Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein−Ligand Docking Results

Tim ten Brink and Thomas E. Exner*

Department of Chemistry and Zukunftskolleg, University of Konstanz, D-78457 Konstanz, Germany

Received November 13, 2008

In this work, we present a systematical investigation of the influence of ligand protonation states, stereoisomers, and tautomers on results obtained with the two protein−ligand docking programs GOLD and PLANTS. These different states were generated with a fully automated tool, called SPORES (Structure PrOtonation and Recognition System). First, the most probable protonations, as defined by this rule based system, were compared to the ones stored in the well-known, manually revised CCDC/ASTEX data set. Then, to investigate the influence of the ligand protonation state on the docking results, different protonation states were created. Redocking and virtual screening experiments were conducted demonstrating that both docking programs have problems in identifying the correct protomer for each complex. Therefore, a preselection of plausible protomers or the improvement of the scoring functions concerning their ability to rank different molecules/states is needed. Additionally, ligand stereoisomers were tested for a subset of the CCDC/ASTEX set, showing similar problems regarding the ranking of these stereoisomers as the ranking of the protomers.

INTRODUCTION

Pharmaceutical research more and more relies on computational methods, from hit identification to lead optimization, in the hope of cutting down the immense consumption of time and money needed for the development of new drugs. Protein−ligand docking, one key method in structure-based drug design, is often used for the prediction of the interaction patterns of small molecules in the active site of a protein as well as in virtual screening of large databases to identify new and hopefully highly active lead structures. A large variety of different approaches for a solution of this problem has been applied[1−6] ranging from fragment-based (e.g., DOCK,[7−10] FlexX,[1,2,11,12] SLIDE,[13] SURFLEX,[14] and GLIDE[15−17]) over stochastic optimization methods for finding the global minimum (e.g., GOLD,[18−20] AutoDock,[21,22] PLANTS,[23−25] ICM,[26] QXP,[27] and PRO_LEADS[28,29]) to multiconformer docking (e.g., FRED[30]).

In the last time, more and more studies have been published comparing different docking tools on a large test set of complex structures.[2,3,31,32] These studies report success rates of 30 to 70%, where the success rate is defined as the percentage of complexes, for which the predicted structure with the lowest energy has a root-mean-square deviation (rmsd) to the crystal structure of less than 2 Å. This shows that, despite the good results of some approaches, a universal docking tool (algorithm and scoring function) that outperforms all others on every system is not available at the moment. Besides insufficient conformational sampling, the scoring functions were identified as the main reason for these failures.[2,3,31,32] Additionally, the importance of the preparation of the protein and especially the ligand structure and its influence on the docking results was stressed (once again) in these studies. Because the position of hydrogen atoms cannot be determined experimentally by X-ray crystallography, protonation and tautomeric states must be predicted on a theoretical basis, and the assignment of atomic hybridization and bond orders are not always straightforward. For scoring functions, which take actual hydrogen positions into account for the calculation of hydrogen-bonding strengths, the correct placement of these atoms is often the key to identifying the correct structure with a docking approach. The number and placement determine hydrogen bonding activity and formal charges and are influenced by the relative orientation of the two complex partners. Some groups commit large amounts of research effort into the preparation of standardized docking test sets. E.g. in the CCDC/ASTEX[33] and the ASTEX diverse[34] sets each complex was manually checked by a human investigator. Because of the high accuracy in these data sets, these are very well suited for testing and comparing different docking approaches and scoring functions. But this effort cannot be devoted to every ligand in a large ligand database of up to millions of compounds. Therefore, fully automatic procedures for the preparations are needed. In most cases in the literature, standard modeling tools are applied. But e.g. for the docking tool GLIDE[15−17] a specific procedure for the setup, either with the GUI MAESTRO[35] or via command line, was developed starting with the routine *PROTPREP* for the protein preparation, which defines the binding site, adds hydrogen atoms, neutralizes side chains (that are not in the binding site and do not participate in salt bridges), and checks for chemical correctness. The routine *LIGPREP* can be used to prepare ligand structures. It is able to add hydrogen atoms, to create up to eight different tautomers for the ligand molecules, to set up stereoisomers, and to change the ionization state of the molecule according to a given pH value. The authors recommend strongly its use and warn of

* Corresponding author phone: +49-7531-88-2015; fax: +49-7531-88-3587; e-mail: thomas.exner@uni-konstanz.de. Corresponding author address: Department of Chemistry and Zukunftskolleg, Deutsch/Englisch oder so lassen, University of Konstanz, D-78457 Konstanz, Germany.

10.1021/ci800420z CCC: $40.75
Published on Web 05/19/2009

the danger of decreasing accuracy if the molecules are not set up with this procedure. The Chemical Computing Group presented PROTONATE 3D[36] which can be used to prepare protein ligand complexes. It adds hydrogen atoms, generates tautomers and rotamers, and determines an accurate protonation for polar groups in the binding site using a free energy optimization. Openbabel[37] and I-interpret[38] were developed for file format conversion, where also the problems of atom/bond-type assignment and sometimes also of adding hydrogen atoms occur. A special feature of I-interpret[38] is the recognition of functional groups, which has the advantage that the process relies less on the geometry of the given molecule. If hydrogen atoms are added, by default the neutral form of all functional groups is generated, but this behavior can be changed manually by the user.

In the paper presented here, we introduce a new approach for the setup of docking experiments called SPORES (Structure PrOtonation and REcognition System). It can be used for the preprocessing of proteins and protein−ligand complexes as e.g. taken from the Protein Data Bank (PDB)[39] or for the setup of 3D ligand databases. Additional to the rule based assignment of atom and bond types and the protonation according to these types, called standard protonation in the following, different protonation and tautomeric states as well as different stereoisomers can be generated. Due to the influence of the protein, the location and number of hydrogen atoms depend on the actual placement of the ligand in the active site and cannot be predicted by the gas phase or by solution properties alone. But as already mentioned above, the hydrogen atoms are crucial for successful docking. In most databases only one configuration is stored for each molecule, but the activity of different isomers can vary extremely, making the generation and testing of all possible stereoisomers obligatory. SPORES can also generate different ring structures. In the publication presented here, we will take the ring conformations from the experimental structure, because we want to focus on the influence of hydrogen atom addition and we tried to minimize the effects of other factors like suboptimal rings.

Surprisingly, only a few papers deal with these factors explicitly. Polgar et al.[40] systematically investigated the influence of protonation using four different docking programs but only with the single target β-secretase (BACE1). Brooks et al.[41] presented a study about the influence of stereoisomers on virtual screening but also for a single target only. More recently, Reulecke et al. presented a novel scoring function, which is mainly based on hydrogen bonds and dehydration.[42] In this study, they test the influence of different ligand protonations on the screening result of three targets. Here, we will use a subset of the CCDC/ASTEX data set consisting of 213 complexes to test our ligand setup system SPORES because the automatically generated structures can be directly compared to the manually created ones. Then, all ligand structures with all possible protonation, tautomeric, and isosteriomeric states will be used in docking and virtual screening experiments using the program GOLD[18−20] and the recently introduced approach PLANTS[23−25] to see if these are able to identify the correct configurations and binding poses and to rank these accordingly.

MATERIALS AND METHODS

General Structure Preparation. All structures used for testing our setup procedure and for the docking experiments were taken from the CCDC/ASTEX clean data set,[33] which contains 235 protein−ligand complexes. The exclusion of all covalently bound complexes reduced the data set to 213. Because of the great care, with which this data set was prepared and manually revised,[33] here we define the stored protonation, tautomeric, and stereoisomeric states of the proteins and the ligands, short the ASTEX protonation, as the optimal one, which is expected to give the best docking results. Additionally, a second set of ligand structures was prepared with our automated procedure SPORES. For this, the manually added hydrogen atoms of the ligand were first removed and then automatically added again with SPORES. SPORES assigns atom and bond types according to the TRIPOS force field[43] convention. Atom types are determined from geometrical properties like the bond length and bond angles. Please note, that at this stage no information about the positions of hydrogen atoms are taken into account. If possible, torsional angles are used for a better discrimination between different hybridizations. After the atom typing, bond types are assigned according to bond length, attached atoms, and their atom types. For conjugated systems, an iterative routine is applied to distribute double and single bonds. In carboxyl, phosphate, and sulfate groups all oxygen atoms are treated equally with a formal charge of q/n_O (q: formal charge of functional group, n_O: number of oxygen atoms), e.g. -0.5 in a carboxyl group, and are connected with a bond of type "*ar*" of the Tripos force field,[43] corresponding to a bond order of 1.5. Afterward the hydrogen atoms are added to the structure. Their coordinates are determined from the heavy atom positions and hybridization. In the standard protonation, amines are protonated to the quaternary form, while carboxyl groups and other acidic groups like phosphate and sulfate groups or hydroxylamines are deprotonated. Other charged groups than the just described ones are avoided if possible.

