American Journal of Modeling and Optimization, 2015, Vol. 3, No. 3, 68-86
Available online at http://pubs.sciepub.com/ajmo/3/3/2

DOI:10.12691/ajmo-3-3-2

How Important is to Account for Water When Modeling Biomolecular Complexes?

María J. R. Yunta*

Departamento de Química Orgánica I, Facultad de Química, Universidad Complutense, Madrid, Spain
*Corresponding author: mjryun@ucm.es

Abstract Taking account of the presence of water molecules is sometimes crucial for free energy calculations to predict binding ability of molecules to receptors, a main subject for drug design. This mini-review seeks to identify the importance of knowing the influence of these molecules in such studies to better achieve correct predictions for drug candidates. Participation of some water molecules need to be considered in docking studies although they have been usually neglected.

Keywords: *molecular modeling, protonation state, computer aided drug design*

Cite This Article: María J. R. Yunta, "How Important is to Account for Water When Modeling Biomolecular Complexes?" *American Journal of Modeling and Optimization*, vol. 3, no. 3 (2015): 68-86. doi: 10.12691/ajmo-3-3-2.

1. Introduction

One common technique used to identify *in silico* the precursors of potential drugs is virtual screening. Computer-aided structure-based methods are aimed at predicting the binding mode of a ligand in the binding site of a protein or any molecular target and at obtaining an estimate of the binding affinity. These methods involve two computational steps: docking and scoring. In the docking step, multiple protein-ligand configurations, called *poses*, are generated. Then, a scoring function is used to calculate the affinity between the receptor and the ligand for each pose. In addition, if multiple ligands are docked, their binding free energies need to be ranked accurately.

Scoring functions can be grouped into three classes: force-field-based, knowledge-based, and empirical scoring functions (see refs 1 and 2 for a review). Force field-based scoring functions apply classical molecular mechanics energy functions. They approximate the binding free energy of protein-ligand complexes by a sum of van der Waals and electrostatic interactions.

A docking program is used to predict the binding pose and energy of a small-molecule model within a selected receptor binding pocket. Traditionally, many ligand models, typically taken from a database of compounds that can be easily synthesized or commercially purchased, are docked into a single static receptor structure, often obtained from NMR or X-ray crystallography. The best predicted ligands are subsequently tested experimentally to confirm binding.

Unfortunately, traditional docking relying on a single receptor structure is problematic. Some legitimate ligands may indeed bind to the single structure selected, but in reality most receptor binding pockets have many valid conformational states, any one of which may be druggable. In a traditional virtual screen, true ligands are often discarded because they in fact bind to receptor conformations that differ markedly from that of the single static structure chosen.

A quantitative understanding of the noncovalent interaction between a ligand and its receptor is a central goal for computational chemistry, and is the route to a deeper appreciation of molecular biochemistry and structure-based drug design. The association of a small organic molecule with a protein frequently induces changes in the conformation of both partners in order to minimise the free energy of association [3]. A measure of strain can be tolerated in the ligand on binding as it seeks to optimise all its interactions with the protein receptor, such as hydrogen bonds and van der Waals contacts. It has been estimated that over 60% of ligands bind in a strained conformation [4], although the level of strain does not appear too great: for example, it was found for a set of 99 drug-like molecules that the majority of bioactive conformations were within 2 kcal/mol of a local energy minimum [5].

A key objective of computational structure-based drug design is the prediction of the structure of protein-ligand complexes. In general, the scoring function of any docking engine is designed to be fast and robust. However, to achieve this goal, important details of the binding process are approximated or even completely ignored. This can be justified especially when one is dealing with virtual libraries of thousands or even millions of compounds to be in silico screened. The rigid receptor approach can be safely pursued in closely related structures where it can perform appropriately, but for cases where ligand sets are diverse or where the protein is flexible, methods that can partially or fully account for protein flexibility may be more suitable.

If a high resolution structure of the receptor is available, and the receptor structure does not change substantially upon ligand binding, the problem can often be reduced to

docking the flexible ligand in the environment of the rigid receptor. When rigid receptor docking fails, one of the simplest approaches is to reduce the van der Waals radii of the protein and/or ligand atoms or delete side chains of residues predicted to be flexible, thus potentially eliminating close contacts [6,7]. However, while this approach may yield the correct ligand binding mode, it may not give insights into specific protein/ligand interactions since the conformation of key residues in the binding site may be inconsistent with the correct ligand structure in the receptor. It is also likely to lead to false positives in virtual screening experiments, as the binding site will effectively increase in size. It is clear that in order to tackle the full protein/ligand structure prediction problem in a robust and accurate manner, it is essential to allow both the structure of the protein and ligand to reorganize. From a computational point of view, this is substantially more challenging than rigid receptor docking, as it involves many more degrees of freedom. There are a number of ways one could account for both ligand and receptor flexibility. In theory, the correct structure of the protein-ligand complex can be generated by running explicit solvent molecular dynamics simulations, starting from an arbitrary initial guess [8,9,10].

2. Interactions Involved in Molecular Recognition and Binding Affinity.

Knowledge-based scoring functions represent the binding affinity as a sum of protein-ligand atom pair interactions. These potentials are derived from the protein-ligand complexes with known structures, where probability distributions of interatomic distances are converted into distance-dependent interaction free energies of protein-ligand atom pairs using the “inverse” Boltzmann law [11].

Molecular recognition between drugs and their receptors is guided by the nature of intermolecular interactions, such as hydrogen bonds, heteroatom electrostatic interactions, π - π interactions, and van der Waals forces. Depending on the specific chemical donor-acceptor combination, and the details of the contact geometry, all of these interactions influence biological activity. Therefore, knowledge of intermolecular interactions and their geometric characteristics enables one to design and manipulate molecular systems, which can be applied in the field of rational drug design. Understanding the relationship between drug structures and biological activities forms the basis for the design of drugs. When the structural group(s) on the drug molecule that interacts with the target is known, structural modifications can be made to increase the affinity towards the desired target, decrease the affinity to an undesired target, alter the drug’s ability to cross a lipid membrane, and so on. Structure-based drug design is perhaps the most elegant approach for discovering compounds exhibiting high specificity and efficacy. In reality, however, drug targets are very complex and this approach has had only limited utility. However, a number of recent successful drugs have in part or in whole emerged from a structure-based research approach.

The three-dimensional crystal structure of a molecule is a free-energy minimum resulting from the optimization of attractive and repulsive intermolecular interactions with varying strengths, directional preferences, and distance-dependence properties. Intermolecular interactions in organic compounds are of two types: isotropic medium-range forces that define the shape, size, and close packing; anisotropic long-range forces that are electrostatic and include hydrogen bonds and heteroatom interactions. The observed three-dimensional architecture in the crystal is the result of interplay between the isotropic van der Waals forces and the anisotropic hydrogen-bond interactions. The distinction between hydrogen bonds and van der Waals interactions lies in their orientational and angular attributes [12].

The nature of intermolecular interactions that mediate molecular recognition for all systems are the same: strong O-H...O, N-H...O, O-H...N hydrogen bonds; weak C-H...O, C-H...N hydrogen bonds; heteroatom interactions X...X, X...O (X = Br, I); π-π interactions; and van der Waals forces [13]. Hydrogen bonds, regarded as the strongest and most directional of intermolecular interactions, have been widely exploited in many fields [14].

At optimal geometry, van der Waals interactions contribute some tenths of a kcal/mol to the hydrogen bond energy whereas electrostatic interaction reduces with increasing distance and with reducing dipole moments or charges involved. For strong donors like O-H or N-H, the electrostatic component is the dominant one, whereas for weakly polarized donor groups like C-H the magnitude of the electrostatic component resembles to van der Waals contribution [15,16].

Hydrogen bonding is a well-known classical structural phenomenon [17]. Knowledge of weak intermolecular interactions enables one to design and manipulate molecular systems and this can be applied in the fields of rational drug design, crystal engineering, supramolecular chemistry, and physical organic chemistry [18]. Also, these secondary interactions have ramifications in the systematic design of new materials possessing novel chemical, magnetic, optical, or electronic properties [19]. Critical assessment of the weak intermolecular interactions is a must as these may exert a substantial effect when added together. In this context, the acceptor capabilities of halogen atoms are controversial and noteworthy [20]. A second type of weak hydrogen bond established in recent years is the hydrogen bond with π acceptors [15].

One of the less known but significant weak interactions in nature is the interaction involving π systems [21] wherein phenyl rings, various heterocycles, C≡C, C=C, and other π-bonded moieties are involved. Two types of interactions can be categorized with π systems: one is a hydrogen bond where the π system acts as an acceptor and the other is the interaction between the π systems (π-π interactions). In this context, energy calculations have shown that there is significant interaction between a hydrogen bond donor and the center of a benzene ring, which is about half the strength of a normal hydrogen bond and contributes approximately 3 kcal/mol of stabilizing enthalpy. The energetics and consequently the chances of the occurrence of this type of interaction are enhanced if the aromatic system contains nitrogen atoms that magnify the π-electron density.

3. The Role of Water Molecules.

Water molecules play a critical role in the formation of protein-ligand interactions and are a significant determinant in the binding free energy. Changes in hydration free energy during complex formation are a crucial element of binding free energies [22,23,24]. With the use of methods for the prediction of binding free energies becoming commonplace in the field of drug design, there is a need for solvation methods that are both quick and highly accurate [25,26,27]. However, it is still challenging to predict accurately not only where water molecules prefer to bind, but also which of those water molecules might be displaceable. The latter is often seen as a route to optimizing affinity of potential drug candidates. For example, using a protocol called WaterDock, the freely available AutoDock Vina tool can be used to predict accurately the binding sites of water molecules [28].

Solvation has been usually taken into account using a distance dependent dielectric function, although solvent models based on continuum electrostatics have also been developed [29,30,31]. Although the use of a continuum solvent model is an improvement over accounting only for direct protein-ligand interactions, as they suggest practical ways to calculate free energies of macromolecular conformations taking into account equilibrium interactions with water solvent an protons, in many cases it is not enough. The continuum solvent models do not account for two explicit water molecule contributions. First, some water molecules serve to bridge interactions through hydrogen bonding between the ligand and protein and therefore affect the binding affinity. Second, some well-defined water molecules can be displaced by the ligand and significantly contributes to the binding affinity due to the increase in translational and rotational entropy upon release.

Explicit solvation models such as free energy perturbation (FEP), thermodynamic integration (TI) [22,32] and the faster linear interaction energy models (LIE) [33,34], offer detail on the distinct nature of water around the solute and are transferable across a wide range of datasets. The caveat of these solvation models for their application in the field of drug design is their lack of throughput. Implicit solvation models offer a quicker alternative to explicit models by replacing the individual water molecules with a continuous medium [35,36,37,38]. For small organic molecules, the loss of molecular detail of the solvent results in relatively small differences between hydration free energy prediction accuracies calculated with implicit solvent models relative to the explicit treatment [39,40,41,42,43]. When applied to complex biomolecular systems, this loss of detail may become problematic at locations where water does not behave as a continuous medium, for example the individual water molecules ordered in concave pockets at the surface of proteins [25,44].

As expected, new docking strategies considering explicit water molecules, partial protein flexibility, and rescoring of docking poses have already been found to improve binding mode prediction of CYP ligands. However, inclusion of explicit water molecules and especially the inclusion of flexibility for the whole enzyme lead to significantly higher complexity. Without proper algorithmic treatment of this complexity, this can lead to a final lower reliability of the docking results. The combination of docking with MD simulations to improve docking predictions and to explore conformational flexibility of substrates and CYP enzymes to a greater extent, as well as the explicit inclusion of water, is seen in an increasing number of studies and shows much promise [45].

At the same time, ordered water molecules play an important role in protein-ligand recognition, either being displaced on ligand binding or bridging groups to stabilize the complex [46].In a recent survey of high-resolution structures, over 85% of complexes had one or more water molecules bridging protein and ligand, with an average of 3.5 per complex [47]. Since the identity of the mediating waters can change from ligand to ligand and since many of the waters observed in an apo-structure are displaced by ligand binding, predicting the role of a particular ordered water molecule to ligand binding remains challenging [26]. The problem of treating ordered waters is acute in molecular docking, which relies on rapid evaluations of discrete states. These rise exponentially with the number of water molecules sampled. Also, it is rarely clear which waters should be treated as displaceable and which should be treated as fixed, despite efforts to categorize them based on environment or crystallographic observation (e.g., thermal factors) [47,48,49].

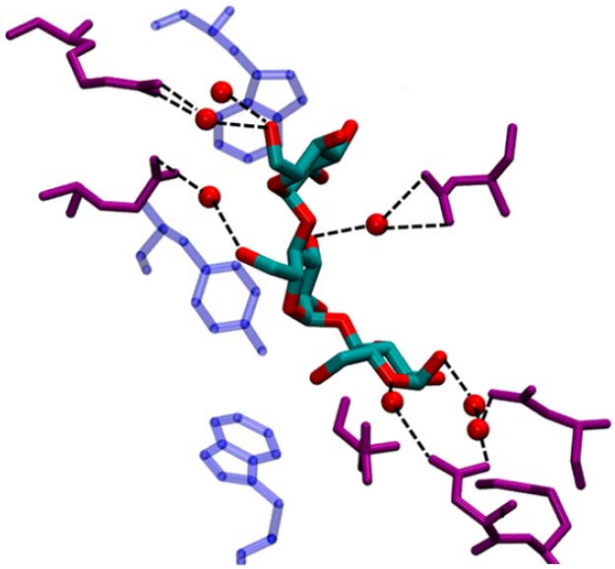

Figure 1. Water molecules mediating in protein-drug interactions [50]

4. Water and Protonation Estates

The protonation of the ligand molecule and the protein binding site has a significant influence on the results obtained by protein-ligand docking. Due to the inability of X-ray crystallography to resolve the hydrogen atom in protein and protein complex structures, the correct protonation for the protein and the ligand has to be assigned on a theoretical basis before the structures can be used. Because of the local environment inside the binding site and because of the influence of the ligand and the protein onto each other, the ligand protonation can differ from the protonation one would expect for the ligand in solution under physiological conditions. Hence for

protein-ligand docking different proton states of the ligand have to be taken into account.

Additionally, the importance of the preparation of the protein and especially the ligand structure and its influence on the docking results has been stressed in several studies. As the position of hydrogen atoms cannot be determined experimentally by X-ray crystallography, protonation and tautomeric states must be predicted on a theoretical basis, and the assignment of atomic hybridization and bond orders are not always straightforward. For scoring functions, which take actual hydrogen positions into account for the calculation of hydrogen-bonding strengths, the correct placement of these atoms is often the key to identifying the correct structure with a docking approach [51].

