



The IUPAC International Chemical Identifier (InChI)

<http://www.iupac.org/inchi>

<http://inchi.sourceforge.net>

InChI FAQ: <http://wwmm.ch.cam.ac.uk/inchifaq>

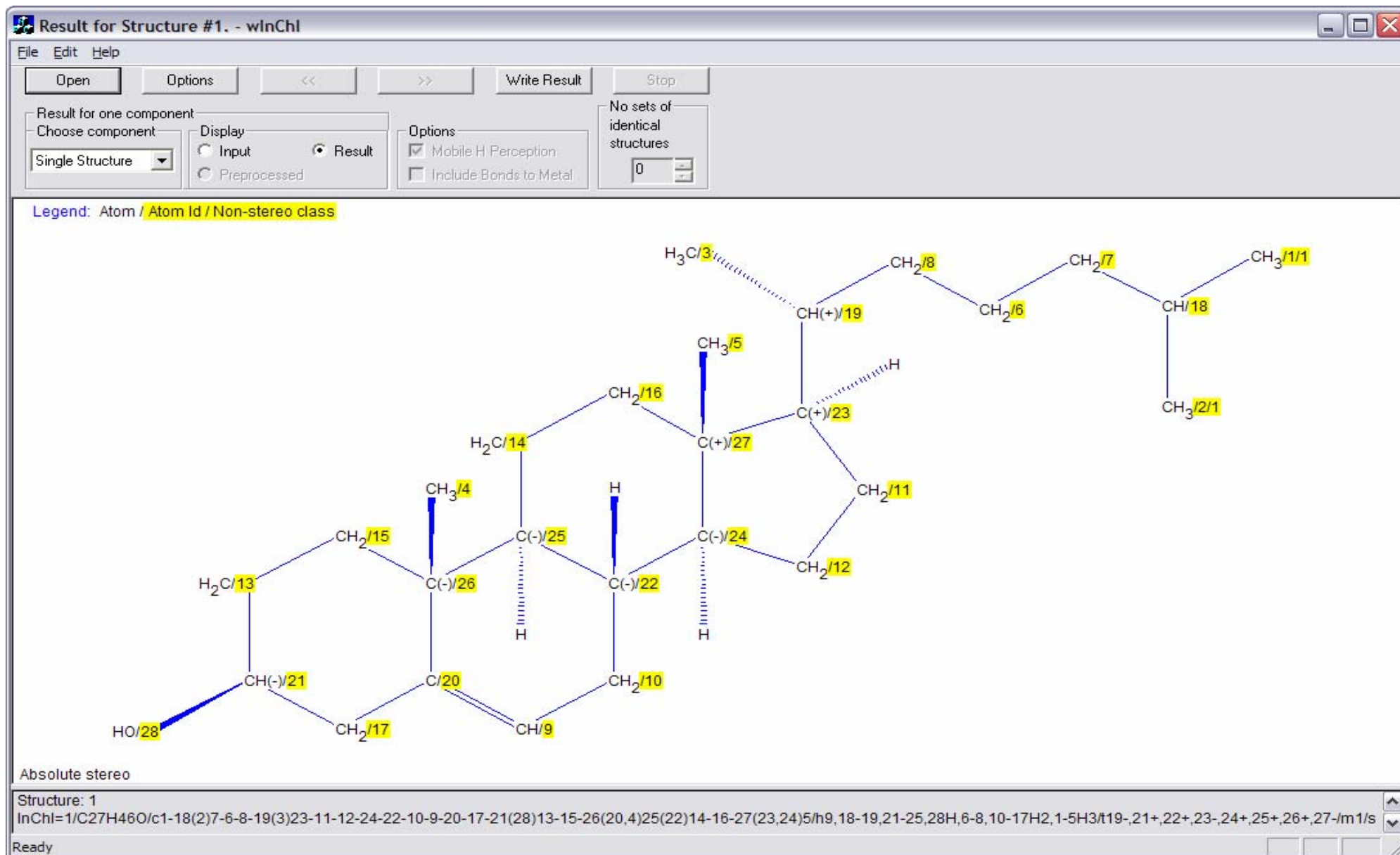
InChI:

