A consistent description of HYdrogen bond and DEhydration energies in protein–ligand complexes: methods behind the HYDE scoring function

Nadine Schneider · Gudrun Lange ·
Sally Hindle · Robert Klein · Matthias Rarey

Abstract The estimation of free energy of binding is a key problem in structure-based design. We developed the scoring function HYDE based on a consistent description of HYdrogen bond and DEhydration energies in protein–ligand complexes. HYDE is applicable to all types of protein targets since it is not calibrated on experimental binding affinity data or protein–ligand complexes. The comprehensible atom-based score of HYDE is visualized by applying a very intuitive coloring scheme, thereby facilitating the analysis of protein–ligand complexes in the lead optimization process. In this paper, we have revised several aspects of the former version of HYDE which was described in detail previously. The revised HYDE version was already validated in large-scale redocking and screening experiments which were performed in the course of the Docking and Scoring Symposium at 241st ACS National Meeting. In this study, we additionally evaluate the ability of the revised HYDE version to predict binding affinities. On the PDBbind 2007 coreset, HYDE achieves a correlation coefficient of 0.62 between the experimental binding constants and the predicted binding energy, performing second best on this dataset compared to 17 other well-established scoring functions. Further, we show that the performance of HYDE in large-scale redocking and virtual screening experiments on the Astex diverse set and the DUD dataset respectively, is comparable to the best methods in this field.

Introduction

The application of computational methods has become standard during the drug discovery process [1]. Virtual screening, which aims to find new bioactive agents for a certain protein target, is one of the first steps in this process. When a three-dimensional structure of the target protein is available, molecular docking is used to predict potential binding modes of several hundreds of thousands of compounds and to estimate their binding strength. In the optimal case, compounds suggested by the docking tool can then be experimentally validated and found to exhibit strong binding constants and the predicted binding mode [2]. The scoring functions integrated into these docking tools have to successfully carry out three tasks to achieve this goal: Firstly, the potential bioactive conformations of the compounds must be selected from a pool of docking poses. Secondly, from the hundreds of thousands of compounds tested during virtual screening, non-binders must be discriminated from true binders. Finally, the binding affinity of a compound must be correctly predicted. To date, the estimation of the free energy of binding is still a

N. Schneider · M. Rarey (✉)
Center for Bioinformatics, University of Hamburg,
Bundesstr. 43, 20146 Hamburg, Germany
e-mail: rarey@zbh.uni-hamburg.de

G. Lange · R. Klein
Bayer CropScience AG, Industriepark Hoechst, G836,
65926 Frankfurt am Main, Germany

S. Hindle
BioSolveIT GmbH, An Der Ziegelei 79, 53757 St. Augustin,
Germany

largely unsolved problem [3]. At the same time, it is probably the most crucial issue that needs to be addressed for all kinds of structure-based design applications.

A lot of different scoring functions have been developed over the last 20 years, relying on very different approaches to solve these problems [4, 5]. They model the interactions, energies or preferred contacts between the protein and its ligand. The first scoring functions mostly consider only favorable contributions to the binding energy [6, 7]. Today it is known that unfavorable contributions also comprise an important part of the free energy of binding [8, 9]. However, quantifying these contributions remains problematic: Most scoring functions are calibrated on experimental binding affinities and/or protein–ligand complexes and herein favorable contributions predominate. Attempts to model these unfavorable contributions have been made using different approaches, by including, for example, artificial "negative data" in the parameterization [10], new terms to model the desolvation penalty [11–14] or other parameters, like logP values, to try and quantify these contributions [15, 16]. In this paper we describe how we model unfavorable contributions to the binding energy in our recently developed scoring function HYDE [17–19].

The first version of the HYDE function was developed by Reulecke et al. [17]. HYDE is based on the estimation of HYdrogen bond and DEhydration energies emerging during protein–ligand binding. Using only these two major contributions of the binding energy, we are able to consistently describe hydrogen bonding and the hydrophobic effect as well as the unfavorable contribution of hydrophilic dehydration. In this study, we revise several aspects of the HYDE function. We retain the basic concept of HYDE [17, 18], while the calculation of the binding energy contribution from polar groups changes substantially. We also re-parameterized the logP increments using a reduced set of atom types. Furthermore, a completely new and faster algorithm is used to calculate molecular surface components. Additional terms concerning the arrangement of waters around both molecules before binding are introduced. Finally, the HYDE function is integrated into an optimization procedure to allow a more accurate prediction of the structure of protein–ligand complexes. All these changes are described in detail in the Methods section. In the Results section, we summarize our results of the revised HYDE function we have obtained in a previous validation study [19]. Additionally, we evaluated the revised HYDE function in the prediction of binding affinities on congeneric compound series and the PDBbind2007 coreset. The results are critically discussed to demonstrate the benefits and the drawbacks of the HYDE scoring function and we compare our results with that of others in this field. Finally, we conclude and give an outlook on the future trend concerning the development of the HYDE scoring function.

Methods

The HYDE scoring function relies on an intuitive concept: Both molecules—protein and ligand—are solvated in aqueous solution in the unbound state. During the binding process, the water molecules around the ligand are stripped off and those in the binding pocket of the protein are squeezed out by the ligand. The hydrogen bonds of the protein and the ligand to water molecules are broken, which leads to an unfavorable enthalpic contribution, even though the water molecules are released to bulk. New hydrogen bonds established between the protein and ligand may counterbalance this energy loss. In addition, hydrophobic moieties of ligand or protein in contact with water molecules lead to a discontinuity in the water hydrogen bond network and, therefore, to an unfavorable energy. The removal of these water molecules from the hydrophobic surfaces and their release to the bulk water induces a gain in energy, the so-called hydrophobic effect [18]. We propose that these processes represent the main contributions to the binding energy and exactly these contributions—hydrogen bonding, the hydrophobic effect and dehydration—are modeled in the HYDE scoring function:

$$\Delta G_{HYDE} = \sum_{atoms\ i} \left(\Delta G_{dehydration}^i + \Delta G_{H-bonds}^i \right) \quad (1)$$

We calculate the change in dehydration ($\Delta G_{dehydration}$) and hydrogen bond ($\Delta G_{H-bonds}$) energy for every atom i in the protein–ligand interface.

Dehydration energy calculation

Whereas the dehydration (desolvation) of hydrophobic atoms contributes favorably to the overall binding energy, the dehydration of hydrophilic groups is foremost energetically unfavorable. In the revised HYDE function, we have developed two separate terms to evaluate the dehydration energy for hydrophobic and hydrophilic atoms respectively:

$$\Delta G_{dehydration}^{i,hydrophobic} = -2.3RT \cdot p\log P^i \cdot \left(acc_{unbound}^i - acc_{bound}^i \right) \quad (2)$$

$$\Delta G_{dehydration}^{i,hydrophilic} = -2.3RT \cdot p\log P^i \cdot f_{bur}^i \cdot f_{water}^i \\ \cdot \sum_{H-bond\ functions\ j} w^j \cdot p_{dehyd}^j \quad (3)$$

Hydrophobic atoms are still treated similarly to the way they were treated in the first version of the HYDE scoring function. We calculate the change in solvent accessible surface Δacc^i [Å^2] of an atom i and multiply it by its logP increment $p\log P^i$ [J/Å^2] to estimate its dehydration energy. We have completely changed the calculation concerning hydrophilic atoms in the revised HYDE function.

Beforehand, the dehydration was estimated according to Eq. 2 for all atoms using a weighted solvent accessible surface area (WSAS) [17]. This meant that for hydrophilic groups, only the parts of the surface area which were located in the preferred direction of a hydrogen bond contributed to the WSAS. In the revised version of HYDE function we have replaced the WSAS by the *molecular or Connolly surface area* [20–22]. The accessibility of hydrophilic atoms is now assessed by testing whether there is sufficient space to accommodate a water molecule in the preferred direction of a hydrogen bond. A similar approach was recently published in the revision of the Autodock force field function [23]. More precisely, we calculate the probability of dehydration p_{dehyd}^j of each hydrogen bond function j (= hydrogen bond donor or acceptor) of a hydrophilic atom. Deeply buried hydrogen bond functions, as well as functions involved in hydrogen bonding, are given a dehydration probability of $p_{dehyd}^j = 1$. Otherwise, the dehydration probability linearly decreases if space for at least one half of the volume of a water molecule is available at the preferred direction of a hydrogen bond. Details of our algorithm concerning the calculation of the surface and accessibility are described below.

Additionally, we introduced weights w^j for multiple hydrogen bonds which can be formed by a single hydrophilic atom. These weights reflect an important finding of our logP study [24]. Atoms which are able to form several hydrogen bonds (e.g. primary amines) were compared to atoms which only can establish one hydrogen bond (e.g. tertiary amines). The results showed that the same contribution to the logP value was made, indicating that the ability of an atom to form multiple hydrogen bonds does not induce a higher hydrophilicity. For this reason, the contributions of hydrogen bond donors/acceptors of a single atom to the dehydration or hydrogen bond energy are weighted according to the following scheme: The geometrically best hydrogen bond gets a weight of 100 %. The weight of the second best is decreased to 20 % and a third hydrogen bond contributes with 10 %. Any further hydrogen bonds have no contribution at all. In the case that donors/acceptors form no hydrogen bonds, their weights are sorted depending on their dehydration probability (from low to high probability).

