Subject: Stereochemical questions
Posted by don on Mon, 07 Nov 2022 16:16:32 GMT
View Forum Message <> Reply to Message

Hi all,
I'm a newb to DataWarrior and have a couple of questions about identifying enantiomers from a compound set.

1) How do I filter a dataset so that the only compounds displayed are single stereoisomers of absolute configuration (filtering out both racemic compounds and compounds with no stereochemistry)?

2) Is there a way one can use DataWarrior to filter a set of compounds so that the only compounds displayed are single enantiomers AND their corresponding enantiomers? In other words, the search would only identify pairs of enantiomers (not single enantiomers, not racemic compounds, not compounds without stereochemistry).

Thanks in advance for your help.
don

Subject: Re: Stereochemical questions
Posted by thomas on Mon, 07 Nov 2022 16:58:48 GMT
View Forum Message <> Reply to Message

Hi Don,

1) You could add a new column that contains the stereo isomer count and relation ("Chemistry->From Chemical Structure->Calculate Properties", then switch to tab "Functional Groups" and select "Stereo isomer count...". This determines the stereo category for every compound with at least one stereo center, e.g. "racemate", "unknown chirality", "4 diastereomers", etc.

With the new column you get a new category filter, where you can select "this enantiomer". Deselect all other categories. The remaining visible rows are all compounds with at least one stereo center and all stereo centers with absolute configuration.

2) Just continue from the state above: Do "File-New From Visible Rows" to get a new Window with the pure enantiomers. Then do "Chemistry->From Chemical Structure->Add Canonical Code...". Unselect 'Distinguish tautomers'. Keep the other two options selected and press OK. You get a new column, which contains a unique code for every different compound neglecting stereo features. Rows that contain enantiomers of otherwise identical molecules will get the same code.

Now create a row list that contains all unique rows considering this canonical code (List->Create Row List From->Unique Rows..." and select the column containing the canonical code). Unique rows are those, which exist only once in the dataset. With the list comes a new filter. In the filter select that list and invert the filter by clicking its 'invert filter' button. Now only those compounds

are visible, which have a different stereo isomer in the same dataset (which, of course, is also visible.

You could sort the table by clicking the column header of the canonical code column.

Interesting question!

Thomas

Subject: Re: Stereochemical questions
Posted by don on Tue, 08 Nov 2022 15:54:52 GMT
View Forum Message <> Reply to Message

Thanks, Thomas! This worked perfectly. I really appreciate the very detailed instructions. Greatly appreciated!
