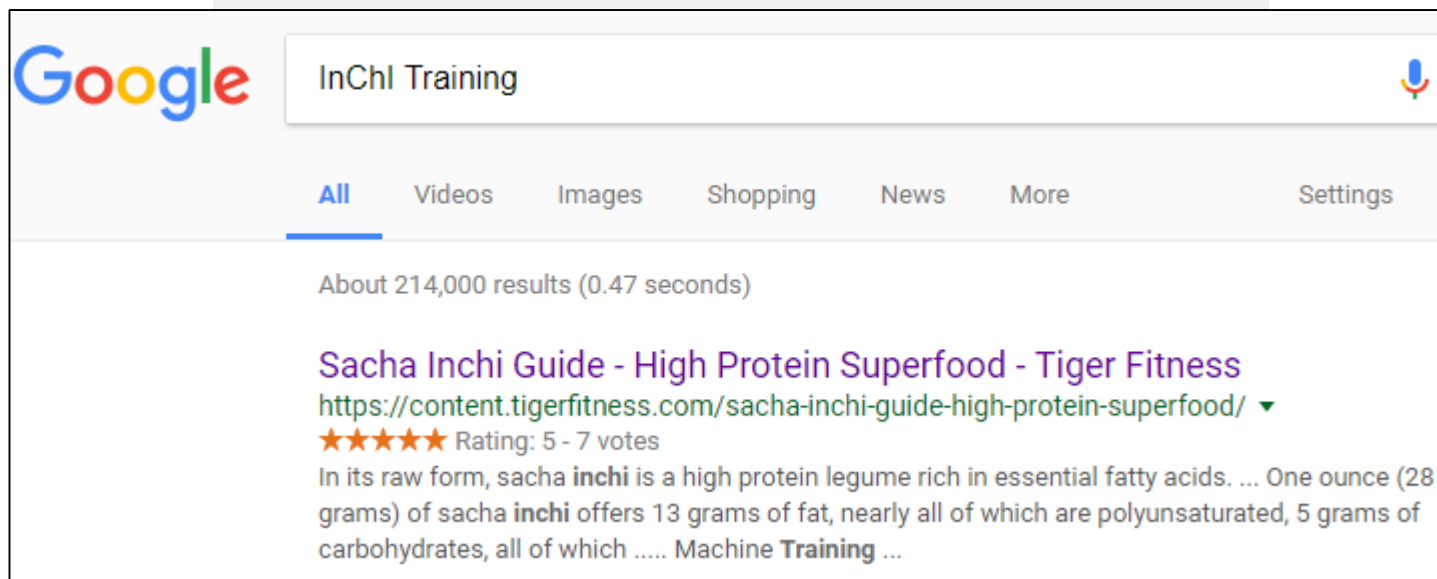


InChI & Working Group Reports:
Training/Education/Outreach

Robert E. Belford
(rebelford@ualr.edu)

*Should there be a working group
on
Training/Education/Outreach?*

(Thursday Break-Out Group: 3:00-4:30 PM)



Google search results for "InChI Training". The search bar shows "InChI Training" and the Google logo. Below the search bar are tabs for "All", "Videos", "Images", "Shopping", "News", "More", and "Settings". The search results show "About 214,000 results (0.47 seconds)". The top result is "Sacha Inchi Guide - High Protein Superfood - Tiger Fitness" with a URL <https://content.tigerfitness.com/sacha-inchi-guide-high-protein-superfood/> and a rating of 5 stars (5 - 7 votes). The snippet below the title reads: "In its raw form, sacha inchi is a high protein legume rich in essential fatty acids. ... One ounce (28 grams) of sacha inchi offers 13 grams of fat, nearly all of which are polyunsaturated, 5 grams of carbohydrates, all of which Machine Training ...".



