

Selecting an Optimal Number of Binding Site Waters To Improve Virtual Screening Enrichments Against the Adenosine A_{2A} Receptor

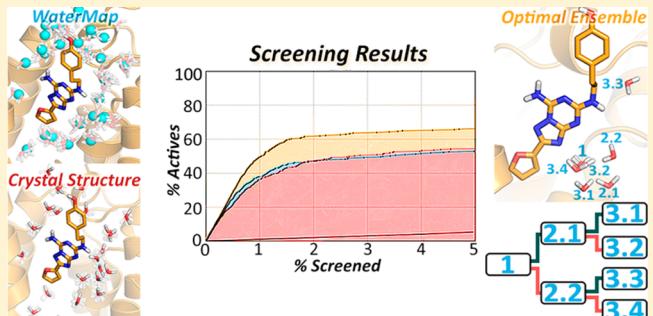
Eelke B. Lenselink,[†] Thijs Beuming,[‡] Woody Sherman,[‡] Herman W. T. van Vlijmen,[†] and Adriaan P. IJzerman*,[†]

[†]Division of Medicinal Chemistry, Leiden Academic Centre for Drug Research, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

[‡]Schrödinger, Inc., 120 West 45th Street, New York, New York 10036, United States

Supporting Information

ABSTRACT: A major challenge in structure-based virtual screening (VS) involves the treatment of explicit water molecules during docking in order to improve the enrichment of active compounds over decoys. Here we have investigated this in the context of the adenosine A_{2A} receptor, where water molecules have previously been shown to be important for achieving high enrichment rates with docking, and where the positions of some binding site waters are known from a high-resolution crystal structure. The effect of these waters (both their presence and orientations) on VS enrichment was assessed using a carefully curated set of 299 high affinity A_{2A} antagonists and 17,337 decoys. We show that including certain crystal waters greatly improves VS enrichment and that optimization of water hydrogen positions is needed in order to achieve the best results. We also show that waters derived from a molecular dynamics simulation — without any knowledge of crystallographic waters — can improve enrichments to a similar degree as the crystallographic waters, which makes this strategy applicable to structures without experimental knowledge of water positions. Finally, we used decision trees to select an ensemble of structures with different water molecule positions and orientations that outperforms any single structure with water molecules. The approach presented here is validated against independent test sets of A_{2A} receptor antagonists and decoys from the literature. In general, this water optimization strategy could be applied to any target with waters-mediated protein–ligand interactions.



INTRODUCTION

Water plays a key role in the molecular recognition of small molecules binding to proteins.^{1–4} In order to bind, both the ligand and the protein must be desolvated, either partially or fully. Solvation has often been described using implicit solvent methodologies such as Poisson–Boltzmann^{5,6} or Generalized Born methods,^{7,8} although more recent work suggests that explicit solvent is needed to more accurately describe the energetics of protein–ligand binding.^{9–13} Often binding site water molecules remain bound in the presence of the ligand, where they can form polar bridging interactions between protein and ligand.¹⁴ As such, it is expected that prediction of ligand poses and affinity values derived from docking studies could benefit from an accurate representation of bridging waters.¹⁵ It is however not straightforward to determine which waters should be selected, and in which orientations, especially when important waters can be missing from the crystal structures due to insufficient resolution or a broad density distribution. Furthermore, even when the positions of key water molecules are known from X-ray crystallography, the orientation of hydrogen atoms are not determined by the density alone, except for cases of extremely high-resolution structures.

For G protein-coupled receptors (GPCRs) the number of crystal structures has rapidly increased in recent years.¹⁶ These structures have revealed water-mediated protein–ligand interactions for a number of GPCRs, including the adenosine A_{2A} receptor in complex with ZM241385,¹⁷ the smoothed receptor in complex with LY2940680,¹⁸ the κ-opioid receptor in complex with JDTic,¹⁹ and the μ-opioid receptor in complex with a morphinan antagonist.²⁰ However, in the majority of GPCR structures water molecules are not observed, most likely as a consequence of their relatively poor resolution, which has been shown to be a determining factor for the number of water molecules observed in crystal structures.²¹ It is therefore possible that additional water-mediated protein–ligand interactions for these targets will be revealed when resolution can be improved.

Despite the limited resolution of GPCR crystal structures, a number of prospective structure-based virtual screening (VS) campaigns have been successfully conducted, such as on the adenosine A_{2A} receptor,^{22,23} the β₂ adrenergic receptor,^{24,25} the dopamine D3 receptor,²⁶ the CXCR4 receptor,²⁷ the

Received: January 22, 2014

Published: May 16, 2014

muscarinic M₂/M₃ receptor,²⁸ and the histamine H₁ receptor.²⁹ However, explicit water molecules were only considered in one case,²² where a substantial increase in enrichment was obtained using three specific water molecules in the first crystal structure of the adenosine A_{2A} receptor.³⁰ Recently, a higher resolution (1.8 Å) crystal structure of the adenosine A_{2A} receptor was published,¹⁷ which revealed an additional 47 water molecules, many of which were near the small molecule antagonist ZM241385. The abundance of water molecules in this crystal structure provides an opportunity for a systematic analysis of the effect of including these individual waters in VS.

Here, we evaluate how the placement and orientation of water molecules can impact docking virtual screening enrichments. We used two methods to place water molecules: 1) oxygen positions were obtained from the crystal structure, and hydrogens were placed using the interactive optimizer implemented in the Protein Preparation Wizard in Maestro^{31,32} or 2) molecular dynamics was used to sample water configurations, and solvent clustering was used to determine regions of highly localized waters.

We analyzed the effects of including individual water molecules and of combining multiple individual waters in an ensemble method using decision trees (DT). We used excluded volume (EV) restraints to assess whether important waters were contributing to favorable enrichments merely by preventing compounds from binding in subpockets (as suggested previously²²) or through more specific polar interactions. The method was further validated using independent test sets derived from previously published antagonists and decoys.^{22,33,34}

METHODS

All calculations were performed using tools in the Schrödinger Suite.³¹ The 4EIY crystal structure of the adenosine A_{2A} receptor was retrieved from the Protein Data Bank (PDB),³⁵ subsequently prepared using the Protein Preparation Wizard,³² and protonation states were assigned using PROPKA.³⁶ Docking was done with Glide 6.0^{37,38} using the SP scoring function with default settings. All enrichment metrics were calculated using the Enrichment Calculator from the Schrödinger Script Center (<http://www.schrodinger.com/scriptcenter/>). BEDROC values were calculated as described by Truchon et al.³⁹ Enrichment Factors (EF) were calculated using the following formula: $EF = (a/n)/(A/N)$, where a is the number of actives found in sample size n , A is the total number of actives, and N is the total number of compounds. All structure based images were generated using PyMOL.⁴⁰

Decoy Set Generation. An active set was generated using ChEMBL_16⁴¹ as a starting point. First, all high affinity A_{2A} receptor ligands (pK_i of 8 or higher) were selected. Based on our prior experience in the field of adenosine receptors, we subjected the selected ligands to manual curation, in order to correct compounds for which the reported K_i values were incorrectly retrieved from the original literature. All ribose-containing ligands were removed from the set, as virtually all of these compounds have agonistic properties, whereas the 4EIY structure is an antagonist-bound receptor. This procedure yielded 299 unique antagonists. Next, 50 decoys were matched to each antagonist based on the rules proposed for the Database of Useful Decoys: enhanced (DUD-e) by Mysinger et al.³⁴ To find decoys matching the antagonists, the Web service was used (<http://dude.docking.org/generate>) as accessed on 25 July, 2013.