The virtual screening data set was directly taken from the PDB, aligned with RELIBASE,[44] and then processed with SPORES as just described.

Generation of Protonation States. SPORES generates different protonation states using a rule based combinatorial method. Hydrogen atoms at carboxylic or phenolic oxygen atoms as well as quaternary amines, aromatic nitrogen atoms, hydroxylamines, and sulfonamides are marked as acidic. While the deprotonated forms of these functional groups are marked as alkaline. Table 1 gives a summary of all acidic and alkaline groups considered. Protonation states are then generated by protonating single alkaline heavy atoms or removing single acidic hydrogen atoms thus creating all possible protonation states according to the predefined classification.

Generation of Tautomers. Tautomers are generated independently from the protonation states. Substructures that can be conveyed in another tautomeric form such as diketons, enols, or lactams are identified. The structures considered for the generation of tautomers are given in Table 2. The atom and bond types are adapted to the hybridization necessary for the new form. Afterward the new structure is reprotonated. In a final step an energy minimization with

Table 1. Different Kinds of Functional Groups Considered in the Generation of Protonation States Shown in Their Standard Protonation

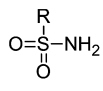
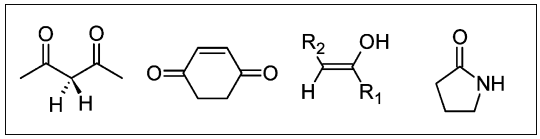
Alcalic groups, protonated in the standard protonation	Quaternary amines	
Weak acidic groups protonated in standard protonation	Phenolic hydroxyl groups, nitrogen in 5-membered aromatic rings, sulfonamides	Ar-OH  
Weak alcalic groups deprotonated in standard protonation	Nitrogen in 6 membered aromatic rings	
Acidic groups deprotonated in standard protonation	Carboxyl groups	

Table 2. Different Kinds of Substructures Considered in the Generation of Tautomers in SPORES



the TRIPOS force field as implemented in SYBYL7.2[45] is conducted to adept the molecule geometry to the new hybridizations. With the iterative application of this procedure like the one described by Milletti et al.,[46] all tautomeric forms of complicated ring systems can be generated as e.g. the one in Figure 1.

Generation of Stereoisomers. Stereoisomers are generated by SPORES in a two step process by first identifying all stereo centers and second generating all possible isomers by combinatorial inversion of single centers. The identification is done by searching for atoms with four neighbors. From these centers, tree searches are started to identify isomorphic neighbors. All atoms with four not isomorphic neighbors are marked as stereo centers. Afterward a combinatorial method is applied to create all possible stereoisomers of the starting molecule. If a center is to be flipped, two neighbor trees are taken and reflected at the plane defined by the stereo center and the first atoms of the two remaining neighbor trees. The neighbor trees to be flipped are selected by the following criteria (see also Figure 2): First, no neighbors, which are part of the same ring system as the center, and second, no neighbors, which contain other stereo centers, are chosen if possible. The second criterion avoids an inversion of these other stereo centers during the reflection. If these rules cannot identify unambiguously two neighbor trees, the sizes of the neighbor trees define the ones to be flipped starting with the smallest one. All generated stereoisomers are minimized using the TRIPOS force field to remove clashes and other unfavorable interactions emerged from the reflection process.

Docking of Protonation States and Stereoisomers. The docking studies were conducted with PLANTS, using standard settings (*speed 1*) and the $PLANTS_{chemplp}$ scoring function,[25] and with GOLD, using automated scaling with disabled ring-corner flips and the GOLDSCORE scoring function.[20] The binding site was taken as defined in the CCDC/ASTEX data set in the redocking experiments and as a sphere enclosing all protein atoms up to 6 Å away from any ligand heavy atom of the superimposed active ligands in virtual screening. Each ligand was docked 25 times to its receptor to minimize heuristic influences. Prior to docking, the ligand structures were randomized with respect to their translational, rotational, and torsional degrees of freedom to eliminate any potential search bias toward the crystallographic conformation. A docking attempt is considered as a success if the best ranked structure has a heavy atom root-mean-square deviation (rmsd) below 2 Å compared to the crystal structure.

RESULTS AND DISCUSSION

Standard Protonation. Our so-called standard protonation is designed to give chemical correct structures with a consistent protonation for aqueous solution and non- or only weak polarized binding sites. When treating the CCDC/ASTEX data set with the automated procedure described above, in 172 cases the standard protonation and the originally stored protonation are identical. Deviations of the standard protonation from the ASTEX protonation, which we assume to be the optimal one for a specific ligand in its docked stage, follow varying reasons. The most important one is that the ASTEX protonation was created with the knowledge of the protein binding site and the ligand's binding pose, which is especially essential if the binding sites is highly polarized or contains a metal ion influencing the relative stability of different protomers. Examples are carbonic anhydrases (e.g., 1a42) with their typical sulfonamide ligands and some other cases in which oxygen or nitrogen atoms are deprotonated either to coordinate metal ions (e.g., 4-nitrocatechol complexed to protocatechuate 3,4-dioxygenase, 1eoc) or to form salt bridges to positively charged groups in protein side chains (e.g., 2-(4-hydroxy-3-nitrophenyl)acetic-acid bound to N1G9 FAB fragment, 1ngp). In other complexes, salt bridges and/or hydrogen bonds are made possible by protonation of the ligand mainly at aromatic nitrogen atoms interacting with hydrogen acceptors in the protein (e.g., tacrine bound to acetylcholinesterase, 1acj). These examples clearly demonstrate the importance of the protein structure on the ligand's protonation. As shown below, the influence on the docking result is substantial, with the correct (ASTEX) protonation achieving the highest success rates. Thus, the identification of the correct out of all possible protonation states during docking is highly aspired. As already mentioned, the influence of the protein can only be predicted with the knowledge of a specific binding mode. Using the known experimental structure for the optimization of the ligand, as done in other approaches,[15] artificially favors known ligands over potential new ones or inactives, for which a binding pose is not known a priori, and is therefore not further pursued here, even if it is possible to do so using SPORES e.g. for rescoring experiments.

In some other cases, the reasons for the ASTEX protonation are not so clearly visible from the binding site and binding pose. In general the ASTEX protonation tends to prefer protonated aromatic nitrogen atoms and has more often tertiary amines, while the SPORES standard protonation leaves aromatic nitrogen atoms uncharged and creates quaternary amines. Nevertheless, we did not change the definition of the standard protonation accordingly because we think that the present standard protonation reflects the chemical correct protonation for most ligand molecules in

Figure 1. Generation of multiple tautomerics forms of the ring system in a guanine ligand (purine nucleoside phosphorylase, 1ulb).

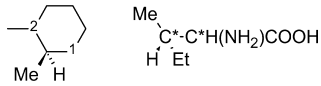


Figure 2. Two examples for the selection of the neighbor trees to be flipped: On the left side, the neighbor trees marked with 1 and 2 are not flipped because they are part of the same ring system as the stereo center. On the right side, the stereo center C_β of the amino acid isoleucine is investigated. In this case, the amino acid backbone will not be chosen to be flipped because it includes another stereo center (C_α atom). None of the other three trees violate a selection rule, and therefore the two smallest ones are chosen. Therefore, in both cases the proton and the methyl group are flipped, which, in these easy cases, would also result from just applying the last rule of taking the smallest substituents.

neutral environment. Additionally, the ASTEX protonation is not consistent for these functional groups, i.e. examples for the protonated and the deprotonated form are among the data set without obvious reasons. Perhaps these inconsistencies are caused by specific preferences of different researchers manually revising the complex structure and can only be resolved by expensive pK_a calculations of the ligands in the active site, which are out of the scope of this publication. Finally, there are some special cases, in which we were not able to find the reasons, why the ASTEX protonation was chosen, e.g. phosphate groups are protonated but do not form hydrogen bonds (MC/PC603 FAB, 2mcp, and thymidylate synthase, 1tdb) to the protein or a hydroxyl group of a carbohydrate is deprotonated to coordinate a metal ion while next to the ligand a protonated amine is present in the binding site (fructose-1,6-biphosphatase, 1frp).