Sacha Inchi Guide - High Protein Superfood

The Truth is:

- Few outside of the Cheminformatics Community are familiar with InChI
- Even within IUPAC many are not familiar
- This is especially true in the chemical education community

InChI Trust WebSite: Key Asset for Training/Education/Outreach



InChI and InChIKeys for chemical structures

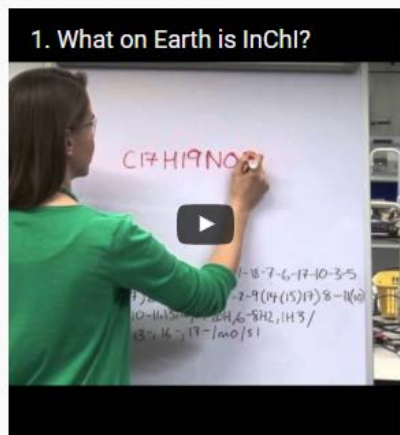
The InChI Trust is a nonprofit charity which works to develop and promote the use of the IUPAC InChI open-source chemical structure representation algorithm.

The InChI with its associated InChIKey was developed as a non-proprietary international standard to represent chemical structures and has just celebrated its tenth anniversary. The Trust aims to develop and improve on the current InChI standard, further enabling the interlinking of chemistry and chemical structures on the web.

Register now for our upcoming "State and Future of the IUPAC InChI" meeting at the NIH 16-18 August 2017 – more details [here](#)

What on Earth is InChI?

This video introduces the InChI standard to represent chemical structures.



The Birth of the InChI

This video describes the background to the development of the InChI standard.



The Googable InChIKey

An InChIKey enables chemical structures to be uniquely identified on the web.



InChI and the Islands

The InChI standard – the International Chemical Identifier – enables the linking of information.



IUPAC Centennial Opportunity for Training/Education/Outreach

IUPAC 100TH ANNIVERSARY TO BE CELEBRATED IN 2019

28 July 2017



Today is the birthday of IUPAC!

Research Triangle Park, NC, USA, July 28, 2017 – The International Union of Pure and Applied Chemistry (IUPAC), the globally-recognized authority on chemical nomenclature and terminology, will celebrate its 100th anniversary on July 28, 2019. The anniversary theme is ***A Common Language for Chemistry***, and while the celebration will recognize the successes of IUPAC's first one hundred years, its purpose is not to solely dwell upon the past, but rather to look to the future of what this international community of chemists, working closely together, can continue to contribute now and into the future in meeting the world's needs through chemical research.



What are the differences between Training/Education/Outreach?

(my shot at this)

Training – Develop understanding of InChI technology and established use scenarios

Education – Develop understanding of how InChI fits into and can impact the practice of science

Outreach – Develop applications of InChI technology to new communities and establish new use scenarios.

Possible Features:

InChI Primers – repository for technical primers targeting a multiplicity of novice and expert needs

InChI Bibliography – wiki based/crowdsourced

Cheminformatic Education Material – Original material and a links to other sites showing applications of InChI

Traditional– Cheminformatics

- Material like the Cheminformatics OLCC

Non Traditional – CER (Chemical Education Research)

- Add a new paradigm to CER utilizing principles of distributed cognition

Traditional
Curriculum Integration:

Quick Overview of
Cheminformatics OLCC and LibreText

- Cheminformatics OLCC
 - Intercollegiate course sponsored by ACS CHED CCCE
- Libretext
 - Formerly ChemWiki/STEMWiki hyperlibrary
 - Open access eTextbook generating service

(Both NSF Funded Projects)

Cheminformatics OLCC

Fall 2015 - 4 schools

Spring 2017 – 10 schools



DivCHED CCCE: Cheminformatics OLCC

Search



Note: Hypothes.is annotations on any page on the web with the following tag
2017OLCCModule2P3
will be aggregated at the bottom of this page.

