

# Advances in the treatment of explicit water molecules in docking and binding free energy calculations

Xiao Hu,<sup>a</sup> Irene Maffucci,<sup>b</sup> and Alessandro Contini<sup>a\*</sup>

<sup>a</sup> Università degli Studi di Milano, Dipartimento di Scienze Farmaceutiche, Sezione di Chimica Generale e Organica “A. Marchesini”, Via Venezian, 21 20133 Milano, Italy

<sup>b</sup> Pasteur, Département de Chimie, École Normale Supérieure, PSL Research University, Sorbonne Universités, UPMC Univ. Paris 06, CNRS, 75005 Paris, France

\* corresponding author: e-mail [alessandro.contini@unimi.it](mailto:alessandro.contini@unimi.it); tel. +39.02.50314480

## Abstract

**Background:** The inclusion of direct effects mediated by water during the ligand-receptor recognition is a hot-topic of modern computational chemistry applied to drug discovery and development. Docking or virtual screening with explicit hydration is still debatable, despite the successful cases that have been presented in the last years. Indeed, how to select the water molecules that will be included in the docking process or how the included waters should be treated remain open questions.

**Objective:** In this review, we will discuss some of the most recent methods that can be used in computational drug discovery and drug development when the effect of a single water, or of a small network of interacting waters, needs to be explicitly considered.

**Results:** Here, we analyse software to aid the selection, or to predict the position, of water molecules that are going to be explicitly considered in later docking studies. We also present software and protocols able to efficiently treat flexible water molecules during docking, including examples of applications. Finally, we discuss methods based on molecular dynamics simulations that can be used to integrate docking studies or to reliably and efficiently compute binding energies of ligands in presence of interfacial or bridging water molecules.

**Conclusions:** Software applications aiding the design of new drugs that exploit water molecules, either as displaceable residues or as bridges to the receptor, are constantly being developed. Although further validation is needed, workflows that explicitly consider water will probably become a standard for computational drug discovery soon.

## 1. Introduction

Since the beginning of modern medicinal chemistry, scientists have relied on models to explain or to improve the activity of drugs. Nowadays, methodology advances in computational chemistry have made molecular models much more accurate, even if they still include many simplifications and approximations. Nevertheless, molecular

This is a post print version of the article published in Curr Med Chem. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

modelling is now an obliged passage of drug discovery and development process.[1–5] The needs for higher accuracy, as well as the constantly increasing computational power, drives developers to create more sophisticated models.

The inclusion of direct effects mediated by water during the ligand-receptor recognition is one of the topics that are currently drawing much of the attention. The analysis of a few thousands of crystallographic protein-ligand complexes showed that one or more water molecules are involved in ligand binding,[6,7] Moreover, Lu and co-workers found that more than 50% of the water molecules observed in 392 high-resolution complexes actually were bridging waters.[8]

The potential role of water in the drug design process has been long recognized. Medicinal chemists often dealt with the possibility to displace a water molecule from the binding site benefitting from the entropy change during the binding process, or to use it as a bridge anchoring the ligand to the receptor.[9] However, making a smart choice between the two strategies is only possible if the binding thermodynamics of the considered water can be reasonably predicted, since no simple experimental methods exist to evaluate the energetics of individual water molecules.

The importance of considering water in computational drug design has already been reviewed and some of the applications that can be used for this purpose have also been discussed.[10–12] However, considering the constant improvements in software design and the awareness of usefulness in applications over years, in this article we intend to provide to medicinal chemists a synoptic discussion of the available computational methods. These include principally those developed in the last decade, which are capable of taking water into account in drug design.

To include explicit water molecules into docking or virtual screening studies, some crucial questions need to be considered: how the water molecules are positioned? Which ones are the most important to be included for later study? How explicit hydration is treated by the docking software of preference? Should the water molecules be fully flexible during docking? How water displaceability and free energy or score contribution is considered?

In the following sections, we will try to address these questions alongside with potential solutions, and provide current options related to docking or virtual screening with explicit hydration.

Indeed, we will analyse some software that can aid the selection of important water, whose coordinates are obtained by experiments or by calculations, for later docking studies. Then, we will discuss how flexible water molecules can be efficiently treated during docking. Finally, we will report on some of the recent methods based on molecular dynamics (MD) simulations that can be used to tackle specific question related to the role played by water molecules in the ligand-receptor recognition process.

## **2. Hydration Sites Selections or Predictions**

### ***2.1. Water selection methods based on experimental data***

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

One of the prevalent and acknowledged water positioning methods is the use of high-resolution X-ray crystallographic structures. This method is generally trustworthy and able to provide precious information on tightly bound water molecules at the binding site.[13,14] Nevertheless, it is important to be aware that X-ray structures, though being the most resolved representation of molecular geometry we can currently generate by experiments, provide rather “static” information, collected at a relatively low temperature and with a well-ordered structural arrangement. Moreover, interpreting the electron density to obtain the averaged position of the atoms provides that the outcome for some water molecules remains dubious, especially at a relatively poor resolution. Additionally, the method inevitably suffers from all the intricate requirements that detain its application to a wider range of biomolecular systems. Nonetheless, well-acknowledged successes related to the HIV-1 protease were accomplished using the information of conserved water molecules experimentally determined at the binding site.[15–20] Thus, it is reasonable to consider that the same logic might apply to other systems. Multiple methods have been developed using this philosophy to estimate conserved water positions using information available in public databases (e.g. Protein Data Bank,[21] PDB, Cambridge Structural Database,[22] CSD). In addition, multiple auxiliary databases have been developed to address the importance of ligand-protein interactions (e.g. Binding MOAD[23,24], PDBbind[25,26], and BindingDB[27,28]).

One of the experiments-based methods, *Consolv*,[29] analyses the environment of explicit water molecules at the active site to estimate their conservativity using single or multiple crystal structure(s). The crystallographic B-factor is considered as one of the environmental evaluation criteria, aside from the number of closest protein atoms around the water molecules (*i.e.* atomic density), the hydrophilicity of the hydration site, and the number of water-protein hydrogen bonds. Moreover, *Consolv* also allows conservation tests among multiple crystal structures using the k-nearest-neighbour genetic algorithm. The method was validated on 7 testing complexes, reaching a 75% prediction accuracy. However, 83% of false positive positions were replaced by polar atoms from the ligands in the original complexes, implying a potential 90% accuracy.

The evaluation of the interactions made by different water molecules found in a protein binding site can also be tackled as a typical protein-ligand scoring problem.[30] Several scoring methods were indeed developed to describe protein ligand interactions,[31] the most common being empirical (or regression-based) [32–34] and force field-based (or physics-based) scoring functions,[35–37] and the same theory can thus be applied to water.

*WaterScore* was indeed developed as a scoring method to estimate conserved water molecules using a logistic regression analysis and multiple structural properties, derived from crystal structures.[38] Three structural properties are considered in the scoring function: the temperature B-factor, the solvent contact surface area, and the number of protein atomic contacts. The method provided a moderate prediction efficiency ranging from 67.4 to 71.7%. The author later published a complete statistic study on energetic and physicochemical properties of tightly bound water molecules. The analysis was carried out on 2332 high-resolution crystal structures, but results were inconclusive,[39] probably due to the modest prediction ability of the method. In

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

another work, *WaterScore*, followed by thermodynamic integration free energy calculations, was successful in determining a tightly bound water molecule in the enzyme Abl-SH3 domain tyrosine kinase in complex with a peptide.[40]

Force-field-like scoring methods were also developed to characterise explicit water in biomolecules. For instance, *HINT*[41] uses a non-Newtonian force field based on experimentally determined LogP between octanol and water ( $\text{LogP}_{o/w}$ ) and a Rank algorithm[42] for water scoring and optimisation. The method was tested on 4 *apo* proteins with 50 active site water molecules. The outcome was then compared to those observed in the *holo* complexes; 76% water positions were accurately predicted.[41] A later study, performed on 9 *apo/holo* structures, predicted 87% of the 68 conserved water molecules.[43] Multiple follow-up works were carried out, covering protein-protein interactions,[44–46] protein-ligand interactions,[47] and the design of inhibitors targeting O-acetylserine sulfhydrylase,[48] thrombin,[49] cystalysin,[50] and human serine racemase.[51]

Though potentially accurate, these methods require extensive amount of inputs and *a priori* knowledge, including extended experimental data. Indeed, one or multiple well-resolved crystal structure(s) should be acquired, generally with a large amount of trials and errors. These requirements probably hindered the applications of some methods in early drug design studies, especially on those biomolecular systems which are difficult to crystallise. Thus, efficient prediction tools which are not very strict on input data, but still provide an acceptable accuracy, are deemed necessary for drug design.

## ***2.2. In silico methods predicting water positions***

Two pathways can be followed to virtually place water at the active sites: the water molecules can be positioned either before or after docking experiments. The former approach is more closely recreating the real solvated binding scenario. The successful docking greatly depends on the ability of the ligand to displace water molecules, or the adaptability to the favourable poses due to effective bridging interactions. On the other hand, water placement after docking is more effective in filling up the space between ligand and receptor, potentially improving the binding free energy by altering the chemical properties about the protein surface. However, this is at the cost of losing some ligand adaptability due to the pre-determined “dry” poses. This could be particularly true if the docked complexes are hold static. Moreover, these “dry” poses could also be incorrect, as demonstrated by Roberts and Mancera, who improved their pose predictions by including explicit water molecules during docking.[52] Nonetheless, the usefulness of these methods should not necessary be excluded, as they might still be worthwhile, for example when an accurate pose predictions has already been achieved, but improved free energy estimations are expected.

*Pre-docking prediction of solvation sites.* Based on the reference interaction site model (RISM) theory[53–56], the three-dimensional RISM with the closure relation by Kovalenko and Hirata (3D-RISM-KH) theory[57–59] has become more widely applied to analyse the solute-solvent interface for water configurations. The RISM theory essentially applies the first principles foundations of statistical mechanics and reduced Ornstein-Zernike integral equation theory[60] to the liquid state of matter.

This is a post print version of the article published in Curr Med Chem. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

The generalisation of the solute-solvent RISM equations,[57] in combination with the closure relation suggested by Kovalenko and Hirata,[58] enables the 3D mapping of solvent distributions around an arbitrarily shaped solute molecule. The detailed theoretical background has been well expounded and reviewed in a variety of reading materials.[58,61–64] In recent years, successful applications of the 3D-RISM-KH theory within multiple solvated biomolecular systems has also been reported.[65–70] This prevalence in applications is likely related to the solid theoretical platform it stands on and the more efficient computational algorithm involving solving solely the integral equations, revealing a comprehensive solvation structure in the statistical-mechanical ensembles. This provides a feasible and powerful alternative for studies related to very complex systems that cannot be easily handled by molecular simulations.

Using the 3D-RISM-KH integral equation theory, multiple thermodynamic properties can be calculated, one of which is the 3D distribution function. *Placevent*, developed by Sindhikara and co-workers, utilises this property to derive discrete solvent positions around the solute of interest.[71] The method was validated on both KNI-272-bound HIV-1 protease and F-type Na<sup>+</sup>-ATPase. All six of the important water molecules of HIV-1 protease were found from a blank test and also those of the ATPase.[71] The 3D-RISM-KH theory combined with *Placevent* became a standard hydration analysis in a variety of studies. Recent applications include facilitating protein NMR shift calculations in explicit solvent,[72,73] solvation of channelrhodopsin cation channels,[74] predicting bound water molecules for BCR-ABL kinase inhibitor design,[75] and providing initial solvated model for MD simulations of ATP-bound Akt1 complex.[76] Stumpe *et al.* also demonstrated that the water densities generated by 3D-RISM-KH remarkably agreed with that from MD simulations.[77] This exhibits the powerfulness of 3D-RISM-KH even when static conformations are used. However, it was noted that the method is not yet suitable for solvation free energy calculation due to intrinsic drawbacks brought about by the closure approximations and the applied force field. Thus, though with good accuracy in depicting water positions, the method is not recommended to be incorporated into free energy calculations.

The continuously increasing power of modern hardware, allows protocols based on MD or Monte Carlo (MC) simulations to become an alternative to predict the position of potential important water molecules within binding site. The inhomogeneous solvation theory (IST or IFST)[78–80] provides the theoretical background to build up solvation energy profiles of systems using molecular simulations. One method adapting this theory, *WaterMap*, uses MD simulations for prior-docking water mapping[81,82] and has become well practiced within the *Schrödinger Small-Molecule Drug Discovery Suite*. [83] The mapping procedure starts with a MD simulation of the restrained protein or complex in explicit solvent. The location and the orientation of every water molecule within the binding site are recorded into a density profile during the simulation. The water molecules are then mapped at gridded positions with a density higher than that of the bulk solvent. Enthalpy and entropy energy terms are then evaluated using a modified IST. Moreover, the *Glide XP* scoring function was modified, namely *WScore*, to incorporate more accurate solvent

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

interactions and some receptor flexibility in docking using *WaterMap* procedures.[84,85] The overall protocol and applications will be thoroughly reviewed in later section.

The applications of *WaterMap* took off since its early release, mainly in investigating structure-activity relationship,[86–88] designing potentially active molecules,[89–91] and exploring druggable sites.[92] Some examples include adenosine A<sub>2A</sub>[87,93] and other G protein coupled receptors,[94] human carbonic anhydrase (HCA),[95] lactate dehydrogenase A (LDH-A),[96] dopamine D<sub>3</sub> receptor,[97] serine proteases,[98] farnesyl pyrophosphate synthase,[99] penicillin-binding protein 3,[100] Janus kinases (JAKs),[90] and proprotein convertase subtilisin-kexin type 9.[101] Most of these studies involve evaluations of solvent thermodynamic properties at the binding site, implying a relatively robust and mature application in the field. Relative binding free energies were also computed with *WaterMap*, but only within congeneric series of ligands for which the desolvation of the binding site, rather than specific ligand interactions, was the main driver of the binding thermodynamics (*e.g.* within highly hydrophobic pockets).[81,102]

One intrinsic drawback of applying MD simulation is related to proper equilibration of buried water molecules.[103–106] The relatively short simulation generally used in *WaterMap* calculations might fail to sample positions about more buried region at the binding site. This is likely severe for ligand-bound complexes. Wang and co-workers have applied a grand canonical Monte Carlo (GCMC) step prior to MD simulation to overcome this demerit.[107] Likewise, other viable methods can also be applied if deemed necessary. Another disadvantage of *WaterMap* is that the simulated system is not allowed for relevant conformational changes during the solvation process, since protein restraints are applied during MD.[11] Hence, some receptor relaxation in solvent might be lost. Moreover, although able to correctly predict the position of water molecules in a network, it cannot calculate the binding free energy of networked waters.

