A Docking Score Function for Estimating Ligand−Protein Interactions: Application to Acetylcholinesterase Inhibition

Jianxin Guo,[†] Margaret M. Hurley,[‡] Jeffery B. Wright,[§] and Gerald H. Lushington*,[†]

Molecular Graphics and Modeling Lab, University of Kansas, Lawrence, Kansas 66045, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005-5067, and U.S. Army RDECOM/Natick Soldier Center, Natick, Massachusetts 01760-5020

Acetylcholinesterase (AChE) inhibition is an important research topic because of its wide range of associated health implications. A receptor-specific scoring function was developed herein for predicting binding affinities for human AChE (huAChE) inhibitors. This method entails a statistically trained weighted sum of electrostatic and van der Waals (VDW) interactions between ligands and the receptor residues. Within the 53 ligand training set, a strong correlation was found ($R^2 = 0.89$) between computed and experimental inhibition constants. Leave-one-out cross-validation indicated high predictive power ($Q^2 = 0.72$), and analysis of a separate 16-compound test set also produced very good correlation with experiment ($R^2 = 0.69$). Scoring function analysis has permitted identification and characterization of important ligand−receptor interactions, producing a list of those residues making the most important electrostatic and VDW contributions within the main active site, gorge area, acyl binding pocket, and periferal site. These analyses are consistent with X-ray crystallographic and site-directed mutagenesis studies.

Introduction

Acetylcholinesterase (AChE) is an enzyme that hydrolyzes the neurotransmitter acetylcholine (ACh) at cholinergic synapses, accomplishing its role at a rate faster than those of most other known enzymes.[1,2] Recent research interest regarding this enzyme is not only due to this high catalytic efficiency but also due to the broad implications of AChE inhibition on human health, agrochemistry, and chemical agents. For example, Alzheimer's disease (AD) is associated with low in vivo levels of acetylcholine; thus, AChE has been targeted in many drug discovery projects aimed at maintaining ACh availability via mild or reversible inhibitors such as tacrine[3] and donepezil,[4] etc. While low-level AChE inhibition is useful for such neurological treatments, higher levels of inhibition can be detrimental. Organophosphorus (OP) compounds, in particular, irreversibly deactivate AChE and may induce failure of cholinergic synaptic transmission, deterioration of neuromuscular junctions, flaccid muscle paralysis, and central nervous system seizures.[5,6] Effective drug design thus requires great care in balancing the level of inhibitive efficacy.

The availability of AChE crystal structures for various species with and without ligands provides a solid basis for structure-based design of novel AChE inhibitors.[7] There are two principle binding sites in the AChE. The catalytic active site is located at the base of a deep gorge in the enzyme. It contains a catalytic triad, Ser203, Glu334, and His447 (huAChE sequence numbering, to

be used throughout unless otherwise specified), and nearby residues (e.g., the choline binding site: Trp86) that collectively effect the ACh catalysis reactions.[8] AChE also has a peripheral anionic site (PAS) located near the enzyme surface at the mouth of the active site gorge. The residue Trp286 plays a very important role in ligand binding in the PAS. Ligand binding to the PAS affects enzymatic activity through a combination of steric blockade of ligands moving through the gorge and allosteric alteration of the catalytic triad conformation and efficiency.[9] The gorge itself is a narrow hydrophobic channel with a length of ∼20 Å, connecting the PAS site to the active site.[10] An acyl binding pocket consists of residues Gly122, Trp236, Phe295, Phe297, and Phe338 and is responsible for interacting with the acetyl group.[11] Early inhibition research was mainly focused on ligands binding in the active site (e.g., tacrine[3], amiridine,[12] etc.). Recent efforts have 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).

