RInChI Status

InChI meeting San Diego

August 23rd to 24th, 2019

David Nicolaides, Gerd Blanke, Günter Grethe, Hans Kraut,
István Öri, Jan Holst Jensen, Jonathan Goodman

Current status

- RInChI 1.0 released in March 2017
- Since summer 2018 the RInChI group has been working on the next release.
  - The currently active members are

David Nicolaides	Biovia, Cambridge, UK
Gerd Blanke	StructurePendium Technology GmbH, Essen
Günter Grethe	Poway, CA
Hans Kraut	InfoChem GmbH, Munich
István Öri	ChemAxon Ltd, Budapest
Jan Holst Jensen	Biochemfusion ApS, Copenhagen
Jonathan M. Goodman	University of Cambridge, UK

Current Status

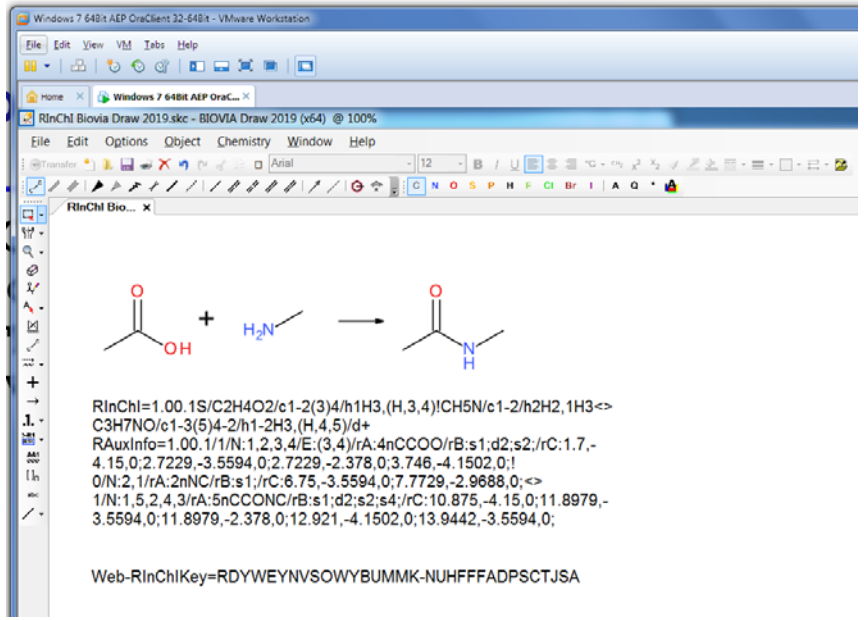
- Public RInChI presentations in 2018/19:
  - Talk at the Biovia community days in Brussels by David Nicolaides and Gerd Blanke, November 2018
    - Presentation of RInChI in Biovia/Draw
  - Talk at the ChemAxon UGM in Budapest by István Öri and Gerd Blanke, May 2019
    - Presentation of RInChI in Marvin by István Öri
  - Talk at the ACS Fall meeting by Gerd Blanke, August 2019
- Publications
  - International chemical identifier for reactions (RInChI), Grethe et al. J Cheminform (2018) 10:22 (May 2018)
    - https://doi.org/10.1186/s13321-018-0277-8