The Chothia-Janin model makes clear that the net contribution of non-covalent interactions, even if zero, must not be ignored because interactions determine the specificity of the complex. A highly specific interaction must reconcile with three criteria, all concerning interface complementarity [52]:

a) Complementarity of ions. If not all charged groups form salt bridges in the interface, the subunit association would require an ionic bond to the solvent (2-6 kcal mol^{-1}) to be broken and, therefore, would highly destabilize the protein-protein complex.

b) Complementarity of hydrogen bonds. A hydrogen bond that is not satisfied within the protein-protein interface would result in a large change in free energy (0.5-6 kcal mol^{-1}) [53].

c) Steric complementarity. Although van der Waals interactions are weak in nature, the number of atoms in the interface is large, and therefore they contribute to the specificity in a non-negligible manner.

5. Water and Binding

Besides all the preceding, there is often a difference between the bound and unbound ligand conformations, and the literature also suggests that while the predicted low energy solvated structure is often the bioactive this is not always true [4,54].

An often neglected aspect in rational drug design is the ubiquitous presence of water inside and around the receptor binding cavity and, in some instances, the free-energy cost of displacing ordered or partially ordered water molecules involved in the residual solvation of active site residues should be considered carefully [44,55,56,57]. Computational methods are being devised to calculate the energy cost of displacing clusters of rotationally and translationally hindered water molecules present in binding site cavities accessible to water [28,58,59,60,61,62].

When a ligand binds to a protein, water molecules surrounding the ligand and protein need to rearrange. These rearrangements contribute to the binding affinity of the ligand. Most water molecules building the hydration shell around the ligand and in the protein active site need to be either removed or displaced. Paying attention to the thermodynamics of these phenomena, for which no relevant experimental data are available, has been done in early theoretical studies of protein folding and protein association [63]. Very recently, molecular dynamics simulation (MDS) has further contributed to our understanding by showing that water molecules in the first hydration layer around a receptor surface are less thermodynamically stable than bulk water [55]. The process of transferring water from the bound state to bulk solvent is associated with a favorable free energy [44,55,58]. This, however, does not hold for water molecules that are deeply buried and tightly bound in crevasses at the surface or inside the active site of proteins. The maximum entropy change of releasing such a water molecule into bulk water is evaluated around 2 $kcal.mol^{-1}$; a gain which is unfavorably outweighed by the loss of enthalpy due to the breaking of hydrogen bonds with the receptor [57]. Because they are highly ordered (immobilized), these water molecules are seen in X-ray crystal structures, and in many cases ligand binding is often mediated by such water molecules [28]. Attempting to displace them with experimental ligands may not only be difficult to achieve but also have unpredictable unfavorable effects [57,64,65].

However, when a ligand is binding to a hydrophobic cleft at the surface or inside the protein the enthalpic contribution to the free energy of binding will increase because of favorable interactions between the ligand and protein. Simultaneously, water molecules that occupy the hydrophobic cleft will be released into the solvent, resulting in a favorable increase in entropy [55].

The interplay between water and the gate-constituting amino acids to the binding site has been analyzed for mechanosensitive channel (MSC) in order to understand the underlying molecular and biophysical mechanisms of the mechano-gating process from force sensing to gate opening, focussing on the investigation of the role of water during the process [66] finding that the presence of water molecules in the gate leads to gate expansion, a point usually neglected in binding studies, although a similar effect would probably influence the ligand aproximation to a buried binding site. Influence of the number of water molecules inside the conducting path for the rapid clearance of drugs from human cells has also been found in studies about MRP1 function [67]

Solvation plays an important role in molecular recognition, but appropriate treatment of solvent effects in scoring functions still remains a major challenge. In many scoring functions these effects are considered only partially, neglected altogether, or included indirectly, as in some knowledge-based scoring schemes. A more rigorous way of treating solvation effects in the estimation of binding affinities has become known as MM-PBSA or MM-GBSA scoring, where MM stands for molecular mechanics, PB and GB for Poisson-Boltzmann and Generalized Born, respectively, and SA for solvent-accessible surface area. The MM-PBSA approach has been pioneered by Kollman et al., and its basis is a thermodynamic cycle for complex formation in aqueous solution [68,69]. The key element is that the electrostatics of (de)solvation and ligand-receptor interactions is treated in a more sophisticated manner using PB or GB instead of simple Coulomb-based terms. The (de)solvation process can be divided into polar and apolar contributions. The associated energies, the polar free energy of solvation and the apolar free energy of solvation, are calculated with the PB or GB approach and using an expression containing a surface area term, respectively [70]. Recently, first

applications of MM-PBSA as a more sophisticated scoring function in the context of SBVS have become known. In contrast to earlier applications, where it was combined with molecular dynamics (MD) simulations, the recent examples demonstrate its value also for 'snapshot scoring,' i.e., the evaluation of the MM-PB(GB)SA expression for one or a few poses per ligand. These poses had been generated using a conventional docking program and not by means of a lengthy MD simulation. Researchers at Wyeth [71] and SGX Pharmaceuticals [72] presented evidence that MM-PBSA scoring can lead to an improvement compared to conventional scoring. It was shown that, given a number of precomputed poses per ligand, re-ranking of the poses with MM-PBSA leads to a better separation between correct and incorrect poses. This improvement was due to a reduction of both false negatives and false positives. Also, it was illustrated that enrichment was significantly higher when MMPBSA was used to rescore larger databases of docked ligands. Treatment of a substantial number of compounds was computationally feasible, as the compute-intensive part of the MD simulations including explicit water had been replaced by pose generation with a fast docking program.

6. Inclusion of Water in Docking Studies

As previously mentioned, a current weakness in docking is the treatment of water-mediated protein-ligand interactions. If one ignores water-mediated interactions during docking then the calculated interaction energy of a given ligand conformation may be too low. If, on the other hand, one retains crystallographically observed water molecules then the binding pose and affinity of a ligand that in reality replaces that water molecule will not be correct. It is notoriously difficult to treat water adequately, as first at all one need to identify possible positions for water molecules where they could interact with the protein and ligand, and subsequently one must be able to predict whether a water molecule is indeed present at that position. Researchers at Astex and the Cambridge Crystallographic Data Centre recently implemented an elegant procedure in the latest version of GOLD to address both these issues [73].

To predict whether a water molecule is present or absent in a protein-ligand complex, it is needed to estimate the free energy change, ΔGb, associated with transferring a water molecule from the bulk solvent to its binding site in a protein ligand complex. ΔGb for a given water molecule w is defined as follows:

$$\Delta G\mathrm{b}(w) = \Delta G\mathrm{p}(w) + \Delta G\mathrm{i}(w) \quad (1)$$

where ΔGp(w) is the free energy associated with the loss of rigid-body entropy on binding to the target. ΔGi(w) contains contributions resulting from the interactions that the water molecule forms with the protein and the ligand (relative to those it forms with bulk solvent) and also reflects any changes in the interactions between the protein and the ligand caused by the introduction of the water molecule. These interactions may be favorable (e.g., hydrogen bonds) or unfavorable (e.g., steric clashes). They refered to ΔGi(w) as the *intrinsic* binding affinity of a water molecule. For a water molecule to bind to a protein-ligand complex, its intrinsic binding affinity needs to outweigh the loss of rigid body entropy on binding. Therefore, to predict water mediation and water displacement, the balance of the two terms in Eq (1) is critical.

The water positions they consider for a given target are taken from a set of complex structures of that target, but one could also use programs to predict potential water-binding sites [74,75]. Each water molecule can then be present ('on') or absent ('off'). If a water molecule is on, it can make favorable interactions with the ligand and protein, but it pays an entropic penalty for loss of translational and rotational degrees of freedom [76]. The value of this penalty was optimized using a training set of 58 protein-ligand complexes. Considering both the training and test sets, on and off status are correctly predicted for 93% of the water molecules. This increases correct pose prediction rates of water-mediated complexes by 10-12 percentage points, but it decreases correct pose prediction rates for nonwater-mediated complexes by 6-7 percentage points. This latter decrease is readily explained when one assumes that prediction of a water molecule where there should not be one leads to an incorrect binding mode. The expectation is that the correlation of calculated and measured affinities will improve with the inclusion of water molecules in the docking runs, which in turn should improve the enrichments obtained in VS experiments, but this remains to be investigated.

Another approach to dealing with water molecules involved in protein-ligand interactions has been incorporated in the FlexX docking program. This method, referred to as the particle concept, includes the calculation of favorable positions of water molecules inside the active site prior to docking. During the incremental construction phase these water molecules are allowed to occupy the precomputed positions if they can form additional hydrogen bonds with the ligand. The method was tested using a data set of 200 protein-ligand complexes and with pose prediction quality as an evaluation criterion. Similar to the observations made for GOLD, it was found that on average the improvement was minor. Nevertheless, in a number of cases the predicted waters corresponded to the crystallographically observed ones, which led to an improvement in the predictions [77].

Another program that needs to be mentioned in this context is the program SLIDE [78]. Prior to docking, a knowledge based approach (CONSOLV [49]) is applied in order to select those water molecules that are likely to remain in their positions upon ligand binding and to determine an energy penalty for their displacement. During docking, overlap between the docked ligand and these water molecules is resolved by iterative translations or annihilation of the water molecules, applying appropriate penalties in due course.

While current sampling and scoring algorithms are often able to predict the correct binding pose [79], satisfactory prediction of binding affinity has yet to be achieved [80]. Other particular challenge in ligand docking studies is the positioning of interface water molecules [79]. That interface water molecules play an important role in ligand binding is evidenced by the fact that many protein/ligand complexes contain structured waters that bridge protein and ligand. For instance in the CSAR dataset 299 out of 341 complexes include waters within hydrogen bonding distance of both protein and

ligand atoms. These water molecules are often absent in experimental structures of the apo protein [81]. Waters stabilize protein/ligand interfaces by providing indirect interactions between protein and ligand through formation of hydrogen bonds with both partners [82]. In empirically derived scoring functions optimized to predict binding affinities [83,84], components such as hydrogen bond energy have been weighted to account for the change in energy compared to hydrogen bonds formed with water [77]. Similarly the ''hydrophobic'' score terms are used to represent desolvation of the protein receptor. Nevertheless, great improvements have been seen in molecular dynamics based binding affinity prediction when water is considered [85,86].

In both self-docking [87] and cross docking published studies [88], correct ligand binding pose prediction can be improved by the presence of conserved crystallographic waters, but although it has been demonstrated that water docking can improve inhibitor placement, it has been not seen a significant improvement in binding affinity prediction. Although water molecules are routinely included in molecular docking methods and protocols because of their important role in mediating ligand protein interactions, it is still unclear that the inclusion of explicit water molecules improves docking accuracy.

The study of Lie et al. [89] showed that water mapping calculations can be used to select key water molecules from experimentally identified water positions for molecular dockings. They have observed that inclusion of all binding site water molecules led to reduced performance and erroneous results. Moreover, an overall improvement in binding pose prediction was achieved when computationally selected water molecules are included during docking simulations. The improvement in the docking performance by including water molecules also depends on protein system, chemical class of ligand, docking method, and scoring function. They have implemented a method for incorporating key water molecules in protein-ligand docking. First, the method fully solvates the ligand with attached water (AW) molecules, and these are then included during the docking calculation, if the interaction energy between the AW molecule and the surroundings is favorable (negative). The loss of rigid-body entropy when a water molecule binds to a protein is taken into account by adding a constant (positive) entropy penalty value per included AW molecule. From the training set consisting of 12 diverse complexes, an optimum is found for the entropy penalty value of Sp = 3. The entropy penalty value of 3 is energetically equivalent to a penalty of 5.5 in a setup where the interaction between the AW molecule and the ligand atom it is attached to contributes to the score. Thus, the entropy penalty is approximately twice the strength of an ideal hydrogen bond with opposite sign in MVD, which is in agreement with the size of entropy penalty for including water molecules in GOLD's displaceable water model [73].

As water is a key structural feature of protein-ligand complexes and can form a complex hydrogen-bonding network between ligand and protein [90,91], water-mediated binding is so common that a study of 392 protein-ligand complexes found that 85% had at least one or more water molecules that bridge the interaction between the ligand and the protein [47]. Furthermore, the displacement of an ordered water molecule can drastically affect a ligand's binding affinity [76,92]. As a result, it is common nowadays to include explicit water molecules in computational drug design [93-95]. The careful consideration of hydration sites has been shown to aid the predictability of 3D QSAR models, [96-98] ensure stable simulations with molecular dynamics [99], and improve the accuracy of rigorous free energy calculations [100]. Continuum solvent models have also been reported to improve with the addition of explicit water molecules [101]. Traditionally, ordered water molecules were ignored in ligand docking studies and ligands were docked into desolvated binding sites. There are now a number of docking protocols that include explicit water molecules and claim to improve accuracy in many cases [77,87,88,102,103,104]. However, it has also been reported that including such water molecules may hamper efforts to predict a ligand's correct binding mode [105]. A popular strategy in rational drug design is to modify a ligand so that it displaces an ordered water molecule into the bulk solvent [92,96,106,107]. This is due to the favorable entropic gain that can result by increasing the water molecule's translational and orientational degrees of freedom. However, the targeted displacement of an ordered water molecule may be unsuccessful [57,76], can also lead to a decrease in affinity if the ligand is unable to replace the water molecule's hydrogen bonds correctly and fulfill its stabilizing role [76,108]. This has important implications for lead optimization and rigorous theoretical studies have investigated how changing a water displacing functional group affects a ligand's affinity [109,110]. In addition, water molecules are important pharmacophoric features of a binding site [111], and the chemical diversity of potential inhibitors generated in silico has been reported to be greatly affected by the targeted displacement of ordered water molecules [112,113,114]. Water molecule locations are typically taken from X-ray crystal structures and may be validated by observing the same position in other crystal structures of the same protein. Nevertheless, there are inherent problems with identifying hydration sites with crystallography. Water molecules can be artifactual, may be too mobile to identify or not observed because of low resolution [115,116,117]. In cases such as homology modeling, there will be no structural knowledge of water molecules. Hence, it is necessary to be able to accurately predict water locations within binding sites.