Molecular modeling has proven increasingly important in helping to design novel enzyme inhibitors. A substantial amount of prior AChE inhibitor research has focused on using ligand-based design methods such as CoMFA.[13−17] Given an accurate receptor structure, molecular docking can also be very useful in characterizing ligand−receptor binding by providing predictions of the bound conformation for the ligand and a scheme for energetically ranking (i.e., scoring) the ligand−receptor interaction. Great successes have been achieved in terms of conformational predictions via flexible docking programs such as Dock,[18] Gold,[19] FlexX,[20] etc. Such conformational predictions are very important to drug design because (1) the binding conformation of

* To whom correspondence should be addressed. Fax: (785) 864-5326. Phone: (785) 864-1140. E-mail: glushington@ku.edu.
† University of Kansas.
‡ U.S. Army Research Laboratory.
§ U.S. Army RDECOM/Natick Soldier Center.

ligands is much easier to validate (i.e., through comparison with experimentally observed structures) than is the binding affinity for different systems and (2) the accurate prediction of the bound conformation is a prerequisite for reliable scoring. Even with good structural predictions, however, the score may not always agree well with experimentally determined affinities mainly because experimental conditions include important dynamic or entropic effects that are difficult to rigorously represent in a general scoring function. To account for such effects empirically, scoring functions are typically trained via diverse sets of previously characterized ligand−receptor interactions. Unfortunately, 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. Indeed, with the AChE system, studies on steroidal alkaloid inhibition show no correlation between the calculated binding energy and experimentally determined activity.[21] Molecular dynamics simulations do provide a natural means for quantifying both the entropic and enthalpic components of binding affinity for ligand−AChE interactions;[20] however, extensive computational demands make molecuar dynamics (MD) simulations prohibitively time-consuming for analysis of large compound collections. In such cases, the best compromise may be to carry out simple docking studies but to explicitly train the scoring function to reproduce behavior in the system of interest. Herein, we present a new scoring function that is based on the ligand−receptor interaction field and is trained specifically to reproduce AChE inhibition.

Methods

The AChE crystal structures used herein were obtained from the Protein Databank (PDB). Our main docking and training activities have been focused on a human AChE (huAChE) structure (code, 1B41)[22] but rely on ligand binding information from a *Torpedo californica* AChE (tcAChE) structure (code, 1EVE) that includes a cocrystallized E2020 inhibitor.[23] Given the absence of a firm understanding of the persistence and roles of individual solvent molecules in and around the AChE binding sites, all waters were removed from the structures. To ascertain the orientation of the ligand E2020 relative to the huAChE structure, the huAChE structure was aligned to the tcAChE in SYBYL[24] by achieving a maximal overlap of C_α atoms for corresponding huAChE/tcAChE residues within the receptor region. The resulting root-mean square deviation (rmsd) between the two aligned huAChE/tcAChE structures is only 0.85 Å for the set of all backbone C_α atoms within the full enzyme subunit, suggesting good overall alignment and substantial structural similarity. Hydrogen atoms were added (via SYBYL) to the resulting huAChE−E2020 complex. The positions of these new protons were then optimized in MOE[25] via molecular mechanics using the MMFF94s force field[26] (all heavy atoms fixed) to avoid bad interatomic contacts. The position of E2020 was then optimized (all receptor atoms fixed) to determine a plausible stable conformational structure for the ligand in the receptor environment. In both of the above simulations, MMFF94s charges were used to account for relevant electrostatics. The

steepest descent minimization algorithm was used for the first 100 steps (unless an rms gradient of less than 100 kcal/(mol·Å) was first achieved), followed by 200 steps of conjugate gradient (unless an rms gradient of less than 1 kcal/(mol·Å) was attained), and finally completed by 1000 steps of truncated Newton (or an rms gradient of less than 0.01 kcal/(mol·Å)). The resulting E2020 structure was then extracted for subsequent docking calculations.

Sixty-nine compounds with IC_{50} data measured with human AChE assay[27−30] were selected for training and testing the scoring function. The activity among these compounds ranges from 0.33 to 30 000 nM (Tables 1 and 2). The active site for the huAChE docking calculations was constructed from the crystal structure by retaining all residues within a radius of 12 Å relative to E2020 (but discarding the original ligand itself). Docking calculations were carried out with the Gold program.[19] A genetic algorithm was used in searching the binding conformation of flexible ligands, using the default parameters in GOLD. A maximum of 20 poses were computed for each compound. Those docked conformations were saved in SDF format and then imported into SYBYL for scoring calculations according to the FlexX and CSCORE modules. The scoring methods available included empirical methods such as ChemScore,[31] FlexX score,[20] and G Score[19] and knowledge-based methods such as PMF score[32] and DrugScore.[33] Multilinear regression (MLR) was used to obtain a consensus score from these methods. One conformation was selected for each compound to give a good compromise between the best consensus score and those with the closest alignment to the original E2020 ligand. Specifically, the pose for the scoring of the activity was selected on the basis of having the highest consensus score (first criterion) and ChemScore (tie-breaking criterion), with the further stipulation that the following knowledge-based criteria (as determined by visual inspection) must be obeyed whenever possible: (1) good $\pi−\pi$ overlap with residue Trp86, as has been found to be critical for E2020 binding;[23] (2) good $\pi−\pi$ overlap with residue Trp286, as has also been found to be very important for E2020 complexation[23]

