



InChI wInChI Spider

Jonathan M Goodman
Yusuf Hamied
Department of Chemistry
University of Cambridge

InChI is..

IUPAC International Chemical Identifier (InChI)

5h2i5
120H5
100rs
C6H18
C7-1-1

- A unique identifier of a chemical structure serving as its digital signature
- A machine-readable string of characters derived solely from a structural representation of a chemical substance
 - A project of IUPAC and the InChI Trust

InChI=1/C6H8O/C1=CC=CC=C1
<(8)5-3(9)4(10)6(11)12-



The purpose of InChI is...

- To streamline naming conventions for chemical compounds and reactions
- To uniquely identify a chemical substance, without ambiguity, providing a precise, robust, structure-derived tag for chemical substances

wInChI

Windows GUI and details of InChI generation - recommended for super-users.

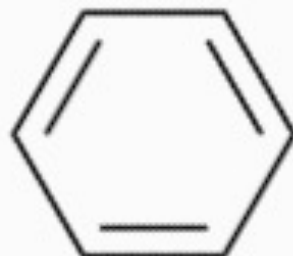
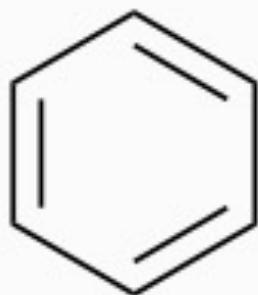
Most mortals will prefer the WebDemo