- . an Open Source, non-proprietary, public-domain identifier for chemicals**
- . a string of characters uniquely representing a specified molecular structure**
- . a precise, robust, IUPAC-approved chemical substance tag**
- . independent of the way the chemical structure is drawn**
- . indexed by internet search engines**
- . usable in both printed and electronic data sources**
- . enables reliable structure recognition and easy linking of diverse data compilations**

InChI software:

- . accepts input in the form of MOLfiles (or SDfiles) and CML files**
- . is integrated with ACD ChemSketch**
- . accepts structures generated by cut-and-paste from ChemDraw, Marvin, and BKChem**
- . deals with organic compounds with *Z/E* and sp^3 configurations, tautomers, and isotopes as well as salts, organometallic compounds and protonated forms**
- . regenerates the original structure from InChI output**

InChI for cholesterol



InChI for Viagra

Result for Structure #1, mobile H. - winChI

File Edit Help

Open Options << >> Write Result Stop

Result for one component

Choose component: Single Structure

Display: Input Result Preprocessed

Options: Mobile H Perception Include Bonds to Metal

No sets of identical structures: 0

Legend: Atom / Atom Id / Non-stereo class / Mobile group id

Structure: 1
InChI=1/C22H30N6O4S/c1-5-7-17-19-20(27(4)25-17)22(29)24-21(23-19)16-14-15(8-9-18(16)32-6-2)33(30,31)28-12-10-26(3)11-13-28/h8-9,14H,5-7,10-13H2,1-4H3,(H,23,24,29)

Ready

InChI for C₆₀ fullerene

Result for Structure #1. - winChI

File Edit Help

Open Options << >> Write Result Stop

Result for one component

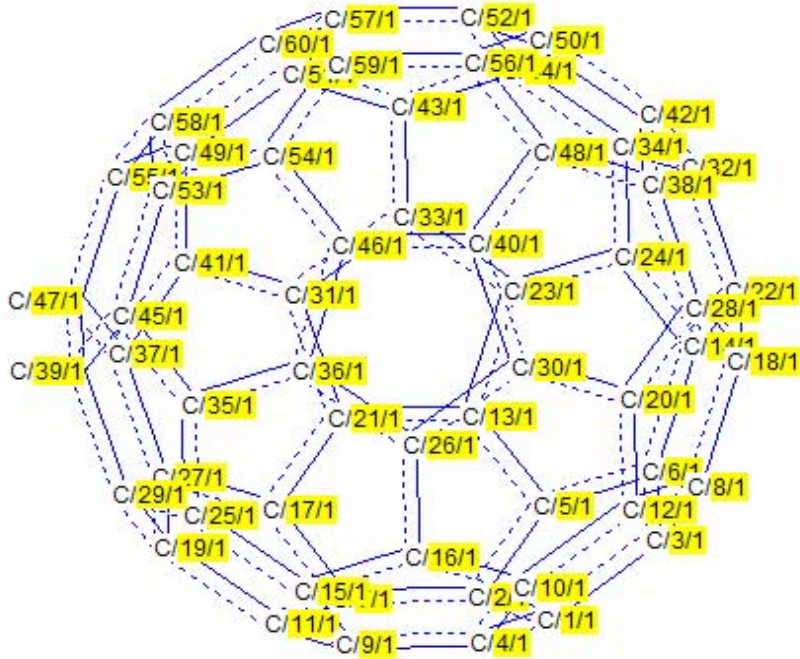
Choose component: Single Structure

Display: Input Result Preprocessed

Options: Mobile H Perception Include Bonds to Metal

No sets of identical structures: 0

Legend: Atom / Atom Id / Non-stereo class



InChI=1/C60/c1-2-5-6-3(1)8-12-10-4(1)9-11-7(2)17-21-13(5)23-24-14(6)22-18(8)28-20(12)30-26-16(10)15(9)25-29-19(11)27(17)37-41-31(21)33(23)43-44-34(24)32(22)42-38(28)48-40(30)46-36(26)35(25)45-39(29)47(37)55-49(41)51(43)57-52(44)50(42)56(48)59-54(46)53(45)58(55)60(57)59