As the affinities, binding modes and chemical diversity of a series of ligands can be greatly affected by the water molecules in a protein binding site, it is important to predict which water molecules are displaced or conserved during the binding process. Some docking procedures, although different in implementation, involve switching explicit water molecules ''on'' and ''off'' [73,77,118]. Other approaches have used the structural features of a water molecule's environment to predict whether it will be displaced or not without any prior knowledge of the ligand. Using a K-nearest neighbors genetic algorithm, Consolv reported 75% accuracy in predicting whether a binding site water molecule would be displaced or not [49]. However, as Consolv used crystallographic temperature factors as structural descriptors, it cannot be applied to predicted water sites. Amadasi and coworkers have combined the HINT force field [119] with the Rank score [120] to classify water molecules into 2 broad categories;

conserved/functionally displaced and sterically displaced/missing [121,122]. Their first study correctly classified 76% of the water molecules tested while their second study reported a classification accuracy of 87%. Their analysis included weakly bound water molecules, which were a maximum of 4 A° away from the protein. On the other hand, WaterScore used water molecules within 7 A° of the bound ligand in protein-ligand binding sites [123]. Using multivariate logistic statistical regression, WaterScore reported 67% accuracy in classifying displaced and conserved waters, although water molecules that were displaced because of steric clashes with the ligand were not included in their analysis. Barillari et al. used the computationally expensive double-decoupling method to calculate the binding energies of 54 water molecules in protein-ligand complexes [46]. They found that water molecules that could be displaced by a ligand were on average less strongly bound than conserved water molecules by 2.5 kcal/ mol.

Despite the positive strides that have been made in understanding the role of ordered waters, no single method is able to answer how displaceable a water molecule is, and what is it likely to be displaced by. When there is limited experimental knowledge of a binding site's solvation structure, addressing these questions becomes even less clear. Trott et al. [124] developed a pipeline that can accurately predict the location of water molecules and predict whether they are likely to be conserved or displaced after ligand binding. They also predict the probability that predicted water molecules will be displaced by polar or non-polar groups.

It is important to remember that when a protein molecule is in solution, its entire surface is covered by water molecules with properties different to those in the bulk [125]. Cryogenic X-ray determinations and molecular dynamics simulations reveal the existence of large networks of water molecules around the surface of proteins [126,127,128]. Most of the water molecules in the vicinity of a protein are loosely bound to it, remain mobile, have short interaction times, and are not readily observed via conventional X-ray crystallography. Although some water molecules that are observed in crystal structures are artifacts of the determination [116], others are clearly tightly bound to the protein surface [47,129,130] particularly in clefts on the protein surface, such as ligand binding sites [131].

Real water molecules observed in the crystal structures of proteins have a tendency to occupy conserved positions in structurally related proteins [132-144], as well as in structures obtained under different conditions [126,145] and/or different bound ligands [146-151]. The most frequent structural change among structurally related ligands bound to the same protein seems to involve different arrangements of water molecules [152]. The importance of water molecules found in the binding site of a protein lies in their ability to mediate the interactions between the ligand and the protein and form hydrogen-bonded networks that can stabilize a protein-ligand complex in solution [81,130,153-160]. Such a hydrogen-bonded network of water molecules may stabilize the complex formed with one ligand but not another, thus contributing to the specificity of ligand recognition [155,161]. Water molecules may also help to stabilize the conformation of the active sites of enzymes [162]. Water molecules have also been used to improve the predictive ability and rationalization of three-dimensional quantitative structure-activity relationship (QSAR) models [97,163] provide a structural rationale for ligand-derived pharmacophore models of binding sites [111,164,165,166] and improve the performance of virtual screening [78,167].

The binding of a ligand to a protein receptor often involves the thermodynamically favorable release of water molecules from the protein surface to the bulk solvent. However, the retention of water molecules that are bound tightly to the protein surface upon ligand binding may be associated with an entropic penalty that is outweighed enthalpically, through favorable hydrogen-bonding interactions to both the protein and the ligand [168-175]. Computer simulations have been used successfully to compute the free-energy changes associated with hydrating binding sites and displacing water molecules upon ligand binding [64,65,171,172,173,175,176,177,178179], as well as the binding of tightly bound water molecules [46]. In addition, several approaches have been taken to predict hydration sites [180,181,182,183], conserved water mediated and polar ligand interactions [49,115,184] water occupancy [185], and the displacement of tightly bound water molecules [66,74,121].

The importance of water molecules is now recognized in structure-based drug design, where the displacing, mimicking, and/or targeting of bound water molecules is performed to improve the binding affinity of ligand molecules [186]. The displacement and mimicking of tightly bound water molecules may result in increased binding affinity through the entropy gain of releasing such ordered water molecules [92,106]. However, this does not always seem to be true [108], and, in some cases, the recruitment of an additional tightly bound water molecule that can bridge the interactions between the ligand and the protein has been determined to decrease the binding affinity of ligands [187]. It has also been observed that natural substrates [188] and designed inhibitors [157] may not necessarily displace tightly bound water molecules but rather preserve water-mediated interactions. In some cases, water molecules have been observed not to mediate any ligand-protein interactions but simply to better define the steric shape of the binding site [189]. The consideration of tightly bound water molecules in de novo drug design methods has also demonstrated their role in modulating the binding modes and chemical diversity of designed ligands by imposing steric and hydrogen-bonding constraints [112,114,180]. A few methods have been developed to place water molecules during ligand-protein docking simulations, with some success [77,191]. However, various studies have provided no conclusive evidence that the inclusion of tightly bound water molecules improves docking accuracy [78,105,165,167,192,193,194,195], mostly because not enough protein targets were investigated to achieve statistical significance. Nonetheless, the first comprehensive docking studies with large datasets of ligand-protein complexes were contradictory, either revealing that including water molecules did not increase accuracy [196] or showing that it did [197]. Recent studies attempted to assess the accuracy of ligand-protein docking in the presence of tightly bound water molecules [103]. The binding modes of various ligands of cytochrome P450 and thymidine kinase were predicted using three different docking programs, with and without water molecules. The

positions of water molecules were obtained directly from the known crystal structures and from predictions using a novel GRID-based method [103]. Docking accuracy improved in the presence of crystallographic water molecules, with a larger improvement measured when predicted water molecules were included. Improvements were also detected in the accuracy of virtual screening [198]. Water molecules were observed not only to mediate the interactions between ligands and proteins, but also to help in the placement of ligands close to the center of the active sites [198]. A simple entropy penalty term has been introduced to account for the unfavorable loss of rotational and translational entropy that accompanies the tight binding of a water molecule to a protein surface [73]. This penalty term was used to predict the displacement of tightly bound water molecules upon ligand binding. An improvement in the accuracy of ligand-protein docking was observed for a large dataset of ligand-protein complexes [73]. This effect was most noticeable in those cases where water molecules were determined to mediate the ligand-protein interaction.

The modelling of protein-protein complexes by means of docking (a computational approach which models the unknown structure of a complex from its constituents) has become increasingly popular, as witnessed by the CAPRI (Critical Assessment of PRedicted Interactions) experiment [199]. Docking approaches have benefited from knowledge obtained by detailed analyses of binding interfaces [200,201]. As discussed in a recent review, water molecules are expected to influence the assembly of biomolecular complexes [202], and, as such, to be important for protein-protein docking. An analysis based on Voronoi volume showed that only upon inclusion of interfacial solvent molecules are protein-protein interfaces as densely packed as protein interiors [203]. So far, however, water has been neglected generally in biomolecular docking. Its role and importance in single proteins have been discussed [203,204] and several case studies have analyzed its conservation in 3D structures of homologues [114,134,139,145,149,151,184,203,205]. There has also been quite some interest in identifying and predicting the positions of water molecules in known structures: this can be quite successfully performed, for example, by GRID [203] or Fold-X [183].

These kind of approaches, however, are not very well suited for docking purposes, since the structure of the complex is not known a priori. Ideally, water should be accounted for directly during the docking process since its presence might affect the resulting models. So far this has only been done for protein-ligand [73,155,194,203,206] and nucleic acid-ligand docking [207]. Only very recently has the role of water molecules at protein-protein interfaces been investigated. A hydrogen bonding potential for water-mediated contacts, in combination with a solvated rotamer library for describing side chain conformations, has been shown to predict rather successfully the positions of water molecules in complexes with known structures [208]. In another study [209], various properties of interfacial water molecules such as residue preference and their number per unit of interface area were investigated.

Nowadays, molecular docking plays an important role in drug design and discovery with the universal application of docking programs, such as Glide [191], Autodock [210], FlexX [211], and GOLD [212]. When these programs are utilized prior to experimental screening, they are usually considered as powerful computational filters to reduce labor and cost. All of these docking programs explore various docked conformations and determine the tightness of interactions between the protein and the ligand, but the performance on predicting the experimentally observed binding poses is not always satisfying. As is widely accepted, the real bottleneck on obtaining the reliable docking result lies in the scoring functions [213-218]. As a matter of fact, considerable efforts have been devoted to the development of approximate computational methods for describing protein−ligand interactions more accurately, but it still lacks a universal scoring function which works reliably for all or most of protein−ligand systems [219,220]. For some particular protein−ligand systems, most of the widely used docking programs are incapable of predicting the correct binding modes, imposing great challenge on the effectiveness of computer-aided drug design. Therefore, improved methods for predicting protein−ligand binding affinities are desperately needed.

It is generally known that the effect of bridging water molecules between the protein and ligand attracts more and more attention recently, as these water molecules are considered to play an important role in mediating the interaction between protein and ligand [81,86,129,131,148,213,221,222,223]. While only a few scoring functions explicitly take the water-mediated protein−ligand interactions into consideration [63,78,194,224] explicitly including the bridging water molecules in molecular docking and scoring function may be crucial for correctly predicting the binding poses as shown in Figure 2 for 1ETR (a) and 1FJS (b) using Autodock [225].

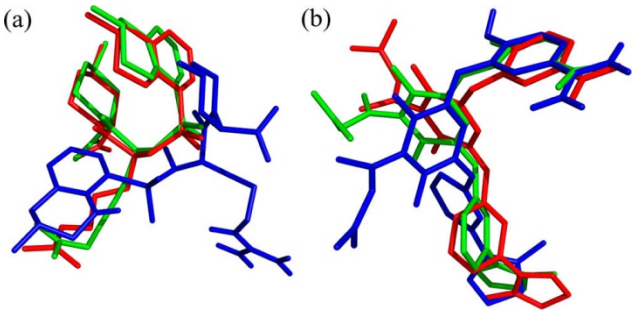

Figure 2. Superposition of best-scored conformations with (in red) and without (in blue) the bridging water molecules compared with the experimentally observed (in green)

It must be pointed out the fact that generating a receptor-ligand structure in silico involves two main components (sometimes inaccurately referred to as ‘‘steps’’): docking and scoring. Docking per se entails conformational and orientational sampling of the ligand within the constraints of the receptor binding site. Scoring function selects the best pose (i.e., ligand conformation, orientation, and translation) for a given molecule and rank orders ligands, if a ligand database is docked/screened. To be successful, docking must accurately predict two things relative to experimentally available information: ligand structure (pose prediction) and its binding propensity (affinity prediction). Aim to address the main challenges of docking: receptor representation (such as structural

waters, side chain protonation, and, most of all, flexibility (from side chain rotation to domain movement)), ligand representation (protonation, tautomerism and stereoisomerism, and the effect of input conformation), as well as accounting for solvation and entropy of binding. These challenges of docking are very well reviewed by Corbeil et al [226].

But in many popular docking methods the ligand is treated as flexible but the protein conformation is kept rigid. This relies on the Lock-and-Key hypothesis for protein ligand binding. However, it is now widely accepted that ligand binding is not a static event but a dynamic process, in which both the ligand and protein may undergo conformational changes. In docking, incorporating protein flexibility exponentially expands the potential search space and quickly becomes impractical. Therefore, properly accounting for receptor flexibility is much more computationally expensive than doing that for ligands [227]. Docking programs have only recently started to offer partial receptor flexibility during docking.

Anderson and co-workers [228] suggest that ligands are more accurately assessed when docked to the minimized ensemble from a single MD simulation, an improvement due to more than just error minimization. They have also developed an efficient method to evaluate and select the most ''contributive'' ensemble members prior to docking for targets with conserved binding site cores [228], and Sanchez-Moreno et al. [229,230] employed a new methodology to account for drug and enzyme flexibility with apparently good results, but not useful for large numbers of drug candidates.

An alternative to MD is to generate receptor ensembles by normal mode analysis. Abagyan and co-workers have demonstrated that the elastic network model (ENM) is a method that may initiate not only local conformational changes, such as those of side chains, but the movement of the protein backbone. They have also demonstrated that the ENM can be significantly more efficient than MD [231]. Gohlke and co-workers [232] have developed an accurate grid-based representation of intermolecular interactions, which evaluates interaction energies via lookup tables even for a moving protein. The efficiency is achieved by adapting a 3D grid with pre-calculated potential field values, derived from the initial receptor conformation, to another conformation by moving the points in space, but keeping the values. Baker and co-workers [233] have commented on the problem of sampling tautomeric and protonation states, given the possible difference of free and bound ligand states in these respects. They suggested enumeration of tautomeric and protonation states as a possible solution but have warned about the potentially prohibitive computational cost. Another suggested alternative included segmentation and incremental construction of the docked ligand, whereby the protonation and tautomerism ''decisions'' are independent and hence decrease the problem size.

As receptors bind to their ligands in solution, the solvation aspects used to be commonly treated implicitly, that is, by the use of implicit solvents, knowledge-based scoring functions or by modification or calibration of other scoring functions. Fong et al. [234] have investigated the inclusion of a desolvation penalty into their QM/MM scoring, using a Generalized Born solvent model, and found that it resulted in improved pose prediction. Cincilla et al. [235] have modified the solvation treatment in the scoring function of AutoDock 3 [210] to improve the predictions of weak complexes containing ligands with polar atoms lacking a matching partner in the binding site. Specifically, they have removed the constant hydrogen bonding energy term for the polar ligand atoms and introduced the Stouten free energy desolvation term. The modified function has also differentiated between the ''polar'' and ''non-polar'' heteroatoms on the basis of hybridization and connectivity. Kuntz and co-workers [236] have used two implicit solvent scoring functions AMBER/GBSA and AMBER/PBSA, implemented in DOCK 6, for docking small molecules to RNA. Sodium ions were used to neutralize the backbone charge and a double shell of explicit water was used to shield the charges. They have found that the quality of pose prediction increased from 70% to 80% for moderately flexible ligands (<7 rotatable bonds) and from 26% to 50-60% for highly flexible ligands (7-13 rotatable bonds). Huang and Wong have tested the performance of a simple implicit solvent method (a distance dependent dielectric model) in comparison to a version of the Generalized Born method (GBMV) and found it to produce better pose prediction results for a fraction of the computational cost [237].

