

# Quantifying the Role of Water in Ligand-Protein Binding Processes

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Joël Wahl

aus Ziefen, Schweiz

2021

Originaldokument gespeichert auf dem Dokumentenserver der Universität  
Basel [edoc.unibas.ch](https://edoc.unibas.ch)

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät  
auf Antrag von

Prof. Dr. Daniel Ricklin, Universität Basel

Dr. Christian Kramer, F. Hoffman – La Roche AG

Basel, den 19.02.19

Prof. Dr. Martin Spiess

Dekan





*Dedicated to my wife Oya  
and my daughter Smyrna*

# Acknowledgements

I want to express my sincere gratitude to Prof. Angelo Vedani, the first supervisor of my PhD thesis, who unexpectedly died in 2016. He taught me the principles of computer-aided drug design and most importantly, he emphasized the importance of always being critical. Not only to the work of others, but also to one's own results. He was greatly missed in the second half of my PhD.

I am very thankful to PD Dr. Martin Smieško, who was willing to take over as my supervisor. I thank for his guidance and fruitful discussions. While he gave me the freedom to develop and test my own ideas, he was always very eager to help and support when it was necessary.

I am grateful to Dr. Christian Kramer, who agreed to serve as the co-referee of my dissertation, despite his busy schedule.

I thank Prof. Daniel Ricklin for being the faculty representative of my PhD application.

I thank Prof. em. Beat Ernst for challenging and interesting scientific discussions. Thanks to his immense knowledge in medicinal chemistry, I could learn a lot for my future path.

A big thank goes to my colleagues and former colleagues of the Group of Molecular Pharmacy and Molecular Modeling. Specially to Christoph Sager, with whom I shared a lot of funny moments, but also interesting scientific discussions.

I am grateful to the SNF for funding and I also thank Claudia Huber for her support with many different administrative aspects concerning my PhD.

I am very thankful to my parents, that always accompanied my way and supported my decisions.

My final and by far biggest gratitude goes to my wonderful wife Oya. Meeting her was actually the most important outcome of my PhD. She is my big support, my inspiration and my motivation. First starting as colleagues, now as husband and wife, I am endlessly thankful to be at her side.

## Table of Contents

<b>1 Introduction.....</b>	<b>1</b>
<b>1.1 The Role of Water in Ligand-Protein Binding Processes .....</b>	<b>1</b>
1.1.2 The Importance of Water in Molecular Recognition .....	2
1.1.3 Water Thermodynamics in Drug Discovery .....	4
1.1.4 Influence of Water on Ligand-Receptor Binding Kinetics .....	9
<b>1.2 Structure-Based Drug Design.....</b>	<b>10</b>
1.2.1 General Overview .....	10
1.2.2 Molecular Force Fields .....	10
1.2.3 Molecular Docking .....	13
1.2.4 Molecular Dynamics Simulations .....	15
1.2.5 Free Energy Perturbation.....	17
1.2.6 Water Models in Molecular Simulations.....	19
1.2.7 Computational Methods for the Thermodynamic Characterization of Water Effects in Drug Discovery .....	21
1.2.8 The Importance of Water in Molecular Docking.....	25
<b>1.3 In silico toxicity prediction .....</b>	<b>26</b>
<b>1.4 Aim of the Thesis .....</b>	<b>27</b>
<b>2 Manuscripts.....</b>	<b>28</b>
<b>2.1 “Thermodynamic Insights into the Effects of Water Displacement and Rearrangement upon Ligand Modifications using Molecular Dynamics Simulations” .....</b>	<b>28</b>
2.1.1 Summary .....	28
2.1.2 Author contributions.....	28
2.1.3 Potential Impact on the Scientific Field .....	28
<b>2.2 “Assessing the Predictive Power of Relative Binding Free Energy Calculations for Test Cases Involving Displacement of Binding Site Water Molecules” .....</b>	<b>40</b>
2.2.1 Summary .....	40
2.2.2 Author contributions.....	40
2.2.3 Potential Impact on the Scientific Field .....	40
<b>2.3 “Incorporating Protein and Ligand Desolvation Effects into a Knowledge- Based Scoring Function” .....</b>	<b>53</b>
2.3.1 Summary .....	53
2.3.2 Author contributions.....	53
2.3.3 Potential Impact on the Scientific Field .....	53
<b>2.4 “Endocrine Disruption at the Androgen Receptor: Employing Molecular Dynamics and Docking for Improved Virtual Screening and Toxicity Prediction” .....</b>	<b>77</b>
2.4.1 Summary .....	77
2.4.2 Author contributions.....	77
2.4.3 Potential Impact on the Scientific Field .....	77
<b>3. Implemented Scripts and Software Tools .....</b>	<b>95</b>
<b>3.1 Technical Details .....</b>	<b>95</b>
<b>3.2 MDWatAnalyzer .....</b>	<b>95</b>
<b>3.3 MDWatAnalyzerFEP .....</b>	<b>96</b>

3.4 DESSCORer.....	97
4. <i>Conclusion and Outlook</i> .....	98
5. <i>References</i> .....	100
6. <i>Publication List</i> .....	113
7. <i>Curriculum Vitae</i> .....	114



# 1 Introduction

## 1.1 The Role of Water in Ligand-Protein Binding Processes

### 1.1.1 The Importance of Water for Protein Stability, Function and Dynamics

"Water is life's matter and matrix, mother and medium. There is no life without water." This famous quote of Albert-Szent Györgyi emphasizes the crucial role of water for the emergence of life (Ball, 2017). Yet, even though the importance of water for biological processes has been known for decades, its role has been underestimated, merely attributing it a passive function, that for example facilitates the diffusion of biomolecules (Chaplin, 2006).

It is only recently that the manifold and active roles that water plays in the machinery of life are gradually unraveled (Levy and Onuchic, 2004). In terms of protein folding, Kauzmann already in 1959 postulated the existence of a hydrophobic driving force, derived from favorable transfer free energies of hydrocarbons from water to organic solvents (Kauzmann, 1959), a mechanism now well-accepted in the field (Dill, 1990).

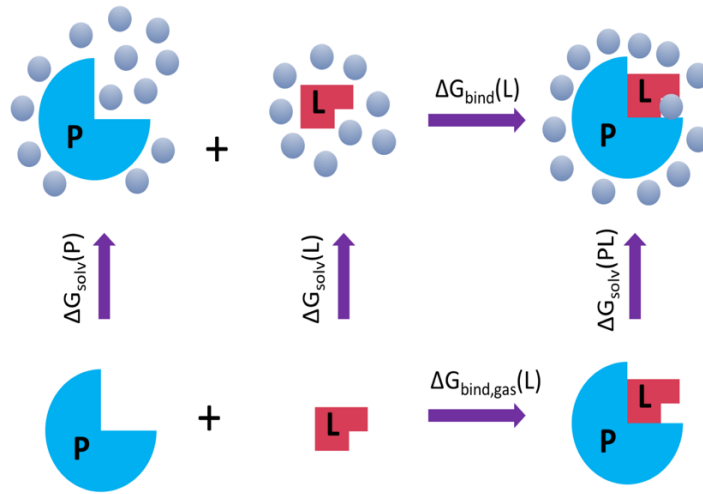
However, water also plays a more direct role in protein stability and function. Buried water molecules with long residence times ( $10^{-8}$  –  $10^{-2}$  s) are integral parts of the protein architecture (Otting et al., 1991), forming tight hydrogen bonds and thereby stabilizing the protein structure. Furthermore, long-range water-mediated interactions between hydrophilic groups guide the folding of proteins and biomolecular recognition processes (Papoian et al., 2004). These examples point out that water is not only involved passively as a solvent, but as an active agent in these processes.

At the surface of proteins, water molecules span networks, glued together by hydrogen bonds (Brovchenko et al., 2005; Nakasako, 2004). The water molecules in these networks have distinct properties compared to bulk water (Chen et al., 2008) with residence times between  $10^{-12}$  and  $10^{-11}$  s. These residence times are longer compared to bulk, but much shorter compared to deeply buried, structural water molecules. These hydration networks facilitate the dynamics of the protein, governing for example domain motions (Nakasako, 2004).

Water also takes place in biochemical reactions as both a weak acid or a weak base (Spyrakakis et al., 2017) and is essential for the functional role of hydrolases, a class of enzymes that cleave chemical bonds using a water molecule as an electrophile (Busto et al., 2010).

### 1.1.2 The Importance of Water in Molecular Recognition

Considering the formation of a complex between a macromolecule and a small molecule (typically a receptor and a ligand, or an enzyme and a substrate) or between two macromolecules, the effect of water on the process is expected to be crucial, since it is the solvent in which the complex formation occurs.



**Figure 1.** Thermodynamic cycle of protein-ligand binding

A simplified depiction of the protein-ligand binding process is given in **Figure 1** (Miller et al., 2012). Whereas the picture is valid for the formation of a molecular complex between two species in general (receptor-ligand, receptor-receptor, receptor-DNA,...), we will refer to receptor-ligand binding processes in the course of this section.

$$\Delta G_{bind}(L) = \Delta G_{solv}(PL) - \Delta G_{solv}(L) - \Delta G_{solv}(P) + \Delta G_{bind,gas}(L) \quad (1)$$

**Equation 1** is an expression for the binding free energy of ligand **L** towards the protein **P** (Miller et al., 2012). The terms  $-\Delta G_{solv}(L)$  and  $-\Delta G_{solv}(P)$  describe the desolvation free energy for the ligand and the receptor respectively. The desolvation free energy is the negative of the free energy of solvation. The free energy of solvation corresponds to the free energy change from transferring a molecule from gas to bulk.  $\Delta G_{solv}(PL)$  describes the solvation free energy of the complex. Hence, high desolvation costs for the ligand and the binding site of the receptor counteract binding. High desolvation penalties usually stem from the stripping of water molecules from charged or polar functional groups (Claveria-Gimeno et al., 2017), whereas non-



polar functional groups are associated with negligible or even positive solvation free energies (Cabani et al., 1981).

The term  $[\Delta G_{\text{solv}}(\mathbf{PL}) - \Delta G_{\text{solv}}(\mathbf{L}) - \Delta G_{\text{solv}}(\mathbf{P})]$  is the overall difference between the solvation free energy of the complex and the summed solvation free energy of the ligand and the receptor. Assuming that no large-scaled structural changes occur upon ligand-binding that would lead to the desolvation or solvation of protein atoms far from the binding site, this difference in solvation free energy mainly occurs because of the displacement of water molecules in the receptor binding site by functional groups of the ligand. The degree to which a receptor binding site is solvated (Matthews and Liu, 2009) and what the energetic cost of displacing these solvent molecules into bulk (“desolvation”) amounts to, strongly depends on the shape and electrostatic properties of the binding site (Baron et al., 2010 ; Haider et al., 2016; Michel et al., 2014; Snyder et al., 2014). In general, displacement of binding-site water molecules is considered one of the determining factors influencing the equilibrium in ligand-protein binding (Friesner et al., 2006; Hummer, 2010; Wang et al., 2011a).

Baron *et al* estimated free energy changes from explicit solvent Molecular Dynamics (MD) simulations for ligand-cavity associations of various model systems and came to the conclusion that water is “an active player in determining binding or rejection” (Baron et al., 2010). Water molecules that are expelled out of a hydrophobic cavity result in a favorable energy change, due to increased water-water interactions in bulk, which is termed as the enthalpy-driven hydrophobic effect (Homans, 2007). Interestingly, the simulations predict an entropic cost for the release of the cavity–water molecules, since they possess highly-correlated interactions in the bulk with other water molecules. On the other hand, desolvation of charged groups leads to an enthalpic penalty due to the loss of the strong charge-water interactions, but to a gain in entropy. This is because of the restriction of rotational and translational degrees of freedom of the water molecules in proximity of charged groups , attributed to the strong charge-dipole interaction (Baron et al., 2010).

Though investigations of such model systems can lead to interesting findings about the influence of water on the thermodynamics of binding, they only partly reflect the much more complex situation of protein-ligand binding (Baron et al., 2010). An interesting motif in protein active sites is hydrophobic enclosure (Friesner et al., 2006). Theoretical studies suggest that such lipophilic and enclosed binding sites lead to a perturbed solvent structure, since the water molecules located there can neither interact strongly with the protein nor form stable clusters with other water molecules (Young et al., 2007). Expelling solvent from such cavities is expected to lead to a vast increase in binding free energy (Young et al., 2007). Mentionable examples are streptavidin and cyclooxygenase-2 (COX-2).

It has been proposed that when a cavity or binding site provides a structural motif that leads to solvent configurations that are energetically too unfavorable, dehydration occurs (Wang

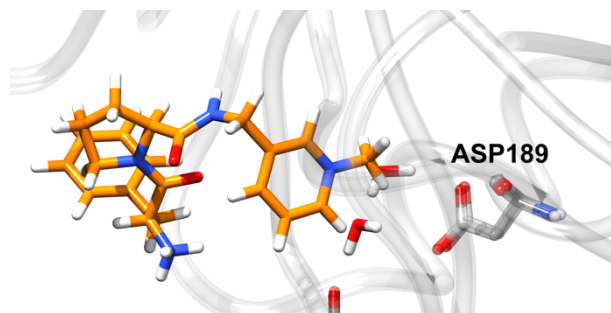
et al., 2011a). Yet, the existence of truly empty cavities inside proteins is a highly controversial matter (Matthews and Liu, 2009). “Nature abhors a vacuum” is a famous statement from Aristotle and contradicting experimental and theoretical results about the water content of various protein cavities were reported (Matthews and Liu, 2009). The excess chemical potential of water in bulk amounts to -6.4 kcal/mol (obtained from MD simulations (Wade et al., 1991)), which is approximately the cost in free energy that has to be paid from transferring a water molecule from bulk into a strictly hydrophobic cavity where there are no other water molecules present. There is indeed experimental data from x-ray crystallography that support the existence of empty cavities that possess the adequate size to host one single water molecule (Liu et al., 2008).

For larger non-polar cavities on the other hand, thermodynamically stable water cluster can be formed (Collins et al., 2005; Yin et al., 2007), whereby the geometries of the clusters are governed by the shape and size of the cavity (Vaitheeswaran et al., 2004). Young *et al* studied different proteins with largely hydrophobic cavities using MD simulations and observed total or partial dewetting of several cavities (Young et al., 2010). Another example is the binding site of mouse urinary protein 1 (MUP-1), which possesses a suboptimally hydrated binding site (Barratt et al., 2005). The association process is driven by the dispersion interactions between protein and ligand, which are not offset by any binding-site desolvation costs. Overall, this leads to a largely favorable enthalpic contribution to the binding free energy (Malham et al., 2005).

### 1.1.3 Water Thermodynamics in Drug Discovery

Due to the active role of water in molecular recognition processes, knowledge about water thermodynamics in ligand-binding is expected to lead to the design of compounds with stronger affinity, and therefore is of high importance in drug discovery. Polar functional group of the ligand are associated with a desolvation cost upon binding (Cabani et al., 1981) and should therefore not be buried in a lipophilic environment in the binding site (Barratt et al., 2006). When this high desolvation cost for strongly polar or charged groups is not compensated by strong interactions with the protein, a loss in binding affinity occurs, as shown by experimental studies on binding thermodynamics in MUP-1 (Barratt et al., 2006; Biela et al., 2012). Klebe and coworkers modified thermolysin inhibitors by changing the location of an ammonium group (Cramer et al., 2017). Their findings were that charged groups that are mainly solvent-exposed, but still close to the protein surface, suffer from partial desolvation and therefore have to be placed with great care in drug design projects.

The general concepts for the desolvation of ligand atoms are also valid for polar protein atoms that are solvated in the unliganded state. If the breakage of the hydrogen bond between protein and water is not compensated by strong interactions between the protein and the ligand, this results in a desolvation penalty (Klebe and Böhm, 1997).

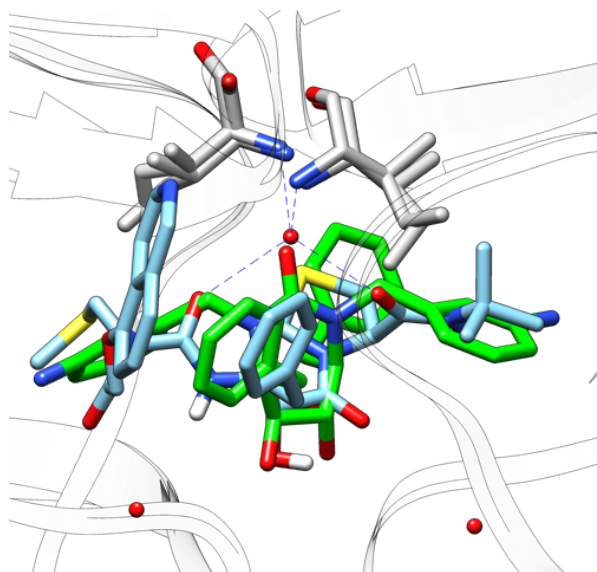


**Figure 2.** Superposition of the apo structure of thrombin (PDB ID: 2UUF) and a thrombin-inhibitor complex (PDB ID: 3QTO). The charged inhibitor displaces two water molecules that were solvating the carboxylate group of ASP 189.

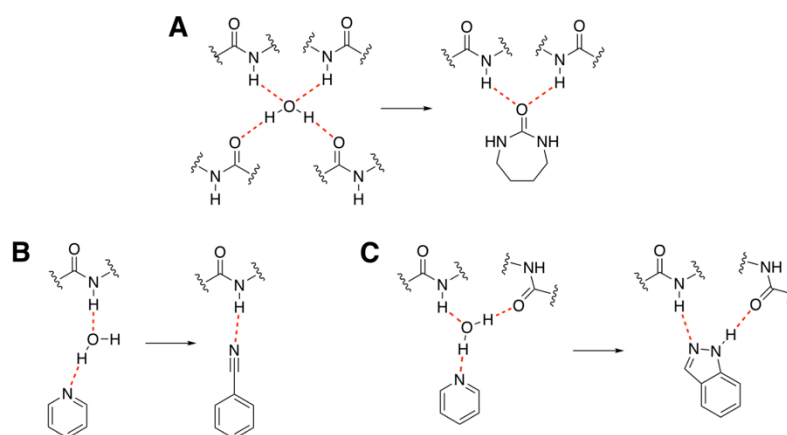
A well-studied example is the ASP189 residue in the binding site of thrombin, where the desolvation cost for the displacement of water molecules around the charged carboxylate has a decisive effect on ligand binding (**Figure 2**) (Biela et al., 2012).

A different aspect of drug design governed by water thermodynamics is the displacement of single, buried water molecules with well-defined positions. An analysis of 392 high-resolution crystal structures of protein–ligand complexes revealed that every ligand in average interacts with 4.6 structural water molecules (Lu et al., 2007). Such interfacial water molecules can increase the specificity and/or affinity of a ligand to its target (Ladbury, 1996). Modifying the chemical structure of a ligand in a way that optimizes interactions with such structural water molecules can be a design strategy to improve affinity. An outcome that is difficult to predict results from attempts to displace such water molecules by ligand functional groups. A showcase example is the displacement of a structural water in the binding site of HIV-1 protease by a new class of inhibitors bearing a cyclic urea (Lam et al., 1996). The carbonyl oxygen of the urea moiety displaces the water molecule and mimics its acceptor functionality by interacting with the backbone NH-groups of Gly49 and Ile 50 (**Figure 3** and **Figure 4A**). This class of inhibitors displays increased potency and bioavailability compared to peptidic inhibitors (Lam et al., 1994).

Another thermodynamically favorable displacement of a crystallographic water molecule that resulted in an increased affinity was reported for p38 $\alpha$  MAP kinase. Replacing a triazine-core by a 5-cyanopyrimidine in the ligand results in the displacement of a water molecule and a more than 10-fold gain in affinity (Liu et al., 2005a). A similar structural motif was exploited for Scytalone dehydratase inhibitors, where also the displacement of a water molecule by a cyano-group resulted in an affinity gain (Chen et al., 1998).



**Figure 3.** Structural water molecule with tetrahedral coordination in the binding site of HIV-1 protease (PDB ID: 1HPX). The cyclic urea inhibitor (green) displaces the water molecule (PDB ID: 1DMP).



**Figure 4.** Different ligand modifications that lead to the displacement of a bridging water molecule while restoring the hydrogen bonds to the receptor.

Yet such an approach is not a guaranteed success. For quinazoline inhibitors of epidermal growth factor receptor kinase (EGFR), substituting one of the ring nitrogen atoms by a nitrile group results in the displacement of a water molecule that bridges the interactions between the ring nitrogen and the protein. Whereas this describes in general the same situation as for the two aforementioned examples, the displacement in this case results in a slight loss in affinity (Wissner et al., 2000).

In general, for the successful displacement for such bridging water molecules in terms of an increase in affinity, a ligand functional group with optimal geometry has to be introduced, that not only displaces the water molecules, but also adequately replaces its interactions with the protein. For a water molecule adjacent to a ring nitrogen and interacting with the protein by accepting a hydrogen bond, a nitrile group is an attractive choice (Davies et al., 2012), as exemplified by various cases. A depiction of the structural features of water displacement by a nitrile group is given in **Figure 4B**. For bridging water molecules adjacent to ring acceptors that interact with the protein as a hydrogen bond donor, Smith and coworkers developed an interesting strategy by extending the pyridyl scaffold to an indazole (Smith et al., 2015). They successfully applied it to a series of Bruton's Tyrosine Kinase (BTK) inhibitors and observed an increase in affinity. The indazole moiety can mimic both the donor and the acceptor functionality of the water molecule (**Figure 4C**) and therefore undergoes tight interactions with the receptors.

All these examples have in common that the ligand modification introduces an additional polar group that displaces the water molecule and compensates for the lost water-protein hydrogen bonds. Yet, there is also the possibility to introduce nonpolar functional groups to the ligand that sterically displace binding site water molecules. Katz and coworkers achieved selectivity for inhibitors of Ser190 trypsin-like serine proteases by displacing a water molecule from the S1 site by ligand halogen atoms (Katz et al., 2001). Depending on the interaction pattern of this S1 water molecule in the corresponding binding sites, displacing this water molecule led to no change in affinity for some kinases and to strong loss in binding affinity for other kinases.

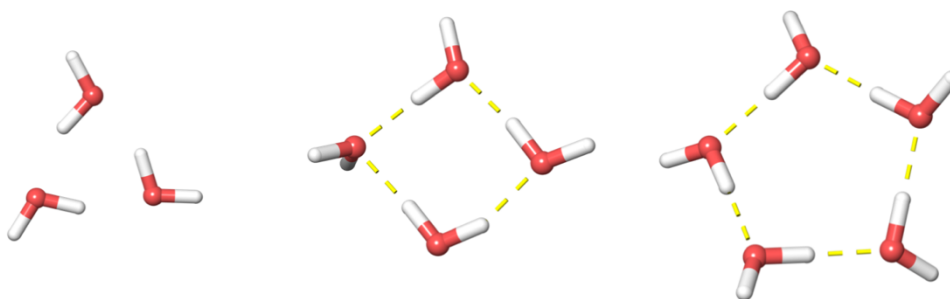
The group of Klebe investigated a different water molecule in the S1 pocket of thrombin and observed a significant affinity gain for its displacement by a nonpolar ligand atom (Baum et al., 2009).

Displacement of two water molecules by substituting a hydrogen atom for a methyl group in a series of Hsp90 inhibitors left the affinity unchanged (Kung et al., 2011).

These examples point out that it is not *a priori* clear if and by which functional group a water molecule can or should be displaced in order to gain affinity or specificity and it is clear that it is the environment of the water molecule (number of interactions, shape of the cavity,...) that influences the outcome of the displacement. To decide in a prospective fashion if a water molecule can and should be targeted for displacement by a ligand functional group, the exact thermodynamic contributions towards the change in binding affinity should be understood and quantified. Several of the above examples have been studied by molecular simulations to assess the exact energetics of the water displacement (Li and Lazaridis, 2003; Michel et al., 2009). The results indicate that expelling water molecules into bulk is energetically unfavorable, due to the loss of strong solute-water interactions. Yet their calculations were conducted for the water

molecules that are tightly bound and displaced by polar functional groups, such as for HIV-1 protease (Li and Lazaridis, 2003) and EGFR, Scytalone dehydratase and p38 $\alpha$  MAP kinase (Michel et al., 2009). In general, it is the fine balance between free energy change for the release of the water molecule into the bulk, the newly formed protein-ligand interactions and the change in ligand solvation free energy upon the structural modification that determines if expelling a water molecule leads to an affinity gain (Michel et al., 2009).

We so far discussed about the non-specific effect of water in protein-ligand binding (protein and ligand desolvation) and about the displacement of single well-defined water molecules by ligand modifications. Yet, an interesting property of water molecules is their ability to interact favorably with other water molecules, leading to specific arrangements in water networks, so called “clusters” (**Figure 5**). Such water clusters display a sheer endless diversity, regarding the number of water molecules involved, their shape and interaction patterns (Maheshwary et al., 2001). Theoretical studies have shown that in nonpolar protein cavities, the lack of favorable solute-solvent interactions can be compensated by the formation of water clusters with strong solvent-solvent interactions, which are the driving forces to make the hydration free energies of such cavities favorable (Vaitheeswaran et al., 2004; Yin et al., 2007). Depending on the hydrogen bond network, such water clusters can profit from cooperative stabilization or anticooperative destabilization (Albrecht et al., 2013; Pérez et al., 2014).



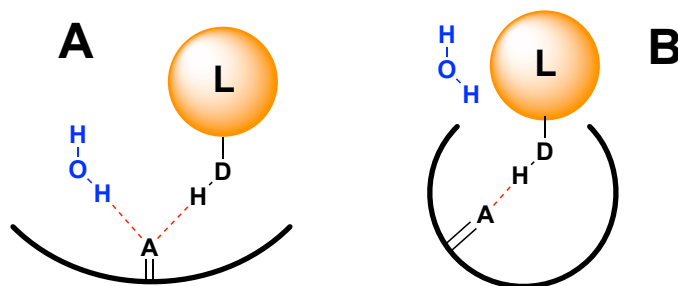
**Figure 5.** Examples of water clusters with three, four and five members.

Yet, there is still not much understanding of the exact role and influence of such water clusters on protein-ligand binding. Studies on inhibitor binding at thermolysin suggest that the quality and completeness of the water network formed around the ligand-receptor complex have a direct impact on the enthalpy and entropy of binding (Betz et al., 2016; Krimmer et al., 2014). Similar conclusions were drawn from a study of carbonic anhydrase inhibitors (Breiten et al., 2013). Whereas it is now understood that these water networks play a role in protein-ligand binding, quantifying their exact contributions is not yet possible.

### 1.1.4 Influence of Water on Ligand-Receptor Binding Kinetics

Recent studies point out that optimizing and investigating the residence times of drugs instead of only targeting the binding affinity in drug design projects can lead to more efficient and safer drugs (Copeland et al., 2006; Lu and Tonge, 2010; Schuetz et al., 2017). Therefore, understanding the exact factors that influence the binding kinetics of drugs can have a big impact in drug design projects. As discussed for the thermodynamics of receptor-ligand binding, binding site water molecules also seem to have a direct influence on the receptor–ligand binding kinetics (Schmidtke et al., 2011; Setny et al., 2013).

Results from molecular simulations of different systems indicate that a complete dewetting of the binding pocket occurs before the entrance of the ligand and that this stage possesses the highest energy barrier and is therefore rate-determining (Dror et al., 2011; Mondal et al., 2014; Setny et al., 2013). It is however to be expected that the shape of the binding site and the extent of solvent-exposure influence the exact contribution of binding site water molecules to the binding kinetics. For the unbinding of dasatinib from the c-Src kinase, MD simulations revealed a concerted and more complex mechanism of binding, but again the hydration of the binding site is a rate-determining step, which was also confirmed for simulations at p38 MAP kinase (Casasnovas et al., 2017). The hydration of the binding site is the first slow step in the unbinding process (Liu et al., 2015). Yet it has also been found that desolvation of highly polar ligand groups can influence the association rate of protein–ligand complexes, as measured for a set of Hsp90 inhibitors (Schuetz et al., 2018).



**Figure 6.** Schematic depiction of a hydrogen bond accessible to solvent (A) and a shielded hydrogen bond (B). The water molecule stabilizes the transition state of the hydrogen bond breakage in A, leading to a lower energetic barrier.

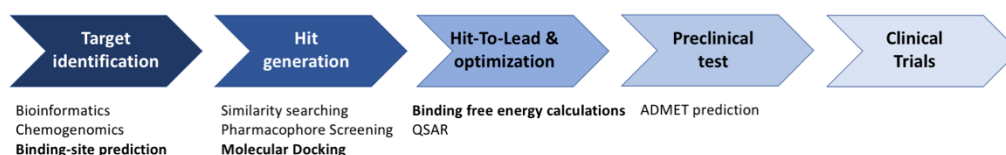
It has been proposed that hydrogen bonds that are shielded from solvent lead to longer residence times of ligands, since the breakage of such hydrogen bonds is associated with a higher energetic barrier (**Figure 6**) (Schmidtke et al., 2011). This argumentation was also used to explain varying residence times of different natural substrates bound to their native lectin binding site, where lectins that have narrow and buried pockets possess ligands with longer residence times (Sager et al., 2017).

## 1.2 Structure-Based Drug Design

### 1.2.1 General Overview

It is widely accepted that computer-aided drug design (CADD) has a big impact on the development and discovery of novel drug candidates and that it is an established methodology extensively used by the pharmaceutical industry (Śledź and Caflisch, 2018). Computational tools have manifold applications in the drug discovery pipeline (**Figure 7**), from target identification, library design, lead identification, lead optimization, ADMET prediction to elucidation of the mode-of-action of a drug (Tang et al., 2006). Structure-based drug design (SBDD) is a branch of CADD for the discovery and optimization of drug candidates guided by the 3D structure of the binding site. The applications of SBDD are the prediction and evaluation of how small molecules bind to their target (binding pose prediction) and the estimation of the strength of the binding (affinity prediction) (Śledź and Caflisch, 2018). A basic requirement for the application of SBDD is the availability of a valid 3D structure of the receptor, either solved by experiment (x-ray crystallography, NMR, EM) or by homology modeling (Levoine et al., 2011).

Increased accuracy of a computational method is usually associated with a higher computational cost, and therefore different methodologies of CADD are applied in different stages of the drug discovery pipeline. Virtual high-throughput screening (vHTS) is performed with computationally less expensive methods such as molecular docking. The goal of vHTS is to enrich chemical libraries with compounds that are more likely to bind and therefore reduce the experimental expenses and in the end lead to more cost-efficient drug-discovery processes (Sliwoski et al., 2014).



**Figure 7.** Simplified depiction of a typical drug-discovery pipeline. SBDD methods are mentioned in bold.

More accurate methods such as relative binding free energy (RBFEE) calculations based on free energy perturbation (FEP) or thermodynamic integration (TI) are applied in lead-optimization phases (Jorgensen, 2004; Sliwoski et al., 2014).

### 1.2.2 Molecular Force Fields

In order to accurately explore the conformational space of small-molecule drugs bound to their targets and to get reliable protein-ligand interaction energies for the estimation of



binding free energies, a method that maps molecular coordinates to a potential energy value is required. Whereas quantum chemical calculations are the method of choice for assessing geometries and energies of chemical structures in terms of accuracy, their high computational cost prevents them from being routinely used in CADD applications (Leach, 2001). Molecular Force Fields provide an alternative that facilitates rapid energy valuation of molecular systems by approximating the energies as a sum of intramolecular and intermolecular contributions. Since the various terms are described by equations taken from classical physics (harmonic oscillator, hard sphere potential, electrostatic potential) this concept is termed as “molecular mechanics”.

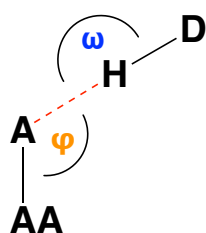
$$\begin{aligned}
 V(r) = & \sum_{bonds} (r - r_0)^2 + \sum_{angles} (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\
 & + \sum_{i>j} (4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}) \quad (2)
 \end{aligned}$$

**Equation 2** describes the typical form of a classical force field. The bond and angle energies are calculated by means of a classical harmonic oscillator, where deviations from the equilibrium bond lengths are associated with a high energetic penalty, increasing quadratically. The torsion term is composed of a summation of cosine functions with varying periodicity, whereas the intramolecular interactions are given by two pair potential functions: the first describes the van der Waals interactions by a so called 12-6 Lennard-Jones potential, whereas the second term accounts for electrostatic interactions by means of a classical point charge Coulomb model.

In order to assess the validity and accuracy of computational algorithms employing force fields, it is crucial to be aware of their strengths and shortcomings. Approximating bond-stretching and angle-stretching energies as harmonic oscillators is only valid for geometries close to energetic minimum. This approximation prohibits bond-breaking and is therefore not suitable to model chemical reactions. This is in general not a big drawback in the field of CADD, where one usually works with geometries close to equilibrium. The bond-stretching and angle-stretching terms are referring to the “hard” degrees of freedom, since they only allow for subtle changes (Leach, 2001). Rotation about a bond is a soft degree of freedom and the associated energy is given by the torsion term. Whereas the functional form of this term can reproduce the potentials obtained from *ab initio* calculations with good accuracy, it needs a lot of parameters to fully cover the complexity of the chemical space and recent advances in force field development aim into this direction (Harder et al., 2016).

Calculating interaction energies between ligands and receptors is of great importance in CADD and therefore an accurate representation of the non-bonded energy terms is an absolute necessity to obtain correct sampling of the energy landscape of a ligand-receptor

complex (Ponder et al., 2010). The simple point-charge model with pair-wise interactions is a simplification of the true charge distribution which is highly anisotropic (Kramer et al., 2012). Hydrogen bonds for example have certain geometric preferences, depending of the type of the acceptor atom (Vedani and Dunitz, 1985). The YETI force field expands the classical force equation by a specific hydrogen bond term, that accounts for the directionality and linearity of the hydrogen bond (**equation 3**).

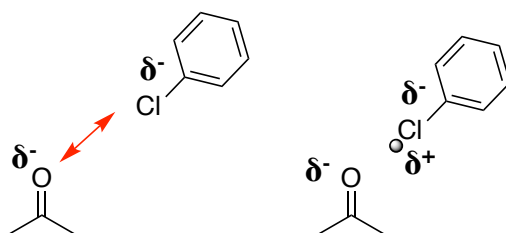


$$E_{\text{hbond}} = \left[ \frac{C}{r_{ij}^{12}} - \frac{D}{r_{ij}^{10}} \right] \cos^2(\varphi_{H\cdots A-AA} - \varphi_0) \cos^2(\omega_{D-H\cdots A} - \omega_0) \quad (3)$$

$$C = -5E_{\text{min}}r_0^{12} \quad (4)$$

$$D = -6E_{\text{min}}r_0^{12} \quad (5)$$

C and D are constants that depend on the atom types of the donor and acceptor atom, the associated well-depth ( $E_{\text{min}}$ ) of the potential and the optimal hydrogen-bond distance ( $r_0$ ) (Vedani, 1988). The two cosine terms have the effect of lowering the energy in terms of deviations from the optimal directionality ( $\varphi_0$ ) and linearity ( $\omega_0$ ). The optimal values for the linearity and directionality were derived from geometric analysis of small-molecule crystal structures (Vedani and Dunitz, 1985).



**Figure 8.** Electrostatic repulsion between an aromatic chlorine and a carbonyl oxygen in the classical point charge model (left) and correct representation of the Cl-O halogen bond by addition of an off-site positive charge.

A further drawback of the simple point-charge model is the insufficient coverage for special interactions such as  $S\cdots O$  chalcogen bonds (Lupyan et al., 2012; Yan et al., 2017) or halogen bonds (Jorgensen and Schyman, 2012; Rendine et al., 2011).

Recent improvements to overcome these limitations are the addition of off-center charges (**Figure 8**) (Jorgensen and Schyman, 2012; Yan et al., 2017) or expanding the point

charges by means of atomic multipoles (Kramer et al., 2012, 2013). A further limitation of the fixed-charge models is the lack of a polarization term (Warshel et al., 2007), and therefore electronic many-body effects such as cooperative hydrogen bond networks (DeChancie and Houk, 2007) are not captured by classical force fields.

The most widely used molecular force fields include the AMBER molecular force fields (Case et al., 2005), CHARMM (Brooks et al., 2009), OPLS (Jorgensen et al., 1996), GROMOS (Oostenbrink et al., 2004) and the MMFF94s (Halgren, 1996). Due to their fast evaluation of molecular energies and the existence of analytical derivatives, force fields find wide application in structure minimization (Kini and Evans, 1991), biomolecular simulations (Mackerell, 2004), conformational searches of small to medium-sized molecules (Grebner et al., 2013; Gürsoy and Smieško, 2017) and molecular docking (Liu and Wang, 2015).

There are ongoing efforts to improve and extend the current force fields for better coverage of small-molecular chemical space (Vanommeslaeghe et al., 2010; Wang et al., 2004), improved parameters for protein conformations (Maier et al., 2015) or the aforementioned developments for more realistic description of charge distributions. Force field parameters for equilibrium bond lengths and angles are usually taken from experimental crystal structures (Wang et al., 2004) or quantum chemical calculations (Halgren, 1996) and the force constants can be fitted to reproduce experimental vibrational spectra (MacKerell et al., 1998). The parameters for the non-bonded terms are optimized to reproduce hydration free energies (Oostenbrink et al., 2004) or interaction-potentials obtained from *ab initio* potentials (Vanommeslaeghe et al., 2010). Torsional parameters are typically fitted to torsional energy profiles from quantum chemical calculations (Halgren and Nachbar, 1996; Vanommeslaeghe et al., 2010; Wang et al., 2004).

### 1.2.3 Molecular Docking

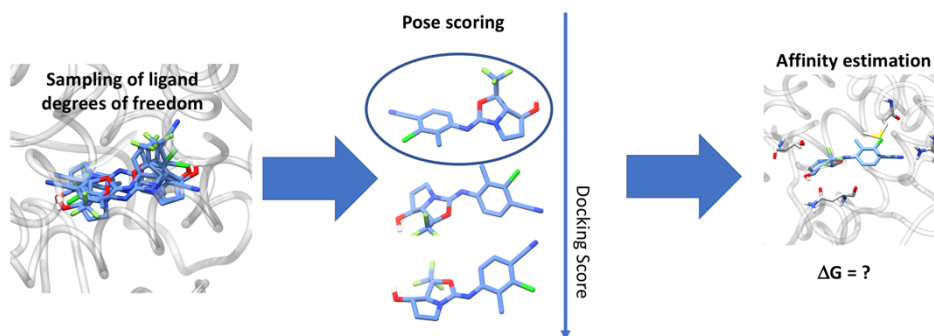
Molecular Docking is a computational tool for studying the interactions of a small molecule with the binding site of a receptor, with the ultimate goal to predict the mode of binding as well as the binding free energy, or affinity of the complex (Meng et al., 2011). While there is an application of docking to the field of protein-protein docking, only algorithms dealing with small-molecule docking are discussed here, though some of the principles and strategies are also valid for other docking problems.

Molecular Docking is one of the key methods in SBDD and has applications in hit identification as well as hit-to-lead and lead optimization phases (Enyedy and Egan, 2008; Kitchen et al., 2004; Meng et al., 2011), since it allows for a rapid assessment of the steric and electrostatic complementarity of a candidate molecule to a receptor binding site. Due to its relatively low computational cost, even screening of large virtual libraries is feasible

(Kontoyianni, 2017). Although reliable prediction of binding affinities with chemical accuracies is still out of reach for molecular docking algorithms (Liu et al., 2017; Wang et al., 2003, 2016), applying it in vHTS applications can lead to libraries that are enriched with true active and this can accelerate the drug discovery pipeline (Sliwoski et al., 2014).

While a plethora of different docking programs and methodologies exist, the common strategy is to sample the ligand degrees of freedom (translational, rotational, torsional) in the binding site of the receptors. The different configurations of these ligand degrees of freedom lead to so called docking poses, which are ranked by means of a scoring function (Meng et al., 2011).

There are different strategies for the exploration of the ligand degrees of freedom in the binding site, ranging from genetic algorithms (Morris et al., 1998), swarm intelligence (Korb et al., 2007) or Monte Carlo (MC) (Hart and Read, 1992). The protein can thereby be kept rigid (flexible ligand–rigid protein docking) or flexible (flexible ligand–flexible protein docking). A third strategy is to dock compounds into an ensemble of different binding–site conformations of the same receptor (ensemble docking) (Amaro et al., 2018). These ensembles can either be taken from experimental crystal structures (Osguthorpe et al., 2012a) or molecular simulations (MD or MC) (Osguthorpe et al., 2012b). A typical workflow of a docking program is depicted in **Figure 9**.



**Figure 9.** Workflow of a typical molecular docking algorithm

Accurate scoring functions are necessary for correct binding pose prediction, estimation of affinities, ranking of compounds by affinity and discriminating binders from nonbinders (Li et al., 2014). Most scoring functions are composed of additive terms that aim to capture the important energetic contributions of protein–ligand binding. Favorable interactions such as hydrogen bonds, salt bridges, hydrophobic interactions and attractive Coulomb interactions lead to better docking scores, while desolvation of polar groups, steric clashes or repulsive coulomb interactions are detrimental for binding. Furthermore, changes in conformational entropy upon binding are estimated by approximations such as functions taking into account the number of rotatable bonds of the ligand (Ben-Shalom et al., 2017). The scoring

functions can be classified into four categories: i) physics-based, ii) empirical, iii) knowledge-based and iv) machine-learning (Liu and Wang, 2015). For physics-based scoring functions, the individual energy terms are taken from the molecular force field and complemented by a desolvation term from an implicit-solvent model (Huang et al., 2006).

Empirical scoring functions have a less rigorous functional form and are calibrated to reproduce experimental affinities (Böhm, 1994). They can include certain force-field terms, but often are composed of additional, physically not always rigorous contributions such as lipophilic contact scores or special hydrogen bond rewards.

Knowledge-based scoring functions estimate the binding affinity by means of pair-wise statistical potentials, derived from analysis of protein-ligand complexes from the PDB (Muegge, 2006). The distance-frequencies of atom-atom contacts between protein and ligand are binned into histograms and the energy function is then derived.

The performance of scoring functions can be evaluated in terms of the docking power, the scoring power, the ranking power and the screening power (Li et al., 2014). The docking power is the ability to select the correct pose (taken from the experimental structure of the protein-ligand complex) from a set of decoy poses. The scoring power describes how well the scoring function can predict experimental binding affinities, if the correct binding pose is given as an input.

The ranking power refers to the ability of correctly ranking a set of compounds that bind to the same receptor. The screening power finally is an indication of how well the scoring function can discriminate true binders from inactive molecules (decoys).

A comparative assessment on several widely-used docking algorithms showed that pose prediction and screening are less of a problem than scoring and ranking (Li et al., 2014). Furthermore, strong binders are often predicted too weak and weak binders too strong and a weakness of docking functions for the binding of rigid molecules to deeply buried pockets was discovered. The authors of the study concluded that non-additive effects and desolvation of polar groups need to be improved in current scoring functions.

#### 1.2.4 Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations are based on Newton's law of motion. They simulate the movement of particles over time (Leach, 2001). In classical MD simulations that are typically applied in medicinal chemistry and structural biology, the moving particles are the atoms that the chemical system is composed of. The result of an MD simulation is a trajectory, that specifies the positions and velocities of the atoms at different points in time.

Newton's second law is given in **equation 6**:

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (6)$$

By knowing the force  $F$  acting on a particle in the direction  $x_i$ , solving the differential equation yields the time-dependent motion of the particle. In classical MD simulations, the forces on the particles can be obtained by means of a molecular force field. The molecular force field describes a potential energy function, mapping the coordinates of a molecular system to a potential energy value. In classical physics, the force is defined as the negative of the gradient of the potential  $V$  (equation 7).

$$F_{x_i} = -\frac{dV(x_i)}{dx_i} \quad (7)$$

In molecular force fields, the force on a particle depends on its position relative to all other particles in the system. As a consequence, the atomic motions are not independent of each other and the equations cannot be solved analytically (Leach, 2001). Therefore, finite difference methods with a given time step  $\Delta t$  are applied, such as the Verlet algorithm (Verlet, 1967). The new positions of the particles are thereby derived from the current step and the previous step. The forces on the particles are calculated at each step.

Since chemical and biological systems are by nature dynamic, MD simulations can provide insights into the role and function of proteins, transports across membranes and molecular recognition (Karplus and Kuriyan, 2005; Nair and Miners, 2014). Furthermore, the theory of statistical mechanics provides a rigorous physical framework to extract thermodynamic quantities from these simulations. Thermodynamic properties are calculated as time averages over the trajectory and the ergodic hypothesis states that ensemble average equal time averages, assuming complete sampling of the phase-space (space of all possible states).

In terms of SBDD, MD simulations can aid binding-site prediction (Feng and Barakat, 2018), refinement of protein structures from homology modeling or low-resolution experiments (Xun et al., 2015), binding pose prediction and refinement (Liu and Kokubo, 2017), binding free energy calculations (Cournia et al., 2017) and assessment of binding kinetics (De Vivo et al., 2016). Such methodologies are now widely established in the field of CADD (De Vivo et al., 2016), supported by the development of novel simulation techniques (Hamelberg et al., 2004; Laio and Parrinello, 2002; Liu et al., 2005b), the increase in computational power (Borhani and Shaw, 2012), increase in experimental structural data available (Hollingsworth and Dror, 2018) and force field development (Harder et al., 2016).

Careful selection of the simulation protocol and underlying parameters is needed for an adequate and reliable application of MD simulations in drug design projects (van Gunsteren

and Mark, 1998). Apart from selecting a simulation package (e.g. AMBER (Case et al., 2005), CHARMM (Brooks et al., 2009), Desmond (Bowers et al., 2006) being popular choices), the force field used for evaluating the forces between the particles has a direct impact on the outcome of the simulation, as shown by different studies (Feig et al., 2003; Martín-García et al., 2015; Watts et al., 2018). Furthermore, a thermodynamic ensemble has to be chosen, specifying if the simulation is run under constant number of particles, volume and temperature (NVT), constant number of particles, pressure and temperature (NPT) or constant number of particles, constant volume and constant energy (NVE). Since the conditions for biological systems are usually room temperature and atmospheric pressure, simulating in the NPT ensemble at  $T=298$  K and  $p = 1$  is a widely used setting when simulating biomolecules.