Representations and Identifiers

All of these are useful
For some applications
the InChI is best

Benzene



CAS: 71-43-2

Wiswesser: L6J

Dyson: B6

SMILES

C1=CC=CC=C1

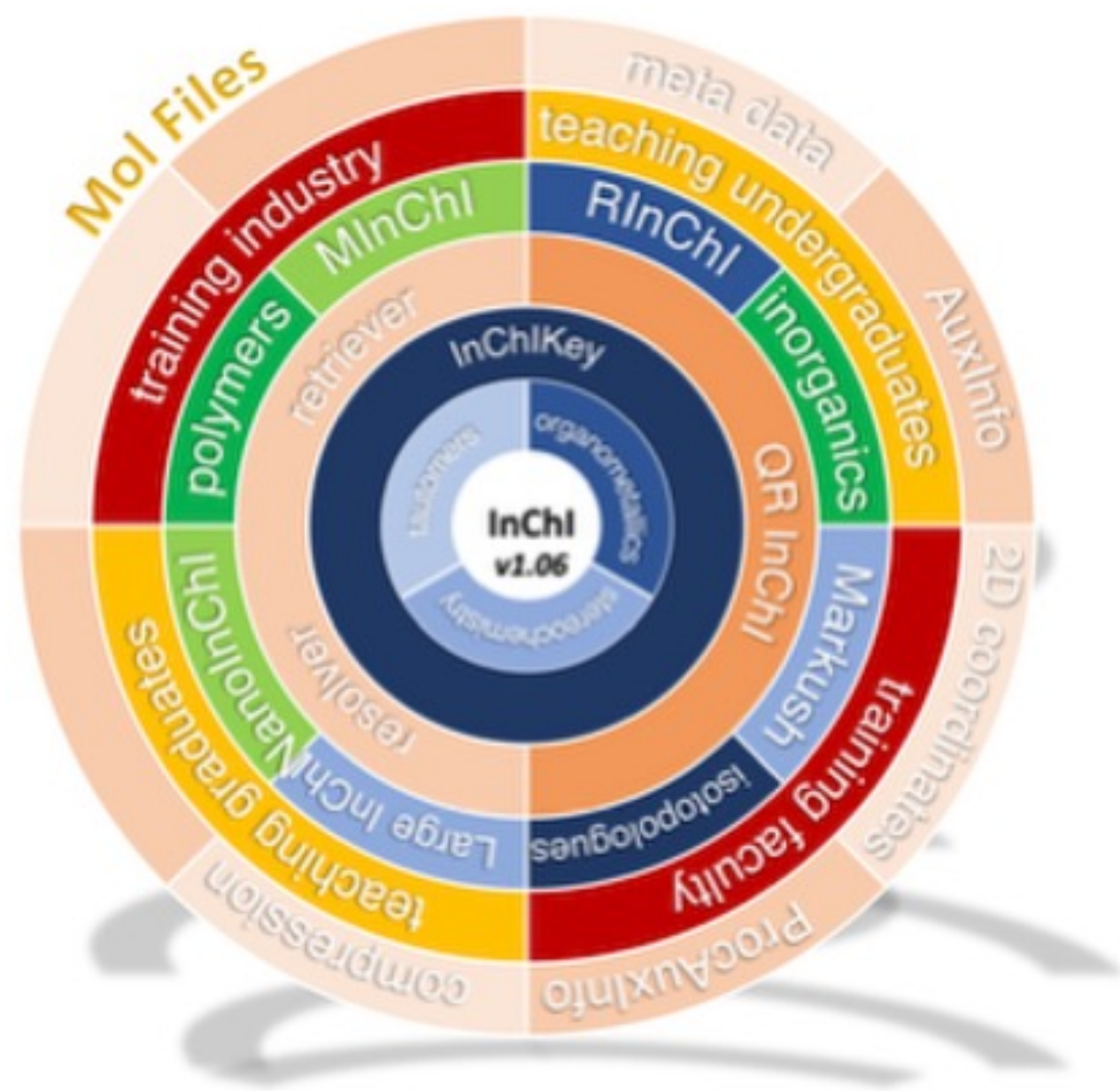
c1ccccc1

InChI

InChI=1S/C6H6/c1-2-4-6-5-3-1/h1-6H

InChIKey

UHOVQNZJYSORNB-UHFFFAOYSA-N



InChI
canonical
representifier
for molecules



InChI is..

IUPAC International Chemical Identifier (InChI)

5h2i5
1201k
100rs
6818
c7-1-

- A unique identifier of a chemical structure serving as its digital signature
- A machine-readable string of characters derived solely from a structural representation of a chemical substance
 - A project of IUPAC and the InChI Trust

InChI=1/C6H8O7<(8)5-3(9)4(10)6(11)12-



The purpose of InChI is...

- To streamline naming conventions for chemical compounds and reactions
- To uniquely identify a chemical substance, without ambiguity, providing a precise, robust, structure-derived tag for chemical substances

winChI

Windows GUI and details of InChI generation - recommended for super-users.

Most mortals will prefer the WebDemo



InChI is..

IUPAC International Chemical Identifier (InChI)

- A unique identifier of a chemical structure serving as its digital signature
- A machine-readable string of characters derived solely from a structural representation of a chemical substance
- A project of IUPAC and the InChI Trust

The purpose of InChI is...

- To streamline naming conventions for chemical compounds and reactions
- To uniquely identify a chemical substance, without ambiguity, providing a precise, robust, structure-derived tag for chemical substances
- To assist in merging and linking chemical databases

InChI is essential...

as the only structure representation standard in the public domain, open-source and freely available to the scientific community



