

Name to structure, Structure to name, chemicalize.org

Daniel Bonniot de Ruisselet



ChemAxon
Solutions for Cheminformatics

Why use chemical names?

- Easier than drawing
- Familiar
- Used in patents, articles, ...

We will talk about:

- Name generation (structure to name, s2n)
- Name import (name to structure, n2s)
- Name extraction (document to structure, d2s)
- Name correction (OCR-error fixing)
- Name highlighting (chemicalize.org)

Structure to name

Usage:

- Plugin in MarvinView and MarvinSketch
- Label updated in real-time in MarvinSketch
- Batch: Save As: IUPAC Name in MarvinView
- Batch: command line (molconvert, cxcalc)
- Instant JChem

Options:

- Strict IUPAC or traditional
- Timeout

Structure to name

Already stable (focus on n2s)

Comparison between 5.1.0 and 5.2.5 on NCI database (260K structures)

- Both over 99.9% named, 4% changed
- More fused names supported (60% to 66%):
 - 5-methyl-6-azatricyclo[8.4.0.0^{2,7}]tetradeca-1(10),2(7),3,5,8,11,13-heptaene
 - 3-methylbenzo[f]quinoline
- Better support for ions (e.g. -olate)
- Stricter IUPAC numbering and priorities
- Overspecific E/Z labels removed

Name to structure

Usage

- Edit/Import Name... in MarvinSketch
- Automatic format recognition
- Paste a name from the clipboard
- Open IUPAC Name file in MarvinView or MarvinSketch (.name extension)
- Batch from command line (molconvert)

Name to structure: evaluation

Molecule->Name->Molecule on NCI

- 5.1.0: 90.0% names imported, 68.7% identical
- 5.2.2: 97.6% names imported, 94.1% identical
- 5.2.5: 97.8% names imported, 95.9% identical

Pubchem data

- 5.1.0: 88.8% names imported, 94.0% identical
- 5.2.2: 98.3% names imported, 95.6% identical
- 5.2.5: 98.3% names imported, 96.1% identical

Name to structure: +33% speed

Customization (new)

Extend name-to-structure conversion using your in-house data

- Dictionary file
- Simple API:
 - Database lookup
 - Webservice lookup
 - ...
- Fully flexible

Document to structure (new)

Goal

- Process documents containing text
- Recognize chemical names
- Convert them to structures
- Return locations, names and structures

Formats

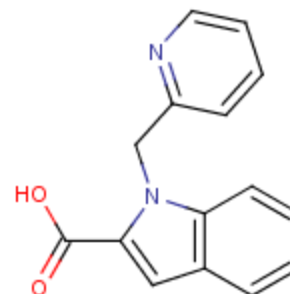
- Implemented: text, html, xml
- Planned: PDF, Doc, ...

OCR error fixing (new)

Scanned texts contain numerous recognition errors

Examples:

- L (small L) instead of 1
- I instead of l (ll, iL)



l-(pyridin-2-ylmethy l)- 1H- indole-2-carboxylic acid

Uses:

- By default in d2s
- Option in n2s (upcoming)

- Adds structural information to existing public webpages
- Popup window with structure image
- Link to structure predictions (logP, pKa, ...)
- Searchable structure->webpage index
- Could be installed natively on custom website, with custom features

Benedict's reagent - Wikipedia, the free encyclopedia - Mozilla Firefox

The chemical names on this page were annotated by the [chemicalize.org](http://www.chemicalize.org) service. You can see the original page: http://en.wikipedia.org/wiki/Benedict%27s_reagent.

Log in / create account

article discussion edit this page history

Benedict's reagent

From Wikipedia, the free encyclopedia

Benedict's reagent (also called **Benedict's solution** or **Benedict's test**) is a chemical reagent named after an American chemist, Stanley Rossiter Benedict.^[1]

Benedict's reagent is used as a test for the presence of all monosaccharides, and generally also reducing sugars. These include [glucose](#), [galactose](#), [mannose](#), [lactose](#) and [maltose](#). Even more generally, Benedict's test will detect the presence of [aldehydes](#) (except aromatic ones), and alpha-hydroxy-ketones, including strictly a reducing sugar, it is an alpha-hydroxy-ketone [mannose](#) by the base in the reagent.^[2]

Benedict's reagent can be prepared from [sodium carbonate solution](#).^[3] It is often used in place of Fehling's solution.^[4] These are precipitated as red copper(I) oxide which is insoluble in water.