The chosen conformations were used to fit an interaction field whose form, basically a variant of the comparative binding energy (COMBINE) method,[34−36] is as follows:

$$pIC_{50} = \sum_i c_i E_i^{ele} + \sum_j d_j E_j^{vdw} \qquad (1)$$

where c_i and d_j are fitted coefficients, E_i^{ele} and E_j^{vdw} are electrostatic and van der Waals (VDW) interactions arising between atoms in the *i*th and *j*th residues and the ligand. In this expression, all receptor residues within 10 Å of the position of the original E2020 ligand (a total of 92 residues) were included in the summation over *i*, and E_i^{ele} and E_j^{vdw} were calculated via an SVL script written for the MOE system. The statistic analysis was performed in Simca-P[37] with partial least squares regression (PLS). Fifty-three of the full 69 compounds were selected as our training set, and the other 16 compounds were used as a test set for validating the predictive power of the new scoring function.

Table 1. huAChE Ligands in the Training Set

-R1	-R2	-R3	-X-	-Y-	ID	IC50 (nM)	Ref	pKi exp	pKi calc
-H	-H	-H	-(CH$_2$)$_2$-	-O-	1	55	[28]	7.26	7.47
-CH$_3$	-H	-H	-(CH$_2$)$_2$-	-O-	2	7.8	[28]	8.11	8.05
-CH$_3$	-OCH$_3$	-H	-(CH$_2$)$_2$-	-O-	3	5.8	[28]	8.24	8.23
-OCH$_3$	-H	-H	-(CH$_2$)$_2$-	-O-	4	7.2	[28]	8.14	7.62
-H	-H	-OCH$_3$	-(CH$_2$)$_2$-	-O-	5	7.1	[28]	8.15	7.55
-H	-NH-CO-CH$_3$	-H	-(CH$_2$)$_2$-	-O-	6	2.8	[28]	8.55	8.40
-H	-NH-SO$_2$-φ	-H	-(CH$_2$)$_2$-	-O-	7	14	[28]	7.85	8.17
-H	-4-morpholino	-H	-(CH$_2$)$_2$-	-O-	8	0.8	[28]	9.10	9.23
-H	-NH$_2$	-H	-(CH$_2$)$_2$-	-O-	9	20	[28]	7.70	8.08
-H	-Br	-H	-(CH$_2$)$_2$-	-O-	10	50	[28]	7.30	7.06
-H	-CN	-H	-(CH$_2$)$_2$-	-O-	11	101	[28]	7.00	7.03
-H	-CO-NH$_2$	-H	-(CH$_2$)$_2$-	-O-	12	8.8	[28]	8.06	7.57
-H	-H	-H	-(CH$_2$)$_3$-	-O-	13	900	[28]	6.05	6.06
-H	-H	-H	-O-CH$_2$-	-O-	14	2600	[28]	5.59	5.65
-H	-H	-H	-NH-CH$_2$-	-O-	15	320	[28]	6.49	7.00
-H	-H	-H	-(CH$_2$)$_2$-	-S-	16	99	[28]	7.00	6.89
-H	-H	-H	-(CH$_2$)$_2$-	-CH=CH-	17	220	[28]	6.66	7.11
-H	-H	-H	-(CH$_2$)$_2$-	-NH-	18	120	[28]	6.92	6.82
-CH$_2$-CH$_2$-CO-NH-		-H	-(CH$_2$)$_2$-	-O-	19	0.57	[29]	9.24	8.89
-NH-CO-CH$_2$-		-H	-(CH$_2$)$_2$-	-O-	20	0.95	[29]	9.02	8.32
-N(CH$_3$)-CO-CH$_2$-		-H	-(CH$_2$)$_2$-	-O-	21	0.48	[29]	9.32	9.19
-H	-NH-CO-CH$_2$-		-(CH$_2$)$_2$-	-O-	22	3.6	[29]	8.44	8.62