TO JOIN

Contact Site Moderator
Dr. Robert E. Belford
rebelford@ualr.edu

[Spring 2017 Course](#)

[Fall 2015 Course](#)

[Spring 2017 Course
Development Site](#)


[About Us](#)

[Forums](#)

[Recent content](#)

2.3 Chemical Representations on Computer: Part III

Download PDF:  [Download OLCC-2017_mod-2_part-3.pdf](#) (554.04 KB)

 [Download OLCC-2017_mod-2_part-3_Questions.docx](#) (3.73 MB)

Sunghwan Kim, National Center for Biotechnology Information

Learning Objectives

- Explain what SMILES, SMARTS and SMIRKS are.
- Explain what InChI and InChIKey are.
- Review SMILES specification rules.
- Compare and contrast SMILES and InChI.
- Demonstrate how to interpret SMILES, SMARTS, InChI strings into their corresponding chemical structures.

[Table of Contents](#)

[What on Earth is InChI?](#)

<http://olcc.ccce.divched.org/2017OLCCModule2P3>

Cheminformatics OLCC material was ported to LibreText

Organic Chemistry Classes at any school using LibreText have access to InChI material from UALR

INTERNATIONAL CHEMICAL IDENTIFIER (InChI) AND InChI KEY

InChI

The IUPAC International Chemical Identifier (InChI)^{1,2,3} was originally developed by the IUPAC and continuing development efforts have been made by the InChI Trust⁴. InChI is non-proprietary, open-source, and freely available to the scientific community. Especially, because the software for generating InChI strings is also freely available, it avoids the interoperability issue that different implementations of SMILES language have.

InChI encodes a chemical structure into "layers". Each layer holds a distinct and separable class of structural information, with the layers ordered to provide successive structural refinement. There are currently six InChI layer types, each different class of structural information: the main layer, a charge layer, a stereochemical layer, an isotopic layer, a fixed-H layer and a connected layer. The main layer, which specifies chemical formula, atoms, and bonds between them, is required for all InChIs. However, the other layers appear only when corresponding input information is provided. Layers and sublayers start with "/" (forward slash) followed by a letter denoting the identity of the layer (except for the chemical formula layer). Below are some examples of InChI.

InChI=1S/CH4/h1H4 (methane)

InChI=1S/C2H6/c1-2/h1-2H3 (ethane)

InChI=1S/C2H6O/c1-2-3/h3H,2H2,1H3 (ethanol)

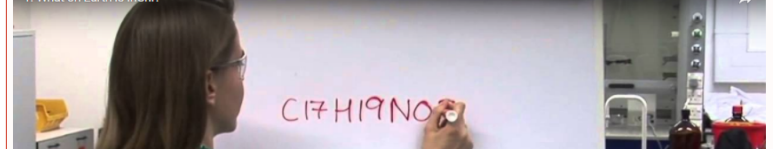
InChI=1S/C3H7NO2/c1-2(4)3(5)6/h2H,4H2,1H3,(H,5,6)/t2-m/s1 (L-alanine)

These InChI strings are not easy for a human to understand (especially compared to SMILES strings). It is because InChI was developed as a "machine-readable" chemical identifier, with an aim to enable a computer to regenerate the corresponding chemical structure from the InChI string generated by another computer. For this reason, InChI is often called as the bar code for chemical structures.

Because the layered structure of InChI allows one to represent a chemical structure with a desired level of details, InChI software may generate different InChI strings for the same molecule. This flexibility may be regarded as an obstacle to standardization and interoperability. In response to this concern, the standard InChI was introduced which contains the same level of structural details and the same conventions for drawing perception, by using standard option settings in InChI software. The standard InChI representations begin with "InChI=1S/", while the non-standard InChI begins with "InChI=1". The digit "1" following "InChI=" is the current InChI version number.