In short, this Web service finds decoys based on the ligand protonation state (pH range 6–8 predicted with Epik^{42,43}) and physicochemical properties (miLogP, rotatable bonds, hydrogen bond acceptors, hydrogen bond donors, and net charge). Next the 25% most dissimilar decoys were selected (based on Tanimoto similarity and ECFP4 fingerprints⁴⁴). Subsequently, ligands and decoys were prepared using LigPrep 2.6 to expand all reasonable tautomer and ionization states at pH 7.0. The complete procedure yielded 345 states (299 compounds, some in multiple protonation states) for the actives and 17337 states for the decoys. The ligand and decoy sets used in this study are available as Supporting Information.

Selection of Water Orientations. Crystal structure waters were selected within a 6 Å radius of the ligand in PDB ID 4EIY, which yielded 22 unique waters. For each of the 22 waters, all orientations were evaluated from the interactive Protein Preparation Wizard (PPW) optimizer, resulting in a total of 345 water orientations. Alternatively, water orientations were also derived from a molecular dynamics (MD) simulation and solvent clustering as implemented in WaterMap v1.4^{9,10}. The WaterMap simulations started with the 4EIY structure from the PDB, from which the solvent was removed but the ligand (ZM241385) was retained.

The WaterMap methodology has been described in detail by Abel et al.,¹⁰ and only brief implementation details are given here. The protein binding site was used in all simulations, truncated 15 Å from the ligand and solvated with a box of TIP4P water molecules extending at least 10 Å from all protein atoms. The OPLS2005 force field was used for all the simulations. No membrane was included in these simulations since the protein heavy atoms were restrained. The solvated structures were then relaxed via a multistage protocol which involved two rounds of restrained minimization, followed by a series of four short (12–120 ps) molecular dynamics runs. These runs gradually increased the temperature of the system from 10 K to 300 K and equilibrated the pressure to 1 atm. The hydration site statistics were then calculated from the 2 ns production simulation with the non-hydrogen protein and ligand atoms harmonically restrained. Water molecules were spatially clustered from the trajectory based on a 1.0 Å cluster radius. All hydration sites extracted from the WaterMap within a 6 Å radius of the ligand were used as an input for the docking. Nineteen WaterMap hydration sites were found within 1 Å of a crystallographic water molecule. The water orientations (134 when the occupancy is 1.0) from each hydration site were extracted from the MD trajectory, and a subset of orientations was selected using the Conformer Cluster script within Maestro. The number of orientations (i.e., the number of cluster members) selected from each of the WaterMap hydration sites was determined using the Kelly criterion, which represents the trade-off between the cluster size and spread of the clusters.⁴⁵ This resulted in a total of 236 different orientations for the 19 WaterMap hydration sites. In addition, we also included 47 extra water orientations for hydration sites that were present within a 6 Å radius of the ligand but were not matched with a water molecule in the crystal structure. The thermodynamic information (entropy, enthalpy, and free energy) of the hydration sites determined by the WaterMap algorithm was not used in this study. To assess the effect of using a different crystal structure of the adenosine A_{2A} receptor, water positions and orientations were also determined for the XAC-bound crystal structure (PDB: 3REY).⁴⁶ Since no water molecules were resolved in this structure, the waters were

determined using WaterMap alone with the procedure as described above. To limit the number of calculations, we did two rounds of docking, first to single water orientations from the hydration site and next to orientations of the top 5 enriched waters. Using this protocol a total of 27 waters and 73 water orientations were considered for 3REY.

Ensemble Methods. Docking scores for all compounds against all models with individual water orientations were obtained and placed into a matrix using a custom Python script. This was done for both crystal structure-derived and WaterMap waters, resulting in a total of 628 ($345 + 236 + 47$) models. Only the best scoring LigPrep state for each compound was considered. We used scikit⁴⁷ to develop Decision Trees (DT) based on these results using the following procedure: Compounds were assigned to be either active (1) or inactive (0), and the docking scores were used as predictor variables. Based on these matrices, different DT were trained and cross-validated using the Matthews Correlation Coefficient and 5-fold cross-validation. Default settings were used, and the depth of the tree was varied (max_depth: 2–6) to avoid overfitting. Based on the cross-validation the final tree depth was set to 3, resulting in 8 leaves. Subsequently, actives and decoys were ranked first based on the statistics of each leaf (percentage ligand) and second based on the docking score of the final node. P-values for the different enrichment metrics were calculated using the Z-scores based on docking to the crystallographic or computed water positions. In principle these DT merge the docking results from all the single water orientations, and the ensemble can be seen as an optimal collection of structures with single water orientations. The workflow for creating decision trees based on different pose files is available upon request.

Excluded Volume Restraints. To assess whether specific polar interactions were responsible for improvements in enrichment, we replaced a subset of waters from the ensemble with excluded volume (EV) restraints. These included waters named 2521-5 (WM), 2522-35 (PPW), 2584-14 (PPW), 2521-9 (WM), and 2521-7 (WM). Restraints were placed on the three water atoms using a van der Waals radius of 1.09 Å for the hydrogen atoms and 1.52 Å for the oxygen atom.

Other Decoy Sets. The ensemble method was further validated using different published decoy sets. The decoy set from DUD-e³⁴ was retrieved and subsequently docked. The decoy sets from Katritch et al.²² and Gatica et al.³³ were first converted to 3D using LigPrep 2.6 with default settings.

Chemical Space Analysis. The 299 ligands from our data set were mapped onto a 2D-plot using the following procedure: MOLPRINT2D fingerprints⁴⁸ were generated within Canvas⁴⁹ and exported to a text file keeping the 1024 most significant bits. This file was used as an input for the t-Distributed Stochastic Neighbor Embedding algorithm (t-SNE)⁵⁰ where a perplexity of 10 was used and the initial dimensions was set to 200. Using these coordinates a “chemical space map” was generated on which the ligands were mapped. These chemical space maps were visualized using Dotmatics Vortex-v2013.12.25046.

■ RESULTS

Crystallographic and Computed Waters. The importance of using water molecules and the dependence on their orientations in VS was evaluated using 299 high affinity antagonists for the adenosine A_{2A} receptor and an appropriate decoy set of 17337 compounds (see Methods). A total of 345

orientations were generated for 22 crystallographic waters with the interactive optimizer in the Protein Preparation Wizard (PPW; Figure 1). Additional water positions and orientations

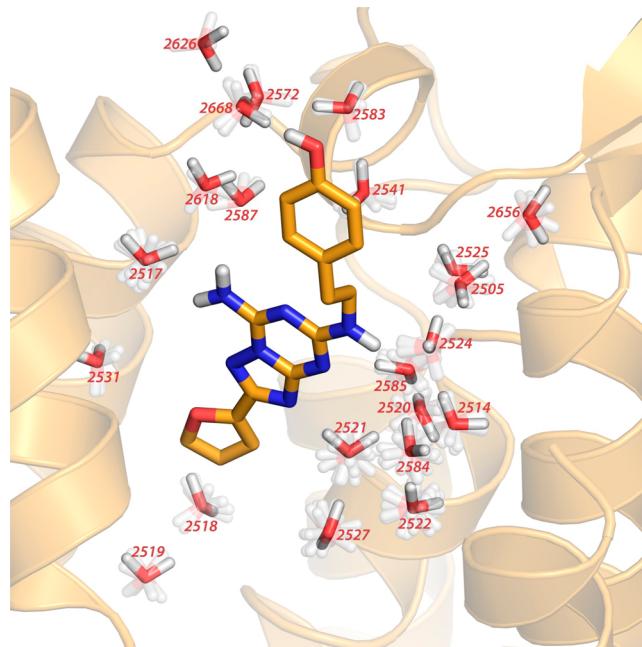


Figure 1. Water molecules from the 4EIY crystal structure used in this study. PDB residue numbers are shown for each water molecule. The default water orientation selected by the Protein Preparation Wizard is rendered opaque. Alternative water orientations are shown transparent.