The grid inhomogeneous solvation theory (GIST) was also proposed.[108–110] GIST allows the calculations of water occupancy, and enthalpic and entropic energetics on a 3D grid, instead of at discrete hydration sites. GIST is implemented in the *AmberTools cpptraj* package[110,111] and was recently combined with *AutoDock4*[112] to incorporate water thermodynamics in virtual screening.[113] The authors demonstrated improvements in docking pose predictions and in virtual screenings for the coagulation factor Xa. A non-commercial alternative to *WaterMap* is the *Solvation Thermodynamics of Ordered Water (STOW)* package.[114] Similarly to *WaterMap*, the method uses a water trajectory from MD simulations and applies the IST to determine the thermodynamic properties of isolated or clustered water molecules. *STOW* has been successfully used in studies on several biological systems including HIV1-protease, Concanavalin A in complex with carbohydrates, and cyclophilin A.[115–118]

Other than applying IST related methods, information on hydration sites can also be derived from averaged Gaussian distribution functions. *WATsite* takes up this strategy to identify important water molecules from a MD simulation implemented within a

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

*Pymol* plugin.[119] The method reads information from a MD trajectory of the protein in the *apo* or in the *holo* state. The protein binding site needs to be defined before running the simulation, and a ligand or a pseudo-ligand must be positioned in the binding pocket. A margin value is also required to define the final box size to enclose the binding site. Successively, for each MD frame, the positions of the oxygen atoms of waters within the binding site are recorded. Gaussian distribution functions are established about each oxygen centre on a 3D grid and averaged over the whole MD simulation. A quality threshold clustering algorithm is then applied to determine locations with peaked water occupancies as the hydration sites. The desolvation free energy is then calculated from solvent entropic and enthalpic surpluses to the bulk solvent for each hydration site. Analogous to *WaterMap* and related methods, a positive desolvation free energy indicates unstable water molecules; replacement of these water molecules by ligands should improve the final binding free energy.[120] This program is able to run all the steps necessary to the definition and visualization of hydration sites, starting from the structure preparation. Moreover, if combined with a pharmacophore modelling tool, *WATsite* can be employed to derive protein-based pharmacophore model. The use of such “water aware” pharmacophores led to higher computational efficiency maintaining the enrichment performance, compared to other models where hydration sites were not considered.[120]

However, *WATsite* is only compatible for *Gromacs* MD trajectory for the hydration site analysis. Alternatively, a similar tool, *WATCLUST*,[121] was developed as a *VMD*[122] plugin to allow analyses on previously loaded MD trajectories. *WATCLUST* can be called from the Extensions → Analysis menu to compute hydration sites around specific residues, together with water site-protein interaction energy and other thermodynamic parameters. Additionally, the obtained water sites can be used to generate water biased grid maps for subsequent docking studies with *AutoDock*.

As an alternative to MD-based methods, MC simulations might be beneficial for those systems whose binding sites are either occupied by a ligand or hardly accessible to solvent. *Just Add Water Molecules (JAWS or JAWM)*,[123] currently implemented in a modified version of *MCPRO* (v. 2.1),[124] is a rather recent MC method that can operate on both *apo* or *holo* proteins. The method requires a 3D grid formed by overlapping spheres centred on each atom of the ligand or of a user-selected group of atoms of the protein. Initially, putative hydration sites are guessed by randomly positioning water molecules on the grid. A MC step is then performed to generate a probability distribution of the water occupancies. An additional degree of freedom,  $\theta$ , representing the “water-likeness” of each hydration site, is introduced in the MC step.[123] The fractional water occupancies obtained in the previous step are then clustered into a defined number of possible hydration sites. In a second step, a biasing potential is applied to each water molecule identified previously, and a new MC simulation is performed for each water to establish the amount of time the molecule is interacting with the system. This procedure allows the estimation of the binding free energy for each water molecule. *JAWS* was initially validated on five biomolecular systems. The method performed well in predicting well-buried water molecules, but it required protocol adjustments for solvent exposed pockets. A consensus analysis of

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

multiple setups was also recommended. Cautions should be taken as *JAWS* is also ligand-dependent, but the method also works with *apo* proteins. The method was further validated on three protein-ligand complexes,[125] and more intensively, on inhibitor binding for p38 $\alpha$  MAP kinase,[126] followed by MC free energy perturbation calculations. *JAWS* has also been applied before an enhanced MC method to confirm the unfavourable placement of water in a cavity.[127]

Grand canonical Monte Carlo (GCMC) technique, another branch of MC methods, was initially introduced in 1974[128] and adapted to hydration analysis in 1990-2000s.[129–132] GCMC allows number of particles in the system to fluctuate during the simulation. This is accomplished by coupling the system to an ideal gas reservoir at a constant chemical potential. Recently, Ross and co-workers developed the method further to overcome its previous limitations.[133] The grand canonical integration (GCI) was incorporated to allow optimisation of the number of solvent molecules through minimising the Helmholtz free energy state. The calculated hydration free energy in bulk solvent was in excellent agreement with experimental data. Additionally, in small cavities at protein-ligand interfaces, GCI calculation of free energy was satisfactorily close to that from replica exchange thermodynamic integration. With this implementation, GCMC can be applied to compute the binding free energies of entire water networks through water titration plots.[134,135] Moreover, this approach can also be used as a water placement tool. It was noted that it is important to achieve full convergence of the number of water molecules for reliable predictions.

*JAWS*, GCMC and double-decoupling (a rigorous alchemical method based on MD and statistical thermodynamics [136]) were compared side-by-side in a later study on zanamivir in complex with N9-neuraminidase.[137] The three methods yielded consistent results in the calculation of water binding free energy, though *JAWS* and GCMC resulted 10 times faster than double-decoupling. However, this higher efficiency was still in the range of 30 hours using 16 2.6 GHz processors.[137] Moreover, issues with *JAWS* can arise if extensive overlaps of population densities occur, possibly leading to difficulties in clustering into separate hydration sites.

Water molecules can also be positioned through docking. *WaterDock* adopted this idea by using *AutoDock Vina*[138] to dock a single water molecule into the binding sites within a cubic space of 15 Å edge length.[139] Both locations and poses of the water molecule were recorded from repeated docking processes. The author reported 97% accurate predictions among 14 structures of OppA-lysine-X-lysine-tripeptide complexes, and 88% within another set of X-ray structures.[139] The prediction of displaceable water by *WaterDock* was validated using the Astex Diverse Set[140] with an accuracy of 75%.[139] A promising perspective regarding this method is the highly efficient calculation due to the application of *AutoDock Vina*; the prediction can be done within seconds. A study on CDK9 applied *WaterDock* for binding site hydration.[141] The applications among the other studies are mainly in the validation domain for conserved water molecules in crystal structures.[142–144]

*WaterDock* has been incorporated into another solvation program, *Dowser++*, to accurately predict water at protein interiors.[145] The latter is a development of the



This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

*Dowser* software,[146] performing the calculation on the whole protein space followed by further analysis. This tandem analyses provided a false-positive rate much lower than using *WaterDock* alone.[145] The authors suggested that the high false positive rate was likely intrinsically determined by the local performance of *AutoDock Vina* at the binding site; the *Vina* methodology is probably not suitable for whole-protein docking.[145] The authors of *WaterDock* also addressed the potential drawback of the program regarding the false-positive rate, recorded at 24% for one test set.[147] Thus, a new version, *WaterDock 2.0*, was introduced with improvements from this standpoint.[147] The solvation behaviour of small molecules was included as an additional criterion based on the hydrogen bond saturation limit. Information on the hydration shell was derived from preliminary MD simulations. It was shown that the method provided a 50% improvement in false positive occurrence within ligand-protein complexes.

The importance of water molecules at binding sites can also be assessed from their desolvation due to their proximity to the solute. *OpenEye's SZMAP*[148] resembles the well-known GRID method[149] to some extent. It provides the option of combining explicit water molecule with a Poisson-Boltzmann solvent continuum[150] (*i.e.* a semi-continuum solvation) and classical statistical mechanics. The method calculates thermodynamic quantities at gridded points using a water probe at multiple orientations, interacting simultaneously with the high dielectric continuum and the solute(s). The polar and apolar sites are distinguished using the free energy difference from charged to uncharged (CH<sub>2</sub>-like) water molecule, using PB calculations with adjusted partial charges on the hydrogens and oxygen. The performance of *SZMAP* was assessed on a variety of crystal complexes.[151] Decent results were obtained in predicting conserved waters and in correlation to experimental B-factor.

Another prediction tool, *SuperStar*,[152–154] empirically applies non-bonded information collected from experimental data, *i.e.* the IsoStar scatterplots based on the CSD and PDB.[154,155] Though designed for predicting protein-ligand interactions, it also provides a high success rate (88%) in mapping explicit hydration in protein-ligand complexes.[153] Multiple applications in mapping functional group interactions were reported so far.[156–160] In one example, *SuperStar* calculations integrated 3D-QSAR results providing a rationale for high affinity ligands interacting with hydrophobic residues through a polar group.[161] It was indeed observed that the relatively apolar side-chains were able to accommodate a water molecule acting as a hydrogen bond acceptor.

*Post-docking prediction of solvation sites.* Though post-docking positioning of water molecules could potentially impair the adaptability of protein-water-ligand interactions, Rossato and co-workers devised an interesting solvation algorithm, *AcquaAlta*, along with molecular docking.[162] The method places water molecules at the ligand-protein interface based on geometric information about water molecules, collected from the CSD. The water propensities also consider interaction energies of water to generic functional groups from *ab initio* calculations.[162] On 20 crystal structures used as test set, *AcquaAlta* succeeded in predicting 76% of water positions.

### **2.3. Applications in prospective virtual screening**

This is a post print version of the article published in Curr Med Chem. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

Several criteria need to be satisfied to implement the above methods in virtual screening applications. Most of the methods have shown fair to good reproducibility of experimental data, *i.e.* crystal-resolved water positions. Some provided solid and well-rationalised computational validations, such as 3D-RISM and *WaterMap*. Only the MC based methods (*JAWS* and GCMC) were validated on fewer experimental complexes, probably due to the relatively high computational costs of these procedures. Compatibility with *apo* structure could be advantageous, as the calculation can be performed just once before docking or virtual screening. However, few of the methods provided statistically significant studies on *apo* proteins, except *WaterMap*.<sup>[84]</sup>

The method also needs to be computational efficient. Calculations involving molecular simulations (*WaterMap*, *JAWS*, GCMC and related methods) are much more expensive in terms of computational time. This could impair their wider application in high-throughput virtual screening, as the meticulous calculation procedure might be more suitable for later lead optimisation. The 3D-RISM method, though it may take hours when considering the whole protein, can be restricted to the binding site to improve its efficiency. Its already well standardised applications are probably related to its conveniently developed connection to the principal MD simulation packages. *SZMAP1.2* calculations can take up to several CPU-hours, but it can be scaled-up drastically with MPI support, if the user can get access to large calculation resources.

Empirical methods seem fair starting points if high-resolution crystal structures are available. This is particularly true for both *HINT* and *WaterScore* methods, as both require pre-deposited water positions to evaluate on. However, *HINT* does not restrict the hydration sites to be experimentally determined. *WaterDock* could be a potentially quick estimator of hydration sites for virtual screening. Nonetheless, its application in a real-case setting is yet to be validated. Summaries of the above methods are shown in Tables 1 and 2.

**Table 1.** The major reviewed methods for explicit water evaluation and prediction

Method	Theory in brief	Experimental data?	Apo structure?	Availability*	Ref.
<i>Consolv</i>	Knowledge-based: the number of closest protein atoms, the hydrophilicity of the hydration site, the number of water-protein hydrogen bonds, and B-factor	B-factor	Yes	Free (SLIDE)	[29]
<i>WaterScore</i>	logistic regression analysis	B-factor, resolved water positions	Yes	N/A	[38]
<i>HINT</i>	A non-Newtonian force field and a Rank algorithm	LogPo/w, resolved water positions	Yes	N/A	[41]
<i>3D-RISM-KH + Placevent</i>	Reduced Ornstein-Zernike integral equation theory in combination with the KH closure relation, then converting solvent distribution function to population function and derive water positions of the highest likelihood	No	Yes	Free (AmberTools)	[57–59], [71]
<i>WaterMap</i>	MD simulation and modified IST	No	Yes	Purchasable license (Schrödinger)	[81,82]
<i>GIST (Amber)</i>	MD simulation and Gridded IST	No	Yes	Free (AmberTools)	[108–110]
<i>STOW</i>	MD simulation and IST	No	Yes	N/A	[114]
<i>WATsite</i>	MD simulations, followed by clustering and Gaussian distribution analysis	No	Yes	Free (PyMOL plugin)	[119]
<i>JAWS</i>	$\lambda$ -dynamics MC simulations, $\theta$ value simulated as water-likeness	No	Yes	Free (ProtoMS and MCPRO)	[123]
<i>GCMC+GCI</i>	Coupling to an ideal gas reservoir at a constant chemical potential, optimise number of solvent through free energy minimisation	No	Yes	Free (ProtoMS)	[128]
<i>SZMAP</i>	Semi-continuum solvation at gridded points; water probe interacting with both the high dielectric continuum and the solute(s)	No	Yes	Purchasable license (OpenEye)	[148]
<i>WaterDock (2.0)</i>	Repeatedly dock a single water molecule into the site of interest	No	Yes	Free (PyMOL plugin)	[139,147]
<i>SuperStar</i>	Knowledge-based: non-bonded contacts	No	Yes	Purchasable license with free trial (CCDC)	[152–154]
<i>AcquaAlta</i>	Water geometric information from CSD and <i>ab initio</i> calculations	No	No	N/A	[162]

\* The software packages that the method incorporated in are shown in parentheses.