-W-	-X-Y-	-Z-	-R	ID	IC50 (nM)	Ref	pKi exp	pKi calc
-	-CH$_2$-CH-	-C(O)-	-CH$_2$-φ	23	250	[27]	6.60	6.82
-C(O)-	-CH$_2$CH$_2$-CH-	-S-	-CH$_2$-φ	24	8	[27]	8.10	8.33
-C(O)-	-CH$_2$C(OH)-CH-	-S-	-CH$_2$-φ	25	43	[27]	7.37	7.33
-C(O)CH$_2$-	-C(OH)CH$_2$CH$_2$-CH-	-S-	-CH$_2$-φ	26	380	[27]	6.42	6.50

Table 1 (Continued)
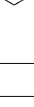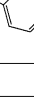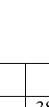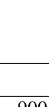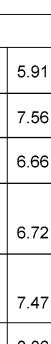
-C(O)CH₂-	-CH₂CH₂CH₂-CH-	-S-	-CH₂-φ	27	110	[27]	6.96	6.64
-C(O)-	-CH₂C(OCH₃)-CH-	-S-	-CH₂-φ	28	120	[27]	6.92	7.22
-C(O)-	-HC=C-	-S-	-CH₂-φ	29	520	[27]	6.28	6.90
-	-C(OH)-CH-	-S-	-CH₂-φ	30	19580	[27]	4.71	4.73
-	-HC=C-	-S-	-CH₂-φ	31	2670	[27]	5.57	5.46
-C(O)-	-CH₂CH₂-CH-	-S-	-(CH₂)₂OCH₂	32	53	[27]	7.28	6.89
-C(O)-	-CH₂CH₂-CH-	-S-	(furan-NO₂)	33	32	[27]	7.49	7.84
-C(O)-	-CH₂CH₂-CH-	-S-	(tetrahydropyran)	34	28	[27]	7.55	7.51
-C(O)-	-CH₂CH₂-CH-	-S-	(tert-butyl ester)	35	79	[27]	7.10	6.67
-C(O)-	-CH₂CH₂-CH-	-S-	-CH₂CH₂-O-φ	36	390	[27]	6.41	6.52
-C(O)-	-CH₂CH₂-CH-	-S-	-CH₂-CN	37	1000	[27]	6.00	6.06

-R1						
-CH₃		38	900	[27]	6.05	5.91
-CH₂CH₃		39	280	[27]	6.55	7.56
-CH₂CH=CH₂		40	540	[27]	6.27	6.66
(cyclopropylmethyl)		41	110	[27]	6.96	6.72
(cyclobutylmethyl)		42	40	[27]	7.40	7.47
-CH₂CH₂-O-CH₂CH₃		43	7	[27]	8.15	8.06
(4-fluorobenzyl)		44	2.6	[27]	8.59	8.39
(methyl benzoate)		45	1000	[27]	6.00	6.07
(thiophene)		46	6	[27]	8.22	7.76
(furan-NO₂)		47	4.5	[27]	8.35	8.33

-R1						
(benzodioxane)		48	100	[30]	7.00	7.32
(methylenedioxybenzene)		49	41.5	[30]	7.38	7.30
(benzoxepine)		50	139	[30]	6.86	6.87
(methoxynaphthalene)		51	50	[30]	7.30	7.60
(dimethoxy-methylindole)		52	120	[30]	6.92	7.61
(phenylpyrazole)		53	22	[30]	7.66	7.89