were computed with WaterMap; 19 of the 22 crystal structure waters were matched within 1.0 Å by a WaterMap hydration site. A total of 236 water orientations were selected from these 19 hydration sites, plus an additional 47 water orientations derived from computed water molecule locations that were not present in the crystal structure (Figure 2). For each of the 628 waters ($345 + 236 + 47$), Glide grids were generated for the protein and a single water. Ligands and decoys were subsequently docked into the grids using the SP scoring function in Glide.

For enrichment calculations, we focused on BEDROC ($\alpha = 160.9$) values,⁴⁰ since we are interested in ranking the active compounds early in the full compound set. The BEDROC values for different waters, either computed (blue) or taken from the crystal structure (red), are shown in Figure 3. Each point represents a different water orientation. The mean and standard deviation for all BEDROC ($\alpha = 160.9$) values were 0.421 and 0.093, respectively. Based on this benchmark, only for 5 of the 19 waters, orientations were found that showed a significantly higher enrichment (Z-scores converted to p-values and $p < 0.05$). These are water sites 2514, 2520, 2521, 2522, and 2584 (numbering according to the residue numbers in the 4EIY PDB file), which are all located in approximately the same part of the binding pocket, formed by TM domains 1, 2, 3, and helix 7, surrounded by residues Ala81,^{3,29} Leu85,^{3,33} Ile274,^{7,39} and His278^{7,43} (Figures 1 and 2).

For the 5 best-performing waters, their default orientation selected by the PPW using the crystal structure waters yielded above average enrichments in most cases, with the only exception being water 2514, where the default orientation

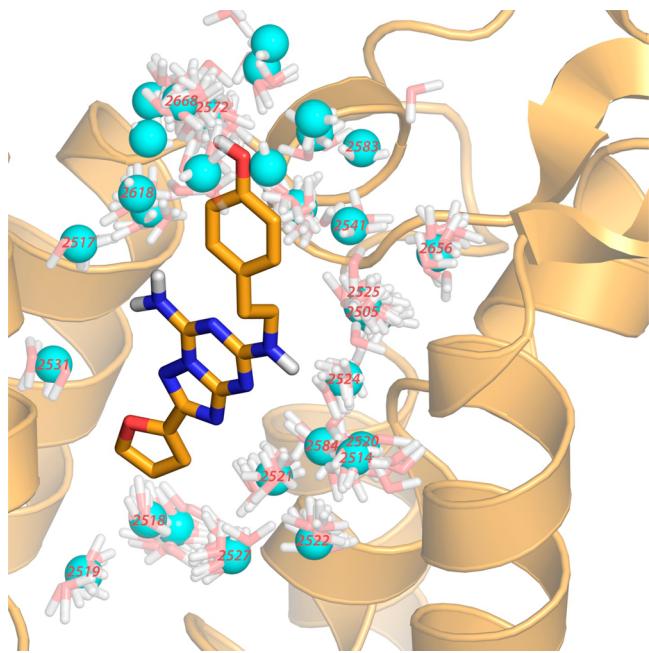


Figure 2. Waters derived from WaterMap analysis. Residue numbers are shown for each water molecule as occurring in the crystal structure. Cyan spheres represent the center of the WaterMap hydration sites.

produced the lowest BEDROC enrichment (0.357). A nondefault water orientation led to improvement in enrichment for all cases. Additionally, for these 5 waters, orientations derived from WaterMap performed equally well or better than crystal structure-derived waters. Overall, the best enrichment was achieved using the computed waters (2514, 2520, 2521, and 2584). The highest achieved BEDROC ($\alpha = 160.9$) enrichments overall from both the crystallographic and computed waters were 0.624 (2520-6) and 0.673 (2521-5), respectively (see Table 1).

Table 1. Docking Results Based on Water Conformations Generated Using the Protein Preparation Wizard (PPW) and WaterMap (WM)^a

water	BEDROC		enrichment factor	
	$\alpha = 160.9$	$\alpha = 20.0$	1%	2%
no water	0.467	0.507	25	19
average WM&PPW	0.421	0.452	23	17
2520-6 (PPW)	0.624*	0.555	36	23
2521-5 (WM)	0.673**	0.553	39	28
ensemble (PPW)	0.772***	0.634	48	25
ensemble (WM)	0.812***	0.674	47	30
ensemble (WM&PPW)	0.818***	0.696	50	31

^aThe optimal enrichment using a single water conformation is reported for both the crystal structure (2520-6) and WaterMap (2521-5) derived waters. Ensemble docking using multiple PPW and WM waters significantly boosts enrichment as compared with the averages for these waters. Significance levels were calculated for the BEDROC ($\alpha = 160.9$) values as compared to the individual waters on which the ensemble was based, with Z-scores converted to p-values: < 0.05 (*), < 0.01 (**), and < 0.001 (***)�. A combination of waters from both methods yielded the best result, and the final Decision Tree was based on this result (Ensemble WM&PPW).

Ensemble of Waters. To test if multiple waters could be incorporated into an ensemble docking workflow, Decision Trees (DTs) were constructed using data from all the docking results. DTs were chosen as a method because of the potential for identifying and classifying ligands interacting with distinct water molecules and orientations. Three DTs were generated for the crystallographic waters, the computed waters, and for a combination of the two methods. Since almost twice as many water orientations were evaluated in the VS as there were ligands, the depths of the DTs were limited. Limiting the interaction depth of the DT was done to avoid both overfitting and to obtain an ensemble with an optimal size (i.e., practical in a prospective virtual screen). Validation of the depth of the