WHAT ON EARTH IS InChI

1. What on Earth is InChI?




2017 Page Views as of June 20

Chemistry	38,202,891
Biology	1,076,501
Physics	394,677
Geology	59,596
Math	263,490
Statistics	245,862
Medicine*	101,677
Social Sci*	26,943
Engineering*	49,269
Agriculture*	1,251
Humanities*	14,219
Photosciences	114,525



University of Arkansas
Little Rock



University of Illinois,
Springfield

3.3: An Introduction to Nomenclature

Last updated: 11:51, 15 Nov 2015

The increasingly large number of organic compounds identified with each passing day, together with the fact that many of these compounds are isomers of other compounds, requires that a systematic nomenclature system be developed. Just as each distinct compound has a unique molecular structure which can be designated by a structural formula, each compound must be given a characteristic and unique name.

Introduction

As organic chemistry grew and developed, many compounds were given trivial names, which are now commonly used and recognized. Some examples are:

Name	Methane	Butane	Acetone	Toluene	Acetylene	Ethyl Alcohol
Formula	CH ₄	C ₄ H ₁₀	CH ₃ COCH ₃	CH ₃ C ₆ H ₅	C ₂ H ₂	C ₂ H ₅ OH

Such **common names** often have their origin in the history of the science and the natural sources of specific compounds, but the relationship of these names to each other is arbitrary, and no rational or systematic principles underlie their assignments.

The IUPAC Systematic Approach to Nomenclature

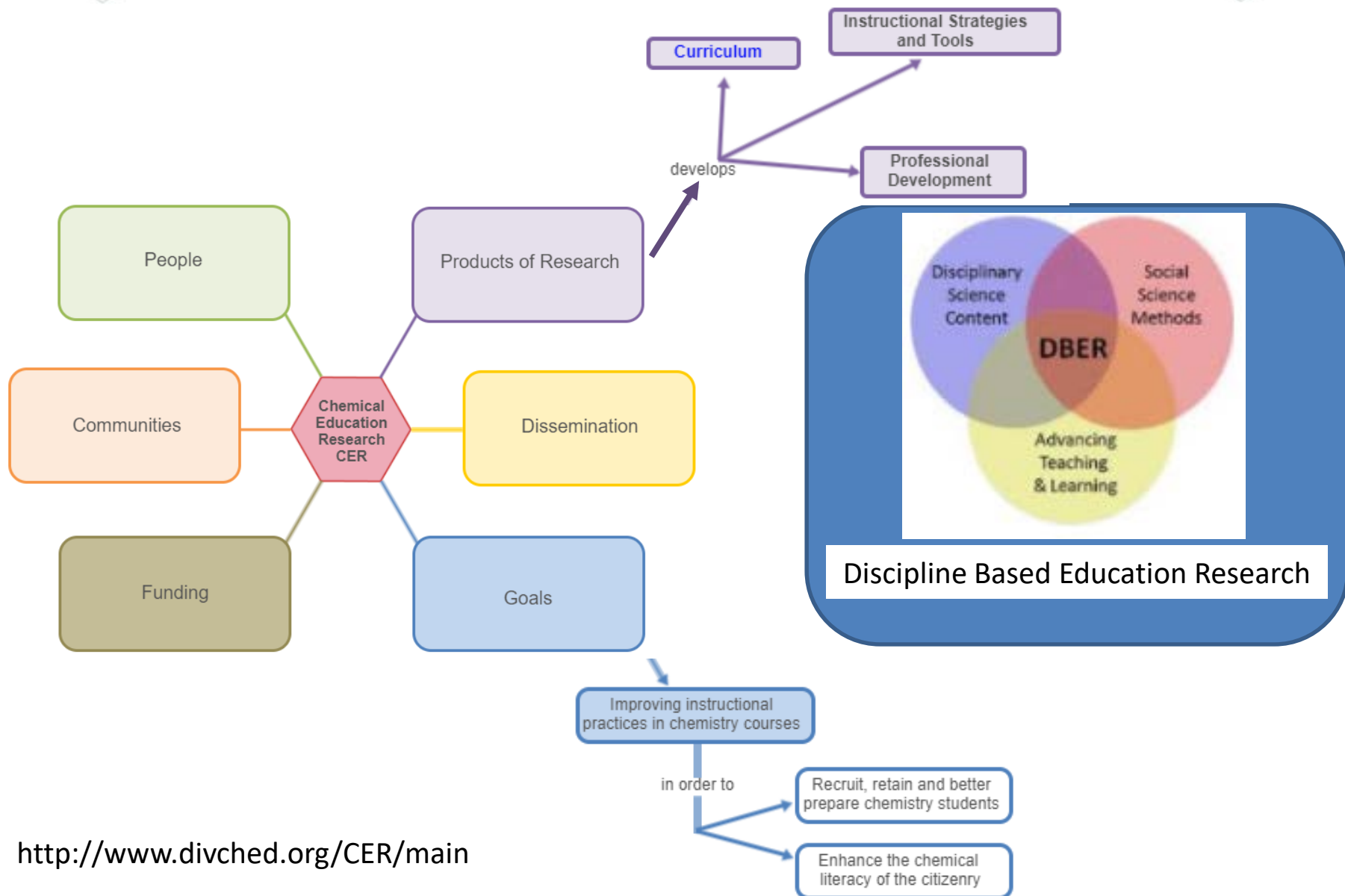
A rational nomenclature system should do at least two things. First, it should indicate how the carbon atoms of a given compound are bonded together in a characteristic lattice of chains and rings. Second, it should identify and locate any functional groups present in the compound. Since hydrogen is such a common component of organic compounds, its amount and locations can be assumed from the tetravalency of carbon, and need not be specified in most cases.

