International Journal of
Molecular Sciences

MDPI

Review

The Advances and Limitations of the Determination and Applications of Water Structure in Molecular Engineering

Balázs Zoltán Zsidó, Bayartsetseg Bayarsaikhan, Rita Börzsei, Viktor Szél, Violetta Mohos and Csaba Hetényi *

Department of Pharmacology and Pharmacotherapy, Medical School, University of Pécs, Szigeti út 12, 7624 Pécs, Hungary; zsido.balazs@pte.hu (B.Z.Z.); bayartsetseg704@yahoo.com (B.B.); rita.borzsei@aok.pte.hu (R.B.); szel.viktor@pte.hu (V.S.); mohos.violetta@aok.pte.hu (V.M.)
* Correspondence: hetenyi.csaba@pte.hu

Abstract: Water is a key actor of various processes of nature and, therefore, molecular engineering has to take the structural and energetic consequences of hydration into account. While the present review focuses on the target–ligand interactions in drug design, with a focus on biomolecules, these methods and applications can be easily adapted to other fields of the molecular engineering of molecular complexes, including solid hydrates. The review starts with the problems and solutions of the determination of water structures. The experimental approaches and theoretical calculations are summarized, including conceptual classifications. The implementations and applications of water models are featured for the calculation of the binding thermodynamics and computational ligand docking. It is concluded that theoretical approaches not only reproduce or complete experimental water structures, but also provide key information on the contribution of individual water molecules and are indispensable tools in molecular engineering.

Keywords: drug design; docking; crystallography; electron microscopy; solvation; free energy

Citation: Zsidó, B.Z.; Bayarsaikhan, B.; Börzsei, R.; Szél, V.; Mohos, V.; Hetényi, C. The Advances and Limitations of the Determination and Applications of Water Structure in Molecular Engineering. *Int. J. Mol. Sci.* **2023**, *24*, 11784. https://doi.org/10.3390/ijms241411784

Academic Editor: Mario Milani

Received: 20 June 2023
Revised: 18 July 2023
Accepted: 20 July 2023
Published: 22 July 2023

1. Introduction

Water is a molecular jolly joker of a living nature. It is a main solvent in bulk solution and cellular interfaces and fills the void spaces in tissues (the mass of the human body is made up of ca. 60% water [1]). Water also acts as an active matrix component and is involved in the stabilization of the biomacromolecules mediating macromolecular interactions, e.g., in signaling pathways [2–5], and in the binding of small molecules to their target structures [6–12]. From a structural point of view, the role of water can be further classified. There are water molecules that form the bulk solvent accounting for 85% of the water content of a cell [13,14], and they might either be exchanged with bound waters or participate in the (de)stabilization of solute complexes. Buried water molecules stabilize solutes internally, giving 10% of the 'dry mass' of proteins. Water molecules also bridge between solute (macro)molecules and fill the void volumes of interaction interfaces [9,13,15]. They can form a hydration shell [16] that is either conserved or displaced upon ligand binding. If hydration shell water molecules are conserved upon ligand binding, they turn into bridges forming a static network of solute–water and water–water hydrogen bonds [13,15,17–24]. Such a static network is characterized by a low mobility and acts by stabilizing complexes. On the other hand, dynamic networks characterized by loosely bound water molecules with a high mobility participate in the complex destabilization or non-selective binding of various ligands.

In molecular engineering, the above structural roles of water can be translated into energetic contributions. For example, in a target–ligand interface (the main stage of drug design), conserved and leaving water molecules can be distinguished upon ligand binding [16,19,25–29]. Conserved waters tend to stay and form bridges in the target–ligand interface and are often referred to as 'happy' waters (Figure 1).

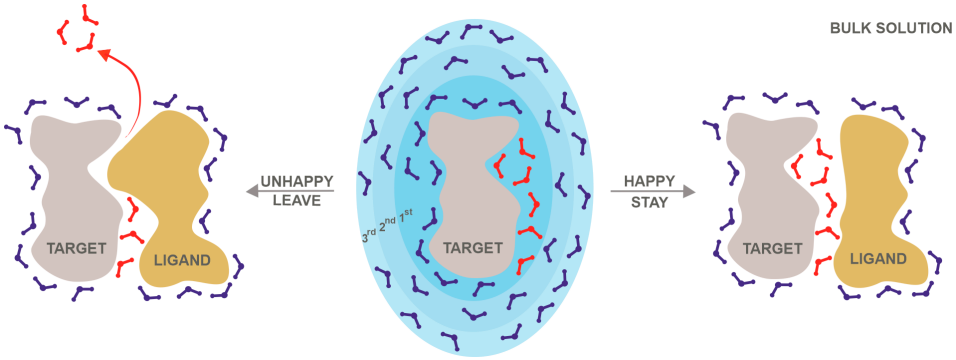

Figure 1. The hydration shells and the possible roles of interface water molecules during ligand binding. The target molecule (grey cloud, in the middle) is covered by hydration shells of surface water molecules (blue sticks), where the fading color of the shells represent the diminishing strength of interaction between the shell (also labelled by a serial number) and the target. Happy interface water molecules (red sticks, on the right) tend to stay, while unhappy water molecules (on the left) are displaced by the ligand (beige cloud) and leave (red arrow) to the bulk solution during the binding process.

There are also 'unhappy' water molecules displaced by the drug molecule during its binding to the target. An unhappy water molecule might offer a possibility for the enthalpic optimization of ligand binding. An indirect, 'unhappy', water-mediated interaction between a ligand and a target might be enthalpically less favorable compared to the direct binding of a ligand to a target amino acid residue after the displacement of an 'unhappy' water molecule [30]. During the displacement of an 'unhappy' water molecule, it moves to the bulk, and this process has a favorable entropic contribution to the free energy change of the binding reaction (ΔG_b) [28]. On the other hand, targeting 'happy' water molecules can be useful, as they often bridge between the target and ligand. A ligand can be optimized to participate in this bridging interaction by adding functional groups with a hydrogen bonding capacity to the drug molecule, providing a favorable enthalpic contribution to the ligand binding, because the additional hydrogen bonding capacity can form hydrogen bonds with 'happy' water molecules, as well as with hydrophilic target amino acids. In drug design, increasing the ligand interactions with happy water molecules and pushing unhappy water molecules away into the bulk (Figure 1) can therefore increase the ligand binding specificity [31–33] and affinity [33–35] to the target. Therefore, the importance of considering water molecules in the drug design process has been long recognized [10,36]. Besides ligand optimization, water molecules have been also utilized to improve docking results (See Section 5 for details [35,37,38]).

Despite the above importance of water in drug design, the determination of the structure and energy contribution of the water networks in intra- and intermolecular interactions is challenging for both experimental and theoretical approaches [13,39,40]. Water molecules are often too mobile, as they can change their positions in the meantime of picoseconds [6] and get lost in the large electron density maps of proteins [39]. On the other hand, theoretical approaches often have a considerable computational cost, calculating all the possible interactions with water molecules in a large simulation box. The present review gives a brief account on the above limitations and advances of the recent experimental and theoretical methodologies for water structure and their applications in the context of molecular engineering focused on biomolecules and drug design.

2. Experimental Determination of Water Structure

The experimental methods of X-ray/neutron crystallography [41], cryo-electron microscopy (Cryo-EM) [42–44], and Nuclear Magnetic Resonance (NMR) spectroscopy [45,46] can be considered as the primary techniques for the determination of molecular structures at the atomic level. While these methods provide a solid background for establishing the

structure–activity relationships [47] of biomolecules and their complexes, they face several challenges in the determination of hydration structures. Most of these difficulties come from the complexity (hydration layers interconnected with hydrogen-bonding networks) and high mobility (dynamic exchange of water molecules between the layers) of hydration structures. Plausibly, water molecules and other non-amino acid moieties, such as ligands, ions, or metals, are not included in the amino acid sequence that is otherwise key information for protein structure determination, indicating the order and covalent links between amino acids.

More than half of the experimental structures published in the Protein Data Bank (PDB, [48,49]) contain at least one water molecule (Figure 2a). Crystallography is the most powerful technique for the exploration of the networks of several water molecules. Cryo-EM, NMR, and other methods can assign far fewer (often individual) water positions (Figure 2a).

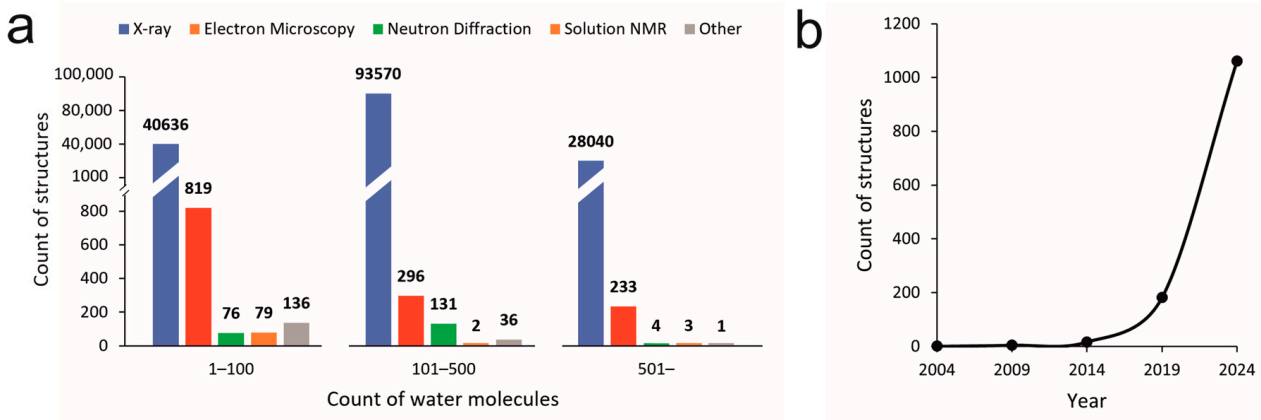

Figure 2. (**a**) Counts of structures containing water molecules resolved by different methods and deposited in the Protein Data Bank. Data were collected by the advanced search module of the PDB database: Entry features > Number of water molecules per deposited model. The number of structures was automatically separated by methods used for resolving. (**b**) The number of water-containing cryo-EM structures resolved in 5-year-long periods of time. The number on the x axis indicates the last year of the period.

However, crystallography provides only a static picture of the structure of solute molecules and their first surrounding water shell [50]. Moreover, the assignation of water positions in electron density maps gained via Fourier Transform from the crystal diffraction pattern is often complicated, even in the first shell. One of the methods is based on the low B factors, by which the waters bound to the protein surface or another water molecule located in the first hydration shell can be confidently identified [51]. In buried regions, such as binding pockets or active sites, even the third hydration layer can be resolved [50,52]. Problematic, partially ordered waters located mainly in the second hydration shell [53] can be assigned using D_2O-H_2O "neutron difference maps" [54]. This method uses the large difference in neutron scattering by deuterated and light waters, resulting in peaks of only water locations, while the scattering of the solute remains the same [51,53,54]. While neutron diffraction is capable of detecting not only oxygen but also hydrogen/deuterium atoms, and has been continuously developed [55], it is still less widespread due to the technical complexity of the method [56] and the limited accessibility of the neutron sources based on only four nuclear reactors worldwide [57], also reflected by the small number of structures [49] resolved by this method (Figure 2a).

