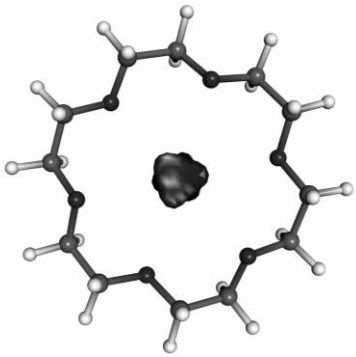
An integral equation theory for ligand design

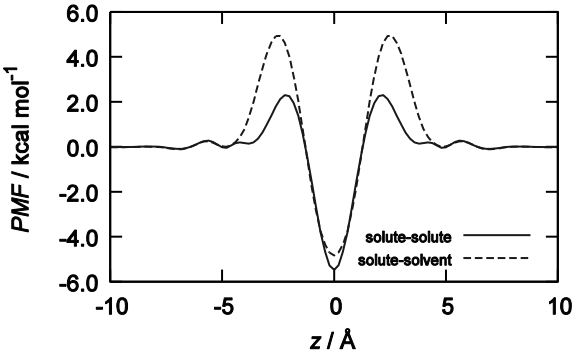
Florian Mrugalla, Christian Schulz, Stefan M. Kast

Physikalische Chemie III, Technische Universität Dortmund

Refining a ligand from an initial hit to a lead molecule is an arduous procedure. Therefore many different methods, experimental as well as computational, were developed to cope with this problem. The computational methods can be divided into classes which differ in their theoretical level, applicable scope, computational demand, and accuracy. Accuracy and computational speed typically diverge for common theoretical modeling approaches. For most rapid virtual high-throughput screening tasks accuracy is gained for structural binding pose predictions only while affinities cannot be determined sufficiently. The most expensive methods, explicit molecular dynamics free energy calculations can yield high affinity accuracy though at the prize of tremendous effort.

As an alternative, integral equation theories that are based on classical density functional theory have the potential to combine computational efficiency with high accuracy [1]. Here we particularly focus on the 3D reference interaction site model (3D RISM) that provides in its basic form solvent site distribution functions around arbitrarily shaped solutes. These equations can be transformed to yield the set of solute-solute pair molecular distribution functions [2] in infinite dilution which represents the proper thermodynamic framework for protein-ligand affinity studies. In particular, the solute-solute equations provide direct, rapid access to the potential of mean force (PMF) between ligands (or fragments and single sites thereof) and a protein host, from which the free energy of binding can be calculated along with maps of potential ligand atom pathways into the binding site.





We describe the mathematical details of the methodology and present several illustrative benchmark applications to host-guest systems. A special benefit of the approach is related to the possibility to compute free energy derivatives with respect to interaction parameters of putative ligand sites [3]. This feature is highly relevant for a classical *computer aided design* process since it allows to rationally modulate chemical properties with the goal to optimize binding affinity, giving rise to an alternative way to score binding poses. This physically appealing characteristic can be efficiently utilized since the methodology is inherently parallel and allows for rapid interactive refinement.

[1] K. F. Schmidt, S. M. Kast, *J. Phys. Chem. B*, **2002**, *106*, 6289-6297.
[2] T. Kloss, S. M. Kast, *J. Chem. Phys.*, **2008**, *128*, 134505.
[3] Y. Pang, P. A. Kollman, *Perspect. Drug Discovery Des.*, **1995**, *3*, 106-122.