Figure 3 gives some examples of structures with common differences between the ASTEX and the standard protonation. As already stressed, most of the differences can be explained by the influence of the protein (Figure 3a-c), but sometimes also more favorable interactions than in the ASTEX protonation were identified with our standard protonation (Figure 3d). Unfortunately, the automated structure recognition and protonation of SPORES has still some limitations with ligands, in which complex conjugated systems are present and interfere with the protonation. Ligands with not optimal bond angles and length can lead to wrong structures (e.g., Figure 3e). But this is the case only in 5 out of the 213 structures, so that in 97.7% chemical correct ligands are produced. If the minimized ligands also available in the CCDC/ASTEX data set are used, only one wrong structure is generated (99.5% correct). Thus, only a minor decrease in the rate of correct structure generation results, when the ligands are extracted directly from the PDB, demonstrating the small sensitivity of the algorithms on the quality of the input structures, which is remarkable taking into account that the resolution of some complexes in the test set is larger than 2.5 Å. These success rates compare very well to other tools, e.g. I-interpret.[38] If using standard hydrogen adding tools, like the ones implemented e.g. in Openbabel[37] or SYBYL,[45] manual correction of atom types, bond types as well as the number of hydrogen atoms is needed in almost all cases (data not shown). The effectiveness of our new method is additionally supported by the fact that even in the (with great care manual revised) CCDC/ASTEX set some errors in ligand-structure setup (atom and bond type definition) can be observed. During our investigations, we identified 8 complexes with minor errors in the atom or bond types (corrected atom and bond types are provided as Supporting Information in mol2 format), from which one, ligand 4-nitro-benzylphosphonobutanoyl-glycine docked into catalytic antibody IGG2A FAB fragment (1yee), is depicted in Figure 3f. Summarizing, this demonstrates that the automatic setup cannot be a full substitute for a human investigator but that the manual setup needs a lot of experience, patience, and care especially for large data sets, which can be supported by an automatic procedure.

The original ASTEX as well as SPORES standard protonation were then redocked into their native protein structures using PLANTS and GOLD. They show a success rate (highest ranked ligand structure within 2.0 Å heavy atom rmsd from the experimental one) of 80% (170 complexes) and 76% (160 complexes) with PLANTS, respectively. With GOLD a similar picture is obtained: 161 (76%) complexes were correctly docked with the ASTEX protonation and 153 (72%) with the automatically assigned standard protonation. Because of the similar behavior of the two docking programs, we will discuss only the results for PLANTS here in detail and give the corresponding numbers for GOLD of the main findings.

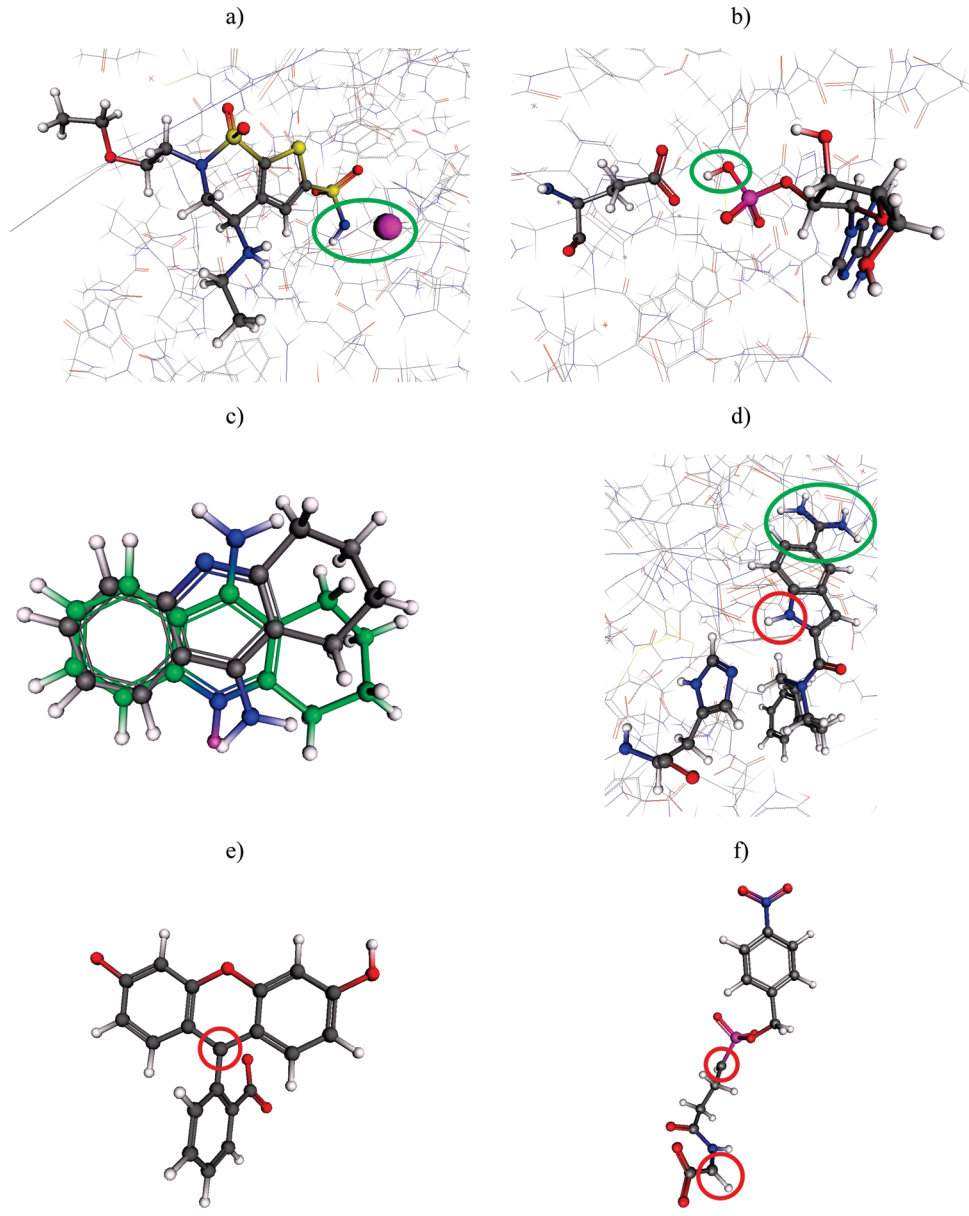


Figure 3. Some cases with differences between the ASTEX and the standard protonation. (a) The ligand of 1a42 (carbonic anhydrase II) is coordinating a zinc ion in the protein binding side (marked in green). Because information of the binding site is taken into account, the deprotonated form, as shown in the figure, is included in the CCDC/ASTEX data set, while the standard protonation uses a protonated sulfonamide. Only with the deprotonated sulfonamide the ligand is docked correctly. (b) The ligand from 6rnt (ribonuclease T1) shows a protonated phosphate group in the ASTEX protonation (marked in green), which is deprotonated in the standard form. This is not essential for the docking pose although the score is better with the ASTEX protonation. (c) Superimposed picture of the best ranked structure of the two protonation states of the ligand from 1acj (acetylcholinesterase). The one protonated at the aromatic nitrogen corresponds to the ASTEX protonation (additional proton is shown in magenta) and is docked correctly, while the other one is rotated nearly 180° to restore the hydrogen bond. (d) The docked structure of complex 1d4p (α-thrombin): While the main interaction is done by the amidinomethyl group (at the top marked in green), the ASTEX (shown in the figure) and the standard protonation differ in the protonation of the nitrogen in the aromatic ring system (marked in red). This leads to a better docking score for the standard protonation due to interaction of the additional proton with the histidine residue shown as ball and stick model on the left. (e) Structure of a fluorescein ligand (FAB, 1flr) in which some double bonds are placed wrongly by SPORES leading to a positively charged carbon atom at the junction of the two ring systems (marked in red). (f) An error in the structures of the CCDC/ASTEX data set (IGG2A FAB, 1yee): The carbon atoms next to the carboxyl and next to the phosphate group (marked in red) are typed as sp^2 but should be sp^3 carbons with two hydrogen atoms.