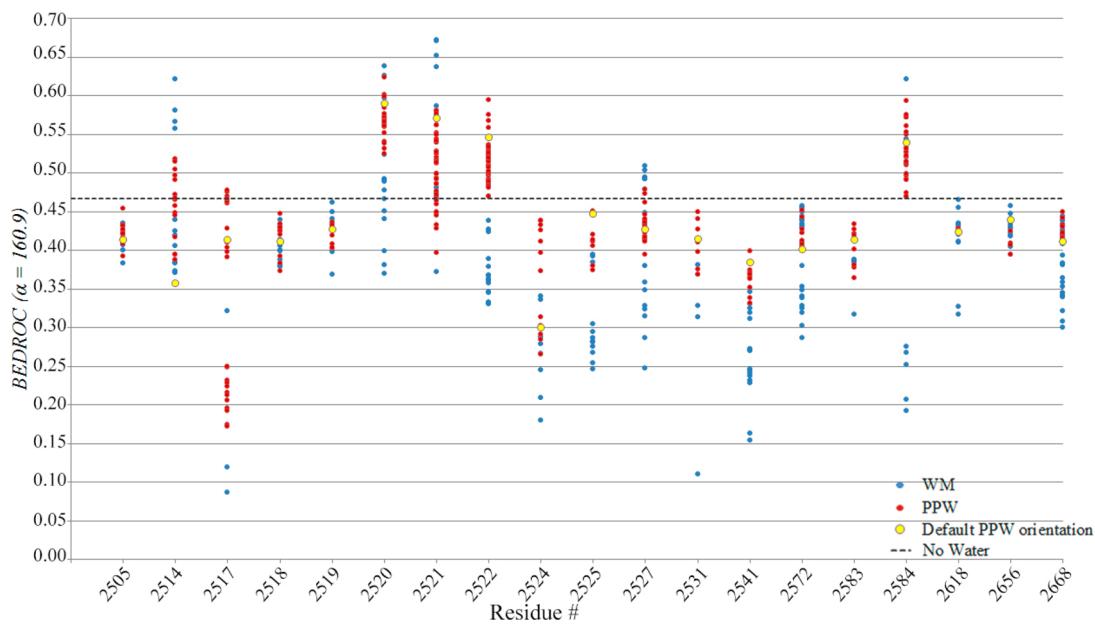


Figure 3. Comparison of virtual screening BEDROC ($\alpha = 160.9$) values for each water molecule. Water orientations were obtained by the Protein Preparation Wizard (red) or WaterMap (blue). The BEDROC value obtained in the absence of any water molecule is indicated by the dashed black line, and the BEDROC value of the default selected water conformation of the Protein Preparation Wizard is indicated by a yellow circle.

DTs was performed using 5-fold cross-validation, setting the final interaction depth to 3, resulting in 8 different leaves.

Results of the DT ensemble were significantly better than inclusion of a single water molecule or no water at all, as shown in Table 1. In all cases the DT-ensembles of waters outperformed the single best water orientation by a significant margin. For example, the best performing DT-ensemble, based on both the crystallographic and computed waters, doubled the enrichment factor in the top 1% of the screen (EF1%) when compared with including no water (50 vs 25). As the EF1% maximum value for this data set is 58, this DT-ensemble reaches enrichments close to the maximum possible value. This is reflected by the large fraction of ligands located in several of the leaves (Figure 4). For example, in leaf 1 88 out of the 90

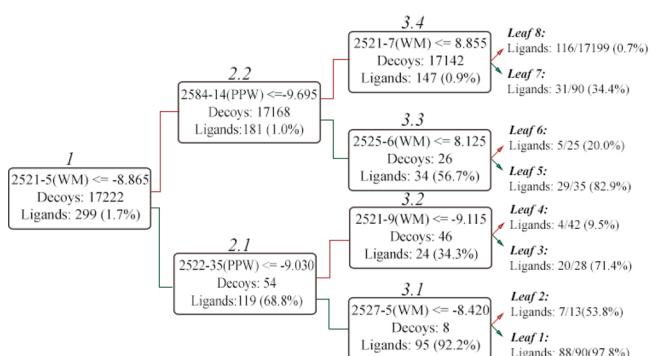


Figure 4. Structure and flowchart of the Decision Tree (DT). Statistics are shown based on the decoy set. Nodes are numbered based on the depth-level to which they belong. The final DT was implemented in a Knime workflow.⁶⁷ Red lines represent links where the criterion was not met, and green lines represent lines where the criterion was met.

compounds are actives (98% of the compounds in this leaf), followed by leaf 5, where 29 out 35 are actives (83%), and finally leaf 3, where 20 ligands out of 28 are actives (71%). The lowest ranked leaf is leaf 8, where 99% of compounds are decoys.

The waters selected in the final ensemble are shown in Figure 5, and the DT nodes they belong to are indicated in parentheses. Most of these waters provided enrichments higher than including no water when considered individually. Indeed, in the DT methodology the first node in the tree is by definition the best performing water in the entire set (2521-5). However, in the lower nodes waters performing worse than including no water are selected (2527-5 (3.1) and 2525-6 (3.3)), as the BEDROC ($\alpha = 160.9$) enrichment with these two waters were 0.359 and 0.286, respectively. While including these waters individually does not produce a high enrichment, they do increase the overall enrichment of the DT by moving some decoys into the lower ranked leaves when presented with a small set of compounds (e.g., see nodes 3.2 and 3.4 in Figure 4).

Excluded Volume. To further test if the hydrogen bonding properties of these water molecules contribute to high enrichments, six of the best performing water molecules from the ensembles were replaced with excluded volume (EV) restraints in Glide, one at a time. Subsequently, the ligand and decoy set was docked into these models using these EV restraints. Significantly lower enrichments were found when using EV restraints instead of water molecules. Especially the early enrichment seems to be benefiting from inclusion of

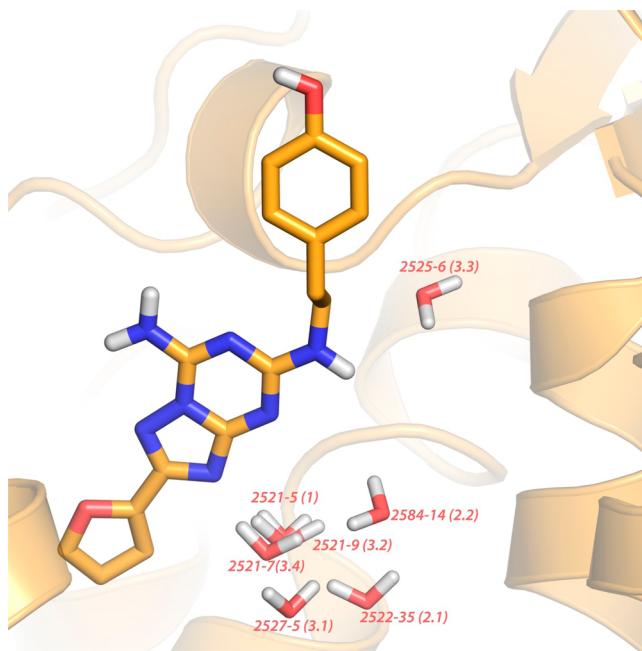


Figure 5. Final selection of water conformations from the Decision Tree (DT). Residue numbers are shown as well as the DT nodes shown in Figure 4, indicated between parentheses.

specific waters (Figure 5); however, for 5 out of the 6 water molecules tested, replacing them with EV restraints still yielded enrichments slightly higher than including no water at all, suggesting that there is also a steric contribution to the enrichment. The exception to this is 2522-35 (PPW), where replacing the water with an excluded volume resulted in a substantial decrease in enrichment: BEDROC ($\alpha = 160.9$) values for 2522-35 (PPW) versus EV restraints were 0.575 and 0.390, respectively (see Figure 6).