Chemical test

To test for the presence of monosaccharides and reducing disaccharides, a small amount of Benedict's reagent is added. During the reaction, the color of the solution changes from blue (with no glucose present), green, yellow, orange, red, and finally brown (with high glucose present).^[4] A color change would signify the presence of glucose. The common disaccharides lactose and maltose are directly detected by Benedict's reagent, because each contains a glucose with a free reducing aldehyde moiety, after isomerization.

[Sucrose](#) (table sugar) contains two sugars (fructose and glucose) joined with by their [glycosidic bond](#) in such a way as to prevent the glucose isomerizing to aldehyde, or the fructose to alpha-hydroxy-ketone form. Sucrose is thus a non-reducing sugar which does not react with Benedict's reagent. Sucrose indirectly produces a positive result with Benedict's reagent if heated with dilute hydrochloric acid

Chemicalized by ChemAxon

Report error

Recap

- Name-to-structure and structure-to-name available and improving fast
- Document-to-structure just released
- Extend using your in-house dictionaries and databases
- Try it, send feedback, we're listening!

Find out more

- **Product descriptions & links**

www.chemaxon.com/products.html

- **Forum**

www.chemaxon.com/forum

- **Presentations and posters**

www.chemaxon.com/conf

- **Download**

www.chemaxon.com/download.html

The collage consists of four screenshots from the ChemAxon website:

- Top-left:** The main website page with navigation tabs (Home, Products, Documentation, Download, Forum, About us, Contact us) and a 'Recent News' section listing updates from 2007.
- Top-right:** The 'Upcoming Meetings' section, listing events such as 'Seminars series: Bridging the corporate and desktop' and 'Software Pipeline Pilot User Group Meeting'.
- Bottom-left:** A 'Marvin and Calculator Plugin Demo' page. It features a rich editing list with bullet points: 'wide range of file types supported: MOL, MOL2, SDF, RXN, RCF (V2000/3000), SMILES, SMARTS/SMARTS (Impure), MW, ICHN, CML, PDB etc.', 'Copy and paste between different editors', 'Addressed groups', 'Pre-loaded structure templates and "My Templates"', and '3D editing'.
- Bottom-right:** The 'Technical Support Forum' page. It includes a search bar and a table of support topics. The table has columns for 'Support', 'Topics', 'Posts', and 'Last Post'.

| Support | Topics | Posts | Last Post |
|---|--------|-------|--|
| Structure editing, viewing and file formats Support for MarvinSketch, MarvinView and MarvinCenter | 827 | 4081 | Fri Feb 02, 2008 9:00 am Global #0 |
| MarvinSpace Development/ discussion area for Open3D, 3D rendering main/3D molecule viewer | 60 | 356 | Thu Jan 31, 2008 9:16 pm Global #0 |
| Structure based prediction and Chemical Terms | 405 | 374 | Fri Feb 01, 2008 10:02 am Global #0 |
| Support for Calculator Plugins operation through cracks, API, Marvin, Instant XChem and Chemical Terms | 351 | 374 | Fri Feb 01, 2008 10:02 am Global #0 |
| Structure search and chemical database Support for XChem Base and XChem Cartridge | 426 | 3952 | Fri Feb 01, 2008 10:43 am Global #0 |
| Instant XChem Discussion area for Instant XChem (Structure database GUI, both Calculator Plugin processor etc) | 173 | 496 | Wed Jan 30, 2008 9:10 pm Global #0 |
| Structure conversion/operation / standardization Support for Standardizer | 76 | 345 | Wed Jan 30, 2008 9:30 pm Global #0 |
| Library enumeration, virtual synthesis and metabolite generation Support for Reactor and Fragmenter | 116 | 577 | Thu Jan 28, 2008 9:02 am Global #0 |
| Virtual screening, clustering and molecular descriptors Support for Screen and Shaper | 308 | 471 | Thu Jan 28, 2008 1:00 am Global #0 |
| License Issues Support for the technical questions related to license handling | 4 | 7 | Thu Jan 28, 2008 10:32 am Global #0 |