Ready

InChI for ferrocene

Result for Structure #1. (Reconnected) - wInChI

File Edit Help

Open Options << >> Write Result Stop

Result for one component

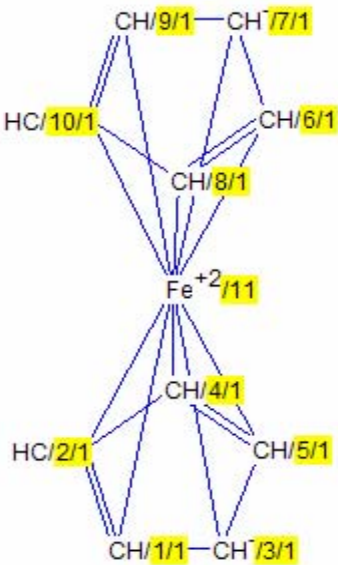
Choose component: Single Structure

Display: Input Result Preprocessed

Options: Mobile H Perception Include Bonds to Metal

No sets of identical structures: 0

Legend: Atom / Atom Id / Non-stereo class



message: type="warning" value="Accepted unusual valence(s): C(5); C-1(4); Metal was disconnected"

Structure: 1
InChI=1/2C5H5.Fe/c2*1-2-4-5-3-1;/h2*1-5H;/q2*-1;+2/rC10H10Fe/c1-2-4-5-3(1)11(1,2,4,5)6-7(11)9(11)10(11)8(6)11/h1-10H

Ready

InChI data layers:

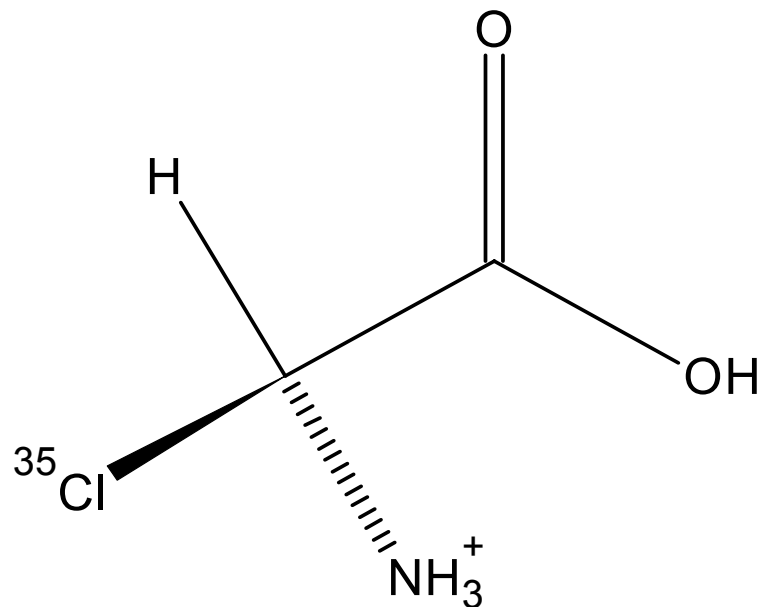
1. Formula (standard Hill sorted)
2. Connectivity (no formal bond orders)
 - i. disconnected metals
 - ii. connected metals
3. Isotopes
4. Stereochemistry
 - i. double bond (*Z/E*)
 - ii. tetrahedral (sp^3)
5. Tautomers (on or off)

Charges are not part of the basic InChI, but are added at the end of the InChI string.