Other Decoys Sets. To further validate the performance of the DT-ensembles, we tested our method on three other previously published ligand and decoy sets (from Katritch et

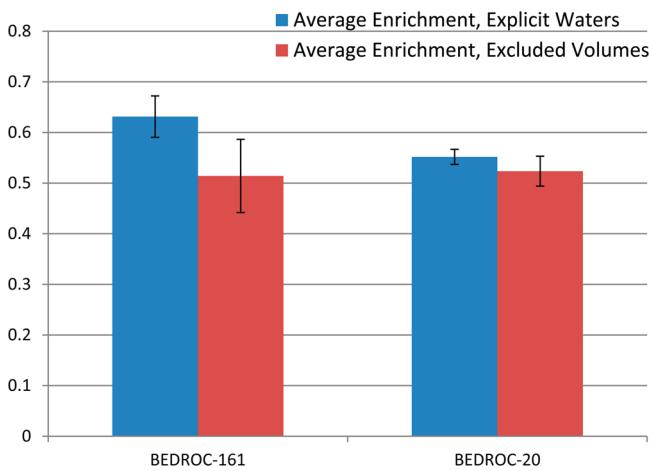


Figure 6. Average BEDROC values ($\alpha = 160.9$ and $\alpha = 20$) for six of the selected water molecules compared with using Excluded Volume restraints in place of these water molecules. The standard deviation within the different water orientations/excluded volumes is shown.

al.,²² the GPCR-decoy database (GDD),³³ and the DUD-e database³⁴). The sizes of these data sets are quite different: 23 ligands and 2000 decoys from Katritch et al., 361 ligands and 14079 decoys from GDD, and 482 ligands and 31550 decoys from DUD-e.

The performance of the DT-ensemble on these three decoy sets is summarized in Table 2. For two data sets the ensemble

Table 2. Evaluation of the Performance of the DT-Based Ensemble Docking on Other Published Decoy Sets^a

data set	BEDROC		enrichment factor		percentage ligand leaf 1
	$\alpha = 160.9$	$\alpha = 20.0$	1%	2%	
Katritch et al. ²²	0.816	0.712	63 (78)	34	83% (15/18)
GLL/ GDD ³³	0.802	0.518	36	22 (11.7)	94% (121/129)
DUD-e ³⁴	0.650	0.545	38 (22)	21	94% (119/125)

^aThe enrichment metric from the original publications is shown in parentheses. Furthermore the percentage ligand present in the best leaf (leaf 1) is shown, with the fraction of ligands shown in parentheses.

method achieved a 2-fold higher enrichment than previously reported. For the GDD set, the DT-ensemble yielded an EF2% of 22, which is an improvement over a previously published value of 11.7%.³³ Additionally, for the DUD-e decoy set, the DT ensemble achieved an EF1% of 38, compared to a published value of 22%.³⁴

For the Katritch et al. set, enrichments were also high, although slightly lower than reported by Katritch et al., with EF1% of 63 for the DT-ensemble versus 78 previously reported.²² In all three decoy sets, the largest contribution to the enrichment came from leaf 1, having ligand percentages between 83% and 94%.

Ligand and Crystal Structure Dependence. In order to assess the sensitivity of the approach to different crystal structures, the analysis was repeated using the XAC-bound structure of the human adenosine A_{2A} receptor (3REY). In this crystal structure there are no crystal structure waters available so only orientations from the WaterMap were used. When the XAC ligand was superimposed on the WaterMap of 4EIY, we found that the waters favorable for enrichment from the 4EIY WaterMap are clashing with one of the bulky propyl side chains of XAC (Supplementary Figure 1), suggesting that 3REY derived waters would not produce similar results. In total, 78 water orientations were taken from the 3REY WaterMap and used to create Glide grids for docking. BEDROC ($\alpha = 160.9$) values for the structure without any water (0.190) are worse than the results obtained for 4EIY (0.467, see Table 1 and Table 3). The best performing water increases enrichment by

Table 3. Performance of the Decoy Set and Method on the Crystal Structure of the Human Adenosine A_{2A} Receptor Bound to XAC (3REY) Including (i) No Water, (ii) the Best Water, and (iii) the Decision Tree-Based Ensemble Method

	BEDROC		enrichment factor	
	$\alpha = 160.9$	$\alpha = 20.0$	1%	2%
no water	0.190	0.224	9	5.8
best water	0.245	0.225	11	6
ensemble	0.267	0.243	12	6.5

approximately 29% to 0.245 (compared to an increase of 44% for 4EIY). In addition, the use of the DT leads to a further increase of 41% to 0.267 (75% for 4EIY). Thus, although the absolute magnitudes and increases in enrichment are not as large as for 4EIY, the procedure of optimizing water orientations and developing a DT leads to substantial gains in enrichment irrespective of the structure. The DT seems to be particularly beneficial for early enrichment; 13 ligands are found in the top 25 for the model without water molecules, 20 for the best water orientation and 25 for the DT. Merging the 3REY and 4EIY data sets into a single DT did not improve results over using 4EIY alone, likely because of the large difference in baseline enrichment for the two structures. The best water of 3REY was located close to the 4EIY hydration site 2518 but with a different orientation than the waters found in the WaterMap of 4EIY (Figure 2 and Supplementary Figure 2).

Analysis of the Chemical Space in Different DT Leaves. In order to assess the results not only based on enrichment performance but also in terms of the diversity of hits recovered, the 299 ligands of our ligand set were mapped on a “chemical space” plot using a nonlinear algorithm called t-SNE⁵⁰ (Figure 7). Three plots were generated for both structures 4EIY (top) and 3REY (bottom), one for the crystal structure without water (left), one for the crystal structure with the optimal single water (middle), and one for the DT based ensemble (right). Based on the data in Figure 7 it is clear that for the DT based ensemble derived from the water orientations of 4EIY yields not only the best enrichment but also the highest diversity. For the 4EIY results it is clear that, in addition to higher overall enrichment, the diversity of hits is gradually increased going from no water, to a single water orientation, to the DT ensemble, as evidenced by the larger spreading of green points in the t-SNE plots. For example, in the DT based ensemble additional ligands in the lower left (e.g., compound 3) and upper right corners are found within the threshold of EF1%.

When we compare the results with the data from 3REY it is clear that the lower enrichment for this structure is due to the fact that only ligands of a single chemotype are under the EF1% threshold (top left cluster, e.g. compound 4), both in the absence of water and when including the single best water. The results improve when the DT based ensemble is used, with additional compounds in other clusters in the EF1%. For example, a ligand similar to XAC (example 6) was observed in the DT ensemble. Notably, all the ligands in the upper left cluster enriched by 3REY are also recovered by the ensemble method of 4EIY.

DISCUSSION

In this study the effects of incorporating explicit water molecules in VS against the adenosine A_{2A} receptor were systematically studied. The adenosine A_{2A} receptor is a suitable target for such a study because 1) a new 1.8 Å high resolution structure with multiple water molecules recently became available,¹⁷ 2) waters interact with ligands in the adenosine A_{2A} receptor as observed in the crystal structure, and 3) it had been shown previously that water molecules are important for pose prediction and VS enrichment.^{22,51} In the first study, performed by Ivanov et al.⁵¹ pose prediction of ZM241385 was evaluated using a lower, 2.6 Å resolution crystal structure (3EML). Inclusion of three water molecules directly interacting with the ligand (equivalent to water 2521, 2584, and 2668 in 4EIY) improved the RMSD from 10.27 to 0.90 Å and proved to

