

HYDE: An integrated description of dehydration and H-bonding within protein ligand interfaces

Gudrun Lange¹, Ingo Reulecke², Matthias Rarey², Robert Klein¹

¹*Bayer CropScience, Frankfurt, Germany*

²*University of Hamburg, Hamburg, Germany*

Scoring functions describe the interaction between molecules such as the binding of ligands to their target protein. They are used to identify the correct pose of a ligand with known inhibition in structure-based drug design. More importantly, they are used in a more automatic approach to score poses of thousands of putative ligands positioned into a target by docking programs and subsequently select those ligands which bind to the target (Virtual Screening). However, this goal was not always achieved and alternative scoring functions are needed (1). A comparison with experimental observations suggests that, for instance, the calculated contributions of interfacial H-bonds seem to be often overestimated while the hydrophobic effect is underestimated in many cases (2). In addition, comparing the size of the experimental ΔG with the size of the contribution attributed to the formation of an interfacial H-bond or the burial of an apolar surface, it becomes obvious that in addition to stabilizing contributions, there must exist a fair amount of counterbalancing destabilizing contributions to ΔG which are in most scoring functions not taken into account.

We believe that the underlying reason for the insufficient understanding of the interaction between molecules in aqueous solution lies in the imperfect description of water and its interaction with functional groups. Thus, we derived new dehydration terms for polar and apolar functions solely based on structural features of the water network and experimental logP values. These dehydration terms contribute stabilizing for apolar atoms (hydrophobic effect) or destabilizing in case of polar atoms and compare very well with experimental values. Our scoring function HYDE combines these dehydration terms with a term for H-bond energies and thus represents a very simple empirical approach describing the physics of protein ligand interactions (3). The balance between the hydrophobic effect and the contribution of H-bonds agrees well with experimental observations. In addition, significant destabilizing contribution to ΔG of individual atoms become apparent which lead to $\Delta G > 0$ if either the pose of a binder is incorrect or the ligand does not bind. The size of these destabilizing contribution explains why a single atom exchange within the binding site can lead to a significant altered affinity. This will be illustrated based on examples taken from the DUD data set (4). The examples show that HYDE is able to distinguish (a) between correct and wrong poses of known binders, (b) between protomers and tautomers of a binder and (c) between binders (Figure 1) and non-binders (Figure 2). This gives rise to drastically improved enrichments (Figure 3). In addition, the target-independent cut-off score allows a much more confident selection of compounds from huge libraries which is particularly important if not many binders to this particular target are known.

1. Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. Prediction of protein-ligand interactions. Docking and scoring successes and gaps. *J. Med. Chem.* **2006**, 49, 5851-5851.
2. Davies, A., D.; Teague, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. International Edition* **1999**, 38, 736.
3. Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an integrated description of hydrogen bonding and dehydration: Reducing false positives in virtual screening using the HYDE scoring function. *ChemMedChem* accepted
4. Huang, N., Shoichet, B.K., Irwin, J.J. Benchmarking sets for molecular docking *J. Med. Chem.* **2006**, 49, 6789-6801.

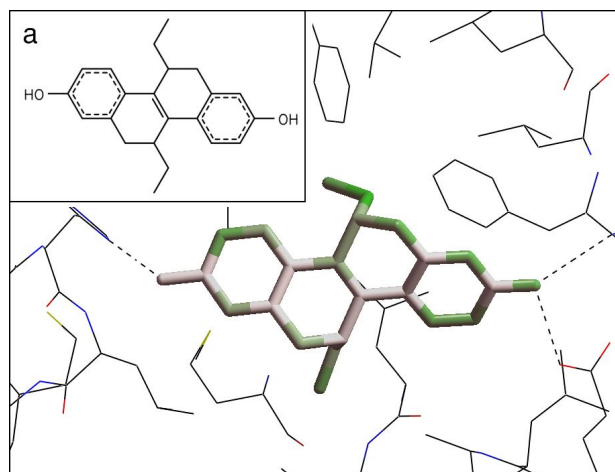


Figure 1: Contribution of individual atoms of the inhibitor in the crystal structure of the estrogen receptor (112i). Green coloured atoms contribute favourable to ΔG .

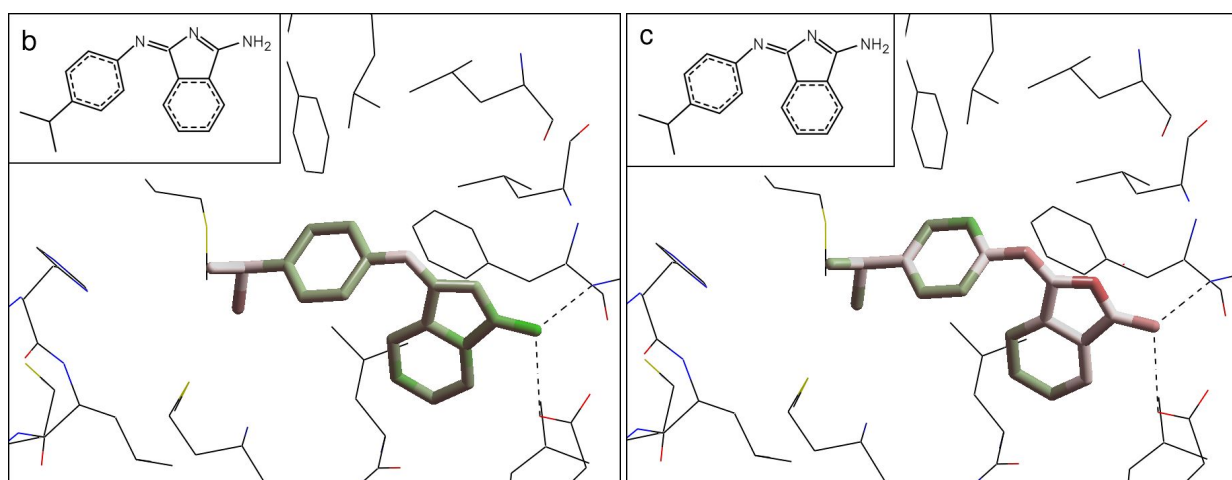


Figure 2: Contribution of individual atoms of decoy ZINC03977652 positioned in the estrogen receptor (112i) according to (a) the FlexX scoring function and (b) the HYDE scoring function. Green coloured atoms contribute favourable and read atoms unfavourable to ΔG .

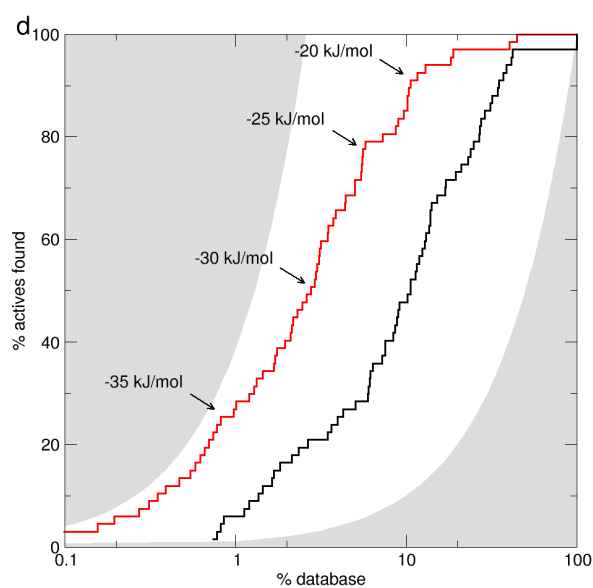


Figure 3: Enrichment plot using the estrogen agonist data set (4) and a random compound library as non-binders.