The factor f_{bur}^i is a scaling factor which takes the buriedness of a hydrophilic group in the unbound state into account. For hydrophilic ligand atoms this factor is set to 1. In the protein, this factor is scaled according to whether the hydrophilic atom is highly exposed or not. The value is calculated based on an approach developed by Stahl [25].

Since HYDE only considers water molecules implicitly, we introduced a correction factor f_{water}^i which accounts for the local arrangement of water in proximity to the hydrogen bond function. This factor is calculated for each polar atom i as follows:

$$f_{water}^i = \sum_{H-bond\ functions\ j\ of\ atom\ i} water_{overlap}^j \cdot water_{interaction}^j$$

(4)

The two factors $water_{overlap}$ and $water_{interaction}$ aim to describe the quality of solvation before binding and its influence on the extent of the unfavorable dehydration energy. The underlying concept of the HYDE function [17, 18] uses an ideal model: each hydrogen bond function is assumed to be saturated by a single water molecule in the unbound state. This scenario is true for isolated hydrogen bond functions. However, in a binding pocket and also for ligands with many adjacent polar groups, the local arrangement of the water molecules is subjected to restrictions. This may lead to a lower dehydration cost of these hydrogen bond functions since they are not ideally satisfied in the solvated state. Our assumption is confirmed by the observation that the logP value of molecules does not linearly decrease with the number of attached polar groups.

water_overlap: overlapping waters.

The *water_overlap* term gives an estimate for the number of water molecules which can be arranged around a hydrophilic atom allowing the dehydration cost to be shared between groups. First, water molecules are placed at the ideal position of a hydrogen bonding partner at the hydrogen bond functions of the polar groups. This is done for the unbound ligand and the empty binding site respectively. For each water molecule i, the overlap with all surrounding water molecules j is calculated:

$$water_{overlap}^i = 1 - \frac{\sum_{surrounding\ waters\ j} \frac{1}{2} \cdot overlap\ volume\ (water^i, water^j)}{volume\ (water)}$$

(5)

Figure 1a shows a schematic of a small hydrophilic pocket where three polar groups interact with the same water molecule. In contrast, Fig. 1b shows the overlap of three ideally placed water molecules in the active site. We calculate the overlap volume of water molecule i with the water molecules j and k respectively (Eq. 5) (Fig. 1c, d). The sum of these volumes is normalized by the volume of a water molecule (radius = 1.4 Å). In this case, the *water_overlap* term of water i would amount to about a third. Consequently the dehydration cost of the polar group of which water molecule i originates is reduced by one-third.
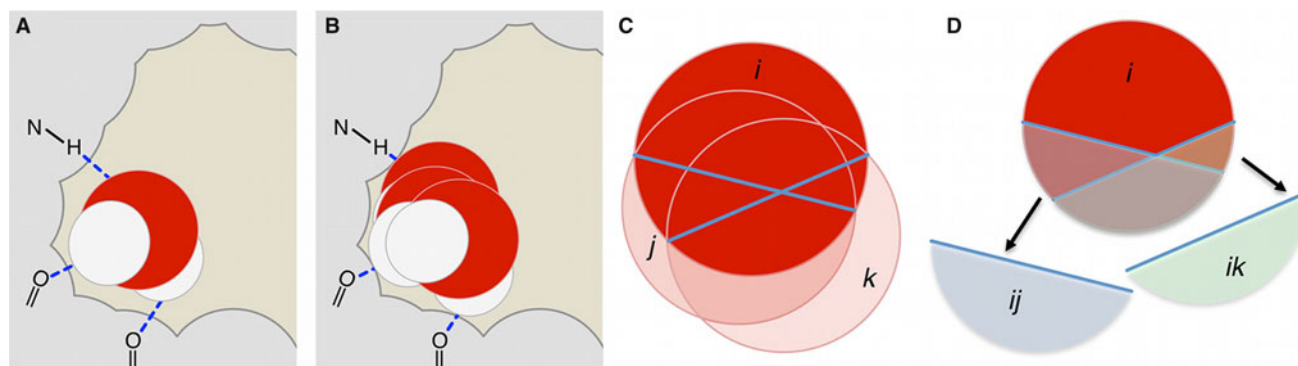
Fig. 1 Lowering dehydration cost of hydrophilic pockets by investigating the number of displaced water molecules. **a** Small hydrophilic pocket, three polar atoms interact with the same water molecule. **b** Overlap of three ideally placed water molecules in the small hydrophilic pocket. **c** Overlap of water molecule i with water molecules j and k. **d** Overlap volume of water molecule i and water molecules j and k

Water$_{interaction}$: conserved waters

The *water$_{interaction}$* factor is complementary to the *water$_{overlap}$* term. It provides an estimate for the saturation of a water molecule interacting with a certain polar group. Depending on the number of hydrogen bonds a water is able to form, the dehydration cost can differ substantially. In most cases, water molecules are highly conserved and displacing these is enthalpically unfavorable. On the contrary, if a water molecule is situated in a small hydrophobic pocket and is only able to form one interaction with a polar group, the dehydration cost at the polar group might be overestimated due to the entropically and enthalpically unfavorable water molecule. For calculating the *water$_{interaction}$* factor a water molecule is placed at the ideal position of a hydrogen bonding partner at each hydrogen bond function. This water molecule is rotated to find the best possible interaction network and to reduce the number of unsatisfied hydrogen bond functions of the water molecule itself. Eventually, the *water$_{interaction}$* factor is determined by a combination of the number of interactions and the number of unsatisfied hydrogen bond functions of the water.

Hydrogen bond energy calculation

The hydrogen bond energy in HYDE takes the following form:

$$\Delta G^i_{\mathrm{H-bond}} = \frac{2.3RT}{F_{sat}(T)} \cdot p \log P^i \cdot f^i_{bur} \sum_{\mathrm{H-bonds}\ j} w^j \cdot f^j_{dev} \qquad (6)$$

To express the complementary nature of the hydrogen bonding and dehydration term, a similar functional form is used. In HYDE, the hydrogen bond energy contribution arises from the fact that not all hydrogen bonds in the hydrogen bond network of bulk water are perfectly realized, thus the energy needed to disrupt these hydrogen bonds is lower than that for an ideal hydrogen bond [18]. We integrate this phenomenon into HYDE by using the saturation factor F_{sat} (see Eq. 6). This factor describes the incomplete saturation of the water hydrogen bond network at a certain temperature. At a temperature of 273 K the saturation factor is $F_{sat}(273\ \mathrm{K}) = 0.89$, while at 310 K it is only about $F_{sat}(310\ \mathrm{K}) = 0.84$ [18]. We use T = 298 K resulting in $F_{sat}(298\ \mathrm{K}) = 0.85$ for estimating the saturation energy for a protein–ligand complex, since most experimental affinity values are measured at room temperature. Consequently, the energy gain of an intermolecular hydrogen bond in HYDE is roughly 17 % (= $1/F_{sat}(298\ \mathrm{K})$) higher than the dehydration cost associated with both of the hydrogen bond functions. The geometrical quality of a hydrogen bond j is accounted for with the factor f^i_{dev}. It is well known that the energy of a hydrogen bond diminishes considerably with the deviation from ideal hydrogen bonding geometry in terms of both the angles and distance between donor and acceptor. In HYDE, the preferred hydrogen bond directions of different atom types are modeled as sections of spherical surfaces. These interaction surfaces represent the optimal location for potential interaction partners and are based on the FlexX interaction model [7] which has been further developed in its current implementation in the LeadIT software package [26]. Hydrogen bonds that deviate from the perfect geometry are linearly scaled until a certain threshold at which HYDE considers the hydrogen bond to no longer be made. The other two factors w^j and f^i_{bur} were already introduced in the dehydration energy calculation of hydrophilic atoms (see Eq. 3).

PlogP re-parameterization

The atom-based logP (plogP) increments used in HYDE were derived from experimental logP values taken from the PHYSPROP database [27]. Nearly all chosen compounds

and experimental values stem from the collection of Hansch and Leo [28–30] meaning most values were determined in the same laboratory and are therefore consistent. Compared to the compounds that were used to derive plogP values for the former version of HYDE [17], we selected only compounds with one heteroatom per molecule. The reason for using only these simple molecules was to avoid proximity effects which are known to influence the logP value of a compound [31]. We considered N, O, S, F, Cl, Br, and I as heteroatoms resulting in a dataset of 445 molecules.

Using these molecules we performed multiple linear regression (MLR) to obtain the plogP increments. We reduced the number of atom types to eight for our re-parameterization. The new atom types resulted from an intensive logP analysis we recently accomplished [24]. Only nitrogen and oxygen atoms were considered as hydrogen bond acceptors or donors. All other atoms—carbon, sulfur, and halogens—were treated as hydrophobic. Overall, a correlation coefficient of $R^2 = 0.94$ is achieved for the training dataset.