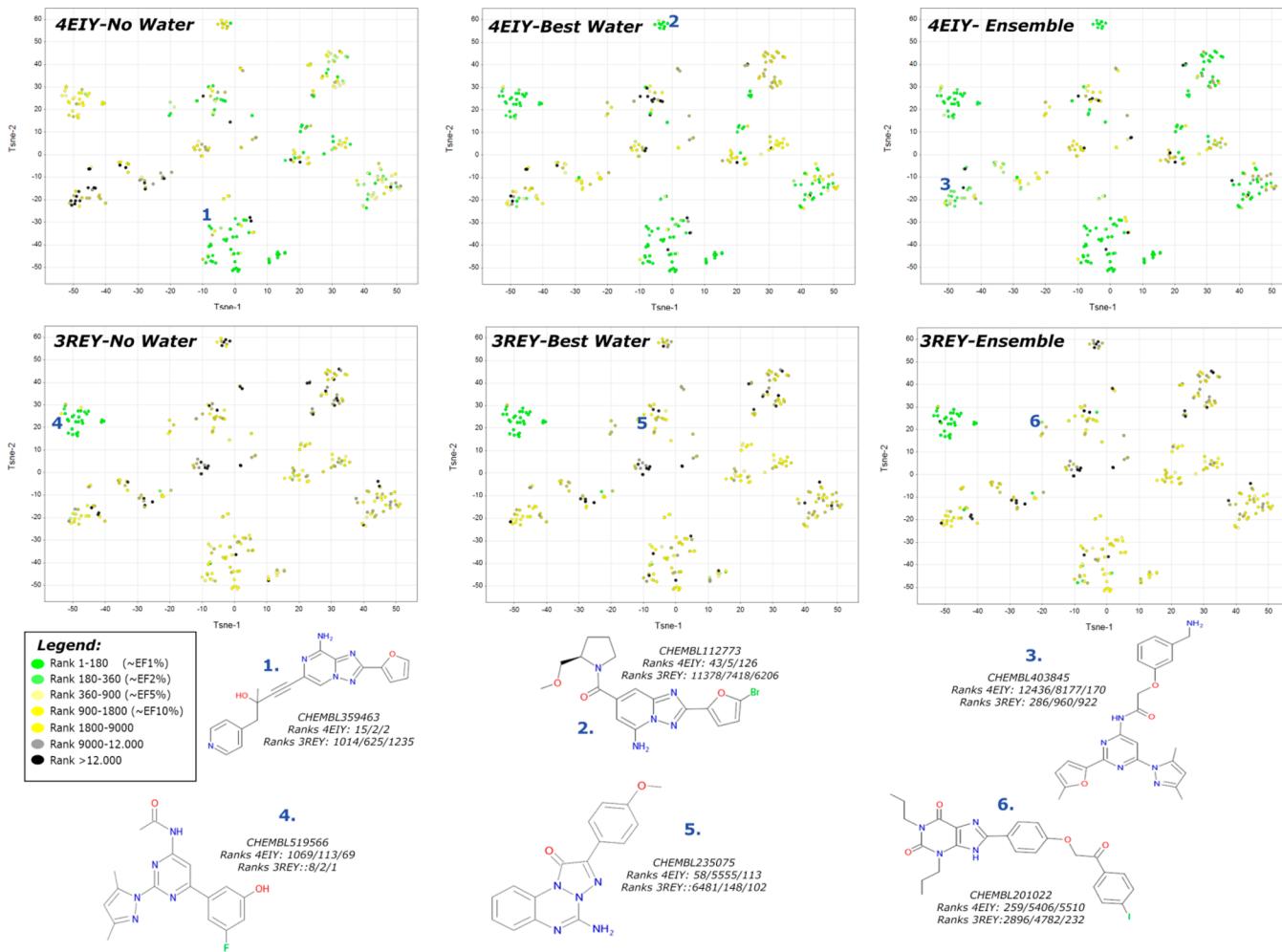


Figure 7. Chemical space of the ligands is represented in t-SNE plots.⁵⁰ Three plots are shown each for 4EIY and 3REY, for the crystal structure (left), optimal single water (middle), and the decision tree based ensemble (right). Ligands are colored based on their ranking in the virtual screen using a gradient from green (high rank) to black (low rank). Several examples of ligands retrieved together with their ranks are shown at the bottom of the picture. Note the largest spread of highly ranked ligands using the 4EIY structure and the decision tree (top right panel).

be superior to placing H-bond restraints on Asn253 (RMSD = 2.09 Å). Both waters 2521 and 2584 were found to be important for VS enrichments in our study as well. The influence of water molecules in the context of virtual screening has been studied previously using the 3EML structure and a small set of ligands/decoys by Katritch et al.²² In that study multiple waters with an increasing B-factor were considered (up to a total of 13), while in this study a total of 626 water orientations were assessed individually. The three waters *wa*, *wa4*, and *wa5* found to be important by Katritch in 3EML are equivalent to waters 2515, 2516, and 2531 in the 4EIY structure, respectively. Of these, the first two were not studied here, as they were further than 6 Å away from the ligand. The only coinciding water studied here, water 2531, was not found to contribute to enrichment. Differences in the docking method, crystal structure, optimization, and decoy set used could explain the difference found for this water. Additionally, Katritch et al. found that the presence of the three selected waters did not contribute significantly to the score of ligands, even though it increased the enrichment in the VS. In the present study it was observed that there are water molecules close to the crystallized ligand that increase enrichment and that these waters make (supposedly) favorable interactions with the ligands, since their replacement with excluded volume

restraints did not lead to the same increase in enrichment. The difference between the results from Katritch et al. and those obtained here suggests that it is not trivial to determine which waters should be selected a priori. In general, waters located in the solvent exposed area of the binding site did not enhance enrichments. Furthermore, our data demonstrates that selecting a water orientation at random is worse than including no water at all, demonstrating that it is important to select the correct water orientations (Table 1).

Differences in water coordination were observed between the crystallographic and computed water molecule positions. A total of 19 of the 22 crystal structure waters were reproduced by the computed water positions. Moreover, several computed water molecules were not present in the high-resolution crystal structure, suggesting that an explicit solvent MD-based method could complement the placement of water molecules even in high-resolution crystal structures. In addition, this method has been shown to be a good predictor of water positions in lower resolution structures.² Our study suggests that MD-based methods are able to substitute for high-resolution crystallography in placing water molecules for VS, and thus the method described here can be applied to any crystal structure where waters influence ligand-protein binding. While the WaterMap program is convenient for generating a cluster of water

positions from an MD, in principle any method that accurately maps the solvation of a binding site could be used as a starting point.^{52–58}

After benchmarking single water orientations, we used a Decision Tree (DT) approach to generate an ensemble of water molecules. The hypothesis behind this is that ligands are more likely to have favorable interactions with several water molecules than decoys would, while not all ligands interact with the same waters *per se*. This was indeed observed here, as most decoys that ever received a good docking score only scored well versus one water in the docking procedure. In fact, in our data set only 22 decoys were found to score above the threshold from the DT for two or more waters, while 144 active ligands were able to do so (Figure 4).

While DTs have been applied to numerous QSAR/multitarget studies,^{59,60} only a few studies have used DTs in ensemble docking. One example is presented in a study by Hritz et al. where the plasticity of CYP2D6 was described by including different protein conformations derived from MD.⁶¹ They addressed the different sites of metabolism for compounds by using a binary DT based on both different molecular descriptors and three distinct structural models. Based on this ensemble method, enrichments were improved considerably. In the current study we obtained significant improvements in enrichments, although we only sampled water orientations instead of protein conformations and did not include any compound descriptors. Hence, incorporation of ensembles of protein structures could further increase both the enrichment and diversity of hits captured. That said, combining results from the 4EIY and 3REY structures did not improve the results here. However, various studies have shown that inclusion of small ensembles of conformational states of the protein can enhance enrichment indeed.^{62–65} A few of them have also explicitly considered both water molecules and protein conformations.^{64,65} In summary, DTs present a powerful and straightforward approach for improving the predictive accuracy of docking methods.