There are various computational methods that help with the assignation of water positions in electron density maps, sometimes equipped with quantum- and/or molecular mechanics refinements [58]. For example, PHENIX [59–61] is a frequently used system for macromolecular crystallographic structure solutions, in which a bulk-solvent determination protocol is based on both maximum-likelihood and least-squares target functions [62]. Coot

performs a cluster analysis on a residual map to find water places [63]. The assigned waters are then checked based on their distance from the hydrogen-bond donors and acceptors, temperature factor, or electron-density level [63]. ARP/wARP [64] includes a fully automated placement for finding ordered water molecules using least-square refinement, in combination with the Fo–Fc difference of the electron density maps [65] (where Fo and F_C are the observed and calculated structure factor amplitudes), and geometric assumptions such as interatomic distances, angles, and van der Waals radii, etc. [66]. Despite the development of new assignation tools, the determination of correct water positions remains problematic, especially if the water structure is disordered surrounding non-polar atoms [67] or has fewer tetrahedral hydrogen bonds [68] at partially occupied solvent sites of low density.

The number of water molecules determined by crystallography mainly depends on the size and form of the system [52], as well as its resolution. By increasing the size of the system, the number of water molecules [41] also increases and the solvent becomes considerably disordered. At the size of the proteins (MW > 30.000), the resolution is usually between 1.5 and 2.5 Å, having a high background noise level in Fourier maps due to the high incoherent scattering cross-section of the numerous hydrogen atoms [41], which makes the solute assignation more difficult [41,69]. Furthermore, in the case of large biomolecules, the number of hydration layers increases, resulting in a weaker and more diffuse solvent density [41]. The other problem is that the electron density of water molecules is similar to that of small iso-electronic ions (e.g., sodium and ammonium), leading to inaccurate assignation. Moreover, the experimental conditions also affect the successful, accurate, and valid water assignation. The limitations of crystallography include the difficulty of the crystallization of biomolecules, especially in the case of large, non-globular, or disordered systems [50]. Furthermore, it is also questionable how the crystallization procedure, such as packing and cryogenic temperature, modifies the native structure of the biomolecule [50] and its hydration shells. It has been proven that the hydration structure of a biomolecule highly depends on temperature [70].

In the case of cryo-EM, high-resolution structural information is gained from thousands of images produced by transmitting an electron beam through the protein sample embedded into a special vitreous environment instead of crystals. Thus, the biomolecules can be studied in a more "native" environment, with different conformational and/or functional states, and this allows for the resolution of structures in a higher molecular weight range than that of X-ray crystallography [71]. Due to the progressive improvement in technological and refining processes, the resolution of cryo-EM maps has been entered into the atomic dimension, where the resolvability of individual atoms, including solvent water atoms, is accessible [72]. The first cryo-EM structure with water molecules was published in 2003 (PDB code: 1uon) [73] at a resolution of 7.6 Å, which is too low for the identification of individual atoms. Due to the 'Resolution Revolution' [74,75], which started in 2013, less than a decade ago, when the first near-atomic resolution cryo-EM structures were published [76–78], the number of water-containing cryo-EM structures has exponentially increased (Figure 2b). This tendency might forecast that cryo-EM structures will catch up to the number of X-ray structures in the next decades, especially in the case of large protein complexes, cellular machines, and viruses [79]. The above computational methods used for the assignation of waters in X-ray crystallography could also be applied to cryo-EM maps. The development of new assignation tools has emerged in this field as well. The assignation of individual atomic positions in cryo-EM can be performed using methods such as SWIM (segmentation-guided water and ion modelling) [80] and UnDowser in MolProbity [81,82].

Unlike crystallography and cryo-EM, NMR spectroscopy is suitable for examining small proteins or oligopeptides in solutions adopting various conformations [50]. Water–protein interactions can be identified by using the nuclear Overhauser effect and/or rotating-frame Overhauser effect between the water protons and protein atom nuclei [83]. Here, individual water molecules can be determined that are located in the first hydration shell and bound to the protein instead of a complex 3D hydration structure [84]. It is notable that

this method is limited by the short-term period of protein–water interactions, the hydrogen exchange with unstable protein moieties, and long-range dipole coupling, and identifies only 1–100 water molecules at best (Figure 2a). Additionally, while crystallography and cryo-EM provide direct information on the positions of water oxygen atoms, solution NMR is based on different principles.

3. Calculation of Water Structure

While there is an impressive, continuous development of experimental structure determination methods, the previous Section also highlighted the limitations of their assignment of the positions of water molecules [9,13,40]. To fill the gap of missing experimental hydration structures, extensive theoretical research has been conducted and resulted in various methods for the calculation of water positions (Table 1).

Table 1. The categorization and performance of theoretical methods of prediction of hydration structure.

Method	Concept	Type [a]	#System/#Water [b]	Match Tolerance (Å)	SR (%)
3D-RISM [c] [85–87]	Knowledge	IF	18/113 [d]	2.5	91
		IF	13/113 [e]	1.5	65
		SF	8/101 [e]	1.5	60
AcquaAlta [c] [88]	Geometry	IF	20/77	1.4	76
Auto-SOL [c] [89]	Geometry	SF	5/1337	1.5	64
AQUARIUS [f] [90]	Knowledge	SF	7/1376	1.4	59
Fold-X [c] [91]	Energy	SF	74/2687	1.0	76
Forli et al., 2012 [c] [92]	Geometry [g]	IF	27/51	2.0	96
HADDOCK [c] [93]	Geometry [g]	IF	27/50	2.0	90
Huggins and Tidor, 2011 [94]	Geometry	IF	5/19	2.0	68
HydraMap [c] [95]	Dynamic	IF	13/113 [e]	1.5	72
		SF	8/101 [e]	1.5	69
HyPred [f] [96]	Dynamic	SF	3/233	1.0	12
MobyWat [c] [13,40]	Dynamic	SF	20/1500	1.5	80
		IF	31/344	1.5	90
Particle concept [h] [97]	Geometry	IF	200/232	1.5	35
Splash'Em [c] [98]	Knowledge	IF	91/230	1.0	62
SZMAP [h] [99]	Knowledge	IF	18/113 [d]	2.5	96
WaterDock [c] [100]	Energy	SF	7/92	2.0	88
WaterFLAP [h] [87,101,102]	Knowledge	IF	18/113 [d]	2.5	98
WaterMap [h] [37,87]	Dynamic	SF	1/11	1.5	82
		IF	18/113 [d]	2.5	96
WarPP [c] [103]	Geometry	IF	1500/20,000	1.0	80
WATGEN [f] [104]	Geometry	IF	126/1264	2.0	88
WATsite [c] [95,105]	Dynamic	IF	13/113 [e]	1.5	75
		SF	8/101 [e]	1.5	77

[a] Water molecules in the target–ligand interface (IF), and on unbound target surface (SF) are considered, respectively. [b] The count of systems/the count of experimental water oxygen positions used in the cited study for method validation. [c] Freeware or free trial for academic use. [d] These data are taken from the comparative paper [87]. [e] These data are taken from the paper [95]. [f] Website no longer available. [g] With docking search. [h] Commercially available.

Despite the flaws of these experimental methods, the validation of theoretical methods still relies on the experimental water oxygen positions. The positions of the predicted water oxygen and experimental water oxygen are compared, and if the distance is within a tolerance threshold, then it is accepted as a successfully predicted water oxygen position. The ratio of the count of the successfully predicted water oxygen positions and all the available experimental water oxygen positions can be considered as a success rate (SR, this number is expressed in percentage after multiplication with 100 in Table 1). The validation and comparison of different theoretical methods can be easily performed based on SR values (Table 1).

Out of the four types (bulk, buried, interface, and surface) of water molecules mentioned in the Introduction (Figure 1), mostly surface and interface water molecules are investigated by theoretical methods (Table 1). Hydrated targets have surface water molecules in their first hydration shell [40] that have a stabilizing function on the macromolecular structure. Target–ligand complexes also have interface water molecules bridging between the target and ligand [13,15,19,25,106]. The prediction of interface water molecules can be accomplished very precisely with SR values even above 90% (Table 1) [9], as these molecules are captured between the target and the ligand, and there is enough space to fit only the water molecules participating in the interaction. Surface water molecules tend to have a higher mobility (B-factors) and can be predicted with SRs of ca. 80% (Table 1) [24,29,40]. That is, the natural uncertainty of surface water positions tends to result in a lower success of their prediction [39,40].

Either static or dynamic methods are used for the prediction of interface or surface water molecules. Static methods assume a static hydration shell and predict the binding sites of the water molecules on the surface of a dry experimental solute structure [40]. Finding a binding site can rely on energy calculations, scoring, prior knowledge, and information on H-bonds, and neural networks have also been applied [107]. Knowledge-based methods rely on the information found mainly in X-ray crystallographic structures (see previous Section). The main limitation of knowledge-based approaches is the assembly of an appropriate test set. The quality of X-ray crystallographic water oxygen positions varies greatly (see previous Section) and the methods perform better on similar structures that are involved in their test sets. Some methods assign a score to the experimental water molecules. Energy calculations may also apply popular docking tools to predict water molecule positions. Energy- or grid-based methods try to locate the energetically favorable positions of water molecules using probes that mimic them. Static methods can accurately identify the water molecules at the interfaces of the proteins and ligands, as these waters are usually static; however, a dynamic exchange of water molecules between the bulk solvent and the protein surface is disregarded by these methods. Generally, static methods do not consider an explicit water model and provide fast results. However, the quickness of these methods often involves a compromise in their precision.

Dynamic methods rely on extensive molecular dynamics (MD) simulations or other global search techniques using an explicit water model and allowing for the mobility of individual water molecules. All atomic movements are recorded into trajectories and the protein–water, ligand–water, and water–water interactions can be followed. This includes a dynamic exchange of water molecules with the bulk solvent and the displacement of water molecules from the binding site by ligands. However, the analysis of each trajectory in a large-scale study using various systems would be time-consuming. To tackle this, the distribution density averages of the water molecules or their occupancy at binding sites might be investigated. Dynamic approaches offer information not only on the location of water molecules, but the displacement of water molecules can be also studied. MD-based thermodynamic analyses or a comparison of the hydration structures of the apo and holo targets can follow the application of these dynamic approaches.

The counts of the systems and water molecules involved in the validation differ in different methods (Table 1). In future studies, the involvement of at least 1000 and 100 experimental (reference) water positions can be recommended for surface and interface

predictions, respectively. Preferably, at least 10 different (protein or complex) systems should be used to have a diverse set of water positions. Notably, the SR depends on the choice of match tolerance, where the highest value is 2.5 Å, but more commonly 1.4–2.0 Å is used, which seems to be the consensus for method validation (Table 1). Naturally, when the match tolerance is set to a higher value (2.0–2.5 Å), the methods achieve better SRs. Notably, the calculated water positions and SR values correspond to a certain biomacromolecular structure (or PDB ID), and the use of high-resolution structures can be recommended for the calculation of the SR. While SRs provide a fair comparison of methods, the number of experimental water oxygen positions used for the method validation and testing is similarly important. Notably, MobyWat, WATGEN, and WarPP use more than 300 experimental water positions for the validation of interface hydration. Auto-SOL, AQUARIUS, Fold-X, and MobyWat use more than 1300 waters to test their surface predictions.

