Subject: Re: Chrial drawn structures in dwar / sdf different
Posted by thomas on Tue, 12 Jun 2018 21:47:24 GMT
View Forum Message <> Reply to Message

In principle you are right. DataWarrior and SD-Files, both distinguish unknown chirality from a racemic.
In molfiles stereo centers are usually represented by an up- or down-bond at the stereo center. In addition the wide-spread molfile version 2 has one flag (wrongly called 'chiral' flag), which defines,
whether the molecule is racemic (flag=0) or the drawn enantiomer (flag=1). Correctly, a racemate would have
a down- or up-bond and the chiral flag would be 0. The molfile also allows an 'either' bond, which would
describe a mixture and thus supposedly also a racemic situation.

When DataWarrior writes an unknown stereo center into the molfile, it doesn't write any stereo information
making the stereo center effectively unknown.

The problem starts when the molfile is read again. If you get SD-files from chemicals' providers, they often
encode racemates of molecules with one stereo center the same way, i.e. they don't use one of the two proper
methods described above to define the racemate. They just don't write any stereo information. Therefore,
molecules with one stereo center, which have no information, are usually meant to be racemates. For this
reason DataWarrior also treats them as racemate. Another ugly example is that many software packages
read molecules as pure enantiomers if a stereo center has an up- or down-bond, even if the chiral flag is
not set. This means, encoding a racemate the right way often creating a wrong result. This may also be
the reason for many software providers to rather write no stereo information in case of racemates.

If you need to write structure files with DataWarrior to read them again from DataWarrior I strongly suggest to
use the native format, because the conversion to and from molfiles is not guaranteed to give the same result.
This is especially true for substructures, which contain query features.

Thomas