### 1.2.5 Free Energy Perturbation

As previously mentioned, statistical mechanics enables the calculation of thermodynamic quantities as time averages from molecular simulations. The Helmholtz free energy of a system is directly obtained by the partition function  $Q$  through **equation 8**:

$$A = k_B T \ln \frac{1}{Q} \quad (8)$$

The partition function  $Q$  is as follows:

$$Q = \iint e^{-E(\mathbf{q}, \mathbf{p})/k_B T} d\mathbf{q} d\mathbf{p} \quad (9)$$

Where  $\mathbf{q}$  and  $\mathbf{p}$  are the positions and momenta of all particles in the system and the integral is over the whole phase space. Inserting **equation 9** into **equation 8** then leads to the following formulation of the Helmholtz free energy:

$$A = k_B T \ln \left[ \frac{\iint e^{E(\mathbf{q}, \mathbf{p})/k_B T} e^{-E(\mathbf{q}, \mathbf{p})/k_B T} d\mathbf{q} d\mathbf{p}}{\iint e^{-E(\mathbf{q}, \mathbf{p})/k_B T} d\mathbf{q} d\mathbf{p}} \right] \quad (10)$$

This expression is valid since the integral on top equals one. Introducing the probability density  $p(\mathbf{q}, \mathbf{p})$  leads to:

$$A = k_B T \ln \left[ \iint e^{E(\mathbf{q}, \mathbf{p})/k_B T} p(\mathbf{q}, \mathbf{p}) d\mathbf{q} d\mathbf{p} \right] \quad (11)$$

Since the total system energy  $E(\mathbf{q}, \mathbf{p})$  occurs in the exponential, system configurations with high energy make a significant contribution to the Helmholtz free energy. However, since simulation algorithms such as MD primarily sample low-energy regions of the phase space, the

sampling problem that prevents free energy values to converge (Cramer, 2005). If we compute the difference in Helmholtz free energy between a system in state A and state B and reformulate the integrals as ensemble average, we can write

$$\begin{aligned}\langle A \rangle_B - \langle A \rangle_A &= k_B T \ln \left( \frac{1}{M_B} \sum_i^{M_B} e^{E_i/k_B T} \right) - k_B T \ln \left( \frac{1}{M_A} \sum_i^{M_A} e^{E_i/k_B T} \right) \\ &= k_B T \ln \langle e^{E_i/k_B T} \rangle_B - k_B T \ln \langle e^{E_i/k_B T} \rangle_A \\ &= k_B T \ln \left( \frac{\langle e^{E_i/k_B T} \rangle_B}{\langle e^{E_i/k_B T} \rangle_A} \right) \quad (12)\end{aligned}$$

If the phase space of the two ensembles would perfectly overlap, this would simply equal to

$$\langle A \rangle_B - \langle A \rangle_A = k_B T \ln \langle e^{(E_B - E_A)/k_B T} \rangle_A \quad (13)$$

**Equation 13** is called the Zwanzig relation (Zwanzig, 1954). The advantage lies in the fact that the equation doesn't contain energies, but energy differences. For two biomolecular systems that are very similar and for example just differ by one atom, most energy terms (solvent-solvent e.g.) remain roughly the same and cancel out. Yet, the accuracy of the Zwanzig relation depends on how well the phase space of systems A and B overlap. To obtain a better overlap, the transformation from system A to B can be split into smaller steps by using a parameter  $\lambda$  with values ranging from 0 to 1 (**equation 14**).

$$E(\lambda) = \lambda E_B + (1 - \lambda) E_A \quad (14)$$

This methodology is called free energy perturbation (FEP). Whereas FEP can be used to predict absolute binding free energy (Singh and Warshel, 2010), it shows its strongest potential in CADD for the prediction of relative binding free energies, meaning the difference in binding free energy  $\Delta\Delta G$  between two ligands (Cournia et al., 2017). This is also due to the fact that relative binding free energies are easier to converge and usually profit from a cancellation of errors (Chodera et al., 2011), since these calculations are often performed for congeneric series of structural analogues in lead optimization stages.

Factors that diminish the accuracy of binding free energy calculations (FEP, thermodynamic integration,...) are shortcomings of molecular force fields (Wang et al., 2015) and insufficient sampling (Mobley et al., 2007). If the change in ligand structures leads to large



conformation rearrangements in the protein, substantially longer sampling times are needed (Lim et al., 2016). Recent developments in enhanced sampling methods such as replica exchange with solute tempering (REST) (Wang et al., 2011b) allow improved accuracy. It is the combination of better force fields, enhanced sampling, increased computational power through GPU and automatized workflows that led to an increased application of relative binding free energy calculations in drug discovery projects (Cournia et al., 2017).

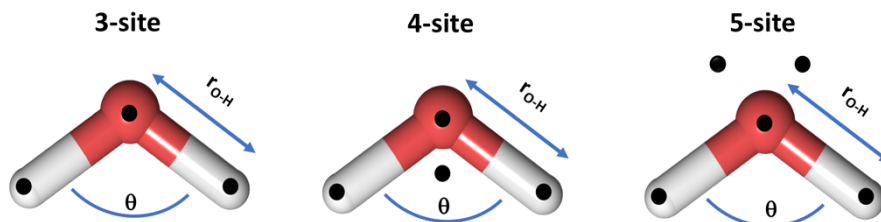
### 1.2.6 Water Models in Molecular Simulations

In general, two approaches exist for the representation of water in molecular simulations: implicit solvent and explicit solvent.

In the implicit representation of the solvent, the water molecules are not explicitly treated, but merely the effect that they have on the behavior of the system is emulated. Liquid water for example has a relative permittivity (dielectric constant) of  $\epsilon = 80.1$  at room temperature (Hobbs et al., 1966), leading to a large shielding of charges according to Coulomb's law. This effect can for example be taken into account by replacing the constant dielectric constant  $\epsilon$  by a distance-dependent dielectric  $\epsilon = \epsilon_{ij}$  or by a sigmoidal function, that has shown to more accurately reproduce the structure of DNA-binding protein in molecular simulations (Guenot and Kollman, 1992). While in the past (early 90s), such simple models were used in molecular simulations due to the computational expense of adding explicit water molecules (Guenot and Kollman, 1992), they nowadays are mainly applied in scoring functions for docking (Friesner et al., 2006), where rapid evaluation of the interaction energies is necessary and the sampling of the solvent needs to be eliminated for computational efficiency.

More realistic implicit representations of solvent that are computationally less expensive than treating the solvent explicitly are the Generalized-Born and the Poisson-Boltzmann methods (Leach, 2001), that allow for estimations of solvation free energies and can be used in combination with molecular mechanics to estimate binding free energies (Genheden and Ryde, 2015).

Nowadays, molecular simulations are mainly conducted with explicit water molecules and it is expected that an accurate representation of the true physical and structural properties of water is a necessity to accurately reproduce the true dynamics of a biomolecular system. In recent decades, different water models with varying complexity have emerged. Depending on the number of interaction sites, widely used water models can be categorized into 3-site, 4-site and 5-site models (**Figure 10**).



**Figure 10.** Schematic depiction of different rigid water models used in molecular simulations with classical force fields. Apart from the number of interaction sites, the water models differ in terms of geometry ( $r_{O-H}$ ,  $\theta$ ) and their charge distributions.

The most widely applied 3-site water models are the TIP3P (Jorgensen et al., 1983) and the SPC (Berendsen et al., 1981). In these models, the interaction sites coincide with the positions of the three atoms of the water molecule. The TIP4P is 4-site model that has an additional dummy atom along the bisector of  $\theta$ . This dummy atom does neither bear a mass nor does it interact with other particles over Lennard-Jones interactions, it solely bears the negative charge of the oxygen (Jorgensen et al., 1983). The interaction site located at the oxygen atom thereby still has the full mass associated with it as well as the Lennard-Jones parameters. In the TIP5P model, two charge-bearing dummy particle are present with the aim to mimic the lone pairs of the oxygen that should lead to a better representation of the tetrahedral interaction geometry of water molecules (Mahoney and Jorgensen, 2000). Water models are usually parametrized against experimental, physical properties of liquid water (density, pressure, radial distribution function, dielectric constant,...) (Berendsen et al., 1981; Jorgensen et al., 1983). For the parametrization, typically molecular simulations are performed and it is evaluated how well the calculated values derived from the simulation reproduce the experimental ones.

With increasing number of particles, the computational cost for evaluating the forces increases, and therefore in terms of speed, the 3-site models perform best and this fact is contributing to their popularity (Wang et al., 2014). The question remains, if this reduced computational cost comes at the expense of accuracy. The TIP3P water model for example shows overestimated mobility, which actually leads to faster dynamics of the system and can increase sampling (Florová et al., 2010). However, simulations of bulk water with TIP4P reproduce diffraction data better and lead to more accurate thermodynamic properties compared to SPC and TIP3P (Jorgensen et al., 1983).

In terms of reproducing positions of crystallographic water molecules, simulations that were performed using a TIP4P water model were shown to be superior compared to TIP3P and TIP5P (Betz et al., 2016). It has also been shown that the choice of the water model can have a

significant impact and the dynamics, structure and thermodynamic properties of binding site solvent in protein–ligand complexes (Fadda and Woods, 2011).

All mentioned water models are associated with major approximations: they are rigid in geometry and have a fixed charge distribution with a point–charge model. Furthermore, the parameters for the water models are usually fitted to reproduce the properties of bulk water, however, due to the different dielectric constant in the protein environment, the dipole moment of water is expected to be significantly different in the interior of proteins (Morozenko et al., 2014). The simple point–charge model and lack of polarization are on the other hand a general shortcoming of the classical force fields used for the simulations. Improved water models with atomic multipoles and polarization terms also need the corresponding force fields (Qi et al., 2015). The AMOEBA force field is a noteworthy example for a polarizable force field with atomic multipoles (Ponder et al., 2010), at a higher computational cost compared to classical force fields.

### 1.2.7 Computational Methods for the Thermodynamic Characterization of Water Effects in Drug Discovery

Due to the crucial role that water plays in molecular-recognition processes, there is a clear need for computational methods that guide drug discovery projects in this aspect. This can range from algorithms that predict hydration free energies of small molecules (Mobley et al., 2009), to scoring functions that account for the solvation and desolvation of polar atoms (Schneider et al., 2013), or to tools that provide a three-dimensional map of the solvent thermodynamics of binding sites to guide drug design (Abel et al., 2008).

Prediction of small–molecular hydration free energies can be obtained by implicit solvent models in combination with force field or *ab initio* calculations (Brieg et al., 2017; Mennucci et al., 2002). For simulations conducted with explicit solvent models, solvation free energies of small molecules are typically calculated by means of free energy calculations such as FEP (Mobley et al., 2009).

To guide ligand design in terms of optimizing the contributions from solvent effects, the distribution and thermodynamic properties of the binding site solvent should be investigated. Arguably the first published approach for determining possible locations and properties of solvent molecules in the binding site is GRID (Goodford, 1985). In GRID, a water probe systematically scans points on a grid and evaluates the associated interaction energies. This results in a 3D contour plot showing potentially favorable and unfavorable locations for water molecules. The advantage of this approach is the low computational cost, but a disadvantage is the focus on solute–probe interactions, neglecting water–water interactions. This makes the approach feasible for identifying structural water molecules with strong interactions

with the protein, but identifying clusters that are stabilized through solvent–solvent interactions is challenging. A similar method is MCSS (multiple copy simultaneous search), which differs in terms of probe representation and functional forms of the interaction energies, but provides similar information about preferred locations for water molecules in biomolecules (Bitetti-Putzer et al., 2001).

SZMAP (solvent–Zap–mapping) is also a method that relies on sampling a grid on the binding site with a water probe. However, the remaining solvent is represented as Poisson–Boltzmann implicit solvent (Bayden et al., 2015), therefore the method accounts for interactions of the water probe with the remaining solvent, although in an implicit manner. Furthermore, SZMAP facilitates the estimation of the free energy difference of the probe compared to bulk, together with the energetic and entropic contributions. These quantities are derived from a partition function, summing over the probe orientations for a given grid point. SZMAP was shown to successfully reproduce the locations of crystallographic water sites and predict if they are conserved or displaced upon the binding of various ligands in a set of HIV-1 protease and FXa inhibitors. The obtained free energies for the water site correlated well with free energies obtained by more rigorous calculations based on MD simulations with thermodynamic integration (Bayden et al., 2015).

3D-RISM is based on the integral equation theory of liquids and uses reference interaction site model (RISM) in three dimensions (Beglov and Roux, 1997) to calculate the distribution and thermodynamic properties of a liquid around a solute. 3D-RISM achieves the complete sampling of solvent in accessible computational time, which is an advantage compared to methodologies based on molecular simulations such as molecular dynamics, where sampling is finite (Sindhikara and Hirata, 2013). 3D-RISM derived free energy densities were used in a partial least squares regression model and successfully captured SAR trends in a series of fXa ligands (Güssregen et al., 2017).

Another possible approach is to extract thermodynamic properties of solvent molecules from Molecular Dynamics Simulations. Inhomogeneous Fluid Solvation Theory (IFST) is a rigorous framework that facilitates the calculation of solvation free energies by quantifying the change in solvent structure due to the insertion of a solute (Lazaridis, 1998). The thermodynamic properties of the solvent can thereby be extracted from the trajectory analysis from molecular simulations such as MD simulations. Separate terms for enthalpic and entropic contributions to the solvation free energy are calculated. Furthermore, solvation energies and entropies are decomposed into solute–solvent terms ( $E_{sw}, S_{sw}$ ) and solvent–solvent reorganization terms ( $E_{ww}, S_{ww}$ ) (Li and Lazaridis, 2003). The second order terms for the solvation energy and entropy are (Li and Lazaridis, 2006):

$$\Delta E_{solv} = E_{ww} + \frac{1}{2}E_{ww} - E_{bulk} \quad (15)$$

$$\Delta S_{solv} = S_{sw} + \Delta S_{ww} \quad (16)$$

Evaluation of the energetic term is straightforward for simulations with classical force fields and can be taken as the force field interaction energies of the solvent with the solute ( $E_{sw}$ ) and the mean solvent–solvent interaction energy ( $E_{ww}$ ).  $E_{bulk}$  is the mean solvent–solvent interaction energy in bulk.

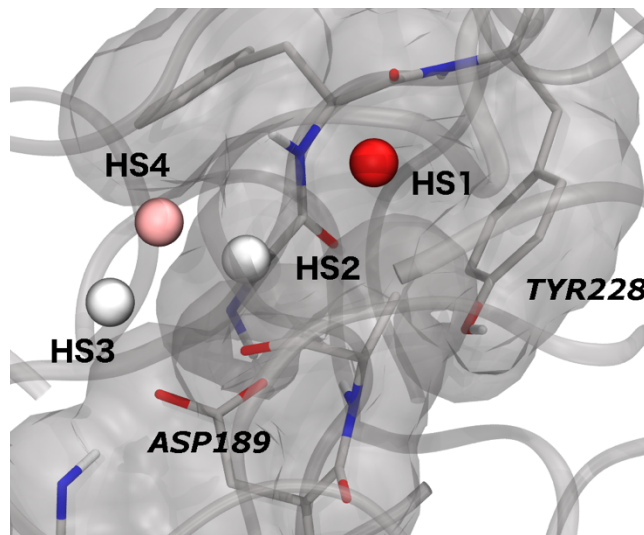
The entropic terms are taken from solute–solvent and solvent–solvent pair correlation functions. The solute–solvent term is given as

$$S_{sw} = -k\rho \int g_{sw}^{tr}(\mathbf{r}) \ln g_{sw}^{tr}(\mathbf{r}) d\mathbf{r} - \frac{N_{wat}}{8\pi^2} \int g_{sw}^{or}(\omega) \ln g_{sw}^{or}(\omega) d\omega \quad (17)$$

Where  $g_{sw}^{tr}$  and  $g_{sw}^{or}$  are the translational and rotational correlation functions respectively. These integrals can be evaluated by numerically integrating over the probability distribution functions of the orientations and the translations (Li and Lazaridis, 2006). These integrals can be evaluated for specific subregions of a biomolecular system, yielding information about the thermodynamic properties of the solvent occupying these regions. This led to the concept of so called hydration sites, which are spherical regions with high water density. They are typically obtained by clustering the water positions from a molecular simulations (Abel et al., 2008; Li and Lazaridis, 2006).

The possibility to obtain a spatial map of the thermodynamic properties of the solvent around a solute has interesting applications in CADD. The best-known tool for integrating IFST into drug discovery projects is the WaterMap software from Schrödinger (Schrödinger Release 2018-4: WaterMap, Schrödinger, LLC, New York, NY, 2018.). The principal idea is to run a simulation of the unliganded state of the protein and analyze the locations and thermodynamic properties ( $\Delta G$ ,  $\Delta E$ ,  $-T\Delta S$ ) of the hydration sites in the binding site. Hydration sites with high  $\Delta G$  values provide hot spots for ligand design, since their displacement expected to result in an affinity gain (Abel et al., 2008). Typical cases are hydration sites in hydrophobic environment (unfavorable  $\Delta E$ ) and hydration sites with few and highly correlated hydrogen bonds (unfavorable  $\Delta E$  and  $-T\Delta S$ ) (Young et al., 2007). WaterMap has been successfully applied to predict SAR trends in congeneric ligand series (Abel et al., 2008), for the prediction of binding sites (Beuming et al., 2012), assessing ligand-binding kinetics (Pearlstein et al., 2013), understand kinase selectivity (Robinson et al., 2010) and develop novel ligands for platelet-derived growth factor receptor (Horbert et al., 2015).

Other implementations of this type of hydration site analysis have been developed and published (Hu and Lill, 2014; Huggins, 2012; Li and Lazaridis, 2012). An outcome of a hydration site analysis is shown in **Figure 11**.



**Figure 11.** Hydration Site Analysis of the S1 pocket of Thrombin. The hydration sites are colored according to their  $\Delta G$  values. HS2 and HS3 possess strong interactions with the carboxylate group and have low free energies. HS1 is surrounded by hydrophobic residues, resulting in a high  $\Delta G$  value.

The shortcoming of WaterMap and related algorithms is that because of the clustering of the water positions into spherical hydration sites, thermodynamic properties of regions with lower solvent density are not assessed. Furthermore, if the density distribution is not radial, clustering into spheres also does not fully capture the whole solvent density even in regions that are highly occupied by water.

A solution to this problem was presented by Nguyen *et al* (Nguyen et al., 2012) in the form of a grid-based formulation of IFST (GIST), where the thermodynamic solvent properties are mapped onto a discrete grid. All of the mentioned algorithms are implementations of IFST, but they differ slightly in the way the thermodynamic properties are calculated. Entropic contributions can be calculated using a histogram-based approach (Abel et al., 2008; Li and Lazaridis, 2006) or a k-nearest neighbor approach (Huggins, 2014; Ramsey et al., 2016) and the latter exhibits improved convergence, especially with a combined translational-rotational distance metric (Huggins, 2014).

Furthermore, the WaterMap method does not involve a scaling of the solvent-solvent interaction energies  $E_{\text{ww}}$  by a factor of 0.5 (Abel et al., 2008), as opposed to the original formulation of IFST (Lazaridis, 1998). This leads to a different physical interpretation of the obtained thermodynamic solvent quantities. In IFST, contributions of regions towards the solvation free energy are calculated (Lazaridis, 1998; Nguyen et al., 2012), whereas in the formulation used by WaterMap, a cavitation term is obtained which describes the change in free

energy between growing the repulsive core of the ligand in the active site versus growing it in the bulk (Abel et al., 2010). The formulation of this so called displaced-solvent functional is only valid for ligands that are complementary to the protein (hydrogen bonds formed in the bulk are conserved in the receptor binding site) (Abel et al., 2010).

The obtained  $\Delta G$  values for the hydration sites in WaterMap often display positive values (Abel et al., 2008), and the interpretation is as follows: if the ligand binds complementary to the protein (hydrophobic ligand parts interact with hydrophobic protein residues and polar ligand functional groups interact with polar protein functional groups), then the displacement of this hydration site is expected to result in an affinity gain. In the original IFST formulation on the other hand, the  $\Delta G$  values are usually negative and this indicates a desolvation cost for the displacement of this hydration site (Li and Lazaridis, 2003). The interaction energy of the ligand with the protein should then be large enough to compensate for this desolvation cost. Both approaches were recently incorporated into scoring functions, aiming at a more accurate description of water thermodynamics and its quantitative contribution towards protein-ligand binding (Balius et al., 2017; Murphy et al., 2016).

### 1.2.8 The Importance of Water in Molecular Docking

We emphasized in the previous sections the important contributions of protein and receptor desolvation and the effect of single well-defined water molecules on the quantification of ligand-protein binding processes. Since molecular docking aims to identify correct ligand binding poses and estimate the associated affinity, accurate scoring functions should include terms that account for these effects. Since scoring functions should be computationally fast, explicitly sampling all solvent degrees of freedom is out of reach. A possible, physically rigorous physics-based formulation is to incorporate solvation/desolvation effects using an implicit solvent model such as Generalized-Born (GB) or Poisson-Boltzmann (Gilson et al., 1997; Zou et al., 1999). Yet, although these methods are computationally cheaper than explicitly sampling the solvent degrees of freedom, they are still too expensive for large screenings of compounds (Liu et al., 2004).

Empirical terms that are physically less restrictive but still aim to capture the underlying concept of ligand and receptor desolvation upon binding can provide a solution to this challenge. Such methods usually assess how the solvent accessibility of ligand and/or protein atoms change upon binding, and the desolvation penalty is then based on the corresponding atom type (Huey et al., 2007; Mysinger and Shoichet, 2010; Schneider et al., 2013).

Though these methods account for general solvation and desolvation effects, they cannot account for specific water molecules with well-defined positions that are known to be

important in molecular recognition (Ladbury, 1996) and it was shown that incorporating such water molecules into docking studies improves pose prediction (Roberts and Mancera, 2008; Thilagavathi and Mancera, 2010). Since such water molecules can either be conserved across a series of ligands or displaced (García-Sosa et al., 2003), a flexible treatment of explicit binding site solvent is necessary for incorporating these water molecules correctly.

Such dynamical treatment of binding site water molecules is for example incorporated in GOLD (Verdonk et al., 2005) or AutoDock (Forli and Olson, 2012). However, there still remains the challenge of accurately quantifying the energetics of the displacement of such water molecules. Recent developments aim to incorporate thermodynamic solvent properties obtained from molecular simulations in combination with inhomogeneous fluid solvation theory. WScore combines Glide (Friesner et al., 2004) with WaterMAP (Abel et al., 2008), whereas GIST (Nguyen et al., 2012) was combined with Dock (Balius et al., 2017) and Autodock (Uehara and Tanaka, 2016).

### 1.3 *In silico* toxicity prediction

Apart from their importance in drug design, computational methods can also be used for the estimation of the toxicology of chemical substances (Raies and Bajic, 2016). A wide range of technologies were applied for toxicity predictions, such as QSAR (Helma, 2006), machine learning (Yang et al., 2018) or structural filters (Sushko et al., 2012). Yet, when the toxicological effect of a compound is initiated by the binding to a biological receptor, also methods that are usually applied in SBDD can be used (Trisciuzzi et al., 2017).

The VirtualToxLab, developed by Prof. Angelo Vedani, is an *in silico* tool for toxicity prediction that estimates the binding affinity of small molecules towards a panel of 16 receptors that are known to trigger adverse effects (Vedani et al., 2012, 2015). The targets involve various nuclear receptors (androgen receptor, estrogen receptor  $\alpha$  and  $\beta$ , thyroid receptor  $\alpha$  and  $\beta$ , progesterone receptor, mineralocorticoid receptor, glucocorticoid receptor, liver X receptor, peroxisome proliferator-activated receptor  $\gamma$ ), members of the cytochrome family (cytochrome P450 1A2, 3A4, 2C9, 2A13, 2D6), the aryl hydrocarbon receptor and the human Ether-a-go-go related gene. The VirtualToxLab relies on molecular docking with flexible ligand and flexible receptor in combination with a physical-scoring function based on the YETI force field, whereas the binding affinities are determined by a Boltzmann-weighted scoring approach (Vedani et al., 2015).



---

## 1.4 Aim of the Thesis

The aim of this thesis is to quantify the contributions of water thermodynamics to the binding free energy in protein-ligand complexes. Various computational tools were directly applied, implemented, benchmarked and discussed.

An own implementation of the IFST formulation was developed to facilitate easy integration in workflows that are based on Schrödinger software. By applying the tool to a well-defined test set of congeneric ligand pairs, the potential of IFST for quantitative predictions in lead-optimization was assessed.

Furthermore, FEP calculations were applied to an extended test set to validate if these simulations can accurately account for solvent displacement in ligand modifications.

As a fast tool that has applications in virtual screening problems, we finally developed and validated a new scoring function that incorporates terms for protein and ligand desolvation.

This resulted in total in three distinct studies, that all elucidated different aspects of water thermodynamics in CADD. These three studies are presented in the next section. In the conclusion, the results and implications of these studies are discussed jointly, as well with possible future developments.

An additional study was focused on virtual screening and toxicity prediction at the androgen receptor, where distinguishing agonists and antagonists poses difficulties. We proposed and validated an approach based on MD simulations and ensemble docking to improve predictions of androgen agonists and antagonists.

## 2 Manuscripts

### 2.1 “Thermodynamic Insights into the Effects of Water Displacement and Rearrangement upon Ligand Modifications using Molecular Dynamics Simulations”

Wahl, J.; Smiesko, M. Thermodynamic Insights into Effects of Water Displacement and Rearrangement upon Ligand Modification Using Molecular Dynamics Simulations. *ChemMedChem* **2018**, *13*, 1325-1335.

#### 2.1.1 Summary

A test set of seven congeneric ligand pairs that contain structural modifications that either lead to the displacement or to a rearrangement of binding site water molecules was constructed from literature data. Inhomogeneous Fluid Solvation Theory (IFST) was applied to calculate the thermodynamic contributions of the change in the water network towards the total change in binding free energy between the ligand pairs. Whereas an approach combining IFST and differences in interaction energies, solvation free energies and binding entropies could qualitatively predict if the change in binding free energy is favorable for a ligand pair, predictions with chemical accuracy ( $<1$  kcal/mol) were not possible for all systems. This highlights that lead optimization strategies that are based on considerations about water thermodynamics are challenging and the outcomes are difficult to predict, even with state-of-the-art methods.

#### 2.1.2 Author contributions

J.W. developed a script that implements the IFST algorithm. The script extracts the positions and thermodynamic properties of hydration sites directly from an MD trajectory. J.W. compiled the test set, designed the study and performed the calculations, all under supervision and guidance of M.S.

J.W. wrote the manuscript. M.S. proofread and improved the manuscript.

#### 2.1.3 Potential Impact on the Scientific Field

The article discusses several theoretical considerations for the application of IFST calculations in lead optimizations, such as the importance of the rearrangement of the water molecules due to structural ligand modifications and the problem of double-counting of water-water interactions between different hydration sites. Furthermore, the article supports the understanding of structure-activity relationships (SAR) that involve the displacement of binding site water molecules.

# Thermodynamic Insight into the Effects of Water Displacement and Rearrangement upon Ligand Modifications using Molecular Dynamics Simulations

Joel Wahl and Martin Smieško<sup>\*[a]</sup>

Computational methods, namely molecular dynamics (MD) simulations in combination with inhomogeneous fluid solvation theory (IFST) were used to retrospectively investigate various cases of ligand structure modifications that led to the displacement of binding site water molecules. Our findings are that water displacement per se is energetically unfavorable in the discussed examples, and that it is merely the fine balance between change in protein–ligand interaction energy, ligand solvation free energies, and binding site solvation free energies that determine if water displacement is favorable or not. We

furthermore evaluated if we can reproduce experimental binding affinities by a computational approach combining changes in solvation free energies with changes in protein–ligand interaction energies and entropies. In two of the seven cases, this estimation led to large errors, implying that accurate predictions of relative binding free energies based on solvent thermodynamics is challenging. Nevertheless, MD simulations can provide insight regarding which water molecules can be targeted for displacement.

## Introduction

Displacement of structural water molecules by small modifications in ligand structures has been an established strategy in medicinal chemistry for improving binding affinities of small molecules.<sup>[1]</sup> Notable examples involve cyclic urea inhibitors binding to HIV protease,<sup>[2]</sup> and nitrile-containing EGFR kinase<sup>[3]</sup> and p38 $\alpha$  MAP kinase<sup>[4]</sup> inhibitors. Whereas in the mentioned cases water displacement either led to an improvement in affinity (HIV-protease, p38 $\alpha$  MAP kinase) or left the affinity unchanged (EGFR kinase), the specific contributions towards the change in affinity cannot be determined experimentally. It is well accepted that displacement of a tightly bound, well-ordered water molecule leads to a gain in entropy,<sup>[5]</sup> and the overall displacement process is only favorable if the ligand compensates for loss of the water–protein interactions.<sup>[6]</sup> However, because in most cases such a clear classification cannot be made,<sup>[7]</sup> there is a demand for computational tools that help in the prediction of thermodynamic effects of water displacement.

Various computational approaches that allow prediction of locations and thermodynamic properties of water molecules in protein binding sites have emerged recently, and a concise overview of these methods is given elsewhere.<sup>[8]</sup> Numerous studies that applied these algorithms focus on running a thermodynamic analysis of the solvent properties of *apo* structures.<sup>[6,9–13]</sup> Structure–activity relationships are then explained


by displacement of strongly or weakly bound solvent molecules, where displacement of weakly bound solvent is expected to increase binding affinities. This treatment of binding site desolvation can help to improve the accuracy of docking algorithms and lead to enhanced scoring functions.<sup>[14–16]</sup>

While this approach has the advantage that it typically only needs one simulation per target structure, it neglects the contribution of water rearrangement upon ligand binding. This approach seems to be the method of choice for structure-based virtual screening applications due to the lower number of simulations required, yet we expect that for lead optimization stages, water rearrangement should be accounted for. Gilson and co-workers proposed the application of grid inhomogeneous solvation theory to assess the change in solvation free energy upon chemical modification of the ligand.<sup>[17]</sup> If ligand A is the initial ligand and B is the modified ligand, and states PA and PB correspond to the ligand–receptor complexes of A and B, then the change in solvation free energy of the complex can be described as  $\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) = \Delta G_{\text{solv}}(\text{PB}) - \Delta G_{\text{solv}}(\text{PA})$ . This approach accounts for reorganization of the solvent. Previous studies used a similar approach employing grid cell theory or 3D-RISM to account for the solvent perturbation upon ligand scaffold modifications.<sup>[18,19]</sup> Changes in solvation energies can be used not only to rationalize differences in binding energies of chemically different ligands, but also for binding pose flips, as shown recently.<sup>[20]</sup>

In the present study, we used hydration site analysis (HSA) based on inhomogeneous fluid solvation theory (IFST)<sup>[21,22]</sup> to estimate binding site solvation free energies for congeneric ligand pairs at several receptors (Table 1). HSA in combination with IFST was applied before by Lazaridis to dissect the effect of water displacement on binding thermodynamics for conca-

[a] J. Wahl, Dr. M. Smieško

Molecular Modeling, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)  
E-mail: martin.smiesko@unibas.ch

 Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/cmdc.201800093>.

**Table 1.** Congeneric ligand pairs with various calculated thermodynamic contributions to the relative binding free energies.

Ligand pair	$\Delta\Delta G_{\text{solv}}$ (AP $\rightarrow$ BP)	$\Delta\Delta G_{\text{solv}}$ (A $\rightarrow$ B)	$\Delta\Delta E_{\text{inter}}$ (AP $\rightarrow$ BP)	$-T\Delta\Delta S_{\text{bind}}$ (AP $\rightarrow$ BP)	$\Delta\Delta G_{\text{bind}}$ (A $\rightarrow$ B) <sub>calc</sub>	$\Delta\Delta G_{\text{bind}}$ (A $\rightarrow$ B) <sub>exp</sub>
Cumming-4j $\rightarrow$ Cumming-4b	1.87	3.14	-7.22	1.88	-6.62	-0.61
Kung-13 $\rightarrow$ Kung-16	3.40	1.05	-3.09	0.07	-0.67	-0.17
Davies-13b $\rightarrow$ Davies-15b	5.75	0.32	-3.11	-0.70	1.62	-1.41
Davies-15a $\rightarrow$ Davies-16	-3.14	-1.18	-0.42	1.29	-1.09	3.25
Woodhead-1 $\rightarrow$ Woodhead-2	15.01	2.60	0.56	-1.48	11.49	2.01
Chen-3d $\rightarrow$ Chen-5d	4.74	-2.52	-11.75	2.78	-1.71	-1.98
Smith-8 $\rightarrow$ Smith-11	6.22	-3.82	-18.42	4.18	-4.20	-1.91

navalin A<sup>[23]</sup> and HIV-1 protease.<sup>[24]</sup> We furthermore employed alchemical free energy perturbation (FEP) calculations to estimate the change in solvation free energies of the ligands in solution. We additionally assessed the change in protein–ligand interaction energies due to the ligand structure modification and added an entropic correction term based on the recently postulated concept of interaction entropy.<sup>[25]</sup> With this analysis, we aim to improve our understanding of the thermodynamics of water displacement upon ligand modifications and to help future decisions about water displacement in lead optimization stages.

We applied the following criteria for selecting appropriate ligand pairs for this study: 1) Small changes in ligand structure that do not alter the binding mode or lead to major changes in the protein structures; 2) High-quality crystal structures for both ligands bound to protein or crystal structure of a closely related analogue; 3) Experimentally determined  $K_d$  or  $\Delta G_{\text{bind}}$  values for both ligands measured in the same lab; 4) Waters of interest should be buried and correspond to structural waters, meaning that their positions are well defined and their average occupancy is one.

## Results and Discussion

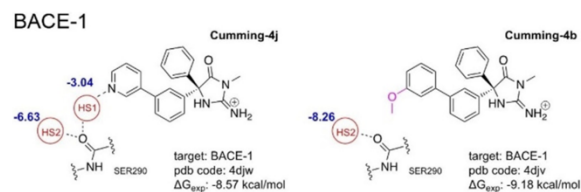
Cumming et al. described novel imidohydantoin BACE-1 inhibitors with nanomolar affinity (Figure 1).<sup>[26]</sup> In the crystal structure of a pyridinyl-containing inhibitor 4j (PDB ID: 4djw), a structural water molecule (HOH1045 in the crystal structure, denoted as W1 in the following) can be observed in the S3 pocket, that bridges an interaction between the ligand nitrogen atom and the Ser190 backbone. By substituting the 3-pyridyl by a 3-methoxy group (inhibitor 4b), this water molecule is

sterically displaced (PDB ID: 4djv). The affinity is thereby increased by nearly a factor of three (from 530 nM to 190 nM, reported by Cumming et al.).<sup>[26]</sup> We ran HSA for both complexes to estimate the desolvation cost caused by the change in scaffold.

### Hydration site analysis

For Cumming-4j, we observe two hydration sites in the proximity of the S3 sub-pocket. HS1 bridges the interaction between the ligand pyridine nitrogen atom and the backbone carbonyl group. The calculated  $\Delta G$  of this hydration site amounts to  $-3.04 \text{ kcal mol}^{-1}$  (Supporting Information Table S1), and hence the desolvation cost for the displacement of HS1 amounts to  $+3.04 \text{ kcal mol}^{-1}$ . The stability of the hydration site can be attributed to strong solute–solvent interactions ( $-16.00 \text{ kcal mol}^{-1}$ ). These are partly offset by unfavorable solvent–solvent interactions ( $+2.71 \text{ kcal mol}^{-1}$ ). The entropy of solvation for HS1 only plays a minor role ( $0.19 \text{ kcal mol}^{-1}$ ). HS2 forms three hydrogen bonds with the protein backbone (Tyr75, Val92, Ser290) and is more stable than HS1 with a  $\Delta G$  value of  $-6.63 \text{ kcal mol}^{-1}$ . Upon replacement of the pyridine moiety by a methoxyphenyl group (Cumming-4b), HS2 is stabilized by  $1.63 \text{ kcal mol}^{-1}$ , primarily because it now can form tighter interactions with the protein.

This effect counteracts the desolvation costs associated with the displacement of HS1. For the sake of simplicity, we term the two inhibitors as A and B in the following and use a similar notation as in a previous study with related content.<sup>[18]</sup>  $\Delta\Delta G_{\text{solv}}(\text{AP} \rightarrow \text{BP})$  corresponds to the difference in protein–ligand complex solvation free energy for the two ligands.  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  corresponds to the change in ligand solvation free energy. The  $\Delta G$  values for the hydration sites correspond to the contribution of this hydration site to the total solvation free energy of the system.<sup>[17,21,27]</sup> Displacing a hydration site with a positive  $\Delta G$  is energetically favorable, and the desolvation cost of the displacement of a hydration site is  $-\Delta G$ . We compare the solvation free energy of a region **R** for both complexes. **R** should ideally be chosen in a way that it captures the total perturbation to the water thermodynamics by going from inhibitor A to B. Because in this example, the change of scaffold occurs in a buried rigid site with only few observed solvent molecules, we simply analyze the two hydration sites HS1 and HS2 for both complexes for estimation of the desolvation costs. As both hydration sites had an occupancy of approxi-



**Figure 1.** Schematic depiction of the binding mode of BACE-1 inhibitors. Hydration sites are given with their corresponding  $\Delta G$  values. The modifications in the ligand structure are highlighted in magenta.



mately 1, we can assume that this approximation accounts for most of the solvent perturbation [Eq. (1)]:

$$\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) = \Delta G_{\text{HS2}}(\text{B}) - [\Delta G_{\text{HS1}}(\text{A}) + \Delta G_{\text{HS2}}(\text{A})] = 1.41 \text{ kcal mol}^{-1} \quad (1)$$

Calculating the contribution of a sub-region **R** to the total free energy of the system needs a correction due to the double-counting of the water–water interactions in the system.<sup>[17]</sup> For Cumming-4j, the region contains the two hydration sites HS1 and HS2. We therefore subtract the interaction energy between HS1 and HS2 to account for the double counting. The interaction energy amounts to  $+0.92 \text{ kcal mol}^{-1}$  and since in the convention of IFST the water–water interaction energies are scaled by a factor of 0.5, we subtract  $0.46 \text{ kcal mol}^{-1}$  from the complex solvation free energy of Cumming-4j (see Experimental Section for explanation). The solvation free energy is thereby lowered, and the cost of water displacement is slightly increased, leading to Equation (2):

$$\Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{AP} \rightarrow \text{BP}) = 1.87 \text{ kcal mol}^{-1} \quad (2)$$

### Ligand desolvation

From the perspective of the protein desolvation energy, going from the 3-pyridyl to the 3-methoxyphenyl scaffold should decrease the binding affinity by a factor of approximately 10. The picture, however, changes when we take into account the difference in the hydration free energies of the two inhibitors. The free energy of hydration is expected to decrease for Cumming-4b, and our FEP calculation confirms this trend. The calculated  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  amounts to  $3.14 \pm 0.08 \text{ kcal mol}^{-1}$  and therefore compensates for the protein desolvation penalty. The modification in the ligand scaffold that leads to the water displacement also leads to an increase of the free energy of solvation, and therefore the costs of the water displacement from the binding site can be compensated.

### Protein–ligand interactions

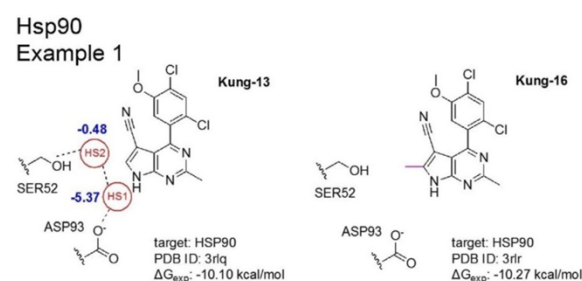
Apart from solvation effects, the change in protein–ligand interaction energies  $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB})$  also plays an important role regarding the thermodynamics of water displacement by functional groups. When a water molecule is sterically displaced, additional interactions between the protein and the displacing functional group of the ligand are formed. In case these interactions are favorable, this leads to an additional compensation of desolvation costs.

The averaged difference in interaction energy between BACE-1 and Cumming-4j (PA) and Cumming-4b (PB), respectively, amounts to  $-7.22 \text{ kcal mol}^{-1}$ . This is accompanied by a higher entropic cost for binding for Cumming-4b ( $-T\Delta\Delta S_{\text{bind}} = 1.88 \text{ kcal mol}^{-1}$ ). In total, this results in an estimated change in binding free energy [see Equation (20), Experimental Section] as shown in Equation (3).

$$\begin{aligned} \Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) &= \Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{AP} \rightarrow \text{BP}) - \Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) \\ &+ \Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) - T\Delta\Delta S_{\text{bind}}(\text{A} \rightarrow \text{B}) = -6.52 \text{ kcal mol}^{-1} \end{aligned} \quad (3)$$

With this approximation, the gain in free energy upon addition of the methoxy group is vastly overestimated. Harder et al. pointed out that accurate sampling of the helix surrounding the S3 pocket is crucial for accurate predictions of binding free energies and that otherwise an overestimation of the affinity of ligands with large substituents occurs.<sup>[28]</sup>

Kung et al.<sup>[29]</sup> used rational design strategies to target crystallographic water molecules for binding heat shock protein 90 (Hsp90; Figure 2). Two water molecules close to the pyrrolopyrimidine core (PDB ID: 3rlq) are sterically displaced by introducing a methyl group (PDB ID: 3rlr). The affinity thereby remains nearly unchanged.



**Figure 2.** Schematic depiction of the binding mode of Hsp90 inhibitors Kung-13 und Kung-16.

HS1 is a stable hydration site with a  $\Delta G$  value of  $-5.37 \text{ kcal mol}^{-1}$  (Supporting Information Table S2). It interacts tightly with the solute ( $-14.89 \text{ kcal mol}^{-1}$ ) and is stabilized by favorable water–water interactions with HS2 ( $-4.70 \text{ kcal mol}^{-1}$ ). HS2 has weaker interactions with the solute ( $-9.94 \text{ kcal mol}^{-1}$ ), but also benefits from the HS1–HS2 hydrogen bonding interaction. Both hydration sites have a small entropic destabilization ( $\sim 0.5 \text{ kcal mol}^{-1}$ ). Because the two hydration sites are fairly buried and not in direct contact with other hydration sites, we assume the solvent reorganization energy upon displacement of HS1 and HS2 to be negligible. We therefore estimate the protein desolvation costs for introduction of the methyl group as in Equation (4):

$$\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) = -[\Delta G_{\text{HS1}}(\text{A}) + \Delta G_{\text{HS2}}(\text{A})] = 5.85 \text{ kcal mol}^{-1} \quad (4)$$

We again use the correction for the double counting of the water–water interactions and obtain Equation (5):

$$\Delta\Delta G_{\text{solv}}^{\text{corr}} = 3.40 \text{ kcal mol}^{-1} \quad (5)$$

Methylation increases the nonpolar surface of the inhibitor, and in this case leads to an increase in the solvation free

energy (decrease in water solubility) of the ligand by  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) = 1.05 \pm 0.03 \text{ kcal mol}^{-1}$  (calculated by alchemical FEP). The change in ligand solvation free energy counteracts the unfavorable change in protein solvation free energy.

$\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB})$  amounts to  $-3.09 \text{ kcal mol}^{-1}$  due to the additional contacts formed by the newly introduced methyl group. Thus the change in  $\Delta\Delta G_{\text{solv}}(\text{AP} \rightarrow \text{BP})$  is fully compensated by  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  and  $\Delta\Delta E_{\text{inter}}(\text{AP} \rightarrow \text{BP})$ . The interaction entropy remains nearly unchanged ( $+0.07 \text{ kcal mol}^{-1}$ ).

$$\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) = -0.67 \text{ kcal mol}^{-1} \quad (6)$$

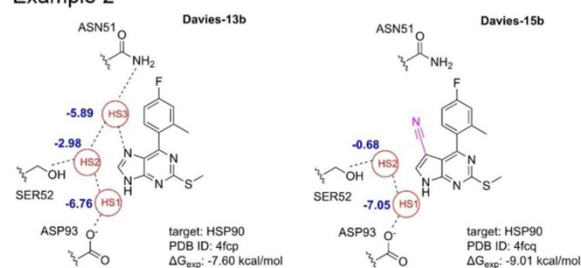
This result is in good agreement with the experimental value of  $-0.17 \text{ kcal mol}^{-1}$ . Even though the two hydration sites do not display perfect saturation of hydrogen bonding with their environment, their displacement is still associated with a desolvation cost. Because the added methyl group of the ligand forms additional interactions with the protein, the cost of water displacement can be compensated, explaining why the affinity remains unchanged upon the change in ligand structure. The addition of a methyl group to an aromatic ring is not expected to lead to a significant change in the dynamics of the protein and the ligand and therefore a more accurate prediction of the change in free energy is possible.

Davies and co-workers targeted crystallographic water molecules during the rational design of Hsp90 inhibitors (Figure 3).<sup>[30]</sup> By changing the scaffold from purine (Davies-13b) to cyanopyrrolopyrimidine (Davies-15b), a crystallographic water molecule bridging the interaction between the inhibitor and ASN51 is displaced, leading to a gain in affinity by one order of magnitude.

