Journal of Molecular Graphics and Modelling 28 (2010) 870–882

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

# Benchmarking docking and scoring protocol for the identification of potential acetylcholinesterase inhibitors

Zaheer-ul-Haq[a,*], Sobia Ahsan Halim[a], Reaz Uddin[a], Jeffry D. Madura[b,1]

[a] Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi-75270, Pakistan
[b] Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Avenue, 308 Mellon Hall, Pittsburgh, PA 15282, USA

ARTICLE INFO

*Article history:*
Received 31 January 2010
Received in revised form 14 March 2010
Accepted 16 March 2010
Available online 4 April 2010

*Keywords:*
Acetylcholinesterase
Comparative molecular docking
Scoring function

ABSTRACT

Acetylcholinesterase (AChE) plays a crucial role in nerve impulse transmission at cholinergic synapses by rapid hydrolysis of the neurotransmitter acetylcholine (ACh). AChE has become an important drug target because partial inhibition of AChE results in modest increase in ACh levels that can have therapeutic benefits, thus AChE inhibitors have proved useful in the symptomatic treatment of Alzheimer's disease. To establish an effective docking protocol for virtual screening of AChE, a comparative molecular docking study was performed. For this purpose six docking/scoring approaches (AutoDock, FlexX, MOE, Surflex-Dock, GOLD and FRED) were compared to determine their ability to reproduce the binding poses in twenty six complexes of AChE. Docking accuracy was evaluated by calculating the RMSD of the docked complexes. FRED was found to be the best in reproducing the experimental pose by placing it near the top of its ranking. The performance of scoring functions was evaluated by identifying known actives out of large database of inactive compounds. A dataset of 5000 "drug like" decoys were retrieved from NCI database and docked into the binding site of AChE with six known inhibitors using FRED in combination with five scoring functions, i.e., Chemgauss2, Chemgauss3, ChemScore, Shapegauss and PLP. The poses obtained by FRED were re-scored using GOLD score, ChemScore and ASP as implemented in GOLD while G_Score, D_Score, ChemScore and PMF as implemented in the CScore module of SYBYL7.3. D_Score presented significantly better enrichment than others and 50% of the active inhibitors were identified in top 20% of the ranked database.



## 1. Introduction

Acetylcholinesterase (AChE) is functionally known for the rapid hydrolysis of the neurotransmitter acetylcholine (ACh) at cholinergic synapses. Medicinally it has been targeted in the treatments for Alzheimer's disease, myasthenia gravis, glaucoma, and in the recovery of victims of nerve agent exposure. Cholinesterase inhibitors which prevent the hydrolysis of acetylcholine have been approved for the symptomatic treatment of the Alzheimer's disease for over a decade. The X-ray structure of *Tc*AChE illustrated that active site of enzyme is a deep and narrow gorge with the catalytic triad composed of residues Ser200, Glu327 and His440, a peripheral anionic site at the entry centered by residue Trp279, and a hydrophobic site consisting of some aromatic residues such as Trp84 and Phe330 [1,2]. From the comparison of published X-ray crystal structures of various *Tc*AChE complexes, it is obvious that side-chain orientation of Phe330 is the site with flexibility, which is responsible for substrate trafficking down the gorge. Three major orientations of Phe330 have been observed (Fig. 1a). The *Tc*AChE-E2020 complex (PDB id: 1EVE) is characterized by the open-gate conformation, while the complex with Tacrine (PDB id: 1ACJ) displays closed one, and half-open conformation is observed in *Tc*AChE-TMTFA (PDB id: 1AMN) [1,2].

As important therapeutic target, AChE has attracted great interest in the search of potent and selective inhibitors. The availability of several experimentally determined three-dimensional structures of AChE co-crystallized with various inhibitors provides an

*Abbreviations:* AChE, acetylcholinesterase; ACh, acetylcholine; ASP, astex statistical potential; CT, catalytic triad; Dm, Drosophila Melanogaster; FRED, fast rigid exhaustive docking; GA, genetic algorithm; GOLD, genetic optimization for ligand docking; Mm, mus musculus; NCI, National Cancer Institute; NAP, (−)-S-3-[1-(dimethylamino) ethyl] phenol; PAS, peripheral anionic site; PDB, protein data bank; PLP, piecewise linear potential; PMF, potential of mean force; RMSD, root mean square deviation; SA, simulated annealing; Tc, torpedo californiica; TMTFA, m-(N,N,N-trimethylammonio)-2,2,2-trifluoroacetophenone; TS, tabu search.

* Corresponding author. Tel.: +92 21 111222292x309 (UAN); fax: +92 21 4819018/19.
*E-mail addresses:* zaheer.qasmi@iccs.edu, zaheer_qasmi@hotmail.com ( Zaheer-ul-Haq), jdmadura@me.com (J.D. Madura).
*URL:* http://www.sites.google.com/site/zaheerqasmi.
[1] Tel.: +1 412 396 6341; fax: +1 412 396 5683.

1093-3263/$ – see front matter. 
doi:10.1016/j.jmgm.2010.03.007

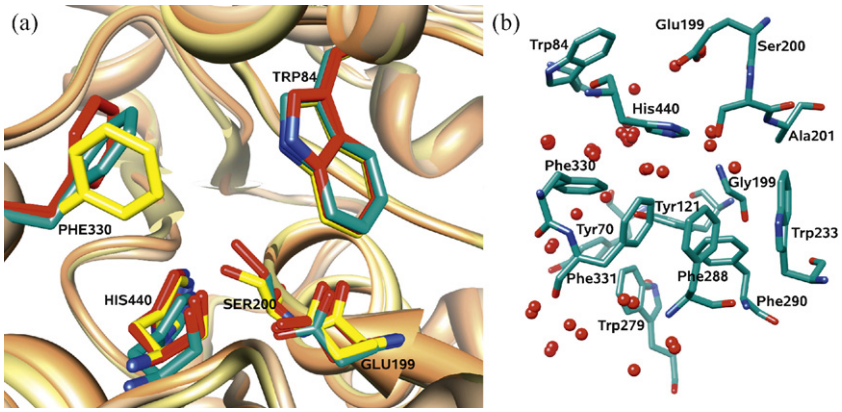



**Fig. 1.** (a) Superimposed view of three most dominant conformations of Phe330 in the binding pocket of *Tc*AChE. The most important amino acid residues Glu199, Ser200, Phe330 and His440 are also overlaid. Three complexes (PDB id: 1ACJ, 1EVE and 1AMN) are presented. 1ACJ (red) is the closed form, 1AMN (Green) shows half-open form while 1EVE (yellow) represents open form of AChE gorge. (b) Conserved water molecules in the active site gorge of AChE.

excellent basis for structure based approaches to discover the new inhibitors.

Virtual screening has become an established and important tool for lead identification as starting point for chemical optimization in drug discovery programs [3–6]. When a high resolution structure of the biological target of interest is available, the most common methodology for performing virtual screening involves the use of docking algorithms in which conformational sampling methods are used to position the ligand into the active site of the target macromolecule. There are several conformational sampling methods, such as Genetic Algorithms, Monte Carlo simulation, and Simulated Annealing used in docking calculations. All sampling methods are guided by a function that evaluates the fitness of interactions between the protein and ligand. In the docking step, many ligand conformations are generated and need to be ranked, i.e., scored. In the scoring step, a scoring function is used to evaluate the protein-ligand affinity. Scoring functions are important because the final predicted conformations are ranked according to the score.

There are three groups of scoring functions: force-field based methods, empirical scoring functions, and knowledge-based potentials. Force-field-based scoring functions [7,8] apply classical molecular mechanics energy functions. They approximate the binding free energy of protein-ligand complexes by a sum of van der Waals and electrostatic interactions. Solvation is usually taken into account using a distance-dependent dielectric function, although solvent models based on continuum electrostatics have been developed. Nonpolar contributions are usually assumed to be proportional to the solvent-accessible surface area [7,8]. A drawback is that the energy landscapes associated with force field potentials are generally rugged, and, therefore, minimization is required prior to any energy evaluation.

Empirical scoring functions estimate the binding free energy by summing interaction terms derived from weighted structural parameters [9]. The weights are determined by fitting the scoring function to experimental binding data of a training set of protein-

**Table 1**
Protein-ligand complexes (**1**–**26**) used for the docking purpose.

| S# | PDB | Water molecules | Source | Ligand binding site | Ligand's rotatable bonds | Resolution (Å) | References |
|---|---|---|---|---|---|---|---|
| 1 | 1ACJ | 634, 643 | *Tc* | CT + PAS | 0 | 2.8 | [46] |
| 2 | 1ACL | 642 | *Tc* | CT + PAS | 11 | 2.8 | [46] |
| 3 | 1AMN | 616, 630, 641 | *Tc* | CT + PAS | 3 | 2.8 | [47] |
| 4 | 1E3Q | 36, 38 | *Tc* | CT + PAS | 12 | 2.85 | [48] |
| 5 | 1E66 | 92, 157 | *Tc* | PAS | 1 | 2.1 | [49] |
| 6 | 1EVE | 1158–1161, 1249, 1254, 1255 | *Tc* | PAS | 6 | 2.5 | [50] |
| 7 | 1GPN | 80, 289 | *Tc* | PAS | 0 | 2.35 | [51] |
| 8 | 1GPK | 265, 529 | *Tc* | PAS | 0 | 2.1 | [51] |
| 9 | 1GQS | – | *Tc* | PAS | 2 | 3.0 | [52] |
| 10 | 1H22 | 67, 218–222 | *Tc* | PAS | 13 | 2.15 | [53] |
| 11 | 1H23 | 70, 161, 266–269 | *Tc* | PAS | 15 | 2.15 | [53] |
| 12 | 1J07 | – | *Mm* | PAS | 12 | 2.35 | [54] |
| 13 | 1N5R | 470 | *Mm* | PAS | 7 | 2.25 | [54] |
| 14 | 1Q84 | – | *Mm* | PAS | 11 | 2.45 | [55] |
| 15 | 1Q83 | 583, 693, 694, 751 | *Mm* | PAS | 12 | 2.65 | [55] |
| 16 | 1QON | – | *Dm* | PAS | 3 | 2.72 | [56] |
| 17 | 2ACK | 627 | *Tc* | CT + PAS | 2 | 2.4 | [57] |
| 18 | 2CEK | 175 | *Tc* | PAS | 11 | 2.2 | [58] |
| 19 | 2CMF | 25, 26 | *Tc* | PAS | 8 | 2.5 | [59] |
| 20 | 2CKM | 32, 37, 54 | *Tc* | PAS | 10 | 2.15 | [59] |
| 21 | 1ODC | 29, 85 | *Tc* | PAS | 11 | 2.20 | [59] |
| 22 | 1UT6 | – | *Tc* | PAS | 9 | 2.40 | [59] |
| 23 | 1W4L | 132, 184, 308, 309 | *Tc* | PAS | 10 | 2.16 | [60] |
| 24 | 1W6R | 107 | *Tc* | PAS | 1 | 2.05 | [60] |
| 25 | 1U65 | 9, 21, 118, 136 | *Tc* | CT + PAS | 5 | 2.61 | [61] |
| 26 | 1ZGB | 89, 109, 128, 141 | *Tc* | PAS | 13 | 2.30 | [62] |

*Tc*: Torpedo Californica; *Mm*: Mus musculus; *Dm*: Drosophila Melanogaster; PAS: Peripheral Anionic Site; CT: Catalytic Triad.

