemical structure of N-*acyl phosphatidylethanolamine*

phosphatidylethanolamine, **N-acylphosphatidylethanolamine**, N ...
Phosphatidylethanolamine – **Structure** and Occurrence .... The activation of **N-acylphosphatidylethanolamine** metabolism in plants seems to be associated with ...
www.lipidlibrary.co.uk/Lipids/pe/index.htm - 26k - Cached - Similar pages -

Molecular composition of the **N-acylphosphatidylethanolamine** of the ...
Its **structure** as an **N-acylphosphatidylethanolamine** (N-acyl-PE) was confirmed by .... Pinus sylvestris L. It has been established by **chemical** and spectral ...
www.springerlink.com/index/W87L34J382582111.pdf - Similar pages -
by AS Isamukhamedov - 1980 - Related articles

The "Lipid Library"

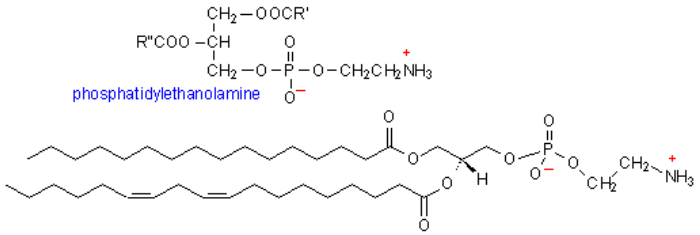
The Lipid Library

PHOSPHATIDYLETHANOLAMINE AND RELATED LIPIDS

STRUCTURE, OCCURRENCE, BIOCHEMISTRY and ANALYSIS

1. Phosphatidylethanolamine – Structure and Occurrence

Phosphatidylethanolamine (once given the trivial name 'cephalin') is usually the second most abundant phospholipid in animal and plant lipids and it is frequently the main lipid component of microbial membranes. It can amount to 20% of liver phospholipids and as much as 45% of those of brain; higher proportions are found in mitochondria than in other organelles. As such, it is obviously a key building block of membrane bilayers. It is a neutral or zwitterionic phospholipid (at least in the pH range 2 to 7) with the structure shown (with one specific molecular species illustrated as an example).



phosphatidylethanolamine

1-hexadecanoyl, 2-(9*Z*,12*Z*-octadecadienoyl)-*sn*-glycero-3-phosphoethanolamine

In animal tissues, phosphatidylethanolamine tends to exist in diacyl, alkylacyl and alkenylacyl forms, and data for the compositions of these various forms from bovine heart muscle are listed in our web pages on **ether lipids**. In addition, as much as 70% of the phosphatidylethanolamine in some cell types (inflammatory cells, neurons and tumor cells) can have an ether linkage.

In general, animal phosphatidylethanolamine tends to contain higher proportions of arachidonic and docosahexaenoic acids than the other zwitterionic phospholipid, phosphatidylcholine. These polyunsaturated components are concentrated in position *sn*-2 with saturated fatty acids most abundant in position *sn*-1, as illustrated for rat liver and chicken egg in **Table 1**. In most other species, it would be expected that the structure of the phosphatidylethanolamine in the same metabolically active tissues would exhibit similar features.

Table 1. Positional distribution of fatty acids in phosphatidylethanolamine in animal tissues.

Position	Fatty acid						
	14:0	16:0	18:0	18:1	18:2	20:4	22:6
Rat liver [1]							

Wikipedia…

N-acylphosphatidylethanolamines

From Wikipedia, the free encyclopedia

N-acylphosphatidylethanolamines (NAPEs) are hormones released by the small intestine into the bloodstream when it processes fat. It travels to the hypothalamus in the brain and suppresses appetite. This could make it useful for treating obesity.[1]

References [edit]

1. ^ "Gut chemical may inspire new way to fight obesity | Science | Reuters". Retrieved on 2008-11-27.

Category: Hormones

Original Source – Full loop

Gut chemical may inspire new way to fight obesity

Wed Nov 26, 2008 3:29pm EST

Email | Print | Share | Reprints | Single Page | Recommend (1) [-] Text [+]

1 of 1 Full Size

By Will Dunham

WASHINGTON (Reuters) - Scientists have identified a fatty substance made in the gut that signals the brain when it's time to stop eating -- a discovery that could inspire new approaches to fighting obesity.

Writing in the journal Cell on Wednesday, U.S. researchers said experiments with mice and rats showed that a naturally occurring fat-derived chemical messenger called NAPE regulated how much the animals ate. It is present in people and may do the same thing, they said.

Gerald Shulman of Yale University and the Howard Hughes Medical Institute and colleagues said that when the rodents were fed a fatty meal, their small intestine made a lot of NAPE and put it into the bloodstream. It then traveled to the brain and shut down hunger signals, they said.

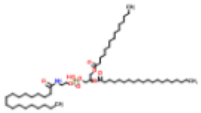
And now a structure…

Add: CIF Spectrum Image Comments

Inherent Properties, Identifiers and References

2D 3D

Quick Links: Permalink Similar Isomers



load save zoom

ChemSpider ID:	21427282
Empirical Formula:	$C_{58}H_{114}NO_9P$
Molecular Weight:	1000.5008
Nominal Mass:	999 Da
Average Mass:	1000.5008 Da
Monoisotopic Mass:	999.82312 Da

Systematic Name:	[(1R)-1-[[hydroxy-[2-(octadecanoylamino)ethoxy]phosphoryl]oxymethyl]-2-pentadecanoyloxy-ethyl] icosanoate
SMILES:	O=C(O[C@@H](COP(O)(=O)OCCNC(=O)CCCCCCCCCCCCCCCCC)COC(=O)CCCCCCCCCCCCCCCCCCCCC
InChI:	InChI=1/C58H114NO9P/c1-4-7-10-13-16-19-22-25-27-28-30-32-35-38-41-44-47-50-58(62)68-55(53-65-57(61)49-46-43-40-37-34-24-21-18-15-12-9-6-3)54-67-69(63,64)66-52-51-59-56(60)48-45-42-39-36-33-31-29-26-23-20-17-14-11-8-5-2/h55H,4-54H2,1-3H3,(H,59,60)(H,63,64)/t55-/m1/s1
InChIKey:	MYGLNICWUBAMIF-KZRJWCEABH

Wikipedia Article(s)

License

N-acylphosphatidylethanolamines (NAPEs) are hormones released by the small intestine into the bloodstream when it processes fat. It travels to the hypothalamus in the brain and suppresses appetite. This could make it useful for treating obesity. Read more... or Edit at Wikipedia...

Associated Data Sources and Commercial Suppliers

Filter

Names and Synonyms

Description

From the Lipid Library

N-Acyl phosphatidylethanolamine in which the free amino group of phosphatidylethanolamine is acylated by a further fatty acid is a common constituent of cereal grains (e.g. wheat, barley and oats) and of some other seeds, but it may occur in other plant tissues, especially under conditions of physiological stress. It has also been found in a number of microbial species.

This phospholipid has been detected in rather small amounts in several animal tissues, but especially brain, nervous tissues and the epidermis, when the N-acyl chain is often palmitic or stearic acid. Under conditions of degenerative stress, it can accumulate in significant amounts, for example as the result of ischemic injury, infarction or cancer.