RESEARCH ARTICLE

Open Access



InChI version 1.06: now more than 99.99% reliable

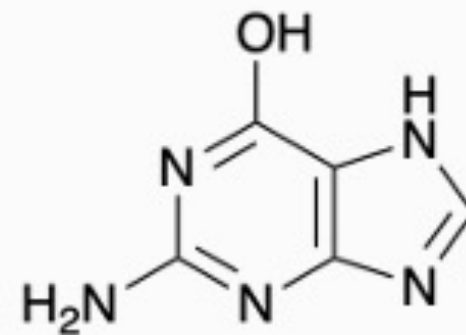
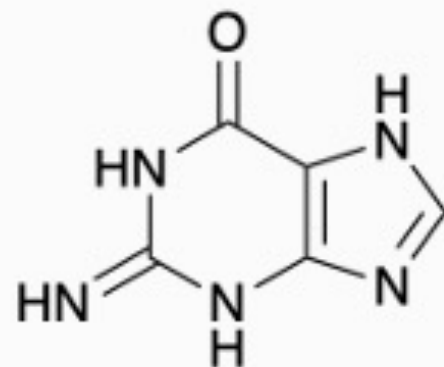
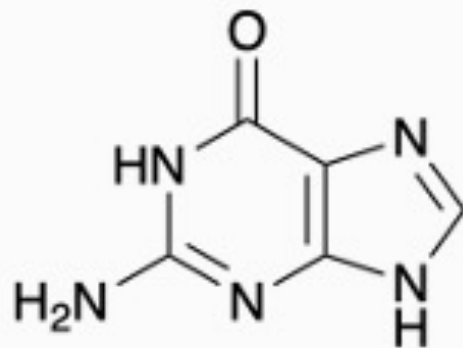
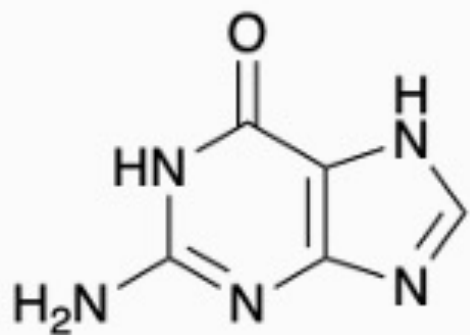
Jonathan M. Goodman¹ , Igor Pletnev^{2,3} , Paul Thiessen⁴ , Evan Bolton⁴  and Stephen R. Heller^{2,4*} 

Abstract

The software for the IUPAC Chemical Identifier, InChI, is extraordinarily reliable. It has been tested on large databases around the world, and has proved itself to be an essential tool in the handling and integration of large chemical databases. InChI version 1.05 was released in January 2017 and version 1.06 in December 2020. In this paper, we report on the current state of the InChI Software, the details of the improvements in the v1.06 release, and the results of a test of the InChI run on PubChem, a database of more than a hundred million molecules. The upgrade introduces significant new features, including support for pseudo-element atoms and an improved description of polymers. We expect that few, if any, applications using the standard InChI will need to change as a result of the changes in version 1.06. Numerical instability was discovered for 0.002% of this database, and a small number of other molecules were discovered for which the algorithm did not run smoothly. On the basis of PubChem data, we can demonstrate that InChI version 1.05 was 99.996% accurate, and InChI version 1.06 represents a step closer to perfection. Finally, we look forward to future releases and extensions for the InChI Chemical Identifier.

