Impact of Ligand Protonation on Virtual Screening against β-Secretase (BACE1)

Tímea Polgár,† Csaba Magyar,‡ István Simon,‡ and György M. Keserü*,†

Gedeon Richter Plc, P.O. Box 27, H-1475 Budapest, Hungary, and Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, P.O. Box 7, H-1518 Budapest, Hungary

Received June 22, 2007

Structure-based virtual screens were carried out against β-secretase (BACE1) to investigate the impact of ligand protonation on screening efficacy. A comparative evaluation of the performance and its dependence on ligand protonation states docking by Surflex, eHiTS, GOLD, and FlexX-Pharm was performed. Virtual screening performed by FlexX-Pharm ($EF_{(1\%)}$=69) and Surflex ($EF_{(1\%)}$=58) provided the best efficiency. Screening protocols by FlexX-Pharm and GOLD were affected by ligand protonation, while performance of Surflex did not depend on ligand protonation.

INTRODUCTION

β-Amyloid peptide is central to the pathophysiology of Alzheimer's disease. Amyloid plaques composed of β-amyloid peptides develop in the brains of Alzheimer's disease patients; therefore, therapeutic strategies to decrease the concentration of β-amyloid peptide in the brain are highlighted.[1,2] β-Secretase (BACE) is a proteolytic enzyme involved in the cleavage of amyloid precursor protein (APP). The proteolysis of APP by β-secretase is followed by subsequent C-terminal γ-secretase cleavage resulting from the neurotoxic Aβ42, β-amyloid peptide. Thus β-secretase is a potential target in the therapy of Alzheimer's disease.[3−5]

So far 32 X-ray structures of β-secretase complexed with peptidomimetic or small molecule inhibitors have been released.[6−19] This enabled a structure-based drug design to shed light on the understanding of molecular recognition between small molecule and peptidomimetic inhibitors and the protein,[20] drawing attention to the conformational changes in the protein matrix[21,22] and protonation states within the active site.[23−26]

Several reports have emphasized the importance of protonation states during structure-based design of BACE1 inhibitors, and some teams have attempted to determine the protonation states of catalytic residues. Huang et al. used protonated Asp-228[25,27] for their calculations, while others suggested Asp-32[23,24,26] to be protonated. Semiempirical, molecular dynamical, docking, and QM/MM calculations were applied to assign the protonation states of the key aspartates in BACE1.[23−26] However, none of them dealt with ligand protonation states in detail.[27−30]

We demonstrated that an effective virtual screening protocol could be developed if Asp-32 was protonated; nevertheless, the effect of ligand protonation states was not studied.[24,31] Despite the extensive investigation of ionization states in BACE1 the effect of ligand protonation remained unexplored. Analyzing the active site of BACE1, however, a significant impact of ligand protonation on screening efficacy can be expected. Recently Knox[32] et al. showed on estrogen receptor alpha that ligand preparation—tautomerism, protonation, stereochemical, and conformational states—had not affected the outcome of virtual screening studies. However, their findings cannot be generalized and should be explored with several targets, scoring functions and docking algorithms. Warren and co-workers screened their ionized ligand databases against 8 different proteins with 10 docking programs and 37 scoring functions. Although the docking programs reproduced experimental binding modes, none of the scoring functions were able to discriminate between highly potent and inactive compounds.[33]

Here, based on our previous results,[24,31] we attempted to investigate the impact of ligand protonation on the results of virtual screening. A comparative study was carried out against 1W51 BACE1 X-ray structure—as the most effective virtual screening protocol was developed against 1W51—by GOLD, FlexX-Pharm, eHiTS, and Surflex. All protomers for each compound were docked and scored, then one representative was selected and used for ranking and enrichment calculations. Our primary objective was to show the dependence of enrichment on protonation states for every docking procedure. Second, through the comparison of virtual screening protocols developed by the various docking algorithms and scoring functions we estimated the ability and performance of the procedures to tackle the problem of ligand protonation.

METHODS

Screening Library. Our screening library includes a subset of World Drug Index (WDI) as inactive molecules that were specifically designated to reduce artificial enrichment.[34,35] WDI was filtered to eliminate compounds outside the range of molecular weight 200 and 800, having clogP larger than 7 and rotatable bonds more than 15. The remaining WDI compounds were subjected to diverse selection based on 2D UNITY[36] fingerprints. Dissimilarity selection performed by the Selector module of Sybyl 7.2[36] resulted in 9950 compounds with a maximum Tanimoto index of 0.69 that was defined as an inactive set. Our active set was compiled by diverse selection—based on 2D UNITY

* Corresponding author phone: +36-1-431-4605; fax: +36-1-432-6002; e-mail: gy.keseru@richter.hu.
† Gedeon Richter Plc.
‡ Hungarian Academy of Sciences.

10.1021/ci700223p CCC: $37.00
Published on Web 10/18/2007

fingerprints—of 50 BACE inhibitors from the total of 218 compounds available in the Prous Integrity Drugs&Biologics database.[37] The maximal Tanimoto similarity index was 0.46 within this set. This library of 10 000 compounds has an active content of 0.5% that mimics real-life screening situations.

All the possible protonation states were generated using Protoplex implemented in Sybyl 7.2. In total 63 509 different protonation states of 10 000 compounds were generated. Protomers included protonation states formed at pH 6.6 at which crystals of 1W51 were grown. Protonation states at pH 4.5 at which optimal BACE1 activity can be reached were also calculated. All data sets are available upon request.

Target Preparation. Enrichment studies were performed on a ligand bound (1W51[17]) structure of β-secretase. Recently released X-ray structures were used for structural comparison. Protonation states in 1W51 were determined at the corresponding pH of crystallization (pH 6.6) using the titration curves calculated by ZAP,[38] thus Asp-32 was protonated while Asp-228 was ionized. The active site of 1W51 was defined as the collection of residues within 6.5 Å of the bound inhibitor. The bound inhibitor was not included in the docking run.

Docking Protocols. *FlexX-Pharm.* Docking by FlexX-Pharm[36] was performed using standard parameters as implemented in the Sybyl 7.2 package. This setup was able to reproduce the structure of druglike BACE1 inhibitor complexes ($n = 20$) with the average rmsd value of 0.74 Å (see the Supporting Information). Thirty docking solutions (poses) for each docked molecule were scored and saved for further analysis. All stored poses were rescored using the CScore module of Sybyl 7.2 comprising three different scoring functions including Dock, Chem, and Gold. ScreenScore was used during ligand building.

