Subject: TPSA
Posted by sansun on Mon, 03 Aug 2020 11:13:30 GMT
View Forum Message <> Reply to Message

I found disagreement between TPSA values calculated from DW and RDKit although both seem to follow same methodology by Peter Ertl et al. in J. Med. Chem. 43, 3714-3717 (2000).

In some case the difference is quite large e.g. 87 vs 252

However, Spearman correlation is > 0.9 for my data set (N > 20000).

Subject: Re: TPSA
Posted by thomas on Tue, 04 Aug 2020 09:24:55 GMT
View Forum Message <> Reply to Message

One reason may be that DataWarrior considers the largest fragment only by removing counter ions, water molecules and other small fragments from the main ingredient before the calculation.

Subject: Re: TPSA
Posted by sansun on Tue, 04 Aug 2020 10:20:26 GMT
View Forum Message <> Reply to Message

I performed calculations after extracting the largest fragments and used the same file for both programs.
However, I find useful information on the following link which says RDkit doesn't use P and S atoms for TPSA calculation.

Does DW uses these atoms?

https://www.rdkit.org/docs/RDKit_Book.html

Subject: Re: TPSA
Posted by thomas on Mon, 17 Aug 2020 12:01:05 GMT
View Forum Message <> Reply to Message

DataWarrior sticks to the original publication, which includes contributions for S and P