The theoretical approaches of the above sections complement well the experimental methods for the atomic-level determination of water structures. In some cases, these methods also offer a complete hydration structure [13,40] of the protein surfaces and interfaces, which is often not achieved using experimental methods due to assignation problems (Section 2). Knowledge of the complete water structure is especially important for the calculation of (single molecular) the energy contribution of the (de)solvation process of drug–target binding (next Sections).

4. Water in the Structure-Based Calculation of Binding Thermodynamics

Water influences the thermodynamics of various biochemical interactions [108,109] important in molecular engineering. For example, ligand binding is described by binding free energy (ΔG_b), a net measure of the strength of target–ligand interactions. During the formation of target–ligand complexes, hydration shells undergo considerable changes (Figure 1) and, therefore, the mediation of the interactions between the drug and target partners is fairly dependent on the water molecules. There are implicit and explicit water models for the calculation of the energetics of the (de)solvation during ligand binding. Implicit solvation models consider the solvent as a continuous medium around solutes and manifest in the formulae, e.g., in electrostatic terms [110]. Explicit water models place numerous water molecules in the simulation box and set various molecular properties for the water prototype used in copies [111,112]. Both types of models have been implemented at the molecular mechanics (MM) and quantum mechanics (QM) levels of calculations.

At the molecular mechanics level, implicit water models such as MM-PB(GB)/SA [113] are widely used and based on the solution of the theoretically accurate, but computationally expensive Poisson–Boltzmann (PB) equation, or a simplified but scalable Generalized Born (GB) equation, to obtain the polar contribution of the solvation free energy change on an ensemble of MD snapshots. The solute cavity formation within the solvent and the van der Waals interactions between the solute and the solvent are represented by a nonpolar term often based on solvent-accessible surface areas (SA) [114]. Docking programs have also implemented implicit water models in their scoring functions due to their simple formulation and low computational costs. Notably, the scoring functions of docking methods require the fastest possible approaches to maintain their high-throughput nature. For example, the popular docking program AutoDock [115] applies the method of Stouten et al. [116], which calculates the solvation free energy as a sum of the atomic contributions with a linear relationship between the percentage of free volume around the atom and its contribution. At the same time, a PB-based distance-dependent dielectric function was also implemented in the Coulomb potential of AutoDock, which dampens the water permittivity value and corrects the screening effects near the solute surfaces [117]. In this way, a continuous transition of the relative permittivity of the medium is considered as we go from the bulk water to the protein surface. Similar implicit solvation terms have also been implemented in other popular docking software such as DOCK [118], MOE [119], and FITTED [120].

While implicit water models are useful for the approximation of long-range electrostatic forces considering the above-mentioned screening effect of solvent dielectric [110], they cannot handle hydration shells (Figure 1) and specific water-mediated linkages. However, the absence or presence of a certain water molecule at the binding site can drastically modify the overall affinity of ligands [121,122]. Thus, an accurate calculation of the binding thermodynamics is a rather impossible undertaking without the representation of individual water molecules.

Explicit water models have been introduced to overcome the above-mentioned limitations of implicit approaches. At the MM level, there are various explicit models, such as SPC [123], TIP3P [124], and TIP4P [124], where the abbreviated names refer to the charge systems and sites (parameters) of the water molecule prototype. Explicit approaches allow for the calculation of the energy contributions of individual waters, e.g., using the statistical mechanical inhomogeneous fluid solvation (IFST, [125]) or grid inhomogeneous solvation (GIST, [126]) theory. In this way, the enthalpic and entropic terms of bound waters can be also considered, like in Wscore [127], DOCK-GIST [35], and AutoDock-GIST [38]. In some cases, the incorporation of explicit waters with the above methods has improved the correlation between the experimentally determined binding affinity and the docking score [38,127], while other works have not observed such improvements [87,128].

In the realm of quantum mechanics, theory permits a more accurate calculation of the charge distribution of molecules compared to MM. The assessment of the electrostatic interaction between the solute and water, in theory, can be included in the self-consistent field (SCF) calculation using dielectric continuum models [110]. However, for realistic solute cavities, it requires a numerical iterative process for every SCF cycle, which is extremely computationally demanding [129]. The Conductor-like Screening Model (COSMO) [130] solves this problem using a Green function description with analytical gradients, making the method practically applicable. COSMO can be considered as an advanced version of the polarizable continuum model (PCM, [131,132]), and is the most accurate implicit solvation model for semi-empirical QM. There is also a universal solvation model based on solute electron density (SMD, [133]), which is usually implemented for more computationally demanding levels of QM. At the semi-empirical level, the combination of PM6s [134–137] and PM7 [137–139] parametrizations, combined with the implicit model of COSMO, is a popular choice for estimating the binding affinities of ligands to targets.

Advances in computational speed and linear scaling methods [140] have allowed for the combination of implicit (COSMO) and explicit models handling long-range electrostatics and individual water contributions, respectively. Such hybrid approaches present a fast QM alternative of MM scoring functions for drug design. For example, Nikitina et al. inserted possible interface waters into hydrogen bond donor–acceptor sites and used the PM3 method [141,142]. Horváth et al. predicted interface waters using a molecular-dynamics-based method, MobyWat [40], and utilized the hybrid water model with PM7 parametrization for the estimation of the binding enthalpies (ΔH_b) of ligands [143]. Here, the inclusion of explicit waters in the hybrid model yielded, e.g., a 3-fold smaller relative error when compared with vacuum calculations (Figure 3). Cavasotto et al. used a single-point PM7 calculation, keeping crystallographic waters in the binding interface in their QM docking scoring methodology to show encouraging enrichment factors on 10 protein targets [144]. The latter studies also extract the binding site environment from the target (similarly to Figure 3) to further reduce the computational time. The ΔG_b was also calculated by Hyslova et al., using a DFT-D3 and PM6-D3X4 combined method with crystallographic waters in the binding pocket, achieving a better fit with the combined implicit/explicit procedure ($R^2 = 0.68$) compared to the implicit alone ($R^2 = 0.49$) [145].

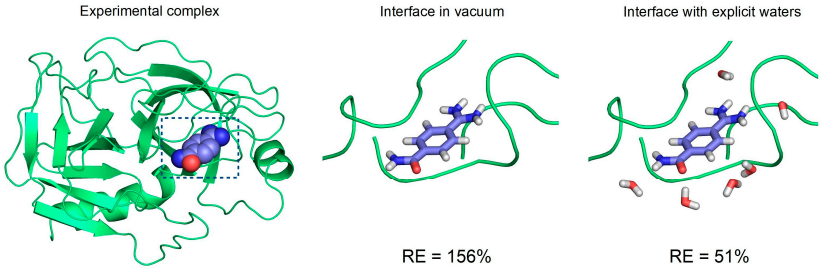

Figure 3. The complex structure of beta-trypsin (on the left, target in cartoon representation) and p-amidinobenzamidine (ligand marked with spheres). The target–ligand interface extracted for ΔH_b calculations is marked with a box. The close-up of the dry (middle) and explicitly hydrated (right, used for hybrid calculations) interface with a ligand in sticks representation. Relative errors (RE) of the calculated binding enthalpy (ΔH_b) values of the dry and hybrid models are indicated below the corresponding interface structure. The RE values were calculated as RE(%) = 100 ($\Delta H_{b,calculated} - \Delta H_{b,experimental}$)/$\Delta H_{b,experimental}$. The coordinates and ΔH_b values were re-used from a previous study [143] (Table 1 and Supporting Supplementary Table S5, $\beta = 0$). The incorporation of explicit water molecules in the ΔH_b calculation considerably reduced the RE in this case.

5. Water in Target-Ligand Docking

Target–ligand complex structures (Figure 1) are key to the engineering of new drugs. Computational docking can supply such atomic-resolution complex structures rapidly and, therefore, it is a widely used [146–148] alternative of experimental structure determination techniques (Section 2) in ligand screening projects [149–151]. Water molecules are active participants in real docking situations, as described in the Introduction (Figure 1). However, the proper use of these water molecules during computational docking is not trivial [36]. The inclusion of happy waters (Figure 1) bridging in the target–ligand interface may help to increase the precision of the docking calculation. On the other hand, if unhappy waters (Figure 1) are included in the interface, they would erroneously block the docking to the target sites used by the ligand in reality. Thus, the misuse of unhappy waters in an interface obviously leads to the mis-docking of the ligand. However, without knowledge of the true hydrated complex structure, it is rather difficult to distinguish between happy and unhappy water molecules in advance. Docking with waters is therefore a true "chicken and egg situation", where let us say water is the chicken and the docked ligand is the egg. Docking is expected to produce a proper target–ligand complex for the decision on the inclusion of happy water molecules in (or the exclusion of unhappy ones from) the docking itself. This awkward situation is reflected in the corresponding literature. Several studies have reported that the incorporation of specific water molecules in the docking process significantly improved the docking performance [127,152–154], while others have found that including water molecules had little effect on this performance [155,156]. Several fast docking tools and strategies [38,92,97,127,157–166] have been developed to incorporate waters in the binding site during docking simulations. Many of these tools work with experimentally determined (known) water positions [160,163–165].

A simple way of incorporating water molecules into docking simulations is to include them as a static part of the target [167], where the positions and orientations of these water molecules are kept restrained during the simulation [162,167]. This strategy is used most in molecular docking studies and has been shown to be effective [168–170]. An improvement to the restrained water model is the displaceable water model, where the included water molecules can be switched on/off automatically during the simulation so that a ligand can keep the favorable water molecules and displace the non-favorable water molecules (GOLD, FlexX, FITTED, and DOCK). These included waters have mostly fixed positions or a limited mobility during the docking process, while some methods allow for the waters to change their positions and orientations through the search algorithm (in FITTED).

Other methods solvate the ligand and then dock the solvated ligand with full flexibility, as the waters are kept or displaced depending on the entropy and/or energy contributions

during the simulation (AutoDock4, MVD, and Glide). Such ligand-centric methods treat water molecules as a flexible part of the ligand, so they present the same flexibility as the ligand itself. RosettaLigand includes water movement both independently and dependently from the ligand during the initial stage, while considering the full flexibility of the target and the ligand through MC search. However, when the method was evaluated on a dataset of 341 diverse protein/ligand complexes from CSAR, no significant improvement was observed in the docking success rate [165]. This could be caused by there being no solvent-specific scoring adjustments in RosettaLigand other than the desolvation energy calculated using an implicit solvent model. Such a desolvation term in a force-field-based scoring function is often calibrated for protein–ligand complexes with no explicit solvent [165].