The IUPAC nomenclature system is a set of logical rules devised and used by organic chemists to circumvent problems caused by arbitrary nomenclature. Knowing these rules

Cheminformatics OLCC
Fall 2015 - 4 schools
Spring 2017 – 10 schools

Many student projects used InChI and
InChI keys
Especially projects with “Smart
Spreadsheets”

NonTraditional CER & DBER Education Research:





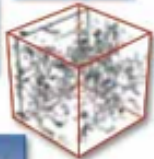


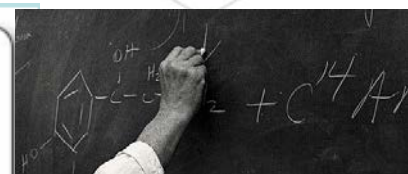
(1917)

Youth want
to Drive
not ride



Science Paradigms

- Thousand years ago:
science was empirical
describing natural phenomena
- Last few hundred years:
theoretical branch
using models, generalizations
- Last few decades:
a computational branch
simulating complex phenomena
- Today: **data exploration (eScience)**
unify theory, experiment, and simulation
 - Data captured by instruments or generated by simulator
 - Processed by software
 - Information/knowledge stored in computer
 - Scientist analyzes database/files using data management and statistics

(2017)

Youth want
to use
digital tech

The FOURTH PARADIGM

DATA-INTENSIVE SCIENTIFIC DISCOVERY

Cheminformatics is changing the fundamental
cognitive artifacts

used to represent, manipulate and communicate chemical information.
The Schema educators teach students to solve problems
are based on those artifacts

Develop projects that engage other (non-cheminformatic) chemistry communities

- IUPAC CCE (Committee On Chemical Education)
- ACS CHED CCCE (Committee on Computers in Chemical Education)
- CHF (Chemical Heritage Foundation)

**Note on Curriculum Integration:
and Outreach**

They will use cheminformatics technologies for purposes that are different than what they were designed for.

Is that good?

Should understanding that be a part of outreach?



IUPAC Centennial
"A Common Language for Chemistry"



Develop Educational Booklet on:
***Evolution of Chemical Nomenclature: From
[?Chemicographs?] to InChI***

**Opportunity for
Training/Education/Outreach**

Could Involve

- Chemical Heritage Foundation
- IUPAC CCE
- ACS CINF/CHED/HIST
- ACS CHED CCCE