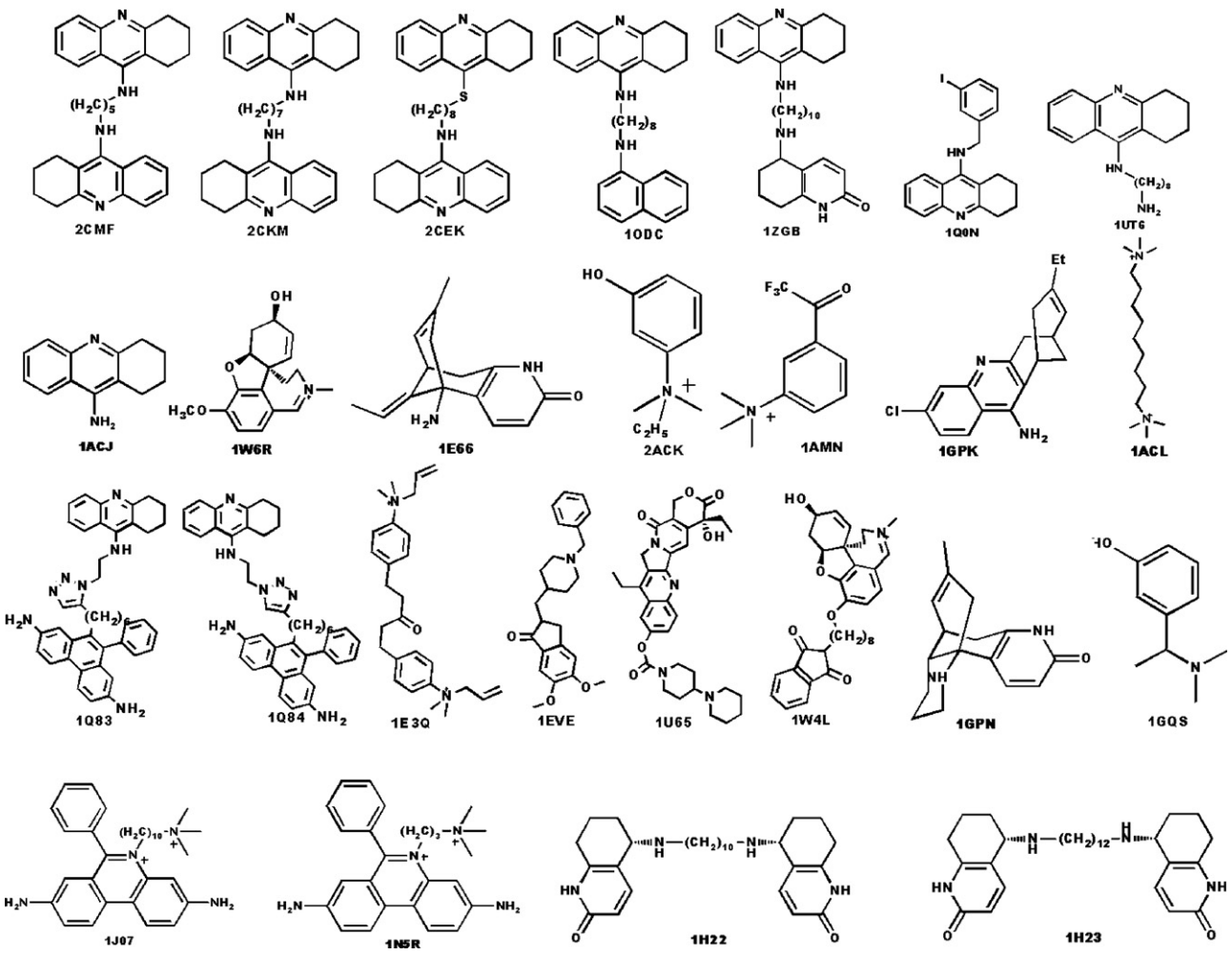



**Fig. 2.** Chemical structure of the ligands 1–26 used in this study.

ligand complexes. The main drawback is that it is unclear whether they are able to predict the binding affinity of ligands structurally different from those used in the training sets.

Knowledge-based scoring functions represent the binding affinity as a sum of protein-ligand atom pair interactions [10]. These potentials are derived from the protein-ligand complexes with known structures, where probability distributions of interatomic distances are converted into distance-dependent interaction free energies of protein-ligand atom pairs. However, 3D structures of protein-ligand complexes do not provide a thermodynamic ensemble at equilibrium, and, therefore, a knowledge-based potential should be considered as statistical preference rather than a potential of mean force. A key ingredient of a knowledge-based potential is the reference state, which determines the weights between the various probability distributions. Before now, several approaches to derive these potentials have been proposed [10–13]. They differ in their definition of the reference state, the protein and ligand atom types, and the list of protein-ligand complexes from which they were extracted.

A number of comparative evaluations of docking programs, conducted over the past several years, indicating that docking programs are generally able to generate ligand poses that are similar to the experimentally determined have been published [14–18]. These studies show that the good performance in reproduction of experimentally determined binding modes did not impart success in virtual screening and the correlation between docking scores and ligand affinity are still challenging [14–18].

Other unresolved issues in automated docking are the consideration of protein and ligand flexibility and the inclusion or omission of explicit water molecules in the ligand binding pocket [19,20]. The complexity of the docking problem increases with the size of ligand and its number of rotatable bonds [21] as rotations around bonds lead to deviations from ideal geometry that results in a small energy penalty when compared to deviations from ideality in bond lengths and bond angles.

The presence of water molecules plays an important role in the accuracy of ligand-protein docking predictions. Water molecules can be involved in protein ligand recognition either by forming mediating hydrogen bonds between the protein and the ligand or by being displaced by the ligand; both of these mechanisms have been shown to be of importance to drug discovery [22–25].

In continuation of our efforts in virtual screening for novel cholinesterase inhibitors [26–29], we performed a comparative evaluation of six widely used docking programs: AutoDock3.0.5 [30], FlexX1.10 [31], MOE2006.08 [32,33], Surflex-Dock [34,35], GOLD3.2 [36,37] and fast rigid body docking program FRED2.2.3 [38,39]. To evaluate the scoring reliability, ten docking scoring functions were utilized to identify the known active inhibitors seeded in a random library of "drug like" compounds. The main goal of this study is the identification of an accurate method for the screening of potential cholinesterase inhibitors.

Although, many comparative molecular docking studies have been reported for AChE [40–44], the reported findings were devoid of the comments regarding conserved water molecules in the active

site of AChE, which can mediate the interaction between amino acid side chains and inhibitors. Similarly probable conformations of Phe330 during docking procedure were also neglected.

In summary, the current study is the most comprehensive comparison of docking and scoring functions reported so far regarding AChE. In this study, open, half open and closed gate conformations of AChE were investigated separately to determine the importance of the side-chain flexibility of Phe330 for ligand trafficking (Fig. 1a). The role of conserved water molecules in docking especially in case of AChE (Fig. 1b), is also analyzed.

## 2. Materials and methods

### 2.1. Selection of testing complexes

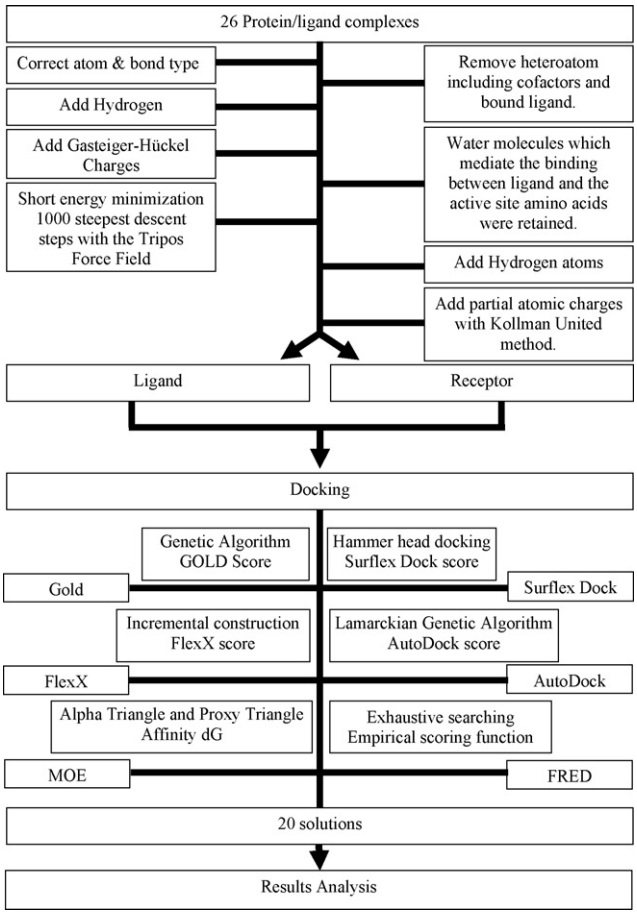
Structures of twenty six protein-ligand complexes were retrieved from the Protein Data Bank (PDB) [45] for the comparative molecular docking studies (Table 1). Structures of the bound ligands are presented in Fig. 2. In order to test the capability of docking programs to deal with the diverse conformations of the same binding pocket and to correctly predict the effects of ligand structure on docking accuracy, a diverse set of ligands were selected. The complexes were chosen according to the following criteria: (i) non-covalent binding between protein and ligand, (ii) crystallographic resolution less than 3 Å (iii) known experimental binding data.

### 2.2. Preparation of ligand and receptor

The SYBYL7.3 package [63] was used to prepare the protein and ligand data required for the docking. The ligands were prepared according to the following procedure. The ligand was extracted from protein structure; this was regarded as the reference structure and was used for the calculation of RMSD values. Correct atom and bond types were defined, hydrogens were added, and partial charges were assigned according to the Gasteiger-Hückel method [64]. Finally a short energy minimization, for 1000 steepest descent steps with the Tripos Force Field [63], was performed to release the internal strain. Ligands with the ionizable group (positively charged quaternary amino group) were protonated in order to comply with the physiological pH.