**Table 2:** Additional validation information and outcomes of the reviewed methods for explicit water evaluation and prediction

Method	Test set	Validations	Prediction accuracy*
<i>Consolv</i>	7 non-homologous proteins (67 active site waters)	Percentage accuracy in total, conserved and displaced water + Matthews coefficient	74.6 %
<i>WaterScore</i>	4 proteins with both free and complexed form of structures (resolution < 2.5 Å)	Probability calculated from best logistic regression model, $1 \times 10^{-20}$ probability threshold	67.4 - 71.7 %
<i>HINT</i>	4 apo proteins	Score and rank water molecules in <i>apo</i> structures and predict their role in ligand binding	76%
<i>3D-RISM-KH + Placevent</i>	1 complexed structure of HIV-1 protease and 2 <i>apo</i> structures for the rotor ring of F-ATP synthase	Blind test of predicting water positions	All the important water molecules were predicted
<i>WaterMap</i>	<i>Apo</i> binding cavities of streptavidin, Cox-2, antibody DB3, and HIV protease	World energy and excessive entropy estimations, and water configuration analysis	N/A
<i>GIST (Amber)</i>	Cucurbit[7]uril and coagulation Factor Xa	Entropy and solvation energy estimations, water configurations, and comparison with hydration-site approach (HSA)	Perform as well as HSA
<i>STOW</i>	HIV-1 protease complexed with inhibitors, concanavalin A-carbohydrate complexes, CypA-CsA complex	Free energy contributions	N/A
<i>WATsite</i>	Factor Xa, HIV-1 protease, and pcDHFR	Pharmacophore-based virtual screening using actives and DUD datasets	N/A
<i>JAWS</i>	5 proteins + ligands	Identification of hydration sites, occupancy probabilities, and binding free energy	Reasonable crystal water position reproducibility
<i>GCMC+GCI</i>	5 proteins, a mixture of <i>apo</i> and <i>holo</i> structures	Reproducibility of consensus crystal water positions	100 % with some false positives
<i>SZMAP</i>	6 proteins with 34 structures in total	Water position prediction, free energy terms calculated at each position, sensitivity of free energy evaluation,	More than half of the crystal water molecules were predicted
<i>Waterdock (2.0)</i>	A set of high resolution X-ray structures, 14 structures of OppA, and the Astex Diverse set	Conserved water prediction for the first two test set, and hydration site prediction for the third one	88%, 97%, and 75 % respectively
<i>SuperStar</i>	50 high-resolution protein structures with a total of 16,474 water positions	Peaks in water propensity maps to crystal resolved positions	88 %
<i>AcquaAlta</i>	20 high-resolution X-ray structures for bridging water prediction and 12 OppA crystal structures for docking assessment	Prediction in bridging water and post-docking process	50 - 100% for binding site waters, and 58.3 - 100 % for bridging water molecules

\* Only results of water position prediction are listed.

### 3. Incorporating Explicit Solvation During Docking

Following the potentially important hydration sites been located, the successive docking procedures should be able to incorporate explicit water molecules accordingly. The treatment of water molecules during the docking process is crucial, as the initial position of waters could be far from ideal to accommodate the foreign compounds. Additionally, potential overlap could occur while the protein-ligand interaction is favoured over solvent-mediated binding. Thus, displaceability and

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

positional flexibility of these water molecules should be considered by the docking software.

Moreover, the scoring of the additional water could require updates to the scoring functions. As some of them are well-characterised for biomolecules and organic compounds, they might not be as explicit-solvent-friendly. Entropy is also important for evaluating free energy contribution of water. The entropic cost of moving a bulk water to the protein was estimated to be up to 2 kcal·mol<sup>-1</sup> at 300 K.[163] This is far from trivial comparing to a solvent free energy benefit of -3.1 to -1.9 kcal·mol<sup>-1</sup> for a single water molecule, calculated from two well studied biomolecular complexes.[99] Finally, other requirements might also be specifically adapted to the strategy taken for including water molecules in docking. Hereafter, we review the docking software that satisfy the above criteria to some extent. We hope this will provide useful information for medicinal chemists upon choice making, and insights for computational developers and chemists to design improved protocols and methods.

### ***3.1. Methods adapted to docking with explicit hydration***

*Schrödinger*[83] has developed *WScore* and *WaterMap* methods, both incorporated in its multi-stage systematic pose sampling methodology,[84,85,164] for a more accurate scoring of docked complexes. To include protein flexibility, Repasky and co-workers introduced an ensemble docking protocol[84] in addition to the growing algorithm of relevant side chain in *Glide*.[164] The growing algorithm already included partial protein flexibility at binding site. The updated protocol uses multiple conformations of receptors from crystal structures and docks all ligands into each structure.[84] However, it is principally feasible to use receptor structures from alternative sources, such as clustered conformations from MD simulations. This can further improve the ligand adaptability in the binding pocket. The best scored pose is taken across the ensemble. A penalty is also added if the best score is considerably better than those of other poses. In addition, water reorganisation upon ligand binding is also considered by sampling on a grid. *Glide XP*,[164] the scoring function that *WScore* derived from, includes special rewarding terms upon replacement of solvent by ligand at the binding site. Such mechanism works like the direct penalisation of bound water molecules. The scoring function is also improved with multiple modified/added terms and penalties, in comparison to *Glide XP*. *WaterMap* is applied to identify tightly bound water molecules in multiple biomolecular systems. The combination of *WaterMap* and *WScore* provided decent improvements in both ROC AUC and enrichment factors in virtual screenings of actives against universal decoys in various systems.[84] However, to our knowledge, no prospective studies were reported so far including the application of the full tandem of *WaterMap* followed by *WScore* evaluation. Notably, efficiency is a fatal drawback of the method as MD simulations are still an essential sampling method for the IST analysis. It was reported that a *WaterMap* calculation for one ligand-complex can take more than 30 hours on 2.2 GHz 8-core processors.[85]

Spatial flexibility of explicit water molecules can also be modelled by considering them as additional ligands or parts of the ligands. *GOLD*[36,165] implements the flexibility using the former method. The program uses an “on-and-off” strategy to account for water occupancy.[166] Free rotation of water molecules is also included

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

in the *GOLD* docking genetic algorithm, but with fixed translation. Therefore, the position of waters is user-defined, which requires sufficient knowledge about the water network in the binding site. Moreover, the scoring functions, *Goldscore* and *Chemscore*, were updated accordingly. The rigid-body entropy of water is included by an additional score penalty with an optimised constant value. Additional binding affinity terms for water molecules are added in the same functional forms to the regular protein-ligand interaction. Furthermore, interactions of each water molecule to other preserved ones are also included. The method reached a 93% successful prediction of the role of water at the binding site, either mediating ligand-to-protein interactions or displaced. Improvements in the prediction of binding mode were also observed at the validation stage.

Multiple studies adopted *GOLD* with explicit hydration into applied molecular docking projects.[167–171] More intriguingly, its application in virtual screening has also been explored. Murray and co-workers combined the virtual screening process with NMR experiments for fragment-based lead development of heat shock protein 90 (Hsp90).[172] Four hits were identified and validated with crystallographic structures. The binding modes were also compared to crystal structures of known active compounds targeting Hsp90 and have shown a good correlation in protein-ligand contacts. One fragment was further optimised into a lead compound with low nM activity in cells. Other retrospective studies were also conducted to validate the screening performance in Cytochrome P450 1A2[173] and RNA.[174]

*RosettaLigand*[175] uses a multi-stage docking protocol based on MC search with full ligand and protein flexibility enabled. The water molecules at the interface can be considered as protein-centric or ligand-centric.[176] The two methods differ in the way water is treated; the former includes water movement independently from the ligand while the latter has an initial stage where water molecules move in accordance with the ligand. Additional water translation and rotation cycles are also included. However, no solvent-specific scoring adjustments are included. The force-field-based scoring function includes a desolvation energy calculated using an implicit solvent model, and the weighting factor is well-characterised for protein-ligand complexes with no explicit solvent.[175,177] Hence, the effect of including water molecules for scoring is undefined. For both the water docking methods, the author demonstrated improvements in both pose prediction and ranking.

Flexibility of explicit water molecules can be included by linking them to ligands. Both *Molegro Virtual Docker (MVD)*[178] and *AutoDock*[179] approached explicit hydration docking from this viewpoint. In both methods, water molecules are attached to the ligand molecules through hydrogen bonding. Thus, the ligands are solvated prior to docking. In *MVD*, attached water molecules (AWM) are added to the ligand according to the hydrogen bond donor/acceptor and hybridisation properties of the heavy atoms they are going to be attached at.[180] The bond distances (2.8 Å) and angles are set at fixed values. Full flexibility is applied to the ligand and to all AWMs during the docking process, while the receptor is kept rigid. Ligand poses are characterised by positions, orientations, and torsional angles, with AWMs added to the torsion tree model with a newly defined hydrogen-bond edge. Any AWM with positive energy contribution is ignored. An additional rigid-body entropic penalty is

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

added at a constant value for each preserved water molecule after the docking run. The method provided a 50% success rate for the 12 tested complexes. This improved to 67% when the top 5 poses were considered. Paulson and colleagues also adopted the same philosophy to include explicit hydration in *Glide*.<sup>[83,181]</sup> Computational investigations suggested that at least two water molecules were required to mediate interactions between thapsigargin and sarco/endoplasmic reticulum calcium ATPase.<sup>[181]</sup> Standard docking procedures failed to reproduce the correct orientation of the ligand.

Similarly to *MVD*, the attachment manoeuvre of water molecules in *AutoDock* occurs by saturating all the hydrogen bond donors and acceptors on the ligand molecule.<sup>[182]</sup> However, a slightly larger H-bond distance (3.0 Å) and a set of empirical values for bond angles are used. The water molecules are represented by a neutrally charged spherical pseudoatom (W atom) combining hydrogen bond acceptor and donor properties. Whether a water molecule is displaced or conserved is continuously evaluated during the docking process by both the enthalpy and entropic contributions. The energy contribution of each W atom consists of weighted van der Waals and hydrogen bond enthalpies, and desolvation entropy. The authors validated the method on 221 complexes of various sizes, obtaining an increase of 11.7% in the accuracy of binding pose predictions compared to docking without water. A standardized protocol to perform docking in the presence of explicit water was also published recently.<sup>[183]</sup> Similar to previous methods, the later applications in molecular docking were mainly incurred with binding mode predictions.<sup>[184–188]</sup>

A special treatment of additional explicit solvent was also developed in *FlexX*. The software applies an incremental construction docking algorithm as its main implementation method.<sup>[32,189]</sup> The water molecules are incorporated using the particle concept as an extension to the original *FlexX* docking procedures.<sup>[190]</sup> The type and number of interactions are predefined for each particle, with restrained orientation of interactions during docking and penalised scoring in the final outcome. The initial hydration sites are approximated through a clustering analysis<sup>[191]</sup> and ghost particles are added. During docking, a ghost is converted to a particle (*i.e.* a water particle) if an interaction with ligand occurs. Conversely, overlapping particles are removed. If one particle has a vacant interaction, a penalty is added implicitly including the rigid-body entropy loss. The method was tested out on 200 complexes from the PDB. Minor improvements in docking were observed, but several correct predictions in the binding mode were still achieved, including the important water molecules in HIV-1 protease.

### ***3.2. Applications in prospective virtual screening***

A synopsis of the docking methods discussed above is reported in Table 3. Most of the methods presented satisfy the criteria to include explicit hydration in docking suggested earlier, with one exception: *RosettaLigand* does not include any updated enthalpic or entropic terms for water molecules. For a method that uses a force-field-based scoring function with a fitted desolvation from implicit solvent, its performance in binding energy evaluation remains unclear.

**Table 3.** The reviewed docking software including explicit water molecules

Method	Water incorporation strategy	Protein Flexibility	Water Flexibility	Enthalpy/ Entropy	Availability	Ref.
<i>WaterMap</i> + <i>WScore</i>	Simultaneous multi-stage docking of water molecule on a grid	Structure ensemble, side chain growing algorithm	Reorganisation on a grid	Rewarding ligand-protein interactions	Purchasable license (Schrödinger)	[81,82], [84,85]
<i>GOLD</i>	Independent from ligand and protein, additional water occupancy and rotation added to docking algorithm	Partial flexibility	On-and-off, free rotation, fixed translation	Constant score penalty for rigid-body entropy	Purchasable license with free trial (CCDC)	[166]
<i>RosettaLigand</i>	Protein-centric: water molecules move independently to ligand	Full flexibility	Full flexibility	No specific treatments	Free* (Rosetta)	[176]
	Ligand-centric: water molecules move with ligand during low resolution sampling, then independently during high resolution sampling	Full flexibility	Full flexibility			
<i>MVD</i>	Ligand pre-solvated with water molecules. Extended ligand torsional tree model for water in guided differential evolutionary algorithm of MVD	Rigid	Full flexibility	As ligand, water attached to ligand through hydrogen bonding, constant rigid-body entropic penalty	Purchasable license with free trial**	[180]
<i>AutoDock</i>	Ligand pre-solvated with water molecules. Special grid map of water-protein interaction for Lamarckian Genetic Algorithm	Rigid	Full flexibility	Van der Waals and hydrogen bond enthalpies, desolvation entropy	Free	[182]
<i>FlexX</i>	The particle concept: water “particles” updated at each phase of the docking process	Rigid	Removable, full flexibility	Geometry dependent scoring, implicit penalty of rigid-body entropy	Purchasable license	[190]

\* *RosettaLigand* is now included in *RosettaScripts*.

\*\* Qiagen Bioinformatics has stopped maintenances of *MVD* after version 6.0.1. A new CLC Drug Design Workbench was developed in replacement.



This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

Current applications of these water-incorporating methods in real life virtual screening are still barren. One successful drug discovery project using *GOLD* was reported, but in combination with NMR screening re-evaluation.[172] The majority of studies, as reviewed above, are limited to molecular docking and, occasionally, lead optimisation. This lack of confidence is probably due to the acceptable performance of the still developing conventional docking methodologies, as well as concerns regarding the potential impairment on computational efficiency. Most methods indeed increase computational expenses drastically. The MD simulation phase in *WaterMap* makes the *Glide* water-docking protocol hardly feasible for virtual screenings where large computational resources are not accessible. Same goes for *RosettaLigand* due to the inclusion of multiple stages of MC simulation and energy minimisation. Nonetheless, with the improving in computational power, the drawbacks due to efficiency impairments will become less significant.

*GOLD* could be a good option if prior knowledge of water network within the binding site has already been acquired. With the water-incorporating strategy of the method, there is likely a limit on the number of water molecules that can be considered without increasing the computational time considerably. An analysis of protein-ligand complexes in PDB suggested that an average of 4.6 water molecules are involved in the bridging process.[8] This might provide some optimism toward a wider application of the method in virtual screening. Combinations of the *GOLD* docking with the water selection or prediction methods discussed in the early sections are also worth for further investigations.

