Converging a knowledge-based scoring function: DrugScore^{2017}

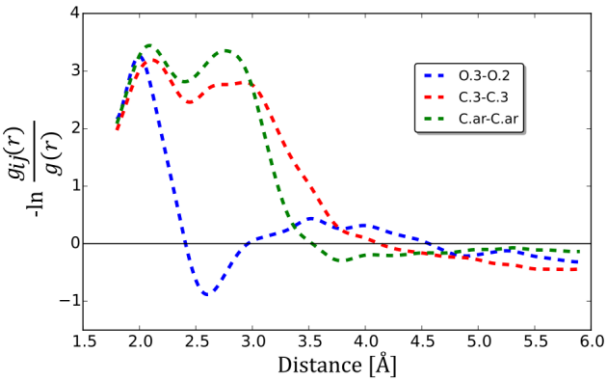
Jonas Dittrich[1], Denis Schmidt[1], Christopher Pfleger[1], Holger Gohlke[1]

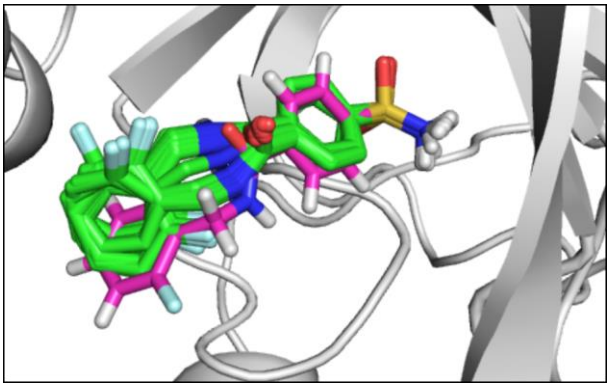
[1]*Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany*

Scoring functions play a vital role in molecular modeling. They are used in docking, virtual screening, *de novo* design, lead optimization, and other applications to evaluate protein-ligand interactions.

In 2000, DrugScore, a knowledge-based scoring function for protein-ligand complexes, was presented by Gohlke *et al.* [1]. DrugScore was successfully applied to score protein-ligand interactions [2] and as an objective function for docking [3]. At that time, about 1300 protein-ligand complexes from the PDB were used to derive distance-dependent pair-preferences for 18 atom types. Since then, structural information in the PDB increased by a factor of 10. This fact provided the incentive for us to re-derive the distance-dependent pair-preferences asking how much structural information it takes to reach a converged knowledge-based scoring function and whether the additional information further improves the predictive power of DrugScore.

DrugScore^{2017} was derived from about 40,000 complexes following the protocol from the previous work [1]. The newly derived pair-preferences cover a broader range of atom types and now include also less common interactions, e.g., involving halogens or metals. For evaluating the predictive performance, we chose two data sets: the CASF-2013 test set [3], providing complexes and associated affinities and the PDBbind "Refined set" [4], comprising more than 4000 crystal structures.





DrugScore^{2017} shows a substantial and significant increase in scoring, ranking, and docking power compared to DrugScore. I) In re-docking experiments, ~ 80 % of the generated poses showed RMSD < 2 Å compared to the crystal structure. II) The Pearson correlation of predicted and experimentally determined binding energies is 0.601 (p < 0.001) on the CASF test set and 0.526 (p < 0.001) on the PDBbind "Refined set", without any further modifications or training. This makes DrugScore^{2017} the best non-trained scoring function tested on the CASF set.

Finally, re-deriving DrugScore^{2017} on several bootstrapped subsets of the currently available data yielded pair-preferences that were indistinguishable from those derived on the full complex dataset, suggesting that the available amount of data may be sufficient to obtain converged pair-preferences.

[1] H. Gohlke, M. Hendlich, G. Klebe, *J Mol. Biol.*, **2000**, *295*, 337-356.
[2] P. Ferrara *et al.*, *J. Med. Chem.*, **2004**, *47*, 3032-3047.
[3] C. A. Sotriffer, H. Gohlke, G. Klebe, *J. Med. Chem.*, **2002**, *45*, 1967-1970.
[4] Y. Li *et al.*, *J. Chem. Inf. Model.*, **2014**, *54*, 1717−1736.
[5] R. Wang *et al.*, *J. Med. Chem.*, **2004**; *47*, 2977-2980.