



# InChI @ Elsevier/Reaxys

August 2017



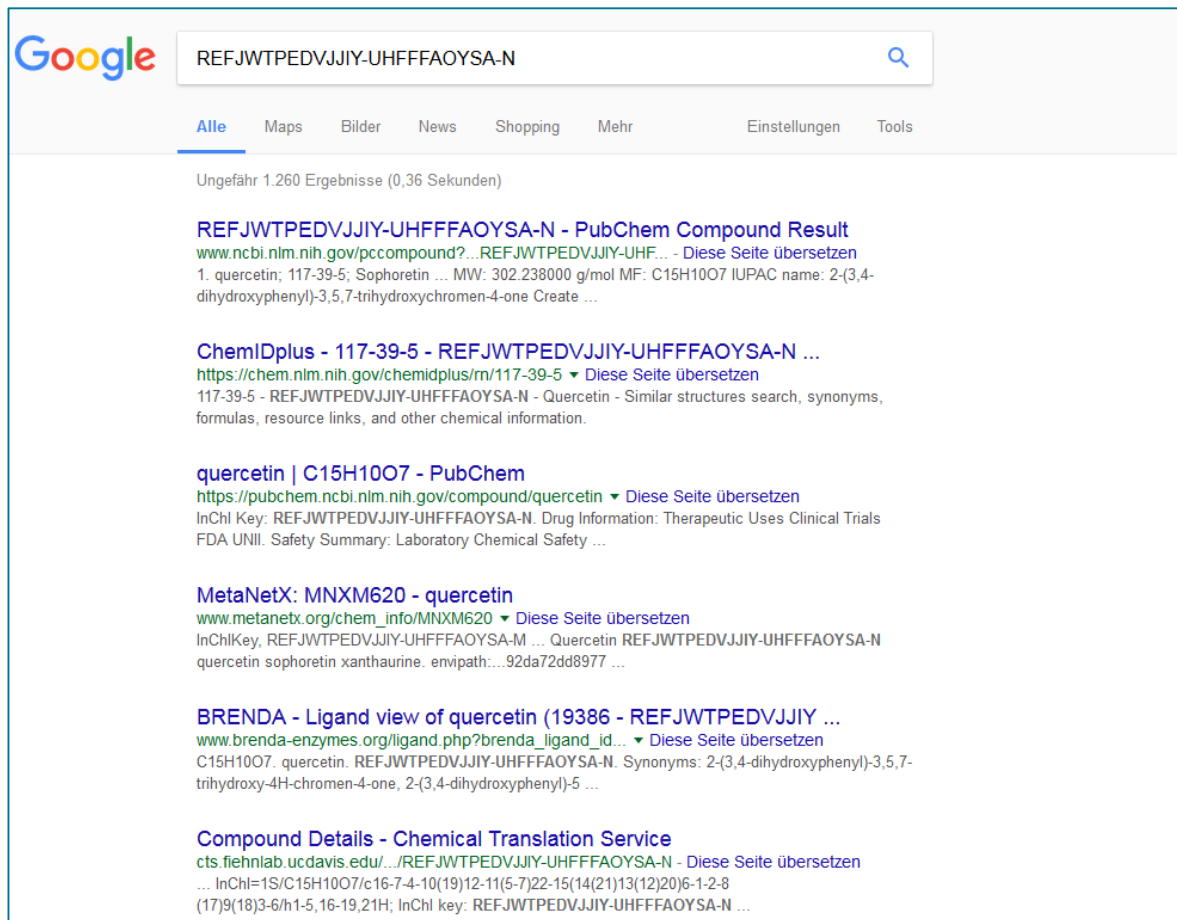
Dr. Juergen Swienty-Busch

August 2017

# Introduction - Why InChI?

- “Identifiers” of chemical structures coded in common, but proprietary formats come with intrinsic issues
  - They may not be unique for a given structure (dependent on applied business rules)
  - Normalization needs to be applied to “register” compounds in databases leading to another, proprietary type of “structure descriptors”
  - proprietary registry numbers based on unique registry strings are too limited and too focused on just one system
    - e.g. Reaxys numbers are simple integers and Reaxys registry strings are not publicly available
- Ideally, there is one unique common identifier for any given substance that is
  - Short
  - Unique
  - Independent
  - human readability is a plus...
  - and that allows to connect disparate data sources (link-able)
- InChI and esp. the InChIKey answers some of those needs...