Surface calculation and accessibility estimation

To determine the dehydration energy of a protein–ligand complex we calculate the degree of buriedness of each atom in the final complex geometry. The change of accessibility of each atom is estimated as the change of its molecular surface area induced by complex formation. In contrast to many other methods/functions, instead of the solvent accessible surface (SAS) [32] of a molecule, we actually calculate the molecular surface or Connolly surface [20–22] (Fig. 2a orange line) to estimate the change in accessibility. Firstly though, we do use the SAS for generating a 3D surface net around both molecules–the protein's binding pocket and the ligand—and assign the underlying molecular surface area increment to each surface node of the net.

The surface net is generated as follows: Firstly, the molecule's SAS [20, 32, 33] is generated using standard van der Waals radii [34] and a surface sphere radius of 1.4 Å (Fig. 2a blue dots). In order to attain a uniform distribution of surface nodes, a 2-stage icosahedron subdivision for each atom is generated which results in 162 surface nodes per atom. These are then scaled to lie on the SAS of the molecule. Hydrogen atoms are only considered implicitly by increasing the vdW radii of heavy atoms by 0.1 per hydrogen atom. All surface nodes buried by neighboring atoms were eliminated. Additionally, we generate surface nodes for representing the re-entrant regions of the molecular surface (see Fig. 2b). The actual underlying molecular surface increment in $Å^2$ is calculated for each surface node. Summing up the underlying surface

areas annotated at each surface node gives the total surface area of a molecule.

We generate a surface net for the ligand and for the binding pocket of the complex. We use the same conformation of the ligand in the unbound and bound state and the protein is treated as rigid. Hence, the change in solvent accessible area for both molecules is only that induced by complex formation.

For hydrophobic atoms, the change of accessibility is calculated by adding the surface increments of surface nodes covered by the heavy atoms of the other molecule in the final complex geometry. Using this value, the dehydration energy of a hydrophobic atom can be estimated in HYDE by using Eq. 2. Figure 3b shows the covered surface nodes of the ligand after complex formation while Fig. 3c shows those of the binding pocket.

To calculate the accessibility of a hydrophilic atom, a hypothetical water molecule is placed at the optimal location for a hydrogen bonding partner (see Fig. 3a). In the bound state, the overlap of this water molecule with all surrounding heavy atoms of the other molecule is calculated. Figure 3b sketches the overlap of two hypothetical waters placed at the ligand's carbonyl group with the binding pocket. If the overlap constitutes more than one half the volume of a water molecule, this hydrogen bond function j is treated as dehydrated ($p_{dehyd}^j = 1$). Otherwise the dehydration probability p_{dehyd}^j is scaled down linearly with respect to the overlap.

Energy estimation for metals ions

In the HYDE scoring function interactions made between metal ions embedded in the binding pocket and ligand metal acceptor atoms are considered as follows:

$$\Delta G_{HYDE}^{metal} = \sum_{metal\,ions\,i} \Delta G_{interaction}^i + \Delta G_{dehydration}^i \tag{7}$$

$$\Delta G_{interaction}^i = \varepsilon_{interact}^{metal} \cdot \sum_{interactions\,j} f_{dev}^j \tag{8}$$

$$\Delta G_{dehydration}^i = \varepsilon_{dehyd}^{metal} \cdot \sum_{coordination\,sites\,j} p_{dehyd}^j \tag{9}$$

Since no reliable logP values are available for metal ions, we investigated the metallo-enzyme complexes contained in the Astex diverse set [35] to empirically derive an energy increment for the metal interaction energy of $\varepsilon_{interact}^{metal} = -20\,kJ/mol$ and $\varepsilon_{dehyd}^{metal} = 10\,kJ/mol$ for the metal dehydration energy. A full coordination of metal ions is crucial for a strong binding affinity. Therefore, we explicitly check for saturation of each coordination site of the metal which is not occupied by a receptor atom. Unsatisfied metal coordination sites, including those covered by apolar atoms, are penalized in HYDE in a
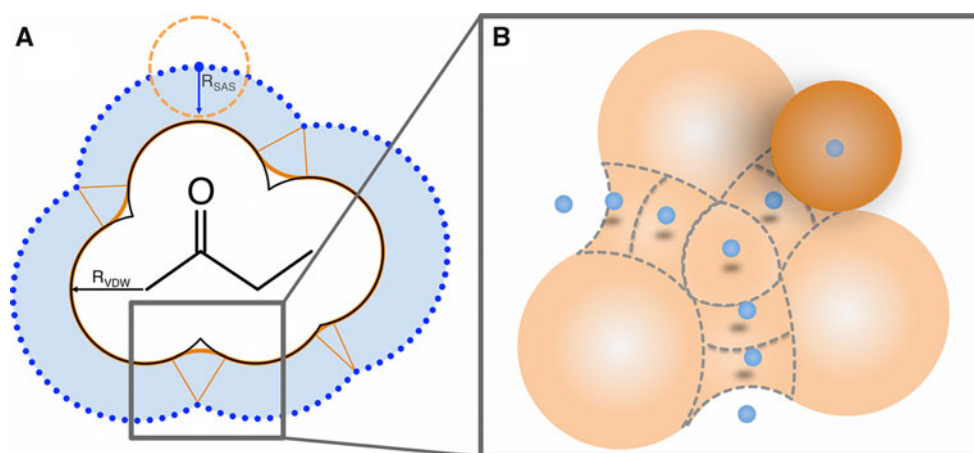
Fig. 2 Surface net generation. **a** *Blue dots* surface net lying on the SAS of the molecule. *Orange line* molecular surface or Connolly surface of the molecule. **b** Detail: Generation of surface nodes for re-entrant regions of the molecular surface. Surface increments are defined by *grey dotted lines*
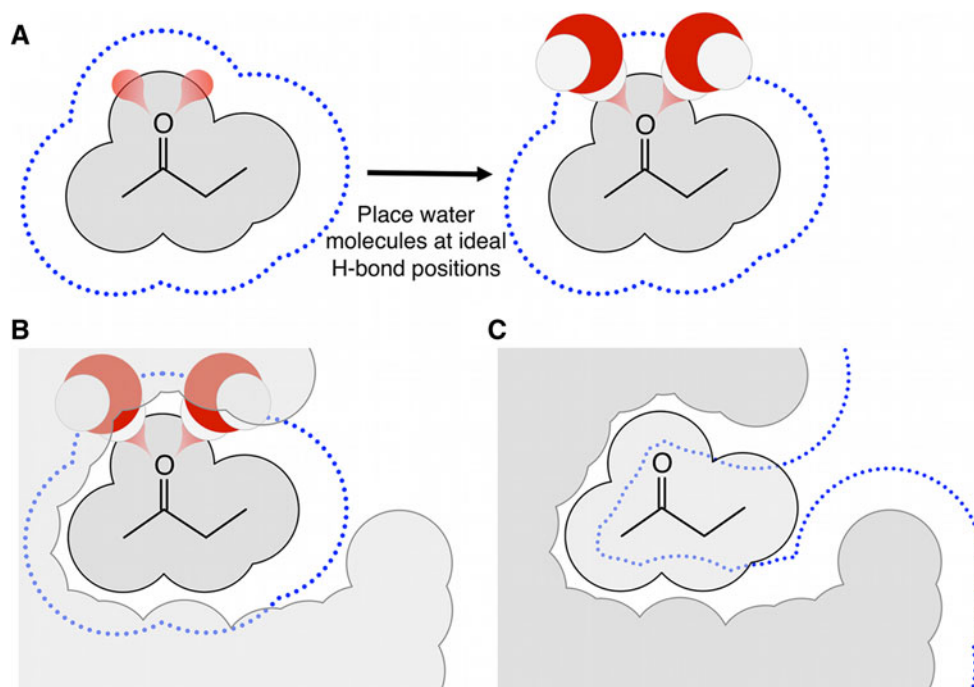

Fig. 3 Accessibility calculation. **a** Generate surface net of the molecule (*blue dots*). For hydrophilic atoms, place water molecules at the preferred hydrogen bonding directions. **b** Change of accessibility of ligand atoms induced by the binding pocket. **c** Change of accessibility of binding pocket atoms induced by the ligand

similar way to unsatisfied hydrogen bond functions. We consider nitrogen, oxygen and sulfur atoms as ligand metal acceptors. They are treated the same way as in hydrogen bonding interactions (see Eq. 5).

The metal interaction geometry is based on the coordination geometry of the metal [36]. An interaction between a metal ion and a ligand metal acceptor is modeled by using overlapping interaction surfaces as already described for hydrogen bonds. Metal interactions that deviate from a perfect geometry are linearly scaled until a certain threshold at which the interaction is no longer considered. Hence, we include a geometrical quality factor f_{dev}^{j} in the estimation of the metal interaction energy (Eq. 8). The calculation of this factor is analogous to the calculation of f_{dev}^{j} for hydrogen bonds (see Eq. 6). To estimate the dehydration energy of a metal ion, the dehydration probability p_{dehyd}^{j} is calculated for each coordination site j of the metal that is not occupied by a receptor atom (Eq. 9).

