Evaluation of Binary QSAR Models Derived from LUDI and MOE Scoring Functions for Structure Based Virtual Screening

Philip Prathipati and Anil K. Saxena*

Medicinal and Process Chemistry Division, Central Drug Research Institute, Chatter Manzil Palace, Lucknow-226 001, India

Received April 11, 2005

In today's world of high-throughput in silico screening, the development of virtual screening methodologies to prioritize small molecules as new chemical entities (NCEs) for synthesis is of current interest. Among several approaches to virtual screening, structure-based virtual screening has been considered the most effective. However the problems associated with the ranking of potential solutions in terms of scoring functions remains one of the major bottlenecks in structure-based virtual screening technology. It has been suggested that scoring functions may be used as filters for distinguishing binders from nonbinders instead of accurately predicting their binding free energies. Subsequently, several improvements have been made in this area, which include the use of multiple rather than single scoring functions and application of either consensus or multivariate statistical methods or both to improve the discrimination between binders and nonbinders. In view of it, the discriminative ability (distinguishing binders from nonbinders) of binary QSAR models derived using LUDI and MOE scoring functions has been compared with the models derived by Jacobbsson et al. on five data sets viz. estrogen receptor αmimics (ERα_mimics), estrogen receptor αtoxins (ERα_toxins), matrix metalloprotease 3 inhibitors (MMP-3), factor Xa inhibitors (fXa), and acetylcholine esterase inhibitors (AChE). The overall analyses reveal that binary QSAR is comparable to the PLS discriminant analysis, rule-based, and Bayesian classification methods used by Jacobsson et al. Further the scoring functions implemented in LUDI and MOE can score a wide range of protein−ligand interactions and are comparable to the scoring functions implemented in ICM and Cscore. Thus the binary QSAR models derived using LUDI and MOE scoring functions may be useful as a preliminary screening layer in a multilayered virtual screening paradigm.

INTRODUCTION

In silico approaches (either ligand based or structure based) for lead generation, involve two kinds of approaches, viz. virtual screening of 3D databases and de novo structure design. While the latter approach allows the creative design of a variety of candidate structures, the former has the advantage of experimental testing without the synthetic efforts.[1] Furthermore as of now several million virtual screening libraries are available for screening, which can be purchased from the respective company after a reliable in silico screening.[2] Though few success stories in lead identification with this kind of an approach have been reported in the literature,[3−12] it is often considered superior to experimental HTS in several respects. For instance, in virtual screening, compounds that have not been synthesized, and even "virtual ligands", can be tested, and it is also possible to apply the technique to target macromolecules for which assays are difficult or expensive in experimental HTS. However one of the major limiting factors of a virtual screening experiment is that all the available methods learn only from known binders and try to quantitatively predict the binding affinity of binders and nonbinders alike.[13,14] Though as of now there are no results indicating how well these methods perform when confronted with prediction of activity of inactive compounds, it is nonetheless expected that these models will fail in predicting nonbinders or inactive molecules. Another major limiting factor in the implementation of virtual screening methods is that the methods that are efficient are not effective (like similarity searching and various ligand based approaches), and those that are effective are not efficient (like very accurate flexible docking and scoring analysis). Thus in view of above, a multilayer virtual screening protocol comprising of a screening layer, a prediction layer, and a validation layer has been envisaged.[14] This multilayered screening approach, integrates several virtual screening techniques in a systematic order according to their computational cost. The first layer is a preliminary *screening layer* that can evaluate a large number of compounds at minimal computational cost. For this part of the calculation, accurate numerical predictions of activity are not essential, but this layer may be used to exclude the compounds that are likely to be poor binders. Some of the methods that can be used in this screening layer are common feature pharmacophores, descriptor based binary QSAR models, and fingerprint based similarity searching. These methods can be used either as a standalone wherein the choice of the method depends on the availability of structural and molecular data or a consensus of all these models may be used. Further as per the existing knowledge in structure based drug design, it is now clear that fast, approximate rigid docking and scoring functions cannot reliably predict the

* Corresponding author phone: +91-0522-2212411-18, extn 4268; fax: +91-0522-2223405; e-mail: anilsak@hotmail.com.

10.1021/ci050120w CCC: $33.50
Published on Web 07/06/2005

Table 1. Ligand Sets and Crystal Structures Used in the Study

target	PDB file	ligand source	no. of binders[a]	no. of nonbinders[a]
ERα	1ere	Sippl, *J. Comput.-Aided Mol Des.* **2000**, *14*, 559 (mimics)	36 (12)	999 (336)
ERα	1ere	Shi et al. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 186 (toxins)	110(36)	999 (336)
AChE	1eve	Contreras et al. *J. Med. Chem.* **2001**, *44*, 2707	54(18)	999 (336)
MMP-3	1hy7	Ha et al. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 395	60(20)	999 (336)
fXa	1g2l	Matter et al. *J. Med. Chem.* **2001**, *44*, 2707	129(43)	999 (336)

[a] The number of test set compounds is indicated in the parentheses.

binding free energies and that more correct prediction still require more complex computational methods based on simulations.[15−18] Subsequently, several recent studies suggest that scoring functions should rather be regarded as filters for distinguishing binders from nonbinders than the accurate predictions of their binding free energies.[13,15,18−21] Hence the predictions based on rigid docking and scoring may be used as a part of the preliminary screening layer of the multilayered virtual screening protocol, in cases where the 3D structure of the protein, information on the ligand-binding site is available and for proteins which are not very flexible. Consequently several improvements have been made in this area, which have significant implications in virtual screening. These include the use of multiple, rather than single scoring functions and the subsequent application of either consensus[22−25] or multivariate statistical methods[13] or both[25] to improve the discrimination between binders and nonbinders.

In view of the above we have evaluated the suitability of a new classification modeling algorithm 'binary QSAR' to build classifiers which are able to discriminate between active and inactive compounds, using four hitherto unexplored scoring functions for discriminant ability, three of which were implemented in LUDI and one implemented in MOE, and the results are presented in this manuscript.

MATERIALS AND METHODS

Data Sets. To evaluate the discriminative ability of the binary QSAR algorithm and the suitability of the four new scoring functions, five data sets which were the subject of a previous assessment study by Jacobsson et al.[13] have been used (Table 1, Figure 1). In addition to the known binders of four different targets namely estrogen receptor α (ERα), matrix metalloprotease 3 (MMP-3), factor Xa (fXa), and acetylcholine esterase (AChE), this data set includes 999 diverse ligands extracted from MDDR, selected using 2D fingerprint-based clustering.[13] The partitioning into training and validation sets was done by sorting the known actives according to activity and inactives according to molecular weight and setting apart every third compound as a validation compound in each catagerory.

Preparation of the Proteins and the Ligands. The protein−ligand complexes were downloaded from the Web site (http:/www.compumine.com/research/scoring.html). In MOE, the proteins and ligands were used without any changes, and the charges were assigned to both the proteins and ligands automatically. In InsightII, the CVFF force field[26] charges were assigned to both the proteins and ligands. There were few residues and ligands for which automated charge assessment was a problem, so the structures were manually corrected, the bond orders were suitably modified, and the charges were then assigned.

Binary QSAR. The binary QSAR[27−29] is a relatively new methodology that has been introduced recently by Labute,[27] in which biological activity is expressed in a "binary" format (1 = active and 0 = inactive) and is correlated with molecular descriptors. In this method the set of computed molecular descriptors is transformed into a set of decorrelated and normalized set of variables, and the probability distribution is estimated based on Bayes' theorem.[30] The quality of a binary QSAR model implemented in the software program is generally measured using three parameters of performance, which are estimated using a normal validation and the cross-validation procedure.[27−29] The three parameters are (1) accuracy on active compounds, *c0*/*m0*; (2) accuracy on inactive compounds, *c1*/*m1*; (3) overall accuracy on all of the compounds, (*c0* + *c1*)/(*m0* + *m1*), where *m0* represents the number of active compounds, *m1* the number of inactive compounds, *c0* the number of active compounds correctly labeled by the QSAR model, and *c1* the number of inactive compounds correctly predicted by the QSAR model. These parameters are derived using only the model performance on the training set, and hence they represent the statistical robustness of the training set only.

For the sake of comparison to the results obtained by Jacobsson et al.[13] these addition parameters for the test set were also computed: (1) the overall classification accuracy of a prediction model, Accuracy = tp + tn/tp + fp+ tn +fn, (2) measure of the accuracy of predicting an active class, Precision = tp/tp + fp, and (3) measure of the ability of a prediction model to select instances of a certain class, Recall = tp/tp + fn and enrichment factor EF = precision/[(tp+fn)/(tp+fp+tn+fn)]; wherein tp = number of true positives, tn = number of true negatives, fp = number of false positives, and fn = number of false negatives.

As it has been already indicated,[13] none of the above parameters is an absolute measure of classification performance by itself but should be seen together. In general it has been suggested that a classifier can only be seen as really successful if high precision is accompanied by high recall. The expectation values of the precision and recall for the active class of a random classifier is given by the ratio of actives in the entire set and the ratio between the number of predicted actives and the total number of compounds, respectively. The *enrichment factor* (EF) another commonly used parameter in virtual screening was also estimated. The EF indicates the relative enrichment of active compounds in the set of instances predicted to be active in relation to the fraction of active compounds in the original data set and is considered a good basis of comparison between different methods for the same ligand set and target protein, even though the actual values will depend on the number of actives included in the validation set.