# InChI at ScienceDirect



Google

REFJWTPEDVJJIY-UHFFFAOYSA-N

Alle Maps Bilder News Shopping Mehr Einstellungen Tools

Ungefähr 1.260 Ergebnisse (0,36 Sekunden)

**REFJWTPEDVJJIY-UHFFFAOYSA-N - PubChem Compound Result**  
[www.ncbi.nlm.nih.gov/pccompound?..REFJWTPEDVJJIY-UHF...](http://www.ncbi.nlm.nih.gov/pccompound?..REFJWTPEDVJJIY-UHF...) - Diese Seite übersetzen  
1. quercetin; 117-39-5; Sophoretin ... MW: 302.238000 g/mol MF: C15H10O7 IUPAC name: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one Create ...

**ChemIDplus - 117-39-5 - REFJWTPEDVJJIY-UHFFFAOYSA-N ...**  
<https://chem.nlm.nih.gov/chemidplus/rn/117-39-5> ▾ Diese Seite übersetzen  
117-39-5 - REFJWTPEDVJJIY-UHFFFAOYSA-N - Quercetin - Similar structures search, synonyms, formulas, resource links, and other chemical information.

**quercetin | C15H10O7 - PubChem**  
<https://pubchem.ncbi.nlm.nih.gov/compound/quercetin> ▾ Diese Seite übersetzen  
InChI Key: REFJWTPEDVJJIY-UHFFFAOYSA-N. Drug Information: Therapeutic Uses Clinical Trials FDA UNII. Safety Summary: Laboratory Chemical Safety ...

**MetaNetX: MNXM620 - quercetin**  
[www.metanetx.org/chem\\_info/MNXM620](http://www.metanetx.org/chem_info/MNXM620) ▾ Diese Seite übersetzen  
InChIKey, REFJWTPEDVJJIY-UHFFFAOYSA-M ... Quercetin REFJWTPEDVJJIY-UHFFFAOYSA-N quercetin sophoretin xanthaurine. envipath:...92da72dd8977 ...

**BRENDA - Ligand view of quercetin (19386 - REFJWTPEDVJJIY ...**  
[www.brenda-enzymes.org/ligand.php?brenda\\_ligand\\_id...](http://www.brenda-enzymes.org/ligand.php?brenda_ligand_id...) ▾ Diese Seite übersetzen  
C15H10O7. quercetin. REFJWTPEDVJJIY-UHFFFAOYSA-N. Synonyms: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-chromen-4-one, 2-(3,4-dihydroxyphenyl)-5 ...

**Compound Details - Chemical Translation Service**  
[cts.fiehnlab.ucdavis.edu/.../REFJWTPEDVJJIY-UHFFFAOYSA-N](http://cts.fiehnlab.ucdavis.edu/.../REFJWTPEDVJJIY-UHFFFAOYSA-N) - Diese Seite übersetzen  
... InChI=1S/C15H10O7/c16-7-4-10(19)12-11(5-7)22-15(14(21)13(12)20)6-1-2-8 (17)9(18)3-6/h1-5,16-19,21H; InChI key: REFJWTPEDVJJIY-UHFFFAOYSA-N ...

# InChI @ Reaxys – Content Integration

The screenshot displays the Reaxys web interface. At the top, the navigation bar includes 'Quick search', 'Query builder', 'Results' (highlighted), 'Synthesis planner', and 'History'. On the right, there are 'Sign in' and help icons. Below the navigation bar, a header indicates '2 Substances out of 9,599 Documents, containing 1,055 Reactions, 398 Targets'. A left sidebar contains a 'Filters and Analysis' section with a '2' indicator and a list of filter categories: 'By Structure', 'Measurements', 'Highest Clinical Phase', 'Targets', 'Parameters', 'Substance Classification', 'Molecular Weight', 'Availability', 'Availability in Market', 'Available Data', 'Document Type', 'Publication Date', and 'Patent Assignments'. The main content area shows a search result for 'aspirin' with its chemical structure and a list of options: 'Find Similar', 'View related Markush', 'Copy structure to query', and 'Open in database'. A secondary menu is open over the 'Open in database' option, listing 'eMolecules', 'LabNetwork', 'PubChem', and 'SigmaAldrich'. Below the options menu, the 'Identification' section for aspirin is visible, listing various identifiers and properties.