Keywords: InChI, InChIKey, PubChem, RInChI

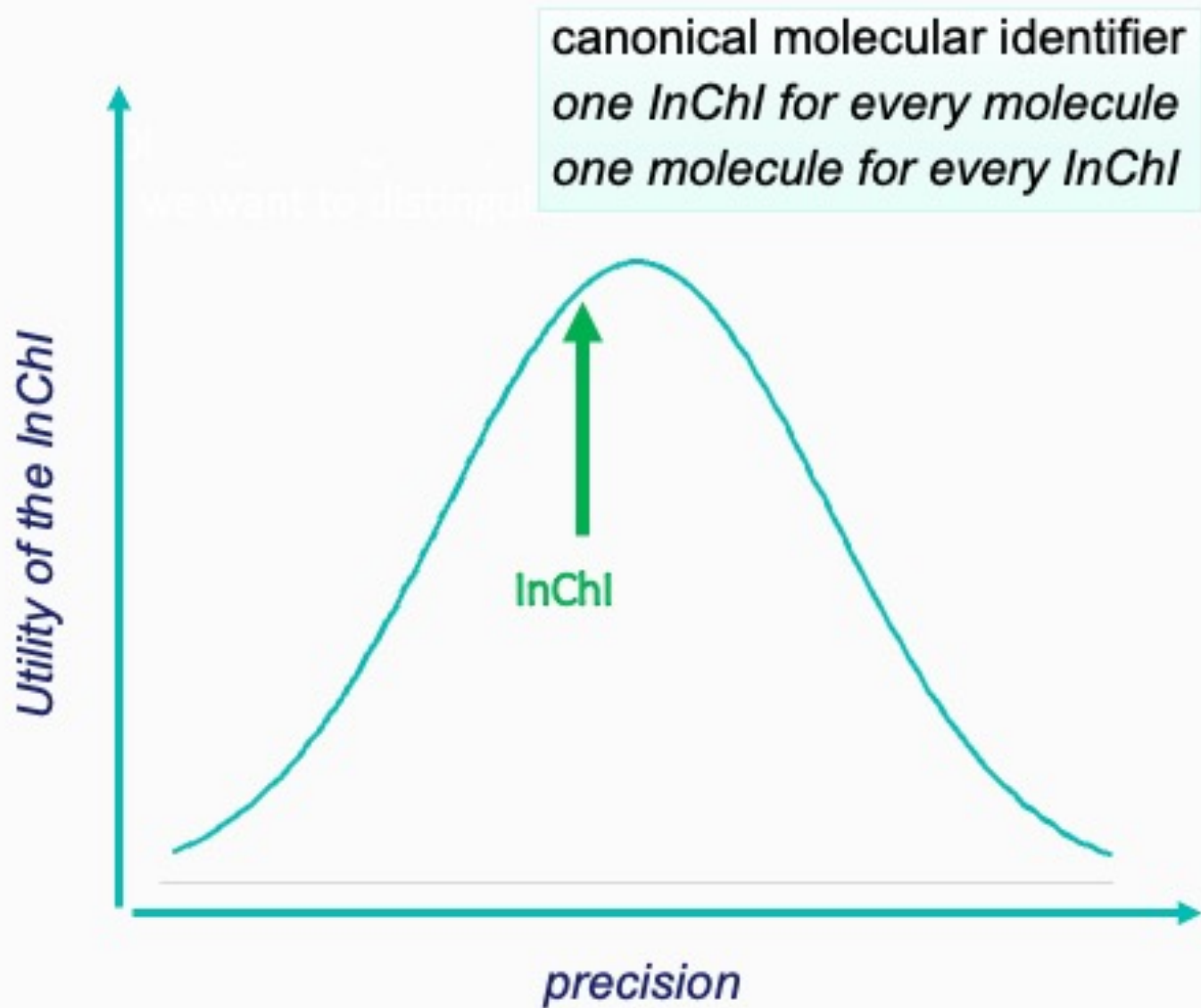
Guanine



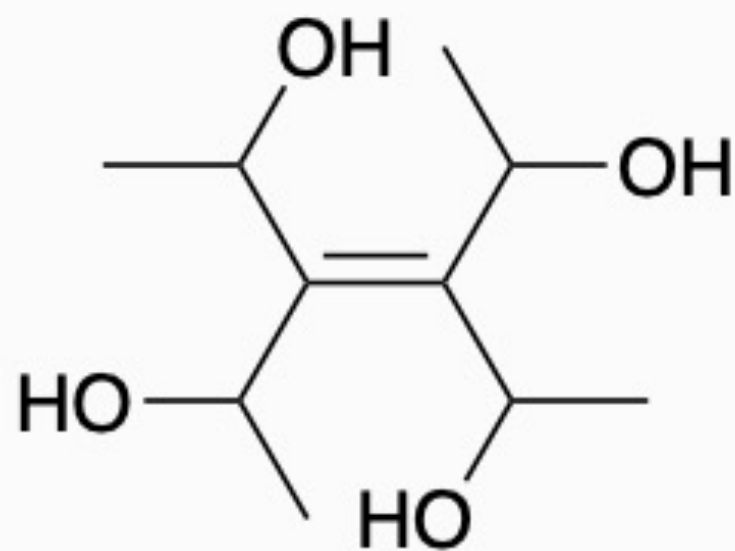
InChI=1S/C5H5N5O/c6-5-9-3-2(4(11)10-5)7-1-8-3/h1H,(H4,6,7,8,9,10,11)