Hydrogen bond network and geometry optimization

In the HYDE function, no terms are included to assess the steric arrangement of a protein–ligand complex or the strain energy of the ligand. Furthermore, the HYDE scoring functions only tolerates small deviations from ideal hydrogen bond geometries. To ensure a protein–ligand complex is properly prepared for scoring with HYDE, two optimization procedures can be employed prior to scoring. First, the hydrogen bond network within the protein and between the protein and ligand can be optimized using ProToss [37] and the stringent definition of hydrogen bond geometries in HYDE. Second, an optimization/minimization of the ligand in the active site can be carried out which takes clashes between the ligand and the protein as well as within the ligand, plus the relaxation of the ligand strain energy, into consideration. This procedure uses a numerical optimization algorithm for a local optimization or, alternatively, a stochastic Monte-Carlo optimization strategy with simulated annealing for searching a global optimum. Due to the high computational cost, an approximate HYDE function is used in both optimization strategies [19]. In addition to approximate terms of the HYDE function the optimization uses a (12,6)-Lennard-Jones term and an estimate of the torsional strain energy of the ligand which is taken from the FlexX approach [7]. It was found to be important to consider these terms in the optimization to eliminate unfavorable complex geometries, as they are not part of the HYDE energy estimate.

Visualization of the HYDE score: HYDE coloring scheme

To facilitate the easy detection of favorable and unfavorable contributions to binding affinity, we use the intuitive atom-based HYDE coloring scheme [17]—available in the HYDE module of the LeadIT software [26]. A coloring scale from dark green for the most favorable score contributions through white for neutral to red for unfavorable contributions is applied to the atoms. For example, atoms involved in hydrogen bond interactions with good geometry, metal coordination or the hydrophobic effect, are colored in green. In contrast, atoms in unfavorable regions, such as donor–donor, acceptor–acceptor or polar–apolar contacts, are marked in red. White atoms do not contribute to the binding affinity. Figure 4 shows an example of a favorable interaction and an unfavorable contact with CPK coloring (Element color mode) and in the HYDE coloring scheme (HYDE color mode). Note that hydrogen bonds which deviate too far from the ideal hydrogen bond geometry in terms of both angles and distance are also considered unfavorable in HYDE and are therefore also colored red.

This coloring scheme allows direct visualization of the impact individual atoms have on the binding energy. We often choose, however, to map the scores of protein atoms onto their nearest ligand atom neighbor and color only the ligand atoms according to this accumulated score, to thus facilitate the identification of potential optimization sites at the ligand during the lead optimization process.

Results and discussion

The performance of the revised HYDE scoring function has been evaluated in several different aspects: Firstly, we assess the ability of HYDE to predict experimental binding constants of protein–ligand complexes. Here, two smaller series of congeneric compounds binding to thrombin and p38 MAP kinase respectively were analyzed in detail. Furthermore, the performance of HYDE was benchmarked on the PDBbind2007 coreset [3, 38, 39] and compared with the first version of HYDE [17], as well as with other well-established scoring functions. Secondly, HYDE is used as a post-docking rescoring function in cognate docking experiments to identify the bioactive conformation of a ligand from the pool of docking poses produced by FlexX [7, 26]. The results are compared with FlexX, as well as with GOLD [40–42] and PLANTS [43–45] which were also evaluated on the Astex diverse set [35]. Finally, in a large-scale virtual screening experiment using the Directory of Useful Decoys (DUD) [46] the ability of HYDE to discriminate between binders and non-binders is assessed and compared to other popular docking methods.
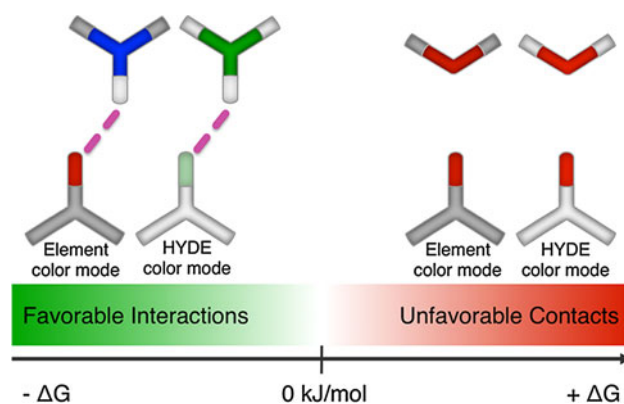

Fig. 4 Hyde coloring scheme: *Green atoms* contribute favorably to ΔG_{HYDE}. *Red atoms* contribute unfavorably to ΔG_{HYDE}. *White atoms* are energetically negligible. On the *left*, a hydrogen bond with ideal geometry is depicted; both atoms—donor and acceptor—are *colored green*. On the right, two hydrogen bond acceptor atoms (ether and carbonyl) making an unfavorable contact are both *colored red*

Binding affinity prediction: congeneric series

In this section, the binding affinities of compounds in two congeneric inhibitor series—one for thrombin and one for p38 MAP kinase—are estimated using the revised HYDE scoring function. Using detailed examples, we demonstrate in depth how the atom-based HYDE score and color scheme highlight the features of binding. We also use these examples to assess the ability of HYDE to predict binding affinity.

Thrombin

The crystal structures of five thrombin inhibitors (2ZFF, 2ZDV, 2ZF0, 2ZC9, 2ZDA) [47] were scored with HYDE. These D-Phe-Pro-based inhibitors differ only in the moiety binding to the S1 pocket of thrombin. In four complexes, a hydrophobic phenyl meta-substituted with H, CH_3, F or Cl occupies the S1 pocket. All five inhibitors are depicted in Fig. 5. The free energy of binding ΔG_{exp} of these structures (Fig. 5) is measured by isothermal titration calorimetry (ITC) [47]. Figure 5 also shows the five inhibitors in the HYDE coloring scheme with the predicted HYDE score ΔG_{HYDE}. For all of the compounds the HYDE score agrees well with the experimental binding affinity.

Some of the atoms in four of the inhibitors contribute unfavorably to the overall energy. We exemplarily use the thrombin complex 2ZC9 to give a more detailed explanation of the atom-based score contributions of the HYDE scoring function (see also Fig. 6). The D-Phe moiety of the inhibitor binds in the S3/S4 pocket of thrombin. The Pro moiety can be found in the S2 pocket and the *m*-chlorophenyl is situated in the S1 pocket (see Fig. 6 left).

In Fig. 6a, an unfavorable contribution to the HYDE score is shown. In this case, a hydrogen bond is formed between the amide nitrogen of the ligand and the backbone carbonyl of SER214. This hydrogen bond deviates from the ideal hydrogen bond geometry as the out-of-plane angle of the carbonyl lone-pair plane is 49°. HYDE tolerates a deviation up to 20° from the ideal angle and so considers this deviation too large (Fig. 6a). The hydrogen bond deviation factor f_{dev} (see Eq. 6) is 0.5 for this hydrogen bond which means that the hydrogen bond energy contribution is reduced by a half (−8.2 kJ/mol). Consequently, this hydrogen bond cannot compensate the desolvation costs of both hydrogen bonding partners and in fact turns out to make a destabilizing contribution to the overall energy of +6.4 kJ/mol.

Figure 6b shows an example of a favorable score contribution from the hydrophobic effect. The meta-substituted chlorine atom fits perfectly in the small hydrophobic subpocket of the S1 pocket leading to the full desolvation of this subpocket (−3.2 kJ/mol) and the chlorine itself (−3 kJ/mol).

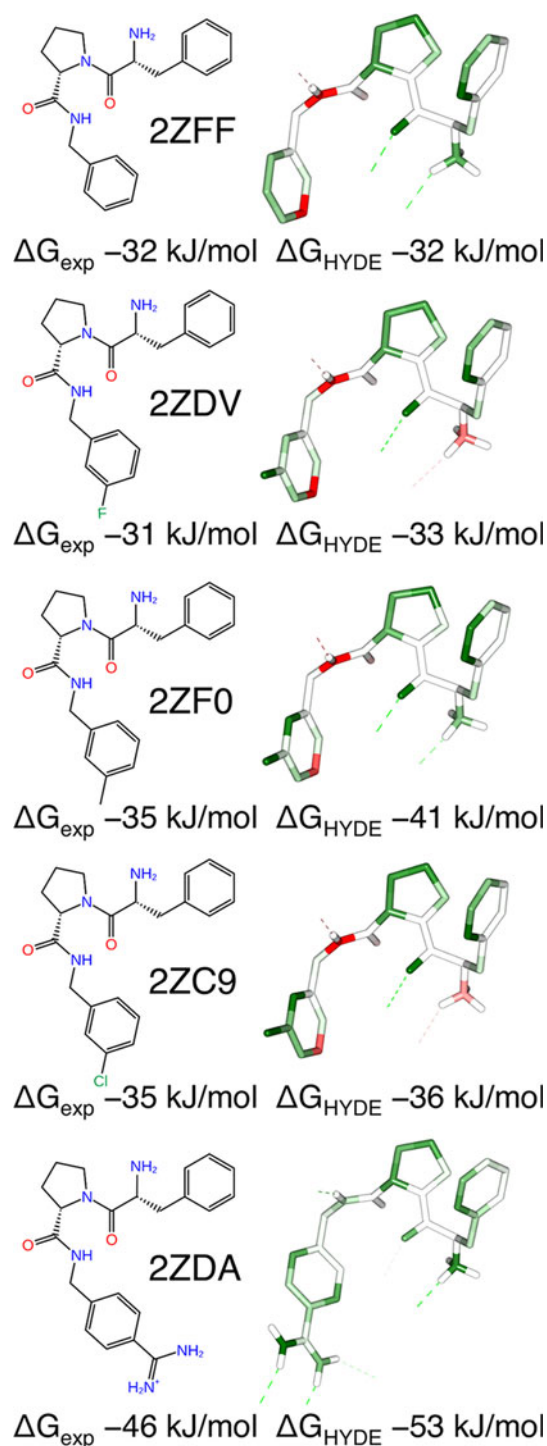**Fig. 5** Congeneric series of thrombin inhibitors. The PDB code for each inhibitor complex is shown in the *middle*. On the *left* the experimental binding affinity (ΔG_{exp}) is shown below the inhibitors. On the *right*, the inhibitors are depicted with the HYDE coloring scheme while the HYDE score (ΔG_{HYDE}) is shown below each inhibitor