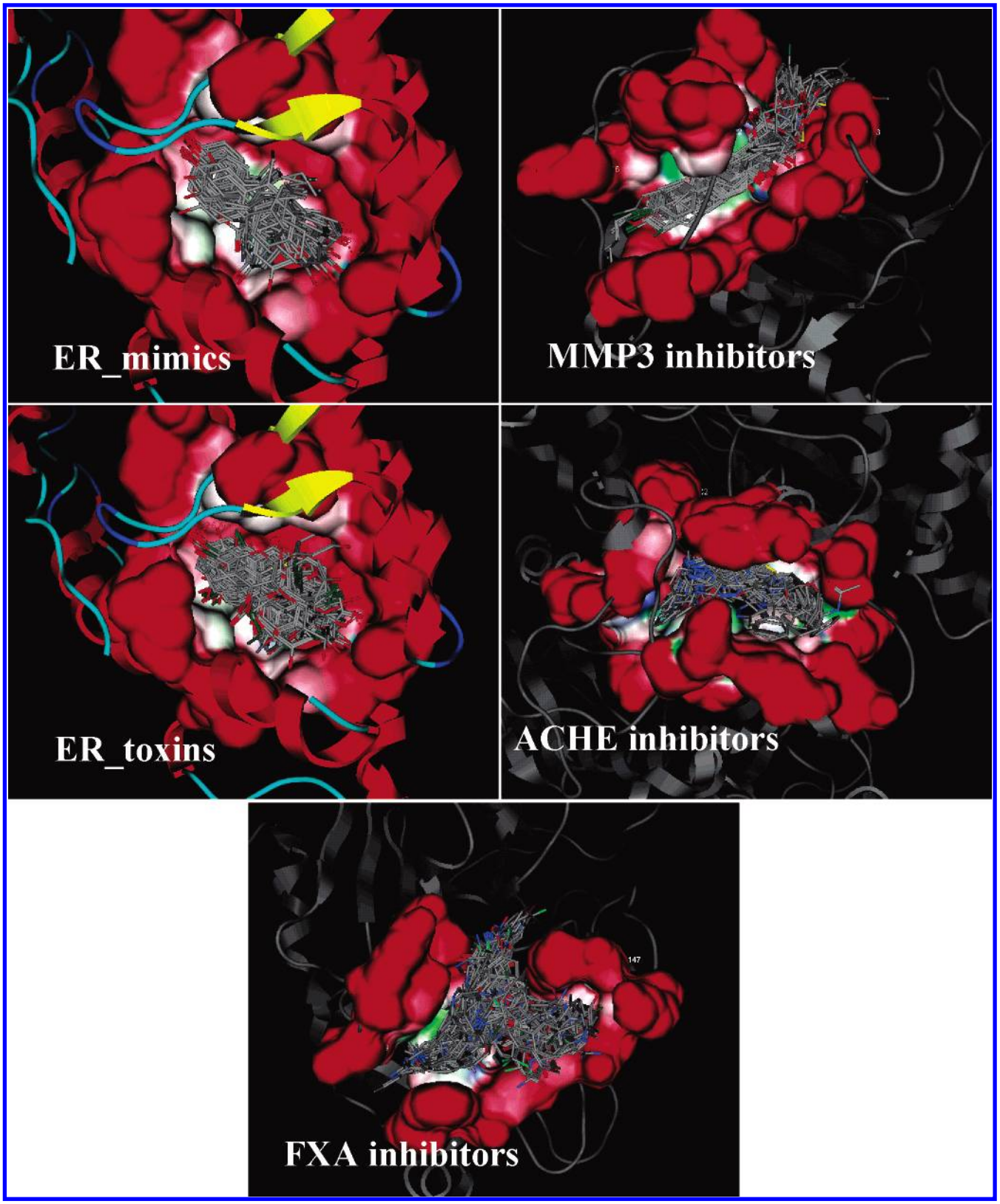

Figure 1. The binding mode of the active molecules docked into their respective targets.

LUDI Scoring Function. This pioneering empirical scoring function developed by Böhm[31−33] dissects the protein−ligand binding free energy as a resultant of hydrogen bonds, hydrophobic effects formed between the complex, and the torsional entropy loss (Table 2). The contribution of each hydrogen bond is scaled by a distance- and angle-dependent function in order to penalize the deviations from an ideal geometry; further "neutral" and "ionic" hydrogen bonds are treated separately. The hydrophobic effect, which calculates the buried hydrophobic molecular surface, counts of all the rotatable single bonds (rotors) in the ligand, which is supposed to be related to the torsional entropy loss of the ligand upon protein−ligand complex formation. The last term is a regression constant. This scoring function is calibrated

Table 2. Individual Terms Implemented in LUDI Scoring Functions

title	description
score	total score
# ionic	number of ionic interactions
Ionic score	contribution of ionic interaction term to the scoring function
Rot score	contribution of the entropy (i.e flexibility α no. of rotatable bonds) to the scoring function
aro/aro score	contribution of π-π interactions to the scoring function
# HB	number of hydrogen bonds
HB score	score from hydrogen bond term of scoring function
Lipo score	score from lipophilic term of scoring function
contact	percentage of fragment in contact with receptor

by fitting known dissociation constants of 87 protein−ligand complexes.[30] One of the advantages of this approach is that it is purely geometrical and therefore avoids costly calculations of potential energy functions.

LUDI provides three scoring functions, namely LUDI_1, LUDI_2, and LUDI_3 set as Energy_Estimate_1, Energy_Estimate_2, and Energy_Estimate_3, respectively. In LUDI_1 scoring function, each fragment is evaluated as a function of the potential number of hydrogen bonding and hydrophobic contacts it can make with a receptor. The LUDI_2 scoring function has been reported to be the only scoring function that has weight factors before each term, which were derived by fitting to experimentally observed binding affinities. LUDI_3 is similar to the LUDI_2 scoring function except that it has an additional term to evaluate the contribution of aromatic−aromatic $\pi-\pi$ interactions to the binding free energy.

MOE Scoring Function. The MOE ligand−protein scoring algorithm[34,35] is based upon an empirical scoring function consisting of (1) a directional hydrogen-bonding term (direct bonds, water-mediated, transition-metal contacts), (2) a directional hydrophobic interaction term, and (3) an entropic term (ligand rotatable bonds immobilized in binding). It also computes an entropic penalty term for those atoms in the ligand, which become immobilized upon binding to the receptor. However like most other empirical scoring it also ignores the conformational strain because it assumes that the conformation given is accessible.The final pK_i prediction is based on the scoring function coefficients that have been trained on a data set of ligand−receptor complexes collected from the protein data bank and their published binding ΔG values. Further this tool can be used to visualize intermolecular contacts of the following types: direct hydrogen bonds, water-mediated hydrogen bonds, transition-metal interactions, and hydrophobic interactions.

RESULTS AND DISCUSSION

Accuracy of the Docked Pose. To assess the accuracy of the docking poses, the MOE 2004 software was used to pictorially evaluate the accuracy of the pose as seen in the X-ray crystallography derived complexes with the pose predicted by the ICM docking algorithm used by Jacobsson et al. Further the binding affinities between the complexes derived using X-ray crystallography and docking studies using ICM and/or experimentally observed binding affinities with the binding affinity predicted from the ICM docking study were compared wherever possible.

Estrogen Receptor Mimics and Toxins. The estrogen receptor (ER) is located in the cell in parallel with other steroid hormone receptors. When bound to a hormone like estrogen, it acts as a transcription factor i.e., it regulates the reading of DNA and production of proteins.[36,37] It exists as dimers with the two possible subunits (α and β): ERα (αα), ERβ ($\beta\beta$), and ERαβ (αβ). Different tissues express the combinations of these subunits in different proportions, and every combination has a different affinity to estrogen response elements and to the sequence on DNA that leads to transcription of particular genes on activation of the estrogen receptor.[38] The design of estrogen receptor mimics or selective estrogen receptor modulators (SERMs) and estrogen receptor toxins or endocrine disruptors is of current interest because of their important role in new drug discovery and environmental toxicity, respectively. The SERMs are a class of medications that acts on the estrogen receptor and the characteristic that distinguishes these substances from receptor agonists and antagonists is that their action is different for various tissues, thereby granting the possibility to selectively inhibit or stimulate estrogen-like action in various tissues. Predominately, they are used in the treatment of ovulation related disorders, osteoporosis, and breast cancer.[39,40]

Some of the major 17β-o-esterdiol-estrogen receptors as reveled by X-ray crystallography[41] (1ERE) are shown in Figure 2[A]. The phenolic hydroxyl of the A-ring is located between Helix3 and Helix6 and makes direct hydrogen bonds to the carboxylate of GLU353, the guanidinium group of ARG394, and a water molecule. The 17β-hydroxyl (O 17) of the D ring makes a single hydrogen bond with HIS524 in Helix11. The remainder of the molecule participates in a number of hydrophobic contacts that are concentrated over the A, A/B interface, and D-rings. The A-ring as well as the planar A/B-ring interface is sandwiched between the side chains of ALA350 and LEU387 on its b face and PHE404 on its a face. At the other end of the binding cavity, the D ring makes nonpolar contacts with ILE424, GLY521, and LEU525. The pose of the estrogen molecule predicted by ICM docking is in close agreement with the ligand−receptor interaction picture described above (Figure 2[B]). Some of the interactions that are common are the single hydrogen bond of 17-β hydroxyl (O 17) of the D ring makes a single hydrogen bond with HIS524 in Helix11 and the hydrogen bonding interactions of the phenolic hydroxyl of the A-ring with the carboxylate of GLU353 (Figure 2[C]). A comparable prediction of binding affinities of 17β-o-esterdiol for its receptor (8.3 and 9.1, respectively, for the complex obtained using X-ray crystallography and docking studies using ICM) further validates the docking pose obtained by the ICM docking algorithm.

