

Counting constitutional isomers from the molecular formula

Egon Willighagen 

Published December 17, 2006

Citation

Willighagen, E. (2006, December 17). Counting constitutional isomers from the molecular formula. *Chem-bla-ics*. <https://doi.org/10.59350/r9gwr-k2s81>

Keywords

Cheminf, Cdk

Abstract

Update: check these two papers.

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.

Update: check [these two](#) papers.

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]{.chem:inchi xmlns:chem="http://www.blueobelisk.org/chemistryblogs/"}), 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]{.chem:inchi xmlns:chem="http://www.blueobelisk.org/chemistryblogs/"} is one of the isomers.

So, my question: is anyone aware of free code (in order of preference: 1. LGPL, 2. BSD/MIT,

1. opensource, 4. free) to calculate or estimate the number of constitutional isomers for a certain molecular formula. An estimate would already be nice. Ideally, I would implement this bit of code into the CDK, but otherwise, just knowing the number of isomers for C₈H₈NBr would be nice :)

Additionally, any relevant, recent literature recommendations are most welcomed. I am aware of the use of polynomials, but literature I have seen so far just focuses on molecules of a certain architecture, and it not able to come up with a guess based on the molecular formula alone.