The receptors were prepared according to the following procedure. Initially the PDB heteroatom records (HETATM) including cofactors and bound ligands were manually removed from the coordinate files. Water molecules which mediate the binding between ligand and the active site amino acids were retained (Table 1). These conserved water molecules (Fig. 1b) play important role as a lubricant to allow large amplitude fluctuations of the loop structures forming the gorge wall of AChE active site [65]. Such fluctuations are required to facilitate the traffic of substrate, products and water molecules to and from the active site. Furthermore these water molecules also mediate interactions between the ligand and the protein and form hydrogen-bonded networks that can stabilize a protein-ligand complex in solution [65]. Hydrogen atoms were added to the receptor atoms and partial atomic charges using Kollman method [63] were assigned to the protein. Aspartate and glutamate were deprotonated, while lysine and arginine were protonated (in accordance with the physiological pH). Histidine 440 was kept neutral (H on δN).

For virtual screening, all three orientations of the Phe330 side chain, i.e., open, half open and close gate, were taken into account (Fig. 1). Three sets of proteins were prepared. The *Tc*AChE-E2020 complex (1EVE) with the open-gate conformation, *Tc*AChE-THA (1ACJ) with closed conformation and *Tc*AChE-TMTFA (1AMN) with the half-open conformation was used separately to determine the importance of side-chain flexibility of Phe330 for ligand trafficking.




**Scheme 1.** Representation of re-docking protocol.

Important water molecules which mediate the binding between ligand and the active site amino acids were retained during virtual screening (Table 1).

### 2.3. Molecular docking

Docking experiments were carried out for all 26 protein-ligand complexes using the docking algorithms employed in AutoDock3.0.5 [30], FlexX1.10 [31], MOE2006.08 [32,33], Surflex-Dock [34,35], GOLD3.2 [36,37] and fast rigid body docking program FRED2.2.3 [38,39]. AutoDock with Lamarckian Genetic Algorithm, FlexX with Incremental Construction, MOE with two docking placement methods Alpha Triangle matcher and Proxy Triangle, Surflex-Dock with Hammerhead incremental construction, GOLD with Genetic Algorithm, and FRED with exhaustive searching were used to dock the ligands in the binding site of AChE. The docking simulations were performed on Xeon 3.0 GHz Linux work station running under SUSE 9.2 and CPU times were recorded. Docking speed is a critical issue in screening large libraries of organic compound and may be important even in a careful study of a small set of ligands. All the docking experiments, reported here, were performed with the default parameters. The time to dock one ligand was approximately 1–2 min for AutoDock, FlexX, FRED and Surflex-Dock and 2–5 min for GOLD and MOE. Scheme 1 represents the re-docking procedure used in this study.

#### 2.3.1. AutoDock 3.0.5

AutoDock 3.0.5 uses a force-field based empirical free energy scoring function [30]. The first three terms are a subset of AMBER (Assisted Model Building with Energy Refinement) poten-

tial energies, including Lennard–Jones 12–6 repulsion/dispersion interactions, the directional 12–10 hydrogen bonding interactions, and the Coulombic electrostatic potential. The remaining energy terms are measures of the unfavorable entropy of ligand binding due to the restriction of conformational degree of freedom and the desolvation effects. The desolvation parameter was assigned to protein using ADDSOL modules of AutoDock. The Lamarckian Genetic Algorithm (LGA) [66] was used as a search engine. The active site was defined using AutoGrid. The grid size was set to 70 Å × 70 Å × 70 Å points with grid spacing of 0.375 Å centered on the ligand center of mass. Step sizes of 1.0 Å for translation and 50 degree for rotation were chosen, a maximum number of energy evaluations was set to 250,000. A maximum number of 27,000 LGA operations were generated on a single population of 50 individuals. Operator weights for crossover, mutation, and elitism were set to 0.80, 0.02, and 1, respectively. For each docking simulation twenty different poses were generated.

*2.3.2. FlexX 1.12*

FlexX uses a pure empirical scoring function similar to that developed by Böhm and coworkers [31]. Binding free energy of a protein/ligand complex is estimated as the sum of free energy contributions from hydrogen bonding, ion-pair interactions, aromatic or lipophilic interactions. A scaling function is used to penalize deviations from ideal geometry. Standard docking conditions of the FlexX program were adopted as implemented in the SYBYL 6.9 package. A receptor description file was defined from the PDB coordinates through the FlexX graphic interface. The active site includes protein residues around 6.5 Å radius sphere centered on the center of mass of the ligand. Based on energy, twenty top ranked poses were saved.

*2.3.3. MOE2006.08*

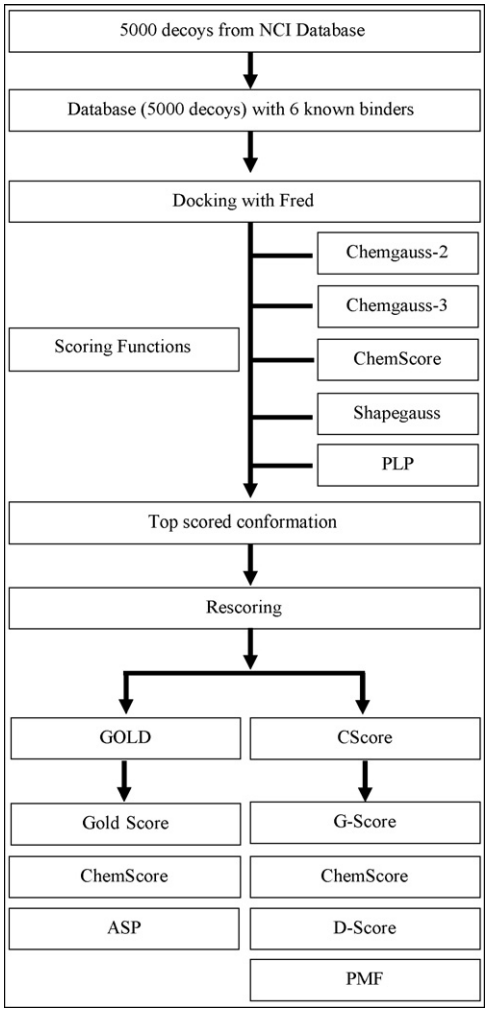
For the docking, the protein was minimized using the MMFF94 force field [67,68], keeping all the heavy atoms fixed until a RMSD gradient of 0.05 kcal $mol^{-1}$ $Å^{-1}$ was reached. The active site was generated for each enzyme using MOE alpha site finder. The docking simulations were performed using two docking placement methods (i) Alpha Triangle placement method and (ii) Proxy Triangle placement method in order to identify the most appropriate docking method in combination with the Affinity dG scoring method [32,69]. Twenty top ranked docked poses were saved for further study.

*2.3.4. Surflex-Dock*

Surflex-Dock [34,35] uses an empirical scoring function and a patented search engine to dock ligand into the protein's binding site. Surflex-Dock employs an idealized active site ligand called a protomol as a target to generate putative poses of molecules or molecular fragments. These putative poses are scored using Hammerhead scoring function which also serves as an objective function for local optimization of poses. Flexible docking proceeds either by incremental construction from high scoring fragment as in Hammerhead or by crossover procedure that combines pieces of poses from intact molecule. The docking procedure was started with the protomol generation. The protomol was generated using ligand based approach. Proto_thresh was set to 0.2 and proto_bloat was left at 0 as a default parameter. For each protein-ligand pair, twenty top ranked docked solutions were saved.

*2.3.5. GOLD 3.2*

GOLD uses a Genetic Algorithm [36] to explore the full range of ligand conformational flexibility with the partial flexibility of protein [37]. The GOLD scoring function includes the terms for hydrogen-bonding, vdW, and intramolecular energies. The vdW interactions for the protein/ligand complex are described by 8–4 potential. A 12–6 potential is used for the ligand steric energies that also include the torsional energies. The active site with a 10 Å radius sphere was defined. For each independent GA run, a maximum number of 100,000 GA operations were performed on a single population of 100 individuals. Operator weights for crossover, mutation, and migration were set to 95, 95, and 10, respectively. To allow for poor nonbonded contacts at the start of each GA run, the maximum distance between hydrogen donors and fitting points was set to 4.0 Å, and non-bonded vdW energies were cut-off at 2.5 Å. Twenty top ranked poses were saved.




**Scheme 2.** Representation of virtual screening and re-scoring protocol used in this study.

*2.3.6. FRED 2.2.3*

In FRED [38,39] the first stage in docking is a shape fitting process, which takes a set of ligand conformers as input and tests them against a "bump map" (a Boolean grid with true values where ligand atoms can potentially be placed). Orientations that clash with the protein or are distant from the active site are rejected. The crude docking solutions are further tested against a pharmacophore feature if specified, and any poses that do not satisfy the pharmacophore are rejected. Poses surviving the shape fitting routine can then be passed to three scoring function filters in the screening process. Various options are available for optimization with respect to the built-in scoring functions: optimization of hydroxyl group rotamers, rigid body optimization, torsion optimization, and reduction of the number of poses that are passed on to the next scoring function. FRED offers a number of scor-

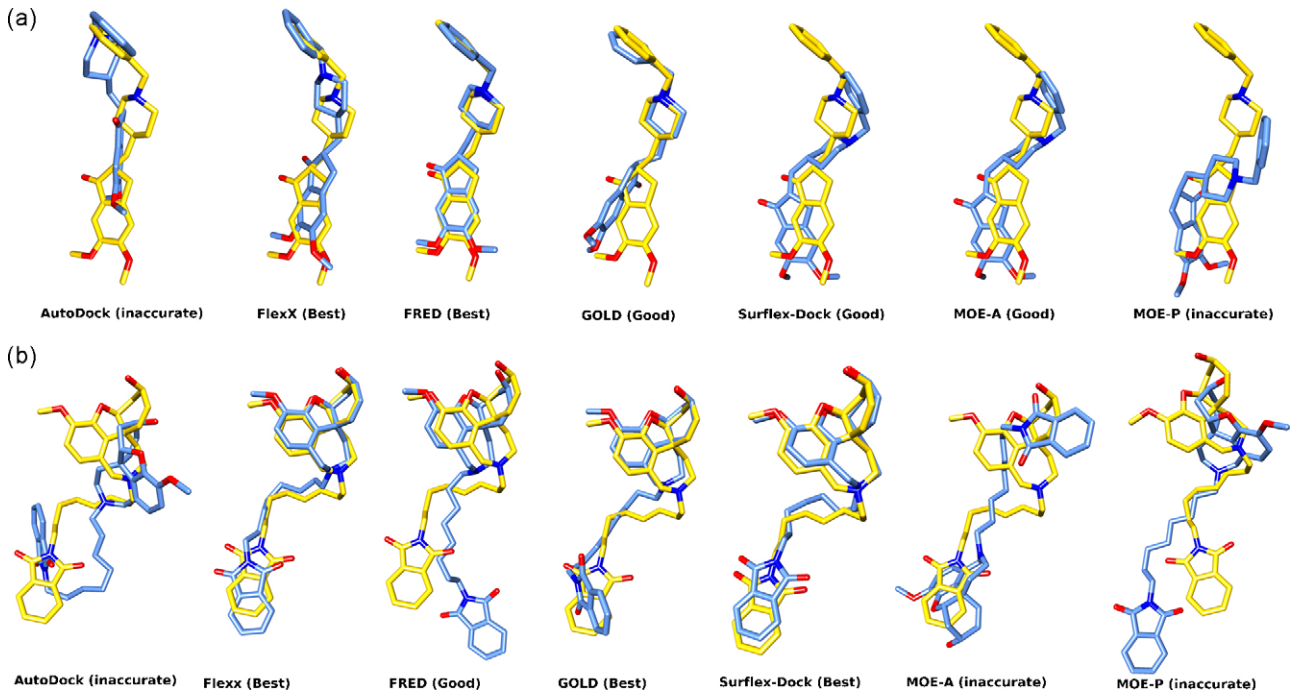



**Fig. 3.** Docking results for (a) 1EVE and (b) 1W4L generated by the six docking/scoring approaches. The pose selection criteria are given in parenthesis. The experimental conformation (yellow) and docked conformation (blue) are presented in stick model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

ing functions including a Gaussian type fitting (Chemgauss3 and its earlier version Chemgauss2 which is a proprietary function of OpenEye), ChemScore, PLP, ScreenScore and Shapegauss. Qualitatively, the Gaussian scoring function has favorable values when the ligand and protein has a high surface contact and little volume overlap.

