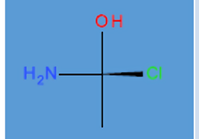
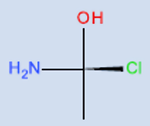
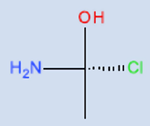
The Status of the IUPAC InChI Chemical Structure Standard – Today and the Future

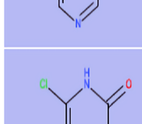
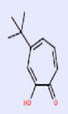
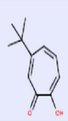
Stephen Heller, NIST & InChI Trust

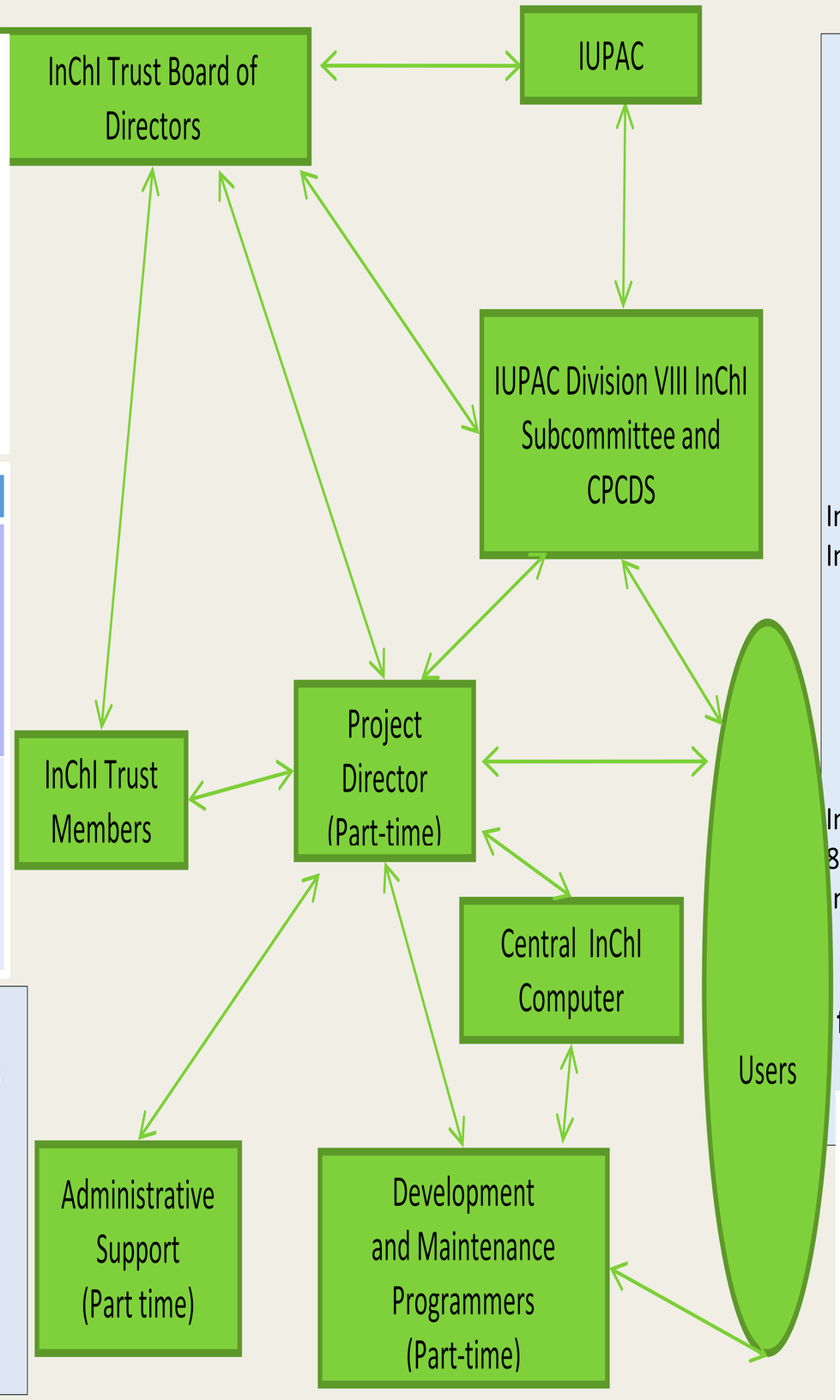
Objectives: Develop a computer algorithm to produce unique label – Chemical Identifier (InChI)

Create Open Source, non-propriety Identifier Used in printed and electronic source

Enable easier linking to diverse data and information compilations; Not a replacement for any existing internal structure representations

Example	Meaning	InChI	MInChI
AND Enantiomer 	**Mixture of**  and 	Standard InChI=1S/C2H6ClNO/c1-2(3,4)5/h5H,4H2,1H3/t2-/m0/s1 (Note: absolute stereochemistry!) Parameter /SUCF or /SREL InChI=1/C2H6ClNO/c1-2(3,4)5/h5H,4H2,1H3/t2-/s2 Parameter /SRAC InChI=1/C2H6ClNO/c1-2(3,4)5/h5H,4H2,1H3/t2-/s3	MInChI=0.0.1S/C2H6ClNO/c1-2(3,4)5/h5H,4H2,1H3/t2-/m0/s1 & C2H6ClNO/c1-2(3,4)5/h5H,4H2,1H3/t2-/m1/s1 /n{1&2}