Similarly the important hydrogen bonding interactions are also seen to have conserved between the DES (the most active molecule in the ER_toxin data set)−estrogen receptor complex as obtained by X-ray crystallography[40] (3ERD) and docking studies using ICM (Figure 3[A−C]). Further the comparable prediction of the binding affinity of DES for its receptor (9.3 and 9.1 for the complexes obtained using X-ray crystallography and docking studies using ICM, respectively) also substantiates the validity of the docking algorithm.

Factor Xa Inhibitors. The activated factor X is a drug target for control of coagulation[43] because of its role in

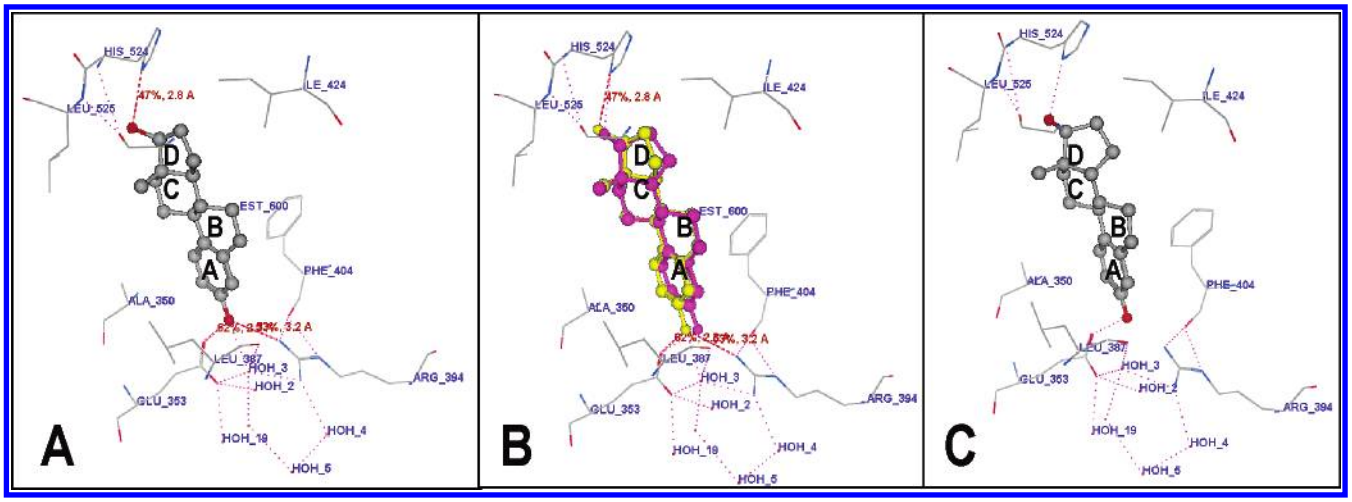

Figure 2. (A) The major interactions of 17β-o-esterdiol with the estrogen receptor as seen in 1ERE. (B) The overlay of the docked and cocrystallized structure of 17β-o-esterdiol in blue and yellow, respectively. (C) Figure depicting some of the conserved hydrogen and hydrophobic binding interactions between the docked structures of 17β-o-esterdiol with the estrogen receptor.

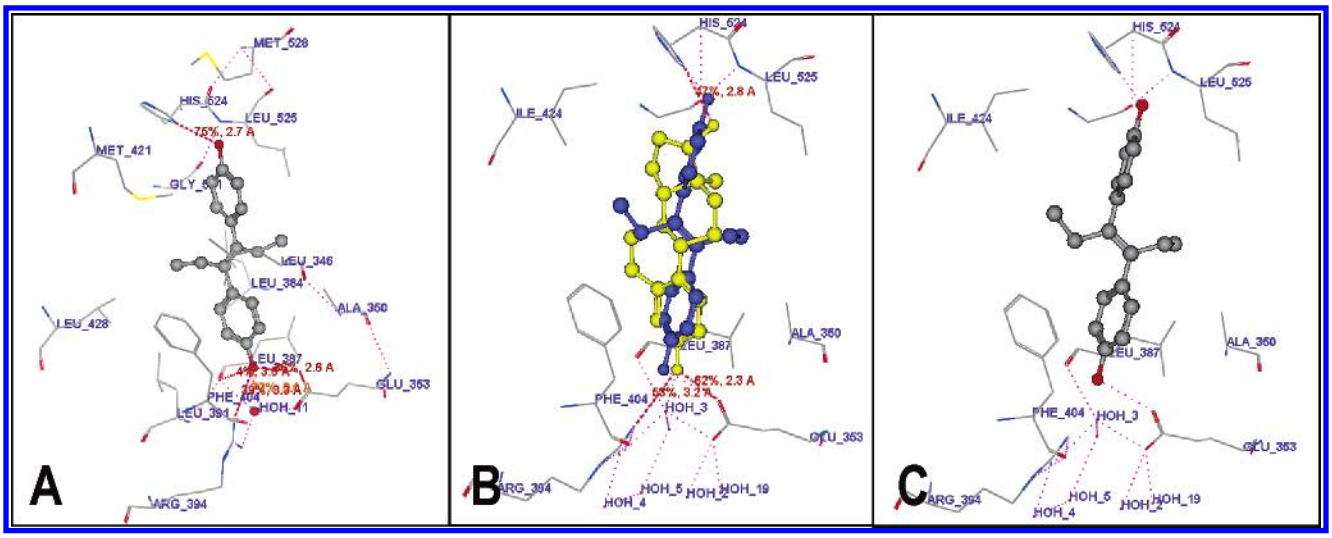

Figure 3. (A) The major interactions of DES with the estrogen receptor as seen in 3ERD. (B) The overlay of the docked and cocrystallized structure of DES in blue and yellow, respectively. (C) Figure depicting some of the conserved hydrogen and hydrophobic binding interactions between the docked structures of DES with the estrogen receptor.

converting prothrombin to thrombin in blood coagulation. It has been reported that synthetic inhibitors are effective antithrombotic agents in animal models, with a low risk of bleeding.[44] Blood coagulation factor X knockout mice experiments have been suggested to form an effective model for the bleeding disorders observed in severe factor X deficiency in humans.[45]

The active site of fXa as delineated from the X-ray crystallographic complex[46](1G2L) comprises of three main regions for interactions: S1 pocket, S4 pocket, and catalytic triad site (Figure 4[A]). The S1 pocket is a narrow cleft with planar hydrophobic walls and ASP194 at the bottom. A positively charged group is favored in this pocket to make a salt bridge with the side chain of the aspartic acid. The S4 pocket consists of three aromatic amino acids: PHE177, TRP220, and TYR100. A hydrophobic group as well as a positively charged group that makes a cation interaction is favored in the S4 pocket. The catalytic triad site (CT site) is composed of the triad residues (i.e. SER200, HIS57, and ASP102) and neighboring residues (i.e. GLU22 and ARG143). The side chains of SER195, GLN192, and ARG143 are fully exposed in the solvent for easy access by an inhibitor. The docking mode of one of the most active fXa inhibitors (F1_47) has revealed a conservation of the major interactions at the S1 and S4 pockets; however, the interactions with the catalytic triad are lost (Figure 4[B,C]). However very poor predictions of binding affinities were obtained where the experimentally determined affinity and predicted affinity values using MOE were 4.244 and 7.6, respectively, thus indicating the poor performance of the scoring function implemented in MOE.

Further the experimentally determined and predicted binding affinity values of F1_47 for its receptor using MOE scoring function were 8.1549 and 6.4, respectively. Again indicating the poor performance of the scoring function implemented in MOE as well as inaccuracy of the docking pose due to the rigid docking algorithm.

MMP-3 Receptor Inhibitors. The zinc- and calcium-dependent families of proteins called the MMPs (matrix metalloproteases) are collectively responsible for the degradation of the extracellular matrix.[47] The enzymes are synthesized as zymogens, and these enzymes under physiological conditions are selectively regulated by endogenous inhibitors.[48]An imbalance between the active enzymes and

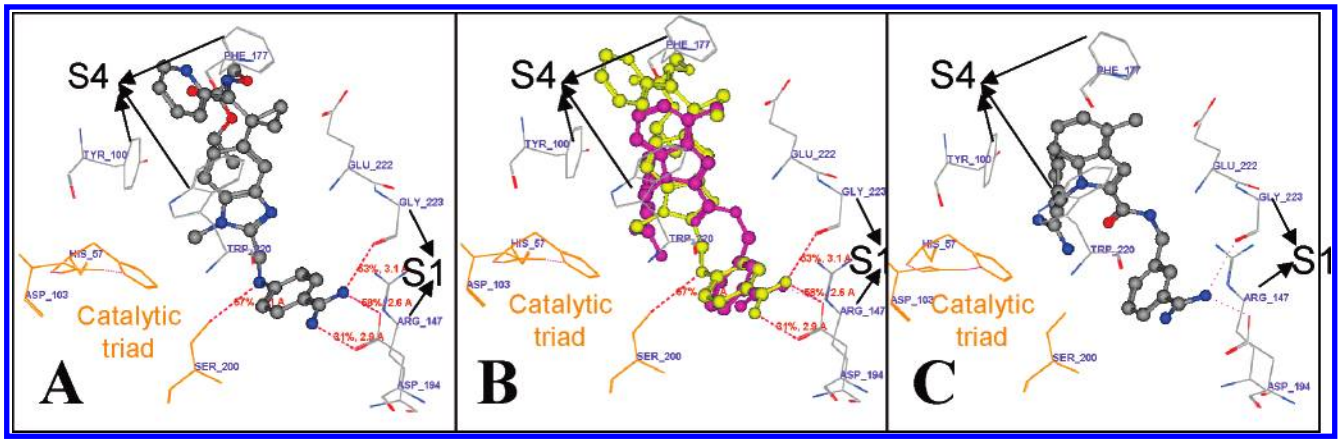

Figure 4. (A) The major interactions of BIBT0871 with the Factor Xa as seen in 1G2L. (B) The overlay of the docked structure of F1_47 and cocrystallized structure of BIBT0871 in blue and yellow, respectively. (C) Figure depicting some of the conserved hydrogen and hydrophobic bonding interactions between the docked structures of F1_47 with the fXa.