#### **4. MD-based Binding Free Energy Calculation with Explicit Solvent**

MD simulations are often exploited to refine the docking poses, since they can include explicit solvation and full protein flexibility. Moreover, the post-processing of MD trajectories can be applied to compute relative binding energies, for instance rescoring results of a previous docking run. In the process of ligand recognition and, consequently, in drug design, the flexibility of proteins plays a crucial role. When a ligand binds to a protein, the receptor can undergo a delicate structural fitting process (*i.e.* the induced-fit effect), thus resulting in a tightly bound complex. Although multiple docking methods have implemented full or partial receptor flexibility, the treatment is still rudimental compared to MD simulations.[192–197]

Being computationally more expensive than docking, MD-based binding predictions are currently possible only in medium to low throughput screening.[12] However, rigorous MD methods such as alchemical transformations are often used to explain *a posteriori* the role played by water in protein-ligand binding.

Double-decoupling[136,198] and free energy perturbation (FEP)[199,200] are the most widely used techniques to dynamically evaluate the role of specific water molecules upon ligand binding. These methods are based on MD and statistical mechanics and belong to the alchemical free energy calculations,[201] in which the absolute or relative binding energies of a molecule are computed by “switching it off” from its surrounding, or by converting it to related species, respectively. Consequently, both double-decoupling and FEP can be applied to evaluate the free energy of binding of a water molecule into a protein binding site. In the first method,

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

the electrostatic and van der Waals interactions of the water molecule with the surroundings are gradually switched off during two sets of simulations: one for transferring the water from bulk to the gas phase, and the other for transferring the water from the protein binding pocket to the gas phase.[136,198] Since this method is rather time-consuming, its application is only feasible when dealing with a very limited number of structures, where precision and detailed thermodynamic information are a priority for the calculation. This and related approaches might be helpful in retrospective studies of compounds showing peculiar structure-activity relationships.

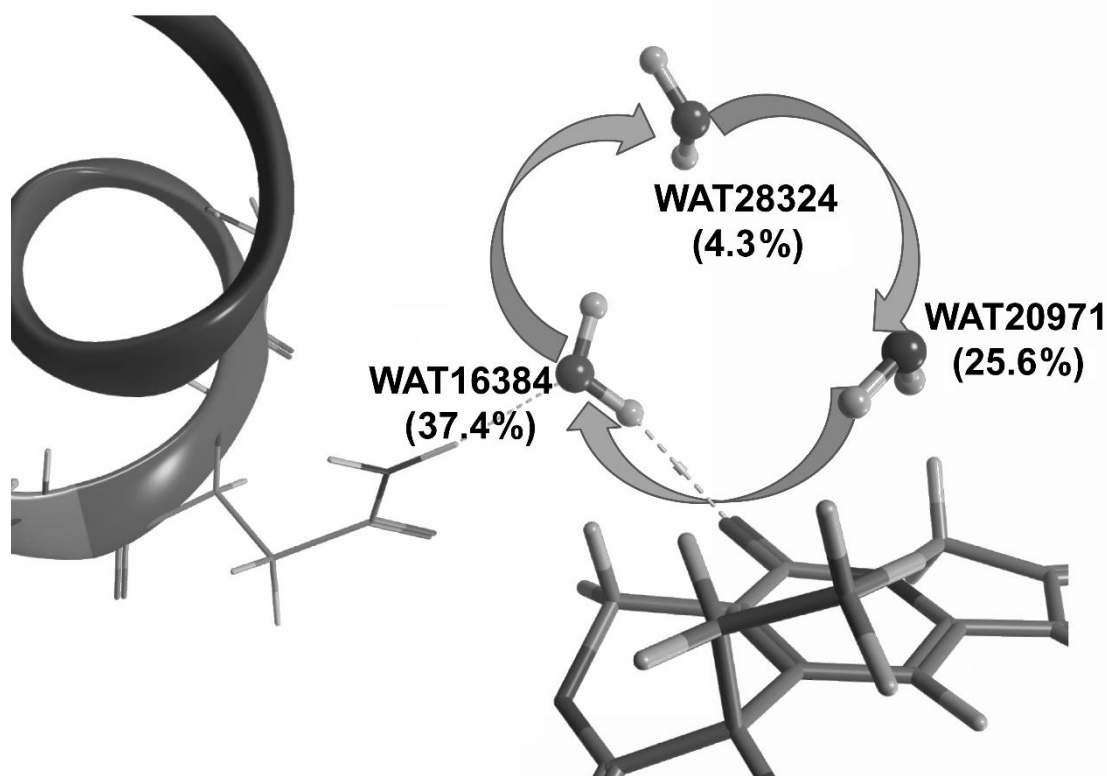
FEP methods allow the computation of the free energy of binding through finite or infinitesimal alchemical changes of the system energy function.[199] Like double-decoupling, these methods are very accurate but extremely time-consuming. However, recent studies involving the combination of FEP and replica exchange solute tempering methods (FEP+) demonstrated the applicability and the utility of this technique in the drug discovery process.[202–204] When simulations are conducted in explicit solvent, or if important water molecules are retained during the FEP step, this approach can consider the contribution of water to the binding free energy. For example, Ciordia *et al.* have recently conducted both retrospective and prospective evaluations of binding energies to optimize a novel series of amidine containing spirocyclic BACE1 inhibitors.[204] During the FEP analyses, conserved crystallographic waters were kept in the binding site and their contribution to the binding energy was included in the calculation. Alternatively, Lesenlink *et al.* used *WaterMap* to generate water molecules to be considered in FEP+ calculations to predict the binding affinities of 45 ligands to four GPCRs.[205] In another work, classical MD and FEP calculations in explicit TIP3P water were performed to identify new spleen tyrosine kinase inhibitors.[206]

Another renowned method to predict relative binding energies is the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) approach,[207,208] or its approximation using Generalized Born (MM-GBSA). The method is a good trade-off between accuracy and efficiency.[174,209–214] MM-PB/GBSA is based on the assumption that the solvent can be macroscopically described as a continuum dielectric medium.[215] Nevertheless, “bulk-like” water and “ligand-like” water could be distinguished by keeping selected explicit solvent molecules during the MM-PB/GBSA analysis.[216] This allows the contribution of water-mediated hydrogen bonds to the receptor-ligand binding energy to be considered explicitly. In some cases, this approach satisfied the expectations,[217–222] but sometimes it revealed to be useless or even detrimental.[223,224]

A possible source of uncertainty is given by the way in which water molecules to be included in the calculations are selected. When experimental information is available, the most intuitive approach is to include those molecules which are known to bridge the protein with one or more ligands.[221,223,224] In most of the cases, only conserved and stable water molecules are detected in crystallographic structures. For example, the experimental activity rank of three disaccharides binding the *Pseudomonas aeruginosa* Lectin I (PA-IL) was correctly reproduced only by including a bridging water molecule in MM-PBSA analysis.[221] However, the

selection of waters should be deliberated with caution, possibly considering other parameters such as the experimental B-factor,[225] hydrogen bonds,[219] or water density/occupancy analyses.[226] In a study on factor Xa complexed to six ligands, the inclusion of water molecules with oxygen B-factor lower than 100 resulted in 0.93 and 0.97 correlations between binding energies predicted by PB and GB models, respectively, and experiments. Conversely, the default implicit approach gave an inverse correlation.[225]

It should be noted that X-ray water positions can derive from the average electron density of several molecules competing for the same site.[217,219,227] For instance, a water molecule bridging the ligand to Asn722 was observed in the crystal structure of topoisomerase I in complex with topotecan (PDB code: 1K4T).[228] However, MD simulations conducted on the complex showed that the water bridge was shared principally by three water residues (Fig. 1).[217]



**Figure 1.** Representation of the swap of bridging water molecules observed during the MD simulation of topoisomerase I in complex with topotecan (PDB code 1K4T).

For this system, the inclusion of the crystallographic water in MM-PBSA or MM-GBSA calculations led to a failure in ranking a series of camptothecin derivatives. However, good correlations between calculations and experiments were obtained by selecting a fixed number of water molecules that were the closest to the ligand in every frame of the MD trajectory that was going to be considered in the MM-PBSA or MM-GBSA analyses ( $r^2 = 0.51$  and  $0.87$  for the two methods, respectively).

We developed this idea into the Nwat-MMGBSA (or Nwat-MMPBSA, depending on the nature of the implicit solvent model) protocol that, besides topoisomerase I, was

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

validated on systems such as penicillopepsin,  $\alpha$ -thrombin and avidin,[217] as well as on ranking different protein-protein complexes.[218] We observed an increase in the correlation between predicted binding energies and experiments of 15-50% compared to standard MM-PB/GBSA,[217] with differences principally depending on the role played by water in the ligand-receptor interaction. Aldeghi and co-workers validated the method on 57 bromodomain systems against alchemical absolute binding free energy (ABFE) calculations.[229] The authors found that ABFE calculations still were the most accurate, but MM-PBSA including explicit ligand hydration shells was a valid alternative, especially considering the much lower computational cost. In another work, Kannan and co-workers observed that MM-PBSA calculations provided better correlations with experimental  $IC_{50}$  for a set of Mnk1 and Mnk2 inhibitors only when up to 10 water molecules were included in the calculation.[230] Interestingly, this number is lower than that found in our previous studies (30-50),[217,218] possibly because of the smaller interface between Mnk1/Mnk2 kinases and the respective Type II inhibitors, compared to other systems. We recently integrated the Nwat-MMGBSA protocol into a complete workflow for virtual screening, automatically performing library setup, docking, ligand parameterization and MD simulations, prior to Nwat-MMGBSA rescoring.[231] The workflow was tested on two systems, AmpC  $\beta$ -lactamase and Rac1. Nwat-MMGBSA rescoring improved the AUC of ROC curves by 16 and 20% for the two systems, respectively, compared to standard MM-GBSA rescoring.

It should be noted that computed binding energies generally increase when explicit waters are included in MM-PB/GBSA calculations. Therefore, the out-coming energies should be interpreted as scores and not as absolute binding free energies. This is also due to the entropic term that is usually neglected in the calculation, principally due to the significant increase in computational cost required by the normal mode analysis.[225,229,232–235] Nevertheless, Aldeghi and co-workers[229] observed an improvement in correlation to experiments when entropy was considered by using a new and rather cheap approach proposed by Duan and co-workers.[236] However, the method was tested in standard MM-PBSA calculations and its combination with the inclusion of explicit waters still needs to be validated.

An attempt in using MM-GBSA to evaluate the effect of water displacement was also reported.[237] Ryde and co-workers used MM-GBSA to reproduce the binding free energies of nine phenol analogues to ferritin. All the water molecules present in the binding site were included in the calculation as part of the receptor. In addition, the effect of displaced water molecules was considered by computing the affinities of binding-site waters both before and after the ligand binding. To obtain the most accurate free energy estimations, the authors converted both the ligand and the water molecules into non-interacting ghost molecules, when unbound, to avoid the inclusion of the continuum solvent contribution to the energy. This approach then required two single-average calculations, one for the complex and one for the receptor. However, the authors evidenced the estimation of the free energy of an unbound water molecule as the principal drawback of the method. They suggested the use of the experimental value of  $-47 \text{ kJ} \cdot \text{mol}^{-1}$  to obtain correct results on their reported example, but different values might be needed for different systems.

## Conclusions

Modern drug design relies on the complete understanding of molecular recognitions occurring between a target and a ligand. This implies generating as much information as possible on both the structural and chemical-physical features of both partners. Nevertheless, this might not be enough, since most of the biological processes occur in water. The behaviour of water as a bulk solvent is generally considered by all computational methods applied to drug design. However, given the differences between bulk and binding site solvent properties and the role played by water in the binding process, considering water as the “third actor” in the ligand-receptor interaction is less trivial. This should indeed be done at different stages of the drug design process. First, detailed hydration pattern within the binding site can be obtained by combining experimental information with the results of MD simulations. Moreover, important water molecules might be considered in virtual screening, with the non-negligible toll taken in terms of simulation time.

Software applications such as *Consolv*,[29] *WaterScore*[38] and *HINT*[41] were developed to aid the selection of water molecules that might be particularly important in ligand recognition (Table 1). However, whenever experimental information on conserved water is lacking, relevant hydrations sites can also be predicted. Some predictive methods, like *Placevent*,[71] relies on the 3D-RISM equations, other on IST,[78–80] like the well-known *WaterMap*,[82] or its non-commercial alternative *STOW*. [114] IST-based methods require information on water dynamics, generated by a preliminary MD simulation, but MC is also feasible in water placement applications. For instance, *JAWS* is a rather recent MC method that works on both *apo* or *holo* proteins.[123] Modern implementations of the GCMC technique can be used either as a water placement tool, or to compute the binding free energies of single or networked water molecules.[134,135] Other methods, such as *WaterDock*[139,147] or *SZMAP*[148] can be used to place water in a putative binding site before attempting a drug design study or a virtual screening.

Relevant docking procedures with the option of incorporating explicit solvation were also discussed in this review (Table 2), together with some applications in drug design. *GOLD*[36,165] is one of the most widely applied tools with robust applications and good successful rates in hydrated docking. It models water molecules as additional ligands and uses a “on-and-off” strategy and additional score penalty to evaluate water displacement.[166] *Schrödinger*[83] incorporated its *WScore* and *WaterMap* methods in a multi-stage methodology to score docking complexes.[84,85,164] *RosettaLigand* performs a MC search and proceed in a way either water motions are more connected to the receptor or the ligand.[175] *FlexX* treats water molecules as additional particles, as an extension to its original procedure.[190] Other packages, including *AutoDock*[179] and *MVD*,[178] can tackle explicit hydration by attaching water to the ligand molecules through hydrogen bonding.[180]

Finally, this review also covered the use of MD simulations to elucidate questions related to the active presence of water at binding interfaces. Indeed, the conformational space accessed by the receptor, ligand and solvent can be explored by

This is a post print version of the article published in Curr Med Chem. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

MD simulations more extensively than by docking. Moreover, the stability of water mediated interactions can be investigated and their contribution to binding explicitly evaluated. Alchemical transformations are the method of choice whenever precise and accurate thermodynamic information are desired.[136,198–200] Conversely, when a trade-off between speed and accuracy is needed, methods derived from the popular MM-PBSA technique can also be used to improve the scoring of ligands whose binding is actively facilitated by water.[238,237] Since computationally demanding, MD methods are still confined to the evaluation of a limited number of cases, but the fast development of computer hardware will probably lead to wider applications, potentially in real life drug discovery projects.