For docking, multi-conformer libraries of the input ligand database were generated by OMEGA (OpenEye Scientific Software) [70] with the default settings. The receptor files were created by using the supplied protein structure in combination with a shape based site detection algorithm and the position of a known bound ligand. The receptor file was docked with multiple conformer libraries for each protein-ligand complex using five scoring functions, i.e., Chemgauss3, Chemgauss2, ChemScore, PLP and Shapegauss. Twenty top ranked docking poses were saved and analyzed.

### 2.4. RMSD and rank

RMSD values were calculated to quantify the difference between the crystal ligand coordinates and the predicted coordinates by each docking program. In this study, the prediction was classified according to the following criteria: (i) Good (RMSD $\leq$ 2 Å), (ii) Fair (RMSD >2 and <3 Å) (iii). Inaccurate when docking solution is far from active site or in an inverted or incorrect position.

The purpose of re-docking the inhibitors into their crystal binding site was to identify the most accurate docking method which can be subsequently used for virtual screening. In our case FRED was the most accurate in regenerating the best docking poses (see Section 3). Therefore, for the virtual screening, FRED was employed as the docking engine for the 5000 decoys gathered from NCI database [71], along with six known inhibitors of AChE in combination with five scoring functions (Chemgauss2 (earlier version), Chemgauss3 (latest version), ChemScore, PLP and Shapegauss). To evaluate the robustness of scoring functions in finding known inhibitors embedded in random decoys, the top-ranked FRED docked poses were re-scored using GOLD and CScore [72,73] module implemented in SYBYL7.3. GOLD was used with GOLD score, ChemScore and ASP while CScore was used with four scoring functions, i.e., G_Score, D_Score, ChemScore and PMF. The goal of re-scoring was to scrutinize which scoring functions are able to correctly rank the known inhibitors among the top positions from the random library of "drug like" compounds.

### 2.5. Virtual screening

The NCI database contains about 0.2 million compounds [71]. A set of 5000 compounds was randomly selected from the NCI database according to the Lipinski rule of five [74]. Hydrogens were added and partial charges were assigned according to the Gasteiger-Hückel method [64] and a short energy minimization for 1000 steepest descent steps with the Tripos Force Field was performed and multi conformer libraries of the input ligand database were generated by OMEGA [70]. Six AChE inhibitors (Tacrine, TMTFA, Edrophonium, Rivastigmine, NAP and Donepezil) were

**Table 2a**
The % of good, fair and inaccurate pose for twenty six AChE-ligand complexes predicted by different docking approaches in the presence of water molecules.

| Pose category | FRED (%) | Surflex-Dock (%) | GOLD (%) | FlexX (%) | MOE (%) | | AutoDock (%) |
|---|---|---|---|---|---|---|---|
| | | | | | Alpha triangle | Proxy triangle | |
| Good | 42 | 35 | 30 | 23 | 11 | 19 | 19 |
| Fair | 31 | 27 | 35 | 15 | 35 | 46 | 38 |
| Inaccurate | 27 | 38 | 35 | 62 | 54 | 35 | 43 |

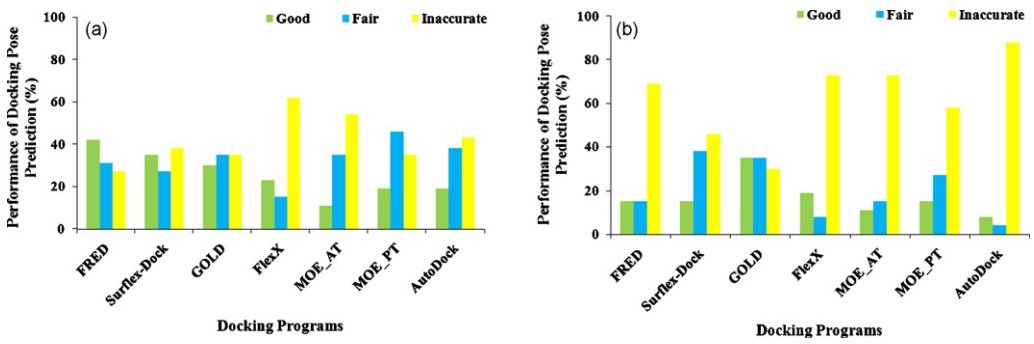



**Fig. 4.** Performances of docking pose prediction of all six docking/scoring approaches. (a) In the presence of interacting water molecules. (b) In the absence of interacting water molecules.

**Table 2b**
Docking accuracy as a function of rotatable bonds of ligand in the presence of water.

| No. of rotatable bonds | No. of ligands | Pose category | FRED (%) | Surflex-Dock (%) | GOLD (%) | MOE (%) | | FlexX (%) | AutoDock (%) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Alpha triangle | Proxy triangle | | |
| 0–3 | 9 | Good | 66 | 22 | 77 | 22 | 33 | 44 | 44 |
| 5–8 | 4 | | 75 | 25 | 0 | 25 | 0 | 25 | 25 |
| 9–15 | 13 | | 15 | 46 | 8 | 8 | 15 | 8 | 8 |
| 0–3 | 9 | Fair | 22 | 11 | 11 | 22 | 11 | 0 | 0 |
| 5–8 | 4 | | 0 | 50 | 75 | 0 | 75 | 25 | 25 |
| 9–15 | 13 | | 46 | 30 | 38 | 38 | 61 | 23 | 38 |
| 0–3 | 9 | Inaccurate | 11 | 66 | 11 | 55 | 55 | 55 | 55 |
| 5–8 | 4 | | 25 | 25 | 25 | 50 | 25 | 50 | 50 |
| 9–15 | 13 | | 38 | 23 | 53 | 53 | 23 | 69 | 53 |

**Table 3a**
The % of good, fair and inaccurate pose for twenty six AChE-ligand complexes predicted by different docking approaches in the absence of water molecules.

| Pose category | FRED (%) | Surflex-Dock (%) | GOLD (%) | FlexX (%) | MOE (%) | | AutoDock (%) |
|---|---|---|---|---|---|---|---|
| | | | | | Alpha triangle | Proxy triangle | |
| Good | 15 | 15 | 35 | 19 | 11 | 15 | 8 |
| Fair | 15 | 38 | 35 | 8 | 15 | 27 | 4 |
| Inaccurate | 69 | 46 | 30 | 73 | 73 | 58 | 88 |

added to this database library. The protocol used for virtual screening is schematically represented in Scheme 2.

### 2.6. Enrichment factor

The results obtained by virtual screening, were further analyzed in terms of enrichment factors. Enrichment factor (EF) is a common metric used when comparing virtual screening results. EF is defined as:

$$\mathrm{EF} = \frac{(\mathrm{HITS}_{\mathrm{sampled}}/\mathrm{HITS}_{\mathrm{total}})}{(N_{\mathrm{sampled}}/N_{\mathrm{total}})}$$

Here, $N_{\mathrm{total}}$ is the number of ligands in the docked database, $N_{\mathrm{sampled}}$ is the number of ligands in the docked database to be examined, $\mathrm{HITS}_{\mathrm{total}}$ is the total number of the known active ligands, and $\mathrm{HIT}_{\mathrm{sampled}}$ is the number of known active ligands found in the top $N_{\mathrm{sampled}}$ ligands of docked database.

## 3. Results and discussion

### 3.1. Performance of docking accuracy and correct pose prediction

Correct prediction of protein-ligand interaction geometries is essential for the success of virtual screening approaches in

**Table 3b**
Docking accuracy as a function of rotatable bonds of ligand in the absence of water.

| No. of rotatable bonds | No. of ligands | Pose category | FRED (%) | Surflex-Dock (%) | GOLD (%) | MOE (%) | | FlexX (%) | AutoDock (%) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Alpha triangle | Proxy triangle | | |
| 0–3 | 9 | Good | 11 | 11 | 55 | 11 | 22 | 33 | 22 |
| 5–8 | 4 | | 25 | 0 | 50 | 25 | 25 | 0 | 0 |
| 9–15 | 13 | | 8 | 23 | 15 | 8 | 8 | 15 | 0 |
| 0–3 | 9 | Fair | 22 | 22 | 0 | 22 | 33 | 11 | 0 |
| 5–8 | 4 | | 0 | 50 | 25 | 0 | 25 | 0 | 25 |
| 9–15 | 13 | | 15 | 46 | 61 | 15 | 23 | 8 | 0 |
| 0–3 | 9 | Inaccurate | 66 | 66 | 44 | 66 | 44 | 55 | 77 |
| 5–8 | 4 | | 75 | 50 | 25 | 75 | 50 | 100 | 75 |
| 9–15 | 13 | | 76 | 30 | 23 | 76 | 69 | 76 | 100 |