HS1 interacts with the ASN51 side chain, the ring nitrogen atom of the ligand Davies-13b and HS2. These strong interactions lead to a favorable  $\Delta E$  value of  $-6.39 \text{ kcal mol}^{-1}$  and a  $\Delta G$  value of  $-5.89 \text{ kcal mol}^{-1}$  (Supporting Information Figure S3). Displacement of water from HS1 is therefore associated with a considerable desolvation penalty. Upon displacement of HS1, HS2 is destabilized due to the loss of the hydrogen bonding interaction with HS1. This destabilization increases the desolvation cost [Eq. (7)].

$$\begin{aligned} \Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) &= [\Delta G_{\text{HS2}}(\text{B}) + \Delta G_{\text{HS3}}(\text{B})] \\ &- [\Delta G_{\text{HS1}}(\text{A}) + \Delta G_{\text{HS2}}(\text{A}) + \Delta G_{\text{HS3}}(\text{A})] = 7.90 \text{ kcal mol}^{-1} \end{aligned} \quad (7)$$

### Example 2



**Figure 3.** Schematic depiction of the binding mode of Hsp90 inhibitors Davies-13b and Davies-15b.

Correcting for the double counting of the water–water interactions yields Equation (8):

$$\Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB}) = 5.75 \text{ kcal mol}^{-1} \quad (8)$$

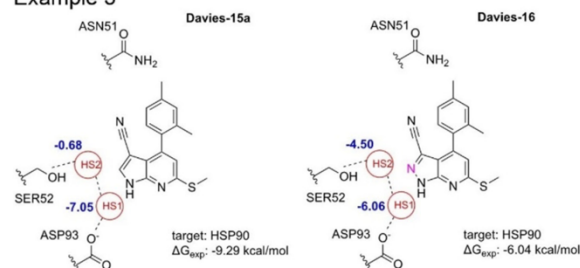
The change in ligand solvation free energy in solution  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  amounts to  $0.32 \pm 0.05 \text{ kcal mol}^{-1}$  according to the alchemical FEP calculation. Regarding protein–ligand interaction energies, the added nitrile group forms additional contacts with the protein (hydrogen bond to ASN52) and leads to a more favorable protein–ligand interaction energy ( $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = -3.11 \text{ kcal mol}^{-1}$ ) and an increase in interaction entropy ( $-0.7 \text{ kcal mol}^{-1}$ ). This in total results in the following predicted change in binding free energy [Eq. (9)]:

$$\begin{aligned} \Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) &= \Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB}) \\ &- \Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) + \Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = -1.62 \text{ kcal mol}^{-1} \end{aligned} \quad (9)$$

The estimated change in binding free energy is in disagreement with the experimental value of  $-1.41 \text{ kcal mol}^{-1}$ . Yet, this example displays the effect of solvent rearrangement due to a change in ligand structure. Displacing a hydration site that interacts favorably with neighboring hydration sites is expected to increase their free energy due to weaker water–water interactions ( $E_{\text{ww}}$ ). This effect can only be accurately quantified by calculating the solvation energies for both states PA and PB. Once again, the data indicate not that the water displacement itself is energetically favorable, but that counteracting terms (change in ligand free energy of solvation and protein–ligand interaction energy) can compensate for the protein desolvation costs.

In another example from the same series of Hsp90 inhibitors,<sup>[30]</sup> going from a cyanopyrrolopyrimidine (Davies-15a) to cyanopyrazolopyrimidine (Davies-16) led to a drastic decrease in affinity by two orders of magnitude (Figure 4). The authors suggest that the additional ring nitrogen atom engages in electronic repulsion with the water oxygen atom of HS2 and therefore leads to a destabilization of the water network.<sup>[30]</sup> We applied a HSA to validate this hypothesis, but found that the presence of the additional ring nitrogen stabilizes the water network. This example does not involve displacement of a

### Example 3



**Figure 4.** Schematic depiction of the binding mode of Hsp90 inhibitors Davies-15a and Davies-16.

water molecule, but rather uncovers the effect of solvent stabilization/destabilization.

HS2 interacts strongly with the ring nitrogen of Davies-16 and therefore has a considerably lower  $\Delta G$  relative to Davies-15a (Supporting Information Table S4), where it only weakly interacts with the pyrrole core, whereas HS1 suffers from a slight stabilization of  $1 \text{ kcal mol}^{-1}$  for the pyrazole core. In total, the addition of the ring nitrogen leads to a decrease in the complex solvation free energy by  $\sim 3 \text{ kcal mol}^{-1}$  [Eq. (10)]:

$$\Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB}) = -3.14 \text{ kcal mol}^{-1} \quad (10)$$

This is counteracted by the change in ligand solvation free energy ( $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) = -1.18 \pm 0.03 \text{ kcal mol}^{-1}$ ). The protein–ligand interactions stay nearly unchanged [Eqs. (11) and (12)]:

$$\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = -0.42 \text{ kcal mol}^{-1} \quad (11)$$

$$\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) = -1.09 \text{ kcal mol}^{-1} \quad (12)$$

Our calculation cannot successfully reproduce the experimentally observed change in binding free energy ( $3.25 \text{ kcal mol}^{-1}$ ). Still, the IFST calculations indicate that the pyrazole core stabilizes the water network around the ligand, and questions the explanation of the weaker affinity presented in the original study.

Woodhead and co-workers discovered a high-affinity 2,4-dihydroxyphenyl inhibitor (Woodhead-1). By exchanging a hydroxy group with a methoxy (Woodhead-2), HS1 and HS2 are displaced and the affinity drops by a factor of 30 (Figure 5).

Compared with examples 1 and 2, HS1 and especially HS2 are considerably better stabilized due to strong interactions with the hydroxy group of the ligand (Supporting Information Table S5). Both hydration sites interact strongly with the protein ( $-18.17$  and  $-19.78 \text{ kcal mol}^{-1}$ ) and favorably with each other ( $-3.84 \text{ kcal mol}^{-1}$ ), and displacement of these two hydration sites is associated with a vast desolvation cost [Eq. (13)]:

$$\Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB}) = 15.01 \text{ kcal mol}^{-1} \quad (13)$$

This example shows how much the presence of the ligand affects the thermodynamic properties of the water network. HS1 and HS2 correspond to the same hydration sites that are discussed in Hsp90 example 1. Whereas HS1 and HS2 cannot

undergo an optimal interaction pattern in Hsp90 example 1 and therefore have higher free energies, they form a strong hydrogen bonding network in Hsp90 example 4. This explains why the scaffold modification in Hsp90 example 4 results in a much higher desolvation cost of the ligand–receptor complex relative to example 1, where the addition of a methyl group leaves the affinity unchanged. In example 1, the hydration sites are exposed to a hydrophobic part of the ligand (Kung-13). In example 4, on the other hand, the ligand (Woodhead-1) provides hydrogen bonding donor functionalities that lead to low  $\Delta G$  values of the two hydration sites.

Methylation of one of the two hydroxy groups is expected to make the compound more lipophilic and therefore increases the free energy of solvation for the isolated ligand, and indeed  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  obtained by FEP amounts to  $+2.60 \pm 0.07 \text{ kcal mol}^{-1}$ .

In terms of protein–ligand interaction energies, the difference is marginal ( $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = 0.56 \text{ kcal mol}^{-1}$ ). The binding entropy is slightly more favorable ( $-T\Delta\Delta S_{\text{bind}} = -1.48 \text{ kcal mol}^{-1}$ ), leading to a total change in binding free energy of  $\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) = \Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB}) - \Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) + \Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = 11.49 \text{ kcal mol}^{-1}$ .

Whereas our estimation of  $\Delta\Delta G_{\text{bind}}$  correctly predicts an affinity loss for this ligand modification, the absolute number is vastly overestimated.

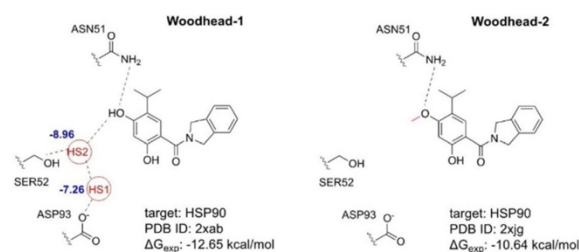
### Scytalone dehydratase

Chen et al. developed scytalone dehydratase inhibitors and observed that changing the benzotriazine core to a cyanocinnoline facilitated the displacement of a structural water molecule that resulted in a considerable gain in binding affinity (Figure 6).<sup>[31]</sup>

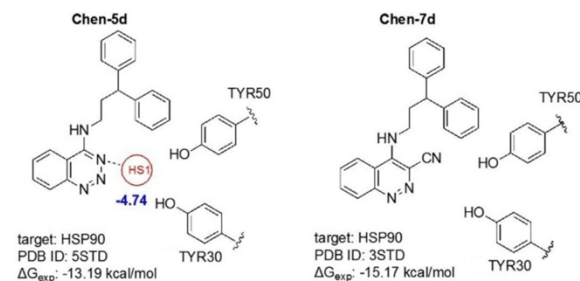
HS1 interacts with the benzotriazine nitrogen of inhibitor Chen-5d and the side chains of Tyr30 and Tyr50. Displacement of this hydration site is associated with a desolvation cost of  $4.74 \text{ kcal mol}^{-1}$  (Supporting Information Table S6; [Eq. (14)]:

$$\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) = 4.74 \text{ kcal mol}^{-1} \quad (14)$$

Furthermore, the solvation free energy for Chen-7d obtained from the FEP calculation is lower:  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) = -2.52 \pm 0.09 \text{ kcal mol}^{-1}$ . Therefore, regarding the solvation free energies, adding the cyano group creates a desolvation penalty for



**Figure 5.** Schematic depiction of the binding mode of Hsp90 inhibitors Woodhead-1 and Woodhead-2.



**Figure 6.** Schematic depiction of the binding mode of scytalone dehydratase inhibitors.



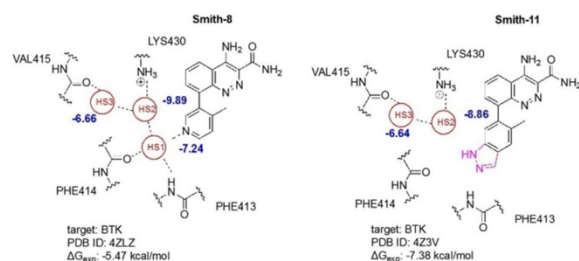
the removal of the stable hydration site HS1 and furthermore, leads to a decrease in ligand solvation free energy; both effects are detrimental for ligand binding. However, Chen-7d interacts more strongly with the protein, which is expected upon addition of the cyano group.  $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB})$  equals  $-11.75 \text{ kcal mol}^{-1}$ , and this explains the affinity gain of the cyanocinnoline inhibitor. A slight loss in binding entropy occurs ( $2.78 \text{ kcal mol}^{-1}$ ; [Eq. (15)]):

$$\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) = \Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) - \Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) + \Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = -1.71 \text{ kcal mol}^{-1} \quad (15)$$

The experimental  $\Delta\Delta G_{\text{bind}}^{\text{exp}}(\text{A} \rightarrow \text{B})$  is  $-1.98 \text{ kcal mol}^{-1}$  and therefore in excellent agreement with our calculation. Gerogiou et al. used the same inhibitor pair to evaluate the water displacement energetics using grid cell theory.<sup>[18]</sup> With their approach, they obtained a  $\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB})$  value of  $4.4 \text{ kcal mol}^{-1}$ , which is in very close agreement with our value of  $4.74 \text{ kcal mol}^{-1}$ . Furthermore, they report a  $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB})$  value of  $-13.6 \text{ kcal mol}^{-1}$ , confirming the tight interactions formed by the cyano group.

### Bruton's tyrosine kinase

Smith et al. reported the discovery of small-molecule inhibitors of Bruton's tyrosine kinase (BTK).<sup>[32]</sup> Extending the pyridyl scaffold (Smith-8) to an indazole (Smith-11) leads to a gain in inhibitory potency expressed by a lower  $\text{IC}_{50}$  value (Figure 7). A network of three hydration sites is present in the vicinity of the pyridine ring and one hydration site (HS1) is displaced by the indazole group.



**Figure 7.** Schematic depiction of the binding mode of BTK inhibitors.

In the presence of inhibitor Smith-8, all three hydration sites exhibit low  $\Delta G$  values, due to strong interactions with the solute as well as favorable water–water interactions. Upon displacement of HS1, HS2 shows weaker water–water interactions (Supporting Information Table S7), but has more favorable interactions with the protein. In total, there is a net destabilization of  $\sim 1 \text{ kcal mol}^{-1}$  for HS2 due to displacement of HS1, whereas HS3 does not seem to be affected significantly by the presence of HS1. The total change in binding site solvation free energy  $\Delta\Delta G_{\text{solv}}^{\text{corr}}(\text{PA} \rightarrow \text{PB})$  is  $6.22 \text{ kcal mol}^{-1}$ .

Estimation of the change in ligand solvation free energy by means of FEP yielded  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) = -3.82 \pm 0.06 \text{ kcal mol}^{-1}$ .

The ring extension leads to significantly stronger interactions with the protein ( $\Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) = -18.42 \text{ kcal mol}^{-1}$ ), partly cancelled out by an entropic penalty ( $-T\Delta\Delta S_{\text{bind}} = 4.18 \text{ kcal mol}^{-1}$ ), and this finally results in a calculated  $\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) = -4.20 \text{ kcal mol}^{-1}$ . The experimental difference in binding free energy is  $-1.91 \text{ kcal mol}^{-1}$ .

### Implications for ligand design

It is clear that the method of choice for calculating accurate relative free binding energies are alchemical free energy calculations,<sup>[33,34]</sup> as they are based on a rigorous thermodynamic framework. The goal of our study was not to postulate a new way of predicting relative free binding energies, but was merely to gain additional insight into the thermodynamics of water displacement by studying several cases from the literature. While alchemical free energy perturbations can give accurate predictions, they often do not facilitate understanding of the underlying factors that led to the change in binding affinity. Furthermore, FEP calculations that involve water displacement from deeply buried pockets can suffer from insufficient sampling of the buried water molecules,<sup>[35]</sup> which can lead to trapping of water molecules in unfavorable states.<sup>[36,37]</sup>

We saw that for all discussed examples, the displacement of the water molecule was associated with a desolvation cost, even for water molecules that do not display full hydrogen bond saturation with the protein (Hsp90, example 1). For systems where an isolated water molecule is displaced (scytalone dehydratase), estimating the desolvation cost is straightforward. On the other hand, if the water molecule of interest is in close contact with other water molecules, two cases must be distinguished: 1) The water molecule interacts favorably with the neighboring water molecules (Hsp90 example 1,2,3,4, BTK) or it interacts with the same functional group of the protein as another water molecule (BACE-1). In the first case, displacement of the water molecule destabilizes the neighboring hydration sites due to a loss in water–water interactions, except for the case that the incoming functional group can compensate for this by mimicking the favorable interactions with the remaining water network. In case the interactions are not compensated, the destabilization of the water network increases the costs of the binding site desolvation.

The opposite is true if a water molecule is displaced that has unfavorable water–water interactions with its neighbors. In the example of BACE-1, the water molecule that gets displaced had been interacting with the same backbone carbonyl group as another water molecule. This arrangement leads to a slight H–H repulsion and a displacement of one of these waters leads to a stabilization of the other. This counteracts the desolvation costs.

The strength of IFST is that it splits the energies of the hydration site into solute–water and water–water terms. It can therefore be easily seen if the hydration site interacts favorably or unfavorably with the other solvent, and this information can guide the decision on water displacement.

In two of the discussed examples, a nitrile group was used as a ring extension replacing a ring nitrogen group involved in



a bridged ligand–water–protein interaction. The geometry of the nitrile group allows displacement of the water molecule<sup>[30]</sup> and furthermore can replace the acceptor functionality of the water oxygen atom by being an even stronger acceptor. Furthermore, it can form additional hydrophobic contacts perpendicular to the  $C\equiv N$  bond and is therefore an attractive group for water displacement. It also does not add rotatable bonds to the ligand. An important aspect of the nitrile group is its effect on the ligand solvation free energy. In one case (Hsp90, example 2), the ligand solvation free energy was increased, whereas for the scytalone dehydratase example, it was decreased. Due to its electron-withdrawing properties, the exact effect of the nitrile substituent on the solvation free energy seems to be dependent on the ring system.

An interesting ring extension that leads to water displacement was shown in the example of BTK.<sup>[32]</sup> The indazole ring can replace the acceptor as well as the donor functionality of the water molecule, and can form strong, direct contacts with the protein.

The remaining examples of scaffold changes led to “unspecific” water displacement. In these cases, the hydrogen bonding interactions of the water molecule are not compensated by the displacing functional group (BACE-1, Hsp90 example 1,4). Still, in the first two cases, the displacement was accompanied by no change or a slight increase in affinity. For BACE-1 and Hsp90 example 1, the displaced hydration sites displayed the highest free energies of all discussed hydration sites, indicating that they are quite unstable. The additional contacts formed by the methyl and methoxy groups can successfully compensate these low desolvation costs. Displacement of tightly interacting water molecules by such nonpolar substituents on the other hand is not expected to be successful,<sup>[6]</sup> as shown for Hsp90 example 4.

## Conclusions

We collected examples from the literature, where small ligand modifications led to the displacement of water molecules or a rearrangement of the water network. For all examples, thermodynamic data are present, and the binding modes are supported by high-resolution crystal structures. To our knowledge, such a high-quality dataset has not yet been summarized, and we expect it to be of interest in the field. Apart from the practical insight, it can also be used to benchmark computational methods that predict thermodynamic properties of binding site water molecules.

Our IFST calculations showed that various factors determine if displacement of a binding site water molecule results in a loss or gain in affinity. Not only do these examples provide valuable insight into displacement costs and rearrangement of water molecules, they are also a good benchmark for assessing the predictive power of IFST and other methods for sample cases relevant for lead optimization.

Whereas we could predict if the change in ligand structure should lead to an increase or decrease in binding free energies, quantitatively predicting the exact  $\Delta\Delta G_{\text{bind}}$  value with chemical accuracy remains challenging, implying that dissect-

ing the various terms contributing to a change in binding affinity is difficult, and determining the exact quantitative effect of water molecules still needs improvement in the field. Recent thermodynamic studies on congeneric thrombin inhibitors showed that in the context of water displacement, small changes in ligand structures can lead to dramatic changes in thermodynamic signatures.<sup>[38,39]</sup> This points out that lead optimization guided by water thermodynamics is difficult, and that most approaches presented in the literature<sup>[6,12,13,18]</sup> are too simplistic for quantitative predictions. A prediction of a change in binding free energy upon scaffold minimization can be feasible if the dominating terms are only the change in interaction energy and change in solvation free energy. However, our examples showed that this is not even given for highly congeneric ligands with the same binding mode and only minor structural differences.

In terms of qualitative predictions, information on binding site water thermodynamics can surely guide lead optimization stages by shedding light on which water molecules are worth attempting a displacement and which functional group should be chosen. Furthermore, in absence of information on water positions from crystal structures, molecular simulations can predict the location of structural water molecules.<sup>[40]</sup>

Regarding the limitations of IFST, an important question is associated with the treatment of the second-order water–water entropy,  $S_{\text{ww}}$ . Huggins has shown that this term is small for buried, isolated water molecules.<sup>[27]</sup> Therefore, we regarded it as appropriate to approximate  $\Delta S = S_{\text{sw}} - S_{\text{bulk}}$ . However, the second-order entropy can become significant for highly hydrated binding sites at the protein surface.<sup>[41]</sup>

Another important aspect is the need for solute heavy-atom restraints in the context of IFST. Deriving the solvation free energies from a restrained MD simulation and the interaction energies and ligand solvation free energies from flexible simulations is expected to bias the results. The importance of flexibility on the outcome of a hydration site analysis has been previously described,<sup>[42]</sup> but there is clearly a need for more work and effective protocols that account for flexibility in the context of IFST calculations.

The presented calculations suffer from the general weaknesses of simulations based on molecular force fields, such as neglect of polarization or the simplistic point charge models. Regarding the representation of water molecules, it is questionable if simple models such as TIP3P or TIP4P, that were parametrized to reproduce properties of bulk water, yield accurate energies for structural waters, which can be expected to be polarized differently.<sup>[43]</sup> These shortcomings can explain why reaching chemical accuracy with our presented estimation is not achievable for all cases; however, for the two examples were the deviation between estimated and experimental change in binding free energy was 5 kcal mol<sup>−1</sup> or more, alternative explanations have to be considered. Insufficient sampling can be an important drawback. Regarding the example containing the BACE-1 inhibitors, a ring flip can occur for Cumming-4j, leading to a change in binding interactions, binding entropy, and a different solvation pattern as in the crystal structure.

Overall, valuable insight into the thermodynamics of water displacement and the impacts on ligand design were gained. We demonstrated that displacement of solvent molecules can stabilize or destabilize the remaining water network. These effects can only be investigated by running simulations of both states (PA, PB). In our opinion, these effects have not been discussed in detail in previous studies, but they have to be taken into account for a realistic view on the role of water in ligand binding.

In terms of solvation free energies, the solvent displacement per se leads to a desolvation penalty in the discussed examples and the balance between the solvation free energies of the ligands in solution, the solvation free energies of the ligand–receptor complexes and the change in protein–ligand interactions determine the impact of the ligand modification on the binding free energy. All input structures and results of the HSA are made publicly available, and the corresponding link is given in the Supporting Information (S8).

## Experimental Section

### Hydration site analysis (HSA)

Hydration sites correspond to spherical regions (typically with a radius of 1.0 Å) with high water density. Hydration sites are identified by a clustering procedure applied on the water positions from an MD simulation with restrained solute. We used the same clustering approach as described elsewhere.<sup>[6]</sup> The thermodynamic properties of the hydration sites were calculated by means of IFST, a theoretical framework introduced by Lazaridis.<sup>[21]</sup> IFST calculates the perturbation of the solvent due to the presence of the solute and facilitates calculation of energies and entropies of solvation. It can be directly applied to trajectories from MD or Monte Carlo (MC) simulations. The energetic contributions to the free energy of solvation are calculated by taking the interaction energy of the hydration site with solute and remaining solvent and comparing it with the interaction energy of a water in bulk. The entropy is calculated by solute–solvent and solvent–solvent correlation functions.<sup>[21]</sup> In our implementation, the entropic contribution was calculated by means of the k-nearest neighbor method<sup>[27]</sup> using the total distance metric for the solute–water entropy<sup>[44]</sup> and taking into account the C2v symmetry of the water molecule. Only the first order water–solute entropy was calculated, and we approximated  $\Delta S_{\text{ww}}$  as  $-S_{\text{bulk}}$  as suggested by Huggins.<sup>[44]</sup> The value of  $S_{\text{bulk}}$  depends on the water model and was taken from the literature.<sup>[45]</sup>  $\Delta E$  of a hydration site was calculated as  $\Delta E = E_{\text{sw}} + \frac{1}{2}E_{\text{ww}} - E_{\text{bulk}}$ .  $E_{\text{sw}}$  is the solute–water interaction,  $E_{\text{ww}}$  the water–water interaction.  $E_{\text{bulk}}$  was derived from a simulation of a TIP4P-Ew water box and amounts to  $-11.412 \text{ kcal mol}^{-1}$ .

Calculating the contribution of a region **R** to the total solvation free energy of a system requires a correction due to double counting of the water–water interactions in the sub-region.<sup>[17]</sup> Let us consider a region with two structural water molecules (hydration sites, HS2 and HS2).  $E_{\text{ww}}$  for this region is composed of the sum of the interactions of HS1 and HS2, HS1 and the remaining solvent, and of HS2 with HS1 and HS2 with the remaining solvent. In that case, the interaction between HS1 and HS2 is double counted and we have to subtract it from the total water–water interactions of the region **R**.

The thermodynamic properties ( $\Delta G$ ,  $\Delta E$ ,  $\Delta S$ ) of the hydration sites are calculated by averaging over all water molecules occupying the site during the simulation. We considered the hydration site approach to be a valid method for studying the binding site solvation free energies of the system. The drawback of hydration sites is usually that they cannot fully account for the full water density in a region, contrary to grid-based approaches.<sup>[17]</sup> See the Supporting Information (S9) for an evaluation of how much water density is not accounted for by a HSA for the different binding sites. In our examples, we were interested in the impact of small ligand modifications on the binding site solvation free energies. In this case, only the change of solvation free energy in proximity of the modification has to be calculated because we are interested in relative differences between pairs of ligands  $\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB})$ . Because we specifically looked for published examples where the structural change of the ligand leads to the displacement of a structural water molecule with a well-defined position, a hydration site is expected to capture the full contribution of this water molecule to the total solvation free energy of the system. We used the following strategy: if an isolated water molecule (a water molecule not directly interacting with any other water molecules) was displaced, the change in the complex solvation free energy  $\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) \Delta G_{\text{bind}}(\text{A}) = \Delta G_{\text{solv}}(\text{PA}) - \Delta G_{\text{solv}}(\text{A}) - \Delta G_{\text{solv}}(\text{P}) + \Delta G_{\text{bind,gas}}(\text{A})$  was taken as the  $\Delta G$  of this hydration site. If the displaced water molecule was interacting with other molecules, then the change in  $\Delta G$  of these hydration sites was also taken into the equation, to account for the solvent rearrangement.

In the original formulation of IFST proposed by Lazaridis, positive  $\Delta G$  values of hydration sites occur rarely, in contrast to other formulations of the theory,<sup>[6]</sup> where positive free energies of hydration sites occur frequently. This is due to the scaling of the water–water interactions by a factor of 0.5 in the original formulation and we adapted the same convention. Regarding isolated hydration sites, a positive  $\Delta G$  would imply that this hydration site makes an unfavorable, positive contribution to the total solvation free energy of the system and is expected to be vacated. In this case, a positive  $\Delta G$  of a hydration site in the given formulation of IFST would rather point out to force field inaccuracies or is due to approximations made by the theory. However, regarding water networks, a hydration site can interact with neighboring hydration sites and add a net stabilization to the water network. Even when this hydration site has a positive  $\Delta G$ , its stabilizing effect can in total lead to a lower solvation free energy and in such an example a positive free energy is not an artifact.

### Estimation of thermodynamic contributions to the binding free energy

The thermodynamic cycle for protein–ligand binding is depicted in Figure 8, and the binding free energy for ligand A binding to protein P is expressed as shown in Equation (16):

$$\Delta G_{\text{bind}}(\text{A}) = \Delta G_{\text{solv}}(\text{PA}) - \Delta G_{\text{solv}}(\text{A}) - \Delta G_{\text{solv}}(\text{P}) + \Delta G_{\text{bind,gas}}(\text{A}) \quad (16)$$

The change in binding free energy by going from ligand A to B is then expressed in Equation (17):

$$\begin{aligned} \Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) &= [\Delta G_{\text{solv}}(\text{PB}) - \Delta G_{\text{solv}}(\text{PA})] - [\Delta G_{\text{solv}}(\text{B}) - \Delta G_{\text{solv}}(\text{A})] \\ &+ [\Delta G_{\text{bind,gas}}(\text{B}) - \Delta G_{\text{bind,gas}}(\text{A})] \end{aligned} \quad (17)$$

Where the first term describes the change in the solvation free energy of the complex, the second term the change in solvation



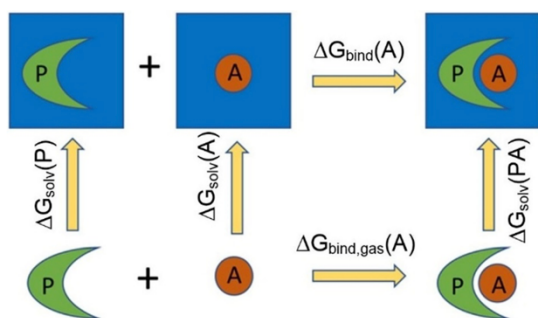


Figure 8. Thermodynamic cycle of protein–ligand binding.

free energy of the ligand and the third term the change in gas phase binding free energy. The method of choice for computationally obtaining accurate estimates for relative binding free energies are relative binding free energy (RBFE) calculations such as free energy perturbation (FEP) or thermodynamic integration (TI) and the authors point out to a recently published review about the subject.<sup>[36]</sup>

So-called end-state methods on the other hand, such as molecular mechanics generalized Born surface area solvation (MM-GBSA)<sup>[46]</sup> estimate the various contributions to the binding free energy using force-field energies in combination with implicit solvent models.

The IFST methodology<sup>[21]</sup> uses an explicit solvent representation and facilitates calculation of the contribution of sub-volumes of the system to the total free energy of solvation.<sup>[17]</sup> Considering a congeneric pair of ligands A and B, then the structural modification leads to a perturbation of the solvent structure in close spatial proximity. One therefore has to define a monitored region **R** that captures the full change in complex solvation free energy by going from ligand A to B.<sup>[17,37,47]</sup> This can be expressed as in Equation (18):

$$\begin{aligned}\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) &= \Delta G_{\text{solv}}(\text{PB}) - \Delta G_{\text{solv}}(\text{PA}) \\ &\approx \Delta G_{\text{solv},\text{R}}(\text{PB}) - \Delta G_{\text{solv},\text{R}}(\text{PA})\end{aligned}\quad (18)$$

In principle, grid-based methods (grid inhomogeneous solvation theory (GIST)<sup>[17]</sup> or grid cell theory (GCT)<sup>[37]</sup>) facilitate flexible definitions of the monitored region, whereas a HSA only accounts for the contributions of regions with high solvent density.

The accurate choice of the monitored region is nontrivial and depends on the problem at hand. Introducing a charged group close at the protein surface is expected to create long-range perturbations and only a sufficiently large region **R** is expected to capture the full contribution.

The other extreme would be the addition of a nonpolar substituent in a highly rigid, buried pocket that leads to the displacement of an isolated, structural water molecule. In this case, the solvent rearrangement is small and the change in the complex solvation free energy can be approximated as the cost of displacing this specific water molecule and therefore, IFST in combination with HSA can be expected to lead to a reasonably good estimation. If ligand B displaces the hydration site, then the relative change in complex solvation free energy is:

$$\Delta\Delta G_{\text{solv}}(\text{PA} \rightarrow \text{PB}) = -\Delta G_{\text{HS}}\quad (19)$$

Where  $\Delta G_{\text{HS}}$  is the solvation free energy of the displaced hydration site.

In the discussed examples, all displaced water molecules were located in buried pockets; however, not all of them were isolated. In case the water molecule displaced by ligand B interacts with neighboring water molecules in the presence of ligand A, then the change in thermodynamic properties of the corresponding hydration sites must be calculated, as shown in various examples.

To finally estimate the change in binding free energy by going from ligand A to B, we calculated the change in ligand solvation free energy  $\Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B})$  by means of FEP. The change in gas-phase binding free energy  $\Delta\Delta G_{\text{bind,gas}}(\text{A} \rightarrow \text{B})$  was approximated as the change in protein–ligand interaction energy  $\Delta\Delta E_{\text{inter}}(\text{A} \rightarrow \text{B})$  plus the change in interaction entropy  $-T\Delta\Delta S_{\text{inter}}(\text{A} \rightarrow \text{B})$ .<sup>[25]</sup>

This leads to the following equation for the approximation of the relative binding free energies [Eq. (20)]:

$$\begin{aligned}\Delta\Delta G_{\text{bind}}(\text{A} \rightarrow \text{B}) &= \Delta\Delta G_{\text{solv}}(\text{AP} \rightarrow \text{BP}) - \Delta\Delta G_{\text{solv}}(\text{A} \rightarrow \text{B}) \\ &+ \Delta\Delta E_{\text{inter}}(\text{PA} \rightarrow \text{PB}) - T\Delta\Delta S_{\text{bind}}(\text{A} \rightarrow \text{B})\end{aligned}\quad (20)$$

This approach is similar in flavor to common end-state methods such as MM-GBSA<sup>[46]</sup> applied to relative binding free energies and is inherently different from rigorous relative binding free energy calculations like FEP or thermodynamic integration (TI), that yield directly the free energy without any decomposition.

There is controversy regarding the usefulness of free energy decompositions,<sup>[48,49]</sup> as not all of the calculated thermodynamic contributions toward the binding affinity are state functions. Solute and solvent degrees of freedom can be separated,<sup>[50,51]</sup> and IFST gives the solvation free energy of a solute in a fixed conformation, sampled over the solvent degrees of freedom. In our approach, we just derived the complex solvation energy obtained from IFST for the supposedly dominant solute conformation (from the relaxed crystal structure), and this undoubtedly leads to errors in the prediction, because the interaction energies, entropies and ligand solvation free energies are computed from simulations with flexible solute. The influence of protein flexibility on the outcomes of IFST calculations has not yet been extensively investigated to our knowledge and definitely leaves space for further improvements to strengthen the applicability of IFST for quantitative predictions.

### Structure preparation

Input structures for the MD simulations were retrieved from the RCSB Protein Data Bank (PDB). Pre-processing of all structures was performed with the Protein Preparation Wizard, which is part of the Schrödinger Small-Molecule Drug Discovery Suite.<sup>[52]</sup> The crystallographic water molecules were kept, and all hydrogen atoms were added to the system and ligand and protein protonation states and tautomers were assigned for pH 7.4. The hydrogen bond network was optimized, and a short relaxation employing the OPLS\_2005 force field was performed. Because for the ligand pairs Davies-13b and Davies-15b no crystal structures were available, we took the structure of a close analogue of Davies-15b (PDB ID: 4fcq) and changed one methyl group into a fluorine atom. To get the structure of Davies-13b, we deleted the nitrile group and added the ring nitrogen and the water molecule. This seems justified because the crystal structure of another purine fragment (PDB ID: 4fcq) confirms the conserved binding mode and the presence of a water molecule for this scaffold family. Davies-16 was mod-

elled by converting the ring carbon of Davis-15b into a nitrogen and converting the fluoro substituent to a methyl.

For the scytalone dehydratase inhibitors, the input structure for Chen-7d is present (PDB ID: 3STD). There is no structure present for inhibitor Chen-5d, but for a close analogue (PDB ID: 5STD), where the presence of the water molecule is proven. To prepare an input structure for Chen-5d, the nitrile group was deleted and the ring system was modified. The water molecule was added at the position obtained from the structure 3STD and the hydrogen network was optimized.

### Experimental binding affinities

The experimental binding affinities for all ligand pairs were taken from the literature.  $\Delta G_{\text{exp}}$  was approximated by  $-RT \ln K_i$  or directly taken from the isothermal titration calorimetry (ITC) data if available. For inhibitors of Bruton's tyrosine kinase, we used the Cheng–Prusoff equation to convert the  $IC_{50}$  values into  $K_i$  values.

### MD simulations

All molecular dynamics simulations were carried out with Desmond.<sup>[53,54]</sup> The default relaxation protocol was applied, consisting of the following steps: 1) 100 ps Brownian dynamics in the NVT ensemble at 10 K and solute heavy-atom restraints ( $50 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ); 2) 12 ps simulation in the NVT ensemble, keeping the restraints and temperature at 10 K; 3) 12 ps simulation in the NPT ensemble, keeping restraints and temperature at 10 K; 4) 12 ps simulation in the NPT simulation with solute heavy atom restraints at 300 K; and 5) 24 ps simulation in the NPT ensemble at 300 K without restraints. Production runs were performed in the NPT ensemble at 300 K and 1.01 bar. The TIP4P-Ew water model was used, and counterions were added to neutralize the net charge of the system. The temperature and pressure were controlled using a Langevin thermostat (relaxation time: 1.0 ps) and a Langevin barostat (relaxation time: 2.0 ps). Electrostatic interactions were treated using the particle mesh Ewald calculation. The cutoff for nonbonded interactions was set at  $9.0 \text{ \AA}$ . The production runs of 40.0 ns were run by applying a constraint of  $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  to the solute heavy atoms for the IFST calculations. The coordinates were written every 8 ps, resulting in 5000 snapshots, which were then used for calculations of the solvent thermodynamics. For calculations of the protein–ligand interaction energies, the same settings were used, but without applying restraints. For the protein–ligand interaction energies and interaction entropies, 100 000 snapshots were saved and analyzed.

### FEP calculations

FEP calculations of the ligand modification in solution were performed by the LigandFEP<sup>[55,56]</sup> module of Desmond<sup>[53,54]</sup> using the FEP/REST (replica exchange with solute tempering) technology. The simulations were split into 12  $\lambda$  windows and the production runs lasted 5 ns. Relative solvation free energies were obtained by taking the difference between the free energy of the transformation of A to B in solution and the transformation in gas phase:

$$\Delta \Delta G_{\text{solv}}(A \rightarrow B) = \Delta G_{\text{aq}}(A \rightarrow B) - \Delta G_{\text{gas}}(A \rightarrow B) \quad (21)$$

The TIP4P-Ew water model was employed.

### Acknowledgements

J.W. is funded by the Schweizerischer Nationalfonds (SNF), grant number 200020\_159952.

### Conflict of interest

The authors declare no conflict of interest.

**Keywords:** binding affinity • inhomogeneous fluid solvation theory • ligand optimization • molecular dynamics • solvent displacement

- [1] C. Bissantz, B. Kuhn, M. Stahl, *J. Med. Chem.* **2010**, *53*, 5061–5084.
- [2] P. Y. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bachelier, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, S. Ericksonviitanen, *Science* **1994**, *263*, 380–384.
- [3] A. Wissner, D. M. Berger, D. H. Boschelli, M. B. Floyd, L. M. Greenberger, B. C. Gruber, B. D. Johnson, N. Mamuya, R. Nilakantan, M. F. Reich, R. Shen, H. R. Tsou, E. Upešlacis, Y. F. Wang, B. Q. Wu, F. Ye, N. Zhang, *J. Med. Chem.* **2000**, *43*, 3244–3256.
- [4] C. Liu, S. T. Wroblewski, J. Lin, G. Ahmed, A. Metzger, J. Wityak, K. M. Gillooly, D. J. Shuster, K. W. McIntyre, S. Pitt, D. R. Shen, R. F. Zhang, H. J. Zhang, A. M. Dowejko, D. Diller, I. Henderson, J. C. Barrish, J. H. Dodd, G. L. Schieven, K. Leftheris, *J. Med. Chem.* **2005**, *48*, 6261–6270.
- [5] J. D. Dunitz, *Science* **1994**, *264*, 670.
- [6] R. Abel, T. Young, R. Farid, B. J. Berne, R. A. Friesner, *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.
- [7] D. J. Huggins, W. Sherman, B. Tidor, *J. Med. Chem.* **2012**, *55*, 1424–1444.
- [8] M. S. Bodnarchuk, *Drug Discovery Today* **2016**, *21*, 1139–1146.
- [9] K. Haider, D. J. Huggins, *J. Chem. Inf. Model.* **2013**, *53*, 2571–2586.
- [10] C. N. Nguyen, A. Cruz, M. K. Gilson, T. Kurtzman, *J. Chem. Theory Comput.* **2014**, *10*, 2769–2780.
- [11] R. Abel, N. K. Salam, J. Shelley, R. Farid, R. A. Friesner, W. Sherman, *ChemMedChem* **2011**, *6*, 1049–1066.
- [12] D. J. Sindhikara, F. Hirata, *J. Phys. Chem. B* **2013**, *117*, 6718–6723.
- [13] A. S. Bayden, D. T. Moustakas, D. Joseph-McCarthy, M. L. Lamb, *J. Chem. Inf. Model.* **2015**, *55*, 1552–1565.
- [14] T. E. Balias, M. Fischer, R. M. Stein, T. B. Adler, C. N. Nguyen, A. Cruz, M. K. Gilson, T. Kurtzman, B. K. Shoichet, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6839–E6846.
- [15] S. Uehara, S. Tanaka, *Molecules* **2016**, *21*, 1604.
- [16] R. B. Murphy, M. P. Repasky, J. R. Greenwood, I. Tubert-Brohman, S. Jerome, R. Annabhimoju, N. A. Boyles, C. D. Schmitz, R. Abel, R. Farid, R. A. Friesner, *J. Med. Chem.* **2016**, *59*, 4364–4384.
- [17] C. N. Nguyen, T. K. Young, M. K. Gilson, *J. Chem. Phys.* **2012**, *137*, 044101.
- [18] G. Gerogiokas, M. W. Y. Southey, M. P. Mazanetz, A. Hefetz, M. Bodkin, R. J. Law, J. Michel, *Phys. Chem. Chem. Phys.* **2015**, *17*, 8416–8426.
- [19] T. Hasegawa, M. Sugita, T. Kikuchi, F. Hirata, *J. Chem. Inf. Model.* **2017**, *57*, 2789–2798.
- [20] M. Schaeperl, P. Czodrowski, J. E. Fuchs, R. G. Huber, B. J. Waldner, M. Podewitz, C. Kramer, K. R. Liedl, *J. Chem. Inf. Model.* **2017**, *57*, 345–354.
- [21] T. Lazaridis, *J. Phys. Chem. B* **1998**, *102*, 3531–3541.
- [22] T. Lazaridis, *J. Phys. Chem. B* **1998**, *102*, 3542–3550.
- [23] Z. Li, T. Lazaridis, *J. Phys. Chem. B* **2005**, *109*, 662–670.
- [24] Z. Li, T. Lazaridis, *J. Am. Chem. Soc.* **2003**, *125*, 6636–6637.
- [25] L. Duan, X. Liu, J. Z. H. Zhang, *J. Am. Chem. Soc.* **2016**, *138*, 5722–5728.
- [26] J. N. Cumming, E. M. Smith, L. Wang, J. Misiaszek, J. Durkin, J. Pan, U. Iserloh, Y. Wu, Z. Zhu, C. Strickland, J. Voigt, X. Chen, M. E. Kennedy, R. Kuvelkar, L. A. Hyde, K. Cox, L. Favreau, M. F. Czarniecki, W. J. Greenlee, B. A. McKittrick, E. M. Parker, A. W. Stamford, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444–2449.
- [27] D. J. Huggins, *Biophys. J.* **2015**, *108*, 928–936.

- [28] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, R. A. Friesner, *J. Chem. Theory Comput.* **2016**, *12*, 281–296.
- [29] P.-P. Kung, P.-J. Sinnema, P. Richardson, M. J. Hickey, K. S. Gajiwala, F. Wang, B. Huang, G. McClellan, J. Wang, K. Maegley, S. Bergqvist, P. P. Mehta, R. Kania, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3557–3562.
- [30] N. G. M. Davies, H. Browne, B. Davis, M. J. Drysdale, N. Foloppe, S. Geoffrey, B. Gibbons, T. Hart, R. Hubbard, M. R. Jensen, H. Mansell, A. Massey, N. Matassova, J. D. Moore, J. Murray, R. Pratt, S. Ray, A. Robertson, S. D. Roughley, J. Schoepfer, K. Scriven, H. Simmonite, S. Stokes, A. Surgenor, P. Webb, M. Wood, L. Wright, P. Brough, *Bioorg. Med. Chem.* **2012**, *20*, 6770–6789.
- [31] J. M. Chen, S. L. Xu, Z. Wawrzak, G. S. Basarab, D. B. Jordan, *Biochemistry* **1998**, *37*, 17735–17744.
- [32] C. R. Smith, D. R. Dougan, M. Komandla, T. Kanouni, B. Knight, J. D. Lawson, M. Sabat, E. R. Taylor, P. Vu, C. Wyrick, *J. Med. Chem.* **2015**, *58*, 5437–5444.
- [33] D. L. Mobley, P. V. Klimovich, *J. Chem. Phys.* **2012**, *137*, 230901.
- [34] J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. Branson, V. S. Pande, *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
- [35] M. Maurer, S. B. A. de Beer, C. Oostenbrink, *Molecules* **2016**, *21*, 499.
- [36] Z. Cournia, B. Allen, W. Sherman, *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.
- [37] J. Michel, J. Tirado-Rives, W. L. Jorgensen, *J. Am. Chem. Soc.* **2009**, *131*, 15403–15411.
- [38] B. Baum, M. Mohamed, M. Zayed, C. Gerlach, A. Heine, D. Hangauer, G. Klebe, *J. Mol. Biol.* **2009**, *390*, 56–69.
- [39] A. Biela, M. Khayat, H. Tan, J. Kong, A. Heine, D. Hangauer, G. Klebe, *J. Mol. Biol.* **2012**, *418*, 350–366.
- [40] M. Betz, T. Wulsdorf, S. G. Krimmer, G. Klebe, *J. Chem. Inf. Model.* **2016**, *56*, 223–233.
- [41] C. N. Nguyen, T. Kurtzman, M. K. Gilson, *J. Chem. Theory Comput.* **2016**, *12*, 414–429.
- [42] Y. Yang, M. A. Lill, *J. Chem. Theory Comput.* **2016**, *12*, 4578–4592.
- [43] I. V. Leontyev, A. A. Stuchebrukhov, *J. Chem. Theory Comput.* **2012**, *8*, 3207–3216.
- [44] D. J. Huggins, *J. Chem. Theory Comput.* **2014**, *10*, 3617–3625.
- [45] L. Wang, R. Abel, R. A. Friesner, B. J. Berne, *J. Chem. Theory Comput.* **2009**, *5*, 1462–1473.
- [46] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, T. E. Cheatham, *Acc. Chem. Res.* **2000**, *33*, 889–897.
- [47] S. Ramsey, C. Nguyen, R. Salomon-Ferrer, R. C. Walker, M. K. Gilson, T. Kurtzman, *J. Comput. Chem.* **2016**, *37*, 2029–2037.
- [48] A. E. Mark, W. F. van Gunsteren, *J. Mol. Biol.* **1994**, *240*, 167–176.
- [49] P. E. Smith, W. F. van Gunsteren, *J. Phys. Chem.* **1994**, *98*, 13735–13740.
- [50] H.-X. Zhou, M. K. Gilson, *Chem. Rev.* **2009**, *109*, 4092–4107.
- [51] C. A. Chang, W. Chen, M. K. Gilson, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1534–1539.
- [52] Schrödinger Suite 2016-3, Protein Preparation Wizard; Epik, Schrödinger LLC, New York, NY, **2016**.
- [53] Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, **2017**.
- [54] “Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters”, K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw in Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (SC’06), ACM, New York, NY (USA), **2006**.
- [55] L. Wang, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1937–1942.
- [56] L. Wang, Y. Deng, J. L. Knight, Y. Wu, B. Kim, W. Sherman, J. C. Shelley, T. Lin, R. Abel, *J. Chem. Theory Comput.* **2013**, *9*, 1282–1293.

Manuscript received: February 13, 2018

Revised manuscript received: April 7, 2018

Accepted manuscript online: May 3, 2018

Version of record online: May 30, 2018

## 2.2 “Assessing the Predictive Power of Relative Binding Free Energy Calculations for Test Cases Involving Displacement of Binding Site Water Molecules”

### 2.2.1 Summary

FEP calculations were applied to a benchmark set of 6 congeneric ligand pairs that involve the displacement and/or rearrangement of binding-site water molecules upon structural modification. The calculated relative binding free energies were compared to the experimental values reported in the literature. For some of the systems, large deviations from the experimental  $\Delta\Delta G$  values were observed, together with a high dependency on the initial solvent structure. These results indicate that the insufficient equilibration of binding site solvent is a main reason for these inconsistencies, due to a slow exchange with the bulk.