### **List of Abbreviations**

ABFE = Absolute binding free energy

AUC = Area under the curve

AWM = attached water molecules

CSD = Cambridge structural database

FEP = Free energy perturbation

GB = Generalized Born

GIST = Grid inhomogeneous solvation theory

GCMC = Grand canonical Monte Carlo

GCI = Grand canonical integration

IST = Inhomogeneous solvation theory

MC = Monte Carlo

MD = Molecular dynamics

MM = Molecular mechanics

MM-GBSA = Molecular mechanics - Generalized Born Surface Area

MM-PBSA = Molecular Mechanics - Poisson Boltzmann Surface Area

MPI = Message passing interface

NMR = Nuclear magnetic resonance

PB = Poisson Boltzmann

PDB = Protein data bank

RISM = reference interaction site model

3D-RISM-KH = three-dimensional RISM with the closure relation by Kovalenko and Hirata

ROC = Receiver operating characteristic

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgements

Xiao Hu is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie ITN-European Joint Doctorate grant agreement No. 675527 MOGLYNET programme in Drug Discovery and Development.

### References

- [1] Bruno, A.; Costantino, G.; Sartori, L.; Radi, M. The in Silico Drug Discovery Toolbox: Applications in Lead Discovery and Optimization. *Curr. Med. Chem.* **2017**.
- [2] Vilar, S.; Sobarzo-Sanchez, E.; Santana, L.; Uriarte, E. Molecular Docking and Drug Discovery in  $\beta$ -Adrenergic Receptors. *Curr. Med. Chem.* **2017**, *24* (39), 4340–4359.
- [3] Maheswari, N.; Karthikeyan, C.; Trivedi, P.; Moorthy, N. S. Recent Advances in Protein Tyrosine Phosphatase 1B Targeted Drug Discovery for Type II Diabetes and Obesity. *Curr. Drug Targets* **2017**, *18* (999), 1–1.
- [4] Scotti, L.; T. Scotti, M. Editorial: Multi-Target in Computer-Aided Drug Design Studies. *Curr. Drug Targets* **2017**, *18* (5), 498–499.
- [5] Abdolmaleki, A.; Ghasemi, J.; Ghasemi, F. Computer Aided Drug Design for Multi-Target Drug Design: SAR /QSAR, Molecular Docking and Pharmacophore Methods. *Curr. Drug Targets* **2017**, *18* (5), 556–575.
- [6] Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. Relibase: Design and Development of a Database for Comprehensive Analysis of Protein-Ligand Interactions. *J. Mol. Biol.* **2003**, *326* (2), 607–620.
- [7] Günther, J.; Bergner, A.; Hendlich, M.; Klebe, G. Utilising Structural Knowledge in Drug Design Strategies: Applications Using Relibase. *J. Mol. Biol.* **2003**, *326* (2), 621–636.
- [8] 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* (2), 668–675.
- [9] Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53* (14), 5061–5084.
- [10] Bodnarchuk, M. S. Water, Water, Everywhere... It's Time to Stop and Think. *Drug Discov. Today* **2016**, *21* (7), 1139–1146.
- [11] Yang, Y.; Lightstone, F. C.; Wong, S. E. Approaches to Efficiently Estimate Solvation and Explicit Water Energetics in Ligand Binding: The Use of WaterMap. *Expert Opin. Drug Discov.* **2013**, *8* (3), 277–287.
- [12] de Beer, S.; Vermeulen, N.; Oostenbrink, C. The Role of Water Molecules in Computational Drug Design. *Curr. Top. Med. Chem.* **2010**, *10* (1), 55–66.

- [13] Zhang, X.; Stratton, C. F.; Li, H.; Kaur, T.; Amin, A. High Affinity Inha Inhibitors with Activity against Drug-Resistant Strains of Mycobacterium Tuberculosis. *ACS Chem. Biol.* **2006**, *1* (1), 43–53.
- [14] Howard, E. I.; Sanishvili, R.; Cachau, R. E.; Mitschler, A.; Chevrier, B.; Barth, P.; Lamour, V.; Zandt, M. Van; Sibley, E.; Bon, C.; Moras, D.; Schneider, T. R.; Joachimiak, A.; Podjarny, A. Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase – Inhibitor Complex at 0.66 Å. *Proteins* **2004**, *55*, 792–804.
- [15] Wang, Y.; Freedberg, D. I.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Bhat, T. N.; Erickson, J. W.; Torchia, D. A. Bound Water Molecules at the Interface between the HIV-1 Protease and a Potent Inhibitor, KNI-272, Determined by NMR. *J. Am. Chem. Soc.* **1996**, *118* (49), 12287–12290.
- [16] Fornabaio, M.; Spyraakis, 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* (18), 4507–4516.
- [17] Kageyama, S.; Mimoto, T.; Murakawa, Y.; Nomizu, M.; Ford, H.; Shirasaka, T.; Gulnik, S.; Erickson, J.; Takada, K.; Hayashi, H. In Vitro Anti-Human Immunodeficiency Virus (HIV) Activities of Transition State Mimetic HIV Protease Inhibitors Containing Allophenylnorstatine. *Antimicrob. Agents Chemother.* **1993**, *37* (4), 810–817.
- [18] Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. Structure of HIV-1 Protease with KNI-272, a Tight-Binding Transition-State Analog Containing Allophenylnorstatine. *Structure* **2017**, *3* (6), 581–590.
- [19] Lam, P. Y. S.; 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.; Chang, C.-H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. Rational Design of Potent, Bioavailable, Nonpeptide Cyclic Ureas as HIV Protease Inhibitors. *Science* **1994**, *263* (5145), 380–384.
- [20] Grzesiek, S.; Bax, A.; Nicholson, L. K.; Yamazaki, T.; Wingfield, P.; Stahl, S. J.; Eyermann, C. J.; Torchia, D. A.; Hodge, C. N. NMR Evidence for the Displacement of a Conserved Interior Water Molecule in HIV Protease by a Non-Peptide Cyclic Urea-Based Inhibitor. *J. Am. Chem. Soc.* **1994**, *116* (4), 1581–1582.
- [21] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- [22] Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B* **2016**, *72* (2), 171–179.
- [23] Benson, M. L.; Smith, R. D.; Khazanov, N. A.; Dimcheff, B.; Beaver, J.; Dresslar, P.; Nerothin, J.; Carlson, H. A. Binding MOAD, a High-Quality Protein-Ligand Database. *Nucleic Acids Res.* **2008**, *36* (Database issue), D674–



- [24] Ahmed, A.; Smith, R. D.; Clark, J. J.; Dunbar, J. B.; Carlson, H. A.; Carlson, H. A. Recent Improvements to Binding MOAD: A Resource for Protein-Ligand Binding Affinities and Structures. *Nucleic Acids Res.* **2015**, *43* (Database issue), D465-9.
- [25] Liu, Z.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.; Nie, W.; Liu, Y.; Wang, R. PDB-Wide Collection of Binding Data: Current Status of the PDBbind Database. *Bioinformatics* **2015**, *31* (3), 405–412.
- [26] Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein–Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47* (12), 2977–2980.
- [27] Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A Public Database for Medicinal Chemistry, Computational Chemistry and Systems Pharmacology. *Nucleic Acids Res.* **2016**, *44* (D1), D1045–D1053.
- [28] Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: A Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35* (Database), D198–D201.
- [29] 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* (4), 445–464.
- [30] Renxiao Wang; Yipin Lu, and; Wang\*, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. **2003**.
- [31] Liu, J.; Wang, R. Classification of Current Scoring Functions. *J. Chem. Inf. Model.* **2015**, *55* (3), 475–482.
- [32] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261* (3), 470–489.
- [33] Wang, R.; Lai, L.; Wang, S. Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. *J. Comput. Aided. Mol. Des.* **2002**, *16* (1), 11–26.
- [34] Richard A. Friesner, \*,†; Jay L. Banks, ‡; Robert B. Murphy, ‡; Thomas A. Halgren, ‡; Jasna J. Klicic, ‡,||; Daniel T. Mainz, ‡; Matthew P. Repasky, ‡; Eric H. Knoll, †; Mee Shelley, §; Jason K. Perry, §; David E. Shaw, #; Perry Francis, § and; Shenkin‡, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. **2004**.
- [35] Goodsell, D. S.; Morris, G. M.; Olson, A. J. Automated Docking of Flexible Ligands: Applications of Autodock. *J. Mol. Recognit.* **1996**, *9* (1), 1–5.
- [36] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking 1 Edited by F. E. Cohen. *J. Mol. Biol.* **1997**, *267* (3), 727–748.

- [37] Yin, S.; Biedermannova, L.; Vondrasek, J.; Dokholyan, N. V. MedusaScore: An Accurate Force Field-Based Scoring Function for Virtual Drug Screening. *J. Chem. Inf. Model.* **2008**, *48* (8), 1656–1662.
- [38] García-Sosa, A. T.; Mancera, R. L.; Dean, P. M. WaterScore: A Novel Method for Distinguishing between Bound and Displaceable Water Molecules in the Crystal Structure of the Binding Site of Protein-Ligand Complexes. *J. Mol. Model.* **2003**, *9* (3), 172–182.
- [39] García-Sosa, A. T. Hydration Properties of Ligands and Drugs in Protein Binding Sites: Tightly-Bound, Bridging Water Molecules and Their Effects and Consequences on Molecular Design Strategies. *J. Chem. Inf. Model.* **2013**, *53* (6), 1388–1405.
- [40] García-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. Inform.* **2010**, *29* (8–9), 589–600.
- [41] 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* (1), 289–309.
- [42] 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* (2), 69–82.
- [43] 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* (4), 1063–1067.
- [44] Ahmed, M. H.; Spyrakis, F.; Cozzini, P.; Tripathi, P. K.; Mozzarelli, A.; Scarsdale, J. N.; Safo, M. A.; Kellogg, G. E. Bound Water at Protein-Protein Interfaces: Partners, Roles and Hydrophobic Bubbles as a Conserved Motif. *PLoS One* **2011**, *6* (9), e24712.
- [45] Ahmed, M. H.; Habtemariam, M.; Safo, M. K.; Scarsdale, J. N.; Spyrakis, F.; Cozzini, P.; Mozzarelli, A.; Kellogg, G. E. Unintended Consequences? Water Molecules at Biological and Crystallographic Protein-Protein Interfaces. *Comput. Biol. Chem.* **2013**, *47*, 126–141.
- [46] Parikh, H. I.; Kellogg, G. E. Intuitive, but Not Simple: Including Explicit Water Molecules in Protein-Protein Docking Simulations Improves Model Quality. *Proteins Struct. Funct. Bioinforma.* **2014**, *82* (6), 916–932.
- [47] Bayden, A. S.; Fornabaio, M.; Scarsdale, J. N.; Kellogg, G. E. Web Application for Studying the Free Energy of Binding and Protonation States of Protein-Ligand Complexes Based on HINT. *J. Comput. Aided. Mol. Des.* **2009**, *23* (9), 621–632.
- [48] Salsi, E.; Bayden, A. S.; Spyrakis, F.; Amadasi, A.; Campanini, B.; Bettati, S.; Dodatko, T.; Cozzini, P.; Kellogg, G. E.; Cook, P. F.; Roderick, S. L.; Mozzarelli, A. Design of O-Acetylserine Sulphydrylase Inhibitors by

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Mimicking Nature. *J. Med. Chem.* **2010**, *53* (1), 345–356.
- [49] Mosier, P. D.; Krishnasamy, C.; Kellogg, G. E.; Desai, U. R. On the Specificity of Heparin/Heparan Sulfate Binding to Proteins. Anion-Binding Sites on Antithrombin and Thrombin Are Fundamentally Different. *PLoS One* **2012**, *7* (11), e48632.
- [50] Spyrakis, F.; Cellini, B.; Bruno, S.; Benedetti, P.; Carosati, E.; Cruciani, G.; Micheli, F.; Felici, A.; Cozzini, P.; Kellogg, G. E.; Voltattorni, C. B.; Mozzarelli, A. Targeting Cystalyisin, a Virulence Factor of *Treponema Denticola*-Supported Periodontitis. *ChemMedChem* **2014**, *9* (7), 1501–1511.
- [51] Dellafiora, L.; Marchetti, M.; Spyrakis, F.; Orlandi, V.; Campanini, B.; Cruciani, G.; Cozzini, P.; Mozzarelli, A. Expanding the Chemical Space of Human Serine Racemase Inhibitors. *Bioorganic Med. Chem. Lett.* **2015**, *25* (19), 4297–4303.
- [52] Roberts, B. C.; Mancera, R. L. Ligand–Protein Docking with Water Molecules. *J. Chem. Inf. Model.* **2008**, *48* (2), 397–408.
- [53] Chandler, D.; Andersen, H. C. Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids. *J. Chem. Phys.* **1972**, *57* (5), 1930–1937.
- [54] Hirata, F.; Pettitt, B. M.; Rossky, P. J. Application of an Extended RISM Equation to Dipolar and Quadrupolar Fluids. *J. Chem. Phys.* **1982**, *77* (1), 509–520.
- [55] Hirata, F.; Rossky, P. J. An Extended Rism Equation for Molecular Polar Fluids. *Chem. Phys. Lett.* **1981**, *83* (2), 329–334.
- [56] Hirata, F.; Rossky, P. J.; Pettitt, B. M. The Interionic Potential of Mean Force in a Molecular Polar Solvent from an Extended RISM Equation. *J. Chem. Phys.* **1983**, *78* (6), 4133–4144.
- [57] Kovalenko, A.; Hirata, F. Three-Dimensional Density Profiles of Water in Contact with a Solute of Arbitrary Shape: A RISM Approach. *Chem. Phys. Lett.* **1998**, *290* (1), 237–244.
- [58] Kovalenko, A.; Hirata, F. Self-Consistent Description of a Metal–water Interface by the Kohn–Sham Density Functional Theory and the Three-Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **1999**, *110* (20), 10095–10112.
- [59] Kovalenko, A. Multiscale Modeling of Solvation in Chemical and Biological Nanosystems and in Nanoporous Materials. *Pure Appl. Chem.* **2013**, *85* (1), 159–199.
- [60] Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*, 4th ed.; Academic Press, 2013.
- [61] Kovalenko, A. Three-Dimensional Rism Theory for Molecular Liquids and Solid-Liquid Interfaces. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer Academic Publishers: Dordrecht, 2003; pp 169–275.
- [62] Ram, J. *Equilibrium Theory of Molecular Fluids: Structure and Freezing*