InChI symbolism:

- /c** connectivity-1.1 (excluding terminal H)
- /h** connectivity-1.2 (locations of terminal H, including mobile H attachment points)
- /p** proton balance
- /t** sp^3 (tetrahedral) parity
- /m** parity inverted to obtain relative stereo (1 = inverted, 0 = not inverted, . = unaffected by inversion)
- /s** stereo type (1 = absolute, 2 = relative, 3 = racemic)
- /i** isotopic specification
- /f** chemical formula of the fixed-H structure if different
- /h** connectivity-2 (locations of fixed mobile H)
- /q** charge
- /r** chemical formula of “connected metal” structure

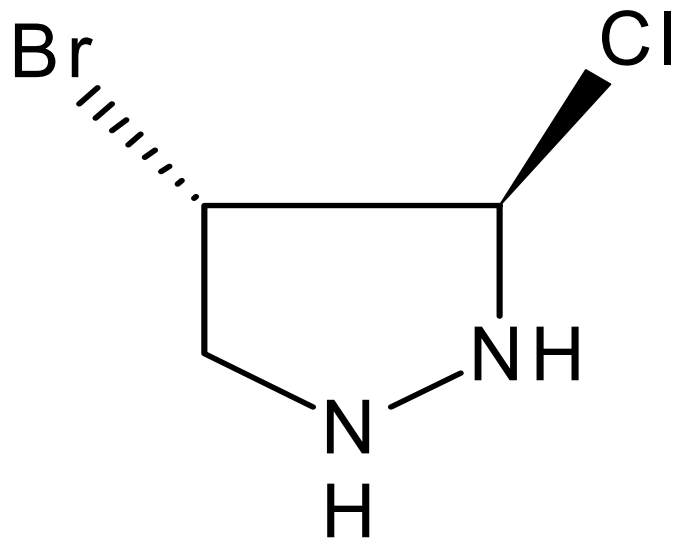
Chloroalaninium InChI layers



InChI=1/C2H4CINO2/c3-1(4)2(5)6/h1H,4H2,(H,5,6)/p+1/t1-/m1/s1/i3+0/fC2H5CINO2/h4-5H/q+1

Main layer:	C2H4CINO2/c3-1(4)2(5)6
Proton balance layer:	p+1
Stereo layer:	t1-/m1/s1
Isotopic layer:	i3+0
Fixed-H layer:	fC2H5CINO2/h4-5H
Charge layer:	q+1

Configuration



Abs. chirality: InChI = 1/C3H6BrClN2/c4-2-1-6-7-3(2)5/h2-3,6-7H,1H2/t2-,3-/m0/s1

Rel. chirality: InChI = 1/C3H6BrClN2/c4-2-1-6-7-3(2)5/h2-3,6-7H,1H2/t2-,3-/s2

Racemic: InChI = 1/C3H6BrClN2/c4-2-1-6-7-3(2)5/h2-3,6-7H,1H2/t2-,3-/s3

Web search for retinol InChI

Google Search: 1/C20H30O/c1-16(8-6-9-17(2)13-15-21)11-12-19-18(3)10-7-14-20(19,4)5/h6,8-9,11-13 - Microsoft Internet Explorer p

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Links

Google +,12-11+,16-8+,17-13+ Search Web 11 blocked AutoFill Options 1 C20H30O c1 16 8 6 9 17 2 13

Address <http://www.google.co.uk/search?hl=en&biw=1243&q=1%2FC20H30O%2Fc1-16%288-6-9-17%282%2913-15-21%2911-12-19-18%283%2910-7-14-20%2819%2C4%285%29h6,8-9,11-13,21H,7,10,14-15H2,1-5H3/b9-6+,12-11+,16> Go

Google Web Images Groups News Froogle^{New!} more »

1/C20H30O/c1-16(8-6-9-17(2)13-15-21)11-12-19-18(3)10-7-14-20(19,4)5/h6,8-9,11-13,21H,7,10,14-15H2,1-5H3/b9-6+,12-11+,16 Search Advanced Search Preferences

Search: the web pages from the UK

"14" (and any subsequent words) was ignored because we limit queries to 32 words.

