Subject: Re: Calculating similarity to one query compound
Posted by thomas on Mon, 22 May 2023 19:55:59 GMT
View Forum Message <> Reply to Message

SMARTS and IDCodes use slightly different concepts. Therefore, a one-to-one conversion is not possible. Not all atom/bond query features, which are available in SMARTS, do exist in OpenChemLib (i.e. DataWarrior) molecule fragments and vice versa. DataWarrior does not support recursive substructures and SMARTS don't support 'exclude groups'. Most query features are translated by DataWarrior close enough, but not necessarily in an exact way. For instance, atoms that are considered aromatic in the Daylight (SMARTS) world are not necessarily considered aromatic in OpenChemLib. And OpenChemLib distinguished aromatic and delocalized atoms. The latter is not existing in SMARTS. During the recent year DataWarrior's SMARTS parser and creator were improved to closer translate features, but it will never be 100%, because the underlying concepts don't match. Another reason is that SMARTS don't support enhanced stereo recognition.

The conversion is done in DataWarrior this way: If you paste a table that contains SMARTS (or open a text file) then DataWarrior should recognize that and a create chemical structure (reaction) column from it, which is column containing IDCodes and is tagged to contain chemical information. To create a SMILES column (or SMARTS, if a structure column contains query features) call "Chemistry->From Chemical Structure->Add SMILES Code".

Your question referred to reactions rather than molecules and I realize that there is no "Chemistry->From Chemical Reaction->Add SMILES Code". You can, however, use "Copy From Reaction->Reaction As->Reaction SMILES". I could easily add the functionality "Chemistry->From Chemical Reaction->Add SMILES Code" with the limitations discussed above, if that would help?