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Transitions. *Physics Reports.* North-Holland May 30, 2014, pp 121–185.
- [63] Kovalenko. Molecular Theory of Solvation: Methodology Summary and Illustrations. *Condens. Matter Phys.* **2015**, 18 (3), 32601.
- [64] Yoshida, N. Role of Solvation in Drug Design as Revealed by the Statistical Mechanics Integral Equation Theory of Liquids. *Journal of Chemical Information and Modeling.* American Chemical Society November 27, 2017, pp 2646–2656.
- [65] Kovalenko, A.; Kobryn, A. E.; Gusarov, S.; Lyubimova, O.; Liu, X.; Blinov, N.; Yoshida, M. Molecular Theory of Solvation for Supramolecules and Soft Matter Structures: Application to Ligand Binding, Ion Channels, and Oligomeric Polyelectrolyte Gelators. *Soft Matter* **2012**, 8 (5), 1508–1520.
- [66] Yoshida, N.; Imai, T.; Phongphanphanee, S.; Kovalenko, A.; Hirata, F. Molecular Recognition in Biomolecules Studied by Statistical-Mechanical Integral-Equation Theory of Liquids. *J. Phys. Chem. B* **2009**, 113 (4), 873–886.
- [67] Phongphanphanee, S.; Yoshida, N.; Hirata, F. On the Proton Exclusion of Aquaporins: A Statistical Mechanics Study. *J. Am. Chem. Soc.* **2008**, 130 (5), 1540–1541.
- [68] Maruyama, Y.; Yoshida, N.; Hirata, F. Electrolytes in Biomolecular Systems Studied with the 3D-RISM/RISM Theory. *Interdiscip. Sci. Comput. Life Sci.* **2011**, 3 (4), 290–307.
- [69] Genheden, S.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Ryde, U. An MM/3D-RISM Approach for Ligand Binding Affinities. *J. Phys. Chem. B* **2010**, 114 (25), 8505–8516.
- [70] Huang, W.; 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-RISM-KH Molecular Theory of Solvation. *J. Chem. Inf. Model.* **2015**, 55 (2), 317–328.
- [71] Sindhikara, D. J.; Yoshida, N.; Hirata, F. Placevent: An Algorithm for Prediction of Explicit Solvent Atom Distribution–Application to HIV-1 Protease and F-ATP Synthase. *J. Comput. Chem.* **2012**, 33 (18), 1536–1543.
- [72] Zhu, T.; Zhang, J. Z. H.; He, X. Automated Fragmentation QM/MM Calculation of Amide Proton Chemical Shifts in Proteins with Explicit Solvent Model. *J. Chem. Theory Comput.* **2013**, 9 (4), 2104–2114.
- [73] He, X.; Zhu, T.; Wang, X.; Liu, J.; Zhang, J. Z. H. Fragment Quantum Mechanical Calculation of Proteins and Its Applications. *Acc. Chem. Res.* **2014**, 47 (9), 2748–2757.
- [74] Watanabe, H. C.; Welke, K.; Sindhikara, D. J.; Hegemann, P.; Elstner, M. Towards an Understanding of Channelrhodopsin Function: Simulations Lead to Novel Insights of the Channel Mechanism. *J. Mol. Biol.* **2013**, 425 (10), 1795–1814.
- [75] Huang, Y.-H.; Henriques, S. T.; Wang, C. K.; Thorstholm, L.; Daly, N. L.; Kaas, Q.; Craik, D. J. Design of Substrate-Based BCR-ABL Kinase Inhibitors

- Using the Cyclotide Scaffold. *Sci. Rep.* **2015**, 5 (1), 12974.
- [76] Lu, S.; Deng, R.; Jiang, H.; Song, H.; Li, S.; Shen, Q.; Huang, W.; Nussinov, R.; Yu, J.; Zhang, J. The Mechanism of ATP-Dependent Allosteric Protection of Akt Kinase Phosphorylation. *Structure* **2015**, 23 (9), 1725–1734.
- [77] Stumpe, M. C.; Blinov, N.; Wishart, D.; Kovalenko, A.; Pande, V. S. Calculation of Local Water Densities in Biological Systems: A Comparison of Molecular Dynamics Simulations and the 3D-RISM-KH Molecular Theory of Solvation. *J. Phys. Chem. B* **2011**, 115 (2), 319–328.
- [78] Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, 102 (18), 3531–3541.
- [79] Lazaridis, T.; Karplus, M. Orientational Correlations and Entropy in Liquid Water. *J. Chem. Phys.* **1996**, 105 (10), 4294–4316.
- [80] Lazaridis, T.; Paulaitis, M. E. Entropy of Hydrophobic Hydration: A New Statistical Mechanical Formulation. *J. Phys. Chem.* **1992**, 96 (9), 3847–3855.
- [81] Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, 130 (9), 2817–2831.
- [82] Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein–ligand Binding. *Proc. Natl. Acad. Sci.* **2007**, 104 (3), 808–813.
- [83] Schrödinger. Small-Molecule Drug Discovery Suite 2017-4. Schrödinger, LLC: New York, NY 2017.
- [84] Repasky, M. P.; Murphy, R. B.; Banks, J. L.; Greenwood, J. R.; Tubert-Brohman, I.; Bhat, S.; Friesner, R. A. Docking Performance of the Glide Program as Evaluated on the Astex and DUD Datasets: A Complete Set of Glide SP Results and Selected Results for a New Scoring Function Integrating WaterMap and Glide. *J. Comput. Aided. Mol. Des.* **2012**, 26 (6), 787–799.
- [85] Murphy, R. B.; Repasky, M. P.; Greenwood, J. R.; Tubert-Brohman, I.; Jerome, S.; Annabhimoju, R.; Boyles, N. A.; Schmitz, C. D.; Abel, R.; Farid, R.; Friesner, R. A. WScore: A Flexible and Accurate Treatment of Explicit Water Molecules in Ligand-Receptor Docking. *J. Med. Chem.* **2016**, 59 (9), 4364–4384.
- [86] Robinson, D. D.; Sherman, W.; Farid, R. Understanding Kinase Selectivity through Energetic Analysis of Binding Site Waters. *ChemMedChem* **2010**, 5 (4), 618–627.
- [87] Higgs, C.; Beuming, T.; Sherman, W. Hydration Site Thermodynamics Explain SARs for Triazolylpurines Analogues Binding to the A2A Receptor. *ACS Med. Chem. Lett.* **2010**, 1 (4), 160–164.
- [88] Fidock, D. A. Drug Discovery: Priming the Antimalarial Pipeline. *Nature*. Nature Publishing Group May 20, 2010, pp 297–298.
- [89] Cuny, G. D.; Yu, P. B.; Laha, J. K.; Xing, X.; Liu, J.-F.; Lai, C. S.; Deng, D. Y.; Sachidanandan, C.; Bloch, K. D.; Peterson, R. T. Structure–activity

- Relationship Study of Bone Morphogenetic Protein (BMP) Signaling Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18* (15), 4388–4392.
- [90] Chrencik, J. E.; Patny, A.; Leung, I. K.; Korniski, B.; Emmons, T. L.; Hall, T.; Weinberg, R. A.; Gormley, J. A.; Williams, J. M.; Day, J. E.; Hirsch, J. L.; Kiefer, J. R.; Leone, J. W.; Fischer, H. D.; Sommers, C. D.; Huang, H. C.; Jacobsen, E. J.; Tenbrink, R. E.; Tomasselli, A. G.; Benson, T. E. Structural and Thermodynamic Characterization of the TYK2 and JAK3 Kinase Domains in Complex with CP-690550 and CMP-6. *J. Mol. Biol.* **2010**, *400* (3), 413–433.
- [91] Knegtel, R. M. A.; Robinson, D. D. A Role for Hydration in Interleukin-2 Inducible T Cell Kinase (Itk) Selectivity. *Mol. Inform.* **2011**, *30* (11–12), 950–959.
- [92] 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 Struct. Funct. Bioinforma.* **2012**, *80* (3), 871–883.
- [93] Bortolato, A.; Tehan, B. G.; Bodnarchuk, M. S.; Essex, J. W.; Mason, J. S. Water Network Perturbation in Ligand Binding: Adenosine A2A Antagonists as a Case Study. *J. Chem. Inf. Model.* **2013**, *53* (7), 1700–1713.
- [94] Tehan, B. G.; Bortolato, A.; Blaney, F. E.; Weir, M. P.; Mason, J. S. Unifying Family A GPCR Theories of Activation. *Pharmacol. Ther.* **2014**, *143* (1), 51–60.
- [95] Breiten, B.; Lockett, M. R.; Sherman, W.; Fujita, S.; Al-Sayah, M.; Lange, H.; Bowers, C. M.; Heroux, A.; Krilov, G.; Whitesides, G. M. Water Networks Contribute to Enthalpy/entropy Compensation in Protein-Ligand Binding. *J. Am. Chem. Soc.* **2013**, *135* (41), 15579–15584.
- [96] Kohlmann, A.; Zech, S. G.; Li, F.; Zhou, T.; Squillace, R. M.; Commodore, L.; Greenfield, M. T.; Lu, X.; Miller, D. P.; Huang, W. S.; Qi, J.; Thomas, R. M.; Wang, Y.; Zhang, S.; Dodd, R.; Liu, S.; Xu, R.; Xu, Y.; Miret, J. J.; Rivera, V.; Clackson, T.; Shakespeare, W. C.; Zhu, X.; Dalgarno, D. C. Fragment Growing and Linking Lead to Novel Nanomolar Lactate Dehydrogenase Inhibitors. *J. Med. Chem.* **2013**, *56* (3), 1023–1040.
- [97] Newman, A. H.; Beuming, T.; Banala, A. K.; Donthamsetti, P.; Pongetti, K.; Labounty, A.; Levy, B.; Cao, J.; Michino, M.; Luedtke, R. R.; Javitch, J. A.; Shi, L. Molecular Determinants of Selectivity and Efficacy at the Dopamine D3 Receptor. *J. Med. Chem.* **2012**, *55* (15), 6689–6699.
- [98] Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. Contribution of Explicit Solvent Effects to the Binding Affinity of Small-Molecule Inhibitors in Blood Coagulation Factor Serine Proteases. *ChemMedChem* **2011**, *6* (6), 1049–1066.
- [99] Ohno, K.; Mori, K.; Orita, M.; Takeuchi, M. Computational Insights into Binding of Bisphosphates to Farnesyl Pyrophosphate Synthase. *Curr. Med. Chem.* **2011**, *18* (2), 220–233.
- [100] Han, S.; Zaniewski, R. P.; Marr, E. S.; Lacey, B. M.; Tomaras, A. P.;

- Evdokimov, A.; Miller, J. R.; Shanmugasundaram, V. Structural Basis for Effectiveness of Siderophore-Conjugated Monocarbams against Clinically Relevant Strains of *Pseudomonas Aeruginosa*. *Proc. Natl. Acad. Sci.* **2010**, *107* (51), 22002–22007.
- [101] Pearlstein, R. A.; Hu, Q. Y.; Zhou, J.; Yowe, D.; Levell, J.; Dale, B.; Kaushik, V. K.; Daniels, D.; Hanrahan, S.; Sherman, W.; Abel, R. New Hypotheses About the Structure-Function of Proprotein Convertase Subtilisin/kexin Type 9: Analysis of the epidermal growth Factor-like Repeat docking site using water map. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (12), 2571–2586.
- [102] Beuming, T.; Farid, R.; Sherman, W. High-Energy Water Sites Determine Peptide Binding Affinity and Specificity of PDZ Domains. *Protein Sci.* **2009**, *18* (8), 1609–1619.
- [103] de Azevedo, W. F. Molecular Dynamics Simulations of Protein Targets Identified in Mycobacterium Tuberculosis. *Curr. Med. Chem.* **2011**, *18* (9), 1353–1366.
- [104] de Azevedo, W. F.; Canduri, F.; Fadel, V.; Teodoro, L. G. V. L.; Hial, V.; Gomes, R. A. S. Molecular Model for the Binary Complex of Uropepsin and Pepstatin. *Biochem. Biophys. Res. Commun.* **2001**, *287* (1), 277–281.
- [105] Cardoso Perez, P.; Caceres, R. A.; Canduri, F.; de Azevedo, W. F. Molecular Modeling and Dynamics Simulation of Human Cyclin-Dependent Kinase 3 Complexed with Inhibitors. *Comput. Biol. Med.* **2009**, *39* (2), 130–140.
- [106] Canduri, F.; Perez, P. C.; Caceres, R. A.; de Azevedo, W. F. Protein Kinases as Targets for Antiparasitic Chemotherapy Drugs. *Curr. Drug Targets* **2007**, *8* (3), 389–398.
- [107] Wang, L.; Berne, B. J.; Friesner, R. A. Ligand Binding to Protein-Binding Pockets with Wet and Dry Regions. *Proc. Natl. Acad. Sci.* **2011**, *108* (4), 1326–1330.
- [108] Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor cucurbit[7]uril. *J. Chem. Phys.* **2012**, *137* (4), 973–980.
- [109] Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. *J. Chem. Theory Comput.* **2014**, *10* (7), 2769–2780.
- [110] Ramsey, S.; Nguyen, C.; Salomon-Ferrer, R.; Walker, R. C.; Gilson, M. K.; Kurtzman, T. Solvation Thermodynamic Mapping of Molecular Surfaces in Amber tools: GIST. *J. Comput. Chem.* **2016**, *37* (21), 2029–2037.
- [111] Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095.
- [112] Morris, G. M.; Ruth, H.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Software News and Updates AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791.