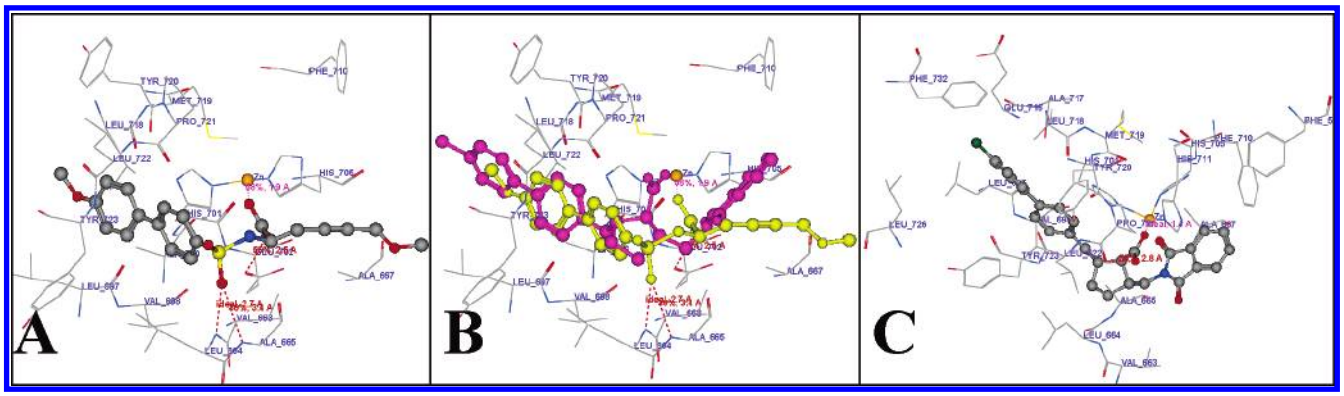

Figure 5. (A) The major interactions of MBS with the MMP-3 as seen in 1HY7. (B) The overlay of the docked structure of M1_59 and cocrystallized structure of MBS in blue and yellow, respectively. (C) Figure depicting some of the conserved hydrogen and hydrophobic bonding interactions between the docked structures of M1_59 with MMP-3.

their natural inhibitors leads to the accelerated destruction of connective tissues associated with the pathology of diseases such as arthritis, cancer, multiple sclerosis, and cardiovascular diseases.[49] The potential for using specific enzyme inhibitors as therapeutic agents to redress this balance has led to intensive research focused on the design, synthesis, and molecular deciphering of low-molecular-mass inhibitors of this family of proteins.

There are three major components in most of the MMP inhibitors for interactions as revealed by X-ray crystallographic analysis[49,50] (Figure 5[A]) namely the zinc binding group (ZBG), the peptidic backbone (consisting of P1−P4), and the pocket occupying side chain. Most MMP inhibitors are classified according to their ZBG. Inhibitor interaction at the zinc active-site plays a critical role in defining the binding mode and relative inhibitor potency. The majority of MMP inhibitors reported in the literature contain an effective zinc-binding group[51−53] e.g. hydroxamic acid, carboxylic acid, or sulfhydryl group that is either generally substituted with a peptide-like structure (P1−P4) that mimics the substrates that cleave or appended to smaller side chains that may interact with specific subsites (e.g., P1′, P2′, P3′, and P4′) within the active site. Although carboxylates exhibit weaker zinc binding properties than hydroxamates, they are known to show better oral bioavailability and are less prone to metabolic degradation. The expected loss of binding affinity after replacement of hydroxamates against carboxylates is faced by an adequate choice of elongated S1′ (a pocket which comprises residues HIS166, TYR168, and TYR155) directed substituents. Some structural requirements essential for achieving high binding affinity and selectivity are as follows: an acidic unit tightly anchored through four contact points, bidentate chelation of Zn2+, carbonyl groups for hydrogen bonding, more than two extra units for hydrogen bonds, and a hydrophobic moiety. It is observed that most of the important interactions as observed in the X-ray crystallography derived complex are also conserved in the docking mode of the most active molecule of the MMP-3 data set (Figure 5[C]). Further both the molecules superimpose to a considerable degree, and the molecule from the MMP-3 data set is observed to make similar interactions (Figure 5[B]).

In this case very good predictions of binding affinities were obtained: the experimentally determined IC50 and the pK_i predictions using MOE were 5.38 and 5.5, respectively, indicating the good performance of the scoring function implemented in MOE.

Further the binding affinity values of M1_59 for its receptor observed experimentally and predicted using MOE were 8.370 and 5.9, respectively, indicating again the poor performance of the scoring as well as the docking algorithm.

AChE Inhibitors. Acetylcholinesterase (AChE) is responsible for the terminating impulse transmission at cholinergic synapses by rapid hydrolysis of the neurotransmitter acetylcholine (ACh).[54] Its key role makes it the target of nerve

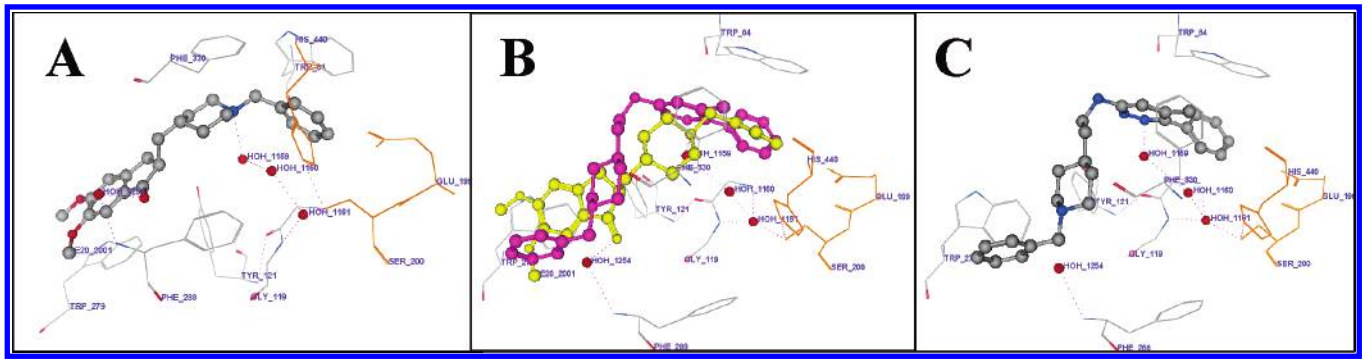

Figure 6. (A) The major interactions of E2020 with the *Tc*AChE as seen in 1EVE. (B) The overlay of the docked structure of A1_47 and cocrystallized structure of E2020 in blue and yellow, respectively. (C) Figure depicting some of the conserved hydrogen bonding and hydrophobic interactions between the docked structures of E2020 with the *Tc*AChE.

gases, pesticides, snake venom toxins, and anti-Alzheimer drugs.[54−59]

There are two principal binding sites in the AChE[60] (Figure 6[A]): the first an active site containing a catalytic triad, SER200, GLU327, and HIS440, and nearby residues such as the choline binding site: TRP84 that collectively effects the ACh catalysis reactions and second a peripheral anionic site (PAS) located near the enzyme surface at the mouth of the active site gorge. The residue TRP279 plays a very important role in ligand binding in the PAS as this interaction affects the enzymatic activity through a combination of steric blockade of ligands moving through the gorge and allosteric alteration of the catalytic triad conformation and efficiency. The gorge itself is a narrow hydrophobic channel with a length of ∼20 Å, connecting the PAS site to the active site. An acyl-binding pocket consists of residues GLY122, TRP236, PHE295, PHE297, and PHE338 and is responsible for interacting with the acetyl group. Early studies were mainly focused on ligands binding in the first active site (e.g., tacrine, amiridine, etc.). The recent efforts have been focused on finding novel ligands that bind to both sites in order to search for more potent reversible inhibitors (e.g., TAK-147, E2020, etc.), selectively favoring the inhibition of AChE rather than the related butyrylcholinesterase (BChE).[61]