A hydration site analysis of all trajectory end-states from the transformations was applied that supports this hypothesis. This was furthermore confirmed by running calculations of control examples from the same ligand series that do not involve a change in binding site solvation. A protocol that adds a GCMC solvent equilibration step before the FEP calculation was tested that leads to better equilibration of the binding site solvent. Yet, this approach still needs improvement.

### 2.2.2 Author contributions

J.W. developed a script to analyze the solvent distributions and thermodynamic properties from FEP calculations. J.W. composed the benchmark set, planned the study and made all calculations, all under the supervision and guidance of M.S.

J.W. wrote the manuscript, M.S. proofread and improved the manuscript.

### 2.2.3 Potential Impact on the Scientific Field

The study points out a shortcoming of common FEP simulation protocols and in detail discusses the problem of insufficient binding site equilibration for buried cavities. State-of-the-art simulation protocols and very long simulation times (50 ns per  $\lambda$ -window, summing up to 600 ns per transformation in total) were applied. The presented benchmark set can serve as a gold standard for FEP protocols that address the issue of binding site solvent equilibration. It is furthermore the first study that tests a GCMC pre-equilibration step in RBFE calculations. Since FEP calculations benefit from increased interest recently, this study should be of great interest to the field.



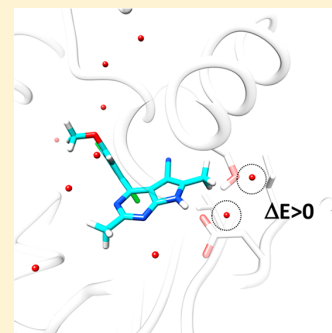


# Assessing the Predictive Power of Relative Binding Free Energy Calculations for Test Cases Involving Displacement of Binding Site Water Molecules

 Joel Wahl<sup>†</sup> and Martin Smieško<sup>\*,†,‡</sup>
<sup>†</sup>Molecular Modeling, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

## Supporting Information

**ABSTRACT:** Improved sampling methodologies, more accurate force fields, and access to longer simulation time scales have led to an increased application of Relative Binding Free Energy (RBFE) calculations in drug discovery projects. In order to assess the strengths and limitations of such tools, adequate benchmark sets are required that challenge the methodology in certain well-defined aspects. We applied Free Energy Perturbation (FEP) calculations to six congeneric ligand pairs taken from the literature, in which addition of a functional group resulted in the displacement of buried binding site water molecules and compared the calculated relative binding free energies with the experimental ones. We started the perturbations from different initial solvation states and registered large inconsistencies (large hysteresis) between the calculated values. We furthermore applied a Grand Canonical Monte Carlo (GCMC) solvent sampling step prior to the FEP calculation that led to a smaller hysteresis for the simulations. By applying a hydration site analysis to the trajectories of the end-states of the perturbation, we could point out that the low accuracy of the predictions as well as the high dependence of the prediction on the chosen initial state is likely caused by the trapping of binding site water molecules and/or insufficient solvation of buried cavities that are formed upon completion of the perturbation. This work highlights that RBFE calculations can suffer from slow solvent exchange of buried parts of the binding sites with the bulk.



## INTRODUCTION

Alchemical Relative Binding Free Energy (RBFE) calculations are increasingly applied in drug discovery processes, facilitated by a variety of improvements in the methodology and successful case studies that highlighted the robustness and usefulness of these methods in hit-to-lead and lead optimization stages.<sup>1–6</sup> Advances include improvements in force fields,<sup>7–9</sup> enhanced sampling methods,<sup>10,11</sup> and accessibility of longer simulation times due to acceleration provided by graphical processing units (GPUs).<sup>12–14</sup>

RBFE calculations usually show the most accurate predictions when they are applied to chemically similar ligands.<sup>6</sup> Therefore, they were so far mainly used for quantifying R-group modifications. Recently developed methodological improvements also facilitate core-hopping studies.<sup>15</sup> The difficulties associated with large changes in ligand structures are related to the problem of the finite sampling of the conformational space,<sup>6</sup> and systems that include protein reorganizations such as the opening of binding cavities were shown to be challenging.<sup>16</sup>

Further sources of error for RBFE calculations are erroneous protonation states of the protein and/or ligand,<sup>17</sup> uncertainties about binding modes<sup>18,19</sup> and binding site solvation.<sup>20</sup> The problem of binding site solvation is related to the slow exchange between bulk and buried binding site solvent molecules and the finite sampling in the simulation time scales that are currently accessible.<sup>1,6,21</sup> Not only can this lead to unrealistic initial

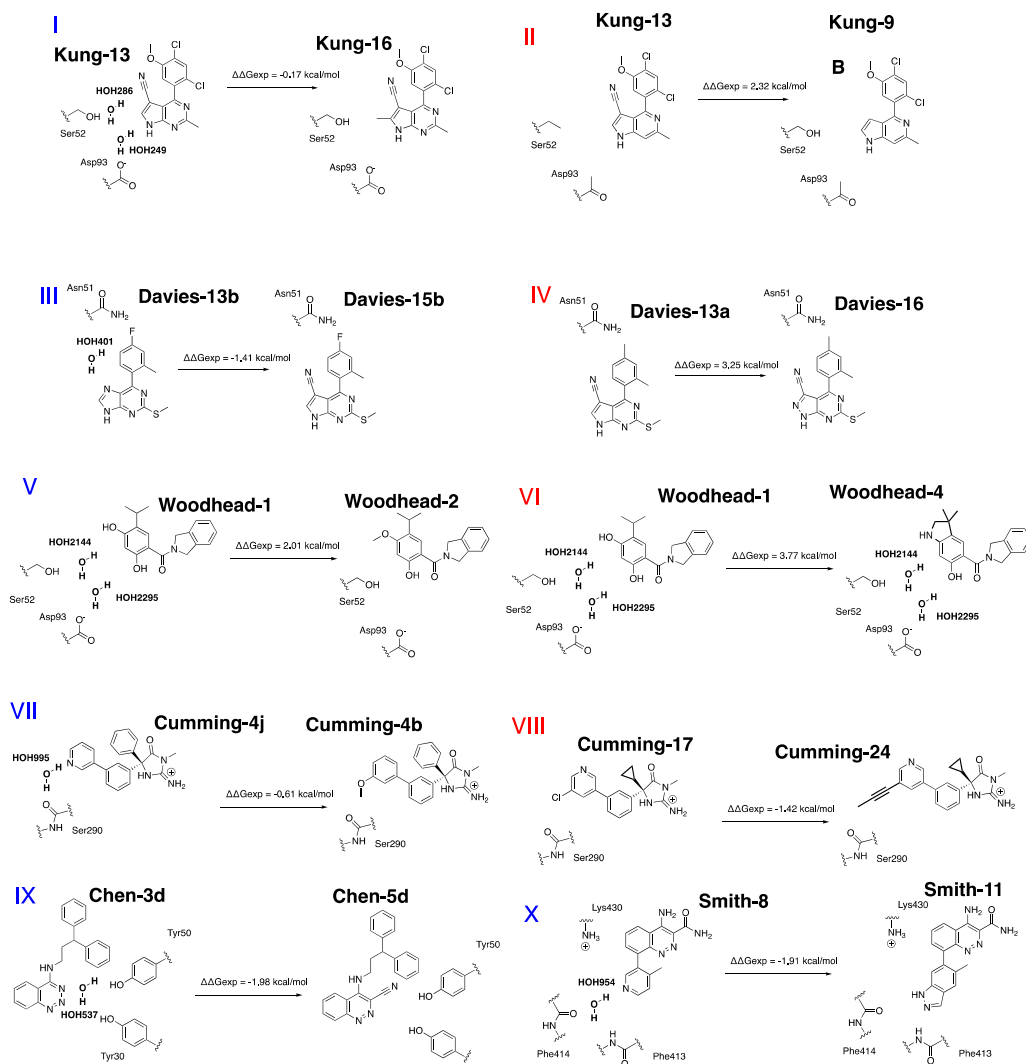
solvation states<sup>20</sup> but can also pose problems for minor ligand modifications where introducing a functional group leads to a displacement or rearrangement of binding site solvent molecules.<sup>22</sup> Michel et al. showed on the basis of three examples that the growth of functional groups that displace buried water molecules leads to a trapping of these water molecules in energetically unfavored states.<sup>22</sup> The accuracy of the predictions was improved when the water molecule was decoupled (double decoupling method) prior to the alchemical ligand transformation. The binding free energy of the displaced water molecule was added as a correction to relative binding free energy. This work highlighted the problem of water trapping in alchemical RBFE calculations. The drawback of the strategy is that prior knowledge about the water displacement is necessary. Furthermore, improper sampling of the phase space could occur due to the simplified two-stepped nature of the simulated thermodynamic cycle.

The problem of insufficient solvent exchange with the bulk was also investigated for the calculation of the absolute binding free energy of camphor to cytochrome P450.<sup>23</sup> The authors employed a GCMC/MD approach to circumvent the problem of trapped water molecules. However, a systematic analysis for

Received: November 19, 2018

Published: January 14, 2019





**Figure 1.** Benchmark set containing six different ligand pairs from four different receptors. Examples I–VI: Hsp90 inhibitors, examples VII–VIII: BACE-1 inhibitors, IX: scytalone dehydratase inhibitors, X: Bruton's tyrosine kinase inhibitors. Examples that involve solvent displacement are colored in blue; examples of the control set that do not involve change in binding site solvation are colored in red.

employing such an approach for the prediction of  $\Delta\Delta G$  values in congeneric ligand series has not been conducted yet.

An alternative to GCMC or the double decoupling of the solvent for RBF calculations is to define four states (Ligand A with the water molecule present, Ligand B with the water molecule absent, Ligand A with the water molecule absent, Ligand B with the water molecule present) and to calculate the free energy differences between all states.<sup>24</sup> The water molecule is thereby converted into a dummy particle.

There remains the question if the recent improvements in enhanced sampling methods and the access to longer simulation time scales can lead to sufficient solvent equilibration. Whereas enhanced sampling methods such as replica exchange with solute tempering (REST)<sup>25</sup> have clear beneficial effects on the accurate simulation of ligand and protein reorganization, their effect on solvent exchange between the binding site and the bulk has not been investigated thoroughly. The REST method used in our study employs a Hamiltonian Replica Exchange

Molecular Dynamics (H-REMD) scheme,<sup>26</sup> which involves scaling of the interactions of the solvent with the solute atoms of the hot region. Furthermore, enhanced sampling employing REST can possibly facilitate the entering or leaving of solvent molecules from the binding site due to larger motions of the solute atoms.

To assess the ability of RBF calculations to yield accurate results for real-world applications, there is a need for benchmark sets that challenge the algorithms in certain clearly defined aspects, with the aim to point out systematic weaknesses.<sup>6</sup> We recently benchmarked the potential of Inhomogeneous Fluid Solvation Theory (IFST) for quantitative predictions of relative binding free energies in congeneric ligand pairs that involve displacement of water molecules from the binding site, using a set of seven well-characterized examples from the literature.<sup>27</sup> The examples were picked carefully and fulfilled the following criteria: (i) high-quality crystal structures available for all ligands or for closely related analogues that confirm the absence/



presence of the water molecules that were targeted for displacement; (ii) availability of thermodynamic data for all ligand pairs with the measurements for the corresponding pairs performed by the same lab; (iii) only minor modifications in ligand structures that led to solvent displacement and/or rearrangement; (iv) displacement and/or rearrangement of solvent occurs in a buried region of the protein and does not cause major conformational changes in the ligand–receptor complex. We identified six ligand pairs in which these criteria are fulfilled. In addition to this, we picked ligand pairs that fulfilled criteria i and ii but do not result in water displacement. These ligand pairs served as control examples, and the corresponding calculations were performed in order to rule out other possible sources of error of the simulation protocol, e.g. protein reorganization, erroneous system setup, or force field inaccuracies.

Due to these characteristics, we regard this set as a valid benchmark set for challenging the ability of state-of-the-art alchemical RBFE calculations to make predictions for congeneric ligand pairs that involve the displacement of binding site water molecules. We applied FEP/REST calculations to these examples and compared the obtained relative binding free energies with the ones from experiment. We furthermore employed hydration site analysis (HSA) to investigate the solvent distribution in the binding site and to identify trapped water molecules. Our results indicate that even state-of-the-art methods with enhanced sampling are still challenged by small-modifications that lead to solvent rearrangements in buried pockets, and even simulation times of 50 ns per  $\lambda$  window do not lead to sufficient solvent equilibration with the bulk for some of the examples. We assess a solvent equilibration protocol using Grand Canonical Monte Carlo (GCMC) and show that it can improve the hysteresis for the systems. This study emphasizes the need for methodological improvements in RBFE calculations and their rigorous validation to extend the applicability and the reliability of the predictions for such cases.

## MATERIALS AND METHODS

**Benchmark Set.** The examples of the benchmark set are shown in Figure 1. For all examples, the ligand on the left corresponds to the smaller ligand with the water molecules present, and the ligand on the right has an added functional group that displaces the water molecule(s).

Examples I–VI correspond to Hsp90 inhibitors. In example I, compound **Kung-13**<sup>28</sup> bears a pyrrolopyrimidine scaffold. The crystal structure of the inhibitor bound to Hsp90 (PDB ID: 3RLQ) reveals the presence of two water molecules. Addition of a methyl-group (compound **Kung-16**) facilitates the displacement of the two water molecules (PDB ID: 3RLR), whereby the affinity roughly remains the same.

Converting the nitrile group of the same inhibitor (**Kung-13**)<sup>28</sup> into a hydrogen (**Kung-9**)<sup>28</sup> in example II leads to a drastic decrease in affinity. This example does not involve a change in binding site solvation and is used as a control example.

In example III, replacing the ring nitrogen of a purine scaffold (**Davies-13b**,<sup>29</sup> PDB ID: 4FCP) with a nitrile group (**Davies-15b**)<sup>29</sup> results in the displacement of a water molecule and a significant gain in affinity.

Example IV is a control example. Changing the cyano-pyrrolopyrimidine scaffold of **Davies-15a**<sup>29</sup> to a cyano-pyrazolopyrimidine (**Davies-16**)<sup>29</sup> results in a significantly higher binding free energy.

Example V involves the conversion of a phenol moiety (**Woodhead-1**,<sup>30</sup> PDB ID: 2XAB) to a methoxyphenyl (**Woodhead-2**, PDB ID: 2XJG). Two water molecules are thereby displaced and the affinity decreased.

Example VI is a control example. Converting the hydroxyl group of **Woodhead-1** from example V<sup>30</sup> to an NH and adding a CH<sub>2</sub> for the closing of a five-membered ring (**Woodhead-4**)<sup>30</sup> does not lead to the displacement of structural binding site water molecules (PDB ID: 2XJJ) but to a lower affinity.

Example VII consists of a pair of BACE-1 inhibitors. A pyridine moiety (**Cumming-4j**,<sup>31</sup> PDB ID: 4DJW) engages in an interaction with a bridging water molecule. Changing the pyridine to a methoxyphenyl (**Cumming-4b**,<sup>31</sup> PDB ID: 4DJV) displaces the water molecule and increases the affinity.

Replacing the chlorine substituent from **Cumming-17h** from the previous series example by a propyne chain (**Cumming-24**)<sup>31</sup> increases the affinity (example VIII). This example is used as a control and does not involve the displacement of a water molecule (PDB ID: 4DJY).

For the inhibitor pair of scytalone dehydratase in example IX, a benztriazine moiety (compound **Chen-3d**)<sup>32</sup> interacts with a water molecule. An added nitrile functionality (**Chen-5d**,<sup>32</sup> PDB ID: 3STD) displaces the water molecule and increases the affinity.

In example X, the pyridine group of inhibitor **Smith-8**<sup>33</sup> (PDB ID: 4ZLZ) is extended to an indazole (**Smith-11**, PDB ID: 4Z3V). A structural water molecule is displaced, and the affinity is increased.

**Structure Preparation.** Preprocessing of all structures was performed with the Protein Preparation Wizard which is part of the Schrödinger Small-Molecule Drug Discovery Suite.<sup>34</sup> The crystallographic water molecules were kept, and all hydrogen atoms were added to the system. Ligand and protein protonation states and tautomers were assigned at a pH of 7.4. The hydrogen bond network was optimized, and a restrained relaxation employing the OPLS\_2005 force field was performed. Missing loops and side chains were added where necessary. For examples I, V, VI, VII, VIII, IX, and XI, crystal structures for both ligands bound to the receptors are available from the PDB. For the remaining ligand pairs, either only one of the ligands had a crystal structure available or only crystal structures of close analogues were available. The structure preparation for these ligand pairs is described in more detail.

For example II, inhibitor **Kung-13** has a crystal structure available (PDB ID: 3RLQ). To generate the structure of **Kung-9**, the nitrile group of **Kung-13** was converted to a hydrogen, the hydrogen bond network was reoptimized, and the restrained relaxation was applied.

In example III, a crystal structure of compound **15a** is available (PDB ID: 4FCQ). The input structure for **Davies-15b** was prepared by changing one of the ring methyl groups into a fluorine atom, followed by restrained minimization. No crystal structure for **Davies-13b** is deposited, but for a purine fragment (PDB ID: 4FCP), indicating that the purine core has the same binding mode as the pyrrolopyrimidine. A water molecule solvating the ring nitrogen of the purine fragment is present. The nitrile group of **Davies-15b** was deleted, and the ring nitrogen was added, together with the water molecule solvating it, to obtain the structure for **Davies-13b**.

A crystal structure for **Davies-15a** in example IV is available (PDB ID: 4FCQ). The input structure for **Davies-15** was obtained by changing the ring C–H group to a nitrogen. This

was followed by a reoptimization of the hydrogen bond network and a restrained minimization.

For **Chen-3d** from example IX, no crystal structure is available. However, there is a crystal structure of a closely related benzotriazine-bearing inhibitor present (PDB ID: SSTD), where the position of the water molecule solvating the ring nitrogen is resolved. To prepare the structure of **Chen-3d**, the nitrile group of **Chen-5d** (which has a crystal structure deposited (PDB ID: 3STD)) was deleted, and the ring nitrogen and the water molecule solvating it were added. The hydrogen bond network was optimized, and a restrained minimization was performed.

**FEP Calculations.** All FEP calculations were conducted with the academic LigandFEP methodology of Desmond,<sup>35,36</sup> using the OPLS\_2005 force field<sup>37</sup> and the TIP4P-Ew water model.<sup>38</sup>

The applied relaxation protocol employed the standard settings for both the bound complex leg and the solvent leg: (i) 100 ps Brownian Dynamics in the NVT ensemble at 10 K and solute heavy atoms restraints (50 kcal/mol/Å<sup>2</sup>), (ii) 12 ps simulation in the NVT ensemble, keeping the restraints and temperature at 10 K, (iii) 12 ps simulation in the NPT ensemble, keeping restraints and temperature at 10 K, (iv) 24 ps simulation in the NPT simulation with solute heavy atom restraints at 300 K, and (v) 240 ps simulation in the NPT ensemble at 300 K without restraints. The temperature and pressure were controlled using a Langevin thermostat (relaxation time 1.0 ps) and a Langevin barostat (relaxation time 2.0 ps). Electrostatic interactions were treated using the particle mesh Ewald calculation. The cutoff for nonbonded interactions was set to 9.0 Å.

The FEP production runs were performed for 50 ns per  $\lambda$ -window for the bound complex leg and 5 ns for the solvent leg. The REST regions included the ligand atoms and the following protein residues in the complex leg: examples I, V, VI: SER52, ASP93; II, III, IV: ASN51, SER52, ASP93; VII, VIII: SER290; IX: TYR30, TYR50; X: PHE413, PHE414, VAL415. In the solvent leg, the ligand atoms define the REST region.

The  $\lambda$ -hopping stage was split into 12 windows. The details of the FEP/REST implementation are given elsewhere.<sup>5,10,11</sup> Enhanced sampling is based on the replica exchange with solute scaling method,<sup>26</sup> which uses the Hamiltonian replica exchange method (H-REMD).<sup>39</sup> For all ligand pairs, the simulations were started from two initial states A and B. State A refers to the protein conformation and binding site solvation with ligand A present, whereas state B is defined analogously for ligand B present. We defined ligand A as the ligand with the binding site water molecules present, which are displaced by a functional group of ligand B.

Error estimates for the predictions were obtained using bootstrapping.<sup>40</sup>

**GCMC.** To obtain a better equilibration of binding-site solvent prior to the  $\lambda$ -hopping stage, the newly available GCMC/MD regime from Desmond was applied.<sup>41</sup> The details of the algorithm have not yet been fully disclosed by the developers, but the solvent equilibration is achieved in two steps: For every  $\lambda$ -window, a GCMC/MD stage is performed with constraints on the solute heavy atoms. The number of water molecules is allowed to vary between the different  $\lambda$  windows in this stage. This is followed by a second GCMC/MD stage without any constraints. The solvent molecules are thereby equalized across the  $\lambda$  windows so that every  $\lambda$  has the same number of particles. The simulation time was 1 ns for both

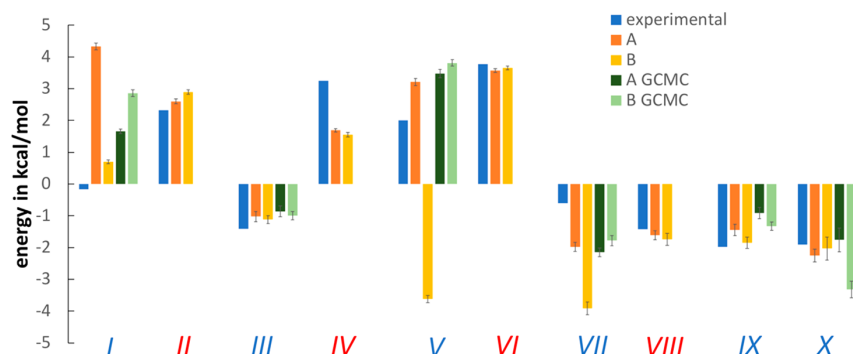
stages. The excess chemical potential of water was set to  $-7.17$  kcal/mol.<sup>42</sup>

**Hydration Site Analysis.** In order to monitor the solvation of the binding site during the FEP calculation, we applied a hydration site analysis (HSA) to the end-state trajectories of the production stage ( $\lambda$ -hopping). Since the  $\lambda$ -hopping stage was split into 12 windows, 12 replicas were run, resulting in 12 trajectories. The trajectories of the replicas that correspond to the two end states ( $\lambda_0$  ( $\lambda = 0$ ) and  $\lambda_{11}$  ( $\lambda = 1$ )) were subject to the HSA. The spatial clustering of the water positions requires them to be in the same frame of reference, which is usually acquired by running the MD simulations with applied constraints to the solute heavy atoms.<sup>43–45</sup> Since FEP calculations run without any constraints, we circumvented this problem by applying a clustering procedure to the trajectory frames.

The whole workflow, including the trajectory processing, trajectory clustering, and HSA, was performed using an in-house script written in Python, making use of the Python API of Schrödinger's Small-Molecule Drug Discovery Suite.<sup>34</sup> The first step consisted of the calculation of the binding site RMSD (root-mean-square deviation) matrix of all frames of the trajectory. The binding site was defined as the ligand heavy atoms and all heavy atoms of residues that have atoms within 5.0 Å of any ligand atom. The RMSD matrix was then used as an input for the clustering with a cutoff criterion of 1.0 Å. The trajectory frame with the highest number of similar frames within an RMSD of 1.0 Å was selected as the center of the first cluster. This procedure was repeated until all frames were assigned to a cluster.

The optimal RMSD cutoff for the clustering is somewhat arbitrary. The number of the clusters should be large enough so that a single cluster does not exhibit too much structural diversity, since this would blur the water density and lead to false negatives, referring to regions with high solvent density that cannot be identified as spherical hydration sites. For all examples, we also visually inspected the trajectories to check if water molecules were occupying regions where no hydration sites were detected. In such cases, the HSA was supplemented by a contour plot of water density in the binding site. A rather small cutoff of 1.0 Å also makes it possible to calculate more reliable estimates for the solvation entropies. If the members of the same cluster exhibit high RMSD values, then the water positions are smeared out which affects the entropy estimations.<sup>46</sup> We therefore regard a cutoff of 1.0 Å as a good compromise, since it still allows the generation of large clusters but at the same time ensures limited structural divergence among the cluster members.

Each cluster was then separately subjected to the HSA, and all frames belonging to the same cluster were aligned by a roto-translational fit to the corresponding cluster center. The size of the hydration site was set to 1.0 Å, and the clustering of the water positions was performed as described elsewhere.<sup>43</sup> The clustering stopped when the occupancy of a hydration site reached less than 0.5 times the number of analyzed frames. For every hydration site, the occupancy, the average water-solute interaction energy ( $E_{sw}$ ), and the average water-water interaction energy ( $E_{ww}$ ) were calculated. The water-water interaction energy involves interactions with all waters of the system. All interaction energies were calculated without cutoffs, employing the minimum image convention. To assess if a hydration site is in an energetically unfavorable state, the energy of solvation was calculated (eq 1)<sup>44,47</sup>



**Figure 2.** Comparison of calculated  $\Delta\Delta G$  values from standard FEP/REST calculations and FEP/REST calculations with GCMC solvent sampling versus experimental  $\Delta\Delta G$  values for simulations starting from initial states A and B. Examples involving solvent displacement are colored in blue. The error bars for the individual simulations are obtained by applying the bootstrapping method.

**Table 1.** Comparison of Experimental and Calculated  $\Delta\Delta G$  Values

	$\Delta\Delta G_A$ (kcal/mol)	$\Delta\Delta G_B$ (kcal/mol)	$\Delta\Delta G_{A\_GCMC}$	$\Delta\Delta G_{B\_GCMC}$	$\Delta\Delta G_{exp}$ (kcal/mol)	hysteresis (kcal/mol)	hysteresis GCMC (kcal/mol)
I	$4.33 \pm 0.11$	$0.70 \pm 0.06$	$1.66 \pm 0.07$	$2.86 \pm 0.11$	$-0.17^a$	3.63	1.20
II	$2.60 \pm 0.08$	$2.89 \pm 0.07$			$2.32^a$	0.29	
III	$-1.02 \pm 0.16$	$-1.12 \pm 0.12$	$-0.86 \pm 0.17$	$-1.00 \pm 0.13$	$-1.41^b$	0.10	0.14
IV	$1.69 \pm 0.05$	$1.55 \pm 0.07$			$3.25^c$	0.14	
V	$3.21 \pm 0.11$	$-3.62 \pm 0.11$	$3.47 \pm 0.13$	$3.81 \pm 0.10$	$2.01^d$	6.83	0.34
VI	$3.57 \pm 0.06$	$3.65 \pm 0.06$			$3.77^d$	0.08	
VII	$-1.98 \pm 0.14$	$-3.91 \pm 0.20$	$-2.15 \pm 0.14$	$-1.78 \pm 0.16$	$-0.61^e$	1.93	0.37
VIII	$-1.61 \pm 0.14$	$-1.74 \pm 0.19$			$-1.42^e$	0.13	
IX	$-1.45 \pm 0.18$	$-1.85 \pm 0.18$	$-0.91 \pm 0.14$	$-1.33 \pm 0.13$	$-1.98^f$	0.40	0.42
X	$-2.25 \pm 0.20$	$-2.03 \pm 0.36$	$-1.76 \pm 0.37$	$-3.32 \pm 0.26$	$-1.91^g$	0.22	1.56

<sup>a</sup>Determined using a competitive assay and a reported accuracy of 10–15%.<sup>28</sup> <sup>b</sup>Determined by ITC.<sup>29</sup> <sup>c</sup>Determined by SPR.<sup>29</sup> <sup>d</sup>Determined by a cell-based assay.<sup>30</sup> <sup>e</sup>Determined by a FRET-peptide substrate hydrolysis assay.<sup>31</sup> <sup>f</sup>Determined by an enzymatic assay.<sup>32</sup> <sup>g</sup>IC<sub>50</sub> obtained by mobility shift assay, converted to affinity by the Cheng-Prusoff equation.<sup>33</sup>

$$\Delta E = E_{sw} + \frac{1}{2}E_{ww} - E_{bulk} \quad (1)$$

where  $E_{bulk}$  is the average interaction energy of a water molecule in bulk, derived from a simulation of pure solvent (−11.412 kcal/mol for the TIP4P-Ew water model). The factor of 0.5 is in accordance with the original formulation of IFST.<sup>44</sup> The solvation entropy was calculated by means of the k-nearest neighbor method<sup>46</sup> using the total distance metric.<sup>48</sup> Only the first-order solute-water entropy was calculated. This leads to the following expression for the solvation free energies of the hydration sites:

$$\Delta G = \Delta E - T\Delta S_{sw} \quad (2)$$

**Simulation Analysis.** Convergence plots for the FEP calculations were generated by the FEPAnalysis tool from Desmond.

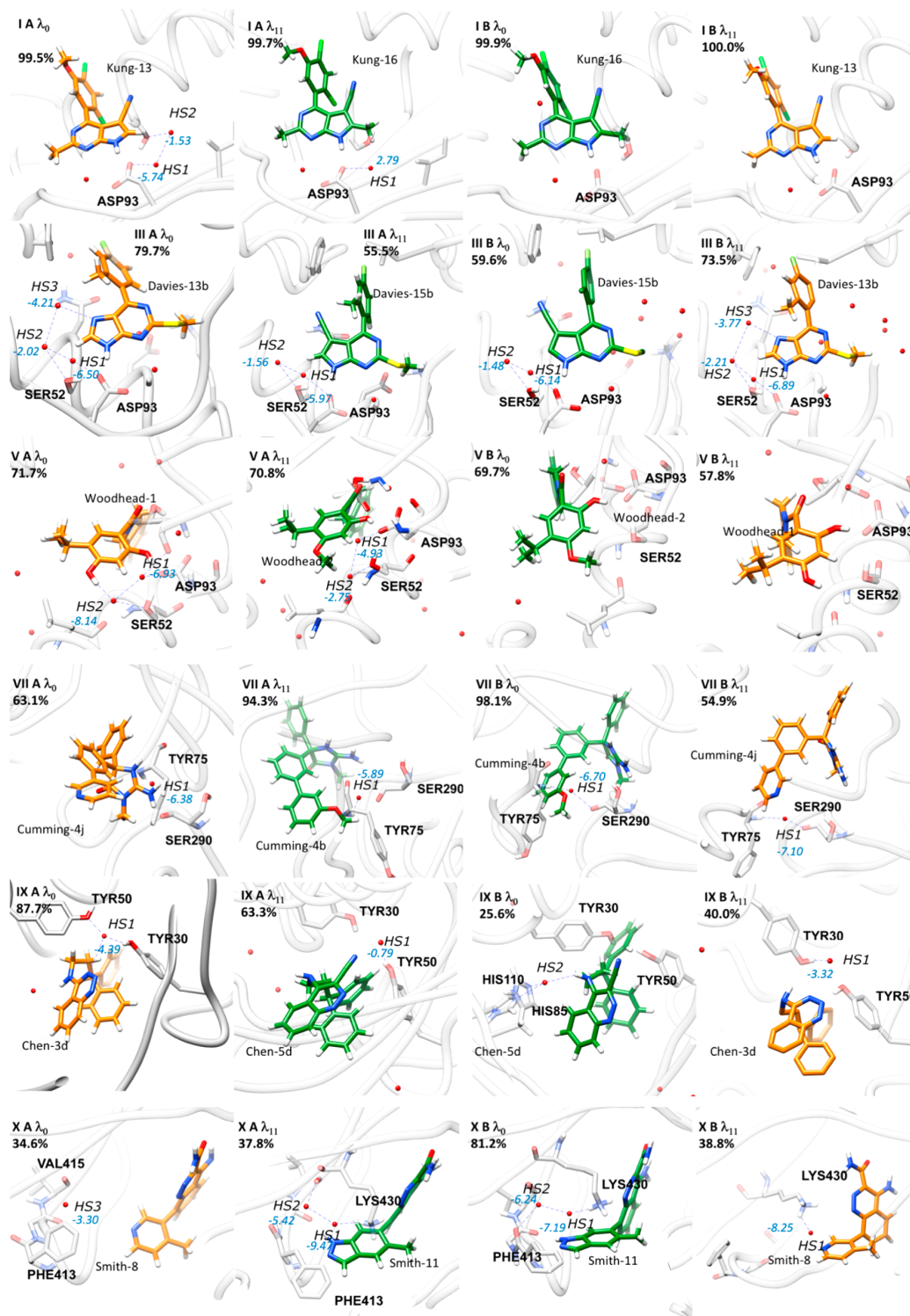
**Graphics.** All molecular graphics were created with the UCSF Chimera package.<sup>49</sup>

## RESULTS AND DISCUSSION

**Accuracy of the FEP/REST Calculations.** The quality of the obtained results can be assessed not only in terms of the accuracy (how much the calculated values deviate from the experimental values) but also in terms of the discrepancy obtained by starting the transformations from two different states.<sup>16</sup> This can be expressed by the so-called hysteresis,<sup>50</sup> which in our examples refers to the absolute difference between  $\Delta\Delta G_A$  and  $\Delta\Delta G_B$ .

Inaccuracies and inconsistency in alchemical RBFE calculations can stem from various sources, such as insufficient sampling in general,<sup>6</sup> force field errors,<sup>5</sup> or uncertainties in system setup,<sup>7</sup> and it is challenging to narrow down the exact cause(s) for the inaccurate results. The hysteresis however is mainly a problem of insufficient sampling.<sup>50</sup> The results of the calculations and a comparison with the experimental values are provided in Figure 2. Six out of the ten examples studied (II, III, VI, VIII, IX, X) showed good agreement (error of less than 1 kcal/mol) with the experimental results as well as small hysteresis (<0.5 kcal/mol) in the standard simulation protocol (no GCMC sampling), whereby all control examples showed high accuracy (except for example IV) and small hysteresis (Table 1). Example IV exhibits a small hysteresis, but  $\Delta\Delta G_A$  and  $\Delta\Delta G_B$  both differ more than 1.5 kcal/mol from the experimental value. Examples I, V, and VII exhibit a hysteresis of larger than 1.9 kcal/mol, indicating a high dependency of the result from the initial state (Table 1).

Since the two initial states for all systems mainly differ in their solvation pattern and not in terms of ligand binding mode and/or protein conformation (as indicated by the experimental crystal structures), we suggest that insufficient equilibration of binding site solvent as the main reason for these discrepancies that are expressed as high hysteresis values. This is further supported by the fact that for all of these three examples with high hysteresis, the control examples did not exhibit a large hysteresis. The control examples refer to ligand pairs from the same system, that have structural modifications at similar atomic positions, but do not involve changes in the solvation pattern.



**Figure 3.** HSA for all examples that involve displacement of binding site solvent molecules. Letters A and B indicate the initial state from which the simulation was started. The given percentage numbers indicate the occupancy of the displayed cluster, obtained from the trajectory clustering. The values in blue are the calculated  $\Delta E$  values of the hydration sites. For clarity, the dummy atoms present in the respective end states are not shown.

The control example for I is IV, the control example for VI is V, and the control example for IX is VII.

In addition, adding a GCMC solvent equilibration stage before the  $\lambda$ -hopping significantly decreased the hysteresis for



examples I, V, and VII yet did not improve the accuracy of the predictions. For example X, adding the GCMC stage led to a higher hysteresis.

**Analysis of Binding Site Solvation for Standard FEP Simulations.** To analyze the binding site solvation and determine examples with trapped water molecules or lacking resolution, we performed HSA for the trajectories of the end states ( $\lambda = 0$  and  $\lambda = 1$ ) of the production runs ( $\lambda$  hopping stage) of all examples that involve a change in the number of binding site water molecules upon ligand modification. The results of the HSA are only reported for the most populated cluster from the trajectory clustering by binding site RMSD for each perturbation. If the most populated cluster has a low occupancy, HSA for additional clusters is provided in the [Supporting Information](#).

In case the binding site solvent achieves full equilibration along the FEP trajectory, the presence and also thermodynamic properties of the hydration sites should be similar for the replicas that correspond to the same bound ligand. Another indication of insufficient equilibration of binding site solvent are  $\Delta E$  or  $\Delta G$  values that are bigger than 0 kcal/mol. This would indicate that it is actually thermodynamically favorable to expel such a water molecule into bulk, but that the diffusion coefficient of the binding site solvent is too slow considering timespan of the trajectory for exchange with the bulk.

The solvation free energy  $\Delta G$  of a hydration site is the actual property of interest that decides if the solvent occupying this specific subregion of the binding site is energetically favorable compared to bulk. However, the obtained  $\Delta G$  values from the IFST procedure described in the [Materials and Methods](#) section have to be assessed with care, primarily due to the way the solvation entropy is treated. We only calculate the first-order entropic term, the solute-water entropy  $\Delta S_{sw}$  and neglect the water–water contributions  $\Delta S_{ww}$ . However, since the true  $\Delta S$  of a hydration site in a protein binding-site is expected to be unfavorable compared to bulk water,<sup>51</sup> hydration sites with positive  $\Delta E$  values will have positive  $\Delta G$  values. For hydration sites with slightly negative  $\Delta E$  values ( $>-2$  kcal/mol), there is still a chance that the overall  $\Delta G$  value is unfavorable, whereas for hydration sites with very low solvation energies ( $<-2$  kcal/mol), the  $\Delta G$  values are expected to be negative, given to the proposed upper boundary of 2 kcal/mol for the entropic cost of water molecules in protein binding sites.<sup>51</sup> Still, for completeness, all thermodynamic properties of the hydration sites ( $\Delta G$ ,  $\Delta E$ ,  $\Delta S$ ) are given in the [Supporting Information \(S3\)](#), while our argumentation is generally based on the  $\Delta E$  values. We would like to point out that the primary goal of the HSA was to compare the solvent distribution of simulations that were started from different initial states and not to make a quantitative assessment of the thermodynamic properties of the binding site solvent.

All examples except for III showed a high dependency on the initial state in terms of the number, exact location, and thermodynamic properties of the identified hydration sites ([Figure 3](#)). In these examples, the initial binding-site solvent pattern in the vicinity of the ligand modification was roughly conserved along the FEP simulation, indicating that solvent exchange with the bulk is not achieved. All identified hydration sites exhibit favorable  $\Delta E$  values except for simulation A of example I, where one hydration site is trapped in an unfavorable energy state with a  $\Delta E$  of 2.79 kcal/mol. In the crystal structure of **Kung-16** bound to Hsp90 (PDB ID: 3RLR), no water molecule can be observed at this hydration site, and also visual

inspection of the electron density (EDS) map does not indicate the presence of this water molecule.

Another example of hydration sites that are conserved in the FEP simulations but absent in the crystal structures occurs in example V, simulation A. In the presence of inhibitor **Woodhead-2** (PDB ID: 2XAB), the two water molecules solvating ASP93 and SER52 are not present in the crystal structure but are conserved in the end-state replica ( $\lambda_{11}$ ) of the  $\lambda$ -hopping and show furthermore favorable solvation energies. They are present in the initial state of the simulation (**Woodhead-1** bound to Hsp90) and are not displaced by the added methyl group.

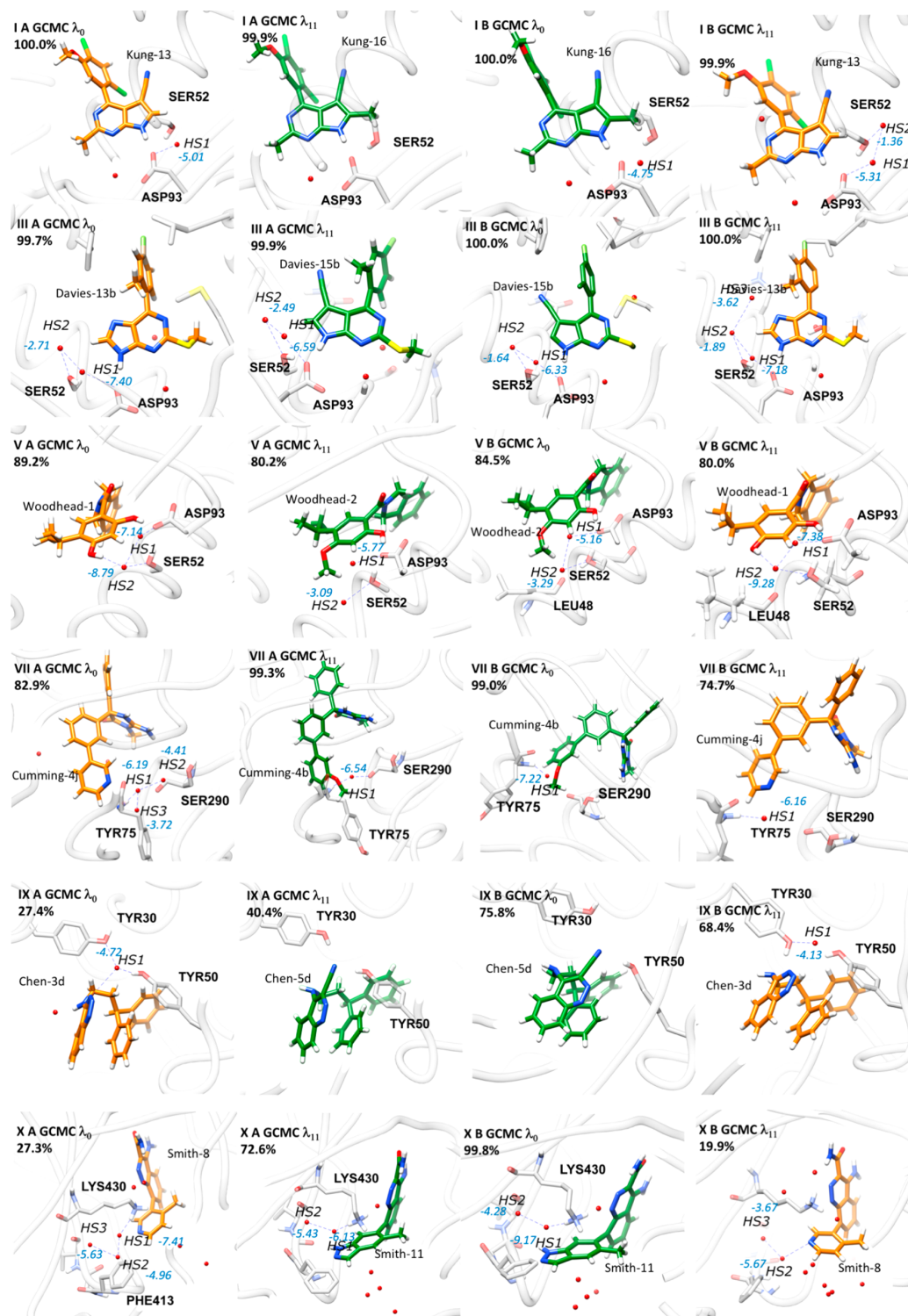
In the end-state replica ( $\lambda_{11}$ ) of simulation B from example IX, the water molecule that is initially solvating the triazine ring of inhibitor Chen-3d is not displaced in the presence of inhibitor Chen-5d. With a  $\Delta E$  value of  $-0.79$  kcal/mol, the actual solvation free energy of this hydration site is probably slightly positive, considering the expected unfavorable entropy of solvation.

The opposite situation occurs in simulation B of examples V and VII. In example V, the inhibitor **Woodhead-2** is in a desolvated state in the end-state replica ( $\lambda_{11}$ ), lacking the two hydration sites observed in the crystal structure (PDB ID: 2XJG). Since the initial state of simulation B corresponds to the crystal structure of **Woodhead-1** bound to Hsp90, these two hydration sites are initially absent and do not get populated in the course of the simulation.

In example VII, the water molecule bridging an interaction between the ring nitrogen of inhibitor **Cumming-4j** and the backbone oxygen of SER290 is neither observed in the  $\lambda_0$ -replica of simulation A nor in the  $\lambda_{11}$ -replica of simulation B, even though the former simulation was started from an initial state with this water molecule present. However, visual inspection of several trajectory frames revealed the presence of a water molecule solvating the ring nitrogen in the  $\lambda_0$  replica of simulation A. Due to the flipping of the pyridyl ring in the course of both simulations, the ligand and the binding site residues exhibit great flexibility, which makes the hydration site analysis challenging in this case.

Therefore, we calculated contour plots of the solvent density on a grid with a resolution of 0.5 Å. A grid cell was considered to be occupied by solvent in a given trajectory frame when it contained a water oxygen atom. The contour plot (S2) reveals a difference in solvent density between the  $\lambda_{11}$ -replica of simulation B and the  $\lambda_0$ -replica of simulation A, with the latter showing increased solvent density around the pyridine nitrogen, indicating better solvation. It is worth mentioning that whereas the cluster centers show distinct conformation of the pyridine rings, all shown clusters contain frames with different ring conformations.

Example X exhibits a complex and fluctuating solvation pattern. For the  $\lambda_0$  replica of simulation A, the two most populated clusters ([Figure 3](#) and [S1](#)) show one hydration site in the subpocket close to the ligand transformation. The third most populated cluster has three hydration sites, which is in agreement with the crystal structure (PDB ID: 4ZLZ). For the  $\lambda_{11}$ -replica replica of simulation A, the two most populated clusters have one and two hydration sites respectively in the vicinity of the indazole ring reported. For simulation B, the  $\lambda_0$ -replica has one predominant cluster with over 80% occupancy and two well-defined hydration sites. The  $\lambda_{11}$ -replica again shows larger binding site flexibility, with three dominant clusters containing one, three, and three hydration sites, respectively



**Figure 4.** HSA for all examples that involve displacement of binding site solvent molecules, obtained from FEP calculations with an additional GCMC step. Letters A and B indicate the initial state from which the simulation was started. The given percentage numbers indicate the occupancy of the displayed cluster, obtained from the trajectory clustering. The values in blue are the calculated  $\Delta E$  values of the hydration sites.

(Figure 3 and S1). Yet, compared to the  $\lambda_0$ -replica, the hydration sites exhibit different thermodynamic properties, indicating that full solvent equilibration is not achieved.