The DT selects nodes (in this case individual waters) based on the Gini criterion, which represents the probability of sending ligand and decoy to the wrong “child”.⁶⁶ While the DT tries to maximize performance irrespectively of the chemical space covered, an interesting addition would be a DT where the splits are based on the diversity of hits covered. Despite the fact that diversity was not considered explicitly in the DT we demonstrated that the diversity of hits covered by the different leaves is higher than of any single leaf, and thus it is advisable to use multiple good performing leaves in a prospective virtual screen if diversity of hits is a goal. Since no normalization is needed for DT they are well suited for data fusion methods where information on multiple sources, such as waters, structures, or chemical descriptors, can be combined. By default the best individual water, 2521-S (WM), was selected as the first node. However, some of the lower nodes of the DT produced a lower enrichment than no water when considered individually; despite this, they increased the enrichment by preferentially removing decoys more than actives. Furthermore, there are 31 ligands that do not satisfy the criteria of the first node but do score well versus a related water orientation 2521-7, even though the two water molecule orientations are quite similar (RMSD 0.92 Å). These results demonstrate the sensitivity of docking results to water positions and highlight the importance of integrating results from different water orientations.

The results for a different crystal structure, bound to the XAC ligand (3REY), show that in principle enrichments can be boosted by including water orientations, independent of the structure. However, compared to 4EIY the enrichments for 3REY were not increased as much in absolute terms, but further improvements might be possible by including more water orientations (76 versus 628). Moreover, differences in the results can be explained by the lower enrichment of the structure without any water (0.19 versus 0.467). In fact, only the ensemble-based method of 3REY was able to rank XAC-like structures within the top 250 screening hits (Figure 7). Waters important for the enrichment of several of the chemotypes in our ligand set were not retrieved using the WaterMap of 3REY, and the waters that were found to be important were located on a hydration center which was not enriching for 4EIY, albeit with a different orientation (Supplementary Figure 2). As we have shown in the results, the ensemble of 3REY is specific for one certain chemotype of ligands. These results show that although higher enrichments are obtained when using an ensemble method, the actual virtual screening results are partially dependent on a proper choice of decoy set and crystal structure.

The external validation of the DT method with different published decoy sets further demonstrated the ability of this ensemble-based approach. A comparable enrichment for the decoy set of Katritch et al.²² was obtained. The difference between an EF1% of 63 and 78 can be explained by the optimized performance of the Katritch method against their small set of ligands and decoys (hence, for Katritch this was a training set), whereas it was used as a test set in the work presented here. The difference in EF1% is a consequence of having either 15 or 17 ligands in the top 20, suggesting both methods perform very well. Furthermore, considerably higher enrichments were achieved for the other two larger decoy sets (GDD and DUD-e).^{33,34} Although there is a partial overlap in the different ligands from the decoy sets, this can be seen as qualitative proof of consistency of the ensemble method across these different decoy sets. For all the previous studies on these decoy sets a lower resolution crystal structure was used,³⁰ and different docking programs were employed. The model used for DUD-e did not consider any waters but instead included an increased dipole moment for Asn253 to favor hydrogen bonding to this residue, as has been published and prospectively validated by Carlsson et al.²³ The model used by Gatica et al. included eight crystallographic waters (excluding these eight waters decreased enrichment by 4.5%). Katritch et al. selected three waters based on B-factors. In general the differences between approaches used in these studies demonstrate the challenges to explicitly account for waters.

CONCLUSIONS

In this study we developed a strategy to incorporate water molecules into a virtual screening protocol even if they are not present in a crystal structure. Since the MD-based water orientations led to comparable or higher enrichments than the crystal structure waters, our method should have broad applicability even when high-resolution crystal structures with accurate water positions are not available. Furthermore, first studying single water orientations and subsequently incorporating them into an ensemble-based approach such as the DT used here proved to be significantly better than screening protocols without water molecules. In this way we also

effectively dealt with incorporating multiple waters and their orientations without having to study different combinations of them, avoiding a potential combinatorial explosion. Replacing the waters from the DT with “excluded volume” restraints did not produce the same increase in enrichment, suggesting that these waters make favorable and specific interactions with ligands. Finally we have shown that with this method high enrichments are also achieved on other, published decoy sets. We propose this ensemble docking strategy as a general approach for VS on targets where a number of known binders exist and where water molecules are mediating protein–ligand interactions.

■ ASSOCIATED CONTENT

Supporting Information

Ligand and decoy sets used in this study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +31 (0)71 527 4651. Fax: +31 (0)71 527 4277. E-mail: ijzerman@lacdr.leidenuniv.nl.

Author Contributions

Eelke B. Lenselink and Thijs Beuming performed the calculations. Eelke B. Lenselink wrote the manuscript. All authors discussed, contributed to, and revised the manuscript.

Notes

The authors declare the following competing financial interest(s): Dr Beuming and Dr Sherman are employees of Schrodinger Inc. In the present study Schrodinger software was used, including WaterMap. Dr van Vlijmen is also an employee of Janssen Research&Development.

■ ACKNOWLEDGMENTS

Adriaan P. IJzerman and Eelke B. Lenselink thank the Dutch Research Council (NWO) for financial support (NWO-TOP #714.011.001).

■ REFERENCES

- (1) Snyder, P. W.; Mecinović, J.; Moustakas, D. T.; Thomas, S. W.; Harder, M.; Mack, E. T.; Lockett, M. R.; Héroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 17889–17894.
- (2) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. Contribution of explicit solvent effects to the binding affinity of small-molecule inhibitors in blood coagulation factor serine proteases. *ChemMedChem.* **2011**, *6*, 1049–1066.
- (3) Lemieux, R. U. How water provides the impetus for molecular recognition in aqueous solution. *Acc. Chem. Res.* **1996**, *29*, 373–380.
- (4) Babine, R. E.; Bender, S. L. Molecular recognition of protein–ligand complexes: Applications to drug design. *Chem. Rev.* **1997**, *97*, 1359–1472.
- (5) Sharp, K. A.; Honig, B. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **1990**, *94*, 7684–7692.
- (6) Nicholls, A.; Honig, B. A Rapid Finite-Difference Algorithm, Utilizing Successive over-Relaxation to Solve the Poisson-Boltzmann Equation. *J. Comput. Chem.* **1991**, *12*, 435–445.
- (7) Ghosh, A.; Rapp, C. S.; Friesner, R. A. Generalized born model based on a surface integral formulation. *J. Phys. Chem. B* **1998**, *102*, 10983–10990.
- (8) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *J. Comput. Chem.* **2002**, *23*, 517–529.
- (9) 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. U.S.A.* **2007**, *104*, 808–813.
- (10) 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.
- (11) Ladbury, J. E. Just add water! The effect of water on the specificity of protein–ligand binding sites and its potential application to drug design. *Chem. Biol.* **1996**, *3*, 973–980.
- (12) Cooper, A. Heat capacity effects in protein folding and ligand binding: a re-evaluation of the role of water in biomolecular thermodynamics. *Biophys. Chem.* **2005**, *115*, 89–97.
- (13) Woo, H. J.; Roux, B. Calculation of absolute protein–ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825–6830.
- (14) Lu, Y. P.; Wang, R. X.; Yang, C. Y.; Wang, S. M. Analysis of ligand-bound water molecules in high-resolution crystal structures of protein–ligand complexes. *J. Chem. Inf. Model.* **2007**, *47*, 668–675.
- (15) Roberts, B. C.; Mancera, R. L. Ligand–protein docking with water molecules. *J. Chem. Inf. Model.* **2008**, *48*, 397–408.
- (16) Katritch, V.; Cherezov, V.; Stevens, R. C. Structure–function of the G protein-coupled receptor superfamily. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53*, 531–556.
- (17) Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; IJzerman, A. P.; Cherezov, V.; Stevens, R. C. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337*, 232–236.
- (18) Wang, C.; Wu, H.; Katritch, V.; Han, G. W.; Huang, X. P.; Liu, W.; Siu, F. Y.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the human smoothed receptor bound to an antitumour agent. *Nature* **2013**, *497*, 338–343.
- (19) Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X. P.; Carroll, F. I.; Mascarella, S. W.; Westkaemper, R. B.; Mosier, P. D.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **2012**, *485*, 327–332.
- (20) Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S. Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **2012**, *485*, 321–326.
- (21) Carugo, O.; Bordo, D. How many water molecules can be detected by protein crystallography? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 479–483.
- (22) Katritch, V.; Jaakola, V. P.; Lane, J. R.; Lin, J.; IJzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J. Med. Chem.* **2010**, *53*, 1799–1809.
- (23) Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-based discovery of A2A adenosine receptor ligands. *J. Med. Chem.* **2010**, *53*, 3748–3755.
- (24) Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5391–5395.
- (25) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6843–6848.
- (26) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **2011**, *7*, 769–778.
- (27) Mysinger, M. M.; Weiss, D. R.; Ziarek, J. J.; Gravel, S.; Doak, A. K.; Karpiak, J.; Heveker, N.; Shoichet, B. K.; Volkman, B. F. Structure-based ligand discovery for the protein–protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 5517–5522.
- (28) Kruse, A. C.; Weiss, D. R.; Rossi, M.; Hu, J.; Hu, K.; Eitel, K.; Gmeiner, P.; Wess, J.; Kobilka, B. K.; Shoichet, B. K. Muscarinic