An inspection of the binding mode of the most active molecule of AChE data set (A1_47) reveals that the binding mode of E2020 that is cocrystallized with AchE and the docked structure of A1_47 (the most active molecule in the *Tc*AChE data set) are similar to each other (Figure 6[B,C]). All the primary interacting functional groups superimpose on each other and consequently make similar interactions. The functional groups that superpose and make similar interactions are as follows: the benzyl moiety of E2020; one face of which displays the classic parallel $\pi-\pi$ stacking with the six-membered ring of the TRP84 indole and on the opposite face, the benzyl group makes a classic aromatic hydrogen bond (H bond) with a water molecule (WAT 1160) and superimposes with the 4-methyl-3-phenylpyridazine moiety of A1_47 and makes the corresponding interactions; the piperidine nitrogen which interacts with some of the residues of catalytic triad (Ser200 O); and the oxyanion hole, consisting of residues GLY118 N, GLY119 N, and GLY201 N, through a bridgehead water molecule (WAT 1161). Further this functional group makes a cation interaction with the phenyl ring of PHE330 and also makes an in-line 2.9 Å H bond with WAT 1159, which, in turn, makes H bonds with Tyr121 OH, with WAT 1158 and with WAT 1160, superimposes well on the corresponding basic nitrogen of the 4-methyl-3-phenyl-pyridazine moiety of E2, and is seen to make similar interactions and finally indanone ring of the dimethoxyindanone moiety, of E2020 ring stacks against the indole ring of TRP279 in the peripheral binding site, by a classical $\pi-\pi$ interaction. WAT 1249 lies in the plane of the indanone moiety, and H bonds to the methoxy group of E2020 O25; it is also H-bonded to GLU185 O1 of the symmetry-related crystal-lattice copy of the enzyme. The carbonyl on the five-membered ring of the indanone only interacts with AChE via edge-on van der Waals contacts with the aromatic rings of PHE331 and PHE290. It also makes indirect contact, via WAT 1254, with PHE288 N. The benzene ring part of the phenyl piperazine moiety of the E2 superimposes on the indanone moiety of E2020 and may make similar interactions with it.

Again in this case though a good pose prediction was obtained, the prediction of binding affinities of E2020 was rather dismal: the experimentally determined pK_i and the pK_i predictions using MOE were 8.475 and 5.8, respectively, indicating the very poor performance of the scoring function implemented in MOE.

Further the binding affinity values of A1_47 for its receptor observed experimentally and predicted using MOE were 8.370 and 6.597, respectively, indicating again the poor performance of the scoring as well as the docking algorithm.

Thus the overall analysis reveals that the moderately accurate docking poses were observed in the Jacobsson et al. data set, with the most active molecule of the data set overlapping well with the ligands conformation that was derived using X-ray crystallography, thus validating the use of this data set for the present analysis.

Comparison of the Individual Terms Used in LUDI and MOE Scoring Functions. A simple theoretical assessment of the scoring functions indicates that none of the individual scoring functions are completely suitable for scoring all the ligand−receptor interactions that have presently been characterized (Table 3). While MOE has the additional advantage of scoring the metal−ion complex interactions as part of its hydrogen bonding scoring function term, the LUDI_3 scoring function has a unique descriptor, which can score the $\pi-\pi$ interactions between the ligand and the protein.

In general it was observed that there were significant differences of the three LUDI functions incorporated in insightII. In theory, according to the LUDI manual all the three scores were derived using different energy functions. A comparison of the correlation coefficient values indicate

Table 3. Assessment of the Availability of Various Terms To Score the Various Ligand−Receptor Interactions

ligand−receptor interactions	LUDI-1	LUDI-2	LUDI-3	MOE
hydrophobic	√	√	√	√
hydrogen bonding	√	√	√	√
ionic interactions	√	√	√	
bridge waters	√	√	√	√
π-π			√	
metal ion complex				√
steric contacts	√	√	√	
loss in entropy (bond rotation)	√	√	√	√

Table 4. Correlation Coefficient Values between the Actual Activity and Activities Predicted Using Scoring Functions

	MOE	LUDI-1	LUDI-2	LUDI-3
$AChE_{obs}$	−0.261	0.138	−0.035	−0.191
fXa_{obs}	0.239	−0.094	−0.118	0.136
ER_{obs}	0.396	0.514	0.486	0.434
$MMP3_{obs}$	0.340	0.284	0.350	0.248

a significant difference in final score values of the three different scoring functions, LUDI-1 is similar to LUDI-2 but both are very dissimilar to LUDI-3. While the HB scores, Ionic scores, and Rot bond scores of the three different scoring functions are similar, a significant difference in the Lipo scores of the three different scoring functions was observed.

Quantitative Regression Models. None of the final scoring functions were able to explain the variation in activities as exemplified in the poor correlation coefficients (<0.6) between the binding affinities and final scores. The inability of the present scoring functions to predict the binding affinities of known binders is in agreement with the comparative scoring function analyses, where it was shown that it is difficult to find a good correlation between the binding affinity predicted using the empirical scoring functions and observed binding affinity (Table 4). Such poor performance may be attributed to the rigid docking algorithm used to obtain the pose and more significantly because of the inaccuracy of the empirical scoring functions used, which are typically trained via diverse sets of previously characterized ligand−receptor interactions. Further, no finite training set is likely to provide a perfect representation for all systems of interest because of the varying physicochemical conditions present in different receptors. Though molecular dynamics simulations do provide a natural means for quantifying both the entropic and enthalpic components of binding affinity for ligand−AChE interactions, however the extensive computational demands make molecular dynamics (MD) simulations prohibitively time-consuming for analysis of large compound collections. In such cases, it has been suggested that the best compromise may be to carry out simple docking studies and to explicitly train the scoring function to reproduce behavior in the system of interest.[62,63]

Hence all the individual terms of the scoring functions were used to train a fXa regression model using the PLS method. This model, though relatively better at explaining the variance in binding affinity in comparison to the final scoring terms, still demonstrated poor training ($R^2 = 0.461$, $R^2_{cv} = 0.196$) and test set predictions ($R^2_{testset} = 0.220$) (Figure 7) and also performed very poorly against the data set of 999 nonbinders wherein the majority of the molecules were predicted to be active in the range covered by the active

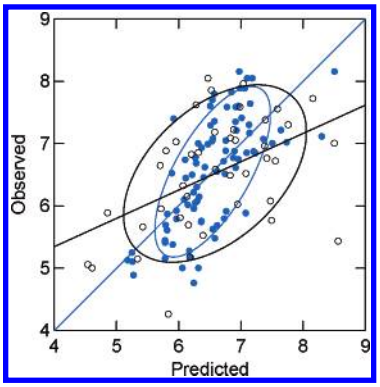

Figure 7. Plot of the observed versus predicted activities for the training (•) and test sets (o).

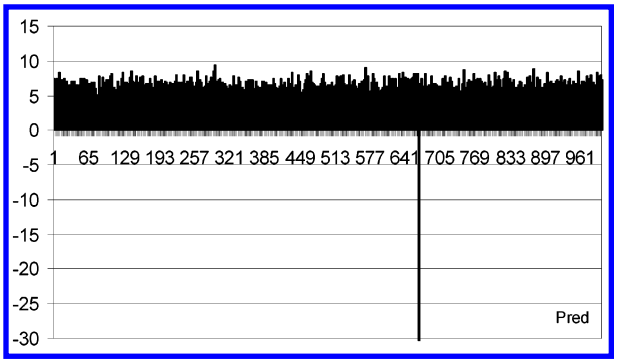

Figure 8. A chart depicting the range covered by the predicted values of the nonbinders.

Table 5. Prediction Statistics of FXa Binders and Nonbinders

	binders	nonbinders
no. of cases	128	999
minimum	4.541	−49.240
maximum	8.562	9.496
range	4.021	58.735
mean	6.558	6.069

molecules (Figure 8, Table 5). Thus the quantitative empirical scoring functions' based models built using only the data of active molecules are neither capable of accurately explaining the variance in the activities of binders nor are capable of distinguishing between binders and nonbinders. This observation is consistent with the fact that the currently available scoring functions cannot accurately capture the important dynamic or entropic effects that are difficult to be rigorously represented,[63−65] and models trained with the data of only actives are not capable of distinguishing between actives and inactives. The observed low predictivity of interaction-energy based models may be due to the scoring functions not being able to accommodate small deviations in atomic coordinates. Such deviations are likely to occur may be due to the small errors in atomic coordinates even in well-refined X-ray structures which have influence on the energy landscape of ligand binding.[64] Further there may be errors in correct placement of the ligand in the binding pocket and in optimization of its noncovalent interactions with the protein. In addition the problem is acquainted with the limitations in understanding of the physics and thermodynamics of ligand binding, especially solvation and entropic effects, which are at the moment fully neglected by interaction-energy based models.[65]

Table 6. Comparison of the Binary QSAR Discriminatory Power with Three Other Multivariate Statistical Methods viz. PLS Discriminant Analysis, Rule-Based Methods, and Bayesian Classification, on the AChE Data Set

data set	method	accuracy of active	accuracy of inactive	overall accuracy	accuracy	precision	recall	EF
AChE	MOE binary (p=0.5)	0.222	0.994	0.954	0.934	0.308	0.222	6.000
	MOE binary (p=0.0513)	0.222	0.994	0.954	0.860	0.196	0.556	3.824
	MOE binary (p=0.5)[a]	0.861	0.972	0.917	0.611	0.078	0.860	4.252
	C-W20				0.869	0.220	0.611	4.30
	Bayes, pact = 0.0513				0.903	0.250	0.444	4.90
	PLS-DA				0.681	0.110	0.722	2.1
	PLS-DA, limit 0.2				0.766	0.084	0.500	2.1
	ICM score, $T = -26.64$				0.484	0.084	0.907	1.6

[a] Equal number of active and inactive compounds considered.