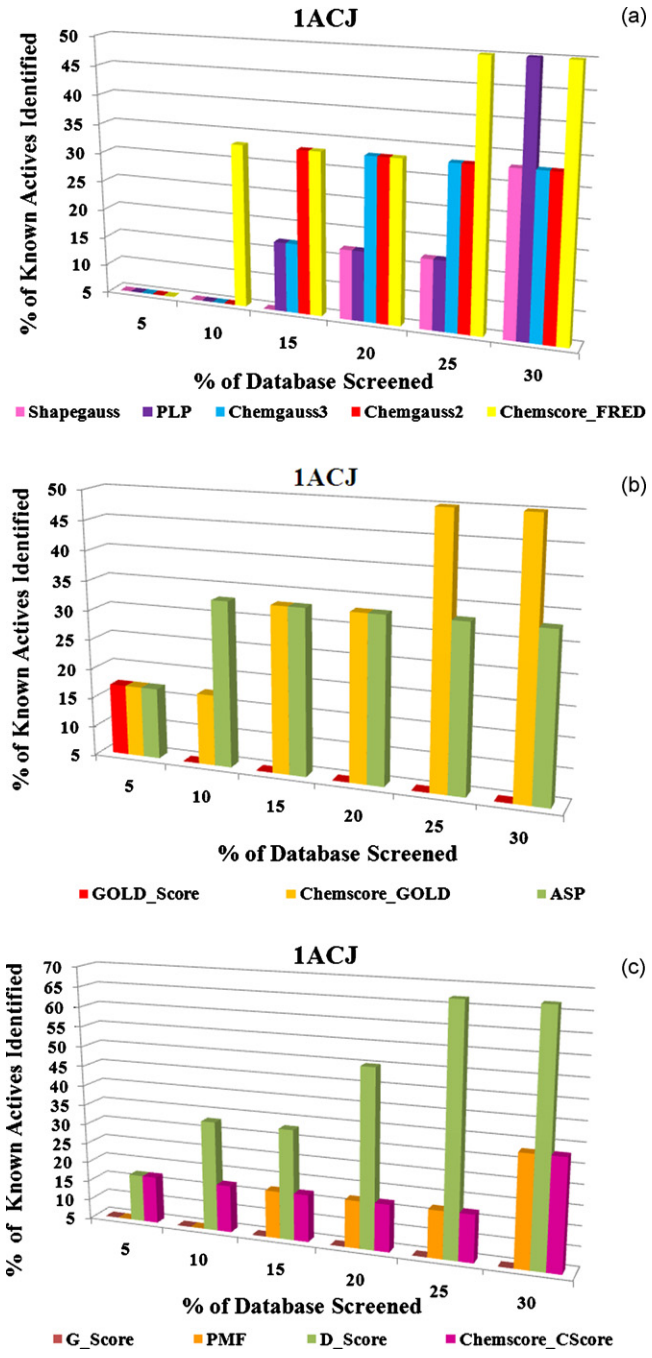



**Fig. 5.** Percent of known actives found (*y*-axis) vs. percent of ranked database screened (*x*-axis) for (a) Chemgauss2, Chemgauss3, ChemScore, Shapegauss, PLP, (b) GOLD score, ChemScore_GOLD, ASP, (c) G_Score, D_Score, ChemScore_CScore and PMF for the closed form of AChE (1ACJ).

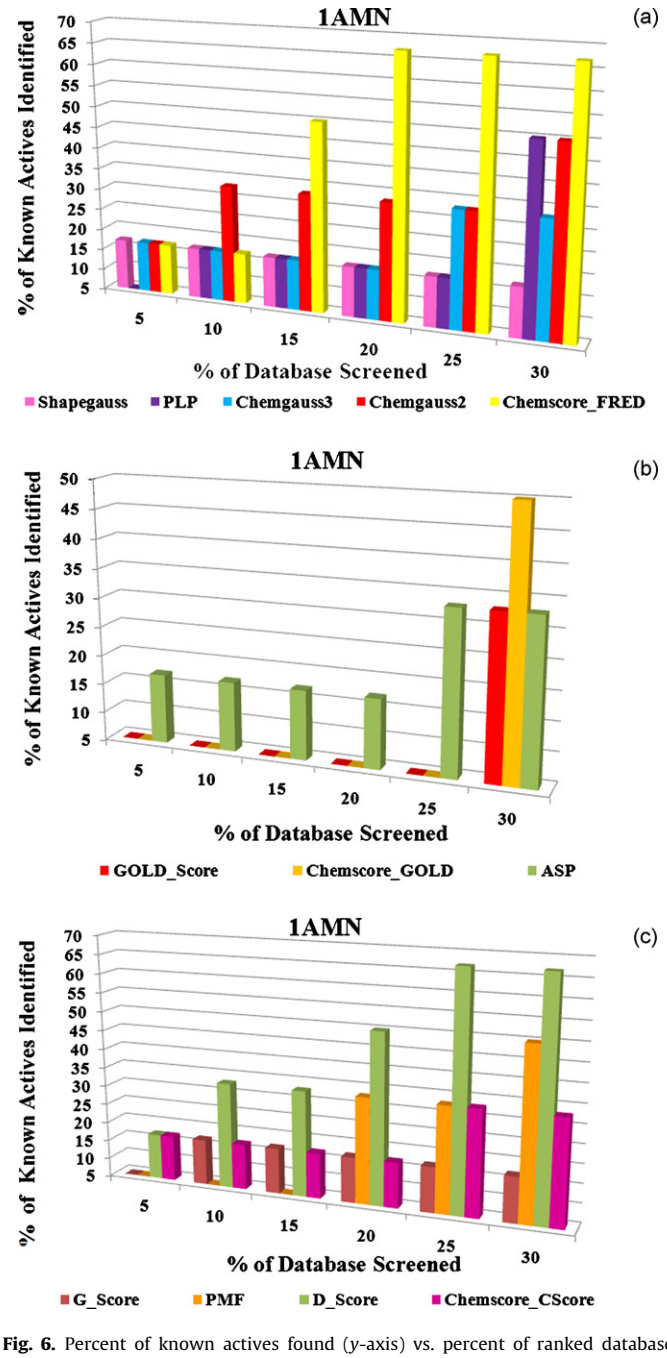



**Fig. 6.** Percent of known actives found (*y*-axis) vs. percent of ranked database screened (*x*-axis) for (a) Chemgauss2, Chemgauss3, ChemScore, Shapegauss, PLP, (b) GOLD score, ChemScore_GOLD, ASP, (c) G_Score, D_Score, ChemScore_CScore and PMF for the closed form of AChE (1AMN).

structure-based drug design. It requires a docking tool that is able to generate suitable conformations of the ligand within the protein binding site and reliable energetic evaluation indicating the quality of interaction. The RMSD values between crystal and predicted structures are used to confirm whether a correct docking position was obtained by the docking simulation.

Here six docking protocols were used to predict bound conformations for twenty-six AChE/ligand crystal complexes (Table 1). Each docking protocol returned twenty docked poses for each ligand; RMSD values were computed for all returned poses. Fig. 3 shows the top ranked solutions generated with all these methods.

Table 2a summarizes the categories in which the dock poses fell. The results showed relatively poor performance ranging from 11 to 42% (generate good pose). These low percentages indicate that it was difficult for the docking programs to place the correct binding pose in the top ranking position. Among all the docking methods used in this study, FRED was found to be the best in generating poses closest to co-crystallized ligands in the 20 Å deep binding pocket of AChE, followed by Surflex-Dock and GOLD. Fig. 4 shows that FRED, Surflex-Dock and GOLD performed well with 42%, 35% and 30% of good poses (RMSD <2.0 Å) in top rank respectively. In comparison with FRED, Surflex-Dock and GOLD, other programs like FlexX, MOE and AutoDock showed relatively poor performance in generating the poses close to the crystal structure (RMSD <2.0 Å). FlexX did

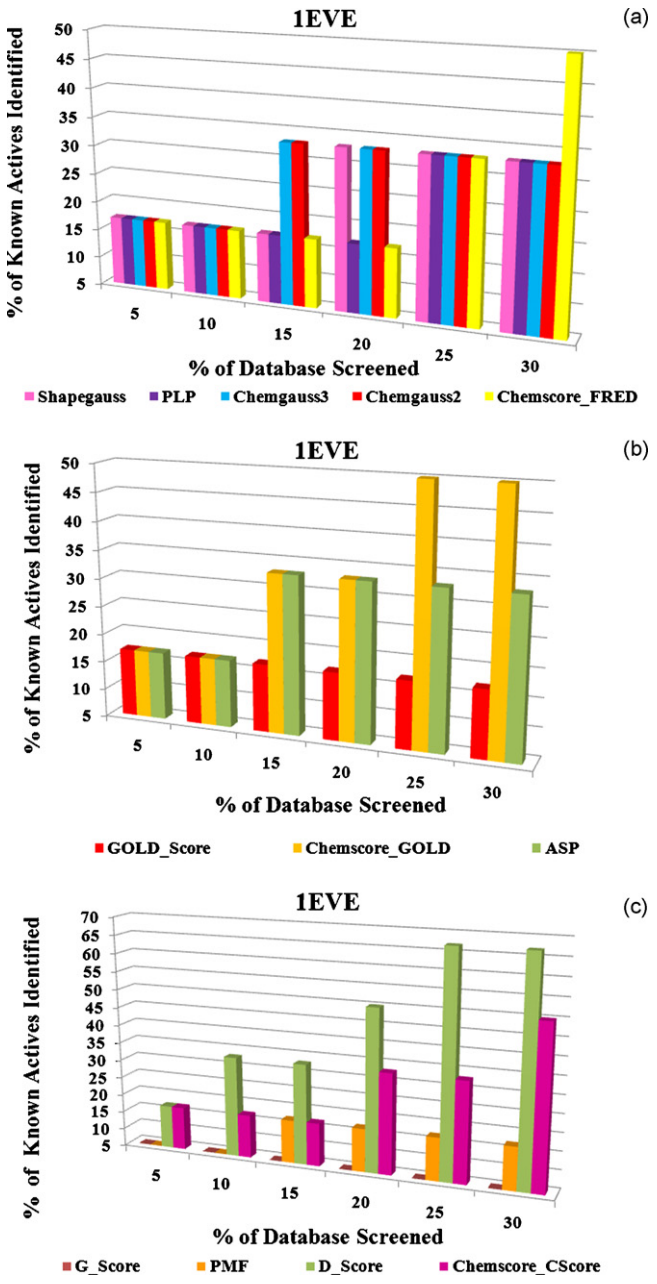



**Fig. 7.** Percent of known actives found (*y*-axis) vs. percent of ranked database screened (*x*-axis) for (a) Chemgauss2, Chemgauss3, ChemScore, Shapegauss, PLP, (b) GOLD score, ChemScore_GOLD, ASP, (c) G_Score, D_Score, ChemScore_CScore and PMF for the closed form of AChE (1EVE).

not significantly predict an accurate binding mode. Only 23% of the small molecules were re-docked accurately (RMSD ≤2.0 Å) in the binding pocket similar to their cognate ligand's orientation. The most probable reasons are flexibility of Trp279 and Phe330 in the active site region and the structural complexity and flexibility of ligands. Ligand with more than 10 rotatable bonds tends to generate wrong solutions due to the additional degrees of freedom during the base placement which seems to cause the wrong solution [74].

