Subject: Re: mean from 'n' in pKaInWater.dwar
Posted by thomas on Mon, 24 May 2021 10:35:43 GMT
View Forum Message <> Reply to Message

Hi Peter,

the pKa in water file was given to me about 20 years ago. Allegedly, these are all literature values, but I don't have an explanation for the columns. I later intended to use it as training file for developing a pKa-prediction algorithm, what never materialized. However, in order to use the input file added some manual curation, e.g. to assign basic/acidic atoms, and put the molecule in the correct protonation state, e.g. protonate the first amine, if the value referred to the second protonation step. Tautomeric states were normalized, but not necessarily to the dominant tautomer. If there were multiple values reported for the same molecule and same atom, then a mean was calculated. The comments 'a1' 'a2', 'b1', typically referred to multiple (de-)protonatable atoms. If a 'b' was mentioned, the value often was a pKb (14-pKa). However, this was not consistent. Thus, there was some judgement involved, whether it would be a pKa or pKb.

Hope, this helps.

Thomas