In summary, the HSA revealed several examples where the obtained positions and thermodynamic properties of the binding site water molecules show high dependency on the initial state, indicating insufficient equilibration, in spite of applying enhanced sampling and quite long simulation times (50 ns). The HSA is however difficult to interpret for examples with high flexibility of the binding site and/or ligand.

**Analysis of Binding Site Solvation for FEP Simulations with GCMC Solvent Sampling.** The HSA of the FEP/GCMC simulations of the same examples is shown in Figure 4. Addition of the GCMC stage for solvent sampling leads in general to a better equilibration of the binding site solvent compared to the standard simulation protocol. The problem of desolvated polar ligand groups in the simulations V B and VII B was vastly improved by the GCMC protocol. In the  $\lambda_{11}$ -replica of simulation V B, the two hydration sites solvating the hydroxyl group of inhibitor **Woodhead-1** are present. Importantly, waters in these hydration sites were populated by the GCMC step and did not enter by chance in the course of the FEP simulation. The locations and thermodynamic properties of the hydration sites of both simulations of example V are similar, indicating a low dependency on the initial state, which is in contrast to the results from the standard FEP calculation. In the  $\lambda_0$ -replica of simulation VII A, corresponding to **Cumming-4j** bound to BACE-1, two hydration sites are detected in the vicinity of the pyridyl ring, whereas again the pyridyl seems to be desolvated in the  $\lambda_{11}$ -replica of simulation VII B. However, the solvent density contour plot (S2) shows that the region around the pyridine ring is well solvated. Again, these additional water molecules were added in the GCMC stage.

The discrepancy of the obtained hydration sites in dependence of the initial state of the simulation is also improved by the GCMC protocol in example IX. The triazine nitrogen is solvated in the respective replicas of both simulations ( $\lambda_0$  for IX A GCMC and  $\lambda_{11}$  for IX B GCMC), and no trapping of a water molecule occurs in the  $\lambda_{11}$ -replica of IX A GCMC. Again, this solvent distribution was already achieved after the GCMC step.

The only example where the discrepancy between simulations A and B in terms of binding site solvation still exists is example I. There is both a discrepancy between the number of hydration sites in the presence of inhibitor **Kung-13** (one hydration site detected for I A GCMC  $\lambda_0$  and two hydration sites for I B GCMC  $\lambda_{11}$ ), as well as in the presence of inhibitor **Kung-16** (no hydration site detected for I A GCMC  $\lambda_{11}$  and one hydration sites for I B GCMC  $\lambda_0$ ). In this case, the GCMC equilibration itself showed dependency on the initial state, placing two water molecules in this cavity for **Kung-13** in simulation A but only one water molecule in simulation B. For **Kung-16**, one water molecule was placed in the cavity for simulation A, and no water molecule was placed in simulation B.

In example III, the GCMC presampling only has a minor effect on the obtained solvent distributions. The purine nitrogen seems to be desolvated in the  $\lambda_0$ -replica of simulation A due to the absence of HS3. However, the density plot (S2) reveals that this region actually is solvated. HS3 shows occupancy values of around 0.5 in all simulations of the bound purine fragment **Davies-13b**. In the current case, the occupancy of HS3 is slightly below the occupancy cutoff of 0.5 that was employed in the HSA.

In example X, the GCMC presampling does not result in a better solvent equilibration, and the hysteresis is even higher than for the FEP/REST protocol without the GCMC stage.

The improved solvent sampling prior to the  $\lambda$ -hopping stage leads in general to a smaller dependency on the chosen initial water positions. This is even more of importance in examples where—in contrary to our study—no crystallographic water positions are available.<sup>52</sup> This is of great benefit for the robustness of RBF calculations. In terms of the accuracy of the predictions, avoiding states of water molecules trapped in unfavorable environments and states of desolvated protein/ligand groups due to slow exchange of binding water molecules with the bulk reduces the number of wrong predictions. However, not in all examples did the additional solvent-sampling stage lead to a full independency from the initial state.

Whereas improving the solvent sampling generally lowered the hysteresis of the different simulations, the accuracy of the predicted  $\Delta\Delta G$  values depends on other factors such as the force field.<sup>7</sup> Furthermore, in some of the examples there is a discrepancy between the predicted number and locations of binding site water molecules from the FEP/GCMC protocol and the crystal structure, which can also diminish the predictive power of the calculations in case the water positions from the GCMC sampling are erroneous. In some examples, the standard FEP calculations without GCMC presampling showed differences between the obtained hydration depending on the initial solvation state but still exhibited a small hysteresis and high accuracy (examples IX, X). This is a probably an example of getting the “right answers for the wrong reasons”, due to cancellation of errors.

#### Convergence of Relative Binding Free Energies.

Another way to assess the quality of the simulation results is to check the convergence of the FEP simulations by plotting the calculated  $\Delta\Delta G$  values against the simulation time (S4). Simulations I A, V A, VII A, VII B, and X B are not converged after 50 ns. They all correspond to ligand modifications that involve solvent displacement. Interestingly, whereas simulation V B shows a vast disagreement with experiment in terms of the estimated change in binding free energy, it exhibits good convergence already after 10 ns. For the control examples, II A and B converge after 10 ns. IV A and B show convergence after approximately 20 ns, VI A and VI B show convergence after 30 ns, and VIII A and VIII B show convergence after 40 ns. Therefore, the control examples all converged to a deviation of approximately 0.1 kcal/mol within the 50 ns time frame. The GCMC presampling improved the convergence for simulations V A, VII A, and VII B but not for the remaining examples of the test set.

To assess the influence of longer simulation times on the accuracy and the hysteresis of the calculations, we compared the results obtained after 5 ns and after 50 ns (S5). For the standard FEP/REST simulations, the hysteresis is decreased for all examples, whereas the accuracy is increased in examples III A, IX A, X A, and X B. The hysteresis values for the simulations employing the GCMC equilibration step are less sensitive to prolonging of the simulation time (S5). In terms of the accuracy, examples VII A, VII B, X A, and X B show significant improvement for a sampling time of 50 ns compared to 5 ns.

**Conformational Flexibility in the Binding Site of BACE-1.** BACE-1 inhibitors that are structurally similar to the ones from example VII have been previously studied extensively by means of FEP calculations, in general yielding a good agreement with the experimental values.<sup>53,54</sup> These studies



however did not address the solvent equilibration in the binding site. Whereas we could show that improved solvent equilibration can lead to a smaller hysteresis for the given example, the agreement with the experimental binding affinity remains poor. Keränen et al. proposed that insufficient sampling of the movement of the 10s loop and the closing and opening of the binding site flap can have a detrimental effect on the accuracy of the predictions. We monitored the loop movement and the flap opening/closing for the standard FEP/REST simulations of example VII (S6). Whereas both states were sampled for the flap in all end-state replicas, only the closed state of the 10s loop is sampled, which could be a possible explanation for the low accuracy obtained for that example. One strategy to overcome this is to repeat the FEP calculations by using a different random number seed to generate the initial velocities.<sup>54</sup> It has been furthermore shown that the OPLS3 force field leads to a more realistic representation of structural rearrangements in the S3 loop, compared to older versions of the OPLS force field.<sup>7</sup> We employed the OPLS\_2005 and not the OPLS3 force field in our study, and this can also contribute to the inaccuracy of the prediction of this particular system.

## CONCLUSION

We applied state-of-the-art RBFE calculations toward a benchmarking set consisting of six ligand modifications that led to the displacement of buried water molecules, along with four control examples where ligand modifications left the binding site water network unchanged. Comparison with experimental  $\Delta\Delta G$  values revealed that cases with water displacement are indeed challenging and that in several cases, the predicted  $\Delta\Delta G$  value highly depends on the initial solvation state of the binding site. Analyzing the solvent distribution in the binding site for the two end-states of each trajectory ( $\lambda = 0$  and  $\lambda = 1$ ) of the production phase from all examples by means of HSA revealed the problem of trapped water molecules on one hand and lacking resolution on the other hand. This indicates that current simulation protocols and even long time scales (50 ns in our case) do not lead to sufficient equilibration of binding site solvent molecules in some cases, despite applying enhanced sampling methodologies such as REST.

This equilibration can be improved by adding a GCMC solvent sampling stage before the  $\lambda$ -hopping, as introduced in the new Desmond/GCMC module by Schrödinger.<sup>41</sup>

To detect if insufficient exchange of binding-site solvent with bulk is a problem in real-world examples, where no experimental data on water positions is available, HSA can be a solution to detect hydration sites with unfavorable thermodynamic properties ( $\Delta E > 0$ ). Furthermore, inconsistencies for the obtained  $\Delta\Delta G$  values when conducting the perturbations from different initial states can be an indication for insufficient sampling not only in general<sup>16</sup> but also more specific to insufficient solvent sampling. We therefore recommend to carefully check perturbations that occur in solvated, buried regions of the binding site for trapped water molecules. In such a case, one should question the outcome of the RBFE calculation. Running an MD simulations of the receptor of interest prior to the RBFE calculation and analyzing the mobility of the binding site solvent molecules and their exchange with the bulk can also be employed for assessing if the system is problematic in this regard.

Convergence for the examples that involved changes in the number of binding site solvent molecules was more challenging compared to the control examples. The latter all converged within the 50 ns simulation time. The problem of convergence in

RBFE calculations was previously addressed for ligand modifications that led to large reorganizations in the protein,<sup>16</sup> and our study extends these findings toward changes in binding site solvation for buried pockets. All ligand modifications investigated in our study consisted of addition/deletion of not more than three heavy atoms, and typically such small changes are not expected to raise convergence problems.

A comparison between the high-occupancy clusters centers of the FEP transformations of the test set and the initial structures is provided in the Supporting Information (S7). It confirms that for examples I, III, and V, the mobility of the ligand as well as the binding site is limited. These examples are especially suited for studying the effect of the solvent equilibration, since sampling of ligand and protein should not be of an issue. Examples IX and VII exhibit flexibility of the ligand, whereas example X shows also high flexibility in the binding site, making a proper analysis of binding site solvation patterns challenging.

Solvent sampling using GCMC is an interesting approach to solve the issue of solvent equilibration. The GCMC stage lowered the hysteresis significantly in half of the studied examples. One example already exhibited a small hysteresis in the standard simulation protocols, whereas two examples did not benefit from the solvent equilibration step. Furthermore, it also led to a faster convergence of the simulations in average. However, we think that this approach needs further testing and improvement in the future. In one of the examples, the number of water molecules placed by the GCMC protocol depended on the initial state, and maybe longer sampling times in this GCMC/MD solvent equilibration stage are necessary to fully converge the number of the binding site water molecules.

Therefore, we think that the currently employed protocol in the Desmond GCMC/MD module has room for optimization, in order to give access to converged solvent distributions even for challenging systems.

A more elaborate approach for the incorporation of GCMC sampling into RBFE calculations was recently published by Bruce Macdonald et al.<sup>55</sup> Instead of running the GCMC as a pre-equilibration step, it is directly incorporated into the alchemical free energy calculation. It furthermore allows for the extraction of the energetic contribution of the change in the water network toward the total change in binding free energy.

We hope that our study increases the awareness of the community toward the importance of the correct assessment of binding site solvation in RBFE calculations, and we think that our test set is composed of suitable benchmark systems to validate new approaches and protocols for the prediction of  $\Delta\Delta G$  between ligand pairs or series. The prepared structures and input files for the LigandFEP module from Desmond are available for download (S8).

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00826.

Additional results from hydration site analysis, contour plots of water density, comparison of cluster centers with starting structures of simulations, convergence plots, analysis of structural flexibility of BACE-1, and link to download structures and input files to reproduce results obtained in this study (PDF)



## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: martin.smiesko@unibas.ch.

### ORCID

Martin Smiesko: 0000-0003-2758-2680

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

J.W. is funded by the Schweizerischer Nationalfonds SNF, grant number 200020\_159952.

## REFERENCES

- (1) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911–2937.
- (2) Perez, A.; Morrone, J. A.; Simmerling, C.; Dill, K. A. Advances in Free-Energy-Based Simulations of Protein Folding and Ligand Binding. *Curr. Opin. Struct. Biol.* **2016**, *36*, 25–31.
- (3) Kuhn, B.; Tichý, M.; Wang, L.; Robinson, S.; Martin, R. E.; Kuglstatter, A.; Benz, J.; Giroud, M.; Schirmeister, T.; Abel, R.; Diederich, F.; Hert, J. Prospective Evaluation of Free Energy Calculations for the Prioritization of Cathepsin L Inhibitors. *J. Med. Chem.* **2017**, *60*, 2485–2497.
- (4) Lenselink, E. B.; Louvel, J.; Forti, A. F.; van Veldhoven, J. P. D.; de Vries, H.; Mulder-Krieger, T.; McRobb, F. M.; Negri, A.; Goose, J.; Abel, R.; van Vlijmen, H. W. T.; Wang, L. L.; Harder, E.; Sherman, W.; IJzerman, A. P.; Beuming, T. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. *ACS Omega* **2016**, *1*, 293–304.
- (5) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lypyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, S.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.
- (6) Mobley, D. L.; Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annu. Rev. Biophys.* **2017**, *46*, 531–558.
- (7) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lypyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296.
- (8) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D., Jr. CHARMM General Force Field (CGenFF): A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2009**, *31*, 671–690.
- (9) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (10) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of CDK2 Inhibitors. *J. Chem. Theory Comput.* **2013**, *9*, 1282–1293.
- (11) Wang, L.; Berne, B. J.; Friesner, R. A. On Achieving High Accuracy and Reliability in the Calculation of Relative Protein–Ligand Binding Affinities. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1937–1942.
- (12) Rodrigues, C. I.; Hardy, D. J.; Stone, J. E.; Schulten, K.; Hwu, W.-M. W. GPU Acceleration of Cutoff Pair Potentials for Molecular Modeling Applications. In *Proceedings of the 5th Conference on Computing Frontiers*; CF '08; ACM: New York, NY, USA, 2008; pp 273–282, DOI: 10.1145/1366230.1366277.
- (13) Anderson, J. A.; Lorenz, C. D.; Travesset, A. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *J. Comput. Phys.* **2008**, *227*, 5342–5359.
- (14) Mermelstein, D. J.; Lin, C.; Nelson, G.; Kretsch, R.; McCammon, J. A.; Walker, R. C. Fast and Flexible GPU Accelerated Binding Free Energy Calculations within the Amber Molecular Dynamics Package. *J. Comput. Chem.* **2018**, *39*, 1354–1358.
- (15) Wang, L.; Deng, Y.; Wu, Y.; Kim, B.; LeBard, D. N.; Wandschneider, D.; Beachy, M.; Friesner, R. A.; Abel, R. Accurate Modeling of Scaffold Hopping Transformations in Drug Discovery. *J. Chem. Theory Comput.* **2017**, *13*, 42–54.
- (16) Lim, N. M.; Wang, L.; Abel, R.; Mobley, D. L. Sensitivity in Binding Free Energies Due to Protein Reorganization. *J. Chem. Theory Comput.* **2016**, *12*, 4620–4631.
- (17) Hu, Y.; Sherborne, B.; Lee, T.-S.; Case, D. A.; York, D. M.; Guo, Z. The Importance of Protonation and Tautomerization in Relative Binding Affinity Prediction: A Comparison of AMBER TI and Schrödinger FEP. *J. Comput.-Aided Mol. Des.* **2016**, *30*, S33–S39.
- (18) Michel, J.; Verdonk, M. L.; Essex, J. W. Protein-Ligand Complexes: Computation of the Relative Free Energy of Different Scaffolds and Binding Modes. *J. Chem. Theory Comput.* **2007**, *3*, 1645–1655.
- (19) Matricon, P.; Ranganathan, A.; Warnick, E.; Gao, Z.-G.; Rudling, A.; Lambertucci, C.; Marucci, G.; Ezziati, A.; Jaiteh, M.; Dal Ben, D.; Jacobson, K. A.; Carlsson, J. Fragment Optimization for GPCRs by Molecular Dynamics Free Energy Calculations: Probing Druggable Subpockets of the A2A Adenosine Receptor Binding Site. *Sci. Rep.* **2017**, *7*, 6398–6409.
- (20) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of Water Placement on Predictions of Binding Affinities for P38 $\alpha$  MAP Kinase Inhibitors. *J. Chem. Theory Comput.* **2010**, *6*, 3850–3856.
- (21) Maurer, M.; de Beer, S. B. A.; Oostenbrink, C. Calculation of Relative Binding Free Energy in the Water-Filled Active Site of Oligopeptide-Binding Protein A. *Molecules* **2016**, *21*, 499.
- (22) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc.* **2009**, *131*, 15403–15411.
- (23) Deng, Y.; Roux, B. Computation of Binding Free Energy with Molecular Dynamics and Grand Canonical Monte Carlo Simulations. *J. Chem. Phys.* **2008**, *128*, 115103.
- (24) Maurer, M.; Hansen, N.; Oostenbrink, C. Comparison of Free-Energy Methods Using a Tripeptide-Water Model System. *J. Comput. Chem.* **2018**, *39*, 2226–2242.
- (25) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749–13754.
- (26) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **2011**, *115*, 9431–9438.
- (27) Wahl, J.; Smiesko, M. Thermodynamic Insights into Effects of Water Displacement and Rearrangement upon Ligand Modification Using Molecular Dynamics Simulations. *ChemMedChem* **2018**, *13*, 1325–1335.
- (28) Kung, P.-P.; Sinnema, P.-J.; Richardson, P.; Hickey, M. J.; Gajiwala, K. S.; Wang, F.; Huang, B.; McClellan, G.; Wang, J.; Maegley, K.; Bergqvist, S.; Mehta, P. P.; Kania, R. Design Strategies to Target Crystallographic Waters Applied to the Hsp90 Molecular Chaperone. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3557–3562.
- (29) Davies, N. G. M.; Browne, H.; Davis, B.; Drysdale, M. J.; Foloppe, N.; Geoffrey, S.; Gibbons, B.; Hart, T.; Hubbard, R.; Jensen, M. R.; Mansell, H.; Massey, H.; Matassova, N.; Moore, J. D.; Murray, J.; Pratt, R.; Ray, S.; Robertson, A.; Roughley, S. D.; Schoepfer, J.; Scriven, L.; Simmonite, H.; Stokes, S.; Surgenor, A.; Webb, P.; Wood, M.; Wright, L.; Brough, P. Targeting Conserved Water Molecules: Design of 4-Aryl-5-Cyanopyrrolo[2,3-d]Pyrimidine Hsp90 Inhibitors Using Fragment-Based Screening and Structure-Based Optimization. *Bioorg. Med. Chem.* **2012**, *20*, 6770–6789.

- (30) Woodhead, A. J.; Angove, H.; Carr, M. G.; Chessari, G.; Congreve, M.; Coyle, J. E.; Cosme, J.; Graham, B.; Day, P. J.; Downham, R.; Fazal, L.; Feltell, R.; Figueroa, E.; Frederickson, M.; Lewis, J.; McMenamin, R.; Murray, C. W.; O'Brien, M. A.; Parra, L.; Patel, S.; Phillips, T.; Rees, D. C.; Rich, S.; Smith, D. M.; Trewartha, G.; Vinkovic, M.; Williams, B.; Woolford, A. J. Discovery of (2,4-Dihydroxy-5-Isopropylphenyl)-[5-(4-Methylpiperazin-1-ylmethyl)-1,3-Dihydroisindol-2-yl]Methanone (AT13387), a Novel Inhibitor of the Molecular Chaperone Hsp90 by Fragment Based Drug Design. *J. Med. Chem.* **2010**, *53*, 5956–5969.
- (31) Cumming, J. N.; Smith, E. M.; Wang, L.; Misiaszek, J.; Durkin, J.; Pan, J.; Iserloh, U.; Wu, Y.; Zhu, Z.; Strickland, C.; Voigt, J.; Chen, X.; Kennedy, M. E.; Kuvelkar, R.; Hyde, L. A.; Cox, K.; Favreau, L.; Czarniecki, M. F.; Greenlee, W. J.; McKittrick, B. A.; Parker, E. M.; Stamford, A. W. Structure Based Design of Iminohydantoin BACE1 Inhibitors: Identification of an Orally Available, Centrally Active BACE1 Inhibitor. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444–2449.
- (32) Chen, J. M.; Xu, S. L.; Wawrzak, Z.; Basarab, G. S.; Jordan, D. B. Structure-Based Design of Potent Inhibitors of Scytalone Dehydratase: Displacement of a Water Molecule from the Active Site. *Biochemistry* **1998**, *37*, 17735–17744.
- (33) Smith, C. R.; Dougan, D. R.; Komandla, M.; Kanouni, T.; Knight, B.; Lawson, J. D.; Sabat, M.; Taylor, E. R.; Vu, P.; Wyrick, C. Fragment-Based Discovery of a Small Molecule Inhibitor of Bruton's Tyrosine Kinase. *J. Med. Chem.* **2015**, *58*, 5437–5444.
- (34) *Small-Molecule Drug Discovery Suite 2017-4*; Schrödinger, LLC: New York, NY, 2017.
- (35) *Desmond Molecular Dynamics System*; D. E. Shaw Research: New York, NY, 2017.
- (36) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*; SC '06; ACM: New York, NY, USA, 2006; DOI: 10.1145/1188455.1188544.
- (37) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R.; Philipp, D. M.; Repasky, M. P.; Zhang, L. Y.; Berne, B. J.; Friesner, R. A.; Gallicchio, E.; Levy, R. M. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26*, 1752–1780.
- (38) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- (39) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction. *J. Chem. Phys.* **2002**, *116*, 9058–9067.
- (40) Paliwal, H.; Shirts, M. R. A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J. Chem. Theory Comput.* **2011**, *7*, 4115–4134.
- (41) <https://www.schrodinger.com/news/announcing-schrodinger-software-release-2018-2> (accessed July 07, 2018).
- (42) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. Contribution of Explicit Solvent Effects to the Binding Affinity of Small-Molecule Inhibitors in Blood Coagulation Factor Serine Proteases. *ChemMedChem* **2011**, *6*, 1049–1066.
- (43) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.
- (44) Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, *102*, 3531–3541.
- (45) Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. *J. Chem. Theory Comput.* **2014**, *10*, 2769–2780.
- (46) Huggins, D. J. Quantifying the Entropy of Binding for Water Molecules in Protein Cavities by Computing Correlations. *Biophys. J.* **2015**, *108*, 928–936.
- (47) Nguyen, C. N.; Young, T. K.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor Cucurbit[7]Uril. *J. Chem. Phys.* **2012**, *137*, 044101.
- (48) Huggins, D. J. Estimating Translational and Orientational Entropies Using the K-Nearest Neighbors Algorithm. *J. Chem. Theory Comput.* **2014**, *10*, 3617–3625.
- (49) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (50) Wood, R. H.; Muhlbauer, W. C. F.; Thompson, P. T. Systematic Errors in Free Energy Perturbation Calculations Due to a Finite Sample of Configuration Space: Sample-Size Hysteresis. *J. Phys. Chem.* **1991**, *95*, 6670–6675.
- (51) Dunitz, J. D. The Entropic Cost of Bound Water in Crystals and Biomolecules. *Science* **1994**, *264*, 670–6170.
- (52) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of Water Placement on Predictions of Binding Affinities for P38 $\alpha$  MAP Kinase Inhibitors. *J. Chem. Theory Comput.* **2010**, *6*, 3850–3856.
- (53) Ciordia, M.; Pérez-Benito, L.; Delgado, F.; Trabanco, A. A.; Tresadern, G. Application of Free Energy Perturbation for the Design of BACE1 Inhibitors. *J. Chem. Inf. Model.* **2016**, *56*, 1856–1871.
- (54) Keränen, H.; Pérez-Benito, L.; Ciordia, M.; Delgado, F.; Steinbrecher, T. B.; Oehlrich, D.; van Vlijmen, H. W. T.; Trabanco, A. A.; Tresadern, G. Acylguanidine Beta Secretase 1 Inhibitors: A Combined Experimental and Free Energy Perturbation Study. *J. Chem. Theory Comput.* **2017**, *13*, 1439–1453.
- (55) Bruce Macdonald, H. E.; Cave-Ayland, C.; Ross, G. A.; Essex, J. W. Ligand Binding Free Energies with Adaptive Water Networks: Two-Dimensional Grand Canonical Alchemical Perturbations. *J. Chem. Theory Comput.* **2018**, *14*, 6586–6597.

## 2.3 “Incorporating Protein and Ligand Desolvation Effects into a Knowledge-Based Scoring Function”

### 2.3.1 Summary

A scoring function was developed that accounts for protein and ligand desolvation terms, combining water interaction energies that are discretized on a grid together with an approach developed by another group that facilitates fast solvation of ligand polar groups. This newly developed desolvation score was combined with terms from the OPLS\_2005 force field (van der Waals, ligand strain) and the YETI force field (hydrogen bond, ligand-metal interactions). The coefficients for the various contributions were optimized to reproduce experimental binding affinities from the PDBbind refined set and validated against the core subset of the PDBbind. Furthermore, the ability of the scoring function to discriminate correct from wrong docking poses was evaluated by using the docking power test from the Comparative Assessment of Scoring Functions (CASF) benchmark.

### 2.3.2 Author contributions

J.W. designed, implemented and validated the scoring function, all under supervision and guidance of M.S.

J.W. wrote the manuscript.

### 2.3.3 Potential Impact on the Scientific Field

Accurate description of desolvation terms is a necessity for correct pose prediction and affinity prediction in molecular docking. The developed scoring function can serve as a basis for further improvements and it already shows encouraging results in its current status in terms of scoring power.

# **Incorporating Protein and Ligand Desolvation Effects into an Empirical Scoring Function**

*Joel Wahl, Martin Smieško\**

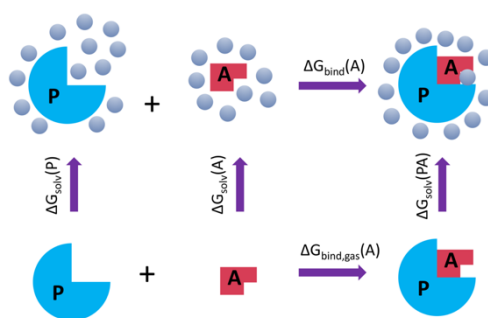
Molecular Modeling, Department of Pharmaceutical Sciences, University of Basel

Klingelbergstrasse 50, Basel CH-4056 Switzerland

## **Introduction**

Algorithms in molecular docking approaches typically rely on a scoring function that maps the three-dimensional (3D) coordinates of a protein-ligand complex to numerical value that serves as an estimate of the associated binding affinity <sup>1-3</sup>. Ideally, a scoring function is able to predict the correct binding pose and the binding free energy. A binding pose is defined as the conformation, position and orientation of the ligand in the binding site. Moreover, in order to be of use in Virtual Screening (VS) applications, a scoring function should be able to correctly rank several binders towards a given receptor, as well as enrich chemical libraries by giving better ranks to known actives compared to a set of decoys <sup>4-6</sup>.

The performance of scoring functions can be measured in terms of scoring power (affinity prediction), docking power (pose prediction), ranking power (ranking ligands to same receptor) and screening power (discriminating actives from decoys) <sup>7</sup>. A rigorous benchmark of several widely-used scoring functions in terms of the above measures revealed primarily weaknesses for scoring and ranking <sup>7</sup>. The authors concluded that the major challenges in molecular docking lie in desolvation effects, polar interactions as well as non-additive contributions <sup>7</sup>. Whereas efforts in all aspects are needed to improve the accuracy of scoring functions, we in the following focus on the treatment of desolvation effects.



**Figure 1.** Thermodynamic cycle of ligand-protein binding.

A schematic depiction of the ligand-protein binding process is given in **Figure 1**. This process can be split into four steps: receptor desolvation, ligand desolvation, complex formation and complex resolution. Accurate estimation of the desolvation costs for the ligand and receptor binding site is therefore expected to be essential in order to obtain valid predictions <sup>8</sup>.

Different types of scoring functions exist (physics-based, knowledge-based, empirical and machine learning) <sup>9</sup>, and this leads to different ways on how solvation terms are incorporated. Physics-based (or force-field based) scoring functions rely on rigorous theoretical frameworks, and desolvation effects are usually incorporated by a solvation model such as Generalized-Born or Poisson-Boltzmann <sup>10,11</sup>. In empirical scoring functions, the functional form is less restricted. Examples for desolvation terms in empirical scoring functions are the fractional desolvation based on Born radii in Dock <sup>12</sup>, the atom-based solvation parameters in combination with an assessment of the desolvation degree in AutoDock <sup>13</sup>, a term based on incremental logP values of atoms types combined with the surface accessible area in HYDE <sup>14</sup>, or the docking of explicit water molecules into the binding site to assess solvent accessibility in Glide <sup>15</sup>.

These methods all capture the general effects of solvation and desolvation. However, there are also specific water molecules in binding sites, that have well-defined positions and are important for the molecular recognition at the binding site <sup>16,17</sup>.

Incorporating such water molecules into docking was shown to improve the accuracy of the pose prediction<sup>18,19</sup>. Furthermore, such waters can be displaced by some ligands while being conserved for others<sup>20</sup>, which requires a flexible treatment in docking algorithms.

Such dynamical treatment of water molecules is incorporated in GOLD<sup>21</sup>, where water molecules are allowed to switch on and off during pose generation. Furthermore, water orientations are optimized on the fly. Displacement of such water molecules is rewarded by a constant entropic penalty. Another philosophy in the way binding site water molecules are treated is the displaceable water force field in AutoDock<sup>22</sup>, where the ligand is docked in a presolvated state and the water sites that are associated with the ligand can be either conserved or displaced. An energetic and entropic score is calculated for every water site.

The most recent advances for the treatment of water molecules in molecular docking are methods that incorporate information about the distribution and thermodynamic properties of the solvent obtained from molecular simulations<sup>23–25</sup>. WScore uses locations and thermodynamic quantities from hydration sites obtained by the WaterMap approach<sup>26</sup>. For every docking pose, it is assessed if the hydration site is displaced by the ligand or conserved. Corresponding energetic and entropic penalties or rewards are thereby assigned. AutoDock-GIST<sup>24</sup> and a closely-related approach<sup>25</sup> use grid-based inhomogeneous fluid solvation theory (GIST)<sup>27</sup>, where the binding site water properties are discretized on a grid. The underlying idea is that binding poses in which ligands overlap with regions with favorable solvation free energies receive an energetic penalty for the receptor desolvation. This approach was tested in a prospective application for compound prioritization and showed a promising hit rate<sup>25</sup>.

The disadvantage of such approaches is that an MD simulation needs to be set up for every receptor structure that is used as a query in a docking workflow. This requires the use of additional software and reduces the easiness of how the calculations

can be set up. Furthermore, most docking programs prioritize a simple functional form for the scoring function that facilitates fast evaluation of various contributions to the estimated binding free energies. There is the possibility that treating the protein desolvation with a rigorous theoretical framework and other terms at a lower level of theory leads to an asymmetry in the various contributions.

In this work we aimed to develop a simple, intuitive and computationally affordable scoring approach that assesses protein and ligand solvation/desolvation by combining water-affinity grids in the binding site with explicit water sites that solvate the polar groups of the ligand <sup>28</sup>. Ligand hydration sites that lie in regions with high affinity for water molecules are kept, whereas others are regarded as displaced. Protein desolvation is assessed by evaluating the number of grid cells with high affinity for solvent that is displaced by the ligand heavy atoms. Finally, an empirical scoring function containing atom type-based desolvation scores together with other terms (protein-ligand interaction energy, hydrogen bond-score and ligand strain) is trained towards the PDBBind <sup>29</sup>. We term our scoring function DESSCORE, emphasizing the explicit inclusion of protein and ligand desolvation. We also evaluated the influence of the desolvation term by training a scoring function that doesn't account for desolvation, which we term as NODESSCORE. Finally, also a naïve scoring based on van der Waals interactions and a lipophilic score is trained and evaluated as a baseline model.

## **Materials and Methods**

### *Desolvation Term*

The simple underlying idea behind our desolvation score is to assess the accessibility of receptor and ligand atoms for solvent, both in the unbound and in the bound state. The degree of solvation for the receptor polar atoms is evaluated by sampling a water probe on a grid that spans the binding site. For every grid cell, 50 random orientations and positions of a SPC <sup>30</sup> water molecule are generated and the respective interaction energies with the protein are calculated on pregenerated grids of

Coulomb and Lennard-Jones potentials<sup>31</sup>. The lowest interaction energy out of these 50 water positions/orientations is then assigned to the corresponding grid cell ( $E_{\min}^{\text{cell}}(\text{SPC})$ ). For every protein oxygen and nitrogen atom, the number of grid cells within a specific radius (3.5 Å in our case) that possess an  $E_{\min}^{\text{cell}}(\text{SPC})$  value below a cutoff of -5.0 kcal/mol are calculated ( $n_{\text{cell},i}$ ). This gives an estimation of the initial degree of solvation for these atoms and we term this as  $SA_{\text{unbound},i}$ , the solvent accessibility of atom  $i$  in the unbound state. For the bound state, grid cells that are within a radius of 2.8 Å from any ligand atom heavy atom are regarded as being displaced.  $SA_{\text{bound},i}$ , the solvent accessibility of atom  $i$  in the bound state, is obtained by subtracting the displaced grid cells from  $n_{\text{cell},i}$ .

For the ligand oxygen and nitrogen atoms, hydration sites are generated using WaterDock2.0<sup>28</sup>. WaterDock2.0 generates hydration sites around the ligand by using solvent distributions around different atom types obtained by explicit solvent molecular dynamics simulations. The solvent accessibility of ligand atom  $i$  in the unbound state corresponds to the number of hydration sites within a radius of 3.5 Å of the ligand. Subtracting the number of hydration sites that are displaced by the protein heavy atoms in the docked pose yields the solvent accessibility for the bound state.

In the current version of the scoring function, only nitrogen and oxygen atoms of the binding site and the ligand contribute to the desolvation penalty. The desolvation penalty of atom  $i$  is defined as (**Equation 1**):

$$E_{\text{desolv},i} = |q_i| * D_i * b_i \quad (1)$$

Where  $q_i$  is the atomic partial charge on atom  $i$ ,  $D_i$  is the degree of desolvation of atom  $i$  and  $b_i$  is the buriedness. The degree of desolvation is given by the following relationship (**Equation 2** and **Equation 3**):

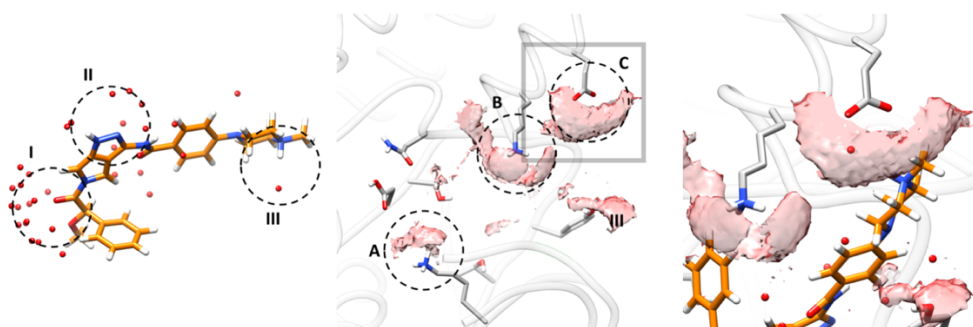
$$D_i = \frac{SA_{\text{unbound},i}}{SA_{\text{max},i}} \frac{SA_{\text{unbound},i} - SA_{\text{bound},i}}{SA_{\text{unbound},i}} \quad (2)$$

$$= \frac{SA_{\text{unbound},i} - SA_{\text{bound},i}}{SA_{\text{max},i}} \quad (3)$$



$SA_{\max}$  is the maximal solvent accessibility observed for this atom type in the set of protein-ligand complexes that was used to train the scoring function.

An illustration of the underlying principle is given in **Figure 2**. The ligand amide oxygen (I) and the ring nitrogen atoms (II) are well solvated, whereas the tertiary amine nitrogen (III) only possesses one hydration site because of the low solvent accessibility (**Figure 2** left). The contour plot of the water probe interaction energies at the receptor binding site (**Figure 2** middle) indicate that two charged amino groups of the lysine (A and B) and the carboxylate of the glutamate residue are accessible for solvent and are expected to be well hydrated in the unbound state. Upon ligand binding (**Figure 2** right), heavy atoms of the bound ligand overlap with the region C around the carboxylate, causing a desolvation penalty.



**Figure 2.** Generated hydration sites for ligand polar atoms (left), contour plot of water probe interaction energy ( $-10$  kcal/mol) at the receptor binding site (middle) and combined ligand and protein hydration for the binding pose (right).

The buriedness was added to **Equation 1** since fully solvent-exposed ligand atoms in the bound state should not receive a desolvation penalty. If a polar ligand atom is fully solvated in the unbound state and located at the surface of the protein in the bound state, it overlaps with a region in the binding site where no strong solute-solvent interactions are present, corresponding to values close to zero in the water grid. In this case, the

ligand atom would receive a desolvation penalty. To prevent this, atoms with low buriedness values don't receive high desolvation penalties.

The buriedness is adapted from the concept of the receptor density<sup>32</sup>. The receptor density of an atom is defined as the number of receptor heavy atoms within a radius of 8.0 Å. Atoms with a receptor density of less than 30 have a buriedness score of 0, whereas atoms with a receptor density of higher than 120 receive a buriedness score of 1. In between, the buriedness score is scaled linearly.

### Scoring Function

The DESSCORE function for the estimation of binding affinities consists of a Coulomb term, a van der Waals term, a directional hydrogen bond term, a term for ligand strain, a metal term and the desolvation terms for oxygen and nitrogen ligand and receptor atoms (**Equation 3** and **Equation 4**):

$$\Delta G = E_{inter} + C_{lipo}E_{lipo} + C_{strain}E_{strain} + C_{des,N} \sum_{i \in atom, ele N, O} |q_i| D_i \quad (3)$$

$$\begin{aligned} E_{inter} = & C_{coul} \sum_{nbond\ pairs} \left( \frac{q_i q_j}{4\pi r_{ij}^2} \right) + C_{vdw} \sum_{nbond\ pairs} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \\ & + C_{hbond} \sum_{hbond} \left[ \frac{C}{r_{ij}^{12}} - \frac{D}{r_{ij}^{10}} \right] \cos^2(\theta_{D-H\cdots A} - \theta_0) \cos^2(\omega_{H\cdots A-AA} - \omega_0) \\ & + C_{ML-cov} \sum_{ligand-metal} \left[ \frac{E}{r_{ij}^{12}} - \frac{F}{r_{ij}^{10}} \right] \quad (4) \end{aligned}$$

Partial charges and Lennard-Jones parameters are taken from the OPLS\_2005 force field<sup>33</sup>. The ligand strain is estimated by optimizing the bound conformation of the ligand and taking the difference in the energy. The lipophilic term was

directly adapted from the ChemScore function<sup>34</sup>. The hydrogen bond term is taken from the YETI force field<sup>35,36</sup>, accounting for the linearity (angle between the donor heavy atom, donor hydrogen atom and the acceptor atom) and for the directionality (angle between the donor hydrogen, acceptor atom and the neighboring atom of the acceptor) by means of two cosine functions (**Equation 4**)<sup>37</sup>. The hydrogen bond term penalizes deviations from optimal linearity ( $\theta_0$ ) and directionality ( $\omega_0$ ). Metal-ligand interactions are included in the Coulombic term, but are also treated by a special term based on a 10-12 potential, which we term as the covalent metal term. The functional form of this term was adapted from the YETI force field<sup>35</sup>. Therefore, both the electrostatic and covalent character of bonds to metals are present in the scoring function<sup>38</sup>.

Both the hydrogen bond term and the metal term of the YETI force field are very sensitive to the distance between donor and acceptor or metal and ligand atom respectively, since the constants C, D, E and F all depend on the well-depth  $E_{\min}$  and the equilibrium distance  $r_0$ . Minimizations performed with the OPLS\_2005 force field can lead to distances that are shorter than the equilibrium distances of these YETI terms. Therefore, the repulsive parts of the hydrogen bond and metal terms were cut, and the maximum contribution was given for distances smaller than  $r_0$ .

### *Preparation of the Training and Test Sets*

The PDBbind refined set v.2016<sup>29</sup> was used as the training set for optimizing the coefficients of the DESSCORE function. All complexes from the training set were subjected to the Protein Preparation Wizard<sup>39</sup> and the protein and ligand protonation states were assigned at a pH of 7.4. This was followed by an optimization of the hydrogen bond network. Since our scoring function considers the hydrogen atoms explicitly and has a directional hydrogen bond term, proper assignment and optimization of the hydrogen positions is a necessary step.

OPLS\_2005 force field parameters were assigned to the protein and the ligand atoms using the system builder from the Desmond simulation package<sup>40</sup>. All complexes that contained ligands which could not be identified by their three-letter code were omitted, since this was a necessary condition in our workflow. Furthermore, all structures that are also part of the PDBbind core set were removed, since they served as the test set for the validation.

### *Training of the Scoring Function*

Various scoring functions were trained on the PDBbind refined set or subsets of it. The coefficients of the X-Score function were optimized using the PDBbind core set<sup>29,41</sup>. The input structures were directly taken as the experimental crystal structures and the coefficients were fitted by multilinear regression (MLR) to the associated experimental binding affinities. AutoDock Vina is another example of a scoring function that was trained using the PDBbind set, yet the details on how the coefficients were optimized are not provided<sup>42</sup>. Other examples include the Vinardo scoring function<sup>43</sup> and different implementations of scoring functions using machine learning<sup>44,45</sup>.

The approach of using linear regression and experimental binding affinities to optimize the coefficients of the various terms results in a scoring function that is primarily trained for scoring power<sup>43</sup>. This implies that upon receiving a valid protein-ligand complex as an input, the scoring function is able to give a good estimate of the associated binding affinity. This task is fundamentally different from the task of picking the correct binding pose from a pool of different poses, which is termed as the docking power.

An interesting showcase is the correlation between the ligand buried solvent accessible surface area upon binding ( $\Delta$ SAS), which shows good correlation with the binding affinities in the PDBbind core set<sup>7</sup>. Since the complexes in the PDBbind all contain real binders, all ligand poses are expected to be fairly complementary to the

binding site, meaning that polar atoms of the protein and the ligand interact via hydrogen bonds or electrostatic interactions, whereas non-polar atoms are forming hydrophobic contacts. Since the desolvation costs of the polar protein and ligand atoms are then compensated by the favorable electrostatic interactions, the shape complementarity can give a good estimate about the binding affinity. However, the  $\Delta$ SAS shows a very bad performance in the docking power test <sup>7</sup>.

On the other hand, if a scoring function is composed of terms that accurately capture the various contributions towards the thermodynamics of protein-ligand binding, then the ability of distinguishing correct from incorrect binding poses should be learned implicitly, even when the scoring function is just trained towards reproducing binding affinities. To verify this, we trained the DESSCORE function for scoring power and evaluated the performance for the docking power in a second step.

In our case, the coefficients of the scoring function were optimized by multiple linear regression against a training set of 2540 protein-ligand complexes from the PDBbind.

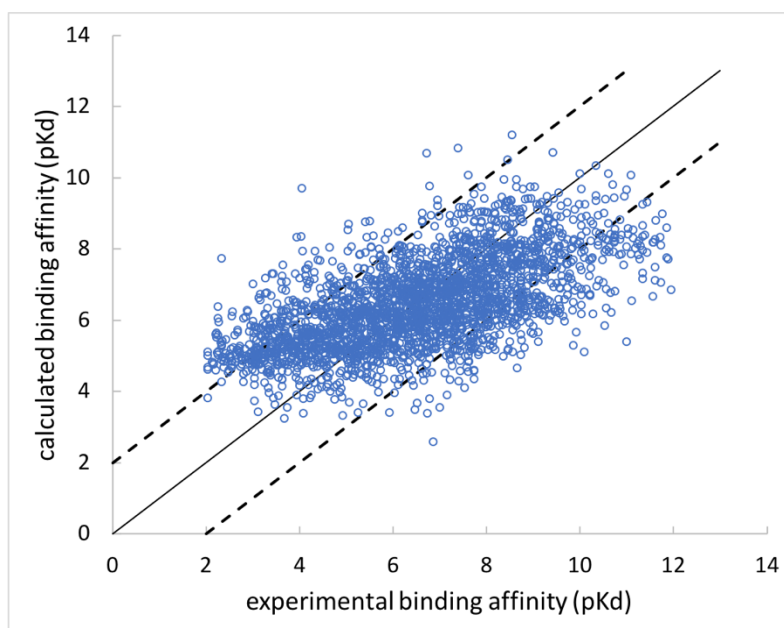
### *Training*

The workflows for the training and the evaluation of the DESSCORE function were implemented in Python 2.7 using the Python API of the Schrödinger Small-Molecule Drug Discovery Suite <sup>39</sup>

## **Results and Discussion**

### *Training*

The correlation plot between experimental and estimated binding affinities by our DESSCORE function for the training set is shown in **Figure 3**.



**Figure 3.** Correlation between experimental binding affinities and binding affinities predicted by DESSCORE for the training set ( $pK_d = -\log K_d$ ). The dashed lines mark the boundaries where the predictions are within an absolute error of 2.0 kcal/mol.

The model reached a pearson correlation coefficient of  $r = 0.60$  while 76.9 % of all compounds were predicted within an accuracy of 2  $\log(K_d)$  units. With an intercept from the MLR of +3.54, DESRES shows the same weakness as most scoring functions, namely an overestimation of the binding affinity for weak binders and an overestimation for strong binders.<sup>7,41</sup> The coefficients obtained by the MLR are shown in **Table 1**.

**Table 1.** Weights for the individual scoring terms obtained from MLR towards the training set. The DESSCORE model includes the desolvation term, whereas NODESSCORE was fitted without taking desolvation explicitly into account.

