

---

Subject: mean from 'n' in pKaInWater.dwar  
Posted by [peter](#) on Fri, 21 May 2021 16:44:42 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Please could someone help me with understanding the reference data pKaInWater.dwar provided with DataWarrior?

In the method column it sometimes has "mean from n". I'm assuming this means that the pKa is the average of n experimental values? Is this correct, and if so, is there any way of obtaining the original experimental values?

Also, in the "type" column it states whether the pKa is thermodynamic or apparent. However, it sometimes has a/apparent a1/apparent b/apparent etc., and I'm not clear what the a and b stand for. I'd guess acid and base, but this doesn't seem quite right. Also, I'm not clear what the difference between a1/apparent and a2/apparent would be; do the numbers indicate different protic sites? Or do they indicate molecules with equivalent smiles but different stereochemistry? Or is it something else?

Many Thanks,

Peter

---

---

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

---

---

Subject: Re: mean from 'n' in pKaInWater.dwar  
Posted by [peter](#) on Tue, 25 May 2021 11:27:04 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hi Thomas,

Thank you very much, that helps a lot, I was thinking it must be something along these lines.

Best Wishes,

Peter

---