For the evaluation of results of FlexX-Pharm using three different scoring functions we first extracted the best poses from the 30 saved ones generated for all compounds in the screening database. Then compounds with the lowest scores were ranked by the same three scoring functions, thereby the pose extraction and ranking were separated resulting in 9 enrichment factors for each docking run. Scoring functions used for pose extraction and ranking are represented successively.

Pharmacophore constraints applied in FlexX-Pharm involved optional interaction constraints for -COO^{Asp32}, -COO^{Asp228}, -CO^{Gly34}, and -CO^{Gly230} groups derived from the pharmacophore hypothesis of Vertex.[20] Accepted poses fulfilled at least two of these constraints simultaneously.

Surflex-Dock. Default options as implemented in Surflex versions 1.33 and 2.03 were used for docking calculations.[36,39−41] This setup was able to reproduce the structure of druglike BACE1 inhibitor complexes ($n = 20$) with the average rmsd value of 3.41 Å (see the Supporting Information). During the protomol generation the proto_bloat parameter was set to 2, thereby an extended active site was made. While version 1.33 does not include postdock minimization capabilities, version 2.03 was used with the "-remin" option, thus a postdock minimization procedure was also applied using BFGS quasi-Newton method and an internal Dreiding force field. Poses were scored by Surflex's built-in scoring function. The default whole molecule construction method was used with the "-multistart 5" option.

eHiTS. eHiTS version 6.2[42] was run following a filtering step. The filter was trained by 9354 decoys derived from WDI. There was no overlap between our test set comprising 9950 molecules and our 9354-membered training set. Four hundred fifteen BACE1 inhibitors were used for filter training as active compounds. Fifty BACE1 inhibitors involved in test calculations were excluded from this 415-membered set. Filter training is not a prerequisite of eHiTS docking, but it was applied to reduce computation time. 1%, 2%, and 5% of the initial database was filtered and then docked using eHiTS. Docked poses were scored by eHiTS scoring function.

GOLD. Docking by GOLD[43−45] 3.1.1 was performed using default parameters, automatic settings with 100% search efficiency. The number of GA runs was set to 6. Docked poses were scored by the GOLDScore implemented in the software.

Enrichment Factor Calculation. The effectiveness of different screening protocols has been investigated in enrichment studies aimed at identifying known actives from the screening library. Enrichment factors were used to assess the quality of the rankings in each docking run

$$EF_{(\%)} = \frac{(N_{active(\%)}/N_{(\%)})}{(N_{active}/N_{all})}$$

where $EF_{(\%)}$ is given as the percentage of the ranked database, $N_{active(\%)}$ is the number of active compounds in a selected subset of the ranked database, $N_{(\%)}$ is the number of compounds in the subset, and N_{active} and N_{all} are the number of active molecules and the number of compounds in the screening database. Enrichment factors were not scaled, i.e., absolute values depend on the ratio of active and inactive molecules and should be compared to the maximum (ideal) achievable EFs. EFs were calculated at 1% of the ranked database ($EF_{(1\%)}$) as a fixed threshold that enables the comparison of enrichment factors when studying the effectiveness of different screening protocols.

Estimation of Dependency of Enrichment Factors on Ligand Protonation. All protomers were docked into the active site by FlexX-Pharm, Surflex, and GOLD. For enrichment factor calculation one protomer was required to be selected. By eHiTS only neutral compounds were docked as the algorithm checks all possible protonation states for both the ligands and the protein. Protomer selection was made based on score values; the best and the worst score values were selected, and average score values were calculated within the groups of protomers. Then enrichment factors were calculated using the best, the worst, and average scoring protomers. Scores calculated for compounds protonated at pH 4.5, pH 6.6 and neutral compounds (compounds having no charged groups) were also used for enrichment factor calculation.

RESULTS AND DISCUSSION

Previously we tested the effect of protein conformation and protonation of the BACE1 active site on screening efficacy.[24,31] A total of nine public domain X-ray structures of BACE1 were clustered and analyzed in detail. Presently available structures (32 structures, June 2007) form four distinct clusters based on main chain conformations. The first

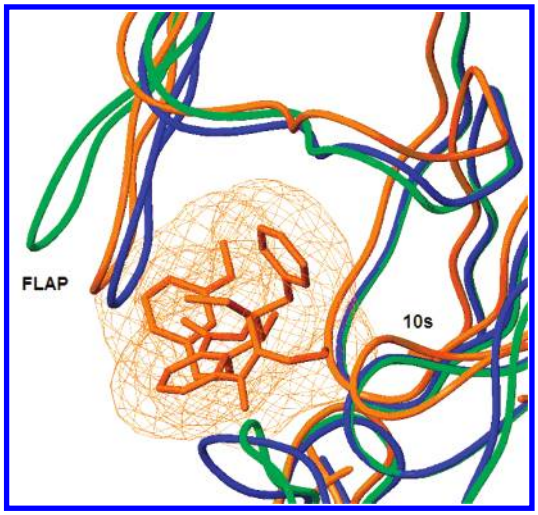

Figure 1. Active sites of BACE1 X-ray structures: 1FKN (blue), 1SGZ (green), and 1W51 (orange).

cluster involves structures 1W50, 2OF0, 2OHK, 2OHL, 2OHM, 2OHN, 2OHP, 2OHQ, 2OHR, 2OHT, and 2OHU where the 10s (9-14) and the FLAP (67-77) loop adopts open conformation. The second cluster is represented by 1W51, 1TQF, 2B8L, 2B8V, 2F3E, 2FDP, 2HIZ, 2HM1, 1YM4, 2IRZ, 2ISO, and 2PH6 where the FLAP loop has a closed conformation, while the 10s loop has an open conformation. Structures 1FKN, 1M4H, 1XN2, 1XN3, 1XS7, 2F3F, 2G94, and 1YM2 are members of the third cluster with a closed FLAP and a 10s loop. 1SGZ and 2OHS belong to the fourth cluster where the 10s loop is closed and the FLAP loop is an open conformation. X-ray structures can be classified into four main groups based on the existence and type of the bound ligands: apostructures (no ligands), fragments, small molecule inhibitor complexes, and peptidomimetic inhibitor complexes. Using apostructures—in which FLAP is open—usually provides lower enrichment because current docking algorithms treat the protein as rigid structures; therefore, conformational changes upon ligand binding are not considered. Similarly, in fragment complexes the FLAP loop remained open. Using X-ray structures with the 10s loop open resulted in improved enrichment.[31] However, this observation should be further explored, and we here restrict ourselves to only mention it as the analysis of its role is beyond the scope of this study. It is interesting to note that McGaughey et al.[21] suggest that the 10s loop is modulated via an induced fit not via direct interactions with catalytic aspartic residues underlining the importance of protein flexibility in virtual screening.

