

Cb comments for InChI's

Egon Willighagen 

Published May 5, 2007

Citation

Willighagen, E. (2007, May 5). Cb comments for InChI's. *Chem-bla-ics*. <https://doi.org/10.59350/z1vb6-w9v12>

Keywords

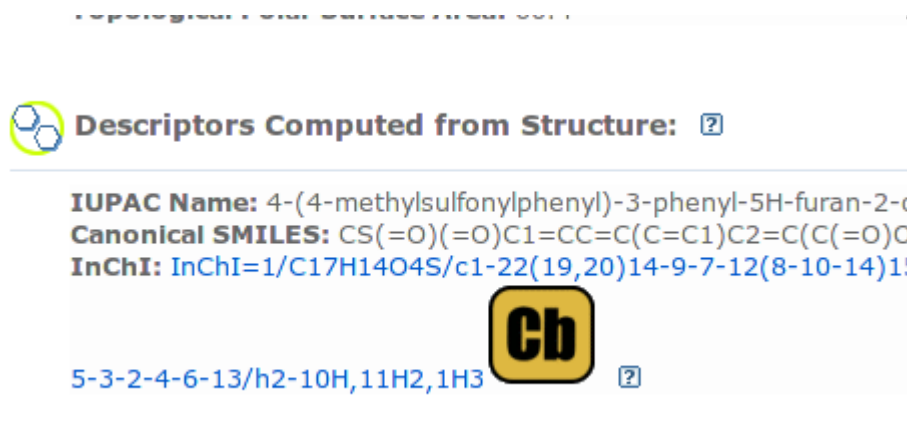
Cb, Inchi, Userscript, Rdf

Copyright

Copyright © Egon Willighagen 2007. 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.

About a year ago [Pedro wrote a Greasemonkey script](#) to add comments from [PostGenomic.com](#) to table of contents of scientific journals. [Noel extended](#) it with support for [Chemical blogspace](#) (see also [this earlier item](#)). Now, the later website is maintained by me, and I [extended the aggregator software with molecule support](#), for example to show *hot molecules on the frontpage* (at some point [my patches will be backported into mainstream](#). Euan, why not invite me to London HQ in, say, June?).

So, when we can show comments from blogosphere for journal articles, why can't we do that for molecules too? Sure we can. Just needs some hacking. Right, and done that today. The scripts works for [PubChem](#):



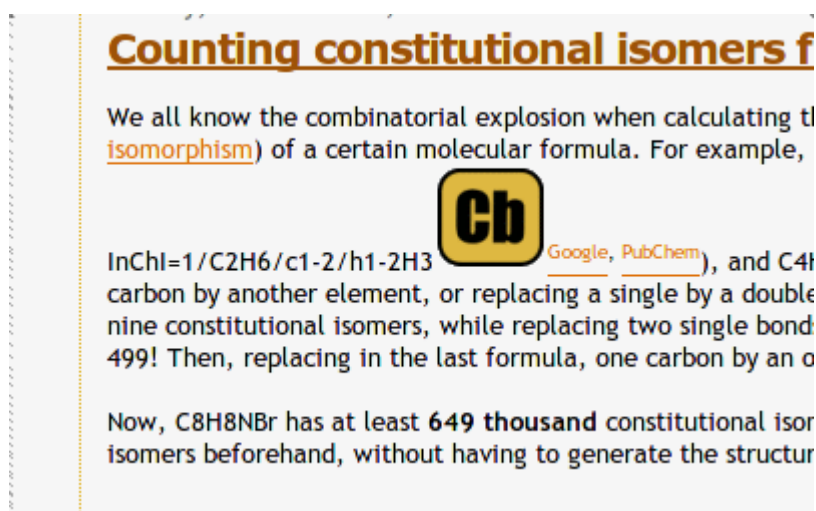
Descriptors Computed from Structure: ?

IUPAC Name: 4-(4-methylsulfonylphenyl)-3-phenyl-5H-furan-2-one
Canonical SMILES: CS(=O)(=O)C1=CC=C(C=C1)C2=C(C(=O)O)C=C2
InChI: InChI=1/C17H14O4S/c1-22(19,20)14-9-7-12(8-10-14)15

5-3-2-4-6-13/h2-10H,11H2,1H3 **Cb** ?

Works for any `<a href>` element with an URL to PubChem like [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pccompound&term=%22InChI=1/CH4/h1H4%22\[InChI\]](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pccompound&term=%22InChI=1/CH4/h1H4%22[InChI]). BTW, while the URL is not very readable, this might actually be a good way to [hide InChIs](#), though I am sure Google will not index this InChI either.

And it also works for [semantically marked up InChI's \(using either microformats or RDFa\)](#) :



Counting constitutional isomers f

We all know the combinatorial explosion when calculating the [isomorphism](#) of a certain molecular formula. For example,

InChI=1/C2H6/c1-2/h1-2H3 **Cb** [Google](#), [PubChem](#)), and C4H10 has 2 carbon by another element, or replacing a single by a double bond has 9 constitutional isomers, while replacing two single bond by a double bond has 499! Then, replacing in the last formula, one carbon by an oxygen atom has 1000 isomers.

Now, C8H8NBr has at least **649 thousand** constitutional isomers beforehand, without having to generate the structures.

You'll notice here that it is friendly with my [Sechemtic script to make links to Google and PubChem](#).

chem-bla-ics

The tools to make this happen involves a new Greasemonkey script (based on Noels code), and a few patches to the Postgenomic.com software. The user script can be downloaded [here](#). An entry on the [Blue Obelisk userscript page](#) will follow; check that page for more goodies.