Subject: Re: Stereoisomers conformer generation and Inchikeys
Posted by nbehrnd on Thu, 02 Sep 2021 20:18:30 GMT
View Forum Message <> Reply to Message

Dear Mattia,

based on visual inspection of the structure formulae, entries #2 and #3 depict the same isomer. Thus, there are three different compounds in the table for which one anticipates three different SMILES strings, InChis, or eventually InChiKeys.

Based on the visual representation, I redrew the structures in PerkinElmer's ChemDraw test page.[1] This page equally permits the export in SMILES, InChi, and InChiKey format. There is no match of the later with those you report as generated by DW. The ChemDraw InChiKeys match the ones you report by RDKit. It is possible to process ChemDraw's SMILES strings into InChiKeys by OpenBabel, too, yet with the same result; match with RDKit/ChemDraw, no match with DW.

Importing the three different ChemDraw SMILES strings into DW yields different structure representations. However already the InChi for these match each other. Since the InChikeys only are a hash of the InChi, it is not surprising that these derivates equally match each other (again). If interested, compare with attached .dwar and Emacs .org file.

Based on these observations, pending correction of this local problem for DW, I would suggest to use the InChiKeys by RDKit. On occasion, implementation of the underlying rules indeed may cause problems which remain unidentified for quite some time until comparison with other programs.[2]

Norwid

[1] https://chemdrawdirect.perkinelmer.cloud/js/sample/index.htm l#
[2] https://mattermodeling.stackexchange.com/questions/6460/rdki
t-and-pysmiles-results-differ-on-some-smiles-strings

Addition:

Since DW is able to generate a random library of molecules and to assign for these SMILES and InChiKeys, I wrote a little Python script to compare InChiKeys by DW with those generated by OpenBabel. (This only is low level of concept/doodle only.) For two runs (10 and 250 molecules), the ratio of SMILES for which both programs assign different InChiKeys over the grand total of structures submitted is a little bit higher, than anticipated. Perhaps a complementary check with RDKit is useful before Thomas steps in.

File Attachments

1) 2021-09-02_InChi_test.dwar, downloaded 300 times
2) 2021-09-02_DW.org, downloaded 319 times
3) dissenter_check.zip, downloaded 247 times