Since cluster 1 contains one apostructure (1W50) and several fragment complexes, the opened FLAP loop discloses this cluster from rigid protein virtual screening. One representative 3D structure from cluster 2−4 (Figure 1) was investigated, and against each of them a virtual screening protocol was developed. The best enrichment factors are summarized in Table 1. Based on these results virtual screening against 1W51 is preferred. Considering that pharmaceutical research programs mainly aim at the identification of non-peptid or non-peptidomimetic BACE1 inhibitors, virtual screening ought to be developed for complexes with small molecule inhibitors such as 1W51.

Virtual screens were carried out by four docking algorithms to study the effect of ligand protonation states. All

Table 1. Screening Efficacy Obtained against Different BACE1 X-ray Structures by FlexX-Pharm

PDB code	resolution (Å)	residues	$EF_{(1\%)}$
1FKN	1.9	1−385	37
1SGZ	2.0	1−385	39
1W51	2.55	43−453	55

Table 2. Summary of Enrichment Factors Calculated at 1% of the Database ($EF_{(1\%)}$), Ranked by GOLDScore and Protein−Ligand van der Waals Energy (S(vdW_ext))

	$EF_{(1\%)}$	
	GOLDScore	S (vdW_ext)
minimum	10	36
maximum	2	30
average	10	40
neutral	24	36
pH 6.6	20	36
pH 4.5	20	40

possible protonation states generated by Protoplex were docked into the 1W51 structure by FlexX-Pharm, GOLD, and Surflex-Dock. By eHiTS only neutral compounds were docked as the algorithm checks all possible protonation states for both ligands and the protein. eHiTS systematically evaluates all possible protonation states for the receptor and ligands, automatically for every receptor−ligand pair. It does this approach through the use of ambiguous properties flags for positions that could be either protonated or deprotonated. Each state is evaluated and scored.

In this study three main factors were analyzed. First, dependence of enrichment factors on various protonation states was studied using the worst, average, and the best scoring protomers. Hereby we attempted to demonstrate if there is any impact of using various protomers on screening efficacy. Second, protonation states were calculated by LigPrep at pH 6.6, at the pH of crystallization of the 1W51 X-ray structure, and also at pH 4.5, at the pH optimum of BACE1 activity. Scores contributing to these ionization states were used for enrichment factor calculation. Third, scores obtained for the neutral compounds, which had been used in our previous paper,[24,31] were used for enrichment factor calculation, as a reference point for comparison.

Next we showed how ligand protonation affected the results given by four docking algorithms: FlexX-Pharm, GOLD, Surflex, and eHiTS. Moreover the performance and its dependence on ligand protonation states of these algorithms was also compared.

Screening by GOLD. The virtual screen carried out by GOLD[45] version 3.1.1 resulted in acceptable enrichment factors. Ranking by GOLDScore yielded a maximal enrichment factor of $EF_{(1\%)} = 24$. However, when poses were ranked by protein−ligand van der Waals energy values (S(vdW_ext)), a radical improvement of $EF_{(1\%)}$ to 40 was observed (Table 2).

Since all protomers were docked, one representative for each molecule with the best, the worst, and the average score value was used for enrichment factor calculation to estimate the dependence of the enrichment factors on protonation states. Comparing enrichment factors calculated for minimum and maximum values of GOLDScores and of protein−ligand van der Waals energy values for each ligand demonstrates that enrichment factors seem to be better for minimum values

than for maximum values. The worse the score is the better the rank is; this is contradictory to our expectations. This anomaly can be explained by improper scoring due to incorrect protonation states; however, its effect was found to be different for ranking by GOLDScore (5-fold) and S(vdW_ext) (1.2-fold). GOLDScore comprises four main terms: protein−ligand hydrogen-bonding energy, protein−ligand van der Waals energy (S(vdW_ext)), ligand internal van der Waals energy, and torsional strain energy. The effect of protonation was obvious when all four terms were included in the ranking function while nearly disappearing for S(vdW_ext). There is a conspicuous difference between $EF_{(1\%)}$s calculated using S(vdW_ext) and GOLDScore. $EFs_{(1\%)}$(S(vdW_ext)) are 4-fold or 15-fold better than $EFs_{(1\%)}$-(GOLDScore).

Ranking by average scores provided similar (or somewhat better) enrichment factors to that given by minima (the worst scores) suggesting that scores of incorrect protonation states tend to distribute closer to the worst score values.

Enrichment factors calculated for ligands protonated at pH 6.6 were improved; for GOLDScore $EF_{(1\%)}$ was 20 and for S(vdW_ext) it was 36. $EF_{(1\%)}$s were doubled for GOLDScore, and the ratio of $EF_{(1\%)}$(S(vdW_ext)) to $EF_{(1\%)}$(GOLDScore) was abated approximately to two. These suppose that compounds protonated at the pH of crystallization tend to score better and can be a better approximation of ligand binding. Enrichment factors calculated for neutral ligands provided comparable results to that given for ligands protonated at pH 6.6. Similar considerations can here also be made to the latter case. However these latter two coincident results were not unexpected as screening compounds were assumed to be mainly neutral at pH 6.6. If ligands were protonated at pH 4.5, which is the pH optimum of BACE activity, $EF_{(1\%)}$s were comparable to that calculated for neutral molecules or compounds protonated at pH 6.6. Ligands protonated at pH 4.5 or neutral compounds lead to a similar ranking; however, GOLDScore was not able to discriminate between these.