In most cases, in which the ASTEX protonation is optimized to coordinate a metal ion in the binding site, the standard protonation is not docked correctly or at least achieves an inferior score (e.g., human carbonic anhydrase II, 1a42 and ADP1 protocatechuate 3,4-dioxygenase, 1eoc). The protonated aromatic nitrogen atoms also proved critical in a number of complexes (e.g., acetylcholinesterase, 1acj, and cytochrome P450cam, 1phd). Thus, the correct protonation, which is assumed to be the one stored in the data set, increases the chance of successful docking for both programs tested. The standard protonation performs better than the ASTEX protonation in cases in which no metal ion is present and the ASTEX protonation uses a tertiary amine while the standard protonation uses a quaternary amine (e.g., MAP kinase P38, 1bl7, and TNF receptor associated factor 2, 1d0l). But in all these cases only the score is better while the rmsd value is very similar. Nevertheless, a revision of the data set toward

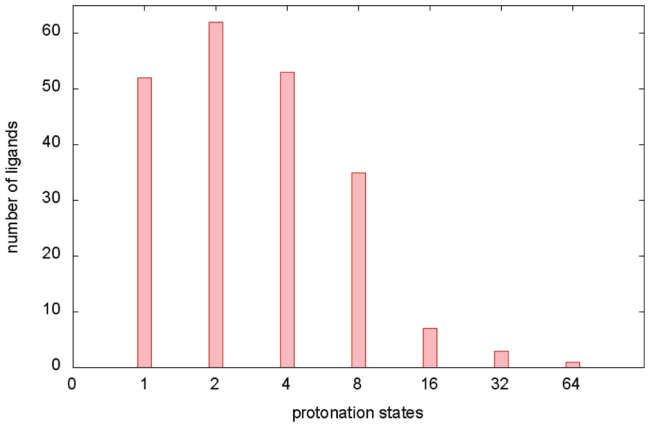


Figure 4. The distribution of the number of protonation states among the complexes.

the standard protonation could also be valuable for these cases (new ligand structures are also given as Supporting Information in mol2 format).

Beside these real improvements in the success rates when using the correct protonation states, in a number of cases one of the docking runs (with ASTEX or standard protonation), gives better results than the other one, even if the two structures (bond and atom types as well as protonation) are exactly the same. This can only be explained by sampling problems, even if we have already performed 25 independent runs to minimize the influence of the stochastic search algorithm. All these ligands are very large and flexible, so that the docking algorithm has problems to find the global optimum. If the optimum is found by one and/or the other protonation (differing only slightly in the location of the hydrogen atoms) is just by chance. In the actual experiments described in this publication, the ligands directly taken from the data set and the ones automatically protonated by SPORES win 6 and 2 times in these sampling-error cases when using PLANTS, respectively. If all complexes in which sampling errors are likely to occur are removed, the success rates increase to 82% (140 of 170 complexes) and 79% (134 of 170 complexes) for the original and the automatically generated data set, suggesting only a small influence of the protonation. Nevertheless, if one looks at the scoring function and the actual rmsd values of all correctly docked cases, many other complexes profit from the correct protonation state as well.

Additional Protonation States. We then continued our investigation by looking at possibilities to include the other protonation states into the docking procedure. All possible protonation states were generated for the ligands of the CCDC/ASTEX clean data set using the described method. This resulted in 1 to 64 different protonation states for the 213 ligands. Table S1 of the Supporting Information gives a summary of the ligands (pdb entries) and the number of protonation states generated, while Figure 4 shows the distribution of the number of protonation states among the ligands.

All protomers are then individually docked into their corresponding protein. If the best ranked structure of each complex is considered without regarding its protonation state, i.e. the structure with the lowest scoring function value taken from the individual docking runs of the same ligand with all generated protonations, 156 (74%) cases are docked correctly with PLANTS and 150 (70%) with GOLD. This reduction of the success rate can be explained by false positive protonation states, in which incorrectly docked structures are scored better because of more hydrogen bridges or more charged interactions. This phenomenon, that scoring functions optimized for pose prediction are not necessary equally well suited to compare different ligands or ligand configurations, is also often seen in virtual screening experiments. Figure 5 shows the docking results of complex recombinant human dihydrofolate reductase with folic acid (1dhf) as an example of a false positive, in which the highly charged, wrongly docked protomer on the left is better ranked than the correctly docked standard protonation on the right.

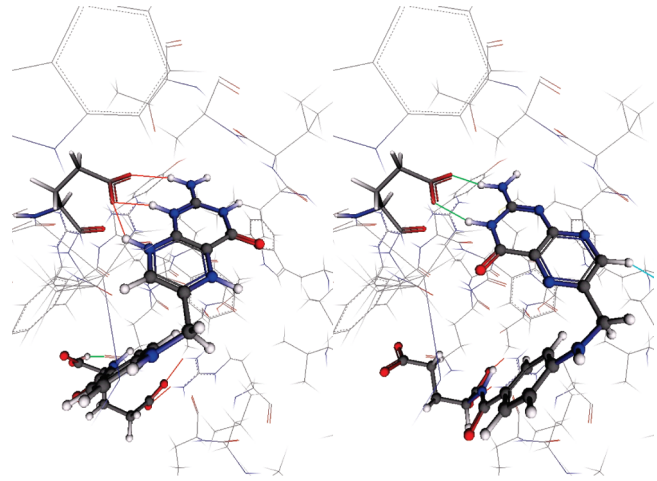


Figure 5. Docking results of 1dhf (dihydrofolate reductase): On the left, the best ranked protonation is shown and, on the, right, the standard protonation, which is identical with the ASTEX protonation. The standard protonation has the correct pose (hydrogen interactions between uncharged donor and charged acceptor shown in green), while the best ranked structure has a wrong pose but is scored better because of more favorable hydrogen-bonding interactions between charged donors and charged acceptors (shown in red).

Many of the best scored incorrectly docked states are highly charged and will not be stable in solution or even in complementary charged active sites. For many additional complexes also a highly charged and, thus, unlikely protomer is ranked best. Even if it is docked correctly, the high score is a fake and would artificially favor the ligands in virtual screening experiments. Therefore, a mechanism must be added to SPORES for removing highly implausible states. One simple possibility is to use the difference between the actual protonation and the standard protonation as a measurement for the stability of the protomer. The standard protonation is designed to represent the ligand in solution and, thus, changing this state is energetically unfavored and must be compensated by stronger intermolecular interactions. The distance to the standard protonation is calculated by comparing each functional group in the ligand separately. Increasing or decreasing the number of hydrogen atoms in such a group increments the difference measure by one, which is then used for an additional term $p = w*d$ in the PLANTS scoring function where p is the penalty added to the scoring value, w is a weighting factor, and d is the difference of the protonation state to the standard protonation. Because the deviation is not dependent on the orientation of the ligand, the term can just be added to the scoring function value obtained in the experiments described above in a

Table 3. Number of Correctly Docked Complexes with PLANTS, Correctly Docked Complexes with Standard Protonation, and ASTEX Protonation Scored Highest Using Different Values Penalizing the Deviation from the Standard Protonation

	no penalty	single penalty	2x penalty	3x penalty	4x penalty	5x penalty	10x penalty
correct	156	157	159	159	159	160	160
standard protonation	90	104	119	127	135	145	160
ASTEX protonation	97	116	128	128	136	142	140

postprocessing step. We used weighting factors of one (single penalty), two (2x penalty), three, four, five, and ten scoring function energy units. E.g. if a protomer under consideration has a protonated carboxyl group (COOH) and a neutral amino group (NH_3), two and four energy units are added to the score for the single and 2x penalty, respectively.