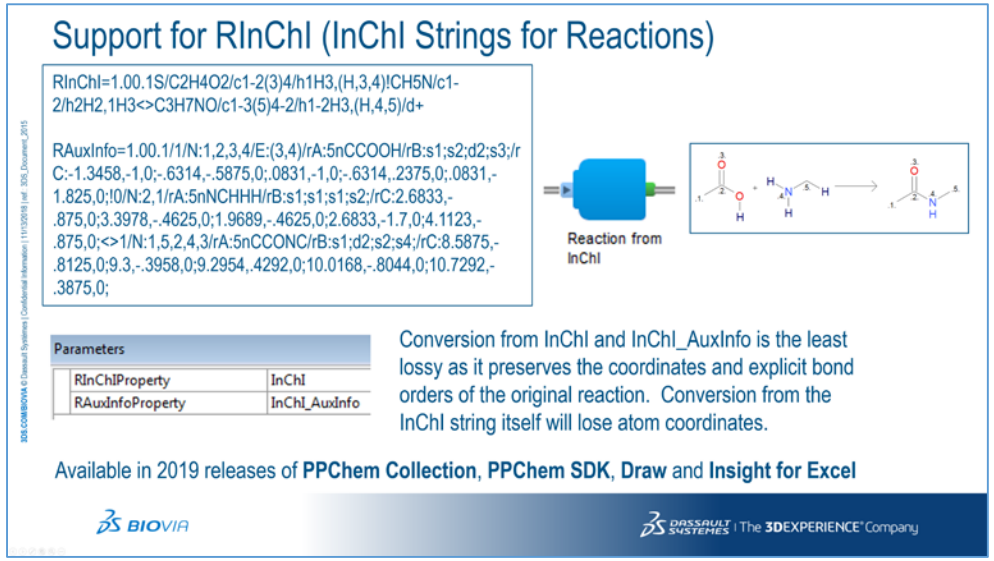
Current Status

- Group meetings
  - Biweekly Skype conferences
  - Group meeting during the InChI Meeting in Cambridge, February 2019
  - Group meeting during the RDKit UGM in Cambridge, October 2018
    - Most of the group members were present
    - Exchange with Jarek Tomczak for the UDM project of Pistoia Alliance
  - Experiences:
    - **A one day personal meeting may save up to 4 months of biweekly phone conferences**

Current Status

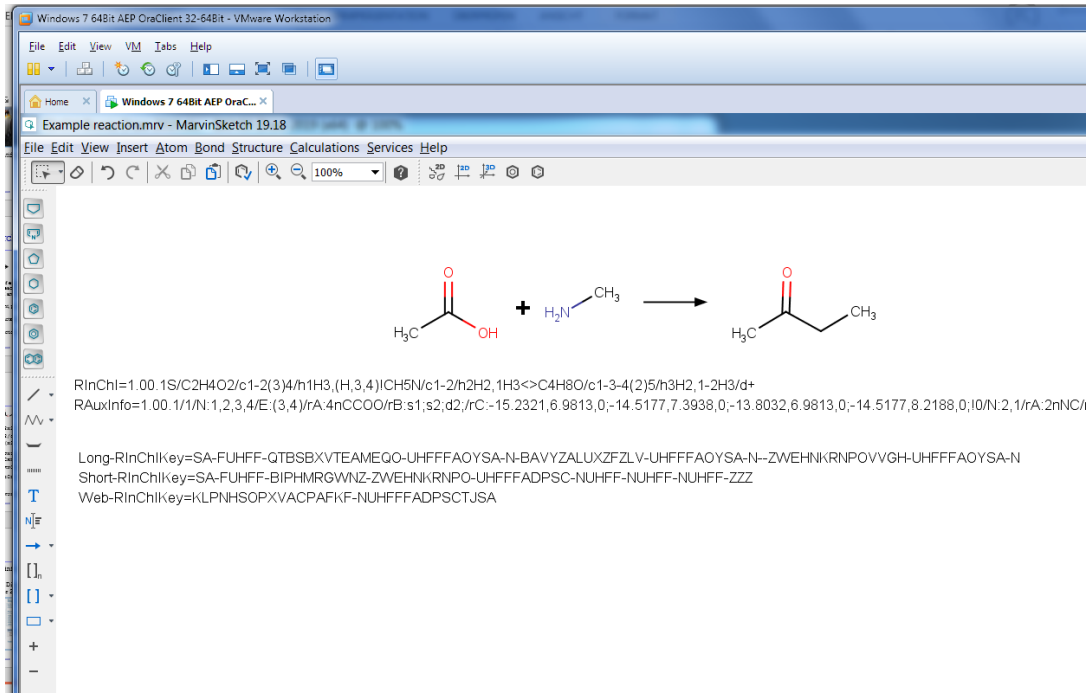
- RInChI implementations by cheminformatics software vendors
  - The Biovia software packages Draw, Direct and Pipeline Pilot (version 2019, released in December 2018) include RInChI





Current Status

- RInChI implementations by cheminformatics software vendors
  - ChemAxon introduced RInChI into its 2019 rlass



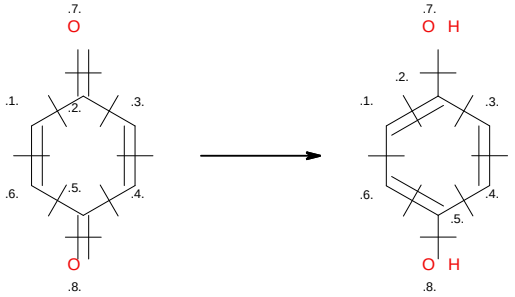
Current Status

- First publisher introducing RInChI
  - Beilstein Institute, Frankfurt (Main), Germany

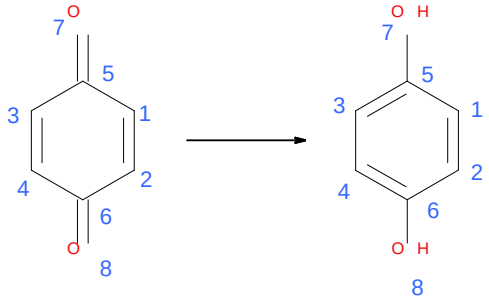
Planned enhancements for the next RInChI release