The problem is that most methods of treating solvation do not take into account the effects that could be exercised by ''structural'' water molecules, that is, those that provide stabilization and/or recognition through specific hydrogen bonding and even van der Waals interactions. Villacanas et al. [238] have reviewed the effect of structural water molecules on docking and concluded that the general feeling in the literature is that explicit water molecules improve docking outcomes, both in pose prediction and virtual screening. Englebienne and Moitessier [239] have shown that the consideration of displaceable water molecules, implemented in FITTED, improves pose prediction, but does not significantly affect scoring accuracy. They have suggested that the latter is most likely the outcome of most scoring functions having been developed for ''dry'' proteins. Recently, Horbert et al. [240] have successfully employed the water map technology [191,241,242] to calculate hydration sites.

7. Identification of Key Water Molecules

Kumar et al. [243] approach is based on previous observations that inclusion of water molecules in docking improves the docking accuracy. It has been observed that inclusion of all active site water molecules creates bias for ligands of a particular chemical class and may not work for others. Also, sometimes water positions are not reliable in crystal structures, especially in structures of low resolution [115,116,244]. Therefore, the appropriate selection of water molecules to be included in docking studies is crucial, existing several computational methods that evaluate solvent energetics for the identification of key water molecules. In several studies, key water molecules for each protein target were identified by running water mapping calculation with SZMAP [see for example [245,246].

They implemented a novel approach to score water mediation and displacement in the protein-ligand docking program GOLD. The method allows water molecules to

switch on and off and to rotate around their three principal axes. A constant penalty, σ_p, representing the loss of rigid-body entropy, is added for water molecules that are switched on, hence rewarding water displacement.

As previously mentioned, water molecules can be involved in protein-ligand recognition either by forming mediating hydrogen bonds between the protein and the ligand or by being displaced by the ligand; both of these mechanisms have been shown to be of importance to drug discovery [170]. For example, the first-generation HIV-1 protease inhibitors were peptidic in nature and all formed hydrogen bonds to a conserved water molecule between the two central "flaps" [247]. Subsequently, it was discovered that it is possible to displace this structural water molecule, which led to new inhibitor series [92]. Similarly, the benzamidine moieties in early factor Xa inhibitors interacted with a conserved water molecule situated above a tyrosine ring in the S1 pocket [248]. More recently, inhibitors binding with neutral moieties in the S1 pocket were shown to displace this water molecule [249,250]. There could be several potential advantages to including water molecules in a protein-ligand docking program. First, if the compound interacts with the water molecule, including it could improve the predicted binding mode. Several studies have been reported in the literature where parallel dockings were done in the absence of water molecules and in the presence of some key water molecules. Some authors have reported significant improvements in docking performance when water molecules were included [167,197], whereas others found that including water molecules had little effect on the quality of the dockings [103,251]. A second potential advantage of addressing water binding in a docking application is that it could distinguish between compounds that can displace a water molecule and compounds that cannot. Finally, correctly scoring water mediation and water displacement in scoring/energy functions could help in ranking compounds and, therefore, increase hit rates obtained from virtual screening.

Various applications have been reported in the literature for predicting potential water binding sites on proteins. For example, AQUARIUS [252] is a knowledge-based approach specifically aimed at identifying water sites in proteins; other applications including GRID [253], MCSS [254], SuperStar [255], and CS-Map [182] can also be used for this purpose. However, such applications do not directly indicate which predicted water molecules are likely to be displaced by a ligand and which are likely to remain bound to the protein. Solving this issue is clearly of importance to structure-based design, as it would indicate whether compounds could be designed to displace the water or to interact favorably with a water molecule.

If a sufficient number of X-ray structures of protein-ligand complexes are available, displaceable and conserved waters can often be identified and a suitable design strategy can be adopted [256]. Consolv was developed by Raymer et al. [49] to automate the process of assigning conserved waters using the distribution of a number of structural parameters describing the water molecules in a training set of 13 diverse proteins. More recently, García-Sosa et al. [74] used a similar set of parameters in WaterScore to distinguish between conserved and displaceable water molecules. When water molecules are known or assumed to play a role in protein-ligand recognition, the most common strategy is to perform separate docking runs in parallel, i.e., one in the absence of water molecules and a second in the presence of one or more water molecules. However, these parallel runs need to be analyzed and some assessment of the cost of displacing a water molecule is required. Hence, it would be preferable if the docking program could assess both the bound and unbound states of water molecules. To address this, FlexX [211] can precalculate energetically favorable water sites [80]; "spherical" water molecules ("particles") can then be switched on at each of these positions during the docking protocol.

In SLIDE [78], Consolv is used to predict water molecules that are likely to be displaced, and these water molecules are removed from the binding site. The remaining water molecules can then be displaced during the docking at the cost of a penalty. AutoDock [257] can use multiple energy grids representing different states of the protein. Österberg et al. [194] created energy maps for different structures of HIV-1 protease, including one structure that contained the key water molecule interacting between the flaps, hence implicitly giving Auto-Dock the option to "choose" between the water-bound and the water-unbound state. What we believe is missing in the above approaches is the concept that a water molecule that is displaced by a ligand gains rigid-body translational and rotational entropy and that this should therefore be *rewarded* in the scoring function used by the docking program. We also feel that predicting the positions of water molecules as well as their occupancies (i.e., whether they are bound or displaced) makes the problem unnecessarily challenging. In most structure-based drug discovery applications, the modeler will have access to knowledge about potential water sites and will be able to make an informed judgment on which water molecules to consider.

In the study by Murray and co-workers [87], the presence of all crystallographic water molecules in the binding site biased the search space by physically restricting the number of possible binding modes. In two studies by Mancera and co-workers, a redocking study [87] and a cross-docking study [257], all water molecules that were capable of interacting with both the ligand and the receptor protein were included by indiscriminately including all water molecules that fulfilled the distance criteria to both ligand and protein, and on that basis, it was concluded that water molecules play a key role in protein-ligand recognition because the efficacy and accuracy of the docking simulations were improved. Thus, the results obtained from all three of the mentioned studies are artificially constrained by including all nearby water molecules [258], all possibly mediating water molecules [87], or all conserved possibly mediating water molecules [258]. However, despite these limitations, it is obvious that water molecules in some cases are important for the binding of a ligand in a protein receptor. Water molecules are thus important in structure-based drug design, where the binding affinity of ligand molecules can be improved by mimicking, displacing, and targeting bound water molecules [186]. When a water molecule is known or assumed to play a role in protein-ligand recognition, a simple way to incorporate it in a docking problem is to perform one docking simulation, where the water molecule is included as a static part of the receptor

structure and another where it is absent. This strategy is feasible if only a few water molecules are potentially important, but when n water molecules are assumed to play a role in protein-ligand recognition, this approach will ultimately sum up to 2n separate docking simulations in parallel. An alternative is a displaceable water model where the individual crystal water molecules from the PDB structure can be toggled on/off automatically during the docking simulation, so that a ligand can keep favorable water molecules and displace nonfavorable water molecules.

In another study with a large set of protein-ligand complexes, the inclusion of all crystallographic water molecules within 6.0 Å of any ligand atom resulted in a large increase in the docking accuracy [259]. However, as the authors acknowledged themselves, by indiscriminately including all water molecules in the binding site, the search space may have been drastically biased, virtually leaving the correct binding mode as the only possible low-energy solution [259]. They have performed a comprehensive survey of the role of water molecules on the accuracy of ligand-protein docking simulations by expanding the number of ligand-protein complexes considered to include all those in the original CCDC/Astex test set [196] that contain water molecules in their crystal structures. This ensures a more-thorough examination of the effect of including water molecules on the accuracy of predictions of binding modes by a standard docking/scoring strategy. We also focus on two particular issues: (1) the influence of the method chosen to optimize the orientation of water molecules prior to docking, and (2) the inclusion of only those water molecules that are in the immediate vicinity of both the ligand and the protein has been determined to decrease the binding affinity of ligands [187]. It has also been observed that natural substrates [188] and designed inhibitors [157] may not necessarily displace tightly bound water molecules but rather preserve water-mediated interactions. In some cases, water molecules have been observed not to mediate any ligand-protein interactions but simply to better define the steric shape of the binding site [189]. The consideration of tightly bound water molecules in de novo drug design methods has also demonstrated their role in modulating the binding modes and chemical diversity of designed ligands by imposing steric and hydrogen-bonding constraints [112,114,190].

8. Conclusions

The principles and methods briefly discussed in this review highlight the role of interactions between a ligand and water molecules which generally represent only a small fraction of the number of interactions the ligand forms in its bound state. Hence, it is not uncommon that a reasonable binding mode can be produced without including even key structural water molecules. Also, the fact that a small drop-off in success rates is usually observed for the displaced and decoy waters test sets is almost inevitable. All water molecules included in these sets are in good positions for a water molecule to interact with the protein and are, therefore, not easily displaced. Additionally, because the water molecules in these sets should be displaced by the ligand, including them can only distract the docking algorithm from identifying the correct binding mode. Each water molecule that is included increases both the search space and the likelihood of obtaining false positives. Therefore, it is important to limit the number of water molecules included in a docking run and only include water molecules that are known to be crucial for ligand binding.

Including water molecules in the docking runs could improve the correlation of the scores with the affinities of the compounds and may improve the enrichments obtained in virtual screening experiments. A protein molecule and a ligand in solution are covered by water molecules. When a ligand and a protein receptor form a complex, desolvation must take place, and the ligand and the protein interact through direct interactions. Also, in some cases, contacts are mediated through discrete water molecules. These water molecules stabilize the protein-ligand complex by forming a hydrogen-bonded network, taking part in the interactions between the ligand and the receptor [46,80,129,132,158,159,260,261,262]. Furthermore, a hydrogen-bonded network of water molecules may stabilize the complex formed with one ligand but not another, thereby contributing to the specificity of ligand recognition. Several protein-ligand docking studies have been performed to elucidate that the presence of water molecules in a ligand binding site plays a key role in protein-ligand recognition. In 2008, Murray and co-workers [254] published a docking study where the inclusion of all crystallographic water molecules within 6.0 Å of any ligand atom resulted in a large increase in docking accuracy. By including all water molecules in the binding site, the search space is however biased toward the correct binding mode, which is also discussed in the paper [259]. In 2008, Mancera and co-workers [87] published a comprehensive redocking study to investigate the importance of water molecules for the accuracy of protein-ligand docking predictions. In that study, any crystal water molecule that is capable of forming mediating hydrogen bonds between the ligand and the receptor, i.e., any crystallographic water within 2.5 Å from any ligand atom and 3.0 Å from any protein atom, was included as a static part of the receptor structure. The study found that the efficacy of the docking simulations and the accuracy of the docking predictions were significantly improved with the inclusion of the crystallographic water molecules in the binding site. The redocking study [87] was in 2010 followed by a cross-docking study [258]. Six different protein targets with between three and 13 available protein-ligand PDB structures were considered. For each of the six targets, a common set of the crystallographic water molecules was found. Also in this study, a significant improvement in the accuracy of the predicted binding modes was observed with the inclusion of the conserved water molecules. Lemmon et al. [263] having find that docking algorithms fail in some cases to predict the correct protein/ligand complex structure, showed that simultaneous docking of explicit interface water molecules greatly improves Rosetta's program ability to distinguish correct from incorrect ligand poses. This result holds true for both protein-centric water docking wherein waters are located relative to the protein binding site and ligand-centric water docking wherein waters move with the ligand during docking.

So, the influence of water in biomolecular complexes formation processes is a crucial point to consider in drug design. Although the improvement achieved to date in the studies related to water influence in drug-receptor interactions is noteworthy, there is still an urgent necessity of further research in order to increase the predictability of the proposed models and to gain a deeper knowledge of all the involved factors.

Most molecular docking programs successfully predict the binding modes of small-molecule ligands within receptor binding sites. However, the current algorithms do not estimate the absolute energy associated with the intermolecular interaction with satisfactory accuracy. The appropriate handling of issues such as solvent effects, entropic effects, and receptor flexibility are major challenges that require attention. Successful molecular docking protocols require a solid knowledge of the fundamentals of the applied methods. Understanding these principles is essential in the production of meaningful results.

Over the past decades, protein-ligand docking has emerged as a particular important tool in drug design and development programs, as it has several strengths, among which the method's ability to screen large compound databases at low cost compared to experimental techniques is particularly notable.

This gain in standing is well portrayed in the rising number of available protein-ligand docking software programs, increasing level of sophistication of its most recent applications, and growing number of users. In spite of the large number of alternatives, we are still far from a perfect docking program. In terms of the searching algorithms, efficiently accounting for protein flexibility remains a challenging task. In terms of the scoring functions features like the presence of structural water molecules and the treatment of entropy, among others, still pose considerable problems for protein-ligand docking. Many protein-ligand docking programs are currently available and new alternatives are continuing to appear every year.

9. Future Perspectives

Biomolecular simulation is extremely computationally demanding, especially when it comes to processing large complex systems. Nonetheless, quantum computers have special capabilities that could make a real difference to our ability to compute and predict the properties of biological systems at the cellular level. Most importantly, it is the ability of a quantum computer to explore many classical paths simultaneously that offers a potential method to overcome the problem of finding the minimum free energy, as opposed to just the minimum energy, in an optimization problem such as protein folding or molecular docking. State-of-the-art MD simulations now routinely include 'repeat' simulations, in which a number of initial conformations are investigated in parallel to check the robustness of any conclusions against thermal noise.

Simulations have comparable advantages and caveats to the other existing experimental techniques, and should not be regarded as any less valid so long as they are used appropriately and the corresponding limitations are clearly stated. To conclude, we can say that the biomolecular sciences need to embrace computer simulation as a useful technique for model building and hypothesis testing, especially given the vast quantities of biomolecular data that are being generated. Most insight will be obtained by combining all available biophysical methods to address a single biological problem, and computer simulation can make a valid and valuable contribution.