Another kind of unfavorable contribution to the HYDE score is shown in Fig. 6c, a polar atom is desolvated by the *m*-chlorophenyl moiety. This desolvation of the polar
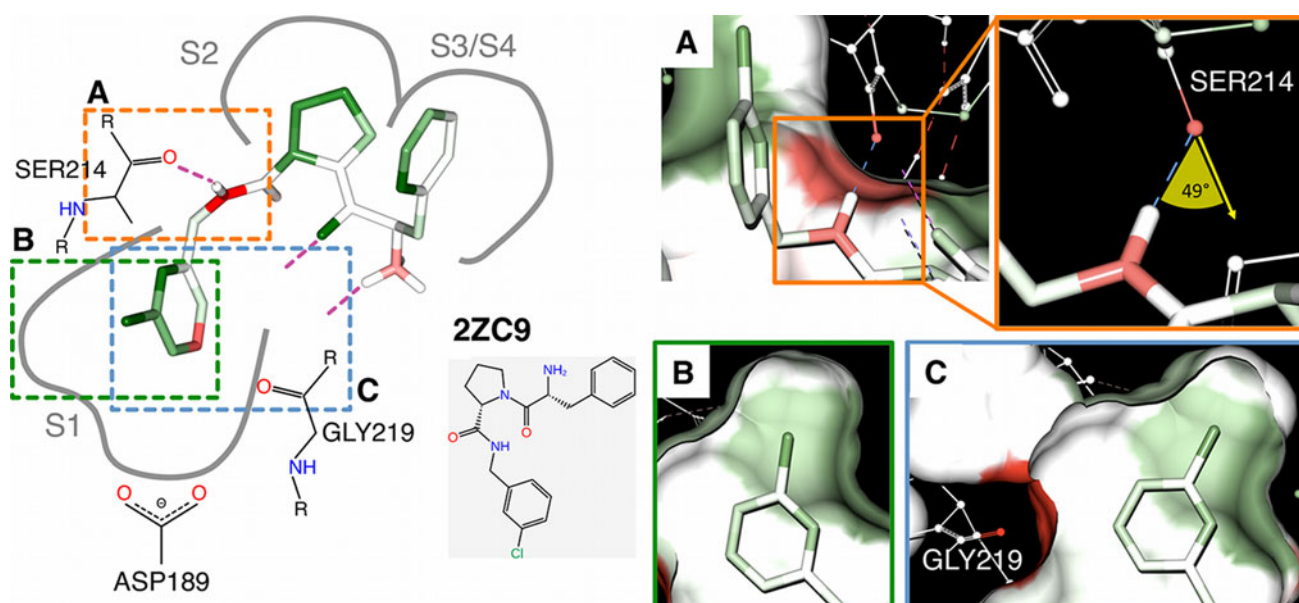
Fig. 6 Details of thrombin complex 2ZC9 as scored with HYDE. On the *left*, the binding pocket of thrombin is schematically depicted with the inhibitor in HYDE coloring scheme. On the *right*, three detailed scenarios are shown (note that the contributions of the protein atoms are not mapped to the ligand atoms in these illustrations): **a** Hydrogen bond deviating from ideal geometry between amide nitrogen of the ligand and backbone oxygen of SER214. The out-of-plane angle of the lone-pair plane amounts to 49°. **b** Hydrophobic effect of the chlorine atom in the small hydrophobic pocket. **c** Desolvation of GLY219 O by the hydrophobic phenyl ring

backbone carbonyl GLY219 is heavily penalized in HYDE (+6.4 kJ/mol) and cannot be compensated by the favorable contribution of the desolvated apolar carbon atom of the ligand (−1.7 kJ/mol). Both contributions are mapped onto the carbon atom of the ligand which is then colored in red (see Fig. 6 left).

p38 MAP kinase

Regan and coworkers have described their development of an inhibitor for p38 MAP kinase from lead structure to a clinical candidate [48]. Two crystal structures were submitted by them to the PDB [49]: the lead structure (1KV1) and the final clinical candidate BIRB (1KV2). We used the structure 1KV2 of the clinical candidate to model five of the intermediate synthesized compounds in the lead optimization process published by Regan et al. [48]. All compounds including the lead and the clinical candidate were scored with HYDE. We achieve a correlation coefficient R_P of 0.88 between the experimental measured affinity and the predicted binding energy for this congeneric compound series.

The lead optimization process is outlined in Fig. 7: the lead structure, the five modeled compounds and the clinical candidate BIRB are shown with the change in experimental binding affinity $\Delta\Delta G_{exp}$ with respect to the binding affinity of the lead structure, and for the modeled compounds. The respective modifications are highlighted.

Additionally, each compound is also depicted with the HYDE coloring scheme and $\Delta\Delta G_{HYDE}$ is given. In all cases except for compound 46, the change in the HYDE score agrees well with the change in experimental affinity. The modifications found in compound 46 lead to a gain in experimental affinity, whereas a small decrease in the binding energy is predicted by HYDE. One of the urea nitrogen atoms is colored red and contributes unfavorably to the binding energy due to its desolvation. In the lead structure both urea nitrogen atoms form a bidentate hydrogen bond with the side chain of GLU71. The introduction of the phenyl ring at the N2 of the pyrazole causes GLU71 to adopt an alternative side chain conformation, thereby allowing the phenyl ring to get in close contact with the alkyl portion of the GLU71 side chain. In addition, this leads to the disruption of the bidentate hydrogen bond of GLU71 with the urea moiety of the inhibitor and the formation of a monodentate hydrogen bond between one urea nitrogen and the carboxylate group of GLU71 [48]. Consequently, the other urea nitrogen becomes desolvated which is then heavily penalized by HYDE. In the case of compound 46, the favorable contribution of the hydrophobic effect of the newly introduced phenyl ring cannot compensate this high desolvation cost together with the loss of binding energy caused by the removal of the chlorine at the other phenyl ring. The reason for this may be that the cost of desolvating the urea nitrogen is currently overestimated by HYDE.

Binding affinity prediction: PDBbind 2007 coreset

We used the PDBbind 2007 coreset [38, 39] to evaluate HYDE on a larger dataset and to compare the performance of HYDE with other well-established scoring functions. Cheng and coworkers [3] assessed the ability of 16 different scoring functions, some of which are highly parameterized on experimental data, to predict experimental binding constants on the PDBbind 2007 coreset. This dataset consists of 195 protein–ligand complexes with high resolution crystal structures (less than or equal to 2.5 Å) and experimentally measured inhibition constant (K_i) or dissociation constant (K_d) values. We processed the crystal structures using the receptor preparation default settings in the LeadIT software [26]. The defaults are as follows: The active site is selected by taking all amino acids, cofactors and ions lying within 6.5 Å of any crystal structure ligand heavy atoms, then a coarse hydrogen bond network optimization of the active site with the crystal structure ligand is carried out by ProToss [37]. Finally, metal coordination geometries are automatically assigned. Some metal coordination geometries were manually

adjusted after close visual inspection of the complexes. All ligands of the dataset were processed using NAOMI [50]. We scored the 195 protein–ligand complexes of the dataset with both the first version of HYDE (HYDE1.0) and the revised version of HYDE (HYDE2.0). We also re-scored the 195 protein–ligand complexes that were optimized with HYDE2.0 and test some combinations of the revised HYDE scoring function terms. All results are summarized in Table 1.

We observed an improved performance of the revised HYDE function (Table 1: HYDE2.0) over the first version (Table 1: HYDE1.0) with a correlation between experimentally measured binding affinity and predicted binding affinity of $R_P = 0.323$. Optimizing with the numerical or stochastic optimization procedure only marginally improved the correlation [Table 1: HYDE2.0, column 7 (optimized structures)].

