

SMILES, CAS and InChI in blogs: Greasemonkey

Egon Willighagen 

Published December 17, 2006

Citation

Willighagen, E. (2006, December 17). SMILES, CAS and InChI in blogs: Greasemonkey. *Chem-bla-ics*. <https://doi.org/10.59350/ygxfh-xfs36>

Keywords

Chemistry, Userscript, Smiles, Pubchem, Inchi

Sunday, December 17, 2006

Counting stereoisomers from the molecular formula

We all know the combinatorial explosion when calculating the number of possible constitutional isomers (see [wp:structural isomorphism](#)) of a certain molecular formula. For example, C₂H₆ has only one constitutional isomer (ethane, InChI=1/C2H6/c1-2/h1-2H3 [Google](#), [PubChem](#)), and C₄H₁₀ has only two. Especially, breaking symmetry by replacing one carbon by another element, or replacing a single by a double bond, increases the number sharply. For example, C₇H₁₆ has only nine constitutional isomers, while replacing two single bonds by two double bonds, creating C₇H₁₀, increases this number to 499! Then, replacing in the last formula, one carbon by an oxygen adds another few, totaling 747 isomers.

Now, C₈H₈NBr has at least **649 thousand** constitutional isomers, and I am quite interested in being able to know the number of isomers beforehand, without having to generate the structures itself (for example, using [CDK's GENMDeterministicGenerator](#)).

InChI=1/C8H8BrN/c9-7-1-2-8-6(5-7)3-4-10-8/h1-2,5,10H,3-4H2 [Google](#), [PubChem](#) is one of the isomers.

Copyright

Copyright © Egon Willighagen 2006. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

As follow up on my [Including SMILES, CML and InChI in blogs](#) blog last week, I had a go at [Greasemonkey](#). Some time ago already, [Flags and Lollipops](#) and [Nodalpoint](#) showed with two cool mashups (one Connotea/Postgenomic and one Pubmed/Postgenomic) that userscripts are rather useful in science too. I can very much recommend the PubMed/Postgenomic mashup, as PubMed has several organic chemistry journals indexed too!

So, how does this relate to my blog of last week? Well, would it not be nice that if your blog uses the markup as suggested in that [blog](#), that you automatically get links to [PubChem](#) and [Google](#)? That is now possible with a small GPL-ed Greasemonkey script called [blogchemistry.user.js](#).

The [Greasemonkey plugin](#) requires [Firefox](#) to be installed. If ready, install the script by clicking this link earlier, and the Greasemonkey will ask you if you want to install the script. After, check the output for this RDFa markup content:

- a SMILES: `[CCO]{.chem:smiles xmlns:chem="http://www.blueobelisk.org/chemistryblogs/"}`
- a CAS registry number: `[50-00-0]{.chem:casnumber xmlns:chem="http://www.blueobelisk.org/chemistryblogs/"}`
- and an InChI: `[InChI=1/CH4/h1H4]{.chem:inchi xmlns:chem="http://www.blueobelisk.org/chemistryblogs/"}`

It should look like the output for this blog item:

Note the superscript PubChem and Google links.