Despite the underlying assumptions and inconsistencies, researchers are dedicated to develop and refine implicit models to be more accurate. Therefore, further research efforts based on implicit solvent models should continue to focus on modifications to overcome such limitations without significantly reducing the computational efficiencies of these models. And the number and types of this improvements is growing everyday. For example, WaterFlap, included in FLAP last release presents an enhanced approach to docking with optional water molecules [264] and WaterMap [265], nowadays included in Schrodinger suite for molecular docking studies, offers a new paradigm for designing optimal ligands based upon their ability to differentially displace and retain specific water molecules in protein binding sites. Mapping the locations and thermodynamic properties of water molecules that solvate protein binding sites offers rich physical insights into the properties of the pocket and quantitatively describes the hydrophobic forces driving the binding of small molecules. WaterMap has been applied with great success to a wide variety of pharmaceutically relevant targets [see for example 266-269].

Additionally, there is increasing effort to incorporate explicit polarization into the general classical mechanics in different forms such as point dipole induction and Drude oscillators to improve the electrostatic representation of biomolecules. Adoption of such polarizable potentials in routine studies remains limited, mostly because of concerns about the computational expense. Advances in computing power and efficient simulation algorithms; however, will continue to reveal shortcomings of oversimplified fixed-charge potentials and remind us of the missing physics. Additionally, development of advanced classical electrostatic model beyond simple polarization is ongoing. In addition to polarization effect, the local charge-transfer (CT) and penetration effects are demonstrated to play important role for short-range molecular interactions in water, aromatics and high-valence ions.

Incorporation of such effects will significantly improve the accuracy in modeling the structural and energetic details of these molecular clusters. Advancements in the electrostatic representation of biomolecules and their solvent environment have already led to successful applications including small molecule solvation, pKas and protein-ligand binding affinity prediction, but computational sampling can however be the next bottleneck in achieving more accurate thermodynamic quantities in complex molecular systems. Advancements in statistical mechanics theories are as important [270]. Although approaches such as free energy perturbation (FEP) and application of Bennett's acceptance ratio (BAR) may require little additional work beyond what is required for molecular dynamics, methods such as thermodynamic integration, lambda dynamics, meta dynamics and the orthogonal space random walk (OSRW) strategy are more time-consuming to implement for polarizable atomic

multipole descriptions of electrostatics. Tenable, but nontrivial, complications arise with the latter methods because of their dependency on the derivative of the potential energy with respect to the state variable λ. For example, to the best of our knowledge, a soft-core method to smoothly decouple atomic multipolar interactions with respect to λ has yet to be described. Given the power of meta dynamics based methods to enhance molecular dynamics sampling and reconstruct the free energy surface along a few collective variables, there is great motivation for force field experts to work closely with developers of the leading statistical mechanics algorithms in the future. For example, the scope of the freely available complex analyzer GIANT [271] is limited to the direct contacts between proteins and small molecules in the current version, but its basic concept is applicable to other various kinds of molecular interactions such as water-mediated interactions. Their developers claim to be planning to take statistics of interactions with metal and water molecules that play important roles for molecular recognitions. In addition, while the interaction patterns defined in GIANT focuses on the relative positions between a protein fragment and a ligand atom, and does not consider the combination of the elements interactions (or the "environment" around the contacting pair), their intentions are to take some statistics of cooccurrences of the interaction patterns in a future work, as the information about environment should be an important factor in the ligand recognition.

Competing Interests

The author declare no competing financial interest.

Acknowledgement

The autor thanks the MCINN Project Consolider Ingenio CSD2010-00065 for financial support.

References

[1] Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed.,* 2002, *41*, 2644-2676.

[2] Bühm, H. J.; Stahl, M. The use of scoring functions in drug discovery applications. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2002; Vol. 18, pp 41-87.

[3] Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discovery*, 2003, *2*, 527-541.

[4] Perola, E.; Charifson, P. S. Conformational Analysis of Drug-Like Molecules Bound to Proteins: An Extensive Study of Ligand Reorganization upon Binding. *J. Med. Chem.*, 2004, *47*, 2499-2510.

[5] Butler, K. T.; Luque, F. J.; Barril, X. Toward accurate relative energy predictions of the bioactive conformation of drugs. *J. Comput. Chem*., 2009, *30*, 601-610.

[6] Carlson, H. A.; McCammon, J. A. Accommodating Protein Flexibility in Computational Drug Design. *Mol. Pharmacol.*, 2000, *57*, 213-218.

[7] Jiang, F.; Kim, S. H. Soft Docking - Matching Of Molecular-Surface Cubes. *J. Mol. Biol*., 1991, *219*, 79-102.

[8] Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation. *Chem. Phys. Lett.*, 1997, *278*, 297.

[9] Wasserman, Z. R.; Hodge, C. N. Fitting an inhibitor into the active site of thermolysin: A molecular dynamics case study. *Proteins-Struct. Funct. Genet.*, 1996, *24*, 227

[10] Dinola, A.; Roccatano, D.; Berendsen, H. J. C. Molecular-Dynamics Simulation Of The Docking Of Substrates To Proteins. *Proteins-Struct. Funct. Genet.*, 1994, *19*, 174.

[11] Sippl, M. J. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. *J. Mol. Biol.*, 1990, *213*, 859-883.

[12] Nangia, A.; Desiraju, G. R. Supramolecular structures - Reason and imagination. *Acta Cryst. A*, 1998, *54*, 934-944.

[13] Nangia, A. Database research in crystal engineering. *Cryst Eng Comm*., 2002, *17*, 93-101.

[14] Rivas, J. C. M. , Brammer, L. Self-assembly of 1-D chains of different topologies using the hydrogen-bonded inorganic supramolecular synthons N-H···Cl2M or N-H···Cl3M. *Inorg. Chem*., 1998, *37*, 4756-4757.

[15] Steiner, T. The hydrogen bond in the solid state. *Angew. Chem. Int. Ed*., 2002, *41*, 48-76.

[16] Steiner, T. Unrolling the hydrogen bond properties of C-HO interactions. *Chem. Commun*., 1997, *1997*, 727-724.

[17] Jones, P. G.; Ahrens, B. Bis(diphenylphosphino)methane and related ligands as hydrogen bond donors. *Chem. Commun*. 1998, *1998*, 2307-2308.

[18] Lommerse, J. P. M.; Stone, A. J.; Taylor, R.; Allen, F. H. The nature and geometry of intermolecular interactions between halogens and oxygen or nitrogen. *J. Am. Chem. Soc*., 1996, *118*, 3108-3116.

[19] Chandrasekhar, V.; Baskar, V.; Kingsley, S.; Nagendran, S.; Butcher, R. J. First characterisation of weak hydrogen bonds in organoantimony compounds: C-H···Cl-Sb-mediated, zig-zag, supramolecular polymeric architecture in $[Ph_2Sb(Cl)\{S_2C_2(CN)_2\}]$. *Cryst. Eng. Comm*., 2001, *17*, 64-66.

[20] Banerjee, R.; Desiraju, G. R.; Mondal, R.; Howard, J. A. K. Organic chlorine as a hydrogen-bridgeacceptor: Evidence for the existence of intramolecular O-H···Cl-C interactions in some gem-alkynols. *Chem. Eur. J*., 2004, *10*, 3373-3383.

[21] Waters, M. L. Aromatic interactions in model systems. *Curr. Opin. Chem. Biol*., 2002, *6*, 736-741.

[22] Gilson, M. K.; Zhou, H. X. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct*., 2007, *36*, 21-42.

[23] Honig, B.; Sharp, K.; Yang, A. S. Macroscopic models of aqueous solutions: biological and chemical applications. *J. Phys. Chem*., 1993, *97*, 1101-1109.

[24] Rashin, A. A. Aspects of protein energetics and dynamics. *Prog. Biophys. Mol. Biol*., 1993, *60*, 73-200.

[25] Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br. J. Pharmacol*., 2008, *153(Suppl. 1*), S7-S26.

[26] Mancera, R. L. Molecular modeling of hydration in drug design. *Curr. Opin. Drug. Discov. Dev*., 2007, *10*, 275-280.

[27] Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug. Discov*., 2004, *3*, 935-949.

[28] Ross, G. A.; Morris, G.M.; Biggin, P. C. Rapid and accurate prediction and scoring of water molecules in protein binding sites. *PLoS ONE*, 2012, *7*, e32036.

[29] Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caisch, A. Exhaustive docking of molecular fragments on protein binding sites with electrostatic solvation. *Proteins: Struct., Funct., Genet.*, 1999, *37*, 88-105.

[30] Zou, X.; Sun, Y.; Kuntz, I. D. Inclusion of solvation in ligand binding free energy calculations using the generalized-Born model. *J. Am. Chem. Soc*,. 1999, *121*, 8033-8043.

[31] Vorobjev, Y. N. Advances in implicit models of water solvent to compute conformational free energy and molecular dynamics of proteins at constant pH. *Adv. Protein Chem. Struct. Biol.*, 2011, *85*, 281-322

[32] Banba, S.; Guo, Z.; Brooks, C. L. III. New free energy based methods for ligand binding from detailes structure-function to multiple-ligand screening. in *Free energy calculations in rational*

drug design. Reddi, M. R.; Erion, M. D. Eds. Kluwer/Plenum Publishers, New York, 2001, pp195-223.

[33] Aqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein. Eng. Des. Sel*., 1994, *7*, 385-391.

[34] Carlson, H. A.; Jorgensen, W. L. An extended linear response method for determining free energies of hydration. *J. Phys. Chem*., 1995, *99*, 10667-10673.

[35] Chen, J.; Brooks, C. L. III; Khandogin, J. Recent advances in implicit solvent-based methods for biomolecular simulations. *Curr. Opin. Struct. Biol*., 2008, *18*, 140-148.

[36] Roux, B.; Simonson, T. Implicit solvent models. *Biophys. Chem.*, 1999, *78*, 1-20.

[37] Simonson, T. Macromolecular electrostatics: continuum models and their growing pains. *Curr. Opin. Struct. Biol*., 2001, *11*, 243-252.

[38] Baker, N. A. Improving implicit solvent simulations: a Poisson-centric view. *Curr. Opin. Struct. Biol*., 2005, *15*, 137-143.

[39] Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting small-molecule solvation free energies: an informal blind test for computational chemistry. *J. Med. Chem*., 2008, *51*, 769-779.

[40] Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. Predictions of hydration free energies from all-atom molecular dynamics simulations. *J. Phys. Chem. B*, 2009, *113*, 4533-4537.

[41] Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small molecule hydration free energies in explicit solvent: an extensive test of fixed-charge atomistic simulations. *J. Chem. Theory. Comput.*, 2009, *5*, 350-358.

[42] Mobley, D. L.; Dill, K. A.; Chodera, J. D. Treating entropy and conformational changes in implicit solvent simulations of small molecules. *J. Phys. Chem. B*, 2008, *112*, 938-946.

[43] Nicholls, A.; Wlodek, S.; Grant, J. A. The SAMP1 solvation challenge: further lessons regarding the pitfalls of parametrization. *J. Phys. Chem. B*, 2009, *113*, 4521-4532.

[44] Li, Z.; Lazaridis, T. Water at biomolecular binding interfaces. *Phys. Chem. Chem. Phys.*, 2007, *9*, 573-581.

[45] de Graaf, C.; Vermeulen, N. P. E.; Feenstra, K. A. Cytochrome P450 in silico: an integrative modeling approach. *J. Med. Chem*., 2005, *48*, 2725-2755.

[46] Barillari, C.; Taylor, J.; Viner, R.; Essex, J. W. Classification of water molecules in protein binding sites. *J. Am. Chem. Soc.* 2007, *129*, 2577-2587.

[47] Lu, Y.; Wang, R.; Yang, C. Y.; Wang, S. Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. *J. Chem. Inf. Model.* 2007, *47*, 668-675.

[48] Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. Enzyme crystal structure in a neat organic solvent. *Proc. Nat. Acad. Sci. U.S.A.* 1993, *90*, 8653-8657.

[49] Raymer, M. L.; Sanschagrin, P. C.; Punch, W. F.; Venkataraman, S.; Goodman, E. D.; Kuhn, L. A. Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbors genetic algorithm. *J. Mol. Biol.* 1997, *265*, 445-464.

[50] Huang, W. J.; Blinov, N.; Wishart, D. S.; Kovalenko, A. Role of Water in Ligand Binding to Maltose-Binding Protein: Insight from a New Docking Protocol Based on the 3D- *Model.*, 2015, *55*, 317-328.

[51] ten Brink, T.; Exner, T. E. Influence of protonation, tautomeric, and stereoisomeric states on protein-ligand docking results. *J. Chem. Inf. Model.*, 2009, *49*, 1535-1546.

[52] Kastritis, P. L.; Bonvin, A. M. J. J. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *J. R. Soc. Interface*, 2012, 10:20120835.

[53] Fersht, A. R.; Shi, J-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter,P.; Waye, M. M. Y.; Winter, G. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature*, 1985, *314*, 235-238.

[54] Bostrom, J.; Norrby, P-O.; Liljefors, T. Conformational energy penalties of protein-bound ligands. *J. Comput. Aided Mol. Des*., 1998 *12*, 383-396.

[55] Roy, S.; Bagchi, B. Free energy barriers for escape of water molecules from protein hydration layer. *J. Phys. Chem. B*, 2012, *116*, 2958-2968.

[56] Ritschel, T.; Philiph, C.; Kohler, P. C.; Neudert, G.; Heine, A.; Diederich, F.; Klebe, G. How to replace the residual solvation shell of polar active site residues to achieve nanomolar inhibition of tRNA-guanine transglycosidase. *ChemMedChem*, 2009, *4*, 2012-2023

[57] Kadirvelraj R, Foley BL, Dyekjaer JD, Woods RJ (2008) Involvement of water in carbohydrate-protein binding: Concanavalin A revisited. J Am Chem Soc 130, 106933-16942.

[58] Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. Grid inhomogeneus solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril. J. *Chem. Phys.*, 2012, *137*, 44-101.

[59] Water…where it matters, when it matters. Available: http://www.eyesopen.com/SZMAP. Accesed 2015 May 13.

[60] A new paradigm in ligand optimization. Available: http://www.schrodinger.com/WaterMap.php. Accesed 2015 May 13.

[61] Genheden, S.; Luchko, T.; Kovalenko, A.; Ryde, U. An MM/3D-RISM approach for ligand binding affinities. *J. Phys. Chem*., 2010, *114*, 8505-8516.

[62] Genheden, S.; Mikulskis, P.; Hu, L.; Kongsted, J.; Söderhjelm, P.; Ryde, U. Accurate predictions of nonnnpolar solvation free energies require explicit consideration of binding-site hydration. *J. Am. Chem. Soc.*, 2011, *133*, 13081-13092.

