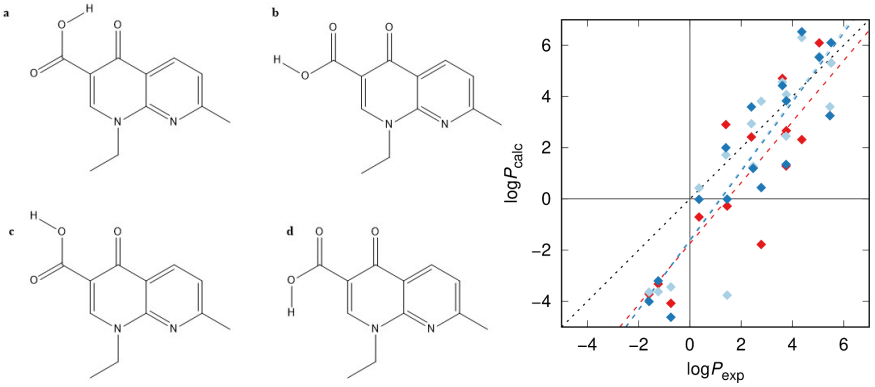
Influence of conformational ensemble models on the prediction of toluene-water partition coefficients

NicolasTielker[1], Michael Strobl[1], Christian Chodun[1], Stefan Güssregen[2], Christoph Riplinger[3], Anneke Dittmer[3], Bernardo de Souza[3], Stefan M. Kast[1]

[1]Fakultät für Chemie und Chemische Biologie, Technische Universität Dortmund, 44227 Dortmund, Germany, [2]Sanofi-Aventis Deutschland GmbH, R&D Integrated Drug Discovery, 65926 Frankfurt am Main, Germany, [3]FAccTs GmbH, 50677 Köln, Germany

In medicinal chemistry the octanol-water partition coefficient (log *P*) is frequently used as a physicochemical property to model the lipophilicity and oral bioavailability of druglike compounds. However, octanol is not an accurate representation of most biomembranes due to its ability to be both an acceptor and a donor in hydrogen bonds. The effect of this on the formation of intramolecular hydrogen bonds can be significant and will not be properly captured if the membrane is modeled as an octanol phase. [1] Furthermore, due to solvent interactions, the balance between intra- and intermolecular hydrogen bonds can drastically influence predicted partition coefficients. Hence, robust solvent-specific conformational search methods are needed.

To make progress in this direction, the SAMPL9 (Statistical Assessment of the Modeling of Proteins and Ligands) blind prediction challenge asked for toluene-water log *P* values of 16 common drugs, using an organic solvent that does not form hydrogen bonds with the solutes. This allows the investigation of radically different solvation properties in the aqueous and the organic phase. [2] To predict the toluene-water partition coefficients we employed the Embedded-Cluster Reference Interaction Site Model (EC-RISM) to represent both solvents in the calculations. [3,4] Reusing the well-established water model developed in earlier challenges and parametrizing a new, united atom toluene model made it possible to utilize the challenge as an external test set.



Finding the most energetically favorable conformation in different solvents is still an area of ongoing research. Investigation of three different conformational workflows revealed three salient points inviting further investigation. First, no single workflow was able to find the energetically most favorable conformation for all compounds. Second, even with some workflows yielding conformational energies >2.5 kcal mol^{-1} than the minimum found by other methods, often these errors were present in both solvents leading to error compensation. Finally, automated tautomer workflows can at times yield constitutional isomers that may be more energetically favorable but would not interconvert *in situ*. Surprisingly, these constitutional isomers still yield good partition coefficients, but other properties such as spectroscopic parameters would be critically affected, so care should be taken that such rearrangements are caught when large databases are investigated.

[1] W. J. Zamora *et al.*, *Phys. Chem. Chem. Phys.* **2023**, *25*, 17952.
[2] https://github.com/samplchallenges/SAMPL9 (last accessed 15/02/2024)
[3] T. Kloss, J. Heil, S. M. Kast, *J. Phys. Chem. B* **2008**, *112*, 4337-4343.
[4] N. Tielker, L. Eberlein, G. Hessler, K. F. Schmidt, S. Güssregen, S. M. Kast, *J. Comput.-Aided Mol. Des.* **2021**, 35, 453-472.