Precision vs Utility

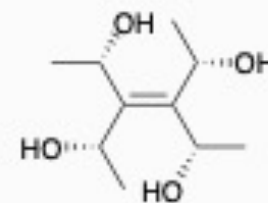
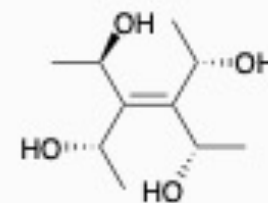
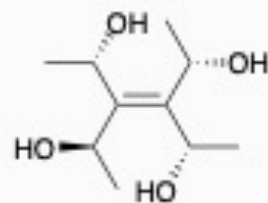
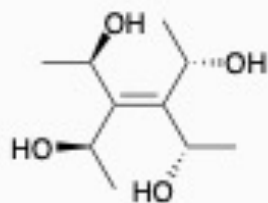
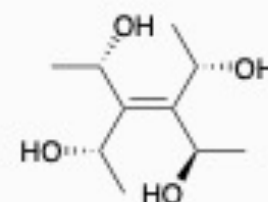
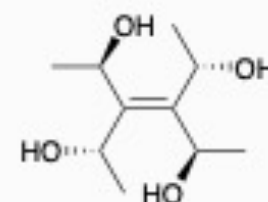
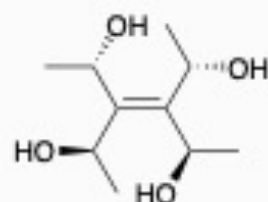
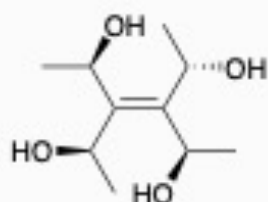
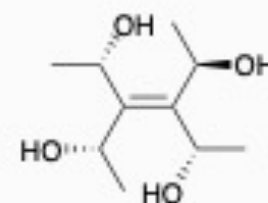
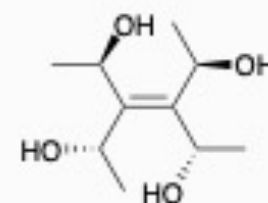
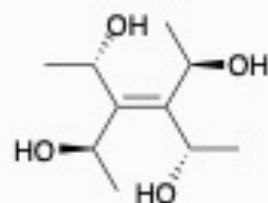
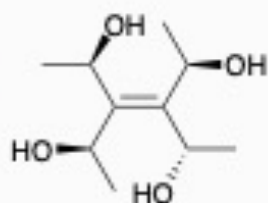
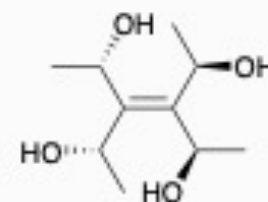
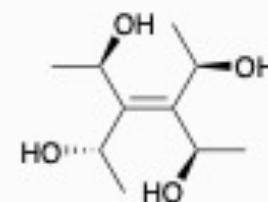
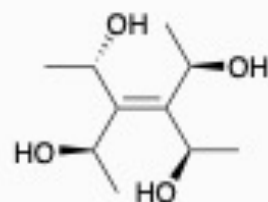
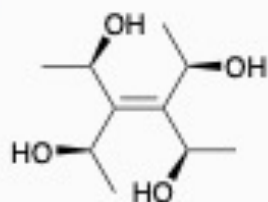
*InChI are most useful if
everybody uses them in
the same way*