[63] Ben Naim, A. Solvent effects on protein association and protein folding. *Biopolymers*, 1990, *29*, 567-596.

[64] Li, Z.; Lazaridis, T. Thermodynamics of buried water clusters at a protein at ligand binding interface. *J. Phys. Chem. B*, 2006, *110*, 1464-1475.

[65] Li, Z.; Lazaridis, T. The effect of water displacement on binding thermodynamics: concanavalin A. *J. Phys. Chem. B*, 2005, *109*, 662-670.

[66] Sawada, Y.; Sokabe, M. Molecular dynamics study on protein-water interplay in the mechanogating of the bacterial mechanosensitive channel MscL. *Eur. Biophys. J*., 2015, *44*, 531-543.

[67] Amram, S; Ganoth, A.; Tichon, O.; Peer, D.; Nachliel, E.; Gutman, M.; Tsfadia, Y. Structural Characterization of the Drug Translocation Path of MRP1/ABCC1. Isr. J. Chem., 2014, 54, 1382-1393.

[68] Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.*, 2000, *33*, 889-897.

[69] Wang, J.; Kang, X.; Kuntz, I. D.; Kollman, P. A. Hierarchical database screenings for HIV-1 reverse transcriptase using a pharmacophore model, rigid docking, solvation docking, and MM-PB/SA. *J. Med. Chem.*, 2005, *48*, 2432-2444.

[70] Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.*, 1994, *98*, 1978-1983.

[71] Rush, T. S. III; Manas, E. S.; Tawa, G. J.; Alvarez, J. C. in *Virtual Screening in Drug Discovery* (Alvarez, J.C. and Shoichet, B.K., Eds.), Taylor & Francis, Boca Raton, FL. 2005, pp. 249-277.

[72] Kroemer, R. T. Structure-based drug design: Docking and scoring. *Current Protein and Peptide Science*, 2007, *8*, 312-328.

[73] 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*, 6504-6515

[74] 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*, 172-182.

[75] Goodford, P. J. A. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.*, 1985, *28*, 849-857.

[76] Clarke, C.; Woods, R. J.; Gluska, J.; Cooper, A.; Nutley, M. A.; Boons, G.-J. Involvement of water in carbohydrate-protein binding. *J. Am. Chem. Soc.*, 2001, *123*, 12238-12247.

[77] Rarey, M.; Kramer, B. Lengauer, T. The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins: Structure, Function, and Bioinformatics*, 1999, *34*, 17-28.

[78] Schnecke, V.; Kuhn, L .A. Virtual screening with solvation and ligand-induced complementarity. *Perspect. Drug Discov. Des*., 2000, *20*, 171-190.

[79] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Protein-ligand docking: current status and future challenges. *Proteins: Structure, Function, and Bioinformatics*, 2006, *65*, 15-26.

[80] Kim, R.; Skolnick, J. Assessment of programs for ligand binding affinity prediction. *J. Comput. Chem*., 2008, *29*, 1316-1331.

[81] Ni, H.; Sotriffer, C. A.;, McCammon, J. A. Ordered water and ligand mobility in the HIV-1 integrase-5CITEP complex: a molecular dynamics study. *J. Med. Chem*., 2001, *44*, 3043-3047.

[82] Sarkhel, S.; Desiraju, G. R. N-HO, O-HO, and C-HO hydrogen bonds in protein-ligand complexes: strong and weak interactions in molecular recognition. *Proteins: Structure, Function, and Bioinformatics*, 2004, *54*, 247-259.

[83] Bohm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known threedimensional structure. *J. Comput. Aided Mol. Des*., 1994, *8*, 243-256.

[84] Jain, A. N. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J. Comp. Aided Mol. Des*., 1996, *10*, 427-440.

[85] Deng, Y.; Roux, B. Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J. Chem. Phys*., 2008, *128*, 115103.

[86] Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Nat. Acad. Sci. USA*, 2007, *104*, 808-813.

[87] Roberts, B. C.; Mancera, R. L. Ligand-protein docking with water molecules. *J. Chem. Inf. Mod*., 2008, *48*, 397-408.

[88] Thilagavathi, R.; Mancera, R. L. Ligand-protein cross-docking with water molecules. *J. Chem. Inf. Mod*., 2010, *50*, 415-421.

[89] Lie, M. A.; Thomsen, R.; Pedersen, C. N. S.; Schiøtt, B.; Christensen, M.H. Molecular Docking with Ligand Attached Water Molecules. *J. Chem. Inf. Model.*, 2011, 51, 909-917.

[90] Roe, S. M.; Prodromou, C.; O'Brien, R.; Ladbury, J. E.; Piper, P.W.; Pearl, L. H. Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin. *J Med. Chem*., 1999, *42*, 260-266.

[91] Sleigh, S. H.; Seavers, P. R.; Wilkinson, A. J.; Ladbury, J. E.; Tame, J. R. Crystallographic and calorimetric analysis of peptide binding to OppA protein. *J. Mol. Biol*., 1999, *291*, 393-415.

[92] 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. () Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science*, 1994, *263*, 380-384.

[93] de Beer, S. B.; Vermeulen, N. P.; Oostenbrink, C. The role of water molecules in computational drug design. *Curr. Top. Med. Chem*., 2010, *10*, 55-66.

[94] Mancera, R. L. Molecular modelling of hydration in drug design. *Curr. Opin. Drug. Discov. Devel*., 2007, *10*, 275-280.

[95] Wong, S. E.; Lightstone, F. C. Accounting for water molecules in drug design. *Exp. Opin. Drug. Discov*., 2011, *6*, 65-74.

[96] Hussain, A.; Melville, J.; Hirst, J. Molecular docking and QSAR of aplyronine A and analogues: potent inhibitors of actin. *J. Comput. Aided Mol. Des.*, 2010, *24*, 1-15.

[97] Pastor, M.; Cruciani, G.; Watson, K. A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure-activity relationship analysis. *J. Med. Chem*. 1997, *40*, 4089-4102.

[98] Taha, M. O.; Habash, M.; Al-Hadidi, Z.; Al-Bakri, A.; Younis, K.; Sisan, S. Docking-based comparative intermolecular contacts analysis as new 3-D QSAR concept for validating docking studies and in silico screening: NMT and GP inhibitors as case studies. *J. Chem. Inf. Model.* 2011, *51*, 647-669.

[99] Wallnoefer, H. G.; Handschuh, S.; Liedl, K. R.; Fox, T. Stabilizing of a globular protein by a highly complex water network: a molecular dynamics simulation study on factor Xa. *J. Phys. Chem. B*, 2010, *114*, 7405-7412.

[100] Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen; W. L. Effects of water placement on predictions of binding affinities for p38a MAP kinase inhibitors. *J. Chem. Theory. Comput*., 2010, *6*, 3850-3856.

[101] Wallnoefer, H. G.; Liedl, K R.; Fox, T. A challenging system: Free energy prediction for factor Xa. *J. Comput. Chem*., 2011, *32*, 1743-1752.

[102] de Graaf, C.; Oostenbrink, C; Keizers, P. H.; van der Wijst, T.; Jongejan, A.; Vermeulen, N. P. E. Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking. *J. Med. Chem*., 2006, *49*, 2417-2430.

[103] de Graaf, C.; Pospisil, P.; Pos, W.; Folkers, G.; Vermeulen, N. P. E. Binding mode prediction of cytochrome P450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking. *J. Med. Chem*., 2005, *48*, 2308-2318.

[104] Santos, R.; Hritz, J.; Oostenbrink, C. Role of water in molecular docking simulations of cytochrome P450 2D6. *J. Chem. Inf. Model*. 2010, *50*, 146-154.

[105] Bellocchi, D.; Macchiarulo, A.; Constantino, G.; Pellicciar, R. Docking studies on PARP-1 inhibitors: insights into the role of a binding pocket water molecule. *Bioorg. Med. Chem.* 2005, *13*, 1151-1157.

[106] 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.

[107] Wissner, A.; Berger, D. M.; Boschelli, D. H.; Floyd, M. B. Jr.; Greenberger, L. M.; Gruber, B. C.; Johnson, B. D.; Mamuya, N.; Nilakantan, R.; Reich, M. F.; Shen, R.; Tsou, H-R.; Upeslacis, E.; Wang, Y. F.; Wu, B.; Ye, F.; Zhang, N. 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.*, 2000, *43*, 3244-3256.

[108] Mikol ,V.; Papageorgiou, C.; Borer, X. The role of water molecules in the structure-based design of (5-hydroxynorvaline)-2-cyclosporin: synthesis, biological activity, and crystallographic analysis with cyclophilin A. *J. Med. Chem*., 1995, *38*, 3361-3367.

[109] Garcia-Sosa, A. T.; Mancera, R. L. Free energy calculations of mutations involving a tightly bound water molecule and ligand substitutions in a ligand-protein complex. *Mol. Inf*., 2010, *29*, 589-600.

[110] 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.

[111] Lloyd, D. G.; Garcia-Sosa, A. T.; Alberts, I. L.; Todorov, N. P.; Mancera, R. L. The effect of tightly bound water molecules on the structural interpretation of ligand-derived pharmacophore models. *J. Comp. Aided. Mol. Des*., 2004, *18*, 89-100.

[112] Garcia-Sosa, A. T.; Firth-Clark, S.; Mancera, R. L. Including tightly-bound water molecules in de novo drug design. Exemplification through the in silico generation of poly(ADP-ribose)polymerase ligands. *J. Chem. Inf. Model*., 2005, *45*, 624-633.

[113] Garcia-Sosa, A. T.; Mancera, R. L. The effect of a tightly bound water molecule on scaffold diversity in the computer-aided de novo ligand design of CDK2 inhibitors. *J. Mol. Mod*. 2006, *12*, 422-431.

[114] Mancera, R. L. De novo ligand design with explicit water molecules: An application to bacterial neuraminidase. *J. Comput. Aided Mol. Des*., 2002, *16*, 479-499.

[115] Carugo, O.; Bordo, D. How many water molecules can be detected by protein crystallography? *Acta Crystallogr. D. Biol. Crystallogr*., 1999, *55*, 479-483.

[116] Davis, A. M.; Teague, S. J.; Kleywegt, G. J. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem., Int. Ed*., 2003, *42*, 2718-2736.

[117] Ernst, J. A.; Clubb, R. T.; Zhou, H.-X.; Gronenborn, A. M.; Clore, G. M. Demonstration of positionally disordered water within a protein hydrophobic cavity by NMR. *Science*, 1995, *267*, 1813-1816.

[118] Huang, N.; Shoichet, B. K. Exploiting ordered waters in molecular docking. *J. Med. Chem*., 2008, *51*, 4862-4865.

[119] Kellogg, G. E.; Semus, S. F.; Abraham, D. J. HINT: a new method of empirical hydrophobic field calculation for CoMFA. *J. Comput. Aided Mol. Des.*, 1991, *5*, 545-552.

[120] Chen, D. L.; Kellogg, G. E. A computational tool to optimize ligand selectivity between two similar biomacromolecular targets. *J. Comput. Aided Mol. Des*., 2005, *19*, 69-82.

[121] Amadasi, A.; Spyrakis, F.; Cozzini, P.; Abraham, D. J.; Kellogg, G. E.; Mozzarelli, A. Mapping the energetics of water-protein and water-ligand interactions with the "natural" HINT forcefield: Predictive tools for characterizing the roles of water in biomolecules. *J. Mol. Biol.* 2006, *358*, 289-309.

[122] Amadasi, A.; Surface, J. A.; Spyrakis, F.; Cozzini, P.; Mozzarelli. A.; Kellogg, G. E. Robust classification of ''relevant'' water molecules in putative protein binding sites. *J. Med. Chem*., 2008, *51*, 1063-1067.

[123] Garcia-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*, 172-182.

[124] 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*, 455-461.

[125] Purkiss, A.; Skoulakos, S.; Goodfellow, J. M. The protein-solvent interface: a big splash. *Philos. Trans. R. Soc. London Ser. A*, 2001, *359*, 1515-1527.

[126] Nakasako, M. Large-scale networks of hydration water molecules around bovine beta-trypsin revealed by cryogenic X-ray crystal structure analysis. *J. Mol. Biol.*, 1999, *289*, 547-564.

[127] Higo, J.; Nakasako, M. Hydration structure of human lysozyme investigated by molecular dynamics simulation and cryogenic X-ray crystal structure analyses: On the correlation between crystal water sites, solvent density, and solvent dipole. *J. Comput. Chem.*, 2002, *23*, 1323-1336.

[128] Yokomizo, T.; Higo, J.; Nakasako, M. Patterns and networks of hydrogen-bonds in the hydration structure of human lysozyme. *Chem. Phys. Lett.*, 2005, *410*, 31-35.

[129] Poornima, C. S.; Dean, P. M. Hydration in drug design 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. *J. Comput-Aided Mol. Des.*, 1995, *9*, 500-512.

[130] Hendlich, M.; Bergner, A.; Günter, J.; Klebe, G. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.*, 2003, *326*, 607-620.

[131] Poornima, C. S.; Dean, P. M. Hydration in drug design. 2. Influence of local site surface shape on water binding. *J. Comput-Aided Mol. Des.*, 1995, *9*, 513-520.

[132] Chung, E.; Henriques, D.; Renzoni, D.; Zvelebil, M.; Bradshaw, J. M.; Waksman, G.; Robinson, C. V.; Ladbury, J. E. Mass spectrometric and thermodynamic studies reveal the role of water molecules in complexes formed between SH2 domains and tyrosyl phosphopeptides. *Struct. Fold. Des.*, 1998, *6*, 1141-1151.

[133] Baker, E. N.; Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.*, 1984, *44*, 97-179.

[134] Sreenivasan, U.; Axelsen, P. H. Buried water in homologous serine proteases. *Biochemistry*, 1992, *31*, 12785-12791.

[135] Loris, R.; Stas, P. P.; Wyns, L. Conserved waters in legume lectin crystal structures. The importance of bound water for the sequence-structure relationship within the legume lectin family. *J. Biol. Chem.*, 1994, *269*, 26722-26733.

[136] Shaltiel, S.; Cox, S.; Taylor, S. S. Conserved water molecules contribute to the extensive network of interactions at the active site of protein kinase A. *Proc. Natl. Acad. Sci., U.S.A.* 1998, *95*, 484-491.