The results of the different penalties are summarized in Table 3. First it becomes evident, that the anticipated effect, that the number of complexes in standard protonation ranked higher than all other protomers increases with increasing penalty, can be reached. This is accompanied by a steady increase in success rate. This shows that even if the ligand is docked wrongly without penalty the energetic difference to the correct complex structure is very low. Thus, the highly implausible, wrongly docked protomers are indeed scored much less favorable when using the penalty function. With a 10x penalty, all complexes are forced into the standard protonation, and the same success rate as obtained when using only this protonation is achieved. If one looks at the number of complexes in ASTEX protonation ranked highest also an increase up to a 5x penalty and then a decrease can be observed. Unfortunately, this maximum in the Astex protonation does not lead to a maximum in the success rates at the same position. As already mentioned, only the success rate of the standard protonation can be obtained. The reason for this is not that the Astex protonation is ranked inferior compared to the standard protonation. In all but 7 cases of the 213 complexes, the ASTEX protonation is identified as the highest ranked one for at least one penalty value. Because of the constant value added for each deviation to the standard protonation, the correct (in the case ASTEX and standard differ) or more likely protonations are equally discriminated than highly implausible ones. Thus, there are still some incorrect protonations or the standard protonation scored highest for small and high penalties, respectively, which leads to the reduced success rates. Examples of these different behaviors can be seen in Figure 6. For further improving the identification of the correct protonation state, the penalty function has to be scaled by an estimate of the energy needed to get to the different states, which can be obtained from pK_a calculations. To get a reasonable size of the data set for the parametrization of this function, additional complexes

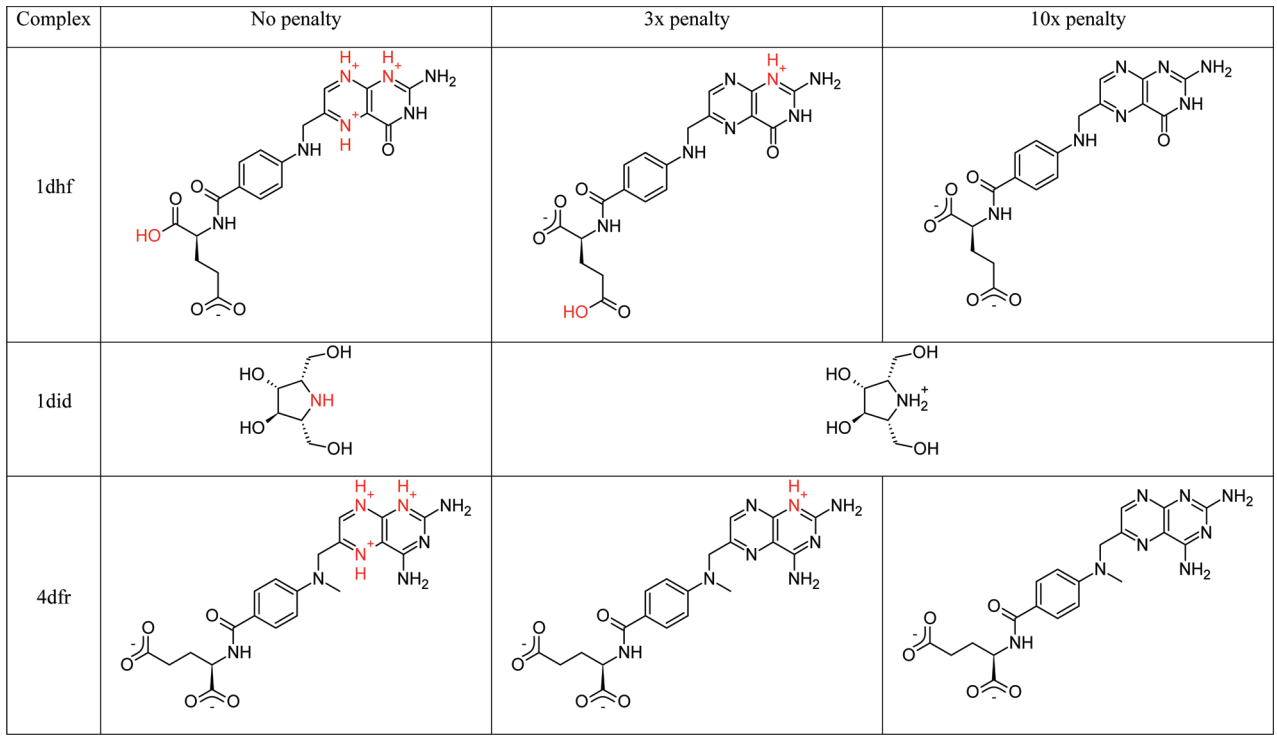


Figure 6. Different behaviors, when adding penalties for differences to the standard protonation (differences in the protonation state with respect to the standard protonation are marked in red): In the first case human dihydrofolate reductase (1dhf), two different implausible protonations are ranked highest and docked incorrectly at penalty values of 0 and 1−4, respectively. Then the correctly docked standard protonation follows, which is identical with the ASTEX protonation. In D-xylose isomerase (1did), the ASTEX protonation is ranked highest without the additional term. But the energy difference of the incorrectly docked standard protonation is very small. Finally, when increasing the penalty, dihydrofolate reductase (4dfr) starts with an implausible (penalty value 0), followed by the ASTEX (penalty values 1−5) and ends with the standard protonation (penalty value 10). In this case, all structures were docked correctly but with ASTEX having the lowest rmsd.

Table 4. Enrichment Factors and AUC Values for the Three Analyses for the Screening of 1mwe (Neuraminidase)[a]

	enrichment factors				
	1%	2%	5%	10%	AUC
standard protonation	10.00 (50)	13.33 (50)	14.66 (20)	9.33 (10)	0.95
complete analysis	12.77 (62.5)	10.51 (49.73)	9.04 (19.89)	5.42 (9.95)	0.82
summarized analysis	13.39 (50)	10.00 (50)	12.00 (20)	8.00 (10)	0.93
single penalty	13.39 (50)	13.33 (50)	12.00 (20)	8.00 (10)	0.94
5x penalty	13.39 (50)	13.33 (50)	12,66 (20)	8.67 (10)	0.94

[a] The optimal enrichment factors are shown in round brackets below the obtained ones.

are needed, in which the standard protonation is not the optimal one. The compilation of such a data set is on its way.

Summarizing it can be stated that the manually generated ligands of the CCDC/ASTEX data set taken separately perform best, which is expected because they were optimized in the context of the nature of the protein binding pocket. Unfortunately, the consideration of all protomers does not lead to a similar success rate. Both docking programs are not able to pick the correct protomers out of all generated ones resulting in a large number of false positives. The additional scoring function term penalizing deviations from the standard protonation is able to remove the implausible protonation states, but it is not possible to identify a penalty value, which always leads to the ASTEX protonation. Therefore, the future step must be to further improve the scoring functions concerning their ability to rank different molecules/states (by e.g. the use of calculated pK_a values) but leaving the good pose prediction accuracy untouched. From this, also virtual screening experiments would benefit (see below).

Virtual Screening with Ligand Protonation States. A virtual screening experiment using the PLANTS software was conducted on the target neuraminidase (1mwe), already used in ref 47, with a data set of 15 known active ligands taken from neuraminidase complexes in the PDB mixed with 735 inactive ligands downloaded from the ZINC database[48] resulting in 2% actives in the test set. The pdb codes of the complexes, from which the active ligands were taken, are available in the Supporting Information. The inactives were chosen to resemble the properties of the actives with an automated procedure according to descriptors like molecular weight, number of hydrogen donors and acceptors, and number and size of rings.[49] First, a cross docking experiment with the active ligands into the chosen receptor showed that all ligands were docked correctly. Then a screening with the ligands standard protonation was conducted. To judge the effectiveness of the screening, enrichment factors $EF_{(\%)}$ for given percentages of the database were calculated. The values for 1%, 2%, 5%, and 10% will be reported. The maximum achievable enrichment factor for each individual case will be given in brackets after the actual obtained one. The results in Table 4 show only a reasonable enrichment. E.g. the enrichment factor at 1% was $EF_{1\%} = 10$ (50) with the standard protonation.