Assessment of the Discriminative Ability of Binary QSAR. In view of the above results, we focused our attention for developing classification models, using the binary QSAR methodology that can be trained on the data of both binders and nonbinders. To compare the discriminatory power of binary QSAR vis-à-vis three other multivariate statistical methods viz. PLS discriminant analysis, rule-based methods, and Bayesian classification used by Jacobssson et al., the entire AChE data set was used as a benchmark. The present studies reveal that the binary QSAR model presents a statistically superior model (accuracy = 0.934, precision = 0.308, recall = 0.22, EF = 6.000) in comparison to the best model (accuracy = 0.903, precision = 0.250, recall = 0.444, EF = 4.90) derived by Jacobsson et al. (Table 6). Since only one complete data set was provided by the original authors, only the results of the same are presented. However it is expected that binary QSAR may, in general, perform slightly better or may be comparable results to the three other multivariate statistical methods used for discriminating between binders and nonbinders. Further to address sensitivity of the binary QSAR models to the unequal number of active and inactive compounds, two strategies have been explored. In the first strategy the a priori probability threshold value was tweaked to the values used by Jacobsson et al., and in the second strategy an equal number of active and inactive data points were considered for analysis. Though considering an equal number of active and inactive data points in the development of the binary QSAR resulted in significant improvement in the discriminative ability related parameters for the training set (accuracy of active = 0.861, accuracy of inactive = 0.972, overall accuracy = 0.917), there was however a substantial decrease in the corresponding parameters for the test set (accuracy = 0.611, precision = 0.078, recall = 0.860, EF = 4.252). Further tweaking the a priori probability, described as 'binary threshold' in MOE, had no impact on the parameters for the training set (accuracy of active = 0.222, accuracy of inactive = 0.994, overall accuracy = 0.954) and did not even improve the discriminative ability for the test set (accuracy = 0.860, precision = 0.196, recall = 0.556, EF = 3.824) (Table 6).

Hence only the original method suggested for binary QSAR was implemented in all the subsequent binary QSAR model development.

Performance of Binary QSAR Models Using Ludi and MOE Scoring Functions. Since binary QSAR uses a PCA to extract the intercorrelating descriptors into principle components, no attempt was made to exclude any of the individual descriptors from the binary QSAR model development. For each of the five target proteins, five different binary QSAR models were constructed and assessed for their discriminative ability. Of these four were derived using the individual scoring terms from different scoring functions implemented in LUDI (LUDI_1, LUD1_2, and LUDI_3) and MOE (MOE), and the fifth was derived using all the individual terms of all the scoring functions (All).

The discriminative ability of both the training and test sets was assessed using various parameters reported in the Materials and Methods section. In short the training set's discriminative ability was assessed using the 'accuracy on active', 'accuracy on inactive', and 'accuracy' parameters, while that of the test set was assessed using parameters 'accuracy_test', 'precision', 'recall', and 'EF'. However to aid in the selection of the best binary QSAR model for each model system, we consider an optimum of the parameters accuracy (training set), precision (test set), and EF (test set).

According to the above criteria, the binary QSAR models derived using MOE scoring function performed well on the ER_mimics and MMP-3 data sets (Table 6), while the binary QSAR model based on all the available scoring functions performed well for AChE and fXa data sets (Table 7). However for the ER_toxin data set the binary QSAR model using LUDI-3 scoring functions showed the best discriminative ability (Table 7). Further the best model identified for each data set was comparable to the corresponding best models reported by Jacobsson et al. (Table 7, Figure 9). In three of the five data sets, the binary QSAR models outperformed the classification models reported by Jacobsson et al. for the corresponding proteins in terms of their discriminatory ability represented by the EF parameter. While the EF values of the best binary QSAR for fXa, MMP-3, AChE are 7.870, 17.650, and 7.020, the EF values of the best models reported by Jacobsson et al. for the corresponding proteins were 7.3, 16.7, and 4.9. For ER_mimcs and ER_toxins the best classifications models reported by Jacobsson et al. were superior to the best binary QSAR models reported in the present paper. The discriminatory ability quantified in terms of the EF values of the best models of Jacobsson et al. for ER_mimics and ER_toxins were 28.8 and 9.6, while the EF values of the best binary QSAR models for the corresponding models are 24.643 and 8.756. In addition a comparison of the binary QSAR models derived using the final scores (all_final) with the binary QSAR models derived using individual terms reveals a mixed performance, with a better discriminant ability for ER_mimics, equal in the case of ER_toxins and MMP-3 and very poor in the case of fXa and AChE. The EF value for the

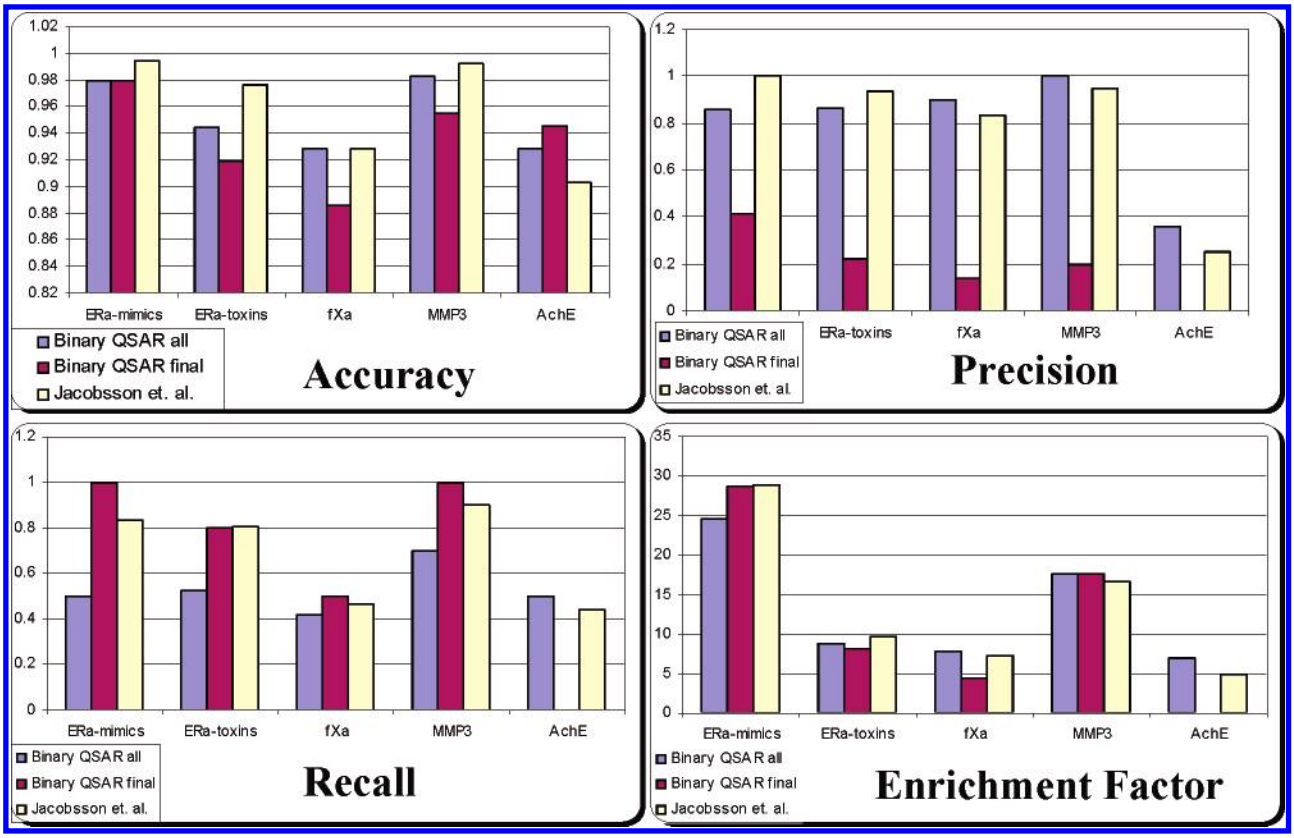

Figure 9. Graphical representation of comparative analysis of accuracy, precision, recall, and enrichment of the three methods for the different target proteins and ligand sets.

Table 7. Performance of the Binary QSAR's Discriminatory Power with Different Scoring Functions for the Training Sets of the Five Different Data Sets

data set	method	accuracy on active	accuracy on inactive	accuracy	accuracy (test set)	precision (test set)	recall (test set)	EF (test set)
ERa-mimics	LUDI-1	0.971	0.583	0.985	0.959	0.400	0.333	11.367
	LUDI-2	0.972	0.583	0.986	0.968	0.556	0.417	15.787
	LUDI-3	0.986	0.708	0.983	0.959	0.429	0.500	12.179
	MOE	0.978	0.375	1.000	0.980	0.857	0.500	24.643
	All	0.993	0.833	0.998	0.962	0.471	0.667	13.529
ERa-toxins	LUDI-1	0.927	0.473	0.977	0.937	0.760	0.528	7.706
	LUDI-2	0.945	0.514	0.992	0.940	0.818	0.500	8.295
	LUDI-3	0.946	0.514	0.994	0.945	0.864	0.528	8.756
	MOE	0.924	0.311	0.992	0.930	0.778	0.389	7.972
	All	0.954	0.635	0.989	0.938	0.724	0.583	7.422
fXa	LUDI-1	0.931	0.488	0.988	0.915	0.867	0.302	7.578
	LUDI-2	0.926	0.453	0.986	0.912	0.813	0.302	7.105
	LUDI-3	0.936	0.500	0.992	0.926	0.857	0.419	7.495
	MOE	0.907	0.279	0.988	0.915	0.789	0.349	6.903
	All	0.944	0.628	0.985	0.928	0.900	0.419	7.870
MMP-3	LUDI-1	0.948	0.500	0.974	0.938	0.429	0.300	7.564
	LUDI-2	0.950	0.550	0.974	0.946	0.533	0.400	9.413
	LUDI-3	0.935	0.375	0.968	0.958	0.667	0.500	11.767
	MOE	0.987	0.775	1.000	0.983	1.000	0.700	17.650
	All	0.984	0.725	1.000	0.946	0.529	0.450	9.344
AChE	LUDI-1	0.983	0.500	0.959	0.934	0.368	0.389	7.184
	LUDI-2	0.983	0.528	0.960	0.934	0.368	0.389	7.184
	LUDI-3	0.986	0.556	0.964	0.926	0.333	0.444	6.500
	MOE	0.998	0.056	0.950	0.943	0.250	0.056	4.875
	All	0.985	0.639	0.967	0.929	0.360	0.500	7.020

binary QSAR models derived from the final scores for ER_mimics, ER_toxins, fXa, MMP-3, and AChE were 28.750, 8.200, 4.372, 17.650, and 0.000, while those of the binary QSAR models derived from individual terms of the scoring function are 24.643, 8.756, 7.870, 17.650, and 7.020 (Tables 8 and 9 and Figure 9).