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- [113] Uehara, S.; Tanaka, S. AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* **2016**, *21* (11), 1604.
- [114] Li, Z.; Lazaridis, T. *Computational Drug Discovery and Design*; Baron, R., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2012; Vol. 819.
- [115] Li, Z.; Lazaridis, T. Thermodynamics of Buried Water Clusters at a Protein-Ligand Binding Interface. *J. Phys. Chem. B* **2006**, *110* (3), 1464–1475.
- [116] Li, Z.; Lazaridis, T. The Effect of Water Displacement on Binding Thermodynamics: Concanavalin A. *J. Phys. Chem. B* **2005**, *109* (1), 662–670.
- [117] Li, Z.; Lazaridis, T. Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease. *J. Am. Chem. Soc.* **2003**, *125* (22), 6636–6637.
- [118] Li, Z.; Lazaridis, T. Water at Biomolecular Binding Interfaces. *Phys. Chem. Chem. Phys.* **2007**, *9* (5), 573–581.
- [119] Hu, B.; Lill, M. A. WATsite: Hydration Site Prediction Program with PyMOL Interface. *J. Comput. Chem.* **2014**, *35* (16), 1255–1260.
- [120] Hu, B.; Lill, M. A. Protein Pharmacophore Selection Using Hydration-Site Analysis. *J. Chem. Inf. Model.* **2012**, *52* (4), 1046–1060.
- [121] López, E. D.; Arcon, J. P.; Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Dumas, V. G.; Marti, M. A.; Turjanski, A. G. WATCLUST: A Tool for Improving the Design of Drugs Based on Protein-Water Interactions. *Bioinformatics* **2015**, *31* (22), 3697–3699.
- [122] Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.
- [123] Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Prediction of the Water Content in Protein Binding Sites. *J. Phys. Chem. B* **2009**, *113* (40), 13337–13346.
- [124] Jorgensen, W. L.; Tirado-Rives, J. Molecular Modeling of Organic and Biomolecular Systems Using BOSS and MCPRO. *Journal of Computational Chemistry*. Wiley Subscription Services, Inc., A Wiley Company December 1, 2005, pp 1689–1700.
- [125] 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* (42), 15403–15411.
- [126] Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of Water Placement on Predictions of Binding Affinities for p38R MAP Kinase Inhibitors. *J. Chem. Theory Comput.* **2010**, *6* (12), 3850–3856.
- [127] Cole, D. J.; Tirado-Rives, J.; Jorgensen, W. L. Enhanced Monte Carlo Sampling through Replica Exchange with Solute Tempering. *J. Chem. Theory Comput.* **2014**, *10* (2), 565–571.
- [128] Adams, D. J. Grand Canonical Ensemble Monte Carlo for a Lennard-Jones Fluid. *Mol. Phys.* **1975**, *29* (1), 307–311.



This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- [129] Resat, H.; Mezei, M. Grand Canonical Ensemble Monte Carlo Simulation of the dCpG/proflavine Crystal Hydrate. *Biophys. J.* **1996**, *71* (3), 1179–1190.
- [130] Marrone, T. J.; Resat, H.; Hodge, C. N.; Chang, C. H.; McCammon, J. A. Solvation Studies of DMP323 and A76928 Bound to HIV Protease: Analysis of Water Sites Using Grand Canonical Monte Carlo Simulations. *Protein Sci.* **1998**, *7* (3), 573–579.
- [131] Speidel, J. A.; Banfelder, J. R.; Mezei, M. Automatic Control of Solvent Density in Grand Canonical Ensemble Monte Carlo Simulations. *J. Chem. Theory Comput.* **2006**, *2* (5), 1429–1434.
- [132] Malasics, A.; Gillespie, D.; Boda, D. Simulating Prescribed Particle Densities in the Grand Canonical Ensemble Using Iterative Algorithms. *J. Chem. Phys.* **2008**, *128* (12), 124102.
- [133] Ross, G. A.; Bodnarchuk, M. S.; Essex, J. W. Water Sites, Networks, and Free Energies with Grand Canonical Monte Carlo. *J. Am. Chem. Soc.* **2015**, *137* (47), 14930–14943.
- [134] Woods, C. J.; Essex, J. W.; King, M. A. Enhanced Configurational Sampling in Binding Free-Energy Calculations. *J Phys Chem B* **2003**, *107* (49), 13711–13718.
- [135] Woods, C. J.; Essex, J. W.; King, M. A. The Development of Replica-Exchange-Based Free-Energy Methods. *J. Phys. Chem. B* **2003**, *107* (49), 13703–13710.
- [136] Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophysical Journal*. 1997, pp 1047–1069.
- [137] Bodnarchuk, M. S.; Viner, R.; Michel, J.; Essex, J. W. Strategies to Calculate Water Binding Free Energies in Protein-Ligand Complexes. *J. Chem. Inf. Model.* **2014**, *54* (6), 1623–1633.
- [138] Trott, O.; Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461.
- [139] 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* (3), e32036.
- [140] Hartshorn, 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* (4), 726–741.
- [141] Randjelovic, J.; Eric, S.; Savic, V. In Silico Design of Small Molecule Inhibitors of CDK9/cyclin T1 Interaction. *J. Mol. Graph. Model.* **2014**, *50*, 100–112.
- [142] Zapata-Torres, G.; Fierro, A.; Barriga-González, G.; Salgado, J. C.; Celis-Barros, C. Revealing Monoamine Oxidase B Catalytic Mechanisms by Means

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

of the Quantum Chemical Cluster Approach. *J. Chem. Inf. Model.* **2015**, *55* (7), 1349–1360.

- [143] Fierro, A.; Edmondson, D. E.; Celis-Barros, C.; Rebolledo-Fuentes, M.; Zapata-Torres, G. Why P-OMe- and P-Cl- $\beta$ -Methylphenethylamines Display Distinct Activities upon MAO-B Binding. *PLoS One* **2016**, *11* (5), e0154989.
- [144] Bartl, J.; Palazzesi, F.; Parrinello, M.; Hommers, L.; Riederer, P.; Walitza, S.; Grünblatt, E. The Impact of Methylphenidate and Its Enantiomers on Dopamine Synthesis and Metabolism in Vitro. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2017**, *79*, 281–288.
- [145] Morozenko, A.; Stuchebrukhov, A. A. Dowser++, a New Method of Hydrating Protein Structures. *Proteins Struct. Funct. Bioinforma.* **2016**, *84* (10), 1347–1357.
- [146] Zhang, L.; Hermans, J. Hydrophilicity of Cavities in Proteins. *Proteins Struct. Funct. Genet.* **1996**, *24* (4), 433–438.
- [147] Sridhar, A.; Ross, G. A.; Biggin, P. C. Waterdock 2.0: Water Placement Prediction for Holo-Structures with a Pymol Plugin. *PLoS One* **2017**, *12* (2), e0172743.
- [148] OpenEye. SZMAP 1.2.1.4. OpenEye Scientific Software: Santa Fe, NM 2015.
- [149] Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28* (7), 849–857.
- [150] Grant, J. A.; Pickup, B. T.; Nicholls, A. A Smooth Permittivity Function for Poisson-Boltzmann Solvation Methods. *J. Comput. Chem.* **2001**, *22* (6), 608–640.
- [151] 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* (8), 1552–1565.
- [152] Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: A Knowledge-Based Approach for Identifying Interaction Sites in Proteins. *J. Mol. Biol.* **1999**, *289* (4), 1093–1108.
- [153] Verdonk, M. L.; Cole, J. C.; Watson, P.; Gillet, V.; Willett, P. SuperStar: Improved Knowledge-Based Interaction Fields for Protein Binding Sites. *J. Mol. Biol.* **2001**, *307* (3), 841–859.
- [154] Boer, D. R.; Kroon, J.; Cole, J. C.; Smith, B.; Verdonk, M. L. SuperStar: Comparison of CSD and PDB-Based Interaction Fields as a Basis for the Prediction of Protein-Ligand Interactions. *J. Mol. Biol.* **2001**, *312* (1), 275–287.
- [155] Bruno, I. J.; Cole, J. C.; Lommerse, J. P. M.; Rowland, R. S.; Taylor, R.; Verdonk, M. L. IsoStar: A Library of Information about Nonbonded Interactions. *J. Comput. Aided. Mol. Des.* **1997**, *11* (6), 525–537.
- [156] Kraemer, O.; Hazemann, I.; Podjarny, A. D.; Klebe, G. Virtual Screening for Inhibitors of Human Aldose Reductase. *Proteins Struct. Funct. Bioinforma.* **2004**, *55* (4), 814–823.

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- [157] Nissink, J. W. M.; Taylor, R. Combined Use of Physicochemical Data and Small-Molecule Crystallographic Contact Propensities to Predict Interactions in Protein Binding Sites. *Org. Biomol. Chem.* **2004**, *2* (22), 3238–3249.
- [158] Matter, H.; Will, D. W.; Nazaré, M.; Schreuder, H.; Laux, V.; Wehner, V. Structural Requirements for Factor Xa Inhibition by 3-Oxybenzamides with Neutral P1 Substituents: Combining X-Ray Crystallography, 3D-QSAR, and Tailored Scoring Functions. *J. Med. Chem.* **2005**, *48* (9), 3290–3312.
- [159] Serrano, D. R.; Mugheirbi, N. A.; O’Connell, P.; Leddy, N.; Healy, A. M.; Tajber, L. Impact of Substrate Properties on the Formation of Spherulitic Films: A Case Study of Salbutamol Sulfate. *Cryst. Growth Des.* **2016**, *16* (7), 3853–3858.
- [160] Saito, R.; Hoshi, M.; Kato, A.; Ishikawa, C.; Komatsu, T. Green Fluorescent Protein Chromophore Derivatives as a New Class of Aldose Reductase Inhibitors. *Eur. J. Med. Chem.* **2017**, *125*, 965–974.
- [161] Söderholm, A. A.; Lehtovuori, P. T.; Nyrönen, T. H. Three-Dimensional Structure–Activity Relationships of Nonsteroidal Ligands in Complex with Androgen Receptor Ligand-Binding Domain. *J. Med. Chem.* **2005**, *48* (4), 917–925.
- [162] Rossato, G.; Ernst, B.; Vedani, A.; Smieško, M. AcquaAlta: A Directional Approach to the Solvation of Ligand-Protein Complexes. *J. Chem. Inf. Model.* **2011**, *51* (8), 1867–1881.
- [163] Dunitz, J. D. The Entropic Cost of Bound Water in Crystals and Biomolecules. *Science* (80-. ) **1994**, *264* (5159), 670–670.
- [164] 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* (21), 6177–6196.
- [165] Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol.* **1995**, *245* (1), 43–53.
- [166] Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Modeling Water Molecules in Protein–Ligand Docking Using GOLD. *J. Med. Chem.* **2005**, *48* (20), 6504–6515.
- [167] Ntie-Kang, F.; Nwodo, J. N.; Ibezim, A.; Simoben, C. V.; Karaman, B.; Ngwa, V. F.; Sippl, W.; Adikwu, M. U.; Mbaze, L. M. a. Molecular Modeling of Potential Anticancer Agents from African Medicinal Plants. *J. Chem. Inf. Model.* **2014**, *54* (9), 2433–2450.
- [168] Pisani, L.; Barletta, M.; Soto-Otero, R.; Nicolotti, O.; Mendez-Alvarez, E.; Catto, M.; Introcaso, A.; Stefanachi, A.; Cellamare, S.; Altomare, C.; Carotti, A. Discovery, Biological Evaluation, and Structure–Activity and –Selectivity Relationships of 6′-Substituted (E)-2-(Benzofuran-3(2H)-Ylidene)-N-Methylacetamides, a Novel Class of Potent and Selective Monoamine Oxidase

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Inhibitors. *J. Med. Chem.* **2013**, *56* (6), 2651–2664.
- [169] Berg, L.; Andersson, C. D.; Artursson, E.; Hörnberg, A.; Tunemalm, A.-K.; Linusson, A.; Ekström, F. Targeting Acetylcholinesterase: Identification of Chemical Leads by High Throughput Screening, Structure Determination and Molecular Modeling. *PLoS One* **2011**, *6* (11), e26039.
- [170] Mouchlis, V. D.; Magrioti, V.; Barbayianni, E.; Cermak, N.; Oslund, R. C.; Mavromoustakos, T. M.; Gelb, M. H.; Kokotos, G. Inhibition of Secreted Phospholipases A2 by 2-Oxoamides Based on  $\alpha$ -Amino Acids: Synthesis, in Vitro Evaluation and Molecular Docking Calculations. *Bioorganic Med. Chem.* **2011**, *19* (2), 735–743.
- [171] Santos, R.; Hritz, J.; Oostenbrink, C. Role of Water in Molecular Docking Simulations of Cytochrome P450 2D6. *J. Chem. Inf. Model.* **2010**, *50* (1), 146–154.
- [172] Murray, C. W.; Carr, M. G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J. E.; Downham, R.; Figueroa, E.; Frederickson, M.; Graham, B.; McMenamin, R.; O'Brien, M. A.; Patel, S.; Phillips, T. R.; Williams, G.; Woodhead, A. J.; Woolford, A. J. A. Fragment-Based Drug Discovery Applied to Hsp90. Discovery of Two Lead Series with High Ligand Efficiency. *J. Med. Chem.* **2010**, *53* (16), 5942–5955.
- [173] Vasanthanathan, P.; Hritz, J.; Taboureau, O.; Olsen, L.; Jørgensen, F. S.; Vermeulen, N. P. E.; Oostenbrink, C. Virtual Screening and Prediction of Site of Metabolism for Cytochrome P450 1A2 Ligands. *J. Chem. Inf. Model.* **2009**, *49* (1), 43–52.
- [174] Li, Y.; Shen, J.; Sun, X.; Li, W.; Liu, G.; Tang, Y. Accuracy Assessment of Protein-Based Docking Programs against RNA Targets. *J. Chem. Inf. Model.* **2010**, *50* (6), 1134–1146.
- [175] Davis, I. W.; Baker, D. RosettaLigand Docking with Full Ligand and Receptor Flexibility. *J. Mol. Biol.* **2009**, *385* (2), 381–392.
- [176] Lemmon, G.; Meiler, J. Towards Ligand Docking Including Explicit Interface Water Molecules. *PLoS One* **2013**, *8* (6), e67536.
- [177] Meiler, J.; Baker, D. ROSETTALIGAND: Protein-Small Molecule Docking with Full Side-Chain Flexibility. *Proteins Struct. Funct. Genet.* **2006**, *65* (3), 538–548.
- [178] Thomsen, R.; Christensen, M. H. MolDock: A New Technique for High-Accuracy Molecular Docking. *J. Med. Chem.* **2006**, *49* (11), 3315–3321.
- [179] Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. Software News and Update a Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* **2007**, *28* (6), 1145–1152.
- [180] 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* (4), 909–917.
- [181] Paulsen, E. S.; Villadsen, J.; Tenori, E.; Liu, H.; Bonde, D. F.; Lie, M. A.;