MOE was used with two docking placement methods: Alpha Triangle (MOE_AT) which is the default method and Proxy Triangle (MOE_PT) method. The results obtained by MOE and AutoDock were similar to the FlexX performance. Both programs were found to be unable to place flexible ligands into the deep binding pocket of AChE.

For the six docking approaches, the fraction of good poses is less than half to about one-fifth of the best poses among the top 30 solutions. More than 50% of the good poses (RMSD ≤2.0 Å) were not ranked in the top 20. This kind of failure reveals the inaccuracy of scoring functions as implemented in docking programs [75]. In some cases it was observed that ligands were placed in the active site of protein but did not adopt the conformation that allows compound to be oriented correctly within the active site resulting in high RMSD values, hence the resultant pose was inaccurately categorized.

In the generation of fair poses (RMSD >2.0 or ≤3.0 Å), MOE_PT was found to be better as compare to the other docking methods (Fig. 4a). Among the best results, MOE_PT had 46% that were classified as fair, while AutoDock had 38%. The Proxy Triangle method was developed for tackling medium to large sized ligands when there are a huge number of conformers to be chosen for docking [32]. Comparatively, the performance of Proxy Triangle was better than Alpha Triangle method. As shown in Fig. 4a, Proxy Triangle was found to be appropriate in generating fair poses as compare to the Alpha Triangle method. The performance of MOE_AT method, GOLD and AutoDock were comparable in finding the fair poses in the top 20 ranked docking solutions. The MOE_AT method, GOLD and AutoDock identified 35%, 35% and 38% of fair solutions (RMSD >2.0 or ≤3.0 Å) as a top 20 rank, respectively. While the performance of FRED, Surflex-Dock and FlexX were significantly reduced.

Structural complexity of the ligands has significant influence on the docking performance. As reflected in Table 2b, well-docked solutions in top 20 ranked poses are found for ligands with 0–3 rotatable bonds. It was observed that docking procedure usually failed to correctly place the ligand in active site as the flexibility of ligand increases. This observation is notoriously associated with the fact that as the number of rotatable bonds increases, the number of possible conformations increases, which in turn affects the accuracy of docking pose predictions.

The ability to generate good poses of molecules with 9–15 rotatable bonds and structural complexity, Surflex-Dock was found to be best. FRED regenerated 66% and 75% of good poses of the ligands with 0–3 and 5–8 rotatable bonds, respectively, while the performance significantly decreased when the complexity and flexibility of molecules increased (9–15 rotatable bonds). Likewise GOLD regenerated 77% of good poses of small ligands with 0–3 rotatable bonds whereas it failed to generate many good poses of the ligands with 9–15 rotatable bonds (Table 2b). This is not surprising since the Genetic algorithm is more likely to fail if the ligand is large or highly flexible [36].

In the generation of fair poses of molecules with 9–15 rotatable bonds, MOE_PT was found to be the most suitable program. GOLD and MOE_PT reproduced 75% fair poses of molecules with 5–8 rotatable bonds.

### 3.2. Role of water molecules in the docking accuracy

In order to find out the role water molecule plays in the docking pose prediction, water molecules were stripped from protein active site and the ligands were docked. Performance of all docking methods to regenerate good poses was significantly decreased due to absence of interactions mediated by these conserved water molecules (Fig. 4b). Even FRED which was found to be relatively more accurate in this study was not able to generate any pose closest to the co-crystallized ligand in the absence of conserved water molecules. Among twenty six ligands, only 15% of the ligands were superimposed on the reference structure with an RMSD less than 2.0 Å (Good pose). Results of the docking calculations in the absence of water molecules are summarized in Table 3a. In comparison, Surflex-Dock and AutoDock identified 15% and 8% good poses, respectively. For both of these docking methods, ability to

**Table 4**
Standard enrichment factor at 5% and 10% of the ranked database for different scoring functions.

| S# | Scoring Functions | 1ACJ | | 1AMN | | 1EVE | |
|---|---|---|---|---|---|---|---|
| | | *5%* | *10%* | *5%* | *10%* | *5%* | *10%* |
| 1 | Chemgauss2 | 0.00 | 0.00 | 3.33 | 3.33 | 3.33 | 3.33 |
| 2 | Chemgauss3 | 0.00 | 0.00 | 3.33 | 1.67 | 3.33 | 3.33 |
| 3 | ChemScore_FRED | 0.00 | 3.33 | 3.33 | 1.67 | 3.33 | 3.33 |
| 4 | PLP | 0.00 | 0.00 | 0.00 | 1.67 | 3.33 | 3.33 |
| 5 | Shapegauss | 0.00 | 0.00 | 3.33 | 1.67 | 3.33 | 3.33 |
| 6 | G_Score | 0.00 | 0.00 | 0.00 | 1.67 | 0.00 | 0.00 |
| 7 | PMF | 0.00 | 0.00 | 0.00 | 0.00 | 0.0 0 | 0.00 |
| 8 | **D_Score** | **3.33** | **3.33** | **3.33** | **3.33** | **3.33** | **3.33** |
| 9 | ChemScore_CScore | 3.33 | 1.67 | 3.33 | 1.67 | 3.33 | 1.67 |
| 10 | GOLD_Score | 3.33 | 1.67 | 0.00 | 0.00 | 3.33 | 1.67 |
| 11 | ChemScore_GOLD | 3.33 | 1.67 | 0.00 | 0.00 | 3.33 | 1.67 |
| 12 | ASP | 3.33 | 3.33 | 3.33 | 1.67 | 3.33 | 1.67 |

Gray bar: Worst results.
Green bar: Good results at 5%.
Orange bar: Good results at 10%.
Yellow bar: Results not improved at 10%.
D_Score results are shown in black bold.

produce good poses was greatly reduced in the absence of interacting water molecules. This is not surprising as the docking programs usually fail to produce correct solutions in the absence of interacting water molecules [22–25]. However, GOLD was found to be similar in identification of good and fair poses either with or without water molecules. Performance of MOE to generate good poses, was found to be relatively similar as identified previously in the presence of water molecules. However, the ability to generate fair solutions was significantly reduced. Similarly the ability of FlexX to generate good and fair poses was reduced in the absence of interacting water molecule. This is due to the reason that in the absence of water molecules the algorithm is not able to complete the correct construction path [76]. This suggests that the water molecules have an important role in the binding of ligands in the active site of AChE that must be taken into account when used for virtual screening. Table 3b shows that in the absence of conserved water molecules, performance of these programs to generate good poses were greatly affected when the complexity and flexibility of the ligand increased.

### 3.3. Evaluation of the performance of scoring functions

Scoring functions are required for two purposes: During the docking process, they serve as fitness functions in the optimization of ligand orientation and conformation, and for comparison with other molecules they are used as estimates of binding affinity for the completely docked molecule. Although, in principle different functions can be used for these two purposes, in most applications the same function has been used. Thus there are various criteria for the quality of a scoring function: its ability to identify the correct binding mode of a ligand out of alternative docking solutions, its ability to rank related ligands with respect to their binding affinity, and its ability to select inhibitors out of a large database of inactive compounds. Here we will focus on the third criterion which is the central issue in virtual screening.

As in the previous section we have discussed the performance of docking accuracy and in this study FRED was found to be the best approach. Therefore, in the next phase FRED was employed as the docking engine for the virtual screening of 5000 compounds gathered from NCI database, in combination with five scoring functions that are FRED's built-in Gaussian type scoring functions Chemgauss3, Chemgauss2 (earlier version), ChemScore, PLP and Shapegauss. Docking solutions obtained were re-scored using GOLD in combination with GOLD score, ChemScore and ASP and CScore module of SYBYL 7.3 in combination with G_Score, D_Score, PMF and ChemScore. Here the performance of ChemScore implemented in FRED, GOLD and CScore are also compared. It is crucial for the scoring functions to rank the known actives above the decoy ligands.

We begin with a summary analysis that compares the performance of different scoring functions averaged over three protein setup (1ACJ, 1AMN and 1EVE). The scoring functions were evaluated on the basis of enrichment in the top 5, 10, 15, 20, 25 and 30% of database. Figs. 5–7 displays the percent of known actives recovered as a function of the percent ranked database sampled for Chemgauss3, Chemgauss2, ChemScore_FRED, PLP, Shapegauss, GOLD score, ChemScore_GOLD, ASP, G_Score, D_Score, PMF and ChemScore_CScore for all three proteins setup.

#### 3.3.1. 1ACJ

When the closed form of AChE (1ACJ) is considered (Fig. 5a–c), all scoring functions implemented in FRED were unable to identify the active compounds embedded in the random database. D_Score and ChemScore_CScore identified 17% of the known binders while the performance of G_Score and PMF were similar to the FRED scoring functions. The performance of GOLD score, ChemScore and ASP as implemented in the GOLD program, is somewhat similar to D_Score and ChemScore_CScore. The known actives found in the top 5% of the ranked database are 17% for all three methods.

When 10% of the database is screened, the performance of ChemScore as implemented in FRED improved significantly while the performance of Chemgauss2, Chemgauss3, PLP and Shapegauss did not improve. D_Score performed well as compared with the G_Score and PMF, while the performance of the ChemScore_CScore was found to be similar as earlier. The performance of GOLD score and ChemScore_GOLD to identify known actives were not improved, while the performance of ASP increased significantly. On average, ChemScore_FRED, D_Score and ASP have similar performance by identifying 33% of known actives in top 10% of ranked database.

Chemgauss2 identified 33% of known actives when 15% of database was screened and performance did not improve up to 30% of database screened. While Chemgauss3 identified 17% and 33% of actives when 15% and 20% of the database screened, respectively. The performance of Chemgauss3 was found to be consistent when upto 30% of database were screened. PLP was able to identify 17% of known actives when 15% of database screened. ChemScore and PLP identified 50% of known actives when 25% and 30% of database screened respectively. Among all the methods scrutinized here, each scoring function could identify active compounds from ligand

database while G_Score performed worst in identification of fraction of active compounds with up to 30% of database screened. On average, D_Score found 67% of known actives in top 30% of ranked database (Fig. 5).

#### 3.3.2. 1AMN

When the half open form of AChE (1AMN) is considered (Fig. 6a–c), performance of Chemgauss2, Chemgauss3, ChemScore_FRED, Shapegauss were improved when 5% of database screened, these soring functions retrieved 17% of known actives. While D_Score, G_Score, PMF and ChemScore_CScore performed as they were in the closed form. The performance of GOLD_Score and ChemScore_GOLD were decreased significantly while ASP identified 17% of known actives. When 10% of the database is screened, further improvement in the performance of Chemgauss2, PLP, GOLD_Score and D_Score was observed. Chemgauss2 and D_Score identified 33% of active inhibitors at this stage. ChemScore_FRED identified 50% of the known actives when 15% of the database is screened. On average, Chemgauss2, PLP, PMF and ChemScore_GOLD identified 50% of known actives when 30% of the database is screened while D_Score identified 50% in top 20% of ranked database. ChemScore_FRED and D_SCORE identified 67% of known compounds when 20% and 25% of the database is screened, respectively. In this case ChemScore_FRED and D_Score achieved better results than other applied scoring functions. Fig. 6 shows the overall performance of all the scoring functions applied to half open form of the protein.

