LogP as an Archetype for modern QSPR

Markus Mühlbacher[1,2], Ahmed El Kerdawy[1], Matthias Hennemann[1],
Johannes Kornhuber[2], Timothy Clark[1]

[1] *Computer Chemistry Center, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstrasse 25, 91052 Erlangen, Germany.*

[2] *Department of Psychiatry and Psychotherapy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Schwabachanlage 6, 91054 Erlangen, Germany*

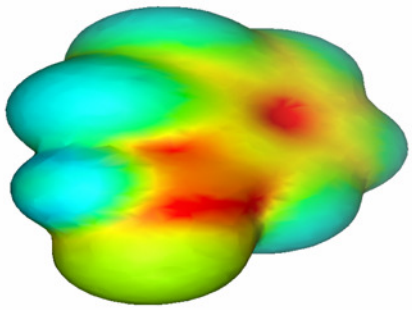
LogP, the distribution coefficient between 1-octanol and water, is a well established and widely used measure for lipophilicity. Due to its immense impact on drug design a large number of *in silico* predictions have been developed up to now. Here we present a new approach for logP modeling, as well as a novel application for conformation sensitive logP predictions.

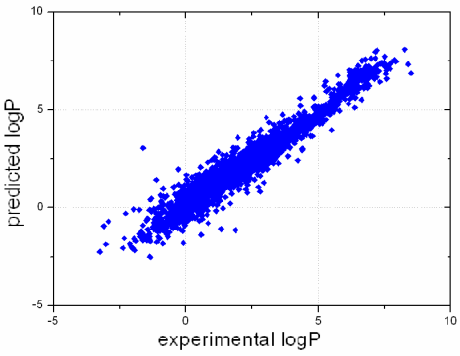
Binned surface-integral models and the descriptors used to construct them are clearly well able to predict logP as accurately as is allowed by the experimental data [1]. Based on the inconsistency of the used dataset (N=11590), we estimate a maximum possible accuracy given by the root mean squared error (RMSE) of 0.48 logP units.

In agreement with this limitation we were able to construct quantitative structure property relationship (QSPR) models with an RMSE of approximately 0.50 log units, using bagged multiple linear regression. These models proved to be extremely robust with respect to rigorous validation (bootstrapping and external validation).

Additionally, logP also embraces the ability of a compound to form hydrophobic interactions. Following this assumption, we used our logP predictions to estimate protein-ligand binding affinities for the Astex validation set [2]. Please note, that in this context the conformation is already given by the crystal-structure. Thus we used single point calculations (instead of full optimization) and calculated the atomic contributions to logP. We summed up all contributions for contact atoms (distance to protein < 4.5 A) to give an estimate for the ability to form hydrophobic interactions.

We analyzed the logP contributions with respect to binding affinities (given by pKi). Without any further fitting this already suggests to be a good estimate for binding affinity. Summarizing, we were able to show that conformation-sensitive logP prediction can be used to estimate binding affinity. In contrast to other scoring functions it is neither hypothesis driven, nor fitted to experimental binding affinity data.





[1] C. Kramer, B. Beck, and T. Clark, *JCIM*, **2010**, *50(3)*, 429-436.
[2] M. Hartshorn, et al., *JMedChem*, **2007**, *50(4)*, 726-741