Then, again the influence of different protonation states, this time on the enrichment, was tested. From all ligand structures the protonation states were generated as described above. This leads to 44 structures from the active molecules and 2721 from the inactive library. In total the complete data set contained 1.6% active ligands and their different protonation states. Two different analyses were conducted: In the first one, called complete analysis in the following, all docked structures of all generated protonation states were sorted according to their function value, and the first percentages of this complete data bank were taken for the EF calculation. In the second, the summarized analysis, only the best ranked protonation state of each ligand was considered. The results are given in Table 4, and Figure 7 shows enrichment plots for all virtual screening experiments. The enrichment factors are reasonable with $EF_{1\%} = 13.39$ (50) for the summarized and $EF_{1\%} = 12.77$ (62.5) for the complete analysis. This shows a slight favor for the summarized analysis state over the complete analysis as well as the standard protonation. In the enrichment plots the better results for the summarized analysis becomes even clearer recovering 50% of the active ligand within the first 4% of the database. In the complete analysis, 50% of the actives are recovered only within 6% of the database. 80% of the actives are recovered within 9% of the database for the summarized analysis, while 37% of the database was needed in the complete analysis. For comparison, when only the standard protonation is taken, 50% and 80% of the actives are found within 3.5% and 6% of the database, respectively. Beside the enrichment factors, AUC (area under curve) values[50] which lie usually between 0.5 (random test) and 1 (ideal test) were calculated to estimate the overall performance of the screening. The AUC value for the standard protonation was 0.95, while the summarized analysis of the protonation states reached 0.93. The complete analysis reached only an AUC value of 0.82. This shows that the standard protonation and the summarized analysis perform very similar, but both suffer from the fact that not the correct protonation state is used in the analysis. The reason in the case of the standard protonation is the insensitivity for the influences of the protein and, in the case of the summarized analysis, unreasonably highly charged and unstable states building a large number of charged hydrogen bonds and, therefore, ranked highest. The inferior results of the complete analysis can be attributed to multiple protonation states of inactive ligands scoring higher than the actives. Using the penalty function as described above has also only a very minor, almost negligible effect on the EF and AUC values (see Table 4 and Figure 7). But we hope that the selection of protomers, once it is optimally parametrized in redocking experiments, will also be beneficial for virtual screening with multiple protomers. Summarizing these results, the EF and AUC values are very insensitive to changes in the docking protocol and are, therefore, not good measurements for the accuracy of the algorithms.

Results for Stereoisomers. 39 and 17 ligands of the CCDC/ASTEX clean data set have one or two stereo centers, respectively. This subset of 56 ligands was used to test the stereoisomer generation capability of SPORES. Visual inspection showed that all stereoisomers were correctly

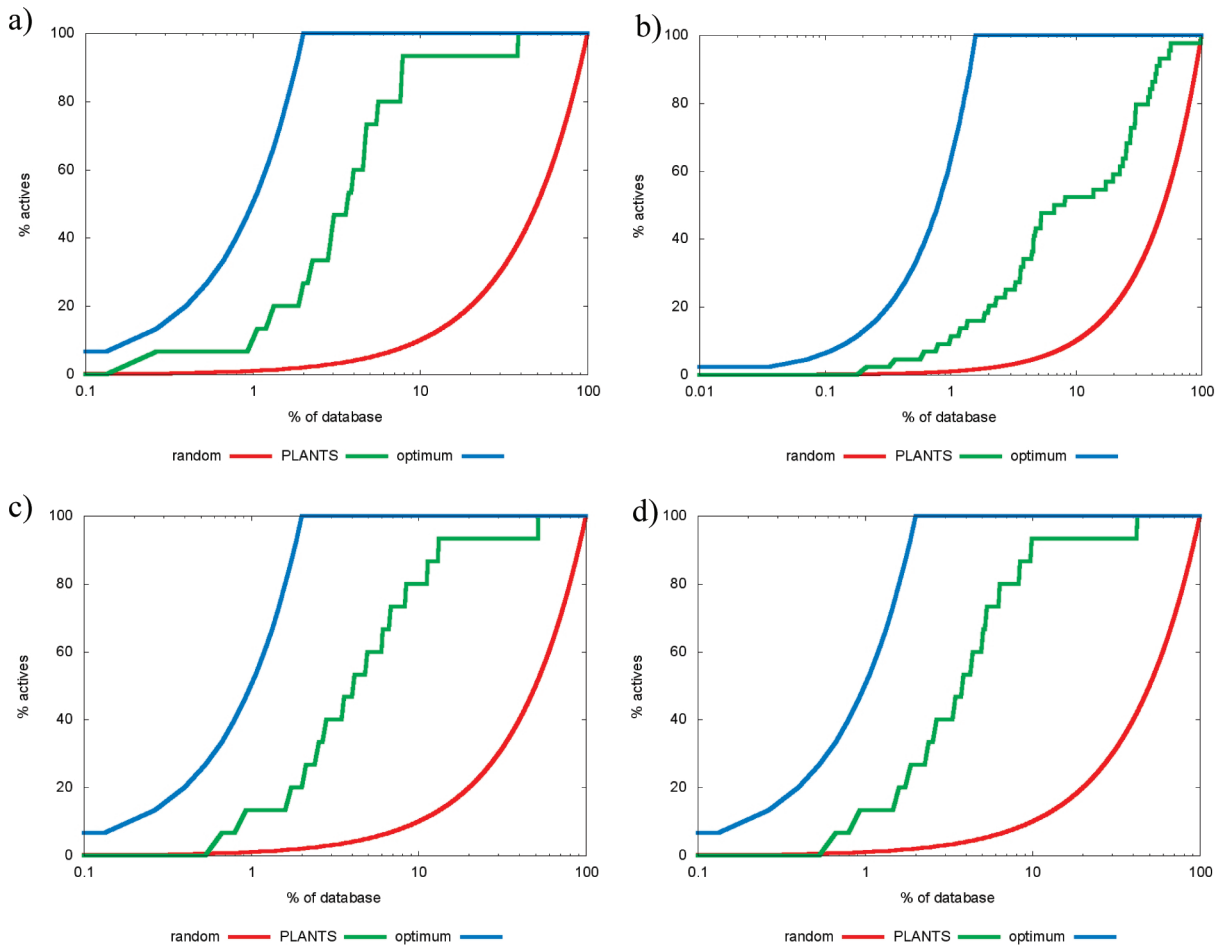


Figure 7. Enrichment plots for the virtual screening experiment: (a) analysis with only the standard protonation of each ligand, (b) analysis with all protonation states (complete analysis), (c) analysis with only the best protonation state of each ligand (summarized analysis), and (d) summarized analysis with a 5x penalty term.

generated with reasonable conformations. Then the influence of the stereoisomers on the docking results obtained with the PLANTS software was analyzed. Here, also the problem like in the virtual screening occurs that only the binding pose of one stereoisomer is known. Therefore, the optimal protonation can only be determined for this stereoisomer based on experimental data. To not artificially favor the correct isomers, we decided to first use the standard protonation for the analysis. Scoring and rmsd values of the lowest energy complex taken from 10 independent docking runs and a visual examination of the binding pose for binding and interaction patterns are used for the evaluation of the docked structures. Even if the rmsd values cannot be used as criterion for docking success for the stereoisomers, which are not included in the experimental structures, they are very useful to see if similar docking poses are possible due to the flexibility of the ligand or only minor differences in the interchanged substituents. The original stereoisomer is only docked in 35 of the 56 cases within 2 Å to the experimental structure, which corresponds to a success rate of 62.5%. For the misdocked complexes, the scoring function is not able to identify the correct pose of the original isomer, and thus it is also not expected that it can differentiate correctly between the different isomers. Therefore, these complexes are not investigated further in the following.

From the correctly docked 35 complexes, in 14 cases the original stereoisomer has a scoring function value (between −88 and −128 energy units for the 14 complexes) of more than 4 energy units lower than the ones of all the other isomers. I.e. the docking algorithm is able to correctly dock the ligand and also to identify the correct stereoisomer. Even if explicit binding data of the stereoisomers is not available to us, in some of the publications for the complexes in this correctly ranked subset information about the relative binding affinity is given. E.g. for 1dwd (human thrombin) a 1000 times larger binding affinity of the correct L-stereoisomer[51] and for 1cbx (carboxypeptidase A) an at least 1 order of magnitude higher affinity[52] are mentioned. Unfortunately, the scoring function is only able to predict the higher affinity of the correct isomer but not the value of the binding affinity differences between isomers, which can be seen by the fact that 4.3 and 11.1 energy units difference are predicted for 1dwd and 1cbx, respectively. For 1ets (bovine thrombin), only one isomer is seen in the crystal even if a racematic mixture is used in the experiment.[53] If one looks at the binding poses, two situations can be differentiated. One example of each is shown in Figure 8. In the first, the binding pattern of the original isomer can be restored by rotation around single bonds, and, therefore, a very similar docking pose can be obtained. In the second example, the stereo center is part of a ring system, and it is not possible to restore the binding pattern resulting in a much lower score.