#### 3.3.3. 1EVE

When the open form of the protein (1EVE) is considered (Fig. 7a–c), Chemgauss2, Chemgauss3, ChemScore_FRED, Shapegauss, D_Score, ChemScore_CScore and ASP performed similarly when 5% of the database is screened. All of these scoring functions identified ~17% of known actives in top 5% of ranked database. While G_Score and PMF were not able to identify any known active at this level. It is worth noting that when the open form of AChE is used, a higher number of scoring functions was able to identify the known actives in top 5% of the screened database. Performance of Chemgauss2 and PLP were found comparable to Chemgauss3 and Shapegauss respectively using the top 30% of the ranked database. While at this level ChemScore_FRED found 50% of the known inhibitors. The performance of G_Score was found to be disappointing among all three protein conformations. In case of 1EVE it was failed to identify any known actives in top 30% of the ranked database. While the performance of D_Score, ChemScore_CScore, ChemScore_GOLD and ASP was significantly increased as 5, 10, 15, 20, 25 and 30% of the database screened. GOLD_Score identified only 17% of the active inhibitors in top 5% of a ranked database and the performance was found to be consistent until 30% of the database is screened. D_Score identified 50% and 67% of known actives respectively when 20% and 25% of the database is screened. ChemScore_FRED and ChemScore_CScore retrieved 50% of known actives when 30% of the database is screened. While ChemScore_GOLD identified 50% of known actives when 25% of the database is screened. Fig. 7 displays the enrichment achieved by all the above mentioned scoring functions with the open form of protein (1EVE).

As depicted in Figs. 5–7, the enrichment bars show that D_Score yields a good performance for all three proteins conformations. While the ability to identify known active compounds in a top ranked database, ChemScore_FRED and ChemScore_GOLD were relatively consistent in their performance. In our case G_Score as implemented in CScore module of SYBYL and GOLD score of GOLD found to be relatively poor. It was identified that on all three proteins conformations D_Score was the best followed by ChemScore_CScore and ASP.

### 3.4. Enrichment factor

Table 4 shows the calculated EF at 5% and 10% of the total decoy set for each scoring function that was used with each of the three protein conformation (1ACJ, 1AMN and 1EVE) used in virtual screening. As illustrated in Table 4, D_Score found to be superior to all other scoring methods for each protein conformation when 5% and 10% of the database is screened. Chemgauss2, Chemgauss3, ChemScore_FRED, PLP, Shapegauss, G_Score and PMF all performed poorly for 1ACJ when 5% of the database is screened. While D_Score, GOLD_Score, ChemScore_GOLD and ASP performed well. G_Score and PMF performed poorly against each of the three protein conformations when calculated using 5% and 10% of the total decoy set. When applied on 1AMN, Chemgauss2, Chemgauss3, ChemScore_FRED, Shapegauss, D_Score, ChemScore_CScore and ASP performed well at 5% while PLP, G_Score and PMF continued to perform poorly. In the half open form of protein, the performance of GOLD_Score was significantly decreased. When the open form of protein 1EVE is considered, all the applied scoring functions showed good results at 5% except G_Score and PMF. Overall D_Score returned the best enrichment results on all three protein set up while G_Score and PMF returned the worst results. GOLD_Score and ChemScore_FRED performed poorly when applied on 1AMN.

## 4. Concluding remarks

Docking and scoring have evolved significantly over the past years. It has become a valuable tool in drug discovery process. Our goal of this study was to explore the feasibility of six different docking approaches: (FRED, GOLD, MOE, AutoDock, Surflex-Dock and FlexX) for our target of interest (AChE) and to find out the best amongst several available combinations. We compared predictive power of each docking and scoring function with reference to reproducing the binding poses and their ranking capability at default parameters, in the presence and absence of active site water molecules. FRED was found to be relatively more useful in docking pose prediction as studied here. A performance of docking pose prediction of all six docking/scoring approaches was improved in the presence of interacting water molecules.

The ability of the scoring functions to select known inhibitors out of a random library of “drug like” compounds is compared and analyzed. FRED was employed as a docking engine for docking of 5000 drug like decoys taken from NCI database. FRED was used with five scoring functions (Chemgauss2, Chemgauss3, ChemScore_FRED, Shapegauss and PLP). Additionally, to see how well other scoring functions perform with FRED’s generated poses, the top ranked solutions obtained by FRED were re-scored with GOLD score, ChemScore and ASP implemented in GOLD program and G_Score, D_Score, PMF and ChemScore implemented in CScore. As seen from the results, D_Score has proven to be a robust and valuable scoring function in the selection of known active compounds embedded in random decoys.

In summary, we performed an exhaustive survey of the contemporary docking and scoring methods with all the possible combinations, i.e., with various scoring functions followed by rescoring. Additionally, role of water molecule in the docking experiments were also discussed in the detailed. In conclusion, we found an efficient docking and scoring protocol for AChE. This set up will definitely aid in our ongoing projects and to the community having same research interests.

## Acknowledgements

We are greatly acknowledged for the technical support provided by Prof. Bernd M. Rode (University of Innsbruck) during this

research. The authors also would like to thank OpenEye Scientific Software for providing us the academic license free of cost for this study.