[137] Sanschagrin, P. C.; Kuhn, L. A. Cluster analysis of consensus water sites in thrombin and trypsin shows conservation between serine proteases and contributions to ligand specificity. *Protein Sci.*, 1998, *7*, 2054-2064.

[138] Krem, M. M.; Di Cera, E. Conserved water molecules in the specificity pocket of serine proteases and the molecular mechanism of Na^+ binding. *Proteins: Structure, Function, and Bioinformatics*, 1998, *30*, 34-42.

[139] Loris, R.; Langhorst, U.; De Vos, S.; Decanniere, K.; Bouckaert, J.; Maes, D.; Transue, T. R.; Steyaert, J. Conserved water molecules in a large family of microbial ribonucleases. *Proteins: Structure, Function, and Bioinformatics*, 1999, *36*, 117-134.

[140] Ogata, K.; Wodak, S. J. Conserved water molecules in MHC class-I molecules and their putative structural and functional roles. *Protein Eng.*, 2002, *15*, 697-705.

[141] Bottoms, C. A.; Smith, P. E.; Tanner, J. J. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.*, 2002, *11*, 2125-2137.

[142] Prasad, B. V. L. S.; Suguna, K. Role of water molecules in the structure and function of aspartic proteinases. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, *58*, 250-259.

[143] Bottoms, C. A.; Schuermann, J. P.; Agah, S.; Henzl, M. T.; Tanner, J. J. Crystal structure of rat R-parvalbumin at 1.05 Å resolution. *Protein Sci.*, 2004, *13*, 1724-1734.

[144] Bottoms, C. A.; White, T. A.; Tanner, J. J. Exploring structurally conserved solvent sites in protein families. *Proteins: Structure, Function, and Bioinformatics*, 2006, *64*, 404-421.

[145] Zhang, X.-J.; Matthews, B. W. Conservation of solvent-binding sites in 10 crystal forms of T4 lysozyme. *Protein Sci.*, 1994, *3*, 1031-1039.

[146] Carrell, H. L.; Glusker, J. P.; Burger, V.; Manfre, F.; Tritsch, D.; Biellmann, J.-F. X-ray analysis of D-xylose isomerase at 1.9 Å: native enzyme in complex with substrate and with a mechanism-designed inactivator. *Proc. Natl. Acad. Sci. U.S.A.*, 1989, *86*, 4440-4444.

[147] Faerman, C. H.; Karplus, P. A. Consensus preferred hydration sites in six FKBP12-drug complexes. *Proteins: Structure, Function, and Bioinformatics*, 1995, *23*, 1-11.

[148] Poornima, C. S.; Dean, P. M. Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J. Comput.-Aided Mol. Des.*, 1995, *9*, 521-531.

[149] Babor, M.; Sobolev, V.; Edelman, M. Conserved positions for ribose recognition: importance of water bridging interactions among ATP, ADP and FAD-protein complexes. *J. Mol. Biol.*, 2002, *323*, 523-532.

[150] Powers, R. A.; Shoichet, B. K. Structure-based approach for binding site idenfication on AmpC beta-lactamase. *J. Med. Chem.*, 2002, *45*, 3222-3234.

[151] Mustata, G.; Briggs, J. M. Cluster analysis of water molecules in alanine racemase and their putative structural role. *Protein Eng.*, 2004, *17*, 223-234.

[152] Boström, J.; Hogner, A.; Schmitt, S. Do structurally similar ligands bind in a similar fashion? *J. Med. Chem.*, 2006, *49*, 6716-6725.

[153] Huang, K.; Lu, W.; Anderson, S.; Laskowski, M.; James, M. N. G. Water molecules participate in proteinase-inhibitor interactions: crystal structure of Leu18, Ala18 and Gly18 variants of turkey ovomucoid inhibitor third domain complexed with *Streptomyces griseus* proteinase B. *Protein Sci.*, 1995, *4*, 1985-1997.

[154] Engh, R. A.; Brandstetter, H.; Sucher, G.; Eichinger, A.; Baumann, U.; Bode, W.; Huber, R.; Poll, T.; Rudolph, R.; von der Saal, W. Enzyme flexibility, solvent and 'weak' interactions characterize thrombin-ligand interactions: Implications for drug design. *Structure*, 1996, *4*, 1353-1362.

[155] Rejto, P. A.; Verkhivker, G. M. Mean field analysis of FKBP12 complexes with FK506 and rapamycin: Implications for a role of crystallographic water molecules in molecular recognition and specificity. *Proteins: Structure, Function, and Bioinformatics*, 1997, *28*, 313-324.

[156] Rutenber, E. E.; Stroud, R. M. Binding of the anticancer drug ZD1694 to E. coli thymidylate synthase: Assessing specificity and affinity. *Structure*, 1996, *4*, 1314-1324.

[157] Finley, J. B.; Atigadda, V. R.; Duarte, F.; Zhao, J. J.; Brouillette, M. J.; Air, G. M.; Luo, M. Novel aromatic inhibitors of influenza virus neuraminidase make selective interactions with conserved residues and water molecules in the active site. *J. Mol. Biol.*, 1999, *293*, 1107-1119.

[158] Palomer, A.; Perez, J. J.; Navea, S.; Llorens, O.; Pascual, J.; Garcia, L.; Mauleon, D. Modeling cyclooxygenase inhibition. Implication of active site hydration on the selectivity of ketoprofen analogues. *J. Med. Chem.*, 2000, *43*, 2280-2284.

[159] Vogt, J.; Perozzo, R.; Pautsch, A.; Prota, A.; Schelling, P.; Pilger, B.; Folkers, G.; Scapozza, L.; Schulz, G. E. Nucleoside binding site of Herpes simplex type 1 thymidine kinase analyzed by X-ray crystallography. *Proteins: Structure, Function, and Bioinformatics*, 2000, *41*, 545-553.

[160] Lemieux, R. U. How water provides the impetus for molecular recognition in aqueous solution. *Acc. Chem. Res.*, 1996, *29*, 373-380.

[161] Wester, M. R.; Johnson, E. F.; Marques-Soares, C.; Dijols, S.; Dansette, P. M.; Mansuy, D.; Stout, C. D. Structure of mammalian cytochrome P4502C5 complexed with diclofenac at 2.1 angstrom resolution: Evidence for an induced fit model of substrate binding. *Biochemistry*, 2003, *42*, 9335-9345.

[162] Pujadas, G.; Palau, J. Molecular mimicry of substrate oxygen atoms by water molecules in the α-amylase active site. *Protein Sci.*, 2001, *10*, 1645-1657.

[163] Wang, T.; Wade, R. C. Comparative binding energy (COMBINE) analysis of influenza neuraminidase-inhibitor complexes. *J. Med. Chem.*, 2001, *44*, 961-971.

[164] Anstead, G. M; Carlson, K. E.; Katzenellenbogen, J. A. The estradiol pharmacophore: Ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids*, 1997, *62*, 268-303.

[165] Grúneberg, S.; Stubbs, M. T.; Klebe, G. Successful virtual screening for novel inhibitors of human carbonic anhydrase: Strategy and experimental confirmation. *J. Med. Chem.*, 2002, *45*, 3588-3602.

[166] Brenk, R.; Naerum, L.; Grädler, U.; Gerber, H.-D.; Garcia, G. A.; Reuter, K.; Stubbs, M. T.; Klebe, G. Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis. *J. Med. Chem.*, 2003, *46*, 1133-1143.

[167] Pospisil, P.; Kuoni, T.; Scapozza, L.; Folkers, G. Methodology and problems of protein-ligand docking: Case study of dihydroorotate dehydrogenase, thymidine kinase, and phosphodiesterase. 4. *J. Recept. Signal Transduct. Res.*, 2002, *22*, 141-154.

[168] Dunitz, J. D. The entropic cost of bound water in crystals and biomolecules. *Science*, 1994, *264*, 670-671.

[169] Dunitz, J. D. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.*, 1995, *2*, 709-712.

[170] 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. *Chem. Biol.*, 1996, *3*, 973-980.

[171] Li, Z.; Lazaridis, T. Thermodynamic contributions of the ordered water molecule in HIV-1 protease. *J. Am. Chem. Soc.*, 2003, *125*, 6636-6637.

[172] Fornabaio, M.; Spyrakis, F.; Mozzarelli, A.; Cozzini, P.; Abraham, D. J.; Kellogg, G. E. Simple, intuitive calculations of free energy of binding for protein-ligand complexes. 3. The free energy contribution of structural water molecules in HIV-1 protease complexes. *J. Med. Chem.*, 2004, *47*, 4507-4516.

[173] Cozzini, P.; Fornabaio, M.; Marabotti, A.; Abraham, D. J.; Kellogg, G. E.; Mozzarelli, A. Free energy of ligand binding to protein: Evaluation of the contribution of water molecules by computational methods. *Curr. Med. Chem.*, 2004, *11*, 3093-3118.

[174] Kříž, Z.; Otyepka, M.; Bartová, I.; Koča, J. Analysis of CDK2 active site hydration: A method to design new inhibitors. *Proteins: Structure, Function, and Bioinformatics*, 2004, *55*, 258-274.

[175] Lu, Y.; Yang, C.-Y.; Wang, S. Binding free energy contributions of interfacial waters in HIV-1 protease/inhibitor complexes. *J. Am. Chem. Soc.*, 2006, *128*, 11830-11839.

[176] Helms, V.; Wade, R. C. Thermodynamics of water mediating protein-ligand interactions in cytochrome P450CAM. A molecular dynamics study. *Biophys. J.*, 1995, *69*, 810-824.

[177] Helms, V.; Wade, R. C. Hydration energy landscape of the active site cavity of cytochrome P450. *Proteins: Structure, Function, and Bioinformatics*, 1998, *32*, 381-396.

[178] Helms, V.; Wade, R. C. Computational alchemy to calculate absolute protein-ligand binding free energy. *J. Am. Chem. Soc.*, 1998, *120*, 2710-2713.

[179] Hamelberg, D.; McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: double-decoupling method. *J. Am. Chem. Soc.*, 2004, *126*, 7683-7689.

[180] Ehrlich, L.; Reczko, M.; Bohr, H.; Wade, R. C. Prediction of protein hydration sites from sequence by modular neural networks. *Protein Eng.*, 1998, *11*, 11-19.

[181] Henchman, R. H.; McCammon, J. A. Extracting hydration sites around proteins from explicit water simulations. *J. Comput. Chem.*, 2002, *23*, 861-869.

[182] Kortveylesi, T.; Dennis, S.; Silberstein, M.; Brown, L.; Vajda, S. Algorithms for computational solvent mapping of proteins. *Proteins: Structure, Function, and Bioinformatics*, 2003, *51*, 340-351.

[183] Schymkowitz, J. W. H.; Rousseau, F.; Martins, I. C.; Ferkinghoff-Borg, J.; Stricher, F.; Serrano, L. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. U.S.A.*, 2005, *102*, 10147-10152.

[184] Carugo, O. Correlation between occupancy and B factor of water molecules in protein crystal structures. *Protein Eng.*, 1999, *12*, 1021-1024.

[185] Carugo, O.; Argos, P. Accessibility to internal cavities and ligand binding sites monitored by protein crystallographic thermal factors. *Proteins: Structure, Function, and Bioinformatics*, 1998, *31*, 201-213.

[186] Marrone, T. J.; Briggs, J. M.; McCammon, J. A. Structure-based drug design: Computational advances. *Annu. Re*V. *Pharmacol.*, 1997, *37*, 71-90.

[187] Holdgate, G. A.; Tunnicliffe, A.; Ward, W. H. J.; Weston, S. A.; Rosenbrock, G.; Barth, P. T.; Taylor, I. W. F.; Pauptit, R. A.; Timms, D. The entropic penalty of ordered water accounts for weaker binding of the antibiotic novobiocin to a resistant mutant of DNA gyrase: A thermodynamic and crystallographic study. *Biochemistry*, 1997, *36*, 9663-9673.

[188] Cherbavaz, D. B.; Lee, M. E.; Stroud, R. M.; Koschl, D. E. Active site water molecules revealed in the 2.1 angstrom resolution structure of a site-directed mutant of isocitrate dehydrogenase. *J. Mol. Biol.*, 2000, *295*, 377-385.

[189] Pickett, S. D.; Sherborne, B. S.; Wilkinson, T.; Bennett, J.; Borkakoti, N.; Broadhurst, M.; Hurst, D.; Kilford, I.; McKinnell, M.; Jones, P. S. Discovery of novel low molecular weight inhibitors of IMPDH via virtual needle screening. *Bioorg. Med. Chem. Lett.*, 2003, *13*, 1691-1694.

[190] García-Sosa, A. T.; Mancera, R. L. The effect of tightly-bound water molecules on scaffold diversity in the computer-aided de novo ligand design of CDK-2 inhibitors. *J. Mol. Model.*, 2006, *12*, 422-431.

[191] 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.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.*, 2004, *47*, 1739-1749.

[192] Rao, M. S.; Olson, A. J. Modelling of factor Xa-inhibitor complexes: A computational flexible docking approach. *Proteins: Structure, Function, and Bioinformatics*, 1999, *34*, 173-183.

[193] Minke, W. E.; Diller, D. J.; Hol, W. G.; Verlinde, C. L. The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: Heat-labile enterotoxin, a multivalent test case. *J. Med. Chem.*, 1999, *42*, 1778-1788.

[194] Österberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: Incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins: Structure, Function, and Bioinformatics*, 2002, *46*, 34-40.

[195] Floriano, W. B.; Vaidehi, N.; Zamanakos, G.; Goddard, W. A., III. HierVLS hierarchical docking protocol for virtual ligand screening of large-molecule databases. *J. Med. Chem.*, 2004, *47*, 56-71.

[196] Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A new test set for validating predictions of protein-ligand interaction. *Proteins: Structure, Function, and Bioinformatics*, 2002, *49*, 457-471.

[197] Yang, J.-M.; Chen, C.-C. GEMDOCK: A generic evolutionary method for molecular docking. *Proteins: Structure, Function, and Bioinformatics*, 2004, *55*, 288-304.

[198] de Graaf, C.; Oostenbrink, C.; Keizers, P. H. J.; van der Wijst, T.; Jongejan, A.; Vermeulen, N. P. E. Prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking. *J. Med. Chem.*, 2006, *49*, 2417-2430.

[199] Mendez, R.; Leplae, R.; Lensink, M. F.; Wodak, S. J. Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures. *Proteins: Structure, Function, and Bioinformatics*, 2005 60, 150-169.