	DESSCORE	NODESSCORE	LipoVdwSCORE
$C_{\text{coul}}$	-2.74E-3	-1.33E-3	0.00
$C_{\text{lipo}}$	6.53E-3	7.81E-3	9.41E-3
$C_{\text{vdW}}$	-3.41E-2	-2.26E-2	-5.09E-3
$C_{\text{hbond}}$	-2.07E-2	-2.82E-3	0.00

<b>C<sub>strain</sub></b>	-6.12E-2	-6.99E-3	0.00
<b>C<sub>desolv</sub></b>	-2.77E-1	0.00	0.00
<b>C<sub>metal</sub></b>	-5.11E-2	-4.63E-2	0.00
<b>Pearson's r</b>	0.60 [0.57, 0.62]	0.60 [0.57, 0.62]	0.55 [0.52, 0.58]
<b>SD</b>	1.63	1.64	1.71
<b>MedAEP</b>	1.10	1.12	1.22
<b>MedAEP<sub>cv</sub></b>	1.12	1.13	1.22

The robustness of the models was evaluated by applying cross-validation based on the Monte Carlo cross-validation (MCCV) <sup>46</sup> scheme. The procedure was as follows: for 10'000 iterations, the whole training set (sample size  $n$ ) was randomly split into an internal calibration set with a sample size  $n_c$  and an internal validation set with sample size  $n_v$ . The ratio was chosen to be  $n_c/n=0.5$  <sup>47</sup>. For every iteration, the median value of the absolute error of the prediction (MedAEP) for the validation set was calculated, as described elsewhere <sup>47</sup>. The mean value of the MedAEP for the 10'000 iterations was then taken as the MedAEP<sub>cv</sub>, the median value of the absolute error for the cross-validation. The MedAEP<sub>cv</sub> can be interpreted as the cross-validated predictive power of the model. This was compared to the MedAEP of the MLR on the full calibration site. The MedAEP is the fitting ability of the model on the calibration set. <sup>47</sup> We obtained values of 1.12 for MedAEP<sub>cv</sub> and 1.10 for MedAEP, showing that our chosen model has a robust predictive power statistic. The confidence interval for the obtained correlation coefficient  $r$  was calculated by applying a Fisher transformation <sup>48</sup>. The 95% confidence interval is  $r=[0.57,0.62]$  (**Table 1**).

In order to evaluate the influence of our developed desolvation term on the estimation of the affinities, we trained a MLR model while setting  $C_{desolv}$  to 0.0 (NODESSCORE). The model reached pearson's correlation coefficient of  $r = 0.60$ . We

again applied cross-validation and obtained values of 1.13 for MedAEP<sub>cv</sub> and 1.12 for MedAEP, again indicating a robust model. Therefore, for the prediction of affinities, including the desolvation term doesn't seem to result in a better fit of the model.

Inspired by the good correlation between the  $\Delta$ SAS and the binding affinity in the CASF benchmark <sup>7</sup>, we also trained a model that was solely based on the LipoScore and on the van der Waals terms. The obtained Pearson's r equaled to 0.55, showing a good correlation (**Table 1**). Again, the values of 1.22 for MedAEP<sub>cv</sub> and 1.22 for MedAEP indicated a robust model. The 95% confidence interval for Pearson's r is [0.52,0.58]. Since the confidence interval marginally overlaps with the one obtained for the full DESSCORE function and with the one of the NODESCORE function, the performance of the simple LipoVdw scoring function is not significantly worse on the provided training set.

#### *CASF2013 Scoring Power Test*

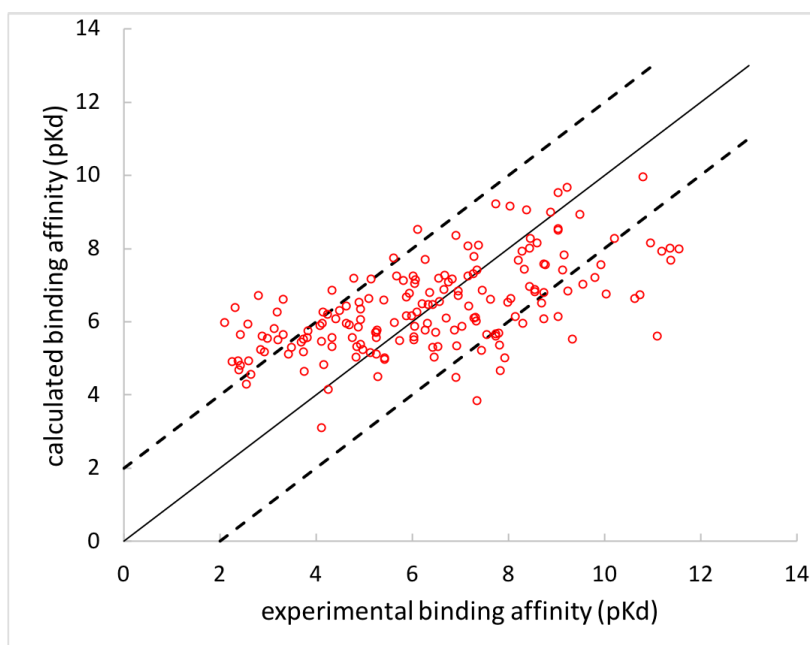
To evaluate the performance of the trained scoring function models in comparison with other published methods, we assessed the scoring power using the CASF2013 benchmark <sup>7</sup>. In the scoring test of this benchmark, the calculated affinities are compared with the experimental values on the PDBbind core set, which is a high quality subset of the PDBbind refined set <sup>49</sup>. The original core set contains 195 structures, divided into 65 clusters. Each cluster contains 3 ligands, classified into a "best", a "median" and a "poorest" binder, whereas the best binder should at least bind 100 times strong as the poorest binder in terms of the dissociation constant  $K_d$  <sup>7</sup>.

In our workflow, 176 structures of the core set were successfully processed, and 55 clusters were present with all 3 member structures processed. Therefore, the results of our benchmarks using the PDBbind core set are based on these 176 substructures.

Our DESSCORE function reached a Pearson's r of 0.59, showing the same performance as for the training set. The standard deviation was 1.85 and 70.5% of all



predictions were within 2.0 logK<sub>D</sub> values. With these values, DESSCORE ranks in the top 5 of the scoring functions tested in the CASF2013, both in terms of Pearson's *r* and standard deviation. The experimental affinities are plotted versus the predicted values in **Figure 4**. Again, NODESSCORE reached the same performance as DESSCORE (*r* = 0.60, SD = 1.85). Interestingly, whereas LipoVdwSCORE performed worse than the two other models for the training set, it reached the same correlation coefficient in the test set (*r* = 0.60, SD = 1.88). A comparison of the performance metrics of the three models is provided in **Table 2**.

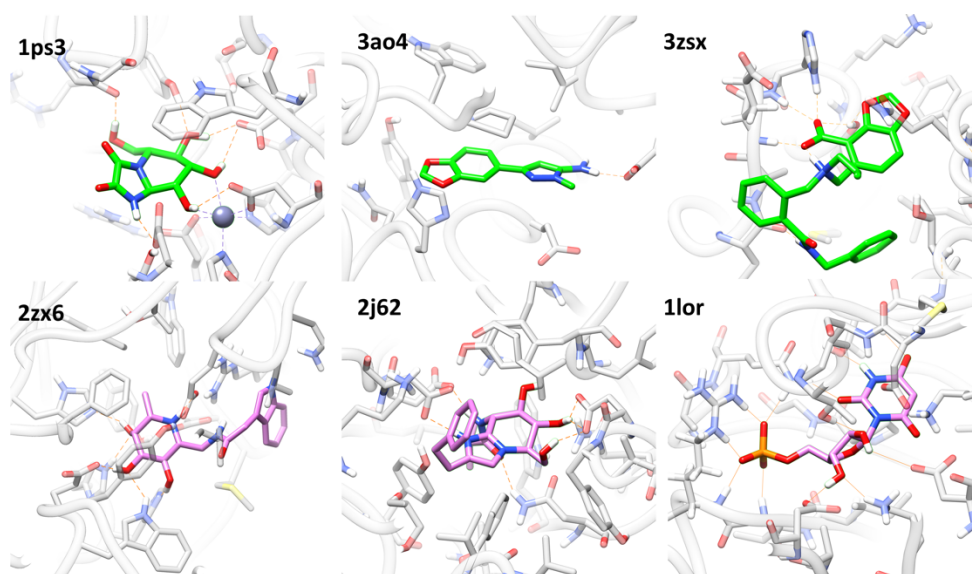


**Figure 4.** Correlation plot between experimental binding affinities and binding affinities predicted by DESSCORE for the test set. The dashed lines give the region where the predictions are within an absolute error of 2.0 kcal/mol.

**Table 2:** Performance of the three scoring function models on the scoring test (PDBbind core set)

	DESSCORE	NODESSCORE	LipoVdwSCORE
<b>Pearson's <i>r</i></b>	0.60	0.60	0.60

SD	1.85	1.85	1.88
----	------	------	------



**Figure 5.** Examples from the CASF2013 scoring power test conducted with DESSCORE that showed large deviations from the experimental values. The examples from the upper row were all predicted too strong, the ones in the lower row were predicted too weak. The deviation was in all cases more than two orders of magnitudes in terms of  $K_d$ .

The binding poses of some examples that exhibited predicted scores that differed by more than two  $\log K_d$  from the experimental values are shown in **Figure 5**. The shown examples with underestimated affinities display cooperative hydrogen bond networks.

#### CASF2013 Ranking Power Test

The CASF-2013 ranking power test is also performed on the PDBbind core set. In this test, the ability of the scoring function to accurately rank the poorest, the median and the best binder within a cluster. The authors of the original benchmark study thereby distinguished between high-level success (correct ranking of all three binders) and low-level success (correct identification of the best binder) <sup>7</sup>.

For our three models, the ranking power was evaluated for 55 complexes of the core set. DESSCORE reached a high-level success rate of 49.0%, which places it in the

middle section of the evaluated scoring functions from the CASF2013 (the best scoring function – X-SCORE – reached 58.5%). Interestingly, with a low-level success rate of 74.5%, DESSCORE is the second best scoring function in this regard (GoldScore@GOLD reached 76.9%). This high gap between high-level success and low-level success means that in a considerable number of cases, the scoring function cannot correctly distinguish poor binders from median binders. This can be a drawback in virtual screening studies, where distinguishing non-binders from potential hits is necessary <sup>7</sup>, since these hits are usually in the medium-level affinity of the target. NODESSCORE reached a high-level success rate of 47.1% and a low-level success rate of 68.6%. LipoVdwScore reached a high-level success rate of 49.0%, whereas the low-level success rate was 66.7 % (**Table 3**).

**Table 3:** Performance of the three scoring function models on the ranking test

	DESSCORE	NODESSCORE	LipoVdwSCORE
<b>high-level</b>	49.0%	47.1%	49.0%
<b>success</b>			
<b>low-level</b>	74.5%	68.6%	66.7%
<b>success</b>			

#### *Original CASF2013 Docking Power Test*

In the CASF docking power test, decoy poses are provided for every complex of the core set of the PDBbind. These decoy poses were created by the GOLD, Surflex and the MOE docking programs <sup>7</sup>. This decoy poses exhibit a wide range of RMSD values (up to 10.0 Å) in reference to the native pose. A successful docking prediction is achieved by a given scoring function if the best-scored pose is within 2.0 Å from the native pose (Top1 success). If such a pose is within the best two scored poses, it is called a Top2 success,

and within the best three it is a Top3 success. Contrary to the scoring power and ranking power test, we conducted this benchmark on the original structures from the PDBbind core set, as it was done for the CASF2013 challenge. If we would take our prepared core set with the improved hydrogen assignment, this would bias the docking scores towards the native ligand, since its tautomer state was assigned in order to optimize the hydrogen bond network.

From this original core set, various structures couldn't be processed by the System builder from Desmond. The results are therefore obtained for a subset (n=120) of the core set.

**Table 4:** Performance of the three scoring function models on the original docking power test

	<b>DESSCORE</b>	<b>NODESSCORE</b>	<b>LipoVdwSCORE</b>
<b>Top1 success</b>	81.7%	80.0%	75.0%
<b>Top2 success</b>	87.5%	87.5%	84.2%
<b>Top3 success</b>	90.8%	91.7%	87.5%

There is no clear difference between the DESSCORE and the NODESSCORE models in terms of success rates on any of the three levels, while they both outperform the LipoVdwSCORE model. Still, the surprisingly good docking power of the naïve model in the LipoVdwSCORE raises questions on how well-suited the chosen decoys are to really challenge the docking power of the scoring functions.

## Conclusion

We developed an empirical scoring function (DESSCORE) that incorporates a simple model for assessing protein and ligand desolvation. We trained the coefficients of this scoring function by MLR using complexes and binding affinities from the refined

set of the PBDbind. A robust model was obtained that exhibited a Pearson's correlation coefficient of  $r = 0.60$  on the training set. In order to assess the influence of the desolvation term, a model was trained that neglected this contribution (NODESSCORE), together with a naïve model that only incorporated the van der Waals and lipophilic terms. Whereas DESSCORE and NODESSCORE showed a superior scoring power for the training set compared to LipoVdwSCORE, they all exhibited the same scoring power on the test set taken from the CASF2013 benchmark.

In the ranking test, all three models were mediocre in the task of accurately ranking the poorest, the median and the strongest binder of a cluster. However, DESSCORE with the added desolvation term was clearly superior in selectively determining the strongest binder in a cluster. The added desolvation term has however no beneficial effect in distinguishing weak from mediocre binders.

Our simple approach for the desolvation comes with various approximations: A linear scaling between the change in accessibility for solvent upon binding and the amount of desolvation seems a crude approximation. If only a small space around the atom is available for solvent, but strong hydrogen bonds to water molecules are still possible, then the magnitude of the true desolvation cost is expected to be limited. Furthermore, our model so far neglects the effect of bridging water molecules between the protein and ligand. In principle, such an approach can be easily incorporated into our workflow. Hydration sites that solvate the ligand polar atoms and overlap with favorable regions from the water grid are potential bridging water molecules. Orientational and translational sampling could then be performed to assess the exact geometric configuration and energetics of these bridging water molecules. Such water molecules can then be assigned special scores.

In the docking power test, the DESSCORE and NODESSCORE models both outperformed the naïve model, even though the latter performed surprisingly well,

especially considering the Top3 success rate. There was however no difference between the DESCSCORE and the NODESSCORE model in the docking power test.

In general, there doesn't seem to be a real benefit in terms of scoring or docking power for the incorporation of the desolvation term presented in this study. The only benefit seems to be in a better discrimination between medium and high-affinity binders. Future improvements should aim for a better response in the low- and medium-affinity range, which would also improve the perspectives for potential virtual screening applications.

An important lesson from this study was to always critically check the performance of a trained model by comparing it with a simple baseline model, even when the values look promising first. This provides an assessment of the real value of the newly added features of a model.

## References

- (1) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr Comput Aided Drug Des* **2011**, *7* (2), 146–157.
- (2) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A Review of Protein-Small Molecule Docking Methods. *J Comput Aided Mol Des* **2002**, *16* (3), 151–166. <https://doi.org/10.1023/A:1020155510718>.
- (3) Ferreira, L. G.; dos Santos, R. N.; Oliva, G.; Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* **2015**, *20* (7), 13384–13421. <https://doi.org/10.3390/molecules200713384>.
- (4) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nature Reviews Drug Discovery* **2004**, *3* (11), 935–949. <https://doi.org/10.1038/nrd1549>.
- (5) Kontoyianni, M. Docking and Virtual Screening in Drug Discovery. *Methods Mol. Biol.* **2017**, *1647*, 255–266. [https://doi.org/10.1007/978-1-4939-7201-2\\_18](https://doi.org/10.1007/978-1-4939-7201-2_18).
- (6) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual Screening Using Protein–Ligand Docking: Avoiding



- Artificial Enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 793–806. <https://doi.org/10.1021/ci034289q>.
- (7) Li, Y.; Han, L.; Liu, Z.; Wang, R. Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J. Chem. Inf. Model.* **2014**, *54* (6), 1717–1736. <https://doi.org/10.1021/ci500081m>.
- (8) Shoichet, B. K. No Free Energy Lunch. *Nature Biotechnology* **2007**, *25* (10), 1109–1110. <https://doi.org/10.1038/nbt1007-1109>.
- (9) Liu, J.; Wang, R. Classification of Current Scoring Functions. *J Chem Inf Model* **2015**, *55* (3), 475–482. <https://doi.org/10.1021/ci500731a>.
- (10) Gilson, M. K.; Given, J. A.; Head, M. S. A New Class of Models for Computing Receptor-Ligand Binding Affinities. *Chemistry & Biology* **1997**, *4* (2), 87–92. [https://doi.org/10.1016/S1074-5521\(97\)90251-9](https://doi.org/10.1016/S1074-5521(97)90251-9).
- (11) Zou, X.; Yaxiong; Kuntz, I. D. Inclusion of Solvation in Ligand Binding Free Energy Calculations Using the Generalized-Born Model. *J. Am. Chem. Soc.* **1999**, *121* (35), 8033–8043. <https://doi.org/10.1021/ja984102p>.
- (12) Mysinger, M. M.; Shoichet, B. K. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. *J Chem Inf Model* **2010**, *50* (9), 1561–1573. <https://doi.org/10.1021/ci100214a>.
- (13) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *Journal of Computational Chemistry* **2007**, *28* (6), 1145–1152. <https://doi.org/10.1002/jcc.20634>.
- (14) Schneider, N.; Lange, G.; Hindle, S.; Klein, R.; Rarey, M. A Consistent Description of HYdrogen Bond and DEhydration Energies in Protein-Ligand Complexes: Methods behind the HYDE Scoring Function. *J. Comput. Aided Mol. Des.* **2013**, *27* (1), 15–29. <https://doi.org/10.1007/s10822-012-9626-2>.
- (15) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. <https://doi.org/10.1021/jm0306430>.
- (16) Böhm, H.-J. The Development of a Simple Empirical Scoring Function to Estimate the Binding Constant for a Protein-Ligand Complex of Known Three-Dimensional Structure. *J Computer-Aided Mol Des* **1994**, *8* (3), 243–256. <https://doi.org/10.1007/BF00126743>.
- (17) Ladbury, J. E. Just Add Water! The Effect of Water on the Specificity of Protein-Ligand Binding Sites and Its Potential Application to Drug Design. *Chemistry & Biology* **1996**, *3* (12), 973–980. [https://doi.org/10.1016/S1074-5521\(96\)90164-7](https://doi.org/10.1016/S1074-5521(96)90164-7).
- (18) Roberts, B. C.; Mancera, R. L. Ligand–Protein Docking with Water Molecules. *J. Chem. Inf. Model.* **2008**, *48* (2), 397–408. <https://doi.org/10.1021/ci700285e>.

- (19) Thilagavathi, R.; Mancera, R. L. Ligand–Protein Cross-Docking with Water Molecules. *J. Chem. Inf. Model.* **2010**, *50* (3), 415–421. <https://doi.org/10.1021/ci900345h>.
- (20) García-Sosa, A. T.; Mancera, R. L.; Dean, P. M. WaterScore: A Novel Method for Distinguishing between Bound and Displaceable Water Molecules in the Crystal Structure of the Binding Site of Protein-Ligand Complexes. *J Mol Model* **2003**, *9* (3), 172–182. <https://doi.org/10.1007/s00894-003-0129-x>.
- (21) Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Modeling Water Molecules in Protein–Ligand Docking Using GOLD. *J. Med. Chem.* **2005**, *48* (20), 6504–6515. <https://doi.org/10.1021/jm050543p>.
- (22) Forli, S.; Olson, A. J. A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J. Med. Chem.* **2012**, *55* (2), 623–638. <https://doi.org/10.1021/jm2005145>.
- (23) Murphy, R. B.; Repasky, M. P.; Greenwood, J. R.; Tubert-Brohman, I.; Jerome, S.; Annabhimoju, R.; Boyles, N. A.; Schmitz, C. D.; Abel, R.; Farid, R.; et al. WScore: A Flexible and Accurate Treatment of Explicit Water Molecules in Ligand–Receptor Docking. *J. Med. Chem.* **2016**, *59* (9), 4364–4384. <https://doi.org/10.1021/acs.jmedchem.6b00131>.
- (24) Uehara, S.; Tanaka, S. AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* **2016**, *21* (11). <https://doi.org/10.3390/molecules21111604>.
- (25) Balius, T. E.; Fischer, M.; Stein, R. M.; Adler, T. B.; Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T.; Shoichet, B. K. Testing Inhomogeneous Solvation Theory in Structure-Based Ligand Discovery. *PNAS* **2017**, *114* (33), E6839–E6846. <https://doi.org/10.1073/pnas.1703287114>.
- (26) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130* (9), 2817–2831. <https://doi.org/10.1021/ja0771033>.
- (27) Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor Cucurbit[7]Uril. *J Chem Phys* **2012**, *137* (4). <https://doi.org/10.1063/1.4733951>.
- (28) Sridhar, A.; Ross, G. A.; Biggin, P. C. Waterdock 2.0: Water Placement Prediction for Holo-Structures with a Pymol Plugin. *PLOS ONE* **2017**, *12* (2), e0172743. <https://doi.org/10.1371/journal.pone.0172743>.
- (29) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47* (12), 2977–2980. <https://doi.org/10.1021/jm030580l>.
- (30) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry*

and Biochemistry Held in Jerusalem, Israel, April 13–16, 1981; Pullman, B., Ed.; The Jerusalem Symposia on Quantum Chemistry and Biochemistry; Springer Netherlands: Dordrecht, 1981; pp 331–342. [https://doi.org/10.1007/978-94-015-7658-1\\_21](https://doi.org/10.1007/978-94-015-7658-1_21).

(31) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated Docking with Grid-Based Energy Evaluation. *Journal of Computational Chemistry* **1992**, *13* (4), 505–524. <https://doi.org/10.1002/jcc.540130412>.

(32) O’Boyle, N. M.; Brewerton, S. C.; Taylor, R. Using Buriedness to Improve Discrimination between Actives and Inactives in Docking. *J Chem Inf Model* **2008**, *48* (6), 1269–1278. <https://doi.org/10.1021/ci8000452>.

(33) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *Journal of Computational Chemistry* **2005**, *26* (16), 1752–1780. <https://doi.org/10.1002/jcc.20292>.

(34) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical Scoring Functions: I. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J Comput Aided Mol Des* **1997**, *11* (5), 425–445. <https://doi.org/10.1023/A:1007996124545>.

(35) Vedani, A. YETI: An Interactive Molecular Mechanics Program for Small-Molecule Protein Complexes. *Journal of Computational Chemistry* **1988**, *9* (3), 269–280. <https://doi.org/10.1002/jcc.540090310>.

(36) Vedani, A.; Dobler, M.; Hu, Z.; Smieško, M. OpenVirtualToxLab—A Platform for Generating and Exchanging in Silico Toxicity Data. *Toxicology Letters* **2015**, *232* (2), 519–532. <https://doi.org/10.1016/j.toxlet.2014.09.004>.

(37) Vedani, A.; Dunitz, J. D. Lone-Pair Directionality in Hydrogen-Bond Potential Functions for Molecular Mechanics Calculations: The Inhibition of Human Carbonic Anhydrase II by Sulfonamides. *J. Am. Chem. Soc.* **1985**, *107* (25), 7653–7658. <https://doi.org/10.1021/ja00311a071>.

(38) Vedani, A.; Huhta, D. W. A New Force Field for Modeling Metalloproteins. *J. Am. Chem. Soc.* **1990**, *112* (12), 4759–4767. <https://doi.org/10.1021/ja00168a021>.

(39) *Small-Molecule Drug Discovery Suite 2017-4*, Schrödinger, LLC, New York, NY, 2017.

(40) *Schrödinger Release 2017-4: Desmond Molecular Dynamics System*, D. E. Shaw Research, New York, NY, 2017. *Maestro-Desmond Interoperability Tools*, Schrödinger, New York, NY, 2017.

(41) Wang, R.; Lai, L.; Wang, S. Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. *J Comput Aided Mol Des* **2002**, *16* (1), 11–26. <https://doi.org/10.1023/A:1016357811882>.

(42) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J Comput Chem* **2010**, *31* (2), 455–461. <https://doi.org/10.1002/jcc.21334>.

- (43) Quiroga, R.; Villarreal, M. A. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. *PLoS One* **2016**, *11* (5). <https://doi.org/10.1371/journal.pone.0155183>.
- (44) Ballester, P. J.; Mitchell, J. B. O. A Machine Learning Approach to Predicting Protein–Ligand Binding Affinity with Applications to Molecular Docking. *Bioinformatics* **2010**, *26* (9), 1169–1175. <https://doi.org/10.1093/bioinformatics/btq112>.
- (45) Li, L.; Wang, B.; Meroueh, S. O. Support Vector Regression Scoring of Receptor-Ligand Complexes for Rank-Ordering and Virtual Screening of Chemical Libraries. *J Chem Inf Model* **2011**, *51* (9), 2132–2138. <https://doi.org/10.1021/ci200078f>.
- (46) Landrum, G. *RDKit: Open-Source Cheminformatics Software*. [Http://Www.Rdkit.Org](http://www.rdkit.org).
- (47) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, *55* (12), 2562–2574. <https://doi.org/10.1021/acs.jcim.5b00654>.
- (48) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A Fast Method of Molecular Shape Comparison: A Simple Application of a Gaussian Description of Molecular Shape. *Journal of Computational Chemistry* *17* (14), 1653–1666.
- (49) Haque, I. S.; Pande, V. S. PAPER--Accelerating Parallel Evaluations of ROCS. *J Comput Chem* **2010**, *31* (1), 117–132. <https://doi.org/10.1002/jcc.21307>.
- (50) Shao, J. Linear Model Selection by Cross-Validation. *Journal of the American Statistical Association* **1993**, *88* (422), 486–494. <https://doi.org/10.1080/01621459.1993.10476299>.
- (51) Konovalov, D. A.; Llewellyn, L. E.; Vander Heyden, Y.; Coomans, D. Robust Cross-Validation of Linear Regression QSAR Models. *J. Chem. Inf. Model.* **2008**, *48* (10), 2081–2094. <https://doi.org/10.1021/ci800209k>.
- (52) Nicholls, A. Confidence Limits, Error Bars and Method Comparison in Molecular Modeling. Part 1: The Calculation of Confidence Intervals. *J Comput Aided Mol Des* **2014**, *28* (9), 887–918. <https://doi.org/10.1007/s10822-014-9753-z>.
- (53) Liu, Z.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.; Nie, W.; Liu, Y.; Wang, R. PDB-Wide Collection of Binding Data: Current Status of the PDBbind Database. *Bioinformatics* **2015**, *31* (3), 405–412. <https://doi.org/10.1093/bioinformatics/btu626>.

## 2.4 “Endocrine Disruption at the Androgen Receptor: Employing Molecular Dynamics and Docking for Improved Virtual Screening and Toxicity Prediction”

Wahl, J; Smieško, M. Endocrine Disruption at the Androgen Receptor: Employing Molecular Dynamics and Docking for Improved Virtual Screening and Toxicity Prediction. *Int. J. Mol. Sci.*, **2018**, *19*, 1784.

### 2.4.1 Summary

The absence of an experimentally resolved 3D structure of the androgen receptor (AR) bound to an antagonist is a clear hurdle in SBDD approaches at this receptor. To facilitate more confident *in silico* predictions of AR antagonism, an approach based on 100 ns MD simulations of AR complexes (taking the agonist AR structure as the initial structure) bound to different antagonists was applied, with the goal to obtain binding site conformations that are able to accommodate AR antagonists. With the antagonist protein ensemble generated by this protocol, a docking approach could successfully discriminate AR antagonists from a set of decoys. However, the antagonist protein structures can also accommodate AR agonists with good docking scores. Therefore, molecular docking towards such an ensemble of AR antagonist structures cannot lead to a discrimination between agonist and antagonist ligands.

### 2.4.2 Author contributions

J.W. designed the study and run all calculations. All under the supervision of M.S.

J.W. wrote the manuscript. M.S. proofread and improved the manuscript.

### 2.4.3 Potential Impact on the Scientific Field

The prediction of AR antagonism has a high relevance for the field of cancer therapy, but also for the prediction of endocrine disruption. Our study showed that MD simulations can help to obtain protein structures of functional states in the absence of experimental structures. This finding can possibly be extended to other systems.

Finally, the findings from the study are expected to support future drug design and toxicology campaigns involving AR antagonism.








## Article

# Endocrine Disruption at the Androgen Receptor: Employing Molecular Dynamics and Docking for Improved Virtual Screening and Toxicity Prediction

Joel Wahl and Martin Smieško \* 

Molecular Modeling, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland; joel.wahl@unibas.ch

\* Correspondence: martin.smiesko@unibas.ch; Tel.: +41-61-207-1563

Received: 4 May 2018; Accepted: 6 June 2018; Published: 15 June 2018



**Abstract:** The androgen receptor (AR) is a key target for the development of drugs targeting hormone-dependent prostate cancer, but has also an important role in endocrine disruption. Reliable prediction of the binding of ligands towards the AR is therefore of great relevance. Molecular docking is a powerful computational method for exploring small-ligand binding to proteins. It can be applied for virtual screening experiments but also for predicting molecular initiating events in toxicology. However, in case of AR, there is no antagonist-bound crystal structure yet available. Our study demonstrates that molecular docking approaches are not able to satisfactorily screen for AR antagonists because of this reason. Therefore, we applied Molecular Dynamics simulations to generate antagonist AR structures and showed that this leads to a vast improvement for the docking of AR antagonists. We benchmarked the ability of these antagonist AR structures discriminate between AR antagonists and decoys using an ensemble docking approach and obtained promising results with good enrichment. However, distinguishing AR antagonists from agonists with high confidence is not possible with the current approach alone.

**Keywords:** androgen receptor; ensemble docking; virtual screening; molecular dynamics simulations; agonists; antagonists

## 1. Introduction

The androgen receptor (AR) is the main biomolecular target involved in the development and progression of the hormone-dependent prostate cancer [1]. Furthermore, disruption of the androgen system is associated with decreased sperm count, increased infertility [2], and diabetes mellitus [3]. Therefore, in silico evaluation of small-molecule binding at the AR is of high relevance for the design of novel antiandrogens for the treatment of prostate cancer [1,4–6], but also for the screening of potential androgen-disrupting chemicals [7] causing endocrine disorders [8,9]. Several in silico toxicity prediction models and tools have shown their potential to identify endocrine disrupting chemicals [10–12].

One major challenge associated with the computational prediction of binding modes and affinities of compounds towards the androgen receptor by structure-based methods is the lack of a published structure of antagonist-bound AR and experimental and theoretical studies aimed at elucidating the mechanisms of AR antagonism [13–20].

In its unbound state, the AR is complexed with heat-shock and chaperone proteins that stabilize the apo state of the protein [4]. Agonist binding triggers and stabilizes a conformational change that facilitates coactivator binding to a so-called activation function-2 (AF2) region [4]. A widely proposed mechanism for AR antagonism is the displacement of the Helix-12 (H12) upon ligand binding, leading to distortions in the AF2 site, preventing coactivator binding [13]. This mechanism was recently supported by computational

studies employing Molecular Dynamics (MD) simulations [14–16]. Furthermore, the repositioning of H12 was confirmed to be the mechanism of antagonism for the estrogen receptor (ER) and the peroxisome proliferator-activated receptor (PPAR) [21].

Additional insights into structural determinants for androgen antagonism were gained by structure-activity relationship studies, where the systematic influence of scaffold change on androgen agonism or antagonism was investigated [18–20]. Bohl and coworkers resolved the crystal structures for several steroidal agonists and steroidal and nonsteroidal antagonists towards the wild type (WT) AR and T877A and W741L mutants. They showed how the LBD of the AR can accommodate the B-ring of agonist S-1 by the repositioning of the indole ring of Trp741 [18]. McGinley and coworkers developed pan-antagonists by increasing the size of the B-ring [20]. This was based on the hypothesis that the increased bulkiness of the ring results in displacement of H12. This hypothesis was later disproved by the publication of a crystal structure of a naphthyl-containing WT agonist bound to the T877A form of the androgen receptor [19]. The receptor could accommodate the bulky ring substituent by a rearrangement of the Met745 side chain.

Regarding B-ring containing AR binders, the chemical nature of the linker seems to be the determining factor in converting agonists to antagonist and vice versa. The antagonist behavior of bicalutamide versus the agonist behavior of compound of S-22, an AR agonist containing an oxygen linker, can be explained by the presence of the much bulkier sulfonyl linker in the former [19].

Molecular Dynamics simulations (MD simulations) were applied in various studies to investigate structural determinants of androgen receptor antagonism [14–17]. Zhou and coworkers could show by analyzing replica exchange molecular dynamics simulations (REMD) that AR antagonism correlates with structural fluctuations of helix 12 [17]. Duan et al. applied microsecond long MD simulations and enhanced sampling to reveal the structural dynamics of HFT and bicalutamide binding to AR and also observed increased fluctuations for helix 12 [14]. Liu and coworkers simulated bicalutamide binding towards androgen WT and several mutations and identified a steric clash with helix 12 as the structural determinant for WT AR antagonism [15]. These studies established the applicability of molecular simulations for gaining structural insights into AR antagonism.

However, MD simulations of ligand-receptor complexes need a valid input structure of the bound ligand in the binding site and furthermore, due to their computational expense, cannot be applied to large libraries of compounds that are usually used in virtual screening campaigns. Structure-based virtual screening (SBVS) [22,23] based on molecular docking [24] can bridge this gap by generating conformations of the bound ligand (“poses”) at the binding site of a receptor, the docking poses are accompanied by a score which is an estimate of the affinity of the scored ligand towards the receptor. Ligand enrichment is then based on the docking scores. SBVS is now a widely used method in the field and has displayed several success stories in hit identification [22,25].

Furthermore, molecular docking can be applied to toxicology profiling of chemical compounds by assessing how well a chemical entity fits into the binding site [7,26]. Molecular docking studies require a representative receptor conformation for the prediction of binding poses and binding affinities of the screened ligands. This structure can for example be obtained by protein crystallography. For the AR, where a crystal structure in the antagonist conformation is not available, this is not feasible. By only applying docking to the agonist-bound structure, the outcomes of the virtual screening campaigns are expected to be biased towards compounds that act in an agonistic manner. One possibility to create an antagonistic structure of the AR is to apply homology modeling [27–29], taking experimental antagonistic structures of related Nuclear Receptors (NR) such as the estrogen receptor (ER). Whereas this strategy led to the successful discovery of novel antiandrogens, the docking accuracy of these antagonist structures were not evaluated.

Bisson et al. used a biased probability Monte Carlo procedure and iteratively refined receptor conformations that are biased to yield good discrimination of agonists and antagonists [30]. We used a different approach and applied MD simulations of various antagonists bound to WT AR. We used representative structures of these simulations (obtained by clustering of the trajectories) in an ensemble

docking approach and validated if this strategy can be used for improved screening for AR antagonists, an application that is highly relevant for drug discovery and toxicology studies. In an ensemble docking approach, flexible ligand—rigid protein (the degrees of freedom of the receptor are not altered) docking is applied to different protein conformations obtained by crystal structures or from a Molecular Dynamics trajectory [31–33].

To validate the ability of MD simulations to generate antagonist-conformations of the AR, we ran several simulations of antagonist-bound WT AR and used a clustering approach to generate a structural ensemble for docking towards the AR. We then docked androgen antagonists towards this antagonist-ensemble and calculated the performance metrics compared to a set of decoys. Furthermore, we also constructed an agonist ensemble by simulating agonist-bound AR complexes and to assess if the performance for agonist screening is improved compared to the ensemble obtained from crystal structures.

A generally similar strategy was applied by Swift and coworkers [34] where AR conformations were generated by MD simulations and the best scoring ensembles in discriminating actives and decoys based on the ROC AUC values were identified by different search methods. However, they only run a simulation of one receptor-ligand complex and furthermore, no distinction between agonists and antagonists was made.

All evaluations of the docking accuracy were done target the wild type AR, but the strategy can be extended to include AR mutants as well.

The agonist and antagonist structural ensemble presented in our study have the potential to improve virtual screening campaigns and toxicology predictions for AR antagonism.

## 2. Results

### 2.1. Docking of AR Agonists, Antagonists and Decoys towards Experimental AR Structures

A diverse ensemble (by binding site shape similarity, see Methods Section) of five experimental AR structures (PDB IDs: 3B66, 3G0W, 3V49, 4HLW, and 2PNU) was used for an initial evaluation of the docking performance of Glide SP at the AR (crystal structures ensemble, Table 1).

As a first step, we conducted a cross-docking experiment, where we docked all five ligands from the crystal structures into all five receptor conformations to see if the Glide-SP score assigns the best score for a ligand to the corresponding structure it was extracted from. We determined furthermore how well the experimental pose is reproduced in terms of RMSD (self-docking).

All ligands showed the highest predicted affinity towards their corresponding native (experimental) protein conformation in the cross-docking experiment (Table 1). Furthermore, all native poses were recovered with satisfying accuracy in the self-docking experiments (RMSD values for self-docking: 3B66: 0.22 Å, 2PNU: 0.25 Å, 3G0W: 0.18 Å, 3V49: 0.49 Å, 4HLW: 0.80 Å). The most promiscuous conformation is 2PNU, reporting Glide Scores for all five ligands that were docked. On the other hand, the native ligand of the structure 2PNU (EM5744) corresponds to a bulky ligand (Figure 1) that could only be successfully docked into its native protein conformation. Binding of EM5744 requires a substantial induced fit, induced by rearrangement of Met895 and Trp741. None of the four other protein structures display this structural feature and none of them can therefore accommodate this type of ligand. The magnitude of the induced fit is also expressed by significant differences in binding site volumes of the different structures (2PNU: 261 Å<sup>3</sup>, 3B66: 189 Å<sup>3</sup>, 3G0W: 157 Å<sup>3</sup>, 3V49: 171 Å<sup>3</sup>, 4HLW: 168 Å<sup>3</sup>). The importance of the induced fit for binding at the AR was confirmed in a previous study [26].

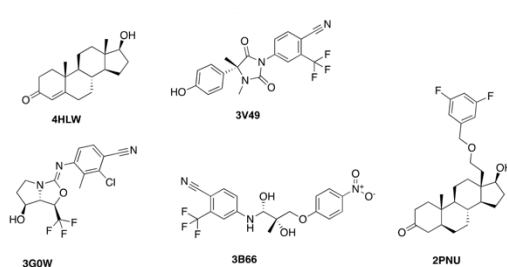
In a next step, an agonist and antagonist set, taken from the NRlist BDB [35], and a set of 3000 decoys, taken from the enhanced directory of useful decoys (DUD-E) [36] were docked into the ensemble and docking performance was evaluated using the area under the receiver operating characteristic (AUC ROC). For the docking of the agonist versus the decoy set, Glide SP reached an



impressive AUC ROC of 0.92 (Table 2), showing that discrimination between AR agonists and decoys is achieved to a high degree.

**Table 1.** Results from the cross-docking experiment. Listed are the PDB IDs of the ligands and corresponding protein structures together with their Glide-Scores.

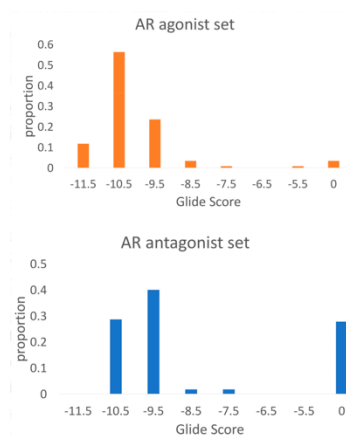
Protein Structure	3B66	3G0W	3V49	4HLW	2PNU
ligand					
3B66	−12.2	−8.5	−8.9	N/A	−10.9
3G0W	−10.8	−11.2	−10.6	−8.4	−10.0
3V49	−10.5	N/A	−12.2	N/A	−9.7
4HLW	−8.8	N/A	−9.4	−11.3	−10.0
2PNU	N/A	N/A	N/A	N/A	−14.2



**Figure 1.** 2D Structures of androgen receptor (AR) agonists bound to the five AR crystal structures.

Taking experimentally validated decoys can be an alternative to the DUD-E, as described by Lagarde [37]. In case of the androgen receptor, the antagonist set can serve as a challenging decoy set. The ROC AUC in this case amounted to 0.80 (Table 2), showing discrimination between agonists and antagonists.

The docking score distributions (Figure 2) for the AR agonists and antagonists show in general lower Glide SP Scores (corresponding to higher affinities) for the AR agonists. 68% of the agonists reached a docking score of −10.0 or lower, compared to 29% of the antagonists. Moreover, 32 out of 115 antagonists (28%) could not be successfully docked into any member of the ensemble, compared to 3 agonists where no docking solution was found.

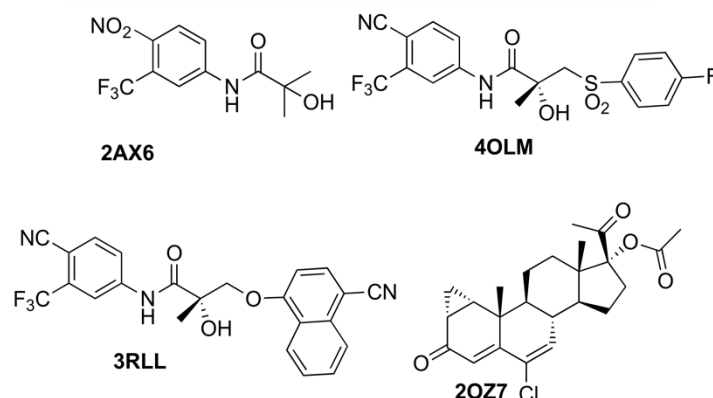


**Figure 2.** Glide docking score distributions for AR agonists and antagonists docked towards a structural ensemble of five agonist AR crystal structures. A docking score of 0 indicates that Glide could not successfully dock the ligand into any member of the ensemble.

As a third evaluation, we used the antagonist set as the active set and the DUD-E set as decoys. In this case, the ROC AUC reached 0.65 (Table 2). This shows that, whereas antagonists docked to the agonist structures of AR were ranked slightly higher than a set of decoys, they, in general, display a low affinity towards the agonist structures and in certain cases cannot fit into the binding site. As a consequence, structure-based virtual screenings or toxicology screenings based on agonist crystal structures are expected to be successful for the identification of AR agonists, but can only insufficiently model AR antagonists, and therefore, the need for AR structures that correspond to the antagonist conformation is high.

## 2.2. Generation of AR Antagonist Structures Using Flexible Docking, Binding Pose Metadynamics and Classical MD Simulations

We used crystal structures of four different AR WT antagonists that act as agonists towards AR receptors with either a T877A or W741L mutation (PDB IDs: 2AX6 (T877A), 2OZ7 (T877A), 3RLL (T877A), 4OLM (W741L)). The antagonist ligand of 2AX6 [18] corresponds to Hydroxyflutamide (HFT), 4OLM [38] corresponds to R-Bicalutamide, and 2OZ7 [39] to cyproterone acetate. The ligand in the structure 3RLL [19] is named as structure 7 in the original publication and will be referred as Duke-7 in the following. Whereas 3 out of the 4 antagonists share the common scaffold from HFT, they differ in terms of presence and size of the B-ring (4OLM, 3RLL) (Figure 3) and the type of the linker. 2OZ7 has a steroid-core and is therefore structurally different.



**Figure 3.** AR antagonists used for the generation of the ensemble of antagonist AR structures.

We mutated the corresponding amino acids back to obtain the WT receptor. These structures will be termed as 2AXA\_WT, 4OLM\_WT, 3RLL\_WT, and 2OZ7\_WT in the following. We then employed induced fit docking (IFD) to dock the antagonists back into the corresponding WT structures and applied binding pose metadynamics [40] (see Methods Section) to assess the structural stability of the obtained binding poses. The most stable pose was then subject to a 100 ns MD simulation. From these trajectories, representative structures were extracted as the cluster centers of the three most populated clusters (see Methods Section). These representative structures then were used for the evaluation of the docking performance and for the selection of the best-performing AR antagonist ensemble (explained in the Methods Section).

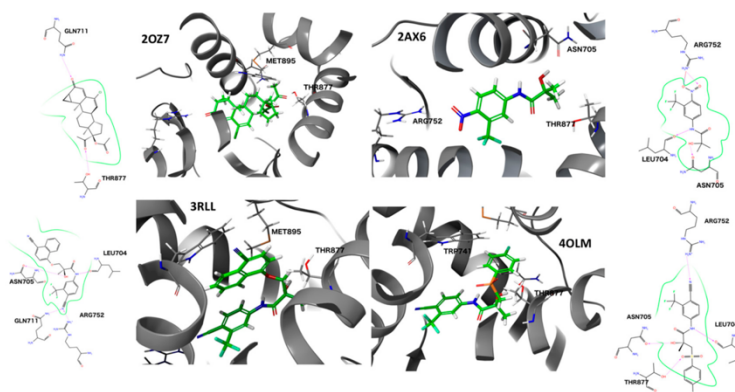
For 2AX6\_WT, the IFD protocol identified four distinct poses: The terminal hydroxyl group can have two different orientations (hydrogen bond either to Thr877 or Asn705). Furthermore, the phenyl ring can have two different orientations as well as the amide group. These four poses were then assessed by the binding pose metadynamics protocol and ranked by their corresponding PoseScores, which is an indication of the stability of the pose [40]. More details are given in the method section. The most stable pose (Figure 4) is characterized by the same binding motif that is present for HFT

bound to AR T877A, exhibiting hydrogen bonding to of the ligand hydroxyl group to the side chain amide oxygen of Asn705.

The same protocol was applied to the other antagonist ligands. For 4OLM\_WT, six different poses were identified by IFD, mainly differing by the orientation of the B-ring. The orientation of the B-ring in the most stable pose (Figure 4) differs vastly from the observed orientation in the crystal structure with the W741L mutation. The presence of the bulky tryptophan side chain makes the accommodation of the B-ring together with the bulky sulfonyl linker more demanding, explaining the occurrence of a different ring- and linker orientation.

Four distinct poses were found for 3RLL\_WT by IFD. The bulkiness of the B-ring doesn't allow alternative orientations and the identified poses are differing in the orientation of the A-ring. The best-scored pose (Figure 4) has the identical binding motif as in the crystal structure bound to T877A AR.

For 2OZ7, seven poses could be identified by IFD, differing in the orientation of the terminal acetate and keto groups. Due to the presence of a hydrogen bond to Thr877, the orientation of the keto group for the identified pose at the WT AR differs from the one that is observed in the crystal structure at the T877A AR (Figure 4).