- Bublitz, M.; Olesen, C.; Autzen, H. E.; Dach, I.; Sehgal, P.; Nissen, P.; Møller, J. V.; Schiøtt, B.; Christensen, S. B. Water-Mediated Interactions Influence the Binding of Thapsigargin to Sarco/endoplasmic Reticulum Calcium Adenosinetriphosphatase. *J. Med. Chem.* **2013**, *56* (9), 3609–3619.
- [182] Forli, S.; Olson, A. J. A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J. Med. Chem.* **2012**, *55* (2), 623–638.
- [183] Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* **2016**, *11* (5), 905–919.
- [184] Shityakov, S.; Salmas, R. E.; Durdagi, S.; Roewer, N.; Förster, C.; Broscheit, J. Solubility Profiles, Hydration and Desolvation of Curcumin Complexed with  $\gamma$ -Cyclodextrin and Hydroxypropyl- $\gamma$ -Cyclodextrin. *J. Mol. Struct.* **2017**, *1134*, 91–98.
- [185] Yuan, H.; Wu, J.; Wang, X.; Chen, J.; Zhong, Y.; Huang, Q.; Nan, P. Computational Identification of Amino-Acid Mutations That Further Improve the Activity of a Chalcone–Flavonone Isomerase from Glycine Max. *Front. Plant Sci.* **2017**, *8*, 248.
- [186] Patel, D.; Antwi, J.; Koneru, P. C.; Serrao, E.; Forli, S.; Kessler, J. J.; Feng, L.; Deng, N.; Levy, R. M.; Fuchs, J. R.; Olson, A. J.; Engelman, A. N.; Bauman, J. D.; Kvaratskhelia, M.; Arnold, E. A New Class of Allosteric HIV-1 Integrase Inhibitors Identified by Crystallographic Fragment Screening of the Catalytic Core Domain. *J. Biol. Chem.* **2016**, *291* (45), 23569–23577.
- [187] Taliani, S.; Pugliesi, I.; Barresi, E.; Salerno, S.; Marchand, C.; Agama, K.; Simorini, F.; La Motta, C.; Marini, A. M.; Di Leva, F. S.; Marinelli, L.; Cosconati, S.; Novellino, E.; Pommier, Y.; Di Santo, R.; Da Settimo, F. Phenylpyrazolo[1,5-A]quinazolin-5(4 H)-One: A Suitable Scaffold for the Development of Noncamptothecin Topoisomerase I (Top1) Inhibitors. *J. Med. Chem.* **2013**, *56* (18), 7458–7462.
- [188] Tamborini, L.; Pinto, A.; Smith, T. K.; Major, L. L.; Iannuzzi, M. C.; Cosconati, S.; Marinelli, L.; Novellino, E.; Lo Presti, L.; Wong, P. E.; Barrett, M. P.; DeMicheli, C.; Conti, P. Synthesis and Biological Evaluation of CTP Synthetase Inhibitors as Potential Agents for the Treatment of African Trypanosomiasis. *ChemMedChem* **2012**, *7* (9), 1623–1634.
- [189] Leach, A. R.; Kuntz, I. D. Conformational Analysis of Flexible Ligands in Macromolecular Receptor Sites. *J. Comput. Chem.* **1992**, *13* (6), 730–748.
- [190] Rarey, M.; Kramer, B.; Lengauer, T. The Particle Concept: Placing Discrete Water Molecules during Protein-Ligand Docking Predictions. *Proteins Struct. Funct. Genet.* **1999**, *34* (1), 17–28.
- [191] Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; Wiley: New York, 1973.
- [192] Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining Docking and Molecular Dynamic Simulations in Drug Design. *Med. Res. Rev.* **2006**, *26* (5), 531–568.

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- [193] Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol.* **2011**, *9* (1), 71.
- [194] Kerrigan, J. E. Molecular Dynamics Simulations in Drug Design. In *In Silico Models for Drug Discovery*; Humana Press, Totowa, NJ, 2013; Vol. 993, pp 95–113.
- [195] Zhang, J.-L.; Zheng, Q.-C.; Chu, W.-T.; Zhang, H.-X. Drug Design Benefits from Molecular Dynamics: Some Examples. *Curr. Comput. Aided. Drug Des.* **2013**, *9* (4), 532–546.
- [196] Zhao, H.; Caflisch, A. Molecular Dynamics in Drug Design. *Eur. J. Med. Chem.* **2014**, *91*, 4–14.
- [197] Borhani, D. W.; Shaw, D. E. The Future of Molecular Dynamics Simulations in Drug Discovery. *J. Comput. Aided. Mol. Des.* **2012**, *26* (1), 15–26.
- [198] 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* (24), 7683–7689.
- [199] Price, D. J.; Jorgensen, W. L. Improved Convergence of Binding Affinities with Free Energy Perturbation: Application to Nonpeptide Ligands with pp60srcSH2 Domain. *J. Comput. Aided. Mol. Des.* **2001**, *15* (8), 681–695.
- [200] Helms, V.; Wade, R. C. Computational Alchemy to Calculate Absolute Protein-Ligand Binding Free Energy. *J. Am. Chem. Soc.* **1998**, *120* (12), 2710–2713.
- [201] Williams-Noonan, B. J.; Yuriev, E.; Chalmers, D. K. Free Energy Methods in Drug Design: The Prospects of “Alchemical Perturbation” in Medicinal Chemistry. *J. Med. Chem.* **2017**, *61* (3), 638–649.
- [202] Abel, R.; Mondal, S.; Masse, C.; Greenwood, J.; Harriman, G.; Ashwell, M. A.; Bhat, S.; Wester, R.; Frye, L.; Kapeller, R.; Friesner, R. A. Accelerating Drug Discovery through Tight Integration of Expert Molecular Design and Predictive Scoring. *Curr. Opin. Struct. Biol.* **2017**, *43*, 38–44.
- [203] Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Acc. Chem. Res.* **2017**, *50* (7), 1625–1632.
- [204] Ciordia, M.; Pérez-Benito, L.; Delgado, F.; Trabanco, A. A.; Tresadern, G. Application of Free Energy Perturbation for the Design of BACE1 Inhibitors. *J. Chem. Inf. Model.* **2016**, *56* (9), 1856–1871.
- [205] Lenselink, E. B.; Louvel, J.; Forti, A. F.; Van Veldhoven, J. P. D.; De Vries, H.; Mulder-Krieger, T.; McRobb, F. M.; Negri, A.; Goose, J.; Abel, R.; Van Vlijmen, H. W. T.; Wang, L.; Harder, E.; Sherman, W.; Ijzerman, A. P.; Beuming, T. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. *ACS Omega* **2016**, *1* (2), 293–304.
- [206] Lovering, F.; Aevazelis, C.; Chang, J.; Dehnhardt, C.; Fitz, L.; Han, S.; Janz, K.; Lee, J.; Kaila, N.; McDonald, J.; Moore, W.; Moretto, A.; Papaioannou, N.; Richard, D.; Ryan, M. S.; Wan, Z. K.; Thorarensen, A. Imidazotriazines:

This is a post print version of the article published in *Curr Med Chem*. 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Spleen Tyrosine Kinase (Syk) Inhibitors Identified by Free-Energy Perturbation (FEP). *ChemMedChem* **2016**, *11* (2), 217–233.
- [207] Massova, I.; Kollman, P. A. Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspect. Drug Discov. Des.* **2000**, *18* (1), 113–135.
- [208] 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. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897.
- [209] Kuhn, B.; Kollman, P. A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models. *J. Med. Chem.* **2000**, *43* (20), 3786–3791.
- [210] Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate - DNA Helices. *J. Am. Chem. Soc.* **1998**, *120* (37), 9401–9409.
- [211] Sun, H.; Li, Y.; Tian, S.; Xu, L.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 4. Accuracies of MM/PBSA and MM/GBSA Methodologies Evaluated by Various Simulation Protocols Using PDBbind Data Set. *Phys. Chem. Chem. Phys.* **2014**, *16* (31), 16719–16729.
- [212] Sun, H.; Li, Y.; Shen, M.; Tian, S.; Xu, L.; Pan, P.; Guan, Y.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 5. Improved Docking Performance Using High Solute Dielectric Constant MM/GBSA and MM/PBSA Rescoring. *Phys. Chem. Chem. Phys.* **2014**, *16* (40), 22035–22045.
- [213] Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *J. Phys. Chem. B* **2013**, *117* (28), 8408–8421.
- [214] Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 6. Capability to Predict Protein–protein Binding Free Energies and Re-Rank Binding Poses Generated by Protein–protein Docking. *Phys. Chem. Chem. Phys.* **2016**, *18* (32), 22129–22139.
- [215] Simonson, T. Binding Free Energies and Free Energy Differences. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243–252.
- [216] Maffucci, I.; Contini, A. Tuning the Solvation Term in the MM-PBSA/GBSA Binding Affinity Predictions. In *Frontiers in Computational Chemistry*; Ul-Haq, Z., Madura, J. D., Eds.; Bentham Science Publishers, 2015; Vol. 1, pp 82–120.
- [217] Maffucci, I.; Contini, A. Explicit Ligand Hydration Shells Improve the Correlation between MM-PB/GBSA Binding Energies and Experimental

This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Activities. *J. Chem. Theory Comput.* **2013**, 9 (6), 2706–2717.
- [218] Maffucci, I.; Contini, A. Improved Computation of Protein-Protein Relative Binding Energies with the Nwat-MMGBSA Method. *J. Chem. Inf. Model.* **2016**, 56 (9), 1692–1704.
- [219] Treesuwan, W.; Hannongbua, S. Bridge Water Mediates Nevirapine Binding to Wild Type and Y181C HIV-1 Reverse Transcriptase-Evidence from Molecular Dynamics Simulations and MM-PBSA Calculations. *J. Mol. Graph. Model.* **2009**, 27 (8), 921–929.
- [220] Wong, S.; Amaro, R. E.; Andrew McCammon, J. MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein-Protein Interface. *J. Chem. Theory Comput.* **2009**, 5 (2), 422–429.
- [221] Nurisso, A.; Blanchard, B.; Audfray, A.; Rydner, L.; Oscarson, S.; Varrot, A.; Imberty, A. Role of Water Molecules in Structure and Energetics of *Pseudomonas Aeruginosa* Lectin I Interacting with Disaccharides. *J. Biol. Chem.* **2010**, 285 (26), 20316–20327.
- [222] 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 (6), 1306–1314.
- [223] Checa, A.; Ortiz, A. R.; De Pascual-Teresa, B.; Gago, F. Assessment of Solvation Effects on Calculated Binding Affinity Differences: Trypsin Inhibition by Flavonoids as a Model System for Congeneric Series. *J. Med. Chem.* **1997**, 40 (25), 4136–4145.
- [224] Greenidge, P. A.; Kramer, C.; Mozziconacci, J. C.; Wolf, R. M. MM/GBSA Binding Energy Prediction on the PDBbind Data Set: Successes, Failures, and Directions for Further Improvement. *J. Chem. Inf. Model.* **2013**, 53 (1), 201–209.
- [225] Wallnoefer, H. G.; Liedl, K. R.; Fox, T. A Challenging System: Free Energy Prediction for Factor Xa. *J. Comput. Chem.* **2011**, 32 (8), 1743–1752.
- [226] Henchman, R. H.; McCammon, J. A. Structural and Dynamic Properties of Water around Acetylcholinesterase. *Protein Sci.* **2009**, 11 (9), 2080–2090.
- [227] Schiffer, C.; Hermans, J. Promise of Advances in Simulation Methods for Protein Crystallography: Implicit Solvent Models, Time-Averaging Refinement, and Quantum Mechanical Modeling. *Methods Enzymol.* **2003**, 374, 412–461.
- [228] Staker, B. L.; Hjerrild, K.; Feese, M. D.; Behnke, C. A.; Burgin, A. B.; Stewart, L. The Mechanism of Topoisomerase I Poisoning by a Camptothecin Analog. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, 99 (24), 15387–15392.
- [229] Aldeghi, M.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Statistical Analysis on the Performance of Molecular Mechanics Poisson-Boltzmann Surface Area versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *J. Chem. Inf. Model.* **2017**, 57 (9), 2203–2221.
- [230] Kannan, S.; Pradhan, M. R.; Cherian, J.; Joseph, T. L.; Poh, Z. Y.; Hai Yan, Y.;



This is a post print version of the article published in *Curr Med Chem.* 2018 May 13. doi: 10.2174/0929867325666180514110824

<http://www.eurekaselect.com/162094/article>

- Melvyn, H.; Boping, L.; Jeffrey, H.; Nacro, K.; Verma, C. S. Small Molecules Targeting the Inactive Form of the Mnk1/2 Kinases. *ACS Omega* **2017**, 2 (11), 7881–7891.
- [231] Maffucci, I.; Hu, X.; Fumagalli, V.; Contini, A. An Efficient Implementation of the Nwat-MMGBSA Method to Rescore Docking Results in Medium-Throughput Virtual Screenings. *Front. Chem.* **2018**, doi: 10.3389/fchem.2018.00043.
- [232] Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the Molecular mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* **2011**, 32 (5), 866–877.
- [233] Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, 51 (1), 69–82.
- [234] Weis, A.; Katebzadeh, K.; Söderhjelm, P.; Nilsson, I.; Ryde, U. Ligand Affinities Predicted with the MM/PBSA Method: Dependence on the Simulation Method and the Force Field. *J. Med. Chem.* **2006**, 49 (22), 6596–6606.
- [235] Case, D. A. Normal Mode Analysis of Protein Dynamics. *Curr. Opin. Struct. Biol.* **1994**, 4 (2), 285–290.
- [236] Duan, L.; Liu, X.; Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein-Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, 138 (17), 5722–5728.
- [237] Mikulskis, P.; Genheden, S.; Ryde, U. Effect of Explicit Water Molecules on Ligand-Binding Affinities Calculated with the MM/GBSA Approach. *J. Mol. Model.* **2014**, 20 (6), 2273.
- [238] Maffucci, I.; Hu, X.; Fumagalli, V.; Contini, A. An Efficient Implementation of the Nwat-MMGBSA Method to Rescore Docking Results in Medium-Throughput Virtual Screenings <https://doi.org/10.26434/chemrxiv.5387575.v1>.