Since the ER binding site was mainly hydrophobic hence almost all the scoring functions performed reasonably well for both mimics and toxins, while in the case of MMP-3 binders MOE had the required scoring function for the important, explicit metal−ion interaction term and hence showed superior classifications ability in comparison to all

Table 8. Performance of the Binary QSAR's Discriminatory Power with Different the Final Terms (all_final) of Scoring Functions of the Five Different Data Sets

data set	method	accuracy on active	accuracy on inactive	accuracy	accuracy (test set)	precision (test set)	recall (test set)	EF (test set)
Era-mimics	all_final	0.375	1.000	0.978	0.980	0.417	1.000	28.750
Era-toxins	all_final	0.243	0.997	0.922	0.919	0.222	0.800	8.200
fXa	all_final	0.291	0.983	0.904	0.886	0.140	0.500	4.372
MMP-3	all_final	0.325	0.991	0.953	0.955	0.200	1.000	17.650
AChE	all_final	0.028	0.998	0.949	0.946	0.000	0.000	0.000

Table 9. Comparison of Discriminative Abilities of the Binary QSAR's Derived Using the Individual Terms of the Scoring Functions, Final Terms of Scoring Functions, and the Classification Models Derived by Jacobsson et al. on Five Different Data Sets

data set	method	accuracy on active	accuracy on inactive	accuracy	accuracy (test set)	precision (test set)	recall (test set)	EF (test set)
ERa-mimics	MOE	0.978	0.375	1.000	0.980	0.857	0.500	24.643
	all_final	0.375	1.000	0.978	0.980	0.417	1.000	28.750
	Jacobsson et al.				0.994	1.000	0.833	28.8
ERa-toxins	LUDI-3	0.946	0.514	0.994	0.945	0.864	0.528	8.756
	all_final	0.243	0.997	0.922	0.919	0.222	0.800	8.200
	Jacobsson et al.				0.976	0.935	0.806	9.6
fXa	all	0.944	0.628	0.985	0.928	0.900	0.419	7.870
	all_final	0.291	0.983	0.904	0.886	0.140	0.500	4.372
	Jacobsson et al.				0.928	0.833	0.465	7.3
MMP-3	MOE	0.987	0.775	1.000	0.983	1.000	0.700	17.650
	all	0.325	0.991	0.953	0.955	0.200	1.000	17.650
	Jacobsson et al.				0.992	0.947	0.900	16.7
AChE	all	0.985	0.639	0.967	0.929	0.360	0.500	7.020
	all_final	0.028	0.998	0.949	0.946	0.000	0.000	0.000
	Jacobsson et al.				0.903	0.250	0.444	4.9

other scoring functions. The poor discriminative power of the empirical scoring functions observed in the present studies and also by Jacobsson et al. for the AChE data set may be due to the difficulties associated in docking of the ligands to this target because of several factors. These include wrong docking modes, nonconsideration of water molecules in a system where most of the ligand−protein hydrogen bonds are mediated through bridgehead water molecules, and scoring based on a rigid body protein docking in protein known to induce significantly under the influence of the ligand.[13] In the present studies the AChE binary QSAR model developed using all the scoring function terms exhibited superior discriminative power (precession=0.36, EF=7.020) in comparison to the best binary classification model (precession=0.25, EF=6.0) reported by Jacobsson et al. This can be attributed to the availability of a scoring function term in LUDI-3 that can measure the very important $\pi-\pi$ stacking interactions.

Further a comparison of the discriminatory power of the binary QSAR models developed using just the final scores (all_final) and the best models selected above (Table 9) substantiates the fact that when regression based empirical scoring functions are used, it is better to consider the individual terms of the scoring functions than the final scores.

Finally, a comparison of the discriminant power of the best models for various data sets obtained by Jacobsson et al. with the best models obtained for the corresponding data sets in the present analysis indicates that the models are comparable in terms of parameters that describe the discriminative ability. Hence the classification models developed using the classification algorithm, binary QSAR, and the scoring functions (as descriptors) implemented in LUDI and MOE were found to be comparable to the models developed by Jacobsson et al. wherein classifiers namely PLS discriminant analysis, rule-based methods, and Bayesian classification, and seven scoring functions implemented in ICM and Cscore were used.

CONCLUSION

The application of binary QSAR based multivariate analysis to LUDI and MOE scoring functions has performed well in discriminating between binders and nonbinders on five different data sets, which were also a subject of such assessments previously. Classification analysis on the AChE data set has indicated that the binary QSAR based classification system is slightly more effective in comparison to the three classifiers used by Jacobsson et al. viz. PLS discriminant analysis, rule-based methods, and Bayesian classification. Further it has been found that either MOE or LUDI or a combination of both are able to discriminate well between binders and nonbinders as observed in the present studies on a wide range of protein−ligand systems viz. ER, AChE, fXa, and MMP3. Further the quantitative model derived using PLS for the fXa data set indicates that quantitative models based on empirical scoring functions are neither capable of accurately explaining the variance in the activities of known binders nor capable of distinguishing between binders and nonbinders. Thus the overall analysis suggests that the models derived using binary QSAR using LUDI and MOE scoring functions may be useful as a screening layer in a multilayered virtual screening paradigm.

ACKNOWLEDGMENT

The authors are thankful to Dr. Micael Jacobsson for providing us the complete AChE data set and for many useful discussions. The technical support of Mr. A. S. Kushwaha and the financial support of CSIR and DST, New Delhi are also gratefully acknowledged. This is CDRI communication number 6748.

REFERENCES AND NOTES

(1) Mizutani, M. Y.; Itai, A. Efficient Method for High-Throughput Virtual Screening Based on Flexible Docking: Discovery of Novel Acetylcholinesterase Inhibitors. *J. Med. Chem.* **2004**, *47*, 4818−4828.

(2) Irwin, J. J.; Shoichet, B. K. ZINC-a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−82.

(3) Schapira, M.; Abagyan, R. A.; Totrov, M. Nuclear hormone receptor targeted virtual screening. *J. Med. Chem.* **2003**, *46*, 3045−3059;

(4) Cavasotto, C. N.; Abagyan, R. A. Protein flexibility in ligand docking and virtual screening to protein kinases. *J. Mol. Biol.* **2004**, *337*, 209−225.

(5) Schapira, M.; Raaka, B. M.; Samuels, H. H.; Abagyan, R. A. Rational discovery of novel nuclear hormone receptor antagonists. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1008−13.

(6) Schapira, M.; Raaka, B. M.; Samuels, H. H.; Abagyan, R. A. In silico discovery of novel retinoic acid receptor agonist structures. *BMC Struct. Biol.* **2001**, *1*, 1.

(7) Liu, B.; Zhou, J. SARS-CoV protease inhibitors design using virtual screening method from natural products libraries. *J. Comput. Chem.* **2005,** *26*, 484−90.

(8) Evers, A.; Klabunde, T. Structure-based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the Alpha1A Adrenergic Receptor. *J. Med. Chem.* **2005**, *48*, 1088−1097.

(9) Liu, Z.; Huang, C.; Fan, K.; Wei, P.; Chen, H.; Liu, S.; Pei, J.; Shi, L.; Li, B.; Yang, K.; Liu, Y.; Lai, L. Virtual Screening of Novel Noncovalent Inhibitors for SARS-CoV 3C-like Proteinase. *J. Chem. Inf. Model.* **2005**, *45*, 10−17.

(10) Li, C.; Xu, L.; Wolan, D. W.; Wilson, I. A.; Olson, A. J. Virtual Screening of Human 5-Aminoimidazole-4-carboxamide Ribonucleotide Transformylase against the NCI Diversity Set by Use of AutoDock to Identify Novel Nonfolate Inhibitors. *J. Med. Chem.* **2004**, *47*, 6681−6690.

(11) Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem.* **2002**, *45*, 2213-21.

(12) Gruneberg, S.; Stubbs, M. T.; Klebe, G. Successful virtual screening for novel inhibitors of human carbonic anhydrase: strategy and experimental confirmation. *J. Med. Chem.* **2002**, *45*, 3588−602.

(13) Jacobsson, M.; Liden, P.; Stjernschantz, E.; Bostrom, H.; Norinder, U. Improving structure-based virtual screening by multivariate analysis of scoring data. *J. Med. Chem.* **2003**, *46*, 5781−9.