It is interesting to note that the van der Waals energy term could rank the molecules significantly better than the GOLDScore, revealing the importance of this contribution to BACE1 inhibitor binding.[27−30] Enrichment factors when ranked by S(vdW_ext) contributing to minima, maxima, or average values did not differ significantly. Moreover ranking of neutral ligands or ligands protonated at pH 4.5 and pH 6.6 resulted in similar enrichment factors to the latter three also. These suggest that protonation states do not affect remarkably the rank provided by a protein−ligand van der Waals energy term.

Screening by FlexX-Pharm. Virtual screening by FlexX version 1.20 was carried out using pharmacophore constraints. A virtual screening protocol was optimized in terms of scoring functions and pharmacophore constraints in our previous study. Here only the best virtual screening protocols, i.e., applying pharmacophore constraints according to the Vertex's pharmacophore and using Dock, Gold, and Chem scores for both extracting poses and ranking, were investigated. This scoring strategy resulted in nine different combinations of the three scoring functions.

Enrichment factors calculated using minima (the best) or maxima (the worst) scores $EF_{(1\%)}$s varied among the nine cases (Figure 2). This denotes that ligand protonation can

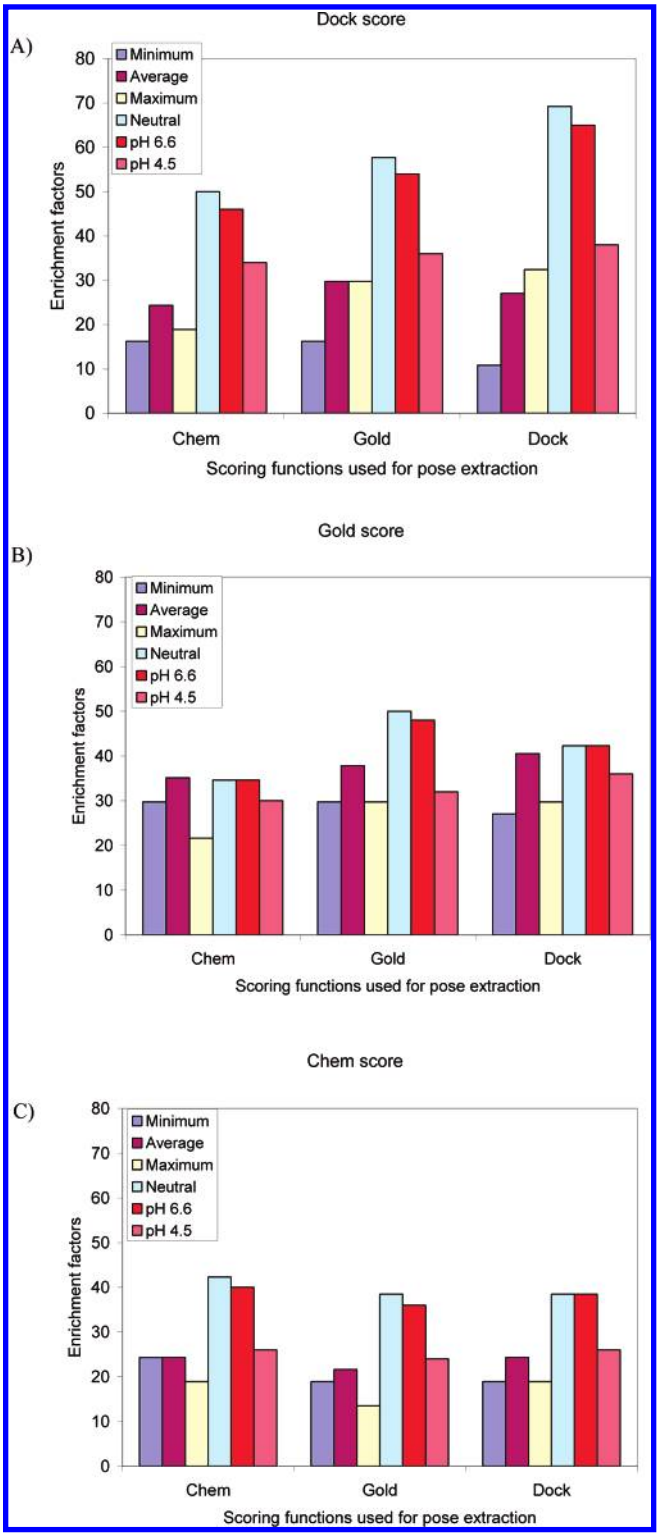

Figure 2. Bar representation of enrichment factors calculated at 1% of the ranked database obtained by FlexX-Pharm. Poses were ranked by (A) Dock, (B) Gold, and (C) Chem scores.

be influential through scoring the rank. If average values of scores contributing to various protonation states were used for $EF_{(1\%)}$ calculation, some improvement was achieved. This can be due to the fact that more or less incorrect states quenched the effect of each other.

Best $EF_{(1\%)}$s were given when neutral molecules or compounds protonated at pH 6.6 were selected for $EF_{(1\%)}$ calculation. The best $EF_{(1\%)}$ was 69 when poses were both extracted and ranked by Dock score (Figure 2). This radical

improvement in $EF_{(1\%)}$s assumes that protonation has a remarkable impact on scoring poses. Incorrect ligand protonation can easily lead to bad ranking and finally yields suboptimal enrichment factors. Neutral compounds and molecules protonated at pH 6.6 provided very similar results again, while protomers calculated at pH 4.5 yielded decreased enrichment.

The DOCK-like function (Dock score) is based on the work of Meng, Shoichet, and Kuntz.[46] It considers both electrostatic and hydrophobic contributions for the binding free energy but does not include entropic terms. A distance dependent dielectric attenuates charge−charge and other polar interactions. Nevertheless, Knegtel et al.[47] report that these terms tend to dominate the nonpolar terms in this type of scoring function. ChemScore[48] is based on a diverse training set of 82 receptor−ligand complexes. It uses four terms that estimate contributions to binding free energy from lipophilic interactions, metal−ligand binding, hydrogen bonding, and loss of ligand flexibility. The GOLD-like (Gold score[49]) function is based on the work of Jones et al. and focuses on hydrogen bonding interactions. A hydrogen bonding term incorporates desolvation of donors and acceptors. A pairwise dispersion term considers the contribution of nonpolar or hydrophobic interactions, while a more typical molecular mechanics equation is used to calculate the internal strain energy of the ligand itself. This type of scoring function performs well in situations where there are significant polar interactions between the ligand and the receptor but has difficulty with ligands that are primarily nonpolar hydrophobic in nature.