**Figure 4.** Docking poses of AR antagonists at AR wild type (WT) obtained from induced fit docking (IFD) and binding pose metadynamics.

The structures with the most stable poses for all complexes were used as input structures for 100 ns MD simulations. We also conducted 100 ns MD simulations of the original PDB complexes (AR mutations) to get an ensemble of agonist structures for comparison and assessment of how the mutation affects the structural dynamics of the receptor.

A comparison of the representative structures (cluster centroids) from the 100 ns MD simulations started from the original PDB structures (agonist activity towards mutants) and from the stable docking poses from the IFD/binding pose metadynamics towards the WT (antagonist activity) is shown in Figure 5.

Structure 2AX6 corresponds to Hydroxyflutamide (HFT) bound to T877A AR. Towards this mutant, HFT acts as an agonist, whereas it has antagonist activity towards the WT AR. The terminal isopropanol group of HFT shows different orientations for the representative MD structures of the wild-type and the T877A structure (Figure 5, panel A). This is probably due to a smaller steric demand of the alanine side chain compared to threonine. This also leads to a different orientation of Ile899 and Met895. Furthermore, the backbone atoms of ILE899 are shifted away from the binding site for the WT structure.

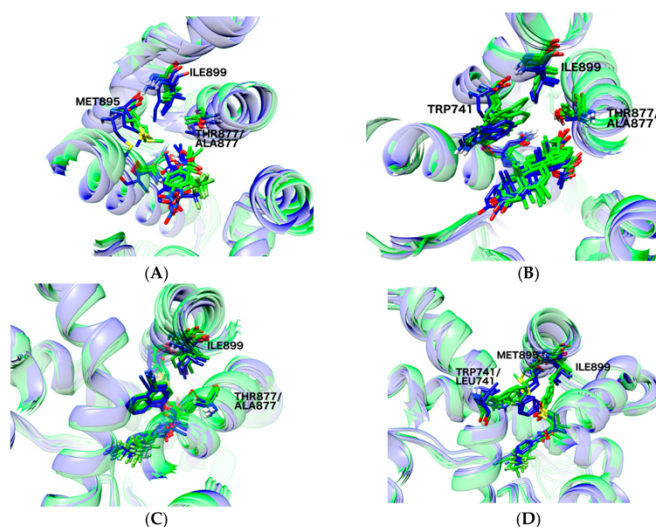
For cyproterone acetate (CPA, Figure 5B), similar observations can be made. For the wild-type simulations, the backbone of Ile899 is shifted away from the binding site. This is accompanied by a



shift of the ring moiety of Trp741. The keto group of CPA interacts with the Thr877 side chain in all snapshots for the WT simulation.

Ligand Duke-7 (Figure 5C) has a bulky B-ring with an oxygen linker. The presence of the bulkier Thr877 side chain in the wild-type forces a shift in the oxygen linker and again results in a slight movement of Ile899 away from the binding site.

The biggest structural differences between WT simulation and mutant simulation are observed for the binding of R-bicalutamide (Figure 5D). The W741L mutation can accommodate the B-ring with the bulky sulfonyl-linker, whereas for the wild type, there is a steric clash with the indole ring of Trp741, and the B-ring undergoes structural change and points away from the binding site, thereby pushing away the backbone atoms of Met895 and Ile899.



**Figure 5.** Cluster centroids from 100 ns MD simulation of four different ligands bound to mutated (agonist activity) and wild type (antagonist activity) AR. The structures from the mutated structures are depicted in blue, the WT structures in green. PDB IDs: (A) 2AX6; (B) 2OZ7; (C) 3RLL; (D) 4OLM.

### 2.3. Docking Towards an Antagonist Ensemble Obtained from MD Simulations

The analysis of the antagonist-bound MD trajectories of the AR clearly indicated structural differences compared with trajectories of mutated AR, for which the antagonists are converted to agonists. However there remains the question if these antagonist-structures created by MD simulations can be used for improved prediction of AR antagonism by molecular docking.

As a first assessment, a set of antagonists and decoys was docked to the antagonist ensemble obtained from the MD simulations. An exhaustive search was then applied to pick a subset of the ensemble, consisting of five structures (see Methods Section). The number was chosen since it provides a good trade-off between covering the necessary conformational flexibility and still achieving results with little computational resources. Analyses with varying ensemble sizes confirmed this choice (results not shown). The ensemble was picked that obtained the best AUC ROC for discrimination between antagonists and agonists (ensemble MD\_antagonist\_1, Table 2).

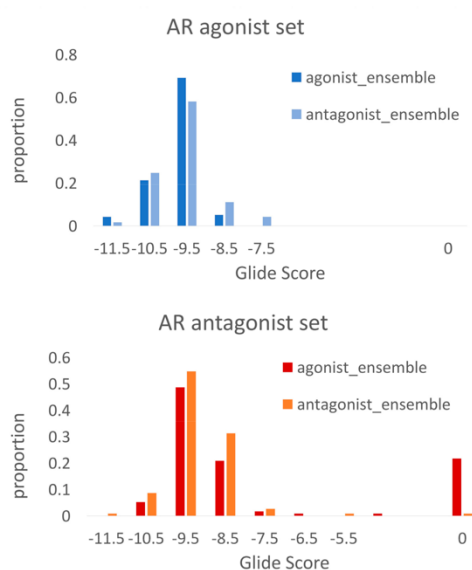
The best ensemble reached a promising AUC-ROC of 0.86, indicating that indeed the antagonist structures created by MD simulations can be used for structure-based virtual screening of AR antagonists. The antagonist ensemble from the MD clearly outperforms the agonist ensemble from the crystal structure in terms of antagonist docking. Only one antagonist molecule could not be successfully docked into the ensemble, a vast improvement compared with 32 unsuccessful docking attempts from the agonist crystal ensemble. The best performing ensemble was composed of the

structures 2AX6\_WT\_1, 2AX6\_WT\_2, 2OZ7\_WT\_1, 2OZ7\_WT\_2, 2OZ7\_WT\_3 (the final number refers to the index of the cluster centroid obtained from the clustering of the trajectories) and it will be termed as the MD\_antagonist\_1 ensemble in the following.

**Table 2.** Docking performances of the different ensembles against various sets of actives and decoys.

Ensemble	ROC AUC	Actives	Decoys
crystal structures	0.92	agonists	DUD-E
crystal structures	0.80	agonists	antagonists
crystal structures	0.65	antagonists	DUD-E
MD_antagonist_1	0.86	antagonists	DUD-E
MD_antagonist_1	0.29	antagonists	agonists
MD_antagonist_1	0.92	agonists	DUD-E
MD_antagonist_1	0.96	agonists	DUD-E
MD_antagonist_1	0.82	agonists	antagonists
MD_antagonist_1	0.69	antagonists	decoys
MD_antagonist_2	0.96	agonists	decoys
MD_antagonist_2	0.74	antagonists	decoys
MD_antagonist_3	0.79	antagonists	decoys
MD_antagonist_3	0.56	antagonists	agonists

As a second evaluation, the agonist ligand set was docked towards the MD\_antagonist\_1 ensemble and the AUC ROC against the decoy set was determined. The obtained value of 0.92 indicates that this ensemble is not selective for AR antagonists, but also shows high promiscuity for the agonist ligands. This can be seen in Figure 6, where the distribution of the Glide scores of the agonists are similar for both docking at the agonist ensemble and at the antagonist ensemble.



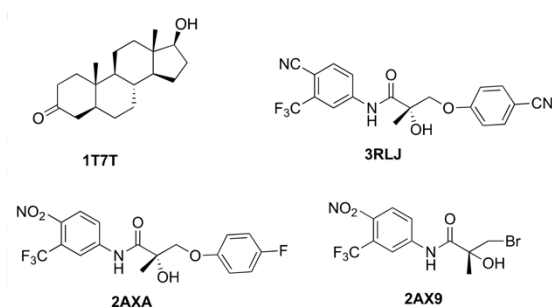
**Figure 6.** Docking score distributions for agonist and antagonist ligands docked to the agonistic and antagonistic receptor ensembles.

#### 2.4. Docking Towards an Agonist Ensemble Obtained from MD Simulations

In order to have consistent structural ensembles for docking of both agonist and antagonist, we also conducted classical MD simulations of four structurally diverse agonists bound to WT AR (Figure 7). The input structures were directly taken as the crystal structures (PDB IDs: 1T7T, 2AXA, 2AX9, 3RLJ). We docked the agonist set to all 12 representative structures derived from clustering

of the MD trajectories (four trajectories and three cluster centers each) and towards the same set of 3000 decoys as for the antagonist experiment. The best ensemble of five structures, showing the best discrimination between agonists and decoys, was again determined by an exhaustive search (ensemble MD\_agonist\_1). The best performing agonist ensemble reached an impressive AUC-ROC of 0.96 (Table 2) and consisted of the structures 1T7T\_1, 1T7T\_2, 2AX9\_1, 2AX9\_2, 3RLJ\_1. All 117 agonist ligands could be successfully docked into the ensemble.

The antagonist ligands reached an AUC ROC of 0.69 (Table 2) at the MD\_agonist\_1 ensemble against the decoy set, implying that the selected ensemble is selective for agonists and shows low promiscuity for the antagonists. This is confirmed by Figure 6, where the Glide scores for the antagonists docked to the antagonist ensemble are, in general, more favorable than the scores at the agonist ensemble. Furthermore, 22 antagonists could not be successfully docked into the agonist ensemble.



**Figure 7.** AR agonists used for the generation of the ensemble of agonist AR structures.

In order to verify if the improved AUC ROC values for the MD-derived ensembles compared to the crystal structures really stem from the sampling of agonist and antagonist-like conformations and not just the better sampling of the receptor conformational space in general, we also determined the AUC ROC against the decoy set of docking the agonist ligands into the antagonist ensemble and vice versa by selecting the best performing agonist ensemble for the docking of the antagonists (ensemble MD\_agonist\_2) and the best performing antagonist ensemble for the docking of the agonists (ensemble MD\_antagonist\_2) (Table 2). The best performing agonist ensemble for the antagonist set had an AUC ROC of 0.74, significantly worse than the one of the best performing antagonist ensembles. On the other hand, the antagonist ensemble achieving the best AUC ROC for the docking of agonists reached an AUC ROC of 0.96, confirming that AR agonists exhibit good docking scores towards both the agonist as well as the antagonist ensemble.

## 2.5. Discriminating Agonists from Antagonists

The above results indicate that ensemble-docking towards MD-derived structures can be used to distinguish agonists from decoys and antagonists from decoys, which is definitely a step forward for structure-based design at the AR. Yet, there remains the question if this strategy can distinguish agonists from antagonists. Therefore, a new search for well-performing ensembles was conducted. For the selection of the antagonist ensemble, the objective function was the summed AUC-ROC for (i) docking the antagonist ligands and the DUD-E decoys and (ii) docking the antagonists and taking the agonists as decoys. The same search was performed for the agonist ensemble, where the agonists were docked and the AUC-ROC was evaluated with taking (i) the DUD-E decoys and (ii) the antagonists as decoys. The best antagonist ensemble with best discrimination against decoys and agonists had AUC-ROC values of 0.79 (decoys) and 0.56 (agonists) (ensemble MD\_antagonists\_3, Table 2). For the agonist ensemble, the AUC-ROC values were 0.96 (decoys) and 0.82 (antagonists) and it coincided with the best performing ensemble picked for maximum AUC-ROC against the decoys



only (ensemble MD\_agonist\_1, Table 2). This ensemble facilitates discrimination of AR agonists against AR antagonists and decoys at the same time. For the antagonist structures, an ensemble that can successfully distinguish antagonists from decoys and agonists could not be found.

This implies that when we conduct a virtual screening campaign or toxicology screening using the presented agonistic ensemble, the compounds with the highest score are the agonist compounds and good selectivity over decoys and antagonists can be achieved. On the other hand, conducting the same screening using the antagonistic ensemble, agonists, and antagonists reach a similar score, making the identification of antagonists based on docking scores impossible.

This is confirmed by the docking score distributions (Figure 6).

### 2.6. Structural Comparison between Antagonist and Agonist Ensembles

The ensembles MD\_antagonist\_1 and MD\_agonist\_1 are depicted in Figure 8. Visual inspection surprisingly revealed only minor differences in the placement of helix12. The major differences are visible for the amino acids Phe876 and Leu701, exhibiting a much bigger conformation diversity for the antagonist ensemble.

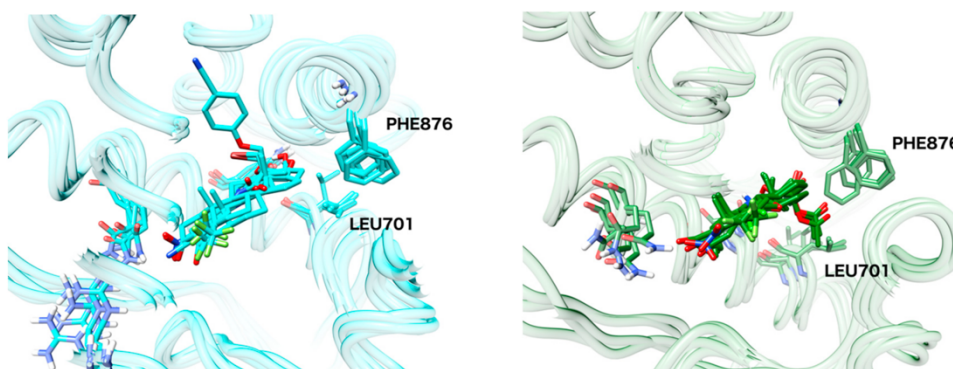


Figure 8. MD\_agonist\_ensemble\_1 (left) and MD\_antagonist\_ensemble\_1 (right).

The movement of Phe876 and Leu701 is most prominent for the member 2OZ7\_WT\_2 of the antagonist ensemble. The movement of these two amino acids facilitates an opening of the binding pocket and therefore, sterically demanding, large androgen antagonists with linear shapes can be docked to this structure, whereas docking them to any structure of the agonist ensemble is impossible and leads to steric clashes. Yet, it still remains unclear how this class of AR antagonists exhibits their mode of action if displacement of helix12 is not the underlying mechanism.

## 3. Materials and Methods

### 3.1. AR Agonists, Antagonists and Decoys

The androgen agonist and antagonist sets were downloaded from the NRlist BDB [35] (downloaded: September 2017). We manually curated some of the compounds by checking the primary literature sources of all agonists and antagonists with reported  $K_d$  values. Some antagonists were wrongly present as agonists in the database. These discrepancies were corrected together with some wrong affinity data. This finally resulted in two sets containing 118 agonists and 115 antagonists.

For the decoys, the enhanced directory of useful decoys (DUD-E) [36] was downloaded for the AR. From this directory, we picked a diverse subset of 3000 compounds to get a ratio of actives to decoys of about 1:30. The diverse set was created by clustering with the SkeletonSpheres Descriptor implemented in the DataWarrior package [41].

### 3.2. Preparation of Experimental AR Structures for Ensemble Docking

All androgen WT crystal structures with a resolution of better or equal than 2.5 Å were downloaded from the protein data bank (PDB, [www.rcsb.org](http://www.rcsb.org)). The structures were refined with the Protein Preparation Wizard [42,43] from the Small-Molecule Drug Discovery Suite from Schrödinger. All hydrogens were added and protonation states were determined for a pH of 7.4. Missing loops and side-chains were added and the hydrogen network was refined, followed by a short minimization. We clustered the structures according to their binding site similarity, as described elsewhere [44]. We identified five clusters from the hierarchical clustering and the centroids of these clusters composed the ensemble used for the docking.

### 3.3. Molecular Dynamics Simulations

Receptor-ligand complexes of AR antagonists bound to mutants and AR WT as well as AR agonists bound to AR WT were refined with the Protein Preparation Wizard [43] (see previous paragraph). All molecular dynamics simulations were carried out with Desmond [45,46]. The default relaxation protocol was applied, consisting of the following steps: (i) 100 ps Brownian Dynamics in the NVT ensemble at 10 K and solute heavy atoms restraints (50 kcal/mol/Å<sup>2</sup>) (ii) 12 ps simulation in the NVT ensemble, keeping the restraints and temperature at 10 K (iii) 12 ps simulation in the NPT ensemble, keeping restraints and temperature at 10 K (iv) 12 ps simulation in the NPT simulation with solute heavy atom restraints at 300 K and (v) 24 ps simulation in the NPT ensemble at 300 K without restraints. Production runs were performed in the NPT ensemble at 300 K and 1.01 bar for 100 ns. The TIP4P-EW water model was used and counterions were added to neutralize the net charge of the system. The temperature and pressure were controlled using a Langevin thermostat (relaxation time 1.0 ps) and a Langevin barostat (relaxation time 2.0 ps). Electrostatic interactions were treated using the particle mesh Ewald summation. The OPLS3 force field was employed [47].

Representative structures for each trajectory were obtained by an affinity propagation clustering based on the heavy atom RMSD values of all heavy atoms within 5.0 Å of the ligand. The centroids of the three most populated clusters were taken as the representative structures.

### 3.4. Binding Pose Metadynamics

Metadynamics facilitates enhanced sampling for Molecular Dynamics simulations by penalizing already visited configurational states of the system, whereby a bias is added set of collective variables (CVs) [48]. Binding Pose Metadynamics [40] is a tool developed by Schrödinger aiming to identify the most stable (and therefore considered best) pose from an induced-fit docking run. The selected collective variable is the RMSD (root mean square deviation) of the ligand atoms and a selected set of binding site atoms from an initial, equilibrated configuration. A set of short metadynamics simulations is run and the free energy of the system is evaluated as a function of the CV. The pose stability is expressed as the thermodynamically most favorable RMSD from the starting configuration. If this RMSD is low, the initial pose can be regarded as stable. The protocol uses the Desmond engine [45] and the OPLS3 force field [47].

### 3.5. Glide Docking

Flexible-ligand rigid-protein docking was performed using the Glide program from Schrödinger [49]. The applied scoring function was Glide SP (standard precision). The ligand structures were prepared using LigPrep [50]. Protonation states were generated at pH 7.0 ± 2.0, and a state penalty was added to the final docking score. A detailed list of the applied settings for the docking is given in the Supporting Information (S2).

### 3.6. Induced-Fit Docking (IFD)

IFD was performed using the induced-fit docking protocol from Schrödinger [51]. The detailed settings are tabulated in the supporting information (S3).

### 3.7. Ensemble Docking

Ensemble docking usually involves flexible ligand—fixed protein docking towards a representative ensemble of protein structures [31,32,34,52,53]. This approach accounts for the protein-flexibility at a moderate computational cost. In our implementation, all ligands were docked to all members of the ensemble and the best Glide-SP score of a ligand across the ensemble was taken as the final docking score. The successful outcome of an ensemble docking depends on the generation and selection of the structural ensemble [34]. Docking to the ensemble should give accurate docking poses [52] and successful discrimination between actives and inactives. In our study, we selected ensembles from a pool of protein conformations generated by MD simulations. We generated a pool of structures from MD simulations of agonists bound to AR and of antagonists bound to AR, which were then used to construct agonist and antagonist ensembles respectively. We chose an ensemble size of five and selected the best ensembles depending on the objective function (successful discrimination of actives and inactives). We chose an exhaustive search procedure for picking the best ensemble, meaning that all possible combinations of five structures were evaluated from the pool.

### 3.8. Evaluation of the Docking Performance

The chosen metric for the docking performance was the ROC AUC, the area under the curve of the receiver operator characteristic curve. For docking and scoring, the AUC gives the probability that a randomly chosen true active is ranked higher than a randomly chosen decoy. A ROC AUC of 1.0 would thereby indicate that all actives are ranked before the inactives/decoys. The ROC AUC metric was criticized for an insufficient incorporation of early enrichment [54]. However it has assets such as statistical reliability [55] and robust behavior towards decoys that are actually active against the target [55] and is in general a widely applied metric for the assessment of docking performance. In addition, since we also look at androgen binding from the point of view of toxicology, where early enrichment is not a main objective, and not only virtual screening, we regarded the ROC AUC as a good measure for the docking performance at the AR. In general, the enrichment obtained by a docking and scoring approach is evaluated by the ability of the methodology to discriminate actives from inactives. For the AR, actives can be furthermore divided into agonist and antagonists. An accurate method would successfully distinguish agonists from decoys, antagonists from decoys and agonists from antagonists.

## 4. Conclusions

Molecular Dynamics simulations in combination with molecular docking was applied to study the ability of structure-based screening methods to predict AR agonism and antagonism. Docking against an ensemble derived from experimental crystal structures of AR in agonist conformation was shown to yield good enrichments for the screening of agonists, but is insufficient for antagonist docking. This was circumvented by using a workflow that combines induced-fit docking of known antagonists towards the AR, assessing the most stable poses and generate structural ensembles by running 100 ns MD simulations from these structures followed by a clustering procedure. To our knowledge, this the first study that benchmarks a docking approach at the AR aiming at prediction of agonism and antagonism.

This MD-derived ensemble of antagonist-like protein conformations was shown to yield superior docking performance in terms of discriminating antagonists from decoys, compared to the ensemble from the crystal structure.



Furthermore, an ensemble of agonist-like protein conformations was created by running MD simulations of bound agonists. This ensemble also showed better enrichments than the ensemble from the experimental crystal structures. A remaining challenge is the discrimination between agonists and antagonists. The antagonist ensemble shows high promiscuity towards both agonists and antagonists, whereas the agonist ensemble is selective for agonists. As a result, screening selectively for agonists is possible, by docking the library only towards the agonist set and taking the highest-scoring compounds as leads. However, selectively screening for antagonists based only on docking is not feasible.

Lagarde et al. recently showed that nuclear receptor agonists and antagonists can be discriminated using a 3D pharmacophore approach [56]. Therefore, substituting the SBVS with a ligand-based method can be expected to yield good results for a selective screening of antagonists. Furthermore, studies employing long-timescale MD simulations showed different protein structural dynamics for agonist-bound AR compared to antagonist-bound AR [14,21] and differing protein-ligand interactions for agonist and antagonists. A hierarchical approach, combining pharmacophore screening, molecular docking, and MD simulations for the best-scoring ligands from the docking can be expected to lead to a focused screening towards AR agonists and antagonists. Such studies can also be extended to different AR mutations for the screening for AR pan-antagonists.

We think that this study improves the potential of SBVS and structure-based design at the androgen receptor and leads to more reliable toxicology prediction of environmental chemicals by better accounting for AR antagonist. We could successfully create an ensemble of antagonist-like AR structures that can accommodate AR antagonists. These structures are made publicly available (see Supporting Information) and are aimed to help future developments of AR antagonists and more confident prediction of AR disrupting chemicals.

We furthermore showed the potential of MD simulations to generate meaningful protein conformations with desired objectives in the absence of experimental crystal structures.

To promote further use of our findings, a link for downloading the agonist ensemble, antagonist ensemble and ligand sets is provided in the Supporting Information (S1).

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/19/6/1784/s1>.

**Author Contributions:** J.W. designed and conducted the study. M.S. supervised the research. M.S. and J.W. wrote the manuscript together.

**Funding:** J.W. is funded by the Schweizerischer Nationalfonds SNF, grant number 200020\_159952.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Heinlein, C.A.; Chang, C.S. Androgen Receptor in Prostate Cancer. *Endocr. Rev.* **2004**, *25*, 276–308. [[CrossRef](#)] [[PubMed](#)]
2. Luccio-Camelo, D.C.; Prins, G.S. Disruption of Androgen Receptor Signaling in Males by Environmental Chemicals. *J. Steroid Biochem. Mol. Biol.* **2011**, *127*, 74–82. [[CrossRef](#)] [[PubMed](#)]
3. Sakkiiah, S.; Wang, T.; Zou, W.; Wang, Y.; Pan, B.; Tong, W.; Hong, H. Endocrine Disrupting Chemicals Mediated through Binding Androgen Receptor Are Associated with Diabetes Mellitus. *Int. J. Environ. Res. Public Health* **2018**, *15*, 25. [[CrossRef](#)] [[PubMed](#)]
4. Tan, M.H.E.; Li, J.; Xu, H.E.; Melcher, K.; Yong, E. Androgen Receptor: Structure, Role in Prostate Cancer and Drug Discovery. *Acta Pharmacol. Sin.* **2015**, *36*, 3–23. [[CrossRef](#)] [[PubMed](#)]
5. Helsen, C.; Van den Broeck, T.; Voet, A.; Prekovic, S.; Poppel, H.V.; Joniau, S.; Claessens, F. Androgen Receptor Antagonists for Prostate Cancer Therapy. *Endocr. Relat. Cancer* **2014**, *21*, T105–T118. [[CrossRef](#)] [[PubMed](#)]
6. Lu, X.; Dun, K.; Wang, Y.; Yang, Y.; You, Q.; Li, Z. Recent Androgen Receptor Antagonists in Prostate Cancer. *Mini. Rev. Med. Chem.* **2014**, *14*, 655–663. [[CrossRef](#)] [[PubMed](#)]
7. Vedani, A.; Dobler, M.; Hu, Z.; Smieško, M. OpenVirtualToxLab—A Platform for Generating and Exchanging in Silico Toxicity Data. *Toxicol. Lett.* **2015**, *232*, 519–532. [[CrossRef](#)] [[PubMed](#)]



8. Takeuchi, T.; Tsutsumi, O.; Ikezuki, Y.; Takai, Y.; Taketani, Y. Positive Relationship between Androgen and the Endocrine Disruptor, Bisphenol A, in Normal Women and Women with Ovarian Dysfunction. *Endocr. J.* **2004**, *51*, 165–169. [CrossRef] [PubMed]
9. Mnif, W.; Hassine, A.I.H.; Bouaziz, A.; Bartegi, A.; Thomas, O.; Roig, B. Effect of Endocrine Disruptor Pesticides: A Review. *Int. J. Environ. Res. Public Health* **2011**, *8*, 2265–2303. [CrossRef] [PubMed]
10. Vedani, A.; Smiesko, M. In Silico Toxicology in Drug Discovery—Concepts Based on Three-Dimensional Models. *Atla-Altern. Lab. Anim.* **2009**, *37*, 477–496.
11. Vedani, A.; Dobler, M.; Lill, M.A. The Challenge of Predicting Drug Toxicity in Silico. *Basic Clin. Pharmacol. Toxicol.* **2006**, *99*, 195–208. [CrossRef] [PubMed]
12. Lill, M.A.; Winiger, F.; Vedani, A.; Ernst, B. Impact of Induced Fit on Ligand Binding to the Androgen Receptor: A Multidimensional QSAR Study to Predict Endocrine-Disrupting Effects of Environmental Chemicals. *J. Med. Chem.* **2005**, *48*, 5666–5674. [CrossRef] [PubMed]
13. Osguthorpe, D.J.; Hagler, A.T. Mechanism of Androgen Receptor Antagonism by Bicalutamide in the Treatment of Prostate Cancer. *Biochemistry* **2011**, *50*, 4105–4113. [CrossRef] [PubMed]
14. Duan, M.; Liu, N.; Zhou, W.; Li, D.; Yang, M.; Hou, T. Structural Diversity of Ligand-Binding Androgen Receptors Revealed by Microsecond Long Molecular Dynamics Simulations and Enhanced Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 4611–4619. [CrossRef] [PubMed]
15. Liu, H.; Han, R.; Li, J.; Liu, H.; Zheng, L. Molecular Mechanism of R-Bicalutamide Switching from Androgen Receptor Antagonist to Agonist Induced by Amino Acid Mutations Using Molecular Dynamics Simulations and Free Energy Calculation. *J. Comput. Aided Mol. Des.* **2016**, *30*, 1189–1200. [CrossRef] [PubMed]
16. Liu, H.-L.; Zhong, H.-Y.; Song, T.-Q.; Li, J.-Z. A Molecular Modeling Study of the Hydroxyflutamide Resistance Mechanism Induced by Androgen Receptor Mutations. *Int. J. Mol. Sci.* **2017**, *18*, 1823. [CrossRef] [PubMed]
17. Zhou, J.; Liu, B.; Geng, G.; Wu, J.H. Study of the Impact of the T877A Mutation on Ligand-Induced Helix-12 Positioning of the Androgen Receptor Resulted in Design and Synthesis of Novel Antiandrogens. *Proteins Struct. Funct. Bioinform.* **2010**, *78*, 623–637. [CrossRef] [PubMed]
18. Bohl, C.E.; Miller, D.D.; Chen, J.; Bell, C.E.; Dalton, J.T. Structural Basis for Accommodation of Nonsteroidal Ligands in the Androgen Receptor. *J. Biol. Chem.* **2005**, *280*, 37747–37754. [CrossRef] [PubMed]
19. Duke, C.B.; Jones, A.; Bohl, C.E.; Dalton, J.T.; Miller, D.D. Unexpected Binding Orientation of Bulky-B-Ring Anti-Androgens and Implications for Future Drug Targets. *J. Med. Chem.* **2011**, *54*, 3973–3976. [CrossRef] [PubMed]
20. McGinley, P.L.; Koh, J.T. Circumventing Anti-Androgen Resistance by Molecular Design. *J. Am. Chem. Soc.* **2007**, *129*, 3822–3823. [CrossRef] [PubMed]
21. Moore, J.T.; Collins, J.L.; Pearce, K.H. The Nuclear Receptor Superfamily and Drug Discovery. *ChemMedChem* **2006**, *1*, 504–523. [CrossRef] [PubMed]
22. Lionta, E.; Spyrou, G.; Vassilatis, D.K.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr. Top. Med. Chem.* **2014**, *14*, 1923–1938. [CrossRef] [PubMed]
23. Shoichet, B.K. Virtual Screening of Chemical Libraries. *Nature* **2004**, *432*, 862–865. [CrossRef] [PubMed]
24. Ferreira, L.G.; Dos Santos, R.N.; Oliva, G.; Andricopulo, A.D. Molecular Docking and Structure-Based Drug Design Strategies. *Mol. Basel Switz.* **2015**, *20*, 13384–13421. [CrossRef] [PubMed]
25. Villoutreix, B.O.; Miteva, R.E.; Miteva, M.A. Structure-Based Virtual Ligand Screening: Recent Success Stories. Available online: <http://www.eurekaselect.com/70496/article> (accessed on 9 April 2018).
26. Smieško, M. DOLINA—Docking Based on a Local Induced-Fit Algorithm: Application toward Small-Molecule Binding to Nuclear Receptors. *J. Chem. Inf. Model.* **2013**, *53*, 1415–1423. [CrossRef] [PubMed]
27. Voet, A.; Helsen, C.; Zhang, K.Y.J.; Claessens, F. The Discovery of Novel Human Androgen Receptor Antagonist Chemotypes Using a Combined Pharmacophore Screening Procedure. *ChemMedChem* **2013**, *8*, 644–651. [CrossRef] [PubMed]
28. Bassetto, M.; Ferla, S.; Pertusati, F.; Kandil, S.; Westwell, A.D.; Brancale, A.; McGuigan, C. Design and Synthesis of Novel Bicalutamide and Enzalutamide Derivatives as Antiproliferative Agents for the Treatment of Prostate Cancer. *Eur. J. Med. Chem.* **2016**, *118*, 230–243. [CrossRef] [PubMed]

29. Shen, H.C.; Shanmugasundaram, K.; Simon, N.I.; Cai, C.; Wang, H.; Chen, S.; Balk, S.P.; Rigby, A.C. In Silico Discovery of Androgen Receptor Antagonists with Activity in Castration Resistant Prostate Cancer. *Mol. Endocrinol.* **2012**, *26*, 1836–1846. [[CrossRef](#)] [[PubMed](#)]
30. Bisson, W.H.; Cheltsov, A.V.; Bruey-Sedano, N.; Lin, B.; Chen, J.; Goldberger, N.; May, L.T.; Christopoulos, A.; Dalton, J.T.; Sexton, P.M.; et al. Discovery of Antiandrogen Activity of Nonsteroidal Scaffolds of Marketed Drugs. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11927–11932. [[CrossRef](#)] [[PubMed](#)]
31. Campbell, A.J.; Lamb, M.L.; Joseph-McCarthy, D. Ensemble-Based Docking Using Biased Molecular Dynamics. *J. Chem. Inf. Model.* **2014**, *54*, 2127–2138. [[CrossRef](#)] [[PubMed](#)]
32. Tian, S.; Sun, H.; Pan, P.; Li, D.; Zhen, X.; Li, Y.; Hou, T. Assessing an Ensemble Docking-Based Virtual Screening Strategy for Kinase Targets by Considering Protein Flexibility. *J. Chem. Inf. Model.* **2014**, *54*, 2664–2679. [[CrossRef](#)] [[PubMed](#)]
33. Osguthorpe, D.J.; Sherman, W.; Hagler, A.T. Exploring Protein Flexibility: Incorporating Structural Ensembles From Crystal Structures and Simulation into Virtual Screening Protocols. *J. Phys. Chem. B* **2012**, *116*, 6952–6959. [[CrossRef](#)] [[PubMed](#)]
34. Swift, R.V.; Jusoh, S.A.; Offutt, T.L.; Li, E.S.; Amaro, R.E. Knowledge-Based Methods to Train and Optimize Virtual Screening Ensembles. *J. Chem. Inf. Model.* **2016**, *56*, 830–842. [[CrossRef](#)] [[PubMed](#)]
35. Lagarde, N.; Ben Nasr, N.; Jérémie, A.; Guillemain, H.; Laville, V.; Labib, T.; Zagury, J.-F.; Montes, M. NReLiSt BDB, the Manually Curated Nuclear Receptors Ligands and Structures Benchmarking Database. *J. Med. Chem.* **2014**, *57*, 3117–3125. [[CrossRef](#)] [[PubMed](#)]
36. Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594. [[CrossRef](#)] [[PubMed](#)]
37. Lagarde, N.; Zagury, J.-F.; Montes, M. Importance of the Pharmacological Profile of the Bound Ligand in Enrichment on Nuclear Receptors: Toward the Use of Experimentally Validated Decoy Ligands. *J. Chem. Inf. Model.* **2014**, *54*, 2915–2944. [[CrossRef](#)] [[PubMed](#)]
38. Hsu, C.-L.; Liu, J.-S.; Wu, P.-L.; Guan, H.-H.; Chen, Y.-L.; Lin, A.-C.; Ting, H.-J.; Pang, S.-T.; Yeh, S.-D.; Ma, W.-L.; et al. Identification of a New Androgen Receptor (AR) Co-Regulator BUD31 and Related Peptides to Suppress Wild-Type and Mutated AR-Mediated Prostate Cancer Growth via Peptide Screening and X-ray Structure Analysis. *Mol. Oncol.* **2014**, *8*, 1575–1587. [[CrossRef](#)] [[PubMed](#)]
39. Bohl, C.E.; Wu, Z.; Miller, D.D.; Bell, C.E.; Dalton, J.T. Crystal Structure of the T877A Human Androgen Receptor Ligand-Binding Domain Complexed to Cypoterone Acetate Provides Insight for Ligand-Induced Conformational Changes and Structure-Based Drug Design. *J. Biol. Chem.* **2007**, *282*, 13648–13655. [[CrossRef](#)] [[PubMed](#)]
40. Clark, A.J.; Tiwary, P.; Borrelli, K.; Feng, S.; Miller, E.B.; Abel, R.; Friesner, R.A.; Berne, B.J. Prediction of Protein–Ligand Binding Poses via a Combination of Induced Fit Docking and Metadynamics Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 2990–2998. [[CrossRef](#)] [[PubMed](#)]
41. Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program for Chemistry Aware Data Visualization and Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460–473. [[CrossRef](#)] [[PubMed](#)]
42. Sastry, G.M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27*, 221–234. [[CrossRef](#)] [[PubMed](#)]
43. Epik, Schrödinger, LLC. *Schrödinger Suite 2016-3 Protein Preparation Wizard*; Epik, Schrödinger, LLC: New York, NY, USA, 2016.
44. Osguthorpe, D.J.; Sherman, W.; Hagler, A.T. Generation of Receptor Structural Ensembles for Virtual Screening Using Binding Site Shape Analysis and Clustering. *Chem. Biol. Drug Des.* **2012**, *80*, 182–193. [[CrossRef](#)] [[PubMed](#)]
45. D.E. Shaw Research. *Desmond Molecular Dynamics System*; D.E. Shaw Research: New York, NY, USA, 2017.
46. Bowers, K.J.; Chow, E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, 11–17 November 2006; ACM: New York, NY, USA, 2006.
47. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296. [[CrossRef](#)] [[PubMed](#)]

48. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566. [[CrossRef](#)] [[PubMed](#)]
49. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [[CrossRef](#)] [[PubMed](#)]
50. LigPrep, Schrödinger, LLC. *Schrödinger Suite 2016-3*; LigPrep, Schrödinger, LLC: New York, NY, USA, 2016.
51. Glide, Schrödinger, LLC. *Schrödinger Suite 2016-3 Induced Fit Docking Protocol*; Glide, Schrödinger, LLC: New York, NY, USA, 2016.
52. Park, S.-J.; Kufareva, I.; Abagyan, R. Improved Docking, Screening and Selectivity Prediction for Small Molecule Nuclear Receptor Modulators Using Conformational Ensembles. *J. Comput. Aided Mol. Des.* **2010**, *24*, 459–471. [[CrossRef](#)] [[PubMed](#)]
53. Korb, O.; Olsson, T.S.G.; Bowden, S.J.; Hall, R.J.; Verdonk, M.L.; Liebeschuetz, J.W.; Cole, J.C. Potential and Limitations of Ensemble Docking. *J. Chem. Inf. Model.* **2012**, *52*, 1262–1274. [[CrossRef](#)] [[PubMed](#)]
54. Truchon, J.-F.; Bayly, C.I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J. Chem. Inf. Model.* **2007**, *47*, 488–508. [[CrossRef](#)] [[PubMed](#)]
55. Nicholls, A. What Do We Know and When Do We Know It? *J. Comput. Aided Mol. Des.* **2008**, *22*, 239–255. [[CrossRef](#)] [[PubMed](#)]
56. Lagarde, N.; Delahaye, S.; Zagury, J.-F.; Montes, M. Discriminating Agonist and Antagonist Ligands of the Nuclear Receptors Using 3D-Pharmacophores. *J. Cheminform.* **2016**, *8*, 43. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 3. Implemented Scripts and Software Tools

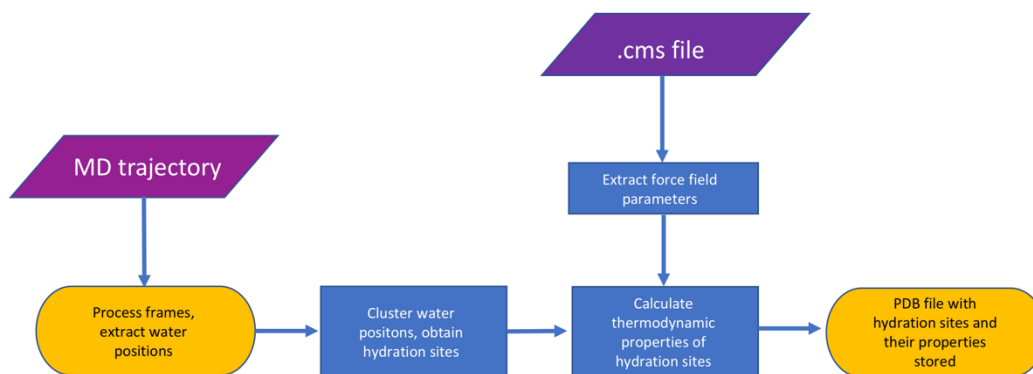
### 3.1 Technical Details

All algorithms were implemented in Python 2.7. The Python API of the Schrödinger Small-Molecule Drug Discovery Suite (Small-Molecule Drug Discovery Suite 2017-4, Schrödinger, LLC, New York, NY, 2017) was extensively used for processing and manipulating chemical structures and molecular trajectories. For numerical calculations, the NumPy package was mainly used, while SciPy and Scikit-learn were applied for data analysis and model training.

Apart from the possibility to use computational chemistry modules provided by the Schrödinger suite, Python was chosen for its readability, widespread use in the community of CADD and the availability of a plethora of libraries for scientific computing.

### 3.2 MDWatAnalyzer

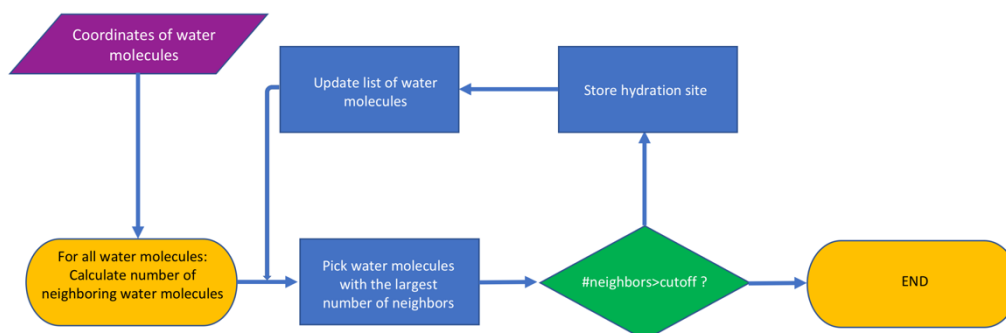
MDWatAnalyzer is a python tool that extracts locations and thermodynamic properties of hydration sites from a Desmond MD trajectory by using IFST.



**Figure 12.** Workflow of MDWatAnalyzer

The general workflow is shown in **Figure 12**. The analysis starts by the construction of a 3D grid containing the region of interest where the solvent properties should be mapped. The positions of all waters that lie in this region are stored and submitted to the clustering procedure that identifies the hydration sites with a radius of 1.0 Å. For every water molecule, the number of neighboring water molecules within 1.0 Å is calculated. The grid-based approach facilitates a fast evaluation of the number of neighbors, since only water molecules occupying neighboring grid cells have to be considered. The water molecule with the highest number of neighbors is

then taken as the center of the first hydration site. All water molecules that are contained in this hydration site are deleted from the grid and the clustering procedure is iteratively repeated until the occupancy of a hydration site is below a defined cutoff. The principle of the clustering algorithm is outlined in **Figure 13**.



**Figure 13.** Workflow of the hydration site clustering algorithm

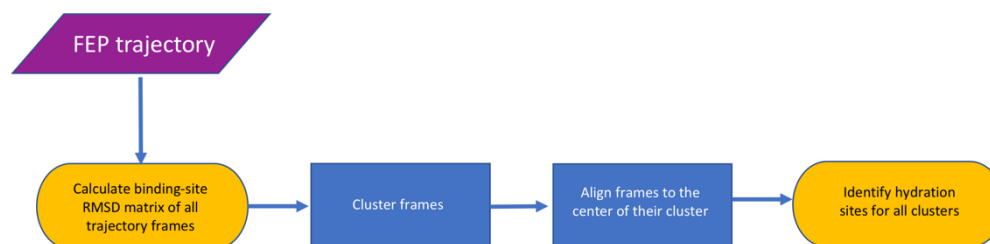
The interaction energies of all water-molecules with the rest of the system are calculated in the minimum image convention is calculated by means of the OPLS\_2005 force field. The force field parameters are extracted from the simulation file (.cms). In addition, the first order solute-water entropy is calculated for the hydration sites and the results are saved to a pdb file and can be used for visual inspection or further analysis.

All distance calculations are performed in NumPy and formulated on the basis of n-dimensional arrays, offering vast performance gains compared to a loop-based approach in native python.

### 3.3 MDWatAnalyzerFEP

The MDWatAnalyzer tool was extended for the application to trajectories from FEP calculations. MDWatAnalyzer requires a trajectory that was run with constraints applied to the solute heavy atoms, since structural rearrangements of the binding site during a flexible MD simulation blur out the solvent density, rendering the identification of hydration sites impossible. Since FEP calculations should be run without any constraints to yield meaningful results, the workflow needed to be adapted and improved.





**Figure 14.** Workflow for the water analysis applied to FEP trajectories

To achieve this, the trajectory frames are first clustered by means of the binding-site RMSD, again with a fixed cutoff. All frames are then aligned by a roto-translational fit to their cluster center. The hydration site analysis is then performed for every cluster (**Figure 14**). Since all trajectory frames that are member of the same cluster share similar binding site conformation, performing the hydration site calculations can be efficiently applied to this structural ensemble. The algorithm of the hydration site clustering and property calculation is the same as for the MDWatAnalyzer tool.

### 3.4 DESSCORer

DESSCORer is a tool that takes a protein-ligand complex as an input and estimates the binding affinity by means of the DESSCORE function, that was developed in the course of this thesis. First, grid potentials are constructed to facilitate the fast evaluation of van der Waals and Coulomb interactions, as described elsewhere (Meng et al., 1992). The force field parameters are again taken from the OPLS\_2005 force field. A water probe is then sampled over these grid points and the interactions of this water probe with the receptor are calculated and stored in a grid, yielding a water grid that is used for the assessment of the protein and ligand desolvation costs. The ligand pose is solvated by an external tool (Sridhar et al., 2017), which places hydration sites at preferred locations around polar atoms.

The program calculates various contributions towards the binding free energy (LipoScore from ChemScore (Eldridge et al., 1997), hydrogen bond and metal terms from YETI (Vedani, 1988) Coulomb and van der Waals interactions and the strain energy of the ligand). The estimated affinity is then given as the output.

## 4. Conclusion and Outlook

Different aspects on the quantification of the thermodynamic contributions of water towards ligand-protein binding processes were elucidated. Considering lead-optimization, information about binding site solvent distribution and thermodynamics can guide decisions on ligand modifications that target the displacement of specific binding site water molecules. Extracting thermodynamic solvent properties from molecular simulations by using inhomogeneous fluid solvation theory can provide information about the desolvation costs that are associated with the displacement of a binding-site water molecule. We have shown that the thermodynamic properties of the water molecules have to be calculated for both states: ligand A in complex with the protein but also for ligand B that bears the added functional group that displaces the binding site water molecule. Only by simulating both states, the rearrangement energy of the remaining solvent is taken into account. This rearrangement energy can be larger than 2 kcal/mol, especially if the displaced water molecule was directly interacting with other binding site solvent molecules.