Example	InChI	MInChI
	InChI=1S/C4H3ClN2O/c5-3-1-6-2-4(8)7-3/h1-2H,(H,7,8)	MInChI=0.00.1S/C4H3ClN2O/c5-3-1-6-2-4(8)7-3/h1-2H,(H,7,8)/n{1&1}
	InChI=1S/C4H3ClN2O/c5-3-1-6-2-4(8)7-3/h1-2H,(H,7,8)	
	InChI=1S/C11H14O2/c1-11(2,3)8-5-4-6-9(12)10(13)7-8/h4-7H,1-3H3,(H,12,13)	MInChI=0.001S/C11H14O2/c1-11(2,3)8-5-4-6-9(12)10(13)7-8/h4-7H,1-3H3,(H,12,13) /n{1&1}
	InChI=1S/C11H14O2/c1-11(2,3)8-5-4-6-9(12)10(13)7-8/h4-7H,1-3H3,(H,12,13)	



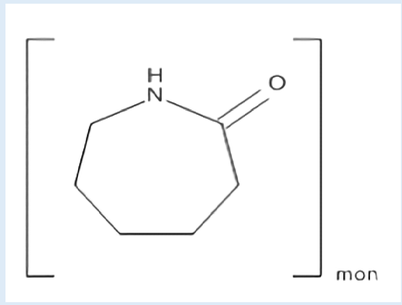
Large Databases with InChIs/InChIKeys

EBI UniChem – 157 million
NIH/NCI – 110 million
NIH/PubChem - 97 million
RSC/ChemSpider – 67 million
Elsevier/Reaxys – 31 million
IUPAC – 0 million

http://www.inchi-trust.org

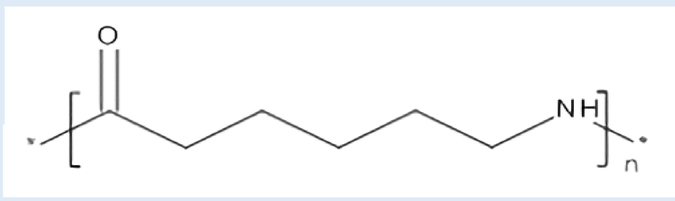
InChI for Polymers

source-based, poly(caprolactam)



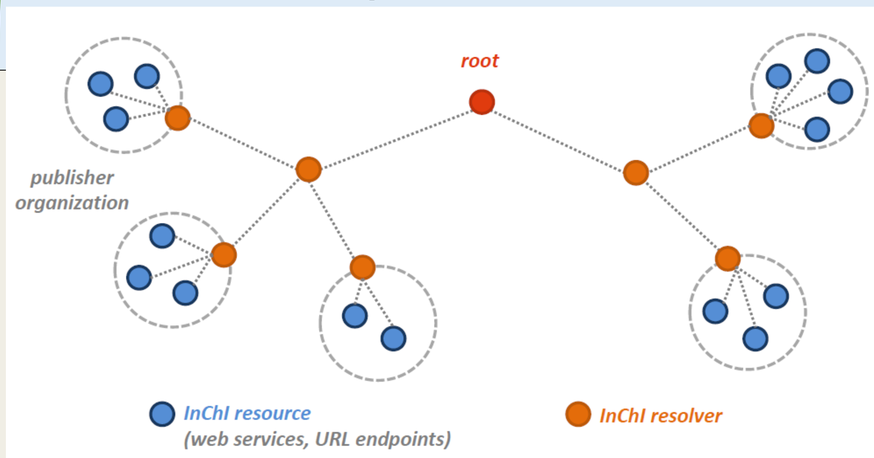
InChI=1B/C6H11NO/c8-6-4-2-1-3-5-7-6/h1-5H2,(H,7,8)/z200-1-8
InChIKey=JBKVHLHDHHXQEQ-ZMQGHSLKBA-N

structure-based representation



InChI=1B/C6H11NO/c8-6-4-2-1-3-5-7-6/h1-5H2,(H,7,8)/z101-1-8(1,2,1,3,2,4,3,5,4,6,5,7)
InChIKey=JBKVHLHDHHXQEQ-DZWZRWJOBA-N

Resolver GOAL: making InChI resources on the Web findable, linkable and browsable by a common, unified protocol



Acknowledgements

(Primarily members for the IUPAC InChI subcommittee and associated InChI working groups)

Steve Bachrach, Colin Batchelor, John Barnard, Evan Bolton, Steve Boyer, Ian Bruno, Steve Bryant, Dominic Clark, Szabolcs Csepregi , Rene Deplanque, Gary Mallard, Nicko Goncharoff, Jonathan Goodman, Guenter Grethe, Richard Hartshorn, Jaroslav Kahovec , Richard Kidd, Hans Kraut, Alexander Lawson , Peter Linstrom, Leah McEwan, Bill Milne, Gerry Moss, Peter Murray-Rust, Heike Nau , Marc Nicklaus, Carmen Nitsche, Matthias Nolte , Igor Pletnev, Josep Prous, Peter Murray-Rust, Hinnerk Rey, Ulrich Roessler, Roger Schenck , Martin Schmidt, Steve Stein, Peter Shepherd, Markus Sitzmann , Chris Steinbeck, Keith Taylor, Dmitrii Tchekhovskoi, Bill Town, Wendy Warr, Jason Wilde, Tony Williams, Andrey Yerin.

Special Acknowledgement: Ted Becker& Alan McNaught for their vision and leadership of the future of IUPAC nomenclature.