The results suggest neutral or compounds protonated at pH 6.6 are favored within the active site of BACE1. There is a possibility that compounds with charged groups over- or underestimate and therefore lead to an obstructed rank in every case. This is further supported by the decreased $EF_{(1\%)}$s calculated for ligands protonated at pH 4.5. Notwithstanding that Dock score, in which nonpolar terms tend to be dominated, provided the best results for neutral compounds.

Comparing the three scoring functions suggests that the Gold score was not so sensitive to protonation as the Dock score, and similar to the Dock score the impact of ligand protonation was remarkable for the Chem score. The reason for the sensitivity of the Dock score to ligand protonation is that polar interactions are dominant in this score, while polar and nonpolar terms are considered in a more balanced way in the Chem score. In summary, ligand protonation has a significant effect on selecting top ranking poses by all of the scoring functions used in FlexX-Pharm.

Screening by Surflex-Dock. Virtual screens were performed by Surflex 1.33 and 2.03. Surflex version 1.33 resulted in a maximal enrichment factor of 36 calculated at 1% of the ranked database. Using Surflex version 2.03 and applying the remin option resulted in improved performance, $EF_{(1\%)}$ of 58, highlighting the importance of postminimization in docking calculations (Table 3).

Docking with Surflex 1.33 provided comparable enrichment factors for all six setups; therefore, screening efficacy does not seem to depend on the ligand protonation. Surflex 2.03 with the application of the remin option ameliorated the enrichment factors. However the enrichment factors corresponding to minimum, maximum, average, neutral, and pH 6.6 remained similar. When molecules were protonated at pH 4.5, $EF_{(1\%)}$ for Surflex 1.33 was improved, while it was obstructed for Surflex 2.03. However, these changes were not significant. Surflex was able to find the correct binding modes and score them resulting in acceptable $EF_{(1\%)}$s even if the ligands were incorrectly protonated.

Table 3. Summary of Enrichment Factors Calculated at 1% of the Ranked Database by Surflex Version 1.33 and 2.03

	$EF_{(1\%)}$	
	Surflex 1.33	Surflex 2.03 -remin
minimum	34	58
maximum	28	44
average	36	56
neutral	30	58
pH 6.6	30	58
pH 4.5	38	44

Table 4. Summary of Enrichment Factors Given by eHiTS, When Docking 1%, 2%, and 5% of the Initial Database

	dock 1%	dock 2%	dock 5%
$EF_{(1\%)}$	30	24	14
no. of actives found by the filter	15	17	17

Surflex-Dock's built-in scoring function is composed of terms that are related to the thermodynamics of protein ligand binding and estimates the parameters of the function based on protein−ligand complexes. The implemented terms are, in rough order of importance, hydrophobic complementarity, polar complementarity, entropic, and solvation terms. By far the most dominant ones are the first two. While the hydrophobic term yields 0.1 unit−contribution to binding free energy−per ideal hydrophobic atom−atom contact, a perfect hydrogen bond yields 1.2 units. Despite the large difference of these latter two contact types the hydrophobic term accounts for a larger total proportion of ligand binding energy on average because there are more hydrophobic contacts than proper polar ones in a typical protein−ligand interaction; therefore, the effect of ligand protonation on enrichment factors was unexpected. Consequently, Surflex was able to find correct poses−leading to remarkable $EF_{(1\%)}$s−independently from the protonation states of the ligands.

In this case only postminimization resulted in a significant increase in $EF_{(1\%)}$s, indicating that the pose has an important influence on scoring. A better pose leads to a better approximation of the free energy of ligand binding.

Screening by eHiTS. Since eHiTS checks all possible protonation states, an analysis such as for GOLD, FlexX-Pharm, and Surflex could not be done. In this case only neutral ligands were docked by eHiTS to compare its effectiveness with GOLD, FlexX-Pharm, or Surflex.

1%, 2%, and 5% of the initial database was prefiltered by the filter training algorithm implemented in eHiTS and then docked. Filtering out 99% of the initial database lead to an enrichment factor of 30 (Table 4). This run tested the performance of the filter itself. Next 2% of the initial database was selected by the filter. Then these compounds were docked by eHiTS (dock2%) and ranked by the built-in scoring function of eHiTS. $EF_{(1\%)}$ was somewhat obstructed. When 5% was docked (dock5%), the enrichment further decreased. These results suggest that the filter could select active compounds effectively, but docking and scoring were not able to top rank them. For dock2% eHiTS ranked 5

inhibitors incorrect; in the case of dock5% 10 inhibitors were misranked.

Herby eHiTS was proved to be a quick and practical tool as there was no need to protonate the input database or even the target protein previously, and in combination of the filter training step the size of the input database could be radically cut down. But it is apparent that the docking algorithm alone could not find all the active compounds that the filter had selected.

Comparison of the Four Docking Algorithms. Four docking algorithms were comparatively evaluated: Surflex, a surface based molecular similarity method; GOLD, a genetic algorithm; FlexX-Pharm, a systematic search using an incremental ligand building with constraints; and eHiTS, an exhaustive search. Surflex score, an empirical scoring function; GOLDScore, a force field based one; Gold, Dock, and Chem scores, two force field based and empirical functions, respectively; ScreenScore, an empirical score; and eHiTS score, a combination of knowledge-based and empirical functions, were compared in this study. Besides their dependency on ligand protonation their efficiencies were also compared against BACE1.[50−54] All raw data sets with the scores are available upon request.

FlexX-Pharm provided the best result, an $EF_{(1\%)}$ of 69. In this case $EF_{(1\%)}$s were remarkably affected by ligand protonation. An incremental ligand building implemented in FlexX searches interacting partners; therefore, the search itself is expected to depend on ligand protonation. However, ligand positions were evaluated through Screenscore,[50] thus the search without the scoring function cannot be evaluated. Results suggest that incremental ligand building in combination with pharmacophore constraints and applying ScreenScore during ligand building can find correct poses. For enrichment factor calculations among Dock, Gold, and Chem scores Dock outperformed the other two. Noteworthy, neutral compounds were better scored suggesting either that neutral compounds are preferably bound or scoring functions fail to score charged groups properly. Scoring functions were able to select correct poses if compounds were not ionized. A possible consequence of this results would be that ionized ligands if bound to BACE1 could not be identified in this way, which is further supported by the $EF_{(1\%)}$s calculated for compounds protonated at pH 4.5 when a significant ratio of the compounds were charged.