In many cases, the experimental water positions are not available or the hydration structure is not complete. In such situations, theoretical methods (Section 3) can supply the water positions for the docking calculations. Solvation before docking and a short molecular dynamics (MD) simulation were performed to improve these water positions, and after the encounter of the interacting partners, the water was removed based on a Monte-Carlo approach in HADDOCK [93,171] and a semi-explicit water model implemented in Rosetta [172]. This method improved the docking results for HADDOCK when compared to docking without explicit water molecules. FITTED [120] treats water molecules as a part of the target, and conserved water molecules are considered by an entropic penalty in the final score. This improves the docking accuracy for HIV-1 inhibitors. GOLD relies on two programs, FlexX [97,173] to pre-calculate the energetically favorable water sites and insert spherical water molecules (particles), and Consolv [174] to predict the water molecules that are likely to be displaced [166]. There is an upper limit of water molecules that is handled by this approach, so as not to increase the complexity and therefore the computational costs above a rational limit [36]. Although GOLD predicted the conservation or displacement of water molecules with a high efficiency, the effect on the ligand binding pose prediction was moderate [166]. Slide [160] enables the virtual screening of a relatively large set of ligands using Consolv [174] for the water prediction.

HydroDock [175], a new approach, separates the chicken and egg situation and solves the hydrated docking using a parallel approach. The ligand is docked into the target without waters (dry docking). Simultaneously, the whole target is filled up with the explicit water molecules predicted by MobyWat [13,40]. Then, the dry docked complex and the hydrated target are merged and the water molecules that clash with the dry docked ligand are removed. The resultant complex is then energy minimized to set the proper orientation of the hydrogen bonds, and a short MD simulation is performed to yield the representative binding mode. HydroDock achieved a high accuracy in the case of ion-channel-bound ligand docking (Figure 4), one of the hardest cases of including water molecules in molecular docking [9,175]. As a specific example, the HydroDock method was validated on the ion channels of the influenza A and SARS-CoV-2 viruses.

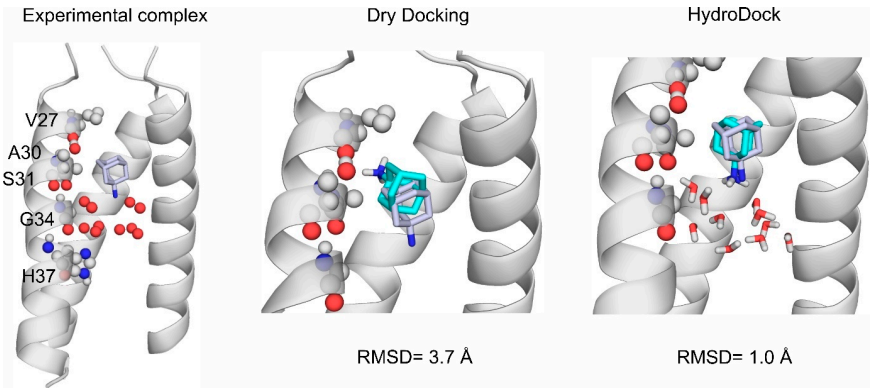

Figure 4. The role of water in ligand binding and the incorporation of explicit water molecules into docking using HydroDock. The influenza virus A ion channel is shown as grey cartoon, target amino acids are shown as spheres and labelled according to the 6bkk [176] PDB structure. Experimental water

oxygen positions are shown as red spheres, and water molecules after HydroDock are shown as red and white sticks [175]. The experimental amantadine structure is shown as grey sticks, and the calculated structures as teal sticks. Dry docking (in the middle) fails to reproduce the experimental ligand binding mode, however, docking with water molecules (on the right) improved similarity with experimental results greatly. Root mean squared deviation (RMSD) is calculated after superimposition of the calculated to the experimental structure, between ligand heavy atoms.

As the inclusion of explicit water molecules increases computational costs [177], the scoring functions of many fast docking methods do not treat explicit water molecules with the proper partial charges and terms for their enthalpic or entropic contributions [177,178]. However, the way that water molecules are treated in the binding site and how their energetic contributions are evaluated is considered to be a key factor greatly affecting the docking performance [127]. To improve this docking performance, the contribution of water-mediated interactions and entropic effects may be considered for individual water molecules. A common modification to scoring functions is to add an entropy penalty, using a positive constant for each included water molecule to model the loss of rigid-body entropy favoring the displacement of the water molecules. However, in the case of large ligands, this approach leads to extremely positive energy contributions, necessitating a modification of the scoring function of AutoDock [179]. Moreover, Friesner and co-workers showed that the contributions of some water molecules to the free energy of binding can be much larger than others [180]. Some attempts have been made to calculate the target surface water sites and thermodynamics prior to the docking process, using third-party tools such as GIST and WaterMap, and to incorporate the solvation information in the scoring function (WScore, AutoDock-GIST, and DOCK-GIST). Although evaluation studies have reported only minor improvements in the success rate of docking for WScore and AutoDock-GIST, such an incorporation of explicit waters into the energy calculation definitely improved the correlation between the experimentally determined binding affinity and the docking score [38,127]. On the other hand, knowledge-based methods such as Consolv [174] (a k-nearest-neighbor-based classifier trained on 5542 molecules taken from 30 independently solved protein structures) can be used to determine the probability of the water molecules in the binding site to be conserved or displaced, as well as their corresponding desolvation penalty values (implemented in Slide) [160]. Instead of on-the-fly energy evaluations, scoring functions with more accurate desolvation functions can be implemented as re-scoring tools after the docking. For example, Wang et al. used molecular-mechanics–Poisson–Boltzmann-surface-area (MM–PBSA) re-scoring to find HIV-1 reverse transcriptase inhibitors [181], and several studies have reported that rescoring using a molecular-mechanics–generalized-Born-surface-area (MM–GBSA) method improved the enrichment of the known ligands for several enzymes and even the identification of substrates [182–184].

In the last decade, targeting protein–protein/DNA/RNA interactions has been considered to be a promising strategy for drug discovery [185–188], and a growing number of docking methods have been specifically developed for this [189]. Their shallow and relatively large interface (more than 1500 Å2 compared to the 300−1000 Å2 range for binding sites) [190] makes it readily accessible to the solvent or water-permeable in the case of nucleic acids. For nucleic acids, unlike proteins, their phosphate groups and corresponding counter ions (such as Mg^{2+} or K^{+}) cause polarization upon the water molecules and functional groups of drugs. Thus, water molecules often play an important role in the ligand recognition and complex stabilization for nucleic acids, as well as proteins. There are several publications that have reported improvement in the success rate of the docking results when water molecules were included for RNA [191], DNA, and protein–protein complexes [192,193]. However, due to their large, solvent-accessible interface, it is extremely challenging to incorporate water molecules into the process of docking macromolecules within a reasonable computation time. Thus, the effect of these water molecules is often ignored in protein–protein/DNA/RNA docking, in which the desol-

vation penalty is estimated as proportional to the solvent-accessible surface area. There have been attempts to incorporate solvation effects in macromolecule docking. For example, HADDOCK explicitly treats water molecules by performing rigid-body docking on solvated macromolecules, followed by a Monte-Carlo (MC) simulation that displaces the waters based on their probabilities to form water-mediated contact, predicted using the Kyte-Doolittle scale [194]. Pavlovicz et al. developed a semi-explicit water model (implemented in Rosetta), in which a modified MC simulation displaces or adds explicit solvent molecules from bulk, followed by an energy evaluation with an implicit solvation energy term. Both attempts have improved the docking and ranking results [171,172].

6. Conclusions

Molecular engineering and drug design have been continuously fueled by the development of experimental structure determination techniques. However, the determination of the position of individual water molecules is often limited by the low resolution of their measurements. Theoretical calculations can supply the atomic-resolution hydration structure of target–ligand interfaces with a high precision, and often complement experimental techniques. The energetic contribution of individual water molecules to the full thermodynamics of target–ligand binding can be also calculated. There has been an improvement in the application of water structures in computational docking, a technique often used in the high throughput virtual screening of ligands in the drug industry. While "docking with waters" is still a problematic "chicken and egg situation", a number of methods have been featured that answer this challenge as well.

Author Contributions: Conceptualization, C.H.; writing—original draft preparation, B.Z.Z., B.B., R.B., V.S., V.M., C.H.; writing—review and editing, B.Z.Z., B.B., R.B., V.S., V.M., C.H.; visualization, B.Z.Z., B.B., R.B., V.S., V.M., C.H.; supervision, C.H.; project administration, B.Z.Z., C.H.; funding acquisition, B.Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Supported by the PTE ÁOK KA-2022-26 grant. The project has been supported by the European Union, co-financed by the European Social Fund. Project name and code: Comprehensive Development for Implementing Smart Specialization Strategies at the University of Pécs, EFOP-3.6.1-16-2016-00004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We acknowledge the support from the Governmental Information Technology Development Agency, Hungary. We acknowledge that the results of this research have been achieved using the DECI resource Archer2 based in the UK at the National Supercomputing Service with support from the PRACE aisbl.

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

References

1. Chumlea, W.C.; Guo, S.S.; Zeller, C.M.; Reo, N.V.; Siervogel, R.M. Total Body Water Data for White Adults 18 to 64 Years of Age: The Fels Longitudinal Study. *Kidney Int.* **1999**, *56*, 244–252. [CrossRef] [PubMed]
2. Elsässer, S.J.; Huang, H.; Lewis, P.W.; Chin, J.W.; Allis, C.D.; Patel, D.J. DAXX Envelops a Histone H3.3–H4 Dimer for H3.3-Specific Recognition. *Nature* **2012**, *491*, 560–565. [CrossRef]
3. Manolaridis, I.; Kulkarni, K.; Dodd, R.B.; Ogasawara, S.; Zhang, Z.; Bineva, G.; O'Reilly, N.; Hanrahan, S.J.; Thompson, A.J.; Cronin, N.; et al. Mechanism of Farnesylated CAAX Protein Processing by the Intramembrane Protease Rce1. *Nature* **2013**, *504*, 301–305. [CrossRef] [PubMed]
4. Musset, B.; Smith, S.M.E.; Rajan, S.; Morgan, D.; Cherny, V.V.; DeCoursey, T.E. Aspartate 112 Is the Selectivity Filter of the Human Voltage-Gated Proton Channel. *Nature* **2011**, *480*, 273–277. [CrossRef] [PubMed]
5. Ostmeyer, J.; Chakrapani, S.; Pan, A.C.; Perozo, E.; Roux, B. Recovery from Slow Inactivation in K+ Channels Is Controlled by Water Molecules. *Nature* **2013**, *501*, 121–124. [CrossRef] [PubMed]
6. Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108*, 74–108. [CrossRef]