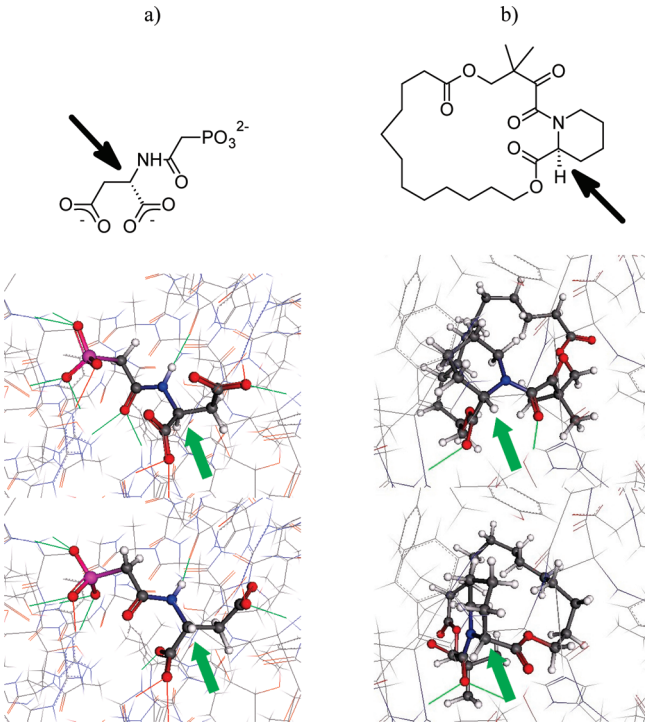


Figure 8. Two examples for the docking of stereoisomers (stereocenters are mark by arrows). Both ligands have two stereoisomers with the original isomer shown on top and the one generated with SPORES below. (a) Ligand of complex 1acm (aspartate transcarbamoylase): The binding pattern of the original stereoisomer (top) can be restored by the generated one. The PLANTS scores are −106.5 for the original and −94.7 for the generated one. (b) Ligand of complex 1fki (FKBP12): The stereo center is part of a ring system, and thus the binding pattern cannot be restored by rotating around single bonds, which results in very different binding poses. In this case the scores are −87.8 for the original isomer and −74.1 for the generated one.

In all except one of the remaining 21 complexes, the scoring function value of the original and the best scored one of the other stereoisomers is less than 3 energy units. This difference is much smaller than the uncertainty in the scoring function, and thus we have to conclude that the scoring function is not able to distinguish between the isomers. For one of the complexes (FAB, 1c12), binding of the other stereoisomer than the one given in the crystal structure is reported, i.e. R,R but probably also S,S is seen in the crystal. In many other examples (1qlq, 1jap, 1lcp, 1lna, 1lst, 1lyl, 1mmq, 2ctc, and 2lgs), the ligand is a natural product or an analogon, often amino acids or derivatives. Here, only the naturally occurring ligands were tested, but probably some activity would remain when inverting the stereo centers. Another reason for the difficulties in scoring is the flip of very small groups, e.g. H and OH. In these cases the differences can be easily balanced by ligand flexibility, resulting in very small rmsd values of all stereoisomers. Only in 1cil (human carbonic anhydrase II), the incorrect isomer gets a score, 8.7 units lower than the original one. This can be explained by the fact that the sulfonamide group in the ligand is not deprotonated in the standard protonation. Therefore, the coordination of this group to the metal ion does not take place so that the docking pose, despite the small rmsd, is wrong. If the ASTEX protonation is used, the difference is decreased to 2.5 units, also placing this complex into the class, where the scoring function is not able to decide on the stereoisomer. Thus, the identification of stereoisomers will probably also benefit from the on-the-fly optimization of the protonation state.

DISCUSSION

The results from the docking experiments with protonation states of the ligands from a standardized test set show that the ligand protonation has a significant influence on the accuracy of docking results. Although the manually revised protonation still outperforms the automatically assigned standard protonation, it is clear that some of the advantages are based on knowledge about the binding pocket and binding pose. SPORES intentionally handles ligands and proteins separately, which is an advantage in the preparation of data sets for virtual screening, where, on one hand, the number of structures prevents manual preparation and, on the other, an equal treatment of active and inactive structures is needed to avoid artificial enrichments. Additional protonation states lead to a reduction of the success rate on this well prepared test set. These problems are mainly caused by the generation method, which does not discriminate between favorable and unfavorable protonation states. Together with the scoring functions this leads in some cases to overfavoring of highly charged ligands because of the possibility of forming charged interactions to the protein thus creating false positive docking results. When comparing the two used docking programs PLANTS and GOLD it is noticeable that PLANTS gives an overall better performance, but the influence of the protonation states is about the same for both programs despite the different approaches in terms of scoring function.[20,23]

In the virtual screening experiments, the enrichment factors for the summarized analysis and the screening with only the standard protonation are directly comparable. For the complete analysis of all protonation states the different ratio of actives to the total number of ligands has to be notified. This leads to higher absolute enrichment factors, but the ratio between the maximum enrichment and the actual enrichment factor is lower. Additionally, the enrichment plots show clearly that the summarized analysis gives better results and is comparable to the screening in which only the standard protonation was used. This highlights once again the problems, which have become obvious in the study conducted on the CCDC/ASTEX data set, that highly charged false positives will be ranked high. Nevertheless, the inactive part of the data set was chosen only to one-dimensional molecular descriptors for similarity, and thus it is possible that some of them are actually active structures, which could have a negative effect on the enrichment.

The stereoisomers do not give such a clear result mainly because experimental data for all stereoisomers (binding constants and complex structure) are not available. Also statistical influences with an increasing number of stereoisomers have to be taken to account. In both cases, protonation states and stereoisomers, the standard structure, standard protonation or original stereoisomer, is more often best ranked if only two different states than if four or more different ones are available. The cases in which the original stereoisomer is not ranked best are mainly flexible molecules or larger molecules in which the inversion of one center only affects a small part of the molecule like the interchange of a hydrogen atom and a methyl, carboxyl, or slightly larger

group. The cases with a good scoring discrimination between the original stereoisomers and the others fall into three categories: smaller molecules where the stereoisomers have differences in hydrogen bonding patterns and salt bridges, less flexible molecules where the whole binding pattern can only be fulfilled by the original stereoisomer, or molecules in which large parts are interchanged during the inversion of stereo centers and thus the binding patterns of the different stereoisomers have large varieties.

CONCLUSION AND OUTLOOK

The automatic ligand preparation is particularly useful for the setup of virtual screening data sets, where the amount of molecules is too high for manual revision of the ligands. SPORES does not need any information about the binding site to create chemical correct ligand structures and is therefore well suited for this task avoiding preferential treatment of the active structures in the screening setup, especially the high accuracy of the generated structures (97.7%) positions SPORES into the top class of today's state-of-the-art setup procedures for docking experiments. Additionally, besides the most probable protonation as defined by a rule based system, SPORES can also generate different protonation and tautomeric states as well as stereoisomers.

The inclusion of different ligand protomers slightly reduces the success rates in docking experiments as well as the enrichment in virtual screening at the moment. But these effects are observed on very well-defined test sets, where the ligand's optimal protonation is either known (in the CCDC/ASTEX data set) or which have been shown to be very well suited for docking using the standard protonation (the target 1mwe was chosen for virtual screening because of the good cross docking results obtained with this protein and the 15 active ligands). But the inclusion of additional protonation states may prove useful when less information is available about the targets or when the performance of the standard protonation is poor for the specific target. The main drawback, the creation of unreasonable protonation states, can be improved by either a fast incremental pK_a calculation or by a rule based approach to avoid too highly charged ligand structures. Until such a procedure is implemented, our recommendation is to use only the standard protonation of SPORES in docking and especially in virtual screening experiments. Only in the case that for a specific target the experimental results cannot be reproduced in redocking experiments, additional protonation states should be produced with SPORES and docked after removing the implausible ones. But, we hope that we can modify these recommendations when a successful automatic selection of possible protonation states and improved scoring functions are available.