Table 1 initially shows that on this dataset, in comparison to other scoring functions, HYDE performs quite poorly and lies in the lower third of the table ranked according to the Pearson correlation coefficient. Some of the scoring functions (Table 1: e.g. PHOENIX or XScore)
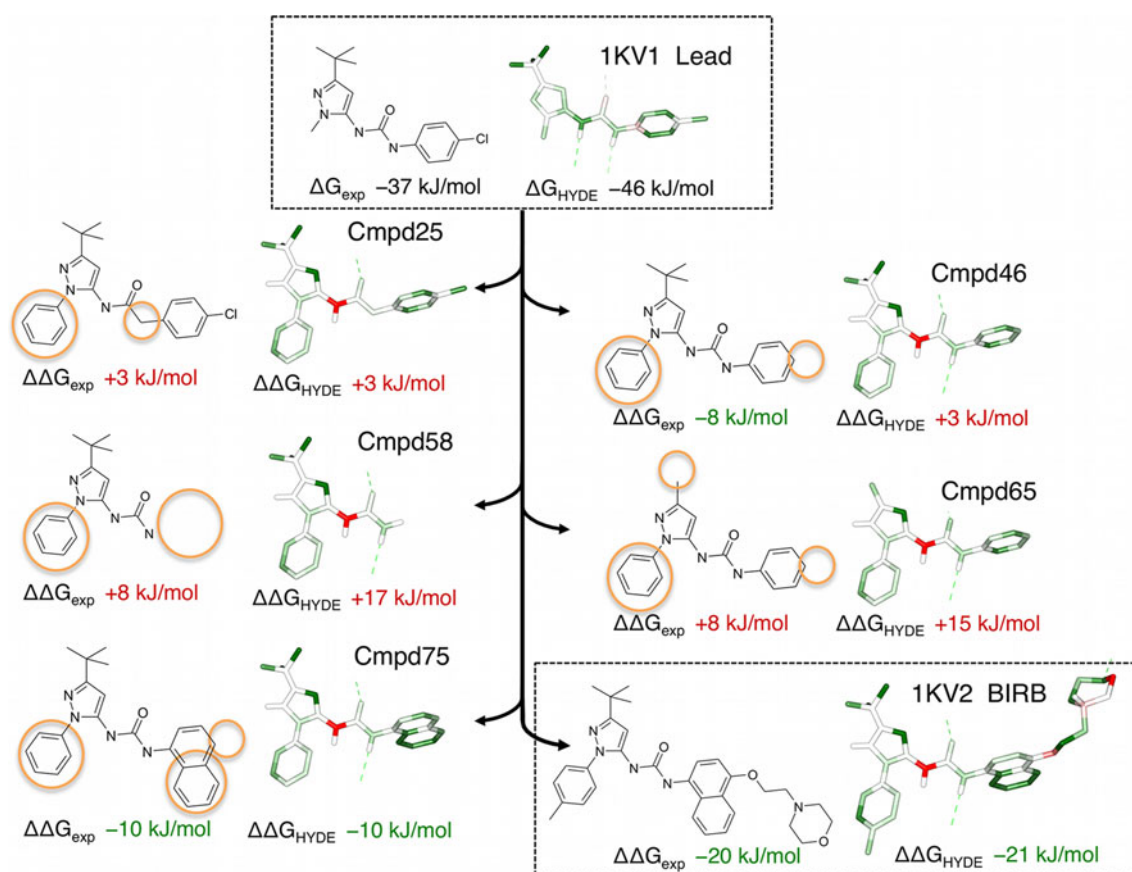

Fig. 7 Congeneric series of p38 MAP Kinase inhibitors. The development from lead structure (1KV1) to clinical candidate (1KV2) is shown. The modifications to the compounds with respect to the lead structure are highlighted with *orange circles*. The change in experimental affinity ($\Delta\Delta G_{exp}$) and Hyde score ($\Delta\Delta G_{HYDE}$) is shown below the compounds

Table 1 Correlation between experimental binding constant and predicted binding affinity for the PDBbind2007 coreset

Scoring function	On original complex structures				On optimized complex structures			
	N^a	R_P^b	SD^c	R_S^d	N^a	R_P^b	SD^c	R_S^d
X-Score::HMScore	195	0.644	1.83	0.705	195	0.649	1.82	0.701
HYDE2.0::HbondsHydrophobic	195	0.620	1.893	0.669	–	–	–	–
PHOENIX	194	0.616	2.16	0.644	–	–	–	–
HYDE2.0::Hydrophobic	195	0.602	1.91	0.642	–	–	–	–
DrugScoreCSD	195	0.569	1.96	0.627	195	0.589	1.93	0.649
SYBYL::ChemScore	195	0.555	1.98	0.585	194	0.622	1.87	0.668
DS::PLP1	195	0.545	2.00	0.588	194	0.529	2.03	0.569
GOLD::ASP	193	0.534	2.02	0.577	194	0.518	2.04	0.558
SYBYL::G-Score	195	0.492	2.08	0.536	195	0.522	2.03	0.579
DS::LUDI3	195	0.487	2.09	0.478	194	0.477	2.10	0.478
DS::LigScore2	193	0.464	2.12	0.507	194	0.479	2.10	0.505
GlideScore-XP	178	0.457	2.14	0.435	187	0.555	2.01	0.556
DS::PMF	193	0.445	2.14	0.448	194	0.471	2.11	0.482
GOLD::ChemScore	178	0.441	2.15	0.452	186	0.528	2.05	0.553
By NHA^e	195	0.431	2.15	0.517	195	0.431	2.15	0.517
SYBYL::D-score	195	0.392	2.19	0.447	195	0.388	2.20	0.443
HYDE2.0	161	0.323	2.26	0.340	189	0.345	2.24	0.378
DS::Jain	189	0.316	2.24	0.346	190	0.339	2.26	0.362
GOLD::GoldScore	169	0.295	2.29	0.322	188	0.329	2.26	0.386
HYDE1.0	150	0.271	2.30	0.321	–	–	–	–
SYBYL::PMF-score	190	0.268	2.29	0.273	180	0.235	2.31	0.235
SYBYL::F-score	185	0.216	2.35	0.243	181	0.238	2.31	0.208

Scoring functions are ranked by the Pearson correlation coefficients obtained on the original complex structures. Results are given for both original and optimized structures

Results (excluding HYDE and PHOENIX [51]) are taken from Cheng et al. [3]

^{a} Number of complexes receiving favorable binding scores by this scoring function

^{b} Pearson correlation coefficients

^{c} Standard deviations in linear correlation (in log K_d units)

^{d} Spearman correlation coefficients

^{e} Using the number of heavy atoms on each ligand as the only variable in correlation

were calibrated on similar datasets to the PDBbind 2007 coreset which may explain their superior performance to HYDE. However, a lower performance than that of using the number of heavy atoms (Table 1: NHA) cannot by explained by training alone. Since we know from experience that HYDE is very sensitive to even small inaccuracies in structural data [19], we examined the dataset more closely and found that structural deficiencies can be observed in many of the complexes. A detailed assessment of the structures including classification criteria can be found in the Supplementary Material (Table S1). Figure 8 shows structural deficiencies of four exemplary complexes. In all four cases, missing electron density for large parts of the ligand can be observed. In two of the complexes, there are alternative conformations for the active site. Sondergaard and coworkers also analyzed the PDBbind 2007

refined dataset for structural artifacts. They found that 36 % of the protein–ligand complexes were influenced by crystal contacts and that the performance of a scoring function will be affected by these [52]. We assume that the hydrogen bond definition used in HYDE, where HYDE penalizes hydrogen bonds deviating from the optimum geometry, introduces noise when using these structures of lower quality.

To test this theory, and to understand better which of the terms in HYDE are influenced the most by the structural quality, we separately tested different components of the HYDE scoring function: the hydrophobic effect, the hydrogen bond energy and the dehydration penalty. Using only the term for the hydrophobic effect, the correlation coefficient vastly increased to $R_P = 0.602$ (Table 1: HYDE2.0::Hydrophobic). Including the hydrogen bond