Web Results 1 - 1 of about 2 for 1/C20H30O/c1-16(8-6-9-17(2)13-15-21)11-12-19-18(3)10-7-14-20(19,4)5/h6,8-9,11-13,21H,7,10,14-15H2,1-5H3/b9-6+,12-11+,16

[PubChem Substance Summary](#)
InChI: InChI=1/C20H30O/c1-16(8-6-9-17(2)13-15-21)11-12-19-18(3)10-7-14-20(19,4)5/h6,8-9,11-13,21H,7,10,14-15H2,1-5H3. Depositor-Supplied Comments: ...
pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?sid=3137 - Similar pages

*In order to show you the most relevant results, we have omitted some entries very similar to the 1 already displayed.
If you like, you can **repeat the search with the omitted results included**.*

Internet

InChI early users:

US National Institute of Standards and Technology – 150,000 structures

US National Center for Biotechnology Information: PubChem project – 800,000+ structures

US National Cancer Institute – 23 million+ structures

US Environmental Protection Agency – DSSTox database – 1450 structures

Kyoto Encyclopedia of Genes and Genomes (KEGG) database – 9584 structures

University of California at San Francisco ZINC database – 3.3 million structures

BRENDA enzyme information system (University of Cologne) – 36,000 structures

Chemical Entities of Biological Interest (ChEBI) database of the European Bioinformatics Institute – 5000 structures

Nature Chemical Biology

Under consideration:

ISI – 2 million+ structures

Beilstein Journal of Organic Chemistry

European Patent Office

US Patent and Trademark Office

Selected literature:

Stephen E. Stein, Stephen R. Heller, and Dmitrii Tchekhovskoi, "An Open Standard for Chemical Structure Representation: The IUPAC Chemical Identifier", in *Proceedings of the 2003 International Chemical Information Conference (Nimes)*, Infonortics, pp. 131-143;
<http://www.hellers.com/steve/resume/p157.html>

Michael Freemantle, "Unique Labels for Compounds", *C&EN*, Vol.80, No. 48, 2 Dec 2002

David Adam, "Chemists synthesize a single naming system", *Nature*, 23 May 2002

Peter Murray-Rust, Henry S. Rzepa and Yong Zhang, "Googling for INChIs; A remarkable method of chemical searching", *W3C Workshop on Semantic Web for Life Sciences*, 27-28 October 2004, Cambridge, Massachusetts USA;
<http://lists.w3.org/Archives/Public/public-swls-ws/2004Oct/att-0019/>

Simon J. Coles, Nick E. Day, Peter Murray-Rust, Henry S. Rzepa, and Yong Zhang, "Enhancement of the chemical semantic web through the use of InChI identifiers", *Org. Biomol. Chem.*, 2005, 3(10), 1832 – 1834

M. D. Prasanna, Jiri Vondrasek, Alexander Wlodawer, T.N. Bhat, "Application of InChI to Curate, Index, and Query 3-D Structures", *PROTEINS: Structure, Function and Bioinformatics*, 2005, 60:1-4

Aurélien Monge, Alban Arrault, Christophe Marot and Luc Morin-Allory, "Analysis of a Set of 2.6 Million Unique Compounds gathered from the Libraries of 32 Chemical Providers", presented at the 10th Electronic Computational Chemistry Conference, April 2005;
<http://www.univ-orleans.fr/icoa/eposter/eccc10/monge/>

Bea Perks, "International chemical identifier goes online", *Chem. World*, 16 May 2005

Stephen E Stein

Stephen R Heller

Dmitrii Tchekhovskoi

(US National Institute for Standards and Technology)

Alan D McNaught

(Royal Society of Chemistry)

IUPAC General Assembly

Beijing

2005

Pages layout for poster

2	1	3
4	5	6
8	7	9
10	12	11
13	15	14