7. Ball, P. Water Is an Activematrix of Life for Cell and Molecular Biology. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13327–13335. [CrossRef]
8. Bellissent-Funel, M.C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A.E. Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016**, *116*, 7673–7697. [CrossRef]
9. Zsidó, B.Z.; Hetényi, C. The Role of Water in Ligand Binding. *Curr. Opin. Struct. Biol.* **2021**, *67*, 1–8. [CrossRef]
10. Bodnarchuk, M.S. Water, Water, Everywhere... It's Time to Stop and Think. *Drug Discov. Today* **2016**, *21*, 1139–1146. [CrossRef]
11. de Simone, A.; Dodson, G.G.; Verma, C.S.; Zagari, A.; Fraternali, F. Prion and Water: Tight and Dynamical Hydration Sites Have a Key Role in Structural Stability. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7535–7540. [CrossRef]
12. Miyano, M.; Ago, H.; Saino, H.; Hori, T.; Ida, K. Internally Bridging Water Molecule in Transmembrane α-Helical Kink. *Curr. Opin. Struct. Biol.* **2010**, *20*, 456–463. [CrossRef]
13. Jeszenői, N.; Bálint, M.; Horváth, I.; Van Der Spoel, D.; Hetényi, C. Exploration of Interfacial Hydration Networks of Target-Ligand Complexes. *J. Chem. Inf. Model.* **2016**, *56*, 148–158. [CrossRef]
14. Pradhan, M.R.; Nguyen, M.N.; Kannan, S.; Fox, S.J.; Kwoh, C.K.; Lane, D.P.; Verma, C.S. Characterization of Hydration Properties in Structural Ensembles of Biomolecules. *J. Chem. Inf. Model.* **2019**, *59*, 3316–3329. [CrossRef] [PubMed]
15. Ahmad, M.; Gu, W.; Geyer, T.; Helms, V. Adhesive Water Networks Facilitate Binding of Protein Interfaces. *Nat. Commun.* **2011**, *2*, 261. [CrossRef]
16. Laage, D.; Elsaesser, T.; Hynes, J.T. Water Dynamics in the Hydration Shells of Biomolecules. *Chem. Rev.* **2017**, *117*, 10694–10725. [CrossRef] [PubMed]
17. Bruce Macdonald, H.E.; Cave-Ayland, C.; Ross, G.A.; Essex, J.W. Ligand Binding Free Energies with Adaptive Water Networks: Two-Dimensional Grand Canonical Alchemical Perturbations. *J. Chem. Theory Comput.* **2018**, *14*, 6586–6597. [CrossRef]
18. Zhong, H.; Wang, Z.; Wang, X.; Liu, H.; Li, D.; Liu, H.; Yao, X.; Hou, T. Importance of a Crystalline Water Network in Docking-Based Virtual Screening: A Case Study of BRD4. *Phys. Chem. Chem. Phys.* **2019**, *21*, 25276–25289. [CrossRef] [PubMed]
19. Venkatakrishnan, A.J.; Ma, A.K.; Fonseca, R.; Latorraca, N.R.; Kelly, B.; Betz, R.M.; Asawa, C.; Kobilka, B.K.; Dror, R.O. Diverse GPCRs Exhibit Conserved Water Networks for Stabilization and Activation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3288–3293. [CrossRef]
20. 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*, 15579–15584. [CrossRef]
21. Brysbaert, G.; Blossey, R.; Lensink, M.F. The Inclusion of Water Molecules in Residue Interaction Networks Identifies Additional Central Residues. *Front. Mol. Biosci.* **2018**, *5*, 88. [CrossRef]
22. Jeszenői, N.; Schilli, G.; Bálint, M.; Horváth, I.; Hetényi, C. Analysis of the Influence of Simulation Parameters on Biomolecule-Linked Water Networks. *J. Mol. Graph. Model.* **2018**, *82*, 117–128. [CrossRef]
23. Kunstmann, S.; Gohlke, U.; Broeker, N.K.; Roske, Y.; Heinemann, U.; Santer, M.; Barbirz, S. Solvent Networks Tune Thermodynamics of Oligosaccharide Complex Formation in an Extended Protein Binding Site. *J. Am. Chem. Soc.* **2018**, *140*, 10447–10455. [CrossRef]
24. Rudling, A.; Orro, A.; Carlsson, J. Prediction of Ordered Water Molecules in Protein Binding Sites from Molecular Dynamics Simulations: The Impact of Ligand Binding on Hydration Networks. *J. Chem. Inf. Model.* **2018**, *58*, 350–361. [CrossRef] [PubMed]
25. Jukič, M.; Konc, J.; Gobec, S.; Janežič, D. Identification of Conserved Water Sites in Protein Structures for Drug Design. *J. Chem. Inf. Model.* **2017**, *57*, 3094–3103. [CrossRef] [PubMed]
26. Wahl, J.; Smieško, M. Thermodynamic Insight into the Effects of Water Displacement and Rearrangement upon Ligand Modifications Using Molecular Dynamics Simulations. *ChemMedChem* **2018**, *13*, 1325–1335. [CrossRef]
27. Hüfner-Wulsdorf, T.; Klebe, G. Protein–Ligand Complex Solvation Thermodynamics: Development, Parameterization, and Testing of GIST-Based Solvent Functionals. *J. Chem. Inf. Model.* **2020**, *60*, 1409–1423. [CrossRef]
28. Krimmer, S.G.; Betz, M.; Heine, A.; Klebe, G. Methyl, Ethyl, Propyl, Butyl: Futile but Not for Water, as the Correlation of Structure and Thermodynamic Signature Shows in a Congeneric Series of Thermolysin Inhibitors. *ChemMedChem* **2014**, *9*, 833–846. [CrossRef]
29. García-Sosa, A.T.; Mancera, R.L.; Dean, P.M. WaterScore: A Novel Method for Distinguishing between Bound and Displaceable Water Molecules in the Crystal Structure of the Binding Site of Protein-Ligand Complexes. *J. Mol. Model.* **2003**, *9*, 172–182. [CrossRef] [PubMed]
30. Chen, D.; Li, Y.; Zhao, M.; Tan, W.; Li, X.; Savidge, T.; Guo, W.; Fan, X. Effective Lead Optimization Targeting the Displacement of Bridging Receptor–Ligand Water Molecules. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24399–24407. [CrossRef]
31. Harriman, G.; Greenwood, J.; Bhat, S.; Huang, X.; Wang, R.; Paul, D.; Tong, L.; Saha, A.K.; Westlin, W.F.; Kapeller, R.; et al. Acetyl-CoA Carboxylase Inhibition by ND-630 Reduces Hepatic Steatosis, Improves Insulin Sensitivity, and Modulates Dyslipidemia in Rats. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1796–E1805. [CrossRef]
32. Collin, M.-P.; Lobell, M.; Hübsch, W.; Brohm, D.; Schirok, H.; Jautelat, R.; Lustig, K.; Bömer, U.; Vöhringer, V.; Héroult, M.; et al. Discovery of Rogaratinib (BAY 1163877): A Pan-FGFR Inhibitor. *ChemMedChem* **2018**, *13*, 437–445. [CrossRef] [PubMed]
33. Beuming, T.; Farid, R.; Sherman, W. High-Energy Water Sites Determine Peptide Binding Affinity and Specificity of PDZ Domains. *Protein Sci.* **2009**, *18*, 1609–1619. [CrossRef] [PubMed]

34. Jung, S.W.; Kim, M.; Ramsey, S.; Kurtzman, T.; Cho, A.E. Water Pharmacophore: Designing Ligands Using Molecular Dynamics Simulations with Water. *Sci. Rep.* **2018**, *8*, 10400. [CrossRef]
35. Balius, T.E.; Fischer, M.; Stein, R.M.; Adler, T.B.; Nguyen, C.N.; Cruz, A.; Gilson, M.K.; Kurtzman, T.; Shoichet, B.K. Testing Inhomogeneous Solvation Theory in Structure-Based Ligand Discovery. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6839–E6846. [CrossRef] [PubMed]
36. de Beer, S.; Vermeulen, N.; Oostenbrink, C. The Role of Water Molecules in Computational Drug Design. *Curr. Top. Med. Chem.* **2010**, *10*, 55–66. [CrossRef]
37. Abel, R.; Young, T.; Farid, R.; Berne, B.J.; Friesner, R.A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831. [CrossRef]
38. Uehara, S.; Tanaka, S. AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* **2016**, *21*, 1604. [CrossRef]
39. Nittinger, E.; Schneider, N.; Lange, G.; Rarey, M. Evidence of Water Molecules—A Statistical Evaluation of Water Molecules Based on Electron Density. *J. Chem. Inf. Model.* **2015**, *55*, 771–783. [CrossRef]
40. Jeszenői, N.; Horváth, I.; Bálint, M.; Van Der Spoel, D.; Hetényi, C. Mobility-Based Prediction of Hydration Structures of Protein Surfaces. *Bioinformatics* **2015**, *31*, 1959–1965. [CrossRef]
41. Savage, H.; Wlodawer, A. Determination of Water Structure around Biomolecules Using X-Ray and Neutron Diffraction Methods. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 1986; pp. 162–183.
42. Halle, B. Protein Hydration Dynamics in Solution: A Critical Survey. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2004**, *359*, 1207–1224. [CrossRef]
43. Frank, J. Averaging of Low Exposure Electron Micrographs of Non-Periodic Objects. *Ultramicroscopy* **1975**, *1*, 159–162. [CrossRef]
44. Henderson, R.; Unwin, P.N.T. Three-Dimensional Model of Purple Membrane Obtained by Electron Microscopy. *Nature* **1975**, *257*, 28–32. [CrossRef]
45. Wüthrich, K. The Way to NMR Structures of Proteins. *Nat. Struct. Biol.* **2001**, *8*, 923–925. [CrossRef]
46. Wüthrich, K. Brownian Motion, Spin Diffusion and Protein Structure Determination in Solution. *J. Magn. Reson.* **2021**, *331*, 107031. [CrossRef] [PubMed]
47. Zsidó, B.Z.; Hetényi, C. Molecular Structure, Binding Affinity, and Biological Activity in the Epigenome. *Int. J. Mol. Sci.* **2020**, *21*, 4134. [CrossRef] [PubMed]
48. Berman, H.M.; Battistuz, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 899–907. [CrossRef] [PubMed]
49. Berman, H.M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]
50. Biedermannová, L.; Schneider, B. Hydration of Proteins and Nucleic Acids: Advances in Experiment and Theory. A Review. *Biochim. et Biophys. Acta (BBA)—Gen. Subj.* **2016**, *1860*, 1821–1835. [CrossRef]
51. Mattos, C.; Ringe, D. Solvent Structure. In *International Tables for Crystallography*; International Union of Crystallography: Chester, UK, 2006; pp. 623–647.
52. Lounnas, V.; Pettitt, B.M. A Connected-Cluster of Hydration around Myoglobin: Correlation between Molecular Dynamics Simulations and Experiment. *Proteins: Struct. Funct. Genet.* **1994**, *18*, 133–147. [CrossRef]
53. Kossiakoff, A.A.; Sintchak, M.D.; Shpungin, J.; Presta, L.G. Analysis of Solvent Structure in Proteins Using Neutron D2O-H2O Solvent Maps: Pattern of Primary and Secondary Hydration of Trypsin. *Proteins: Struct. Funct. Genet.* **1992**, *12*, 223–236. [CrossRef] [PubMed]
54. Shpungin, J.; Kossiakoff, A.A. [24] A Method of Solvent Structure Analysis for Proteins Using D_2O-H_2O Neutron Difference Maps. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 1986; pp. 329–342.
55. Chatake, T.; Fujiwara, S. A Technique for Determining the Deuterium/Hydrogen Contrast Map in Neutron Macromolecular Crystallography. *Acta Crystallogr. D Struct. Biol.* **2016**, *72*, 71–82. [CrossRef] [PubMed]
56. Tanaka, I.; Chatake, T.; Fujiwara, S.; Hosoya, T.; Kusaka, K.; Niimura, N.; Yamada, T.; Yano, N. Current Status and near Future Plan of Neutron Protein Crystallography at J-PARC. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 2020; pp. 101–123.
57. Kono, F.; Kurihara, K.; Tamada, T. Current Status of Neutron Crystallography in Structural Biology. *Biophys. Physicobiol* **2022**, *19*, e190009. [CrossRef] [PubMed]
58. Schiffer, C.; Hermans, J. Promise of Advances in Simulation Methods for Protein Crystallography: Implicit Solvent Models, Time-Averaging Refinement, and Quantum Mechanical Modeling. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 2003; pp. 412–461.
59. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221. [CrossRef]
60. Adams, P.D.; Grosse-Kunstleve, R.W.; Hung, L.-W.; Ioerger, T.R.; McCoy, A.J.; Moriarty, N.W.; Read, R.J.; Sacchettini, J.C.; Sauter, N.K.; Terwilliger, T.C. PHENIX: Building New Software for Automated Crystallographic Structure Determination. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 1948–1954. [CrossRef]