As with the protonation states, the stereoisomers show potential for virtual screening because structure databases, like ZINC, do not provide explicit stereoisomers. One cannot be sure that the isomer, included in the database or which is automatically generated, is the active one for a specific target. If only limited information is available about the binding ligands and their absolute stereochemistry is unknown, all possible stereoisomers of the active structures have to be considered for virtual screening, which can be automatically done using SPORES. Even if the identification of the correct stereoisomer is not perfect, our recommendation is to use all stereoisomers in virtual screening. In contrast to protonation, in the cases where the wrong isomer is identified the scoring function values of the stereoisomers have only minor differences. Therefore, in these cases it is highly probable that both stereoisomers are identified as active and the racemate can be tested experimentally. But in the test examples where the correct isomer is identified, the energy difference can be up to more than 17 energy units. Therefore, active ligands can be missed if only the incorrect stereoisomer is used, which is highly undesired. The number of possible stereoisomers is much smaller than the one of protomers, so that the computational overhead is acceptable.

ACKNOWLEDGMENT

The authors thank Oliver Korb for many constructive discussions as well as his help regarding the PLANTS program and the preparation of the data sets. Jens Gimmler is acknowledged for the careful proofreading of the manuscript. We thank the CUSS cluster and the bwGRiD in Ulm as well as the HPC cluster of the University of Konstanz for allocating computational resources. Additionally, our thanks go to the anonymous reviewers for their helpful suggestion for improving the paper.

Supporting Information Available: pdb codes of the complexes taken from the CCDC/ASTEX data set with the corresponding number of protomers and stereoisomers as well as the ones used to extract the active ligands for the virtual screening set and, additionally, new structure files for ligands, where we have found problems in the CCDC/ASTEX data set. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Drug Discovery* **2004**, *3*, 935–949.
(2) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.
(3) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235–249.
(4) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.* **2002**, *41*, 2644–2676.
(5) Krovat, E. M.; Steindl, T.; Langer, T. Recent advances in docking and scoring. *Curr. Comput.-Aided Drug Des.* **2005**, *1*, 93–102.
(6) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151–166.
(7) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin; Th, E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269–288.
(8) DesJarlais, R. L.; Sheridan, R. P.; Seibel, G. L.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.* **1988**, *31*, 722–729.
(9) Gschwend, D. A.; Kuntz, I. D. Orientational sampling and rigid-body minimization in molecular docking revisited: On-the-fly optimization and degeneracy removal. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 123–132.
(10) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
(11) Rarey, M.; Wefing, S.; Lengauer, T. Placement of medium-sized molecular fragments into active sites of proteins. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 41–54.

(12) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
(13) Zavodszky, M. I.; Sanschagrin, P. C.; Korde, R. S.; Kuhn, L. A. Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 883–902.
(14) Jain, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46*, 499–511.
(15) *Glide, version 4.0*; Schrödinger, LLC: New York, NY, 2005.
(16) 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.
(17) 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.
(18) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
(19) Jones, G.; Willett, P.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(20) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.
(21) Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins* **1990**, *8*, 195–202.
(22) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(23) Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of ant colony optimization to structure-based drug design. In *Ant Colony Optimization and Swarm Intelligence, 5th International Workshop, ANTS 2006, LNCS 4150*; Dorigo, M., Gambardella, L. M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., Eds.; pp 247−258.
(24) Korb, O.; Stützle, T.; Exner, T. E. An ant colony optimization approach to flexible protein-ligand docking. *Swarm Intell.* **2007**, *2*, 115–134.
(25) Korb, O.; Stützle, T.; Exner, T. E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49*, 84–96.
(26) Totrov, M.; Abagyan, R. Flexible protein ligand docking by global energy optimization in internal coordinates. *Proteins* **1997**, *1*, 215–220.
(27) McMartin, C.; Bohacek, R. S. QXP: Powerful, rapid computer algorithms for structure-based drug design. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 333–344.
(28) Westhead, D. R.; Clark, D. E.; Murray, C. W. A comparison of heuristic search algorithms for molecular docking. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 209–228.
(29) Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins* **1998**, *33*, 367–382.
(30) McGann, M.; Almond, H.; Nicholls, A.; Grant, J. A.; Brown, F. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76–90.
(31) Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of docking performance: Comparative data on docking algorithms. *J. Med. Chem.* **2004**, *47*, 558–565.
(32) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graphics Modell.* **2003**, *21*, 289–307.
(33) Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A new test set for validating predictions of protein-ligand interaction. *Proteins* **2002**, *49*, 457–471.
(34) Hartshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T. M.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein-ligand docking performance. *J. Med. Chem.* **2007**, *50*, 726–741.
(35) *Maestro, version 7.5*; Schrödinger, LLC: New York, NY, 2006.
(36) Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 187–205.
(37) Guha, R.; Howard, M. T.; Hutchison, G. R.; Murray-Rust, P.; Rzepa, H.; Steinbeck, C.; Wegner, J. K.; Willighagen, E. The Blue Obelisk - Interoperability in Chemical Informatics. *J. Chem. Inf. Model.* **2006**, *46*, 991–998.
(38) Zhao, Y.; Cheng, T.; Wang, R. Automatic perception of organic molecules based on essential structural information. *J. Chem. Inf. Model.* **2007**, *47*, 1379–1385.
(39) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–240.
(40) Polgar, T.; Magyar, C.; Simon, I.; Keserue, G. M. Impact of ligand protonation on virtual screening against beta-secretase (BACE1). *J. Chem. Inf. Model.* **2007**, *47*, 2366–2373.
(41) Brooks, W. H.; Daniel, K. G.; Sung, S. S.; Guida, W. C. Computational validation of the importance of absolute stereochemistry in virtual screening. *J. Chem. Inf. Model.* **2008**, *48*, 639–645.
(42) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an integrated description of hydrogen bonding and dehydration: decreasing false positives in virtual screening with the HYDE scoring function. *ChemMedChem* **2008**, *3*, 885–897.
(43) Clark, M.; Cramer, R. D.; Van Opdenbosch, N. Validation of the general purpose tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(44) Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. Relibase - Design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.* **2003**, *326*, 607–620.
(45) *SYBYL 7.2*; Tripos Inc.: 1699 South Hanley Road, St. Louis, Missouri 63144, USA, 2008.
(46) Milletti, F.; Storchi, L.; Sforna, G.; Cross, S.; Cruciani, G. Tautomer enumeration and stability prediction for virtual screening on large chemical databases. *J. Chem. Inf. Model.* **2009**, *49*, 68–75.
(47) Birch, L.; Murray, C. W.; Hartshorn, M. J.; Tickle, I. J.; Verdonk, M. L. Sensitivity of molecular docking to induced fit effects in influenza virus neuraminidase. *J. Comput.-Aided Mol. Des.* **2003**, *16*, 855–869.
(48) Irwin, J. J.; Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Comput. Sci.* **2005**, *45*, 177–182.
(49) Korb, O.; Stützle, T.; Exner, T. E. Structure-based virtual screening using PLANTS. 2009, . in preparation.
(50) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J. P.; Bertrand, H. O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* **2005**, *48*, 2534–2547.
(51) Banner, D. W.; Hadvary, P. Crystallographic analysis at 3.0 Å resolution of the binding to human thrombin of four active site-directed inhibitors. *J. Biol. Chem.* **1991**, *266*, 20085–20093.
(52) Mangani, S.; Carloni, P.; Orioli, P. Crystal structure of the complex between carboxypeptidase A and the byproduct analog inhibitor L-benzylsuccinate at 2.0 Å resolution. *J. Mol. Biol.* **1992**, *223*, 573–578.
(53) Brandstetter, H.; Turk, D.; Hoeffken, H. W.; Grosse, D.; Stuerzebecher, J.; Martin, P. D.; Edwards, B. F. P.; Bode, W. Refined 2.3 Å x-ray crystal structure of bovine thrombin complexes formed with the benzamidine and arginine-based thrombin inhibitors NAPAP, 4-TA-PAP and MQPA. A starting point for improving antithrombotics. *J. Mol. Biol.* **1992**, *226*, 1085–1099.

CI800420Z