Reaxys® Quick search Query builder Results Synthesis planner History Sign in

Filters and Analysis 2 Substances out of 9,599 Documents, containing 1,055 Reactions, 398 Targets Reaxys - 2

Options aspirin  
C<sub>9</sub>H<sub>8</sub>O<sub>4</sub> 180.16 779271 50-78-2

Options

- Find Similar
- View related Markush
- Copy structure to query
- Open in database

Physical Data - 545  
Spectra - 166  
Other Data - 3,500

Preparations - 96  
Reactions - 1,055  
Targets - 398  
Documents - 9,599

eMolecules  
LabNetwork  
PubChem  
SigmaAldrich

aspirin

Identification

Reaxys ID:	779271	Substance type:	isocyclic
Chemical Names:	aspirin, acetylsalicylic acid, acetyl	Linear Structure Formula:	C <sub>6</sub> H <sub>4</sub> OOCCH <sub>3</sub> COOH
CAS Registry Number(s):	50-78-2	No of references:	9600
Molecular Formula:	C <sub>9</sub> H <sub>8</sub> O <sub>4</sub>	Highest Clinical Phase:	Marketed
Molecular Weight:	180.16		
InChI Key:	BSYNRYMUTYPSO		

Substance Label - 98

# InChI @ Reaxys – Link-Outs

The screenshot displays the Reaxys web interface. At the top, the navigation bar includes 'Quick search', 'Query builder', 'Results' (underlined), 'Synthesis planner', and 'History'. On the right, there are 'Sign in' and a help icon. A left sidebar contains a 'Filters and Analysis' menu with various filter categories like 'By Structure', 'Measurement pX', and 'Availability'. The main content area shows search results for '2 Substances' out of 9,599 documents. The first result is 'aspirin' (C<sub>9</sub>H<sub>8</sub>O<sub>4</sub>), with a chemical structure and a 'Synthesize' button. A 'Substance Availability' popup window is open, listing suppliers such as Accelrys' ACD, eMolecules, CambridgeSoft ACX, Sigma Aldrich, Labnetwork, and PharmaPendium. The popup also displays the InChIKey (BSYNRYMUTXBXSQ-UHFFFAOYSA-N) and other substance details like 'Substance type: isocyclic' and 'No of references: 9600'.

Reaxys

Quick search Query builder Results Synthesis planner History Sign in ?

2 Filters and Analysis

By Structure Measurement pX Highest Clinical Phases Targets Parameters Substance Classes Molecular Weight Availability Availability in other databases Available Data Document Type Publication Year Patent Assignee

2 Substances out of 9,599 Documents, containing 1,055 Reactions, 398 Targets Reaxys - 2

Limit To Exclude Export No of References Heatmap

Options aspirin C<sub>9</sub>H<sub>8</sub>O<sub>4</sub> 180.16 779271 50-78-2

Identification Physical Data - 545 Preparations - 96 >  
Druglikeness Spectra - 166 Reactions - 1,055 >  
Bioactivity (All) Other Data - 3,500 Targets - 398 >  
Documents - 9,599 >

Synthesize

Substance Availability

- Accelrys' ACD
- eMolecules
- CambridgeSoft ACX
- Sigma Aldrich
- Labnetwork
- PharmaPendium

isocyclic acid, acetylsalicylic Acid

Substance type: isocyclic  
Linear Structure Formula: C<sub>6</sub>H<sub>4</sub>OOCCH<sub>3</sub>COOH  
No of references: 9600  
Highest Clinical Phase: Marketed

InChIKey: BSYNRYMUTXBXSQ-UHFFFAOYSA-N

Thank you very much for your attention...