Virtual screening with Surflex resulted in the second best results, an $EF_{(1\%)}$ of 58. The search strategy of Surflex employs an idealized ligand, called protomol, that utilizes various molecular fragments. Molecular fragments are tessellated in the active site and optimized based on the scoring function. The search algorithm makes use of the morphological similarity function that is evaluated between the protomol and the putative ligands. It was supposed that this similarity was not sensitive enough to differentiate between protomers; therefore, enrichment factors did not vary significantly if the best, the worst, or the average scores were used for ranking. We assume that the dependency on ligand protonation is further attenuated through the Surflex score. In the Surflex score by far the most important terms are the hydrophobic and the polar ones, in order of rough importance. However a slight decrease was found for the case of maximum, when the best scores were used for ranking, which was unexpected. In Surflex 2.03, postdock minimization was done before scoring. This postprocessing step approximately doubled the results revealing the importance of the precise ligand positions, though enrichment factors given for Surflex 1.33 without postminimization were also acceptable. Surflex and its scoring function was proven to be a robust procedure that did not show notable dependency on ligand protonation. In this case neutral molecules or compounds protonated at pH 4.5, pH 6.6 provided comparable enrichment factors to those calculated for minima, maxima, or average values of scores.

GOLD provided the third best enrichment factors. GOLD utilizes a genetic algorithm and a scoring function with a significant polar term so we expected the dependency of enrichment factors on ligand protonation. The best $EF_{(1\%)}$ was 24 given for neutral compounds when ranking by GOLDScore. Maxima, minima, or average values of GOLDScore provided low enrichment factors, so such as for FlexX-Pharm neutral ligands or compounds protonated at pH 6.6 gave acceptable enrichment factors. This further highlights the inability of these scoring functions to top rank protonated molecules. Interesting to note that when ranking by the protein−ligand hydrophobic term of GOLDScore only, enrichment factors were about doubled, and the dependence of enrichment factors on ligand protonation was diminished. This is a further indication—mentioned for Surflex score—of the relevance of the hydrophobic contribution to the binding free energy. This latter had already been observed in several studies[27−30] calculating linear interaction energy or new linear binding affinity values and also underlies our findings. Scoring functions seems to perform better with significant hydrophobic terms when docking to BACE1. Moreover it suggests that ligand protonation has a slight impact on enrichment factors when using scoring functions like these.

Finally, eHiTS, an exhaustive search, combined with prefiltering was tested with its built in eHiTS scoring function. Here the docking algorithm checks all protomers so ligand protonation was not done beforehand. The best $EF_{(1\%)}$ was 30, which was acceptable and given when 99% of the initial screening database was filtered out. This $EF_{(1\%)}$ measures the performance of the filter itself. To test the docking algorithm and the scoring function 2% and 5% were selected by the filter, then docked, and ranked. However the $EF_{(1\%)}$s were obstructed suggesting that the docking algorithm can dock and score BACE1 inhibitors less effectively. Although eHiTS was proved to be a practical tool because previous ligand or protein protonation can be avoided, the docking algorithm could not outperform the efficacy of the other three virtual screening protocols.

CONCLUSION

Four docking algorithms were comparatively evaluated in virtual screening experiments against BACE1. The effect of ligand protonation on screening efficacy was estimated over the comparison of the four docking algorithms. To the best of our knowledge no studies dealt with the effect of ligand protonation on virtual screening results against BACE1; therefore, we analyzed its effect on various virtual screening protocols.

Virtual screening by FlexX-Pharm provided the best results and showed dependency on ligand protonation. While Surflex gave enrichment factors comparable to FlexX-Pharm, it did

not depend on ligand protonation. Comparing these protocols FlexX-Pharm included scores more dependent on polar interactions than Surflex. So the dependency on ligand protonation was realized through mainly the characteristics of the scoring functions. Scoring functions with relevant hydrophobic contribution showed reduced dependency on ligand protonation and indicated its relevance in the thermodynamics of binding. GOLD showed remarkable dependency on ligand protonation and further underlined the importance of the hydrophobic contribution, while eHiTS did not outperform the other three.

In this study we showed that ligand protonation can affect the results of virtual screening against BACE1; however, it is much more dependent on the characteristics of the scoring functions used. Dock, Gold, Chem scores, and GOLDScore were not able to select good poses from all the protomers, while the Surflex score was able to find a reliable pose resulting in acceptable enrichment factors. In all cases neutral molecules or compounds protonated at pH 6.6 were ranked well.

Based on these results obtained for BACE1 we suggest the investigation of the effect of ligand protonation in conjunction with docking algorithms and scoring functions for each different target. Dealing with all possible protomers does not obviously improve screening efficacy but definitely increases computational resources allocated to the screening campaign.

ACKNOWLEDGMENT

The authors thank Dr. György Ferenczy for fruitful discussions, Tripos Inc. for evaluation licenses of Surflex, and SimBioSys Inc. for evaluation licenses of eHiTS. Financial support (for Cs.M. and I.S.) under grant number GVOP-3.2.1.-2004-04-0195/3.0 is also acknowledged.