- Defined for next release
  - Technical issues
  - Additional input and output formats
  - Reaction mapping (MapAuxInfo)
  - Workaround for stereochemistry restrictions
  - Support for AI/ML in reaction prediction
    - Class code layer for reaction similarity clustering and pathway optimization, InfoChem tool

Atom mapping (Example Quinone reduction)



Mapping as defined in RXN file



The numbers in blue represent the InChI numbering

Quinone reduction

```
RInChI=1.00.1S/C6H4O2/c7-5-1-2-
6(8)4-3-5/h1-4H<>C6H6O2/c7-5-1-2-
6(8)4-3-5/h1-4,7-8H/d+
```

Mapping for each atom (including layer and molecule)

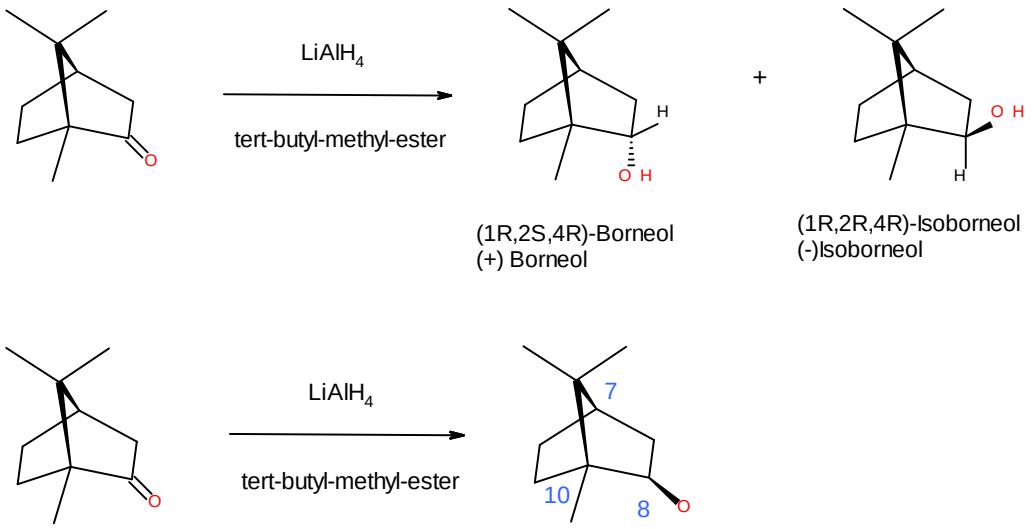
```
2-1-1 <> 3-1-1      2-1-5 <> 3-1-5
2-1-2 <> 3-1-2      2-1-6 <> 3-1-6
2-1-3 <> 3-1-3      2-1-7 <> 3-1-7
2-1-4 <> 3-1-4
```

Skip the trailing 2 and 3 as they are defined by the reaction separator “<>” and the 4th layer is not available

```
1-1 <> 1-1      1-5 <> 1-5
1-2 <> 1-2      1-6 <> 1-6
1-3 <> 1-3      1-7 <> 1-7
1-4 <> 1-4
```

```
MapAuxInfo=1.00.1/1-1<>1-1;1-2<>1-2;1-
3<>1-3;1-4<>1-4;1-5<>1-5;1-6<>1-6;1-7
<>1-7
```

Stereochemistry example



RInChI=1.00.1S/C10H16O/c1-9(2)7-4-5-10(9,3)8(11)6-7/h7H,4-6H2,1-3H3/t7-,10+/m1/s1<>C10H18O/c1-9(2)7-4-5-10(9,3)8(11)6-7/h7-8,11H,4-6H2,1-3H3/t7-,8-,10+/m1/s1/d+/st(3-1-8)mix

- Standard InChI
- Additional stereo layer at the of the RInChI string marked by "/st"
- Use atom identification developed for atom mapping
- Group related centers by brackets
  - Use "mix" to indicate mixtures including racemates
  - Use "pu" to identify pure but unknown centers

Planned enhancements for the next RInChI release

- To be discussed (in the breakout session)
  - How to handle tautomer restrictions
    - Need exchange with Marc's tautomer group
  - Failing reactions
    - Format has not yet been defined
  - Support for AI/ML in reaction prediction
    - Class codes by InfoChem
    - Reaxys formats (discussion with Elena Herzog and Markus Fischer if any and which data formats of Reaxys may be suitable)
    - Reaction properties (ProcAuxInfo)

Planned enhancements for later releases

- Depending on InChI progress
  - Make use of positional isomers and Markush representations for InChI
  - InChI support for enhanced stereochemistry

Thank you