61. Echols, N.; Morshed, N.; Afonine, P.V.; McCoy, A.J.; Miller, M.D.; Read, R.J.; Richardson, J.S.; Terwilliger, T.C.; Adams, P.D. Automated Identification of Elemental Ions in Macromolecular Crystal Structures. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 1104–1114. [CrossRef]
62. Afonine, P.V.; Grosse-Kunstleve, R.W.; Adams, P.D. A Robust Bulk-Solvent Correction and Anisotropic Scaling Procedure. *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *61*, 850–855. [CrossRef]
63. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [CrossRef]
64. Langer, G.; Cohen, S.X.; Lamzin, V.S.; Perrakis, A. Automated Macromolecular Model Building for X-ray Crystallography Using ARP/WARP Version 7. *Nat. Protoc.* **2008**, *3*, 1171–1179. [CrossRef]
65. Lamb, A.L.; Kappock, T.J.; Silvaggi, N.R. You Are Lost without a Map: Navigating the Sea of Protein Structures. *Biochim. et Biophys. Acta (BBA)—Proteins Proteom.* **2015**, *1854*, 258–268. [CrossRef]
66. Lamzin, V.S.; Wilson, K.S. Automated Refinement for Protein Crystallography. In *Methods in Enzymology*; Academic Press: Cambridge, MA, USA, 1997; pp. 269–305.
67. Levitt, M.; Park, B.H. Water: Now You See It, Now You Don't. *Structure* **1993**, *1*, 223–226. [CrossRef]
68. Deng, G.-H.; Shen, Y.; Chen, H.; Chen, Y.; Jiang, B.; Wu, G.; Yang, X.; Yuan, K.; Zheng, J. Ordered-to-Disordered Transformation of Enhanced Water Structure on Hydrophobic Surfaces in Concentrated Alcohol–Water Solutions. *J. Phys. Chem. Lett.* **2019**, *10*, 7922–7928. [CrossRef]
69. Carugo, O. Correlation between Occupancy and B Factor of Water Molecules in Protein Crystal Structures. *Protein Eng. Des. Sel.* **1999**, *12*, 1021–1024. [CrossRef]
70. Reuhl, M.; Vogel, M. Temperature-Dependent Dynamics at Protein–Solvent Interfaces. *J. Chem. Phys.* **2022**, *157*, 074705. [CrossRef] [PubMed]
71. Cheng, Y.; Glaeser, R.M.; Nogales, E. How Cryo-EM Became so Hot. *Cell* **2017**, *171*, 1229–1231. [CrossRef] [PubMed]
72. Pintilie, G.; Zhang, K.; Su, Z.; Li, S.; Schmid, M.F.; Chiu, W. Measurement of Atom Resolvability in Cryo-EM Maps with Q-Scores. *Nat. Methods* **2020**, *17*, 328–334. [CrossRef]
73. Zhang, X.; Walker, S.B.; Chipman, P.R.; Nibert, M.L.; Baker, T.S. Reovirus Polymerase Λ3 Localized by Cryo-Electron Microscopy of Virions at a Resolution of 7.6 Å. *Nat. Struct. Mol. Biol.* **2003**, *10*, 1011–1018. [CrossRef] [PubMed]
74. Kühlbrandt, W. The Resolution Revolution. *Science (1979)* **2014**, *343*, 1443–1444. [CrossRef]
75. Li, X.; Mooney, P.; Zheng, S.; Booth, C.R.; Braunfeld, M.B.; Gubbens, S.; Agard, D.A.; Cheng, Y. Electron Counting and Beam-Induced Motion Correction Enable Near-Atomic-Resolution Single-Particle Cryo-EM. *Nat. Methods* **2013**, *10*, 584–590. [CrossRef]
76. Allegretti, M.; Mills, D.J.; McMullan, G.; Kühlbrandt, W.; Vonck, J. Atomic Model of the F420-Reducing [NiFe] Hydrogenase by Electron Cryo-Microscopy Using a Direct Electron Detector. *eLife* **2014**, *3*, e01963. [CrossRef]
77. Amunts, A.; Brown, A.; Bai, X.; Llácer, J.L.; Hussain, T.; Emsley, P.; Long, F.; Murshudov, G.; Scheres, S.H.W.; Ramakrishnan, V. Structure of the Yeast Mitochondrial Large Ribosomal Subunit. *Science (1979)* **2014**, *343*, 1485–1489. [CrossRef]
78. Liao, M.; Cao, E.; Julius, D.; Cheng, Y. Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy. *Nature* **2013**, *504*, 107–112. [CrossRef]
79. Renaud, J.-P.; Chari, A.; Ciferri, C.; Liu, W.; Rémigy, H.-W.; Stark, H.; Wiesmann, C. Cryo-EM in Drug Discovery: Achievements, Limitations and Prospects. *Nat. Rev. Drug Discov.* **2018**, *17*, 471–492. [CrossRef]
80. Pintilie, G.; Chiu, W. Validation, Analysis and Annotation of Cryo-EM Structures. *Acta Crystallogr. D Struct. Biol.* **2021**, *77*, 1142–1152. [CrossRef] [PubMed]
81. Prisant, M.G.; Williams, C.J.; Chen, V.B.; Richardson, J.S.; Richardson, D.C. New Tools in MolProbity Validation: CaBLAM for CryoEM Backbone, UnDowser to Rethink "Waters," and NGL Viewer to Recapture Online 3D Graphics. *Protein Sci.* **2020**, *29*, 315–329. [CrossRef]
82. Hryc, C.F.; Baker, M.L. Beyond the Backbone: The Next Generation of Pathwalking Utilities for Model Building in CryoEM Density Maps. *Biomolecules* **2022**, *12*, 773. [CrossRef] [PubMed]
83. Armstrong, B.D.; Han, S. Overhauser Dynamic Nuclear Polarization To Study Local Water Dynamics. *J. Am. Chem. Soc.* **2009**, *131*, 4641–4647. [CrossRef]
84. Otting, G. NMR Studies of Water Bound to Biological Molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *31*, 259–285. [CrossRef]
85. 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*, 237–244. [CrossRef]
86. 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*, 10095–10112. [CrossRef]
87. Nittinger, E.; Gibbons, P.; Eigenbrot, C.; Davies, D.R.; Maurer, B.; Yu, C.L.; Kiefer, J.R.; Kuglstatter, A.; Murray, J.; Ortwine, D.F.; et al. Water Molecules in Protein–Ligand Interfaces. Evaluation of Software Tools and SAR Comparison. *J. Comput. Aided Mol. Des.* **2019**, *33*, 307–330. [CrossRef] [PubMed]
88. 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*, 1867–1881. [CrossRef]
89. Vedani, A.; Huhta, D.W. Algorithm for the Systematic Solvation of Proteins Based on the Directionality of Hydrogen Bonds. *J. Am. Chem. Soc.* **1991**, *113*, 5860–5862. [CrossRef]