Supporting Information Available: IDs, scores, and findings on the reproduction of relevant X-ray structures cocrystallized with druglike ligands by FlexX-Pharm. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Vassar, R. β-secretase (BACE) as a drug target for Alzheimer's disease. *Adv. Drug Delivery Rev.* **2002**, *54*, 1589−1602.
(2) John, V.; Beck, J. P.; Bienkowski, M. J.; Sinha, S.; Heinrikson, R. L. Human β-secretase (BACE) and BACE inhibitors. *J. Med. Chem.* **2003**, *46*, 4625−4630.
(3) Eder, J.; Hommel, U.; Cumin, F.; Martoglio, B.; Gerhartz, B. Aspartic proteases in drug discovery. *Curr. Pharm. Des.* **2007**, *13*, 271−285.
(4) Huang, D.; Lüthi, U.; Kolb, P.; Cecchini, M.; Barberis, A.; Caflisch, A. In silico discovery of β-secretase inhibitors. *J. Am. Chem. Soc.* **2006**, *128*, 5436−5443.
(5) Guo, T.; Hobbs, D. W. Development of BACE1 inhibitors for Alzheimer's disease. *Curr. Med. Chem.* **2006**, *13*, 1811−1829.
(6) Hong, L.; Tang, J. Flap position of free memapsin 2 (β-secretase); a model for flap opening in aspartic protease catalysis. *Biochemistry* **2004**, *43*, 4689−4695.
(7) Turner, R. T.; Hong, L.; Koelsch, G.; Ghosh, A. K.; Tang, J. Structural locations and functional roles of new subsites S5; S6; and S7 in memapsin 2 (β-secretase). *Biochemistry* **2005**, *44*, 105−112.
(8) Coburn, C. A.; Stachel, S. J.; Li, Y.; Rush, D. M.; Steele, T. G.; Chen-Dodson, E.; Holloway, M. K.; Xu, M.; Huang, Q.; Lai, M.; DiMuzio, J.; Crouthamel, M.; Shi, X.; Sardana, V.; Chen, Z.; Munshi, S.; Kuo, L.; Makara, G. M.; Annis, D. A.; Tadikonda, P. K.; Nash, H. M.; Vacca, J. P.; Wang, T. Identification of a small molecule nonpeptide active site β-secretase inhibitor that displays a nontraditional binding mode for aspartyl proteases. *J. Med. Chem.* **2004**, *47*, 6117−6119.
(9) Ghosh, A. K.; Devasamudram, T.; Hong, L.; Dezutter, C.; Xu, X.; Weerasena, V.; Koelsch, G.; Bilcer, G.; Tang, J. Structure-based design of cycloamide-urethane-derived novel inhibitors of human brain memapsin 2 (β-secretase). *Bioorg. Med. Chem. Lett.* **2005**, *15*, 15−20.
(10) Stachel, S. J.; Coburn, C. A.; Steele, T. G.; Crouthamel, M.; Pietrak, B. L.; Lai, M.; Holloway, M. K.; Munshi, S. K.; Graham, S. L.; Vacca, J. P. Conformationally biased P3 amide replacements of β-secretase inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 641−644.
(11) Hanessian, S.; Yang, G.; Rondeau, J.; Neumann, U.; Betschart, C.; Tintelnot-Blomley, M. Structure-based design and synthesis of macroheterocyclic peptidomimetic inhibitors of the aspartic protease β-site amyloid precursor protein cleaving enzyme (BACE). *J. Med. Chem.* **2006**, *49*, 4544−4567.
(12) Yang, W.; Lu, W.; Lu, Y.; Zhong, M.; Sun, J.; Thomas, A. E.; Wilkinson, J. M.; Fucini, R. V.; Lam, M.; Randal, M.; Shi, X.; Jacobs, J. W.; McDowell, R. S.; Gordon, E. M.; Ballinger, M. D. Aminoethylenes: A tetrahedral intermediate isostere yielding potent inhibitors of the aspartyl protease BACE-1. *J. Med. Chem.* **2006**, *49*, 839−842.
(13) Ghosh, A. K.; Kumaragurubaran, N.; Hong, L.; Lei, H.; Hussain, K. A.; Liu, C.-F.; Devasamudram, T.; Weerasena, V.; Turner, R.; Koelsch, G.; Bilcer, G.; Tang, J. Design, synthesis and X-ray structure of protein-ligand complexes: important insight into selectivity of memapsin 2 (β-secretase) inhibitors. *J. Am. Chem. Soc.* **2006**, *128*, 5310−5311.
(14) Maillard, M. C.; Hom, R. K.; Benson, T. E.; Moon, J. B.; Mamo, S.; Bienkowski, M.; Tomasselli, A. G.; Woods, D. D.; Prince, D. B.; Paddock, D. J.; Emmons, T. L.; Tucker, J. A.; Dappen, M. S.; Brogley. L.; Thorsett, E. D.; Jewett, N.; Sinha, S.; John, V. Design synthesis and crystal structure of hydroxyethyl secondary amine-based peptidomimetic inhibitors of human β-secretase. *J. Med. Chem.* **2007**, *50*, 776−781.
(15) Rajapakse, H. A.; Nantermet, P. G.; Selnick, H. G.; Munshi, S.; McGaughey, G. B.; Lindsley, S. R.; Young, M. B.; Lai, M,-T.; Espeseth, A. S.; Shi, X.-C.; Colussi, D.; Pietrak, B.; Crouthamel, M.; Tugusheva, K.; Huang, Q.; Xu, M.; Simon, A. J.; Kuo, L.; Hazuda, D. J.; Graham, S.; Vacca, J. P. Discovery of oxadiazoyl tertiary carbinamine inhibitors of β-secretase (BACE-1). *J. Med. Chem.* **2006**, *49*, 7270−7273.
(16) Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A. K.; Zhang, X. C.; Tang, J. Structure of the protease domain of memapsin 2 (β-secretase) complexed with inhibitor. *Science* **2000**, *290*, 150−153.
(17) Patel, S.; Vuillard, L.; Cleasby, A.; Murray, C. W.; Yon, J. Apo and inhibitor complex structures of BACE (β-secretase). *J. Mol. Biol.* **2004**, *343*, 407−416.
(18) Congreve, M.; Aharony, D.; Albert, J.; Callaghan, O.; Campbell, J.; Carr, R. A. E.; Chessari, G.; Cowan, S.; Edwards, P. D.; Frederickson, M.; McMenamin, R.; Murray, C. W.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of β-secretase. *J. Med. Chem.* **2007**, *50*, 1124−1132.
(19) Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to β-secretase. *J. Med. Chem.* **2007**, *50*, 1116−1123.
(20) Bhisetti, G. R.; Saunders, J. O.; Murcko, M. A.; Lepre, C. A.; Britt, S. D.; Come, J. H.; Deninger, D. D.; Wang, T. Vertex Pharmaceuticals Incorporated, Patent WO 02/88101, 2002.
(21) McGaughey, G. B.; Colussi, D.; Graham, S. L.; Lai, M.-T.; Munshi, S. K.; Nantermet, P. G.; Pietrak, B.; Rajapakse, H. A.; Selnick, H. G.; Stauffer, S. R.; Holloway, M. K. β-secretase (BACE-1) inhibitors: accounting for 10s loop flexibility using rigid active sites. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1117−1121.
(22) Moitessier, N.; Therrien, E.; Hanessian, S. A method for induced-fit docking; scoring; and ranking of flexible ligands. Application to peptidic and pseudopeptidic β-secretase (BACE 1) inhibitors. *J. Med. Chem.* **2006**, *49*, 5885−5894.
(23) Yu, N.; Hayik, S. A.; Wang, B.; Liao, N.; Reynolds, C. H.; Merz, K. M., Jr. Assigning the protonation states of the key aspartates in β-secretase using QM/MM X-ray structure refinement. *J. Chem. Theory Comput.* **2006**, *2*, 1057−1069.
(24) Polgár, T.; Keserü, G. M. Virtual screening for β-secretase (BACE1) inhibitors reveals the importance of protonation states at Asp32 and Asp228. *J. Med. Chem.* **2005**, *48*, 3749−3755.
(25) Rajamani, R.; Reynolds, C. H. Modeling the protonation states of the catalytic aspartates in β-secretase. *J. Med. Chem.* **2004**, *47*, 5159−5166.
(26) Park, H.; Lee, S. Determination of the active site protonation state of β-secretase from molecular dynamics simulation and docking experiment: Implications for structure-based inhibitor design. *J. Am. Chem. Soc.* **2003**, *125*, 16416−16422.
(27) Huang, D.; Caflisch, A. Efficient evaluation of binding free energy using continuum electrostatics solvation. *J. Med. Chem.* **2004**, *47*, 5791−5797.