Fig. 8 Examples of electron density for complexes in the PDBbind2007 coreset. Ligands are highlighted with *oranges circles*. **a** Almost no electron density for the ligand (1HK4). **b** No density for the ligand (1GNI). **c** Missing electron density for the ligand and the binding site, plus alternative conformations for PHE146 and ARG145 (1AJP). **d** Missing electron density for the ligand and the binding site, ligand is fragmented and has an alternative conformation for the nitrobenzene, waters are modeled inside the ligand, alternative conformation for LYS33 (1PXO)
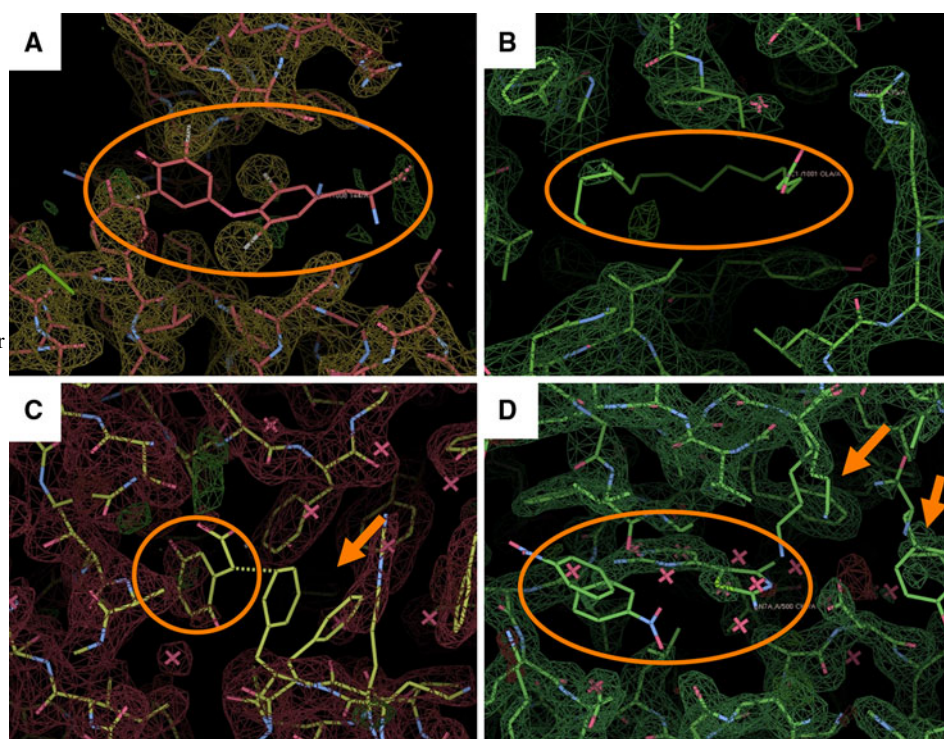
energy term together with the hydrophobic effect lead to a further improvement in the correlation to $R_P = 0.620$ (Table 1: HYDE2.0::HbondsHydrophobic), second best of all the scoring functions. This confirms that it is the polar dehydration penalty relative to the hydrogen bond energy gain that leads to the strong sensitivity of HYDE against structural inaccuracies. Using only these two rather simple components of HYDE, which most notably are *not* calibrated on experimental binding affinities or protein–ligand complexes, we can predict binding affinity better than nearly all of the other highly parameterized scoring functions. Despite these results on the PDBbind 2007 coreset, we found that the polar dehydration term of the HYDE function largely reduces the number of false positives in virtual screening (e.g. on the DUD dataset [46]).

Redocking and virtual screening performance

Recently, we evaluated HYDE in large-scale redocking and screening experiments using a revised version of the Astex diverse set [35] and the Dataset of Useful Decoys (DUD) [46], respectively [19]. Here, we show the performance of HYDE in cognate docking on the original Astex diverse set and compare it to FlexX [7, 26] and two other methods, PLANTS [43–45] and GOLD [40–42], which also used this dataset for validation. Furthermore, we compare the virtual screening performance of HYDE with several well-established structure-based methods on the DUD. This section provides a comparison of HYDE to other methods rather

than a detailed analysis of our results. A very detailed study of both datasets using HYDE and the exact set-up of the experiments can be found in [19].

The Astex diverse set contains 85 high-quality crystal structures of relevant protein–ligand complexes [35]. The PDB crystal structures of the complexes were processed using the receptor preparation in the LeadIT software [26]. The hydrogen bond network of the complexes is pre-optimized. All amino acids, cofactors and ions lying within 6.5 Å of any crystal ligand heavy atom were included in the binding site definition. The automatically assigned metal coordination geometry for ions by LeadIT was manually corrected in some of the complexes (for more information see the Supplementary Material of [19]). The reference ligands were converted from the *mol* format to *mol2* format using NAOMI [50]. Random start conformations were generated for all reference ligands from SMILES format with CORINA 3.48 [53, 54]. We generated 200 docking poses for each ligand using the latest version of the FlexX docking algorithm [7] which is included in the LeadIT software suite (version 2.1.1) [26]. Table 2 shows our result in comparison to other methods.

A good performance of HYDE in cognate docking is achieved using the stochastic optimizer (HYDE2.0 Table 2). We performed three iterations due to the stochastic nature of the optimizer and yielded a success rate of 76 % for the best scored pose with RMSD better than 2 Å. By considering the 20 best scored poses this success rate is even enhanced to 94 %. Comparing these results to the

Table 2 Cognate docking results on the Astex diverse set

	RMSD [Å] ≤		
	1.0	1.5	2.0
HYDE[a]			
Top 1	53 (2)	67 (2)	76 (1)
Top 20	80 (1)	90 (1)	94 (1)
FlexX			
Top 1	35	56	69
Top 20	73	85	87
GOLD::GoldScore[b]			
av best25	–	–	74–81
GOLD::ChemScore[b]			
av best25	–	–	68–79
PLANTS::CHEMPLP[b]			
av best25	–	–	76–89
PLANTS::PLP[b]			
av best25	–	–	75–86

Results are given as percent

[a] Due to the stochastic nature of the optimization process the results are averaged over three iterations, standard deviations are given in brackets

[b] Results taken from Korb et al. [45]

performance of the FlexX score the success rates are 7 percentage points for both the best scored pose and best 20 poses. The results of HYDE are comparable to the performance of PLANTS and GOLD on this dataset (Table 2). However, it is important to again note that HYDE is not calibrated on any protein–ligand complexes.

The Directory of Useful Decoys (DUD) [46] contains 40 different relevant protein targets and a number of experimental validated binders for each of the targets. Appropriate decoys for each target, being physically similar but topological distinct to the binders, were chosen from the ZINC database [55]. Hence, this dataset presents a challenging large-scale virtual screening test set. Recently, we published our results on the DUD, comprising a detailed analysis on exemplary targets of this dataset [19]. Here, we compare our results on this dataset to that of other well-established structure-based methods. In summary the workflow is as follows: For all compounds—actives and decoys-docking poses were generated using the LeadIT software [26]. The best 40 poses according to the FlexX score were kept for rescoring with HYDE. Both optimization steps—ProToss optimization of the hydrogen bond network followed by numerical optimization of the complex geometry—were employed during the rescoring. Figure 9 shows the comparison of the performance of HYDE in virtual screening on the DUD with other. All results shown are based on rigid protein structures—results including a minimization of the protein are not shown in
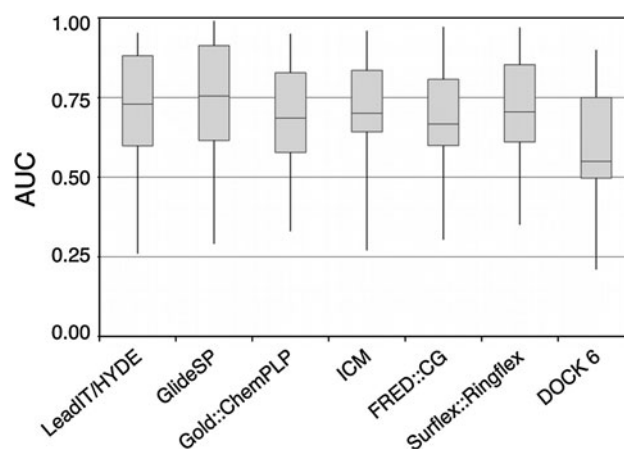

Fig. 9 Virtual screening results of different methods using the DUD benchmark set. Shown in the *boxplot* are the median AUC (*grey line* in the *box*), the lower and upper quartile (*box*) and the minimum and maximum AUC value (*black vertical lines*). AUC values are taken from [19] (LeadIT/HYDE), [56] (GlideSP), [57] (Gold::ChemPLP), [58] (ICM), [59] (FRED::CG, Surflex::Ringflex) and [60] (DOCK 6)

this comparison, as this version of HYDE did not include a protein minimization. The combination of LeadIT and HYDE achieved a median AUC of 0.73 across all 40 complexes [19] which is comparable with the other best performing methods on this dataset. This result was also achieved using a fully automated process without any manual correction of the input data.

Conclusions

In this paper we described the further development of the HYDE scoring function. Several aspects of first version of HYDE were revised even though the overall concept has been retained. The plogP increments were re-parameterized and the number of plogP atom types was vastly reduced in comparison to the first version. We introduced new terms to better describe the dehydration of hydrophilic atoms and to allow scoring of metal ions. In addition, the HYDE scoring function is now embedded in a target function for optimization of the complex conformation.

Looking at two examples in detail, we have shown that HYDE is able to predict the experimental binding affinity of congeneric series of compounds and rank them in the correct order. The evaluation of HYDE on a larger, very diverse dataset again highlighted the sensitivity of HYDE to inaccuracies in the input data. We found that especially the polar dehydration term of HYDE causes this sensitivity, since small structural inaccuracies in the data can lead to a highly amplified penalty. Comparing the ability of HYDE to estimate the binding energy of protein–ligand complexes to that of other well-established scoring functions, we found that HYDE performed as one of the best. This

promising result was obtained using two components of the HYDE scoring function: the hydrophobic effect and the hydrogen bond energy. Moreover no parameterization of HYDE on experimental affinities or protein–ligand complexes is necessary to achieve this result. Although this is a really satisfying achievement, we would still prefer to evaluate HYDE on more meaningful datasets, such as a congeneric compound series complete with crystal structures and binding affinity data all generated in the same laboratory ensuring consistency throughout. This would allow us to draw more reliable conclusions about the performance of our methods.

In addition, we also demonstrated that when HYDE is applied as a rescoring function in cognate docking or virtual screening, we are able to improve upon the results of the native scoring function. On the Astex diverse set, we obtain a success rate of up to 76 % defined as finding a docking pose with an RMSD below 2 Å at the first rank. This increases to 94 % if we take the 20 best scored poses into account. In the virtual screening experiment on DUD, designed to test the discrimination of true binders from decoys, HYDE performs as well the other best methods, achieving a median AUC of 0.73.