[200] Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, 2002, *47*, 409-43.

[201] van Dijk, A. D. J.; Boelens, R.; Bonvin, A. M. J. J. Data-driven docking for the study of biomolecular complexes. *FEBS J.*, 2005, *272*, 293-312.

[202] Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature*, 2005, *437*, 640-647.

[203] van Dijk, A. D. J.; Bonvin, A. M. J. J. Solvated docking: introducing water into the modelling of biomolecular complexes. *Bioinformatics*, 2006, *22*, 2340-2347.

[204] Raschke,T. M. Water structure and interactions with protein surfaces. *Curr. Opin. Struct. Biol*., 2006, *16*, 152-159.

[205] Houborg, K.; Harris, P.; Petersen, J.; Rowland, P.; Poulsen, J-C. N.; Schneider, P.; Vind, J.; Larsen, S. Impact of the physical and chemical environment on the molecular structure of Coprinus cinereus peroxidase. *Acta Crystallogr. D*, 2003, *59*, 989-996.

[206] Yang, J. M.; Chen, C. C. GEMDOCK: A generic evolutionary method for molecular docking. *Proteins: Structure, Function, and Bioinformatics*, 2004, *55*, 288-304.

[207] Moitessier, N.; Westhof, E.; Hanessian, S. Docking of Aminoglycosides to hydrated and flexible RNA. *J. Med. Chem*., 2006, *49*, 1023-1033.

[208] Jiang, L.; Kuhlman, B.; Kortemme, T.; Baker, D. A 'solvated rotainer' approach to modeling water-mediated hydrogen bonds at protein-protein interfaces. *Proteins: Structure, Function, and Bioinformatics*, 2005, *58*, 893-904.

[209] Rodier, F.; Bahadur, R. P.; Chakrabarti, P.;l Janin, J. Hydration of protein-protein interfaces. *Proteins: Structure, Function, and Bioinformatics*, 2005, *60*, 36-45.

[210] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, 1998, *19*, 1639−1662.

[211] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.*, 1996, *261*, 470-489.

[212] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, 1997, *267*, 727−748.

[213] Wang, R. X.; Lu, Y. P.; Wang, S. M. Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem*., 2003, *46*, 2287−2303.

[214] Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. Comparison of automated docking programs as virtual screening tools. *J. Med. Chem*., 2005, *48*, 962−976.

[215] Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of docking performance: comparative data on docking algorithms. *J. Med. Chem.*, 2004, *47*, 558−565.

[216] Kontoyianni, M.; Sokol, G. S.; McClellan, L. M. Evaluation of library ranking efficacy in virtual screening. *J. Comput. Chem.*, 2005, *26*, 11−22.

[217] Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins: Structure, Function, and Bioinformatics*, 2004, *56*, 235−249.

[218] Wang, R. X.; Lu, Y. P.; Fang, X. L.; Wang, S. M. An extensive test of 14 scoring functions using the PDB-bind refined set of 800 protein-ligand complexes. *J. Chem. Inf. Comput. Sci*., 2004, *44*, 2114−2125.

[219] Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem*., 2000, *43*, 4759−4767.

[220] Paul, N.; Rognan, D. ConsDock: A new program for the consensus analysis of protein-ligand interactions. *Proteins: Structure, Function, and Bioinformatics*, 2002, *47*, 521−533.

[221] McConkey, B. J.; Sobolev, V.; Edelman, M. The performance of current methods in ligand-protein docking. *Curr. Sci.*, 2002, *83*, 845−856.

[222] Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision Glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem*., 2006, *49*, 6177−6196.

[223] 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.

[224] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins: Structure, Function, and Bioinformatics*, 1999, *34*, 17−28.

[225] Liu, J.; He, X.; Zhang, J. Z. H. Improving the Scoring of Protein−Ligand Binding Affinity by Including the Effects of Structural Water and Electronic Polarization. *J. Chem. Inf. Model.*, 2013, *53*, 1306−1314.

[226] Corbeil, C. R.; Therrien, E.; Moitessier, N. Modeling reality for optimal docking of small molecules to biological targets. *Curr. Comp. Aided Drug Des*., 2009, *5*, 241-263.

[227] B-Rao, C.; Subramanian, J.; Sharma, S. D.. Managing protein flexibility in docking and its applications. *Drug Discov. Today*, 2009, *14*, 394-400.

[228] Bolstad, E. S.; Anderson, A. C. In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking. *Proteins: Structure, Function, and Bioinformatics*, 2009, *75*, 62-74.

[229] Sánchez-Moreno, M.; Gómez-Contreras, F.; Navarro, P.; Marín, C.;Ramírez-Macías, I.; Rosales, M. J.; Campayo, L.; Cano, C.; Sanz, A. M.; Yunta, M. J. R. Imidazole-containing phthalazine derivatives inhibitFe-SOD performance in Leishmania species and are active in vitro against visceral and mucosal leishmaniasis. Parasitology, 2015, 142, 1115-1129.

[230] Sánchez-Moreno,M.; Gómez-Contreras,F.; Navarro, P.; Marín, C.; Olmo, F.; Yunta, M. J. R.; Sanz, A. M.; Rosales, M. J.; Cano, C.; Campayo, L. Phthalazine Derivatives Containing Imidazole Rings Behave as Fe-SOD Inhibitors and Show Remarkable Anti-T. cruzi Activity in Immunodeficient-Mouse Mode of Infection. *J. Med. Chem*. 2012, *55*, 9900−9913

[231] Rueda, M.; Bottegoni, G.; Abagyan, R. Consistent improvement of cross-docking results using binding site ensembles generated with elastic network normal modes. *J. Chem. Inf. Model*., 2009, *49*, 716-725.

[232] Kazemi, S.; Kruger, D. M.; Sirockin, F.; Gohlke, H. Elastic potential grids: Accurate and efficient representation of intermolecular interactions for fully flexible docking. *ChemMedChem*, 2009, *4*, 1264-1268.

[233] Davis, I. W.; Raha, K.; Head, M. S.; Baker, D. Blind docking of pharmaceutically relevant compounds using RosettaLigand. *Protein Sci*., 2009, *18*, 1998-2002.

[234] Fong, P.; McNamara, J. P.; Hillier, I. H.; Bryce, R. A. Assessment of QM/MM scoring functions for molecular docking to HIV-1 protease. *J. Chem. Inf. Model*., 2009, *4*, 913-924.

[235] Cincilla, G.; Vidal, D.; Pons, M. An improved scoring function for suboptimal polar ligand complexes. *J. Comput. Aided Mol. Des*., 2009, *23*, 143-152.

[236] Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.;Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I.D. DOCK 6: Combining techniques to model RNA-small molecule complexes. *RNA*, 2009, *15*, 1219-1230.

[237] Huang, Z.; Wong, C. F. Docking flexible peptide to flexible protein by molecular dynamics using two implicit-solvent models: An evaluation in protein kinase and phosphatase systems. *J. Phys. Chem. B*, 2009, *113*, 14343-14354.

[238] Villacanas, O.; Madurga, S.; Giralt, E.; Belda, I. Explicit treatment of water molecules in protein-ligand docking. *Curr. Comp. Aided Drug Des.*, 2009, *5*, 145-154.

[239] Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 4. Are popular scoring functions accurate for this class of proteins? *J. Chem. Inf. Model*., 2009, *49*, 1568-1580.

[240] Horbert, R.; Pinchuk, B.; Johannes, E.; Schlosser, J.; Schmidt, D.; Cappel, D.; Totzke, F.; Schächtele, C.; Peifer, C. Optimization of potent DFG-in inhibitors of platelet derived growth factor receptorβ (PDGF-Rβ) guided by water thermodynamics. *J. Med. Chem*., 2015, *58*, 170-182.

[241] Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. Force field validation using protein side chain prediction. *J. Phys. Chem. B*, 2002, *106*, 11673−11680.

[242] Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Bioinf.*, 2004, *55*, 351−367.

[243] Kumar, A.; Zhang, K. Y. J. Investigation on the effect of key water molecules in docking performance in CSARdock exercise. *J. Chem. Inf. Model*., 2013, *53*, 1880-1892.

[244] Yu, B.; Blaber, M.; Gronenborn, A. M.; Clore, G. M.; Caspar, D. L. D. Disordered water within a hydrophobic protein cavity visualized by x-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.*, 1999, *96*, 103−108.

[245] SZMAP, version 1.2.0.7; OpenEye Scientific Software, Inc., Santa Fe, NM, USA, 2012; www.eyesopen.com. Accessed 2015, May, 21

[246] Bayden, A. S.; Moustakas, D. T.; Joseph-McCarthy, D.; Lamb, M. L. Evaluating free energies of binding and conservation of crystallographic waters using SZMAP. *J. Chem. Inf. Model*., 2015, *55*, 1552−1565.

[247] Chen, Z.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L.; Culberson, C.; Shafer, J. A.; and Kuo, L. C. Crystal structure at 1.9-A resolution of human immunodeficiency virus (HIV) II protease complexed with L-735, 524, an orally bioavailable inhibitor of the HIV proteases. *J. Biol. Chem.*, 1994, *269*, 26344-26348.

[248] Klein, S. I.; Czekaj, M.; Gardner, C. J.; Guertin, K. R.; Cheney, D. L.; Spada, A. P.; Bolton, S. A.; Brown, K.; Colussi, D.; Heran, C. L.; Morgan, S. R.; Leadley, R. J.; Dunwiddie, C. T.; Perrone, M. H.; Chu, V. Identification and initial structure-activity relationships of a novel class of nonpeptide inhibitors of blood coagulation factor Xa. *J. Med. Chem.*, 1998, *41*, 437-450.

[249] Herron, D. K.; Goodson, T., Jr.; Wiley, M. R.; Weir, L. C.; Kyle, J. A.; Yee, Y. K.; Tebbe, A. L.; Tinsley, J. M.; Mendel, D.; Masters, J. J.; Franciskovich, J. B.; Sawyer, J. S.; Beight, D. W.; Ratz, A. M.; Milot, G.; Hall, S. E.; Klimkowski, V. J.; Wikel, J. H.; Eastwood, B. J.; Towner, R. D.; Gifford-Moore, D. S.; Craft, T. J.; Smith, G. F. 1,2-Dibenzamidobenzene inhibitors of human factor Xa. *J. Med. Chem.*, 2000, *43*, 859-872.

[250] Yee, Y. K.; Tebbe, A. L.; Linebarger, J. H.; Beight, D. W.; Craft, T. J.; Gifford-Moore, D.; Goodson, T., Jr.; Herron, D. K.; Klimkowski, V. J.; Kyle, J. A.; Sawyer, J. S.; Smith, G. F.; Tinsley, J. M.; Towner, R. D.; Weir, L.; Wiley, M. R. N(2)-Aroylanthranilamide inhibitors of human factor Xa. *J. Med. Chem.*, 2000, *43*, 873-882.

[251] Birch, L.; Murray, C. W.; Hartshorn, M. J.; Tickle, I. J.; Verdonk, M. L. Sensitivity of molecular docking to induced fit effects in influenza virus neuraminidase. *J. Comput.-Aided Mol. Des.*, 2002, *16*, 855-869.

[252] Pitt, W. R.; Goodfellow, J. M. Modelling of solvent positions around polar groups in proteins. *Protein Eng.*, 1991, *4*, 531-537.

[253] Goodford, P. J. A computational-procedure for determining energetically favorable binding-sites on biologically important macromolecules. *J. Med. Chem.*, 1985, *28*, 849-857.

[254] Miranker, A.; Karplus, M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins: Structure, Function, and Bioinformatics*, 1991, *11*, 29-34.

[255] Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: A knowledge based approach for identifying interaction sites in proteins. *J. Mol. Biol.*, 1999, *289*, 1093-1108.

[256] Gunther, J.; Bergner, A.; Hendlich, M.; Klebe, G. Utilising structural knowledge in drug design strategies: applications using Relibase. *J. Mol. Biol.*, 2003, *326*, 621-636.

[257] Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins: Structure, Function, and Bioinformatics*, 1990, *8*, 195-202.

[258] Thilagavathi, R.; Mancera, R. L. Ligand-Protein cross-docking with water molecules. *J. Chem. Inf. Model*., 2010, *50*, 415-421.

[259] Hatshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T. M.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein-ligand docking performance. *J. Med. Chem*., 2007, *50*, 726-741.

[260] Ball, P. Water as an active constituent in cell biology. *Chem. Rev*., 2008, *108*, 74-108.

[261] Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.*, 2002, *99*, 5982-5987.

[262] Levy, Y.; Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct*., 2006, *35*, 389-415

[263] Lemmon, G., Meiler, J. Towards ligand docking including explicit interface water molecules. *PLosOne*, 2013, *8*, e67536.

[264] http://www.moldiscovery.com/ (accesed 10/15/2015).

[265] http://www.schrodinger.com/WaterMap.php (accesed 10/15/2015)

[266] Weldon DJ, Shah F, Chittiboyina AG, Sheri A, Chada RR, Gut J, Rosenthal PJ, Shivakumar D, Sherman W, Desai P, Jung JC, Avery MA. Synthesis, biological evaluation, hydration site thermodynamics, and chemical reactivity analysis of α-keto substituted peptidomimetics for the inhibition of Plasmodium falciparum. Bioorg Med Chem Letts, 24 (5): 1274-1279, 2014.

[267] Horbert R, Pinchuk B, Johannes E, Schlosser J, Schmidt D, Cappel D, Totzke F, Schächtele C, and Peifer C. Optimization of Potent DFG-in Inhibitors of Platelet Derived Growth Factor Receptorβ (PDGF-Rβ) Guided by Water Thermodynamics. J Med Chem, 58 (1): 170-182, 2015.

[268] Pearlstein RA, Sherman W, Abel R. Contributions of water transfer energy to protein-ligand association and dissociation barriers: Watermap analysis of a series of p38α MAP kinase inhibitors. Proteins, 81 (9): 1509-1526, 2013.

[269] Beuming T, Che Y, Abel R, Kim B, Shanmugasundaram V, Sherman W. Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. Proteins, 80: 871-83, 2012.

[270] Ren, P.; Chun, J.; Thomas, D.G.; Schnieders, M. J.; Marucho, M.; Zhang, J.; Baker, N. A. Biomolecular electrostatics and solvation: a computational perspective. *Q. Rev. Biophys*., 2012, *45*, 427-491.

[271] http://giant.hgc.jp/. (accesed 10/19/2015).