(28) Rajamani, R.; Reynolds, C. H. Modeling the binding affinities β-secretase inhibitors: application to subsite specificity. *Bioorg. Med. Chem. Lett.* **2004**, *47*, 5159−5166.
(29) Tounge, B. A.; Rajamani, R.; Baxter, E. W.; Reitz, A. B.; Reynolds, C. H. Linear interaction energy models for β-secretase (BACE) inhibitors: Role of van der Waals; electrostatic; and continuum-solvation terms. *J. Mol. Graphics Modell.* **2006**, *24*, 475−484.
(30) Holloway, K. M.; McGaughey, G. B.; Coburn, C. A.; Stachel, S. J.; Jones, K. G.; Stanton, E. L.; Gregro, A. R.; Lai, M.-T.; Crouthamel, M.-C.; Pietrak, B. L.; Munshi, S. K. Evaluating scoring functions for docking and designing β-secretase inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 823−827.
(31) Polgár, T.; Keserü, G. M. Ensemble docking into flexible active sites. critical evaluation of FlexE against JNK-3 and β-secretase. *J. Chem. Inf. Model.* **2006**, *46*, 1795−1805.
(32) Knox, A. J. S.; Meegan, M. J.; Carta, G.; Lloyd, D. G. Considerations in compound database preparation-"hidden" impact on virtual screening results. *J. Chem. Inf. Model.* **2005**, *45*, 1908−1919.
(33) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, E. C.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem*. **2006**, *49*, 5912−5931.
(34) Pan, Y.; Huang, N.; Cho, S.; MacKerell, A. D., Jr. Consideration of molecular weight during compound selection in virtual target-based database screening. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 267−272.
(35) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein-ligand docking: Avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−806.
(36) *SYBYL 7.2*; Tripos Inc.: 1699 South Hanley Road, St. Louis, MO 63144-2319.
(37) Prous Integrity Database, Prous Science. http://www.prous.com (accessed month year).
(38) OpenEye Scientific Software, Inc. 3600 Cerrillos Road, Suite 1107, Santa Fe, NM 87507, U.S.A.
(39) Jain, A. N. Surflex-dock 2.1: robust performance from ligand energetic modelling, ring flexibility, and knowledge-based search. *J. Comput.-Aided Mol. Des*. **2007**, *21*, 281−306.
(40) Jain, A. N. Surflex: a fully automatic flexible molecular docking using a molecular similarity-based search engine. J. *Med. Chem*. **2003**, *46*, 499−511.
(41) Pham, T. A.; Jain, A. N. Parameter estimation for scoring protein ligand interactions using negative training data. *J. Med. Chem*. **2006**, *49*, 5856−5868.
(42) SimBioSys Inc. 135 Queen's Plate Drive, Toronto, Ontario M9W 6V1, Canada. http://www.simbiosys.ca (accessed month year).
(43) Jones, G.; Willett, P.; Glen, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J. Comput.-Aided Mol. Des*. **1995**, *9*, 532−549.
(44) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol*. **1997**, *267*, 727−748.
(45) Cole, J. C.; Willem, J.; Nissink, M.; Taylor, R. Protein-ligand docking and virtual screening with GOLD. In *Virtual screening in drug discovery*; Alvarez, J., Shoichet, B., Eds.; Taylor&Francis LLC: New York, 2005; pp 379−411.
(46) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem*. **1992**, *13*, 505−524.
(47) Knegtel, R. M. A.; Bayada, D. M.; Engh, R. A.; von der Saal, W.; van Geerestein, V. J.; Grootenhuis, P. D. J. Comparison of two implementations of the incremental construction algorithm in flexible docking of thrombin inhibitors. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 167−183.
(48) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.
(49) Jones, G.; Willett, P.; Glen, R. C. *J.* Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol*. **1995**, *245*, 43−53.
(50) Schulz-Gasch, T.; Stahl, M. Scoring functions for protein-ligand interactions: A critical perspective. *Drug Discovery Today*: *Technol*. **2004**, *1*, 231−239.
(51) 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.
(52) Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L. Comparative study of several algorithms for flexible ligand docking. *J. Comput.-Aided Mol. Des*. **2003**, *17*, 755−763.
(53) McGovern, S. L.; Shoichet, B. K. Information decay in molecular docking screens against holo; apo; and modeled conformations of enzymes. *J. Med. Chem*. **2003**, *46*, 2895−2907.
(54) Gohlke, H.; Hendlich, M.; Klebe, G. Predicting binding modes; binding affinities and 'hot spots' for protein-ligand complexes using a knowledge-based scoring function. *Perspect. Drug Discovery Des.* **2000**, *20*, 115−44.

CI700223P