Another advantage of HYDE has been illustrated in several detailed examples: the comprehensible atom-based score contributions can be translated into a very intuitive coloring scheme, which allows easy detection of favorable and unfavorable contributions in the protein–ligand complex.

The development of the HYDE scoring function is still ongoing. We intend to include receptor flexibility during the optimization process to better handle inaccuracies in crystal structures. We are currently working on an improved model of water to replace the correction factor, whilst consideration of the conformation of the ligand in the unbound state is also certainly of interest for future work.

References

1. Jorgensen WL (2004) The many roles of computation in drug discovery. Science 303:813–1818
2. Matter H, Sotriffer C (2011) In: Sotriffer C (ed) Virtual screening: principles, challenges and practical guidelines, 1st edn. Wiley-VCH, Weinheim
3. Cheng T, Li X, Li Y, Liu Z, Wang R (2009) Comparative assessment of scoring functions on a diverse test set. J Chem Inf Mod 49:1079–1093
4. Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR (2008) Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. Br J Pharmacol 153:7–26
5. Sotriffer C, Matter H (2011) In: Sotriffer C (ed) Virtual screening: principles, challenges and practical guidelines, 1st edn. Wiley-VCH, Weinheim
6. Böhm HJ (1994) The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. J Comput Aided Mol Design 8:243–256
7. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. J Mol Biol 261:470–489
8. Savage HJ, Elliott CJ, Freeman CM, Finney JL (1993) Lost hydrogen bonds and buried surface area: rationalising stability in globular proteins. J Chem Soc, Faraday Trans 89:2609–2617
9. Bissantz C, Kuhn B, Stahl M (2010) A medicinal chemist's guide to molecular interactions. J Med Chem 53(14):5061–5084
10. Pham TA, Jain AN (2006) Parameter estimation for scoring protein–ligand interactions using negative training data. J Med Chem 49:5856–5868
11. Krammer A, Kirchhoff PD, Jiang X, Venkatachalam CM, Waldman M (2005) LigScore: a novel scoring function for predicting binding affinities. J Mol Graph Model 23:395–407
12. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. J Med Chem 49: 6177–6196
13. Sotriffer CA, Sanschagrin P, Matter H, Klebe G (2008) SFC-score: scoring functions for affinity prediction of protein–ligand complexes. Proteins 73:395–419
14. Mysinger MM, Shoichet BK (2010) Rapid context-dependent ligand desolvation in molecular docking. J Chem Inf Model 50:1561–1573
15. Kellogg GE, Burnett JC, Abraham DJ (2001) Very empirical treatment of solvation and entropy: a force field derived from Log Po/w. J Comput Aided Mol Des 15:381–393
16. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. J Comput Aided Mol Des 16:11–26
17. Reulecke I, Lange G, Albrecht J, Klein R, Rarey M (2008) Towards an integrated description of hydrogen bonding and dehydration: reducing false positives in virtual screening with the hyde scoring function. ChemMedChem 3(6):885–897
18. Lange G, Klein R, Albrecht J, Rarey M, Reulecke I (2010) European patent specification EP2084520
19. Schneider N, Hindle S, Lange G, Klein R, Albrecht J, Briem H, Beyer K, Claußen H, Gastreich M, Lemmen C, Rarey R (2012) Substantial improvements in large-scale redocking and screening using the novel HYDE scoring function. J Comput Aided Mol Des 26:701–723
20. Richards FM (1977) Areas, volumes, packing, and protein structures. Ann Rev Biophys Bioeng 6:151–176
21. Connolly ML (1983) Solvent-accessible surfaces of proteins and nucleic acids. Science 221:709–713
22. Connolly ML (1983) Analytical molecular surface calculation. J Appl Cryst 16:548–558
23. Stefano Forli, Olson AJ (2012) A force field with discrete waters and desolvation entropy for hydrated ligand docking. J Med Chem 55:623–638

24. Schneider N, Klein R, Lange G, Rarey M (2012) Nearly no scoring function without a Hansch-analysis. Mol Inf 31:503–507
25. Stahl M (2000) Modifications of the scoring function in FlexX for virtual screening applications. Perspect Drug Discov 20:83–98
26. LeadIT. BioSolveIT GmbH, Sankt Augustin. http://www.biosolveit.de/leadit/. Accessed 12 June 2012
27. Physprop database. http://www.syrres.com/esc/physprop.htm. Accessed 12 June 2012
28. Hansch C, Leo AJ (1985) Medchem project issue no. 26. Pomona College, Claremont, CA
29. Hansch C, Leo AJ (1987) The log P database. Pomona College, Claremont, CA
30. Hansch C, Leo A, Hoekman D (1995) Exploring QSAR. Hydrophobic, electronic, and steric constants. American Chemical Society, Washington, DC
31. Leo AJ (1993) Calculating log Poct from structures. Chem Rev 93:1281–1306
32. Lee B, Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55:379–400
33. Shrake A, Rupley JA (1973) Environment and exposure to solvent of protein atoms, lysozyme and insulin. J Mol Biol 79:351–371
34. Bondi A (1964) Van der Waals volumes and radii. J Phys Chem 68:441–451
35. Hartshorn MJ, Verdonk ML, Chessari G, Brewerton SC, Mooij WTM, Mortenson PN, Murray CW (2007) Diverse, high-quality test set for the validation of protein–ligand docking performance. J Med Chem 50:726–741
36. Seebeck B, Reulecke I, Kämper A, Rarey M (2008) Modeling of metal interaction geometries for protein–ligand docking. Protein Struct Funct Bioinform 71:1237–1254
37. Lippert T, Rarey M (2009) Fast automated placement of polar hydrogen atoms in protein–ligand complexes. J Cheminf 1:13
38. Wang R, Fang X, Lu Y, Wang S (2004) The PDBbind database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures. J Med Chem 47:2977–2980
39. Wang R, Fang X, Lu Y, Yang CY, Wang S (2005) The PDBbind database: methodologies and updates. J Med Chem 48:4111–4119
40. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245:43–53
41. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748
42. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein–ligand docking using GOLD. Proteins 52:609–623
43. Korb O, Stützle T, Exner TE (2006) PLANTS: application of ant colony optimization to structure-based drug design. Lect Notes Comput Sci 4150:247–258
44. Korb O, Stützle T, Exner TE (2007) An ant colony optimization approach to flexible protein–ligand docking. Swarm Intel 1(2): 115–134
45. Korb O, Stützle T, Exner TE (2009) Empirical scoring functions for advanced protein–ligand docking with PLANTS. J Chem Inf Mod 49:84–96
46. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. J Med Chem 49(23):6789–6801
47. Baum B, Mohamed M, Zayed M, Gerlach C, Heine A, Hangauer D, Klebe G (2009) More than a simple lipophilic contact: a detailed thermodynamic analysis of nonbasic residues in the s1 pocket of thrombin. J Mol Biol 390:56–69
48. Regan J, Breitfelder S, Cirillo P, Gilmore T, Graham AG, Hickey E, Klaus B, Madwed J, Moriak M, Moss N, Pargellis C, Pav S, Proto A, Swinamer A, Tong L, Torcellini C (2002) Pyrazole urea-based inhibitors of p38 MAP kinase: from lead compound to clinical candidate. J Med Chem 45:2994–3008
49. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242
50. Urbaczek S, Kolodzik A, Fischer JR, Lippert T, Heuser S, Groth I, Schulz-Gasch T, Rarey M (2011) NAOMI—on the almost trivial task of reading molecules from different file formats. J Chem Inf Mod 51:3199–3207
51. Tang YT, Marshall GR (2011) PHOENIX: a scoring function for affinity prediction derived using high-resolution crystal structures and calorimetry measurements. J Chem Inf Mod 51:214–228
52. Sondergaard CR, Garrett AE, Carstensen T, Pollastri G, Nielsen JE (2009) Structural artifacts in protein–ligand X-ray structures: implications for the development of docking scoring functions. J Med Chem 52:5673–5684
53. Sadowski J, Gasteiger J, Klebe G (1994) Comparison of automatic three-dimensional model builders using 639 X-ray structures. J Chem Inf Comput Sci 34:1000–1008
54. CORINA. Molecular Networks GmbH, Erlangen, Germany. http://www.molecular-networks.com/products/corina. Accessed 12 June 2011
55. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182
56. Repasky MP, Murphy RB, Banks JL, Greenwood JR, Tubert-Brohman I, Bhat S, Friesner RA (2012) Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. J Comput Aided Mol Des 26:787–799
57. Liebeschuetz JW, Cole JC, Korb O (2012) Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Mol Des 26:737–748
58. Neves MAC, Totrov M, Abagyan R (2012) Docking and scoring with ICM: the benchmarking results and strategies for improvement. J Comput Aided Mol Des 26:675–686
59. McGann M (2011) FRED pose prediction and virtual screening accuracy. J Chem Inf Mod 51(3):578–596
60. Brozell SR, Mukherjee S, Balius TE, Roe DR, Case DA, Rizzo RC (2012) Evaluation of DOCK 6 as a pose generation and database enrichment tool. J Comput Aided Mol Des 26:749–773