(14) So, S. S.; Karplus, M. Evaluation of designed ligands by a multiple screening method: application to glycogen phosphorylase inhibitors constructed with a variety of approaches. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 613−47.

(15) Kontoyianni, M.; Sokol, G. S.; McClellan, L. M. Evaluation of library ranking efficacy in virtual screening. *J. Comput. Chem.* **2005**, *26*, 11−22.

(16) Marsden, P. M.; Puvanendrampillai, D.; Mitchell, J. B.; Glen, R. C. Predicting protein−ligand binding affinities: a low scoring game? *Org. Biomol. Chem.* **2004**, *21*, 3267−73.

(17) Wang, R.; Lu, Y.; Fang, X.; Wang, S. An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein−ligand complexes. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2114−25.

(18) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins: Struct., Funct. Genet.* **2004**, *57*, 225−42.

(19) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins: Struct., Funct. Genet.* **2004**, *56*, 235−49.

(20) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L 3rd. Assessing scoring functions for protein−ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032−47.

(21) Krovat, E. M.; Langer, T. Impact of scoring functions on enrichment in docking-based virtual screening: an application study on renin inhibitors. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1123−9.

(22) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.

(23) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100−5109.

(24) Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Modell.* **2002**, *20*, 281−295.

(25) Klon, A. E.; Glick, M.; Davies, J. W. Combination of a naive Bayes classifier with consensus scoring improves enrichment of high-throughput docking results. *J. Med. Chem.* **2004**, *47*, 4356−9.

(26) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins: Struct., Funct. Genet.* **1988**, *4*, 31−47.

(27) Labute, P. Binary QSAR: a new method for the determination of quantitative structure−activity relationships. In *Proceedings of the Pacific Symposium on Biocomputing'99*; Altman, R. B., Dunker, A. K., Hunter, L., Klein, T. E., Lauderdale, K., Eds.; World Scientific: New Jersey, pp 444−455.

(28) Labute, P.; Nilar, S.; Williams, C. A probabilistic approach to high throughput drug discovery. *Comb. Chem. High Throughput Screening* **2002**, *5*, 135−45.

(29) Gao, H.; Williams, C.; Labute, P.; Bajorath, J. Binary quantitative structure−activity relationship (QSAR) analysis of estrogen receptor ligands. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 164−8.

(30) Feller, W. *An Introduction to Probability Theory and its Applications*; Wiley & Sons Inc.: New York, 1950; Vol. 1.

(31) Bohm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein−ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243−56.

(32) Bohm, H. J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 309−23.

(33) Accelrys molecular modeling software package InsightII 2000, Webpage. http://www.accelrys.com/products/insight/sbd_modules.html

(34) Chemical Computing Group, MOE, Quebec, Canada, 2004.

(35) Chemical Computing Group, Svl Exchange Webpage. http://svl.chemcomp.com/, 2004.

(36) Tsai, M. J.; O'Malley, B. W. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* **1994**, *63*, 451−486.

(37) Beato, M.; Herrlich, P.; Schutz, G. Steroid hormone receptors: many actors in search of a plot. *Cell* **1995**, *83*, 851−857.

(38) Dutertre, M.; Smith, C. L. Molecular mechanisms of selective estrogen receptor modulator (SERM) action. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 431−7.

(39) Gradishar, W. J.; Jordan, V. C. Clinical potential of new antiestrogens. *J. Clin. Oncol.* **1997**, *15*, 840−852.

(40) Jordan, V. C. Antiestrogenic action of raloxifene and tamoxifen: today and tomorrow. *J. Natl. Cancer Inst.* **1998**, *90*, 967−971.

(41) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustaffson, J.-A.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389*, 753−758.

(42) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **1998**, *95*, 927

(43) Choi-Sledeski, Y. M.; Kearney, R.; Poli, G.; Pauls, H.; Gardner, C.; Gong, Y.; Becker, M.; Davis, R.; Spada, A.; Liang, G.; Chu, V.; Brown, K.; Collussi, D.; Leadley, R., Jr.; Rebello, S.; Moxey, P.; Morgan, S.; Bentley, R.; Kasiewski, C.; Maignan, S.; Guilloteau, J. P.; Mikol, V. Discovery of an orally efficacious inhibitor of coagulation factor Xa which incorporates a neutral P1 ligand. *J. Med. Chem.* **2003**, *46*, 681−4.

(44) Ieko, M.; Tarumi, T.; Takeda, M.; Naito, S.; Nakabayashi, T.; Koike, T. Synthetic selective inhibitors of coagulation factor Xa strongly inhibit thrombin generation without affecting initial thrombin forming time necessary for platelet activation in hemostasis. *J. Thromb. Haemostasis* **2004**, *2*, 612−8.

(45) Dewerchin, M.; Liang, Z.; Moons, L.; Carmeliet, P.; Castellino, F. J.; Collen, D.; Rosen, E. D. Blood coagulation factor X deficiency causes partial embryonic lethality and fatal neonatal bleeding in mice. *Thromb. Haemostasis* **2000**, *83*, 185−90.

(46) Nar, H.; Bauer, M.; Schmid, A.; Stassen, J. M.; Wienen, W.; Priepke, H. W.; Kauffmann, I. K.; Ries, U. J.; Hauel, N. H. Structural basis for inhibition promiscuity of dual specific thrombin and factor Xa blood coagulation inhibitors. *Structure (Cambr.)* **2001**, *9*, 29−37.

(47) Nagase, H. Matrix Metalloproteinases. In *Zinc Metalloproteases in Health and Disease*; Hooper, N. M., Ed.; Taylor & Francis: London, 1996; pp 153−204.

(48) Nagase H. Activation mechanisms of matrix metalloproteinases. *Biol. Chem.* **1997**, *378*, 151−60.

(49) Natchus, M. G.; Bookland, R. G.; Laufersweiler, M. J.; Pikul, S.; Almstead, N. G.; De, B.; Janusz, M. J.; Hsieh, L. C.; Gu, F.; Pokross, M. E.; Patel, V. S.; Garver, S. M.; Peng, S. X.; Branch, T. M.; King, S. L.; Baker, T. R.; Foltz, D. J.; Mieling, G. E. Development of new

carboxylic acid-based MMP inhibitors derived from functionalized propargylglycines. *J. Med. Chem.* **2001**, *44*, 1060−1071
(50) Pavlovsky, A. G.; Williams, M. G.; Ye, Q. Z.; Ortwine, D. F.; Purchase, C. F., II.; White, A. D.; Dhanaraj, V.; Roth, B. D.; Johnson, L. L.; Hupe, D.; Humblet, C.; Blundell, T. L. X-ray Structure of Human Stromelysin Catalytic Domain Complexed With Nonpeptide Inhibitors: Implications for Inhibitor Selectivity. *Protein Sci.* **1999**, *8*, 1455−1462.
(51) Morphy, J. R.; Millican, T. A.; Porter, J. R. Matrix metalloproteinaseinhibitors: Current status. *Curr. Med. Chem.* **1995**, *2*, 743−762.
(52) Beckett, R. P.; Davidson, A. H.; Drummond, A. H.; Huxley, P.; Whittaker, M. Recent advances in matrix metalloproteinase inhibitor research. *Drug Discovery Today* **1996**, *1*, 16−26.
(53) Beckett, R. P.; Whittaker, M. Matrix metalloproteinase inhibitors. *Expert Opin. Ther. Pat.* **1998**, *8*, 259−282.
(54) Taylor, P. *The Pharmacological Basis of Therapeutics*, 9th ed.; Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman, A. G., Eds.; McGraw-Hill: New York, 1996; pp 161−176.
(55) Millard, C. B.; Broomfield, C. A. Anticholinesterases: medical applications of neurochemical principles. *J. Neurochem.* **1995**, *64*, 1909−1918.
(56) Davis, K. L.; Powchik, P. Tacrine. *Lancet* **1995**, *345*, 625−630.
(57) Nightingale, S. L. Donepezil approved for treatment of Alzheimer's disease. *J. Am. Med. Assoc.* **1997**, *277*, 10.
(58) Casida, J. E.; Quistad, G. B. Golden age of insecticide research: past, present, or future? *Annu. Rev. Entomol.* **1998**, *43*, 1−16.
(59) Martin, R. J. Modes of action of anthelmintic drugs. *Vet. J.* **1997**, *154*, 11−34.
(60) Kryger, G.; Silman, I.; Sussman, J. L. Structure of acetylcholinesterase complexed with E2020 (Aricept): implications for the design of new anti-Alzheimer drugs. *Struct. Fold. Des.* **1999**, *7*, 297.
(61) Du, D. M.; Carlier, P. R. Development of bivalent acetylcholinesterase inhibitors as potential therapeutic drugs for Alzheimer's disease. *Curr. Pharm. Des.* **2004**, *10*, 3141−56.
(62) Guo, J.; Hurley, M. M.; Wright, J. B.; Lushington, G. H. A docking score function for estimating ligand−protein interactions: application to acetylcholinesterase inhibition. *J. Med. Chem.* **2004**, *21*, 47, 5492−500.
(63) Prathipati, P.; Pandey, G.; Saxena, A. K. CoMFA and docking studies on glycogen phosphorylase *a* inhibitors as antidiabetic agents. *J. Chem. Inf. Model.* **2005**, *45*, 136−45.
(64) Tame J. R. H. Scoring functions: A view from the bench. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 99−108.
(65) Donini O. A. T.; Kollman P. A. Calculation and prediction of binding free energies for the matrix metalloproteinases. *J. Med. Chem.* **2000**, *43*, 4180−4188.

CI050120W