Subject: Re: Bug in how stereochemistry is reported
Posted by thomas on Tue, 09 Mar 2021 15:01:52 GMT
View Forum Message <> Reply to Message

Dear Simon,

I tried to confirm the issue with CC[C@H](O)[C@@H](CC1=CC=CC=C1)C#N with the official 5.2.1 version and with the current dev version. If I paste the smiles into an editor or if I paste a tiny table with the smiles being part of it, DataWarrior always created the correct structure with two stereo centers. Can you please give me a procedure where it doesn't work?

The issue with the chiral flag not set in the SD-file is a V2000 problem, not V3000. The fifth entry in the counts line of a V2000 molfile encodes, whether the structure is a racemate or not (0 or 1). For racemates DataWarrior must normalise the configuration, because DataWarrior stores canonical structure representations. A normalization means here that it determines in a reproducible way, which of the two enantiomers is shown and stored. It cannot just store the input form, because then there would be two different representations encoding the same structure.

What I could do, is to provide an option when loading V2000 SD-files to consider racemic molfiles with encoded stereo centers as enantiomerically pure, just assuming that the chiral flag is meant to be set. This, of course, is dangerious, especially if more than one stereo centers are present. Depending on the source of the molfile, however, one may have good reason to believe that the software 'forgot' to set the chiral flag, e.g. when the structures represent conformers with 3D-coordinates.

What do you think?

Thomas