90. Pitt, W.R.; Goodfellow, J.M. Modelling of Solvent Positions around Polar Groups in Proteins. *Protein Eng. Des. Sel.* **1991**, *4*, 531–537. [CrossRef] [PubMed]
91. Schymkowitz, J.W.H.; Rousseau, F.; Martins, I.C.; Ferkinghoff-Borg, J.; Stricher, F.; Serrano, L. Prediction of Water and Metal Binding Sites and Their Affinities by Using the Fold-X Force Field. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10147–10152. [CrossRef] [PubMed]
92. Forli, S.; Olson, A.J. A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J. Med. Chem.* **2012**, *55*, 623–638. [CrossRef]
93. van Dijk, A.D.J.; Bonvin, A.M.J.J. Solvated Docking: Introducing Water into the Modelling of Biomolecular Complexes. *Bioinformatics* **2006**, *22*, 2340–2347. [CrossRef]
94. Huggins, D.J.; Tidor, B. Systematic Placement of Structural Water Molecules for Improved Scoring of Protein-Ligand Interactions. *Protein Eng. Des. Sel.* **2011**, *24*, 777–789. [CrossRef]
95. Li, Y.; Gao, Y.; Holloway, M.K.; Wang, R. Prediction of the Favorable Hydration Sites in a Protein Binding Pocket and Its Application to Scoring Function Formulation. *J. Chem. Inf. Model.* **2020**, *60*, 4359–4375. [CrossRef]
96. Virtanen, J.J.; Makowski, L.; Sosnick, T.R.; Freed, K.F. Modeling the Hydration Layer around Proteins: HyPred. *Biophys. J.* **2010**, *99*, 1611–1619. [CrossRef]
97. Rarey, M.; Kramer, B.; Lengauer, T. The Particle Concept: Placing Discrete Water Molecules during Protein-Ligand Docking Predictions. *Proteins: Struct. Funct. Genet.* **1999**, *34*, 17–28. [CrossRef]
98. Wei, W.; Luo, J.; Waldispühl, J.; Moitessier, N. Predicting Positions of Bridging Water Molecules in Nucleic Acid–Ligand Complexes. *J. Chem. Inf. Model.* **2019**, *59*, 2941–2951. [CrossRef]
99. Bayden, A.S.; Moustakas, D.T.; Joseph-McCarthy, D.; Lamb, M.L. Evaluating Free Energies of Binding and Conservation of Crystallographic Waters Using SZMAP. *J. Chem. Inf. Model.* **2015**, *55*, 1552–1565. [CrossRef]
100. Ross, G.A.; Morris, G.M.; Biggin, P.C. Rapid and Accurate Prediction and Scoring of Water Molecules in Protein Binding Sites. *PLoS ONE* **2012**, *7*, e32036. [CrossRef]
101. Mason, J.S.; Bortolato, A.; Weiss, D.R.; Deflorian, F.; Tehan, B.; Marshall, F.H. High End GPCR Design: Crafted Ligand Design and Druggability Analysis Using Protein Structure, Lipophilic Hotspots and Explicit Water Networks. *In Silico Pharmacol.* **2013**, *1*, 23. [CrossRef]
102. Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J.S. A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands and Proteins (FLAP): Theory and Application. *J. Chem. Inf. Model.* **2007**, *47*, 279–294. [CrossRef]
103. Nittinger, E.; Flachsenberg, F.; Bietz, S.; Lange, G.; Klein, R.; Rarey, M. Placement of Water Molecules in Protein Structures: From Large-Scale Evaluations to Single-Case Examples. *J. Chem. Inf. Model.* **2018**, *58*, 1625–1637. [CrossRef] [PubMed]
104. Bui, H.-H.; Schiewe, A.J.; Haworth, I.S. WATGEN: An Algorithm for Modeling Water Networks at Protein-Protein Interfaces. *J. Comput. Chem.* **2007**, *28*, 2241–2251. [CrossRef] [PubMed]
105. Hu, B.; Lill, M.A. WATsite: Hydration Site Prediction Program with PyMOL Interface. *J. Comput. Chem.* **2014**, *35*, 1255–1260. [CrossRef] [PubMed]
106. Barillari, C.; Taylor, J.; Viner, R.; Essex, J.W. Classification of Water Molecules in Protein Binding Sites. *J. Am. Chem. Soc.* **2007**, *129*, 2577–2587. [CrossRef]
107. Huang, P.; Xing, H.; Zou, X.; Han, Q.; Liu, K.; Sun, X.; Wu, J.; Fan, J. Accurate Prediction of Hydration Sites of Proteins Using Energy Model with Atom Embedding. *Front. Mol. Biosci.* **2021**, *8*, 756075. [CrossRef]
108. Lazaridis, T.; Karplus, M. Thermodynamics of Protein Folding: A Microscopic View. *Biophys. Chem.* **2002**, *100*, 367–395. [CrossRef]
109. Warshel, A. Energetics of Enzyme Catalysis. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 5250–5254. [CrossRef]
110. Cramer, C.J.; Truhlar, D.G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200. [CrossRef] [PubMed]
111. Spoel, D.; Zhang, J.; Zhang, H. Quantitative Predictions from Molecular Simulations Using Explicit or Implicit Interactions. *WIREs Comput. Mol. Sci.* **2022**, *12*, e1560. [CrossRef]
112. Zhang, J.; Zhang, H.; Wu, T.; Wang, Q.; van der Spoel, D. Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *J. Chem. Theory Comput.* **2017**, *13*, 1034–1043. [CrossRef] [PubMed]
113. Kuhn, B.; Kollman, P.A. A Ligand That Is Predicted to Bind Better to Avidin than Biotin: Insights from Computational Fluorine Scanning. *J. Am. Chem. Soc.* **2000**, *122*, 3909–3916. [CrossRef]
114. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508. [CrossRef]
115. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]
116. Stouten, P.F.W.; Frömmel, C.; Nakamura, H.; Sander, C. An Effective Solvation Term Based on Atomic Occupancies for Use in Protein Simulations. *Mol. Simul.* **1993**, *10*, 97–120. [CrossRef]
117. Mehler, E.L.; Solmajer, T. Electrostatic Effects in Proteins: Comparison of Dielectric and Charge Models. *Protein Eng. Des. Sel.* **1991**, *4*, 903–910. [CrossRef]
118. Allen, W.J.; Balius, T.E.; Mukherjee, S.; Brozell, S.R.; Moustakas, D.T.; Lang, P.T.; Case, D.A.; Kuntz, I.D.; Rizzo, R.C. DOCK 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* **2015**, *36*, 1132–1156. [CrossRef]

119. *Molecular Operating Environment (MOE)*; 2022.02 Chemical Computing Group ULC: Montreal, QC, Canada, 2023; Available online: https://www.chemcomp.com/Research-Citing_MOE.htm (accessed on 18 July 2023).
120. Corbeil, C.R.; Englebienne, P.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 1. Development and Validation of FITTED 1.0. *J. Chem. Inf. Model.* **2007**, *47*, 435–449. [CrossRef]
121. Liu, C.; Wrobleski, S.T.; Lin, J.; Ahmed, G.; Metzger, A.; Wityak, J.; Gillooly, K.M.; Shuster, D.J.; McIntyre, K.W.; Pitt, S.; et al. 5-Cyanopyrimidine Derivatives as a Novel Class of Potent, Selective, and Orally Active Inhibitors of P38α MAP Kinase. *J. Med. Chem.* **2005**, *48*, 6261–6270. [CrossRef]
122. Nasief, N.N.; Tan, H.; Kong, J.; Hangauer, D. Water Mediated Ligand Functional Group Cooperativity: The Contribution of a Methyl Group to Binding Affinity Is Enhanced by a COO—Group Through Changes in the Structure and Thermodynamics of the Hydration Waters of Ligand–Thermolysin Complexes. *J. Med. Chem.* **2012**, *55*, 8283–8302. [CrossRef] [PubMed]
123. Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 331–342.
124. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
125. Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, *102*, 3531–3541. [CrossRef]
126. 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*, 044101. [CrossRef] [PubMed]
127. Murphy, R.B.; Repasky, M.P.; Greenwood, J.R.; Tubert-Brohman, I.; Jerome, S.; Annabhimoju, R.; Boyles, N.A.; Schmitz, C.D.; Abel, R.; Farid, R.; et al. WScore: A Flexible and Accurate Treatment of Explicit Water Molecules in Ligand–Receptor Docking. *J. Med. Chem.* **2016**, *59*, 4364–4384. [CrossRef]
128. Bucher, D.; Stouten, P.; Triballeau, N. Shedding Light on Important Waters for Drug Design: Simulations versus Grid-Based Methods. *J. Chem. Inf. Model.* **2018**, *58*, 692–699. [CrossRef]
129. Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235. [CrossRef]
130. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799–805. [CrossRef]
131. Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3094. [CrossRef] [PubMed]
132. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669–681. [CrossRef]
133. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
134. Dobeš, P.; Řezáč, J.; Fanfrlík, J.; Otyepka, M.; Hobza, P. Semiempirical Quantum Mechanical Method PM6-DH2X Describes the Geometry and Energetics of CK2-Inhibitor Complexes Involving Halogen Bonds Well, While the Empirical Potential Fails. *J. Phys. Chem. B* **2011**, *115*, 8581–8589. [CrossRef]
135. Pecina, A.; Meier, R.; Fanfrlík, J.; Lepšík, M.; Řezáč, J.; Hobza, P.; Baldauf, C. The SQM/COSMO Filter: Reliable Native Pose Identification Based on the Quantum-Mechanical Description of Protein–Ligand Interactions and Implicit COSMO Solvation. *Chem. Commun.* **2016**, *52*, 3312–3315. [CrossRef]
136. Fanfrlík, J.; Bronowska, A.K.; Řezáč, J.; Přenosil, O.; Konvalinka, J.; Hobza, P. A Reliable Docking/Scoring Scheme Based on the Semiempirical Quantum Mechanical PM6-DH2 Method Accurately Covering Dispersion and H-Bonding: HIV-1 Protease with 22 Ligands. *J. Phys. Chem. B* **2010**, *114*, 12666–12678. [CrossRef]
137. Urquiza-Carvalho, G.A.; Fragoso, W.D.; Rocha, G.B. Assessment of Semiempirical Enthalpy of Formation in Solution as an Effective Energy Function to Discriminate Native-like Structures in Protein Decoy Sets. *J. Comput. Chem.* **2016**, *37*, 1962–1972. [CrossRef]
138. Sulimov, A.V.; Kutov, D.C.; Katkova, E.V.; Ilin, I.S.; Sulimov, V.B. New Generation of Docking Programs: Supercomputer Validation of Force Fields and Quantum-Chemical Methods for Docking. *J. Mol. Graph. Model.* **2017**, *78*, 139–147. [CrossRef]
139. Sulimov, A.V.; Kutov, D.C.; Taschilova, A.S.; Ilin, I.S.; Stolpovskaya, N.V.; Shikhaliev, K.S.; Sulimov, V.B. In Search of Non-Covalent Inhibitors of SARS-CoV-2 Main Protease: Computer Aided Drug Design Using Docking and Quantum Chemistry. *Supercomput. Front. Innov.* **2020**, *7*. [CrossRef]
140. Stewart, J.J.P. Application of Localized Molecular Orbitals to the Solution of Semiempirical Self-Consistent Field Equations. *Int. J. Quantum Chem.* **1996**, *58*, 133–146. [CrossRef]
141. Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. Semiempirical Calculations of Binding Enthalpy for Protein-Ligand Complexes. *Int. J. Quantum Chem.* **2004**, *97*, 747–763. [CrossRef]
142. Nikitina, E.; Sulimov, V.; Grigoriev, F.; Kondakova, O.; Luschenka, S. Mixed Implicit/Explicit Solvation Modelsin Quantum Mechanical Calculations OfBinding Enthalpy for Protein–LigandComplexes. *Int. J. Quantum Chem.* **2006**, *106*, 1943–1963. [CrossRef]