The desolvation cost for the displacement of the binding site water molecules have to be compensated by either stronger interactions between the ligand and the protein or lowered ligand desolvation costs. Yet, we could also show that the difference in binding entropy can have a significant impact on the change in affinity. If a ligand binds much tighter with the protein after the displacement of the water molecule, the gained interaction energy is counteracted by an entropic cost, leading to an enthalpy/entropy compensation. This effect is difficult to predict, making quantitative predictions regarding solvent displacement challenging.

For more accurate predictions of the thermodynamic properties of binding site solvent, a framework that allows to apply IFST to simulation trajectories that were run without constraints would provide an improvement. A possible strategy could be to cluster the frames from the simulation and to apply IFST to all clusters.

The gold standard for estimating relative binding free energies between two congeneric ligands are alchemical binding free energy calculations such as free energy perturbation (FEP). In principle, these methods are physically rigorous and calculate the  $\Delta\Delta G$  values directly from the partition function. Contrary to end-state methods that estimate different contributions for the relative binding free energy and take the final values as the sum of the contributions, this is much less error prone. There are two limiting factors for the accuracy of such calculations: the evaluation of energies of the system and the sampling. Whereas the former is associated with general shortcomings of classical molecular force fields, the latter can be overcome with increased computational power and more enhanced sampling techniques. Yet, we have shown that ligand modifications that lead to changes in the number of binding-site solvent molecules are inherently challenging, even with the long simulation times that are available nowadays and

---

despite enhanced sampling. The reason is the slow exchange of buried cavities with the bulk solvent.

This limits the applicability of FEP calculations to such systems. A possible solution is to add algorithms that enhanced the sampling of the binding site solvent, such as Grand Canonical Monte Carlo. We have shown that incorporating a GCMC/MD solvent equilibration step before the actual FEP simulation reduces the dependency of the simulation outcome on the initial solvent structure. Therefore, the combination of FEP with GCMC leads to more robust results for systems that involve a change in binding site solvent molecules upon ligand transformations.

The disadvantage of FEP calculations is that they don't provide information about the different contributions towards the final change in binding affinity. It is inherently then not possible to attribute the exact influence of the solvent displacement on the final  $\Delta\Delta G$ . Yet, a newly published method that directly incorporates GCMC sampling into the FEP simulation stage allows to quantify the contribution of the binding site hydration to the total free energy change (Bruce Macdonald et al., 2018). Improved understanding about the various contributions towards a change in affinity has the advantage that such knowledge can be transferred to other projects in CADD, leading to the generation of some general knowledge that can help to successfully contribute to drug discovery projects. Furthermore, it helps to understand structure-activity relationships, allowing the extrapolation to the whole congeneric ligand series.

A general concern is the accuracy of current force fields, that clearly neglect important aspects such as polarization and charge anisotropy. It can be expected that all methods in CADD would profit from more accurate force fields.

We also aimed at developing a simple method for the estimation of desolvation costs in molecular docking. However, our scoring function model that incorporated this desolvation treatment was not shown to be superior in terms of affinity estimation or pose picking. Only an improved ranking for high-affinity complexes was observed. Further refinement and improvement of the method – for example taking into account bridging water molecules – is necessary to provide an efficient and robust scoring function. The fact that also a naïve scoring function model that solely relies on van der Waals energies and a lipophilic score exhibited good correlations with experimental affinity, should raise awareness about the actual value of a model. An improved scoring function should clearly outperform such a simple model in terms of affinity prediction, pose prediction, ranking and screening.



## 5. References

- Abel, R., Young, T., Farid, R., Berne, B.J., and Friesner, R.A. (2008). Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* *130*, 2817–2831.
- Abel, R., Wang, L., Friesner, R.A., and Berne, B.J. (2010). A Displaced-Solvent Functional Analysis of Model Hydrophobic Enclosures. *J. Chem. Theory Comput.* *6*, 2924–2934.
- Albrecht, L., Chowdhury, S., and Boyd, R.J. (2013). Hydrogen Bond Cooperativity in Water Hexamers: Atomic Energy Perspective of Local Stabilities. *J. Phys. Chem. A* *117*, 10790–10799.
- Amaro, R.E., Baudry, J., Chodera, J., Demir, Ö., McCammon, J.A., Miao, Y., and Smith, J.C. (2018). Ensemble Docking in Drug Discovery. *Biophysical Journal* *114*, 2271–2278.
- Balius, T.E., Fischer, M., Stein, R.M., Adler, T.B., Nguyen, C.N., Cruz, A., Gilson, M.K., Kurtzman, T., and Shoichet, B.K. (2017). Testing inhomogeneous solvation theory in structure-based ligand discovery. *PNAS* *114*, E6839–E6846.
- Ball, P. (2017). Water is an active matrix of life for cell and molecular biology. *PNAS* *114*, 13327–13335.
- Baron, R., Setny, P., and McCammon, J.A. (2010). Water in Cavity–Ligand Recognition. *J. Am. Chem. Soc.* *132*, 12091–12097.
- Barratt, E., Bingham, R.J., Warner, D.J., Laughton, C.A., Phillips, S.E.V., and Homans, S.W. (2005). Van der Waals Interactions Dominate Ligand–Protein Association in a Protein Binding Site Occluded from Solvent Water. *J. Am. Chem. Soc.* *127*, 11827–11834.
- Barratt, E., Bronowska, A., Vondrášek, J., Černý, J., Bingham, R., Phillips, S., and Homans, S.W. (2006). Thermodynamic Penalty Arising from Burial of a Ligand Polar Group Within a Hydrophobic Pocket of a Protein Receptor. *Journal of Molecular Biology* *362*, 994–1003.
- Baum, B., Mohamed, M., Zayed, M., Gerlach, C., Heine, A., Hangauer, D., and Klebe, G. (2009). More than a Simple Lipophilic Contact: A Detailed Thermodynamic Analysis of Nonbasic Residues in the S1 Pocket of Thrombin. *J. Mol. Biol.* *390*, 56–69.
- Bayden, A.S., Moustakas, D.T., Joseph-McCarthy, D., and Lamb, M.L. (2015). Evaluating Free Energies of Binding and Conservation of Crystallographic Waters Using SZMAP. *J. Chem. Inf. Model.* *55*, 1552–1565.
- Beglov, D., and Roux, B. (1997). An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* *101*, 7821–7826.
- Ben-Shalom, I.Y., Pfeiffer-Marek, S., Baringhaus, K.-H., and Gohlke, H. (2017). Efficient Approximation of Ligand Rotational and Translational Entropy Changes upon Binding for Use in MM-PBSA Calculations. *J. Chem. Inf. Model.* *57*, 170–189.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13–16, 1981*, B. Pullman, ed. (Dordrecht: Springer Netherlands), pp. 331–342.
- Betz, M., Wulsdorf, T., Krimmer, S.G., and Klebe, G. (2016). Impact of Surface Water Layers on Protein–Ligand Binding: How Well Are Experimental Data Reproduced by Molecular Dynamics Simulations in a Thermolysin Test Case? *J. Chem. Inf. Model.* *56*, 223–233.

- Beuming, T., Che, Y., Abel, R., Kim, B., Shanmugasundaram, V., and Sherman, W. (2012). Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. *Proteins: Structure, Function, and Bioinformatics* 80, 871–883.
- Biela, A., Khayat, M., Tan, H., Kong, J., Heine, A., Hangauer, D., and Klebe, G. (2012). Impact of Ligand and Protein Desolvation on Ligand Binding to the S1 Pocket of Thrombin. *Journal of Molecular Biology* 418, 350–366.
- Bitetti-Putzer, R., Joseph-McCarthy, D., Hogle, J.M., and Karplus, M. (2001). Functional group placement in protein binding sites: a comparison of GRID and MCSS. *J. Comput. Aided Mol. Des.* 15, 935–960.
- Böhm, H.-J. (1994). The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J Computer-Aided Mol Des* 8, 243–256.
- Borhani, D.W., and Shaw, D.E. (2012). The future of molecular dynamics simulations in drug discovery. *J Comput Aided Mol Des* 26, 15–26.
- Bowers, K.J., Chow, E., Xu, H., Dror, R.O., Eastwood, M.P., Gregersen, B.A., Klepeis, J.L., Kolossvary, I., Moraes, M.A., Sacerdoti, F.D., et al. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. In *In SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, (ACM Press), p.
- Breiten, B., Lockett, M.R., Sherman, W., Fujita, S., Al-Sayah, M., Lange, H., Bowers, C.M., Heroux, A., Krilov, G., and Whitesides, G.M. (2013). Water networks contribute to enthalpy/entropy compensation in protein-ligand binding. *J. Am. Chem. Soc.* 135, 15579–15584.
- Brieg, M., Setzler, J., Albert, S., and Wenzel, W. (2017). Generalized Born implicit solvent models for small molecule hydration free energies. *Phys. Chem. Chem. Phys.* 19, 1677–1685.
- Brooks, B.R., Brooks, C.L., MacKerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., et al. (2009). CHARMM: The Biomolecular Simulation Program. *J Comput Chem* 30, 1545–1614.
- Brovchenko, I., Krukau, A., Smolin, N., Oleinikova, A., Geiger, A., and Winter, R. (2005). Thermal breaking of spanning water networks in the hydration shell of proteins. *The Journal of Chemical Physics* 123, 224905.
- Bruce Macdonald, H.E., Cave-Ayland, C., Ross, G.A., and Essex, J.W. (2018). Ligand Binding Free Energies with Adaptive Water Networks: Two-Dimensional Grand Canonical Alchemical Perturbations. *J. Chem. Theory Comput.*
- Busto, E., Gotor-Fernández, V., and Gotor, V. (2010). Hydrolases: catalytically promiscuous enzymes for non-conventional reactions in organic synthesis. *Chem. Soc. Rev.* 39, 4504–4523.
- Cabani, S., Gianni, P., Mollica, V., and Lepori, L. (1981). Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J Solution Chem* 10, 563–595.
- Casasnovas, R., Limongelli, V., Tiwary, P., Carloni, P., and Parrinello, M. (2017). Unbinding Kinetics of a p38 MAP Kinase Type II Inhibitor from Metadynamics Simulations. *J. Am. Chem. Soc.* 139, 4780–4788.
- Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry* 26, 1668–1688.

- Chaplin, M. (2006). Do we underestimate the importance of water in cell biology? *Nature Reviews Molecular Cell Biology* 7, 861–866.
- Chen, J.M., Xu, S.L., Wawrzak, Z., Basarab, G.S., and Jordan, D.B. (1998). Structure-Based Design of Potent Inhibitors of Scytalone Dehydratase: Displacement of a Water Molecule from the Active Site. *Biochemistry* 37, 17735–17744.
- Chen, X., Weber, I., and Harrison, R.W. (2008). Hydration water and bulk water in proteins have distinct properties in radial distributions calculated from 105 atomic resolution crystal structures. *J Phys Chem B* 112, 12073–12080.
- Chodera, J.D., Mobley, D.L., Shirts, M.R., Dixon, R.W., Branson, K., and Pande, V.S. (2011). Alchemical free energy methods for drug discovery: progress and challenges. *Current Opinion in Structural Biology* 21, 150–160.
- Claveria-Gimeno, R., Vega, S., Abian, O., and Velazquez-Campoy, A. (2017). A look at ligand binding thermodynamics in drug discovery. *Expert Opinion on Drug Discovery* 12, 363–377.
- Collins, M.D., Hummer, G., Quillin, M.L., Matthews, B.W., and Gruner, S.M. (2005). Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. *PNAS* 102, 16668–16671.
- Copeland, R.A., Pompliano, D.L., and Meek, T.D. (2006). Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* 5, 730–739.
- Cournia, Z., Allen, B., and Sherman, W. (2017). Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J Chem Inf Model* 57, 2911–2937.
- Cramer, C.J. (2005). *Essentials of Computational Chemistry: Theories and Models* (Wiley).
- Cramer, J., Krimmer, S.G., Heine, A., and Klebe, G. (2017). Paying the Price of Desolvation in Solvent-Exposed Protein Pockets: Impact of Distal Solubilizing Groups on Affinity and Binding Thermodynamics in a Series of Thermolysin Inhibitors. *J. Med. Chem.* 60, 5791–5799.
- Davies, N.G.M., Browne, H., Davis, B., Drysdale, M.J., Foloppe, N., Geoffrey, S., Gibbons, B., Hart, T., Hubbard, R., Jensen, M.R., et al. (2012). Targeting conserved water molecules: design of 4-aryl-5-cyanopyrrolo[2,3-d]pyrimidine Hsp90 inhibitors using fragment-based screening and structure-based optimization. *Bioorg. Med. Chem.* 20, 6770–6789.
- De Vivo, M., Masetti, M., Bottegoni, G., and Cavalli, A. (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 59, 4035–4061.
- DeChancie, J., and Houk, K.N. (2007). The Origins of Femtomolar Protein–Ligand Binding: Hydrogen Bond Cooperativity and Desolvation Energetics in the Biotin–(Strept)Avidin Binding Site. *J Am Chem Soc* 129, 5419–5429.
- Dill, K.A. (1990). Dominant forces in protein folding. *Biochemistry* 29, 7133–7155.
- Dror, R.O., Pan, A.C., Arlow, D.H., Borhani, D.W., Maragakis, P., Shan, Y., Xu, H., and Shaw, D.E. (2011). Pathway and mechanism of drug binding to G-protein-coupled receptors. *PNAS* 108, 13118–13123.
- Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V., and Mee, R.P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J Comput Aided Mol Des* 11, 425–445.

- Enyedy, I.J., and Egan, W.J. (2008). Can we use docking and scoring for hit-to-lead optimization? *J. Comput. Aided Mol. Des.* *22*, 161–168.
- Fadda, E., and Woods, R.J. (2011). On the Role of Water Models in Quantifying the Binding Free Energy of Highly Conserved Water Molecules in Proteins: The Case of Concanavalin A. *J. Chem. Theory Comput.* *7*, 3391–3398.
- Feig, M., MacKerell, Alexander D., and Brooks, C.L. (2003). Force Field Influence on the Observation of  $\pi$ -Helical Protein Structures in Molecular Dynamics Simulations. *J. Phys. Chem. B* *107*, 2831–2836.
- Feng, T., and Barakat, K. (2018). Molecular Dynamics Simulation and Prediction of Druggable Binding Sites. *Methods Mol. Biol.* *1762*, 87–103.
- Florová, P., Sklenovský, P., Banáš, P., and Otyepka, M. (2010). Explicit Water Models Affect the Specific Solvation and Dynamics of Unfolded Peptides While the Conformational Behavior and Flexibility of Folded Peptides Remain Intact. *J. Chem. Theory Comput.* *6*, 3569–3579.
- Forli, S., and Olson, A.J. (2012). A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J. Med. Chem.* *55*, 623–638.
- Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., et al. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* *47*, 1739–1749.
- Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C., and Mainz, D.T. (2006). Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* *49*, 6177–6196.
- García-Sosa, A.T., Mancera, R.L., and Dean, P.M. (2003). WaterScore: a novel method for distinguishing between bound and displaceable water molecules in the crystal structure of the binding site of protein-ligand complexes. *J Mol Model* *9*, 172–182.
- Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* *10*, 449–461.
- Gilson, M.K., Given, J.A., and Head, M.S. (1997). A new class of models for computing receptor-ligand binding affinities. *Chemistry & Biology* *4*, 87–92.
- Goodford, P.J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* *28*, 849–857.
- Grebner, C., Niebling, S., Schmuck, C., Schlücker, S., and Engels, B. (2013). Force field-based conformational searches: efficiency and performance for peptide receptor complexes. *Molecular Physics* *111*, 2489–2500.
- Guenot, J., and Kollman, P.A. (1992). Molecular dynamics studies of a DNA-binding protein: 2. An evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the *Escherichia coli* trp repressor. *Protein Sci* *1*, 1185–1205.
- van Gunsteren, W.F., and Mark, A.E. (1998). Validation of molecular dynamics simulation. *The Journal of Chemical Physics* *108*, 6109–6116.
- Gürsoy, O., and Smieško, M. (2017). Searching for bioactive conformations of drug-like ligands with current force fields: how good are we? *J Cheminform* *9*.

- Güssregen, S., Matter, H., Hessler, G., Lionta, E., Heil, J., and Kast, S.M. (2017). Thermodynamic Characterization of Hydration Sites from Integral Equation-Derived Free Energy Densities: Application to Protein Binding Sites and Ligand Series. *J. Chem. Inf. Model.* *57*, 1652–1666.
- Haider, K., Wickstrom, L., Ramsey, S., Gilson, M.K., and Kurtzman, T. (2016). Enthalpic Breakdown of Water Structure on Protein Active-Site Surfaces. *J. Phys. Chem. B* *120*, 8743–8756.
- Halgren, T.A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* *17*, 490–519.
- Halgren, T.A., and Nachbar, R.B. (1996). Merck molecular force field. IV. conformational energies and geometries for MMFF94. *Journal of Computational Chemistry* *17*, 587–615.
- Hamelberg, D., Mongan, J., and McCammon, J.A. (2004). Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* *120*, 11919–11929.
- Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., et al. (2016). OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* *12*, 281–296.
- Hart, T.N., and Read, R.J. (1992). A multiple-start Monte Carlo docking method. *Proteins* *13*, 206–222.
- Helma, C. (2006). Lazy structure-activity relationships (lazar) for the prediction of rodent carcinogenicity and Salmonella mutagenicity. *Mol. Divers.* *10*, 147–158.
- Hobbs, M.E., Jhon, M.S., and Eyring, H. (1966). The Dielectric Constant of Liquid Water and Various Forms of Ice According to Significant Structure Theory. *PNAS* *56*, 31–38.
- Hollingsworth, S.A., and Dror, R.O. (2018). Molecular Dynamics Simulation for All. *Neuron* *99*, 1129–1143.
- Homans, S.W. (2007). Water, water everywhere — except where it matters? *Drug Discovery Today* *12*, 534–539.
- Horbert, R., Pinchuk, B., Johannes, E., Schlosser, J., Schmidt, D., Cappel, D., Totzke, F., Schächtele, C., and Peifer, C. (2015). Optimization of Potent DFG-in Inhibitors of Platelet Derived Growth Factor Receptor $\beta$  (PDGF-R $\beta$ ) Guided by Water Thermodynamics. *J. Med. Chem.* *58*, 170–182.
- Hu, B., and Lill, M.A. (2014). WATsite: hydration site prediction program with PyMOL interface. *J Comput Chem* *35*, 1255–1260.
- Huang, N., Kalyanaraman, C., Irwin, J.J., and Jacobson, M.P. (2006). Physics-Based Scoring of Protein–Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening. *J. Chem. Inf. Model.* *46*, 243–253.
- Huey, R., Morris, G.M., Olson, A.J., and Goodsell, D.S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry* *28*, 1145–1152.
- Huggins, D.J. (2012). Benchmarking the thermodynamic analysis of water molecules around a model beta sheet. *J Comput Chem* *33*, 1383–1392.
- Huggins, D.J. (2014). Estimating Translational and Orientational Entropies Using the k-Nearest Neighbors Algorithm. *J. Chem. Theory Comput.* *10*, 3617–3625.
- Hummer, G. (2010). Molecular binding: Under water’s influence. *Nature Chemistry* *2*, 906–907.

- Jorgensen, W.L. (2004). The Many Roles of Computation in Drug Discovery. *Science* *303*, 1813–1818.
- Jorgensen, W.L., and Schyman, P. (2012). Treatment of Halogen Bonding in the OPLS-AA Force Field; Application to Potent Anti-HIV Agents. *J Chem Theory Comput* *8*, 3895–3801.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* *79*, 926–935.
- Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* *118*, 11225–11236.
- Karplus, M., and Kuriyan, J. (2005). Molecular dynamics and protein function. *PNAS* *102*, 6679–6685.
- Katz, B.A., Sprengeler, P.A., Luong, C., Verner, E., Elrod, K., Kirtley, M., Janc, J., Spencer, J.R., Breitenbucher, J.G., Hui, H., et al. (2001). Engineering inhibitors highly selective for the S1 sites of Ser190 trypsin-like serine protease drug targets. *Chemistry & Biology* *8*, 1107–1121.
- Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* *14*, 1–63.
- Kini, R.M., and Evans, H.J. (1991). Molecular modeling of proteins: a strategy for energy minimization by molecular mechanics in the AMBER force field. *J. Biomol. Struct. Dyn.* *9*, 475–488.
- Kitchen, D.B., Decornez, H., Furr, J.R., and Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery* *3*, 935–949.
- Klebe, G., and Böhm, H.J. (1997). Energetic and entropic factors determining binding affinity in protein-ligand complexes. *J. Recept. Signal Transduct. Res.* *17*, 459–473.
- Kontoyianni, M. (2017). Docking and Virtual Screening in Drug Discovery. *Methods Mol. Biol.* *1647*, 255–266.
- Korb, O., Stützle, T., and Exner, T.E. (2007). An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intell* *1*, 115–134.
- Kramer, C., Gedeck, P., and Meuwly, M. (2012). Atomic multipoles: Electrostatic potential fit, local reference axis systems, and conformational dependence. *J. Comput. Chem.* *33*, 1673–1688.
- Kramer, C., Bereau, T., Spinn, A., Liedl, K.R., Gedeck, P., and Meuwly, M. (2013). Deriving Static Atomic Multipoles from the Electrostatic Potential. *J. Chem. Inf. Model.* *53*, 3410–3417.
- Krimmer, S.G., Betz, M., Heine, A., and Klebe, G. (2014). Methyl, Ethyl, Propyl, Butyl: Futile But Not for Water, as the Correlation of Structure and Thermodynamic Signature Shows in a Congeneric Series of Thermolysin Inhibitors. *ChemMedChem* *9*, 833–846.
- Kung, P.-P., Sinnema, P.-J., Richardson, P., Hickey, M.J., Gajiwala, K.S., Wang, F., Huang, B., McClellan, G., Wang, J., Maegley, K., et al. (2011). Design strategies to target crystallographic waters applied to the Hsp90 molecular chaperone. *Bioorg. Med. Chem. Lett.* *21*, 3557–3562.
- Ladbury, J.E. (1996). Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chemistry & Biology* *3*, 973–980.

- Laio, A., and Parrinello, M. (2002). Escaping free-energy minima. *PNAS* *99*, 12562–12566.
- Lam, P.Y., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, Y.N., et al. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* *263*, 380–384.
- Lam, P.Y., Ru, Y., Jadhav, P.K., Aldrich, P.E., DeLucca, G.V., Eyermann, C.J., Chang, C.H., Emmett, G., Holler, E.R., Daneker, W.F., et al. (1996). Cyclic HIV protease inhibitors: synthesis, conformational analysis, P2/P2' structure-activity relationship, and molecular recognition of cyclic ureas. *J. Med. Chem.* *39*, 3514–3525.
- Lazaridis, T. (1998). Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* *102*, 3531–3541.
- Leach, A.R. (2001). *Molecular Modelling: Principles and Applications* (Pearson Education).
- Levoine, N., Calmels, T., Krief, S., Danvy, D., Berrebi-Bertrand, I., Lecomte, J.-M., Schwartz, J.-C., and Capet, M. (2011). Homology Model Versus X-ray Structure in Receptor-based Drug Design: A Retrospective Analysis with the Dopamine D3 Receptor. *ACS Med. Chem. Lett.* *2*, 293–297.
- Levy, Y., and Onuchic, J.N. (2004). Water and proteins: A love–hate relationship. *PNAS* *101*, 3325–3326.
- Li, Z., and Lazaridis, T. (2003). Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease. *J. Am. Chem. Soc.* *125*, 6636–6637.
- Li, Z., and Lazaridis, T. (2006). Thermodynamics of Buried Water Clusters at a Protein–Ligand Binding Interface. *J. Phys. Chem. B* *110*, 1464–1475.
- Li, Z., and Lazaridis, T. (2012). Computing the thermodynamic contributions of interfacial water. *Methods Mol. Biol.* *819*, 393–404.
- Li, Y., Han, L., Liu, Z., and Wang, R. (2014). Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J. Chem. Inf. Model.* *54*, 1717–1736.
- Lim, N.M., Wang, L., Abel, R., and Mobley, D.L. (2016). Sensitivity in Binding Free Energies Due to Protein Reorganization. *J. Chem. Theory Comput.* *12*, 4620–4631.
- Liu, J., and Wang, R. (2015). Classification of current scoring functions. *J Chem Inf Model* *55*, 475–482.
- Liu, K., and Kokubo, H. (2017). Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-docking Study. *J Chem Inf Model* *57*, 2514–2522.
- Liu, C., Wroblewski, S.T., Lin, J., Ahmed, G., Metzger, A., Wityak, J., Gillooly, K.M., Shuster, D.J., McIntyre, K.W., Pitt, S., et al. (2005a). 5-Cyanopyrimidine Derivatives as a Novel Class of Potent, Selective, and Orally Active Inhibitors of p38 $\alpha$  MAP Kinase. *J. Med. Chem.* *48*, 6261–6270.
- Liu, H.-Y., Kuntz, I.D., and Zou, X. (2004). Pairwise GB/SA Scoring Function for Structure-based Drug Design. *J. Phys. Chem. B* *108*, 5453–5462.
- Liu, J., Liu, B., Guo, G., Jing, Y., and Zhao, G. (2015). Discovery of novel androgen receptor antagonists: a hybrid approach of pharmacophore-based and docking-based virtual screening. *Anticancer Drugs* *26*, 747–753.

- Liu, J., Su, M., Liu, Z., Li, J., Li, Y., and Wang, R. (2017). Enhance the performance of current scoring functions with the aid of 3D protein-ligand interaction fingerprints. *BMC Bioinformatics* *18*, 343.
- Liu, L., Quillin, M.L., and Matthews, B.W. (2008). Use of experimental crystallographic phases to examine the hydration of polar and nonpolar cavities in T4 lysozyme. *PNAS* *105*, 14406–14411.
- Liu, P., Kim, B., Friesner, R.A., and Berne, B.J. (2005b). Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *PNAS* *102*, 13749–13754.
- Lu, H., and Tonge, P.J. (2010). Drug-target residence time: critical information for lead optimization. *Curr Opin Chem Biol* *14*, 467–474.
- Lu, Y., Wang, R., Yang, C.-Y., and Wang, S. (2007). Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. *J Chem Inf Model* *47*, 668–675.
- Lupyan, D., Abramov, Y.A., and Sherman, W. (2012). Close intramolecular sulfur-oxygen contacts: modified force field parameters for improved conformation generation. *J. Comput. Aided Mol. Des.* *26*, 1195–1205.
- Mackerell, A.D. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry* *25*, 1584–1604.
- MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* *102*, 3586–3616.
- Maheshwary, S., Patel, N., Sathyamurthy, N., Kulkarni, A.D., and Gadre, S.R. (2001). Structure and Stability of Water Clusters (H<sub>2</sub>O)<sub>n</sub>, n = 8–20: An Ab Initio Investigation. *J. Phys. Chem. A* *105*, 10525–10537.
- Mahoney, M.W., and Jorgensen, W.L. (2000). A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics* *112*, 8910–8922.
- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* *11*, 3696–3713.
- Malham, R., Johnstone, S., Bingham, R.J., Barratt, E., Phillips, S.E.V., Laughton, C.A., and Homans, S.W. (2005). Strong Solute–Solute Dispersive Interactions in a Protein–Ligand Complex. *J. Am. Chem. Soc.* *127*, 17061–17067.
- Martín-García, F., Papaleo, E., Gomez-Puertas, P., Boomsma, W., and Lindorff-Larsen, K. (2015). Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS One* *10*.
- Matthews, B.W., and Liu, L. (2009). A review about nothing: are apolar cavities in proteins really empty? *Protein Sci.* *18*, 494–502.
- Meng, E.C., Shoichet, B.K., and Kuntz, I.D. (1992). Automated docking with grid-based energy evaluation. *Journal of Computational Chemistry* *13*, 505–524.
- Meng, X.-Y., Zhang, H.-X., Mezei, M., and Cui, M. (2011). Molecular Docking: A powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* *7*, 146–157.



- Mennucci, B., Tomasi, J., Cammi, R., Cheeseman, J.R., Frisch, M.J., Devlin, F.J., Gabriel, S., and Stephens, P.J. (2002). Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *J. Phys. Chem. A* *106*, 6102–6113.
- Michel, J., Tirado-Rives, J., and Jorgensen, W.L. (2009). Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc.* *131*, 15403–15411.
- Michel, J., Henchman, R.H., Gerogiokas, G., Southey, M.W.Y., Mazanetz, M.P., and Law, R.J. (2014). Evaluation of Host–Guest Binding Thermodynamics of Model Cavities with Grid Cell Theory. *J. Chem. Theory Comput.* *10*, 4055–4068.
- Miller, B.R., McGee, T.D., Swails, J.M., Homeyer, N., Gohlke, H., and Roitberg, A.E. (2012). MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* *8*, 3314–3321.
- Mobley, D.L., Chodera, J.D., and Dill, K.A. (2007). Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. *J. Chem. Theory Comput.* *3*, 1231–1235.
- Mobley, D.L., Bayly, C.I., Cooper, M.D., Shirts, M.R., and Dill, K.A. (2009). Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* *5*, 350–358.
- Mondal, J., Friesner, R.A., and Berne, B.J. (2014). Role of Desolvation in Thermodynamics and Kinetics of Ligand Binding to a Kinase. *J. Chem. Theory Comput.* *10*, 5696–5705.
- Morozenko, A., Leontyev, I.V., and Stuchebrukhov, A.A. (2014). Dipole Moment and Binding Energy of Water in Proteins from Crystallographic Analysis. *J. Chem. Theory Comput.* *10*, 4618–4623.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* *19*, 1639–1662.
- Muegge, I. (2006). PMF scoring revisited. *J. Med. Chem.* *49*, 5895–5902.
- Murphy, R.B., Repasky, M.P., Greenwood, J.R., Tubert-Brohman, I., Jerome, S., Annabhimoju, R., Boyles, N.A., Schmitz, C.D., Abel, R., Farid, R., et al. (2016). WScore: A Flexible and Accurate Treatment of Explicit Water Molecules in Ligand–Receptor Docking. *J. Med. Chem.* *59*, 4364–4384.
- Mysinger, M.M., and Shoichet, B.K. (2010). Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* *50*, 1561–1573.
- Nair, P.C., and Miners, J.O. (2014). Molecular dynamics simulations: from structure function relationships to drug discovery. *In Silico Pharmacol.* *2*.
- Nakasako, M. (2004). Water–protein interactions from high–resolution protein crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences* *359*, 1191–1206.
- Nguyen, C.N., Kurtzman Young, T., and Gilson, M.K. (2012). Grid inhomogeneous solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril. *J. Chem. Phys.* *137*.

- Oostenbrink, C., Villa, A., Mark, A.E., and van Gunsteren, W.F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* *25*, 1656–1676.
- Osguthorpe, D.J., Sherman, W., and Hagler, A.T. (2012a). Generation of Receptor Structural Ensembles for Virtual Screening Using Binding Site Shape Analysis and Clustering. *Chemical Biology & Drug Design* *80*, 182–193.
- Osguthorpe, D.J., Sherman, W., and Hagler, A.T. (2012b). Exploring Protein Flexibility: Incorporating Structural Ensembles From Crystal Structures and Simulation into Virtual Screening Protocols. *J. Phys. Chem. B* *116*, 6952–6959.
- Otting, G., Liepinsh, E., and Wuthrich, K. (1991). Protein hydration in aqueous solution. *Science* *254*, 974–980.
- Papoian, G.A., Ulander, J., Eastwood, M.P., Luthey-Schulten, Z., and Wolynes, P.G. (2004). Water in protein structure prediction. *PNAS* *101*, 3352–3357.
- Pearlstein, R.A., Sherman, W., and Abel, R. (2013). Contributions of water transfer energy to protein-ligand association and dissociation barriers: Watermap analysis of a series of p38 $\alpha$  MAP kinase inhibitors. *Proteins: Structure, Function, and Bioinformatics* *81*, 1509–1526.
- Pérez, C., Zaleski, D.P., Seifert, N.A., Temelso, B., Shields, G.C., Kisiel, Z., and Pate, B.H. (2014). Hydrogen Bond Cooperativity and the Three-Dimensional Structures of Water Nonamers and Decamers. *Angewandte Chemie International Edition* *53*, 14368–14372.
- Ponder, J.W., Wu, C., Ren, P., Pande, V.S., Chodera, J.D., Schnieders, M.J., Haque, I., Mobley, D.L., Lambrecht, D.S., DiStasio, R.A., et al. (2010). Current Status of the AMOEBA Polarizable Force Field. *J Phys Chem B* *114*, 2549–2564.
- Qi, R., Wang, L.-P., Wang, Q., Pande, V.S., and Ren, P. (2015). United polarizable multipole water model for molecular mechanics simulation. *The Journal of Chemical Physics* *143*, 014504.
- Raies, A.B., and Bajic, V.B. (2016). In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip Rev Comput Mol Sci* *6*, 147–172.
- Ramsey, S., Nguyen, C., Salomon-Ferrer, R., Walker, R.C., Gilson, M.K., and Kurtzman, T. (2016). Solvation thermodynamic mapping of molecular surfaces in AmberTools: GIST. *J Comput Chem* *37*, 2029–2037.
- Rendine, S., Pieraccini, S., Forni, A., and Sironi, M. (2011). Halogen bonding in ligand –receptor systems in the framework of classical force fields. *Physical Chemistry Chemical Physics* *13*, 19508–19516.
- Roberts, B.C., and Mancera, R.L. (2008). Ligand–Protein Docking with Water Molecules. *J. Chem. Inf. Model.* *48*, 397–408.
- Robinson, D.D., Sherman, W., and Farid, R. (2010). Understanding kinase selectivity through energetic analysis of binding site waters. *ChemMedChem* *5*, 618–627.
- Sager, C.P., Eriş, D., Smieško, M., Hevey, R., and Ernst, B. (2017). What contributes to an effective mannose recognition domain? *Beilstein Journal of Organic Chemistry* *13*, 2584–2595.
- Schmidtke, P., Luque, F.J., Murray, J.B., and Barril, X. (2011). Shielded Hydrogen Bonds as Structural Determinants of Binding Kinetics: Application in Drug Design. *J. Am. Chem. Soc.* *133*, 18903–18910.

- Schneider, N., Lange, G., Hindle, S., Klein, R., and Rarey, M. (2013). A consistent description of HYdrogen bond and DEhydration energies in protein-ligand complexes: methods behind the HYDE scoring function. *J. Comput. Aided Mol. Des.* *27*, 15–29.
- Schuetz, D.A., de Witte, W.E.A., Wong, Y.C., Knasmueller, B., Richter, L., Kokh, D.B., Sadiq, S.K., Bosma, R., Nederpelt, I., Heitman, L.H., et al. (2017). Kinetics for Drug Discovery: an industry-driven effort to target drug residence time. *Drug Discov. Today* *22*, 896–911.
- Schuetz, D.A., Richter, L., Amaral, M., Grandits, M., Grädler, U., Musil, D., Buchstaller, H.-P., Eggenweiler, H.-M., Frech, M., and Ecker, G.F. (2018). Ligand Desolvation Steers On-Rate and Impacts Drug Residence Time of Heat Shock Protein 90 (Hsp90) Inhibitors. *J. Med. Chem.* *61*, 4397–4411.
- Setny, P., Baron, R., Michael Kekenos-Huskey, P., McCammon, J.A., and Dzubiella, J. (2013). Solvent fluctuations in hydrophobic cavity–ligand binding kinetics. *Proc Natl Acad Sci U S A* *110*, 1197–1202.
- Sindhikara, D.J., and Hirata, F. (2013). Analysis of Biomolecular Solvation Sites by 3D-RISM Theory. *J. Phys. Chem. B* *117*, 6718–6723.
- Singh, N., and Warshel, A. (2010). Absolute Binding Free Energy Calculations: On the Accuracy of Computational Scoring of Protein-ligand Interactions. *Proteins* *78*, 1705–1723.
- Śledź, P., and Caflisch, A. (2018). Protein structure-based drug design: from docking to molecular dynamics. *Current Opinion in Structural Biology* *48*, 93–102.
- Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe, E.W. (2014). Computational Methods in Drug Discovery. *Pharmacol Rev* *66*, 334–395.
- Smith, C.R., Dougan, D.R., Komandla, M., Kanouni, T., Knight, B., Lawson, J.D., Sabat, M., Taylor, E.R., Vu, P., and Wyrick, C. (2015). Fragment-Based Discovery of a Small Molecule Inhibitor of Bruton's Tyrosine Kinase. *J. Med. Chem.* *58*, 5437–5444.
- Snyder, P.W., Lockett, M.R., Moustakas, D.T., and Whitesides, G.M. (2014). Is it the shape of the cavity, or the shape of the water in the cavity? *Eur. Phys. J. Spec. Top.* *223*, 853–891.
- Spyrakis, F., Ahmed, M.H., Bayden, A.S., Cozzini, P., Mozzarelli, A., and Kellogg, G.E. (2017). The Roles of Water in the Protein Matrix: A Largely Untapped Resource for Drug Discovery. *J. Med. Chem.* *60*, 6781–6827.
- Sridhar, A., Ross, G.A., and Biggin, P.C. (2017). Waterdock 2.0: Water placement prediction for Holo-structures with a pymol plugin. *PLOS ONE* *12*, e0172743.
- Sushko, I., Salmina, E., Potemkin, V.A., Poda, G., and Tetko, I.V. (2012). ToxAlerts: A Web Server of Structural Alerts for Toxic Chemicals and Compounds with Potential Adverse Reactions. *J Chem Inf Model* *52*, 2310–2316.
- Tang, Y., Zhu, W., Chen, K., and Jiang, H. (2006). New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. *Drug Discovery Today: Technologies* *3*, 307–313.
- Thilagavathi, R., and Mancera, R.L. (2010). Ligand–Protein Cross-Docking with Water Molecules. *J. Chem. Inf. Model.* *50*, 415–421.
- Trisciuzzi, D., Alberga, D., Mansouri, K., Judson, R., Novellino, E., Mangiatordi, G.F., and Nicolotti, O. (2017). Predictive Structure-Based Toxicology Approaches To Assess the Androgenic Potential of Chemicals. *J. Chem. Inf. Model.* *57*, 2874–2884.

- Uehara, S., and Tanaka, S. (2016). AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* *21*.
- Vaitheeswaran, S., Yin, H., Rasaiah, J.C., and Hummer, G. (2004). Water clusters in nonpolar cavities. *PNAS* *101*, 17002–17005.
- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., et al. (2010). CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* *31*, 671–690.
- Vedani, A. (1988). YETI: An interactive molecular mechanics program for small-molecule protein complexes. *Journal of Computational Chemistry* *9*, 269–280.
- Vedani, A., and Dunitz, J.D. (1985). Lone-pair directionality in hydrogen-bond potential functions for molecular mechanics calculations: the inhibition of human carbonic anhydrase II by sulfonamides. *J. Am. Chem. Soc.* *107*, 7653–7658.
- Vedani, A., Dobler, M., and Smieško, M. (2012). VirtualToxLab - a platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicol. Appl. Pharmacol.* *261*, 142–153.
- Vedani, A., Dobler, M., Hu, Z., and Smieško, M. (2015). OpenVirtualToxLab—A platform for generating and exchanging in silico toxicity data. *Toxicology Letters* *232*, 519–532.
- Verdonk, M.L., Chessari, G., Cole, J.C., Hartshorn, M.J., Murray, C.W., Nissink, J.W.M., Taylor, R.D., and Taylor, R. (2005). Modeling Water Molecules in Protein–Ligand Docking Using GOLD. *J. Med. Chem.* *48*, 6504–6515.
- Verlet, L. (1967). Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* *159*, 98–103.
- Wade, R.C., Mazor, M.H., McCammon, J.A., and Quiocho, F.A. (1991). A molecular dynamics study of thermodynamic and structural aspects of the hydration of cavities in proteins. *Biopolymers* *31*, 919–931.
- Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry* *25*, 1157–1174.
- Wang, L., Berne, B.J., and Friesner, R.A. (2011a). Ligand binding to protein-binding pockets with wet and dry regions. *PNAS* *108*, 1326–1330.
- Wang, L., Friesner, R.A., and Berne, B.J. (2011b). Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* *115*, 9431–9438.
- Wang, L., Wu, Y., Deng, Y., Kim, B., Pierce, L., Krilov, G., Lupyan, D., Robinson, S., Dahlgren, M.K., Greenwood, J., et al. (2015). Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* *137*, 2695–2703.
- Wang, L.-P., Martinez, T.J., and Pande, V.S. (2014). Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* *5*, 1885–1891.
- Wang, R., Lu, Y., and Wang, S. (2003). Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *Journal of Medicinal Chemistry* *46*, 2287–2303.

- Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S., and Hou, T. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys Chem Chem Phys* *18*, 12964–12975.
- Warshel, A., Kato, M., and Pisiakov, A.V. (2007). Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* *3*, 2034–2045.
- Watts, C.R., Gregory, A., Frisbie, C., and Lovas, S. (2018). Effects of force fields on the conformational and dynamic properties of amyloid  $\beta$ (1-40) dimer explored by replica exchange molecular dynamics simulations. *Proteins* *86*, 279–300.
- Wissner, A., Berger, D.M., Boschelli, D.H., Floyd, M.B., Greenberger, L.M., Gruber, B.C., Johnson, B.D., Mamuya, N., Nilakantan, R., Reich, M.F., et al. (2000). 4-Anilino-6,7-dialkoxyquinoline-3-carbonitrile Inhibitors of Epidermal Growth Factor Receptor Kinase and Their Bioisosteric Relationship to the 4-Anilino-6,7-dialkoxyquinazoline Inhibitors. *J. Med. Chem.* *43*, 3244–3256.
- Xun, S., Jiang, F., and Wu, Y.-D. (2015). Significant Refinement of Protein Structure Models Using a Residue-Specific Force Field. *J. Chem. Theory Comput.* *11*, 1949–1956.
- Yan, X.C., Robertson, M.J., Tirado-Rives, J., and Jorgensen, W.L. (2017). Improved Description of Sulfur Charge Anisotropy in OPLS Force Fields: Model Development and Parameterization. *J. Phys. Chem. B* *121*, 6626–6636.
- Yang, H., Sun, L., Li, W., Liu, G., and Tang, Y. (2018). In Silico Prediction of Chemical Toxicity for Drug Design Using Machine Learning Methods and Structural Alerts. *Front Chem* *6*.
- Yin, H., Hummer, G., and Rasaiah, J.C. (2007). Metastable Water Clusters in the Nonpolar Cavities of the Thermostable Protein Tetrabrachion. *J. Am. Chem. Soc.* *129*, 7369–7377.
- Young, T., Abel, R., Kim, B., Berne, B.J., and Friesner, R.A. (2007). Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *PNAS* *104*, 808–813.
- Young, T., Hua, L., Huang, X., Abel, R., Friesner, R., and Berne, B.J. (2010). Dewetting transitions in protein cavities. *Proteins: Structure, Function, and Bioinformatics* *78*, 1856–1869.
- Zou, X., Yaxiong, and Kuntz, I.D. (1999). Inclusion of Solvation in Ligand Binding Free Energy Calculations Using the Generalized-Born Model. *J. Am. Chem. Soc.* *121*, 8033–8043.
- Zwanzig, R.W. (1954). High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics* *22*, 1420–1426.
- Schrödinger Release 2018-4: WaterMap, Schrödinger, LLC, New York, NY, 2018.

## 6. Publication List

Hu, Z., **Wahl, J.**, Hamburger, M., and Vedani, A. (2016). Molecular mechanisms of endocrine and metabolic disruption: An in silico study on antitrypanosomal natural products and some derivatives. *Toxicol. Lett.* 252, 29–41.

**Wahl, J.**, and Smiesko, M. (2018). Thermodynamic Insights into Effects of Water Displacement and Rearrangement upon Ligand Modification using Molecular Dynamics Simulations. *ChemMedChem*.

**Wahl, J.**, and Smieško, M. (2018). Endocrine Disruption at the Androgen Receptor: Employing Molecular Dynamics and Docking for Improved Virtual Screening and Toxicity Prediction. *International Journal of Molecular Sciences* 19, 1784.

## 7. Curriculum Vitae



# Joel Wahl

## Personal Info

Address:  
St. Alban-Anlage 63  
4052 Basel

Phone:  
+41 079 478 82 21

Mail:  
[joel.c.wahl@gmail.com](mailto:joel.c.wahl@gmail.com)

Linkedin:  
<https://www.linkedin.com/in/joel-wahl-041277106>

Citizenship:  
Swiss

## Skills

### Programming:

- Java
- Python:
  - Numpy
  - Scipy
  - Scikit

### Science:

- Molecular Simulations
- Molecular Docking
- Structure-Based Drug Design
- Virtual Screening

### Scientific Software:

- Schrödinger Small-Molecule Drug Discovery Suite
- Cambridge Crystallographic Data Centre Suite

*Trained chemist with experience and interest in scientific programming and computer-aided drug design. Enthusiastic about solving problems at the interface of software engineering and science by providing tailored software solutions for researchers in the pharmaceutical industry. Proven knowledge in developing and applying state-of-the-art tools for ligand-based and structure-based drug design. Highly-motivated team player that enjoys collaborations with scientists from different fields to achieve common goals.*

## Experience

- 01/2019 – present:** **Idorsia Pharmaceuticals (20%)**  
Scientific Software Engineer, developing algorithms and computational tools for drug discovery
- 06/2017 – 12/2018:** **Idorsia Pharmaceuticals (20%)**  
Assistant in the Scientific Computing Group. Developing a tool for virtual compound screening based on molecular shape. Familiarizing with the in-house Java libraries.
- 01/2015 – 06/2017 :** **Actelion Pharmaceuticals (20%)**  
Assistant in the Research Information Management group. Java programming and improvement of force field parameters for a more accurate description of chemical space. Maintenance and extension of a database containing GPCR ligands.
- 04/2015 – Present:** **University of Basel (80%)**  
Research Assistant in the group of Molecular Modeling at the department of Pharmaceutical Sciences.

## Education

- 04/2015 – present:** **University of Basel**  
PhD in Pharmaceutical Sciences,  
Supervisor: Dr. Martin Smiesko  
Subject: Investigating the Role of Water in Protein-Ligand Binding Processes
- 09/2012 – 04/2014 :** Master of Science in Chemistry, **University of Basel**
- 09/2009 – 04/2012 :** Bachelor of Science in Chemistry, **University of Basel**

