Subject: Re: How to perform general R group searches in DW
Posted by thomas on Tue, 11 Jan 2022 15:41:44 GMT
View Forum Message <> Reply to Message

DataWarrior uses an atom table derived from MDL's original when defining the Molfile format. This table contains R, R1 till R16, A, the amino acids, e.g. 'Ala' and a few others as pseudo atoms. Therefore, you can use these labels, but there is no logic behind them. If a Chembl molecule would contain an R as pseudo atom, you could find it with a substructure search.

You could use the following substructure definintion to exclusively find alkyl alcohols:

This is a C-O with the carbon defined to have no pi electrons and the oxygen set to have no further non-H neighbour. Then there are to exclude groups:

1) The '!C' is an ANY atom except carbon. This prevents any matches to molecules with non-carbon atoms in addiction to the required OH.

2) A C=C double bond to prevent double bonds anywhere. Actually this is not quite correct. Better would a carbon atom with 'at least one pi-electron' as exclude group.

Two R-groups at an otherwise specified molecules could be done analogously.

File Attachments

1) t.png, downloaded 802 times