143. Horváth, I.; Jeszenői, N.; Bálint, M.; Paragi, G.; Hetényi, C. A Fragmenting Protocol with Explicit Hydration for Calculation of Binding Enthalpies of Target-Ligand Complexes at a Quantum Mechanical Level. *Int. J. Mol. Sci.* **2019**, *20*, 4384. [CrossRef] [PubMed]
144. Cavasotto, C.N.; Aucar, M.G. High-Throughput Docking Using Quantum Mechanical Scoring. *Front. Chem.* **2020**, *8*, 246. [CrossRef]
145. Hylsová, M.; Carbain, B.; Fanfrlík, J.; Musilová, L.; Haldar, S.; Köprülüoğlu, C.; Ajani, H.; Brahmkshatriya, P.S.; Jorda, R.; Kryštof, V.; et al. Explicit Treatment of Active-Site Waters Enhances Quantum Mechanical/Implicit Solvent Scoring: Inhibition of CDK2 by New Pyrazolo[1,5-a]Pyrimidines. *Eur. J. Med. Chem.* **2017**, *126*, 1118–1128. [CrossRef]
146. Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **2019**, *20*, 4331. [CrossRef] [PubMed]
147. Śledź, P.; Caflisch, A. Protein Structure-Based Drug Design: From Docking to Molecular Dynamics. *Curr. Opin. Struct. Biol.* **2018**, *48*, 93–102. [CrossRef] [PubMed]
148. Dong, D.; Xu, Z.; Zhong, W.; Peng, S. Parallelization of Molecular Docking: A Review. *Curr. Top. Med. Chem.* **2018**, *18*, 1015–1028. [CrossRef]
149. Ballante, F.; Kooistra, A.J.; Kampen, S.; de Graaf, C.; Carlsson, J. Structure-Based Virtual Screening for Ligands of G Protein–Coupled Receptors: What Can Molecular Docking Do for You? *Pharmacol. Rev.* **2021**, *73*, 1698–1736. [CrossRef]
150. Kitchen, D.B.; Decornez, H.; Furr, J.R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949. [CrossRef]
151. Potlitz, F.; Link, A.; Schulig, L. Advances in the Discovery of New Chemotypes through Ultra-Large Library Docking. *Expert. Opin. Drug Discov.* **2023**, *18*, 303–313. [CrossRef]
152. Lu, S.-Y.; Jiang, Y.-J.; Lv, J.; Zou, J.-W.; Wu, T.-X. Role of Bridging Water Molecules in GSK3β-Inhibitor Complexes: Insights from QM/MM, MD, and Molecular Docking Studies. *J. Comput. Chem.* **2011**, *32*, 1907–1918. [CrossRef] [PubMed]
153. Santos, R.; Hritz, J.; Oostenbrink, C. Role of Water in Molecular Docking Simulations of Cytochrome P450 2D6. *J. Chem. Inf. Model.* **2010**, *50*, 146–154. [CrossRef] [PubMed]
154. Kumar, A.; Zhang, K.Y.J. Investigation on the Effect of Key Water Molecules on Docking Performance in CSARdock Exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1880–1892. [CrossRef] [PubMed]
155. de Graaf, C.; Pospisil, P.; Pos, W.; Folkers, G.; Vermeulen, N.P.E. Binding Mode Prediction of Cytochrome P450 and Thymidine Kinase Protein−Ligand Complexes by Consideration of Water and Rescoring in Automated Docking. *J. Med. Chem.* **2005**, *48*, 2308–2318. [CrossRef] [PubMed]
156. Birch, L.; Murray, C.; Hartshorn, M.; Tickle, I.; Verdonk, M. Sensitivity of Molecular Docking to Induced Fit Effects in Influenza Virus Neuraminidase. *J. Comput. Aided Mol. Des.* **2002**, *16*, 855–869. [CrossRef] [PubMed]
157. Lu, J.; Hou, X.; Wang, C.; Zhang, Y. Incorporating Explicit Water Molecules and Ligand Conformation Stability in Machine-Learning Scoring Functions. *J. Chem. Inf. Model.* **2019**, *59*, 4540–4549. [CrossRef]
158. Sun, H.; Zhao, L.; Peng, S.; Huang, N. Incorporating Replacement Free Energy of Binding-Site Waters in Molecular Docking. *Proteins: Struct. Funct. Bioinform.* **2014**, *82*, 1765–1776. [CrossRef]
159. Mahmoud, A.H.; Masters, M.R.; Yang, Y.; Lill, M.A. Elucidating the Multiple Roles of Hydration for Accurate Protein-Ligand Binding Prediction via Deep Learning. *Commun. Chem.* **2020**, *3*, 19. [CrossRef]
160. Schnecke, V.; Kuhn, L.A. Virtual Screening with Solvation and Ligand-Induced Complementarity. In *Virtual Screening: An Alternative or Complement to High Throughput Screening?* Kluwer Academic Publishers: Dordrecht, The Netherlands, 2010; pp. 171–190.
161. Therrien, E.; Weill, N.; Tomberg, A.; Corbeil, C.R.; Lee, D.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 7. Impact of Protein Flexibility and Water Molecules on Docking-Based Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2014**, *54*, 3198–3210. [CrossRef]
162. Lie, M.A.; Thomsen, R.; Pedersen, C.N.S.; Schiøtt, B.; Christensen, M.H. Molecular Docking with Ligand Attached Water Molecules. *J. Chem. Inf. Model.* **2011**, *51*, 909–917. [CrossRef] [PubMed]
163. Huang, N.; Shoichet, B.K. Exploiting Ordered Waters in Molecular Docking. *J. Med. Chem.* **2008**, *51*, 4862–4865. [CrossRef] [PubMed]
164. Davis, I.W.; Baker, D. RosettaLigand Docking with Full Ligand and Receptor Flexibility. *J. Mol. Biol.* **2009**, *385*, 381–392. [CrossRef]
165. Lemmon, G.; Meiler, J. Towards Ligand Docking Including Explicit Interface Water Molecules. *PLoS ONE* **2013**, *8*, e67536. [CrossRef]
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*, 6504–6515. [CrossRef]
167. Stanzione, F.; Giangreco, I.; Cole, J.C. Use of Molecular Docking Computational Tools in Drug Discovery. *Prog. Med. Chem.* **2021**, *60*, 273–343.
168. Roberts, B.C.; Mancera, R.L. Ligand−Protein Docking with Water Molecules. *J. Chem. Inf. Model.* **2008**, *48*, 397–408. [CrossRef]
169. 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*, 726–741. [CrossRef] [PubMed]

170. Thilagavathi, R.; Mancera, R.L. Ligand−Protein Cross-Docking with Water Molecules. *J. Chem. Inf. Model.* **2010**, *50*, 415–421. [CrossRef] [PubMed]
171. Kastritis, P.L.; Visscher, K.M.; van Dijk, A.D.J.; Bonvin, A.M.J.J. Solvated Protein-Protein Docking Using Kyte-Doolittle-Based Water Preferences. *Proteins: Struct. Funct. Bioinform.* **2013**, *81*, 510–518. [CrossRef]
172. Pavlovicz, R.E.; Park, H.; DiMaio, F. Efficient Consideration of Coordinated Water Molecules Improves Computational Protein-Protein and Protein-Ligand Docking Discrimination. *PLoS Comput. Biol.* **2020**, *16*, e1008103. [CrossRef] [PubMed]
173. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489. [CrossRef]
174. Raymer, M.L.; Sanschagrin, P.C.; Punch, W.F.; Venkataraman, S.; Goodman, E.D.; Kuhn, L.A. Predicting Conserved Water-Mediated and Polar Ligand Interactions in Proteins Using a K-Nearest-Neighbors Genetic Algorithm. *J. Mol. Biol.* **1997**, *265*, 445–464. [CrossRef]
175. Zsidó, B.Z.; Börzsei, R.; Szél, V.; Hetényi, C. Determination of Ligand Binding Modes in Hydrated Viral Ion Channels to Foster Drug Design and Repositioning. *J. Chem. Inf. Model.* **2021**, *61*, 4011–4022. [CrossRef]
176. Thomaston, J.L.; Polizzi, N.F.; Konstantinidi, A.; Wang, J.; Kolocouris, A.; DeGrado, W.F. Inhibitors of the M2 Proton Channel Engage and Disrupt Transmembrane Networks of Hydrogen-Bonded Waters. *J. Am. Chem. Soc.* **2018**, *140*, 15219–15226. [CrossRef]
177. Bello, M.; Martínez-Archundia, M.; Correa-Basurto, J. Automated Docking for Novel Drug Discovery. *Expert. Opin. Drug Discov.* **2013**, *8*, 821–834. [CrossRef]
178. Yuriev, E.; Agostino, M.; Ramsland, P.A. Challenges and Advances in Computational Docking: 2009 in Review. *J. Mol. Recognit.* **2011**, *24*, 149–164. [CrossRef]
179. Hetényi, C.; Paragi, G.; Maran, U.; Timár, Z.; Karelson, M.; Penke, B. Combination of a Modified Scoring Function with Two-Dimensional Descriptors for Calculation of Binding Affinities of Bulky, Flexible Ligands to Proteins. *J. Am. Chem. Soc.* **2006**, *128*, 1233–1239. [CrossRef] [PubMed]
180. Young, T.; Abel, R.; Kim, B.; Berne, B.; Friesner, R. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein–Ligand Binding. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 808–813. [CrossRef] [PubMed]
181. Wang, J.; Kang, X.; Kuntz, I.D.; Kollman, P.A. Hierarchical Database Screenings for HIV-1 Reverse Transcriptase Using a Pharmacophore Model, Rigid Docking, Solvation Docking, and MM−PB/SA. *J. Med. Chem.* **2005**, *48*, 2432–2444. [CrossRef]
182. Huang, N.; Kalyanaraman, C.; Irwin, J.J.; Jacobson, M.P. Physics-Based Scoring of Protein−Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening. *J. Chem. Inf. Model.* **2006**, *46*, 243–253. [CrossRef] [PubMed]
183. Kalyanaraman, C.; Bernacki, K.; Jacobson, M.P. Virtual Screening against Highly Charged Active Sites: Identifying Substrates of Alpha−Beta Barrel Enzymes. *Biochemistry* **2005**, *44*, 2059–2071. [CrossRef]
184. Perola, E. Minimizing False Positives in Kinase Virtual Screens. *Proteins Struct. Funct. Bioinform.* **2006**, *64*, 422–435. [CrossRef]
185. Collie, G.W.; Parkinson, G.N. The Application of DNA and RNA G-Quadruplexes to Therapeutic Medicines. *Chem. Soc. Rev.* **2011**, *40*, 5867. [CrossRef] [PubMed]
186. Dasari, S.; Bernard Tchounwou, P. Cisplatin in Cancer Therapy: Molecular Mechanisms of Action. *Eur. J. Pharmacol.* **2014**, *740*, 364–378. [CrossRef]
187. Howe, J.A.; Wang, H.; Fischmann, T.O.; Balibar, C.J.; Xiao, L.; Galgoci, A.M.; Malinverni, J.C.; Mayhood, T.; Villafania, A.; Nahvi, A.; et al. Selective Small-Molecule Inhibition of an RNA Structural Element. *Nature* **2015**, *526*, 672–677. [CrossRef]
188. Wang, M.; Yu, Y.; Liang, C.; Lu, A.; Zhang, G. Recent Advances in Developing Small Molecules Targeting Nucleic Acid. *Int. J. Mol. Sci.* **2016**, *17*, 779. [CrossRef]
189. Feng, Y.; Yan, Y.; He, J.; Tao, H.; Wu, Q.; Huang, S.-Y. Docking and Scoring for Nucleic Acid–Ligand Interactions: Principles and Current Status. *Drug Discov. Today* **2022**, *27*, 838–847. [CrossRef]
190. Ran, X.; Gestwicki, J.E. Inhibitors of Protein–Protein Interactions (PPIs): An Analysis of Scaffold Choices and Buried Surface Area. *Curr. Opin. Chem. Biol.* **2018**, *44*, 75–86. [CrossRef]
191. 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*, 1134–1146. [CrossRef] [PubMed]
192. Mayol, G.F.; Defelipe, L.A.; Arcon, J.P.; Turjanski, A.G.; Marti, M.A. Solvent Sites Improve Docking Performance of Protein–Protein Complexes and Protein–Protein Interface-Targeted Drugs. *J. Chem. Inf. Model.* **2022**, *62*, 3577–3588. [CrossRef] [PubMed]
193. 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. Bioinform.* **2014**, *82*, 916–932. [CrossRef]
194. Kyte, J.; Doolittle, R.F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [CrossRef] [PubMed]