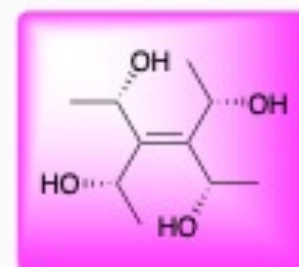
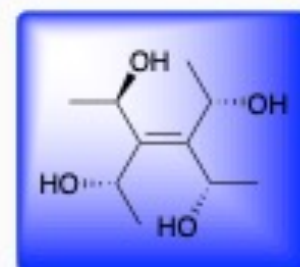
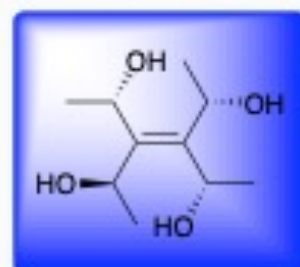
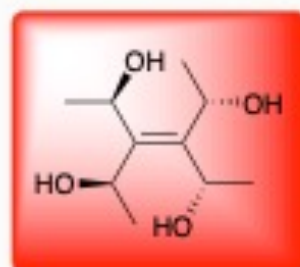
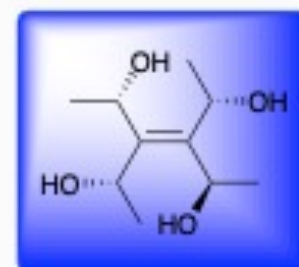
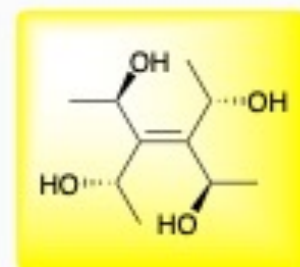
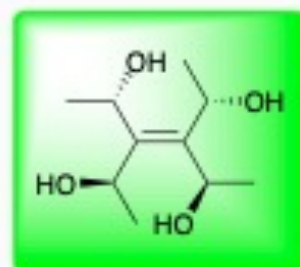
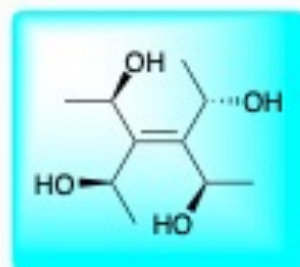
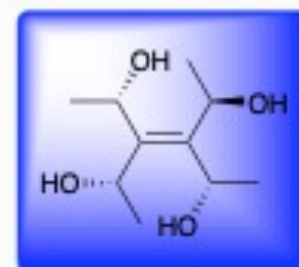
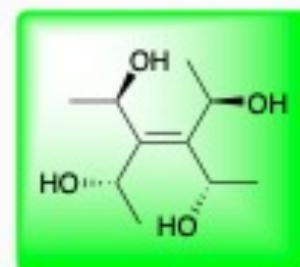
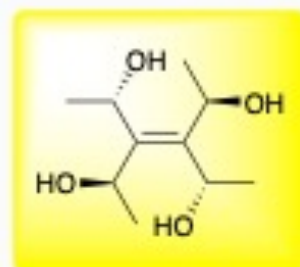
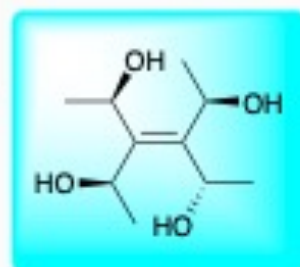
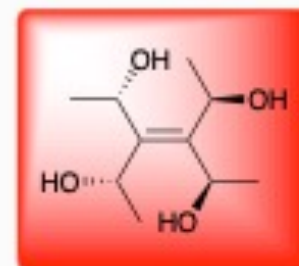
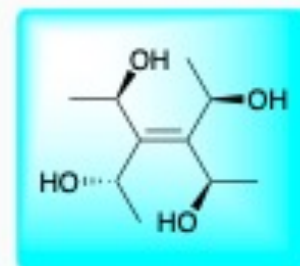
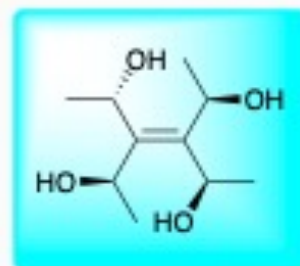
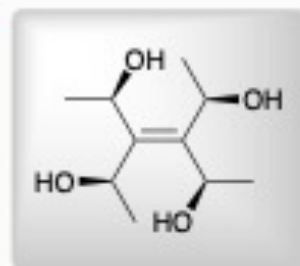
3,4-bis
(1-hydroxyethyl)
hex-3-ene-2,5-
diol



