
Subject: Incorrect dearomatization of indazole structures

Posted by [Andrievs](#) on Tue, 30 Nov 2021 10:48:26 GMT

[View Forum Message](#) <> [Reply to Message](#)

I am having issues with SD files where the structure is an indazole superstructure.
I have attached a SD-file as examples with 3 structures.

The problem is that during the dearomatization process DataWarrior incorrectly places the double bonds resulting in a formation of s stereocenter.

See attached image for visual differences.

Is there a way to have the dearomatization be done correctly with out using additional software?

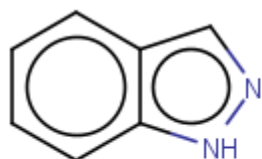
P.s. I am using DataWarrior version 5.5.0

Sincerely, Andrievs

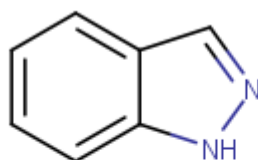
File Attachments

1) [indazole_aromatized.sdf](#), downloaded 586 times

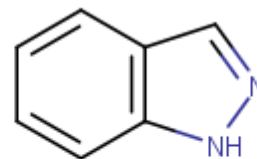
2) [DataWarrior.png](#), downloaded 808 times



**Original
aromatized form**



**Possible kekule
form I**



**Possilbe kekule
form II**

Subject: Re: Incorrect dearomatization of indazole structures

Posted by [thomas](#) on Tue, 30 Nov 2021 13:53:31 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi Andrievs,

Your SD-File includes a non-standard Marvin extension that is not understood by DataWarrior:
MRV_IMPLICIT_H

Marvin uses that feature to define explicit hydrogen atoms, which usually is not needed, because molfiles

usually are in a kekulized form. If, however, molfiles just contain bonds marked to be aromatic, the

Kekulizer needs more information, if multiple potentially protonated nitrogen atoms are present. Standard molfiles simply don't have a place to store that information. The only way to make your SD-files compatible would be to extend DataWarrior's molfile parser to support Marvin's non-standard extensions.

Hope this explains it, Thomas