- receptors as model targets and antitargets for structure-based ligand discovery. *Mol. Pharmacol.* **2013**, *84*, 528–540.
- (29) de Graaf, C.; Kooistra, A. J.; Vischer, H. F.; Katritch, V.; Kuijer, M.; Shiroishi, M.; Iwata, S.; Shimamura, T.; Stevens, R. C.; de Esch, I. J.; Leurs, R. Crystal structure-based virtual screening for fragment-like ligands of the human histamine H(1) receptor. *J. Med. Chem.* **2011**, *54*, 8195–8206.
- (30) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; Lane, J. R.; IJzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217.
- (31) Schrödinger Release 2013-1: Maestro, version 9.4; Schrödinger, LLC: New York, NY, 2013.
- (32) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221–234.
- (33) Gatica, E. A.; Cavasotto, C. N. Ligand and decoy sets for docking to G protein-coupled receptors. *J. Chem. Inf. Model.* **2012**, *52*, 1–6.
- (34) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594.
- (35) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (36) Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704–721.
- (37) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (38) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (39) Truchon, J. F.; Bayly, C. I. Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. *J. Chem. Inf. Model.* **2007**, *47*, 488–508.
- (40) The PyMOL Molecular Graphics System, Version 1.5.0.4; Schrödinger, LLC.
- (41) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100–D1107.
- (42) Shelley, J. C.; Cholletti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681–691.
- (43) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591–604.
- (44) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.
- (45) Kelly, J. L. A new interpretation of information rate. *Inf. Theory, IRE Trans.* **1956**, *2*, 185–189.
- (46) Doré, A. S.; Robertson, N.; Errey, J. C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F. Structure of the Adenosine A2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure* **2011**, *19*, 1283–1293.
- (47) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (48) Bender, A.; Mussa, H. Y.; Glen, R. C.; Reiling, S. Molecular similarity searching using atom environments, information-based feature selection, and a naive Bayesian classifier. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 170–178.
- (49) Schrödinger Release 2014-1: Canvas, version 1.9; Schrödinger, LLC: New York, NY, 2014.
- (50) Van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
- (51) Ivanov, A. A.; Barak, D.; Jacobson, K. A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A2A adenosine receptor crystallographic structure. *J. Med. Chem.* **2009**, *52*, 3284–3292.
- (52) Truchon, J.-F.; Pettitt, B. M.; Labute, P. A Cavity Corrected 3D-RISM Functional for Accurate Solvation Free Energies. *J. Chem. Theory Comput.* **2014**, *10*, 934–941.
- (53) Bren, U.; Janežič, D. Individual degrees of freedom and the solvation properties of water. *J. Chem. Phys.* **2012**, *137*, 024108.
- (54) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Prediction of the water content in protein binding sites. *J. Phys. Chem. B* **2009**, *113*, 13337–13346.
- (55) Huggins, D. J.; Marsh, M.; Payne, M. C. Thermodynamic Properties of Water Molecules at a Protein–Protein Interaction Surface. *J. Chem. Theory Comput.* **2011**, *7*, 3514–3522.
- (56) Nguyen, C. N.; Young, T. K.; Gilson, M. K. Grid inhomogeneous solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit [7] uril. *J. Chem. Phys.* **2012**, *137*, 044101/1–044101/17.
- (57) 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.
- (58) Cui, G.; Swails, J. M.; Manas, E. S. SPAM: a simple approach for profiling bound water molecules. *J. Chem. Theory Comput.* **2013**, *9*, 5539–5549.
- (59) Dudek, A. Z.; Arodz, T.; Galvez, J. Computational methods in developing quantitative structure-activity relationships (QSAR): a review. *Comb. Chem. High Throughput Screening* **2006**, *9*, 213–228.
- (60) van Westen, G. J. P.; Wegner, J. K.; IJzerman, A. P.; van Vlijmen, H. W. T.; Bender, A. Proteochemometric modeling as a tool to design selective compounds and for extrapolating to novel targets. *MedChemComm* **2011**, *2*, 16–30.
- (61) Hritz, J.; De Ruiter, A.; Oostenbrink, C. Impact of Plasticity and Flexibility on Docking Results for Cytochrome P450 2D6: A Combined Approach of Molecular Dynamics and Ligand Docking. *J. Med. Chem.* **2008**, *51*, 7469–7477.
- (62) Osguthorpe, D. J.; Sherman, W.; Hagler, A. T. Exploring protein flexibility: incorporating structural ensembles from crystal structures and simulation into virtual screening protocols. *J. Phys. Chem. B* **2012**, *116*, 6952–6959.
- (63) Osguthorpe, D. J.; Sherman, W.; Hagler, A. T. Generation of receptor structural ensembles for virtual screening using binding site shape analysis and clustering. *Chem. Biol. Drug Des.* **2012**, *80*, 182–193.
- (64) 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.
- (65) Barreca, M. L.; Iraci, N.; Manfroni, G.; Gaetani, R.; Guercini, C.; Sabatini, S.; Tabarrini, O.; Cecchetti, V. Accounting for Target Flexibility and Water Molecules by Docking to Ensembles of Target Structures: The HCV NSSB Palm Site I Inhibitors Case Study. *J. Chem. Inf. Model.* **2013**, *53*, 481–497.
- (66) Breiman, L.; Friedman, J.; Stone, C. J.; Olshen, R. A. *Classification and regression trees*; CRC Press: Boca Raton, FL, 1984.
- (67) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Köster, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. *KNIME: The Konstanz information miner*; Springer: Berlin, Germany, 2008.