**3,4-bis
(1-hydroxyethyl)
hex-3-ene-2,5-
diol**



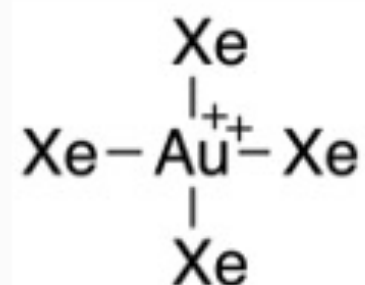
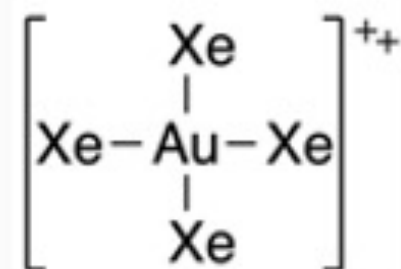
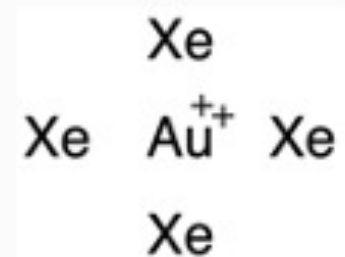
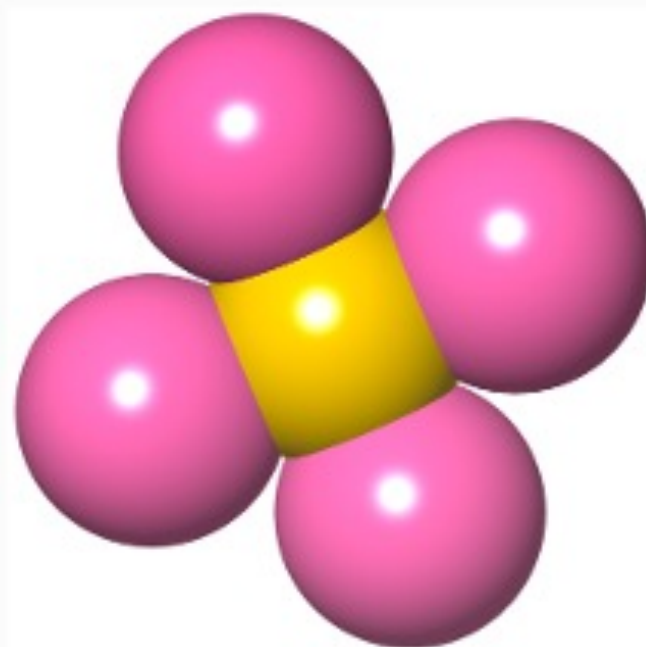
3,4-bis
(1-hydroxyethyl)
hex-3-ene-2,5-
diol



Tetraxenonogold(II)

InChI=1S/Au.4Xe/q+2;;;;

InChI=1/AuXe4/c2-1(3,4)5/q+2



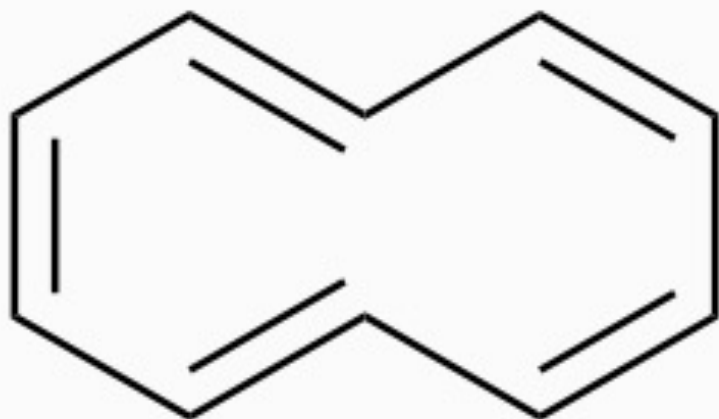
Tractatus Logico-Philosophicus

By
LUDWIG WITTGENSTEIN

With an Introduction by
BERTRAND RUSSELL, F.R.S.



Whereof one cannot speak,
thereof one must be silent.



What is the SMILES ?



InChI is..

IUPAC International Chemical Identifier (InChI)

5h2i5
1201k5
1201k5
1201k5
1201k5
1201k5
1201k5
1201k5
1201k5
1201k5



Monthly InChI
Discussions

Last Thursday
of most months
5 pm UK time

June 25th

tag for chemical substances

winChI

Windows GUI and
details of InChI
generation -
recommended for
super-users.

*Most mortals will prefer
the WebDemo*