## References


[1] D.M. Quinn, Acetylcholinesterase: enzyme structure, reaction dynamics, and virtual transition states, Chem. Rev. 87 (1987) 955–979.
[2] B. Boyd, Ongoing progress in the Alzheimer's disease arena, Drug News Perspect. 13 (2000) 425.
[3] T. Dasgupta, P. Chitnumsub, S. Kamchonwongpaisan, C. Maneeruttanarungroj, S.E. Nichols, T.M. Lyons, J. Tirado-Rives, W.L. Jorgensen, Y. Yuthavong, K.S. Anderson, Exploiting structural analysis, in silico screening, and serendipity to identify novel inhibitors of drug-resistant falciparum malaria, ACS Chem. Biol. 4 (2009) 29–40.
[4] Z. Wang, Y. Lu, W. Seibel, D.D. Miller, W. Li, Identifying novel molecular structures for advanced melanoma by ligand-based virtual screening, J. Chem. Inf. Model. 49 (2009) 1420–1427.
[5] M.H.J. Seifert, M. Lang, Essential factors for successful virtual screening, Mini Rev. Med. Chem. 8 (2008) 63–72.
[6] G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, Drug Disc. Today 11 (2006) 580–594.
[7] N. Majeux, M. Scarsi, J. Apostolakis, C. Ehrhardt, A. Caflisch, Exhaustive docking of molecular fragments with electrostatic solvation, Proteins: Struct. Funct. Genet. 37 (1999) 88–105.
[8] X. Zou, Y. Sun, I.D. Kuntz, Inclusion of solvation in ligand binding free energy calculations using the generalized-born model, J. Am. Chem. Soc. 121 (1999) 8033–8043.
[9] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, 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. 5 (1997) 425–445.
[10] H. Gohlke, M. Hendlich, G. Klebe, Knowledge-based scoring function to predict protein-ligand interactions, J. Mol. Biol. 295 (2000) 337–356.
[11] R.S. DeWitte, E.I. Shakhnovich, SMoG: de novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence, J. Am. Chem. Soc. 118 (1996) 11733–11744.
[12] J.B.O. Mitchell, R.A. Laskowski, A. Alex, J.M. Thornton, B LEEP-potential of mean force describing protein-ligand interactions: I. Generating potential, J. Comp. Chem. 20 (1999) 1165–1176.
[13] I. Muegge, Y.C. Martin, A general and fast scoring function for protein-ligand interactions: a simplified potential approach, J. Med. Chem. 42 (1999) 791–804.
[14] M. Kontoyianni, L.M. McClellan, G.S. Sokol, Evaluation of docking performance: comparative data on docking algorithms, J. Med. Chem. 47 (2004) 558–565.
[15] A.R. Leach, B.K. Shoichet, C.E. Peishoff, Prediction of protein–ligand interactions. Docking and scoring: successes and gaps, J. Med. Chem. 49 (2006) 5851–5855.
[16] E. Perola, W.P. Walters, P.S. Charifson, A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance, Proteins: Struct. Funct. Bioinformatics 56 (2004) 235–249.
[17] G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff, M.S. Head, A critical assessment of docking programs and scoring functions, J. Med. Chem. 49 (2006) 5912–5931.
[18] M.D. Cummings, R.L. DesJarlais, A.C. Gibbs, V. Mohan, E.P. Jaeger, Comparison of automated docking programs as virtual screening tools, J. Med. Chem. 48 (2005) 962–976.
[19] B.J. McConkey, V. Sobolev, M. Edelman, The performance of current methods in ligand-protein docking, Curr. Sci. 83 (2002) 845–856.
[20] C.D. Graaf, P. Pospisil, W. Pos, G. Folkers, N.P. Vermeulen, Binding mode prediction of cytochrome P450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking, J. Med. Chem. 48 (2005) 2308–2318.
[21] G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, Drug. Disc. Today 11 (2006) 580–594.
[22] J.E. Ladbury, Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design, Chem. Biol. 3 (1996) 973–980.
[23] B.C. Roberts, R.L. Mancera, Ligand-protein docking with water molecules, J. Chem. Inf. Model. 48 (2008) 397–408.
[24] C.D. Graaf, P. Pospisil, W. Pos, G. Folkers, N.P.E. Vermeulen, Binding mode prediction of cytochrome P450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking, J. Med. Chem. 48 (2005) 2308–2318.
[25] C.D. Graaf, C. Oostenbrink, P.H.J. Keizers, T. Wijst, A. Jongejan, N.P.E. Vermeulen, Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking, J. Med. Chem. 49 (2006) 2417–2430.
[26] M.I. Choudhary, S.A. Nawaz, M.K. Azim, M.N. Ghayur, M.A. Lodhi, S. Jalil, A. Khalid, A. Ahmed, B.M. Rode, A.H. Gilani, Juliflorine: a potent natural peripheral anionic-site-binding inhibitor of acetylcholinesterase with calcium-channel blocking potential, a leading candidate for Alzheimer's disease therapy, Biochem. Biophys. Res. Commun. 332 (2005) 1171–1179.
[27] M.I. Choudhary, S.A. Nawaz, M.A. Lodhi, M.N. Ghayur, S. Jalil, N. Riaz, S. Yousuf, A. Malik, A.H. Gilani, Withanolides, a new class of natural cholinesterase inhibitors with calcium antagonistic properties, Biochem. Biophys. Res. Commun. 334 (2005) 276–287.
[28] Zaheer-ul-Haq, B. Wellenzohn, K.R. Liedl, B.M. Rode, Molecular docking studies of natural cholinesterase-inhibiting steroidal alkaloids from Sarcococca saligna, J. Med. Chem. 46 (2003) 5087–5090.
[29] Zaheer-ul-Haq, B. Wellenzohn, S. Tonmunphean, A. Khalid, M.I. Choudhary, B.M. Rode, 3D-QSAR Studies on natural acetylcholinesterase inhibitors of Sarcococca saligna by comparative molecular field analysis (CoMFA), Bioorg. Med. Chem. Lett. 13 (2003) 4375–4380.
[30] D.S. Goodsell, G.M. Morris, A.J. Olson, Automated docking of flexible ligands: applications of AutoDock, J. Mol. Recog. 9 (1996) 1–5.
[31] H.J. Böhm, 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. 12 (1998) 309–323.
[32] Molecular Operating Environment version 2006.08, 2006, Chemical Computing Group Inc., Quebec, Canada.
[33] J. Ding, K. Das, H. Moereels, L. Koymans, K. Andries, P.A.J. Janssen, S.H. Hughes, E. Arnold, Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse nonnucleoside inhibitors, Nat. Struct. Biol. 2 (1995) 407–415.
[34] A.N. Jain, Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine, J. Med. Chem. 46 (2003) 499–511.
[35] W. Welch, J. Ruppert, A.N. Jain, Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites, Chem. Biol. 3 (1996) 449–462.
[36] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[37] GOLD version 3.0, 2006, Cambridge Crytallographic Data Center Cambridge, UK.
[38] M.R. McGann, H.R. Almond, A. Nicholls, J.A. Grant, F.K. Brown, Gaussian docking functions, Biopolymers 68 (2003) 76–90.
[39] G.B. McGaughey, R.P. Sheridan, C.I. Bayly, J.C. Culberson, C. Kreatsoulas, S. Lindsley, V. Maiorov, J.F. Truchon, W.D. Cornell, Comparison of topological, shape, and docking methods in virtual screening, J. Chem. Inf. Model. 47 (2007) 1504–1519.
[40] G.P.A. Vigers, J.P. Rizzi, Multiple active site corrections for docking and virtual screening, J. Med. Chem. 47 (2004) 80–89.
[41] Q. Xie, Y. Tang, W. Li, X.H. Wang, Z.B. Qiu, Investigation of the binding mode of (−)-meptazinol and bis-meptazinol derivatives on acetylcholinesterase using a molecular docking method, J. Mol. Model. 12 (2006) 390–397.
[42] J. Guo, M.M. Hurley, J.B. Wright, G.H. Lushington, A docking score function for estimating ligand-protein interactions: application to acetylcholinesterase inhibition, J. Med. Chem. 47 (2004) 5492–5500.
[43] Z. Zhou, A.K. Felts, R.A. Friesner, R.M. Levy, Comparative performance of several flexible docking programs and scoring functions: enrichment studies for a diverse set of pharmaceutically relevant targets, J. Chem. Inf. Model. 47 (2007) 1599–1608.
[44] K. Onodera, K. Satou, H. Hirota, Evaluations of molecular docking programs for virtual screening, J. Chem. Inf. Model. 47 (2007) 1609–1618.
[45] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucl. Acids Res. 28 (2000) 235–242.
[46] M. Harel, I. Schalk, L. Ehret-Sabatier, F. Bouet, M. Goeldner, C. Hirth, P.H. Axelsen, I. Silman, J.L. Sussman, Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 9031–9035.
[47] M. Harel, D.M. Quinn, H.K. Nair, I. Silman, J.L. Sussman, The X-ray structure of a transition state analog complex reveals the molecular origins of the catalytic power and substrate specificity of acetylcholinesterase, J. Am. Chem. Soc. 118 (1996) 2340–2346.
[48] C.E. Felder, M. Harel, I. Silman, J.L. Sussman, Structure of a complex of the potent and specific inhibitor BW284C51 with Torpedo californica acetylcholinesterase, Acta Crystallogr. Sect. D 58 (2002) 1765–1771.
[49] H. Dvir, D.M. Wong, M. Harel, X. Barril, M. Orozco, F.J. Luque, D. Munoz-Torrero, P. Camps, T.L. Rosenberry, I. Silman, 3D structure of Torpedo californica acetylcholinesterase complexed with Huprine X at 2.1 Å resolution: kinetic and molecular dynamic correlates, Biochemistry 41 (2002) 2970–2981.
[50] G. Kryger, I. Silman, J.L. Sussman, Structure of acetylcholinesterase complexed with E2020 (Aricept®): implications for the design of new anti-Alzheimer drugs, Struc. Fold. Des. 7 (1999) 297–307.
[51] H. Dvir, H.L. Jiang, D.M. Wong, M. Harel, M. Chetrit, X.C. He, G.Y. Jin, G.L. Yu, X.C. Tang, I. Silman, D.L. Bai, J.L. Sussman, X-ray structures of torpedo californica acetylcholinesterase complexed with (+)-huperzine a and (−)-huperzine B: structural evidence for an active site rearrangement, Biochemistry 41 (2002) 10810–10818.
[52] P. Bar-On, C. Millard, M. Harel, H. Dvir, A. Enz, J. Sussman, I. Silman, Kinetic and structural studies on the interaction of cholinesterases with the anti-Alzheimer drug rivastigmine, Biochemistry 41 (2002) 3555–3564.
[53] D.M. Wong, H.M. Greenblatt, H. Dvir, P.R. Carlier, Y.F. Han, Y.P. Pang, I. Silman, J.L. Sussman, Acetylcholinesterase complexed with bivalent ligands related to huperzine A: experimental evidence for species-dependent protein-ligand complementarity, J. Am. Chem. Soc. 125 (2003) 363–373.
[54] Y. Bourne, P. Marchot, Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site, EMBO J. 22 (2003) 1–12.

[55] Y. Bourne, H.C. Kolb, Z. Radic, K.B. Sharpless, P. Taylor, P. Marchot, Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 1449–1454.
[56] M. Harel, G. Kryger, T.L. Rosenberry, W.D. Mallender, T. Lewis, R.J. Fletcher, J.M. Guss, I. Silman, J.L. Sussman, Three-dimensional structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors, Protein Sci. 9 (2000) 1063–1072.
[57] R.B.G. Ravelli, M.L. Raves, Z. Ren, D. Bourgeois, M. Roth, J. Kroon, I. Silman, J.L. Sussman, Static Laue diffraction studies on acetylcholinesterase, Acta Crystallogr. Sect. D 54 (1998) 1359–1366.
[58] J.P. Colletier, B. Sanson, F. Nachon, E. Gabellieri, C. Fattorusso, G. Campiani, M. Weik, Conformational flexibility in the peripheral site of Torpedo californica acetylcholinesterase revealed by the complex structure with a bifunctional inhibitor, J. Am. Chem. Soc. 128 (2006) 4526–4527.
[59] E.H. Rydberg, B. Brumshtein, H.M. Greenblatt, D.M. Wong, D. Shaya, L.D. Williams, P.R. Carlier, Y.P. Pang, I. Silman, J.L. Sussman, Complexes of alkylene-linked tacrine dimers with Torpedo californica acetylcholinesterase: binding of bis (5)-tacrine produces a dramatic rearrangement in the active-site gorge, J. Med. Chem. 49 (2006) 5491–5500.
[60] H.M. Greenblatt, C. Guillou, D. Guenard, A. Argaman, S. Botti, B. Badet, C. Thal, I. Silman, J.L. Sussman, The complex of a bivalent derivative of galanthamine with Torpedo acetylcholinesterase displays drastic deformation of the active-site gorge: implications for structure-based drug design, J. Am. Chem. Soc. 126 (2004) 15405–15411.
[61] M. Harel, J.L. Hyatt, B. Brumshtein, C.L. Morton, K.J.P. Yoon, R.M. Wadkins, I. Silman, J.L. Sussman, P.M. Potter, The crystal structure of the complex of the anticancer prodrug 7-ethyl-10-[4-(1-piperidino)-1-piperidino]-carbonyloxycamptothecin (CPT-11) with Torpedo californica acetylcholinesterase provides a molecular explanation for its cholinergic action, Mol. Pharmacol. 67 (2005) 1874–1881.
[62] H. Haviv, D.M. Wong, H.M. Greenblatt, P.R. Carlier, Y.P. Pang, I. Silman, J.L. Sussman, Crystal packing mediates enantioselective ligand recognition at the peripheral site of acetylcholinesterase, J. Am. Chem. Soc. 127 (2005) 11029–11036.
[63] SYBYL Molecular Modeling Software version 6.9, 2003, Tripos Associates St. Louis, MO.
[64] M. Clark, R.D. Cramer Iii, N. Van Opdenbosch, Validation of the general purpose Tripos 5.2 force field, J. Comp. Chem. 10 (1989) 982–1012.
[65] G. Koellner, G. Kryger, C.B. Millard, I. Silman, J.L. Sussman, T. Steiner, Active site gorge and buried water molecules in crystal structure of acetylcholinesterase from torpido californica, J. Mol. Biol. 296 (2000) 713–735.
[66] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function, J. Comp. Chem. 19 (1998) 1639–1662.
[67] T.A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, J. Comp. Chem. 17 (1996) 520–552.
[68] T.A. Halgren, MMFF VI. MMFF94 s option for energy minimization studies, J. Comp. Chem. 20 (1999) 720–729.
[69] H. Edelsbrunner, Weighted Alpha Shapes; Technical Paper of the Department of Computer Science of the University of Illinois at Urbana-Champaign, Urbana, Illinois 61810.
[70] J. Boström, J.R. Greenwood, J. Gottfries, Assessing the performance of OMEGA with respect to retreiving bioactive conformations, J. Mol. Graph. Model. 21 (2003) 449–462.
[71] http://cactus.nci.nih.gov/ncidb2/download.html.
[72] R. Wang, Y. Lu, S. Wang, Comparative evaluation of 11 scoring functions for molecular docking, J. Med. Chem. 46 (2003) 2287–2303.
[73] R.D. Clark, A. Strizhev, J.M. Leonard, J.F. Blake, J.B. Matthew, Consensus scoring for ligand/protein interactions, J. Mol. Graph. Model. 20 (2002) 281–295.
[74] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 46 (2001) 3–25.
[75] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, S. Arthurs, A.B. Colson, S.T. Freer, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Deciphering common failures in molecular docking of ligand–protein complexes, J. Comput.-Aided Mol. Des. 14 (2000) 731–751.
[76] B. Kramer, M. Rarey, T. Lengauer, Evaluation of the incremental construction algorithm for protein-ligand docking, Proteins: Struc. Func. Genet. 37 (1999) 228–241.