Table 2. Testing Set of huAChE Ligands, Showing Corresponding Inhibition Data

-R1	-R2	-R3	-X-	-Y-	ID	IC50 (nm)	Ref	pKi exp	pKi pred
-H	-OCH₃	-H	-(CH₂)₂-	-O-	54	8.3	[28]	8.08	8.40
-H	-NH-CO-φ	-H	-(CH₂)₂-	-O-	55	9.4	[28]	8.03	7.76
-H	-OH	-H	-(CH₂)₂-	-O-	56	26	[28]	7.59	7.01
-H	-H	-H	-(CH)₂-	-O-	57	210	[28]	6.68	6.85
-H	-H	-H	-NH-(CH₂)₂-	-O-	58	810	[28]	6.09	7.43
-H	-H	-H	-(CH₂)₂-	-N=CH₂	59	340	[28]	6.47	7.06
-CH₂CONH-		-H	-(CH₂)₂-	-O-	60	0.33	[29]	9.48	8.18

W	X-Y-Z	-R	ID	IC50 (nm)	Ref	pKi exp	pKi pred
-C(O)-	-CH₂CH(OH)-CH-CH₂-	-CH₂-φ	61	190	[27]	6.72	6.36
-C(O)-	-CH₂-C(OH)-CH₂-	-CH₂-φ	62	90	[27]	7.05	6.41
-C(O)-	-CH₂-CH=CH -	-CH₂-φ	63	750	[27]	6.12	6.37
-	-CH₂-CH-CH₂-	-CH₂-φ	64	30000	[27]	4.52	5.33
-C(O)-	-(CH₂)₂-CH-CH₂-	-CH₂COOCH₃	65	54	[27]	7.27	6.89

-R		ID	IC50 (nm)	Ref	pKi exp	pKi pred
-(CH₂)₂CH₃		66	2570	[27]	5.59	5.23
-(CH₂)₂OCH₃		67	30	[27]	7.52	7.54
-CH₂-φ		68	4.6	[27]	8.34	8.34

	ID	IC50 (nm)	Ref	pKi exp	pKi pred
	69	240	[30]	6.62	5.76

Results and Discussion

Our scoring model built via PLS regression over interactions within the 53-molecule training set appears to be of reasonable quality, with a correlation coefficient of $R^2 = 0.89$ and a leave-one-out cross-validation correlation of $Q^2 = 0.72$. In using the scoring function to evaluate the activity of the 16-molecule testing set, we achieved good predictivity: a correlation of $R^2 = 0.69$ (Figure 1) between the calculated results and experimental values.

To compare the precision and extensibility of our scoring function, we contrasted the above results with predictions made using several commercially available scoring methods, including ChemScore,[31] FlexX score,[20] DrugScore,[33] G Score,[19] and PMF score.[32] The correlation between the experiment and any single scoring method is poor. The PMF score showed the best correlation but was still poor ($R^2 = 0.13$ for the training set). All of the other representations gave even worse correlations: ChemScore = 0.07, FlexX = 0.05, Drug-
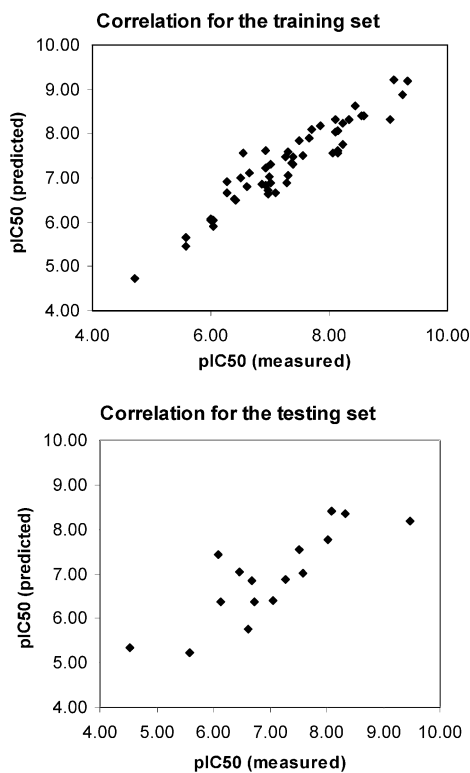
Figure 1. Correlation of the calculated activity (pIC_{50}) with experiment: (top) training set ($R^2 = 0.89$); (bottom) testing set ($R^2 = 0.69$).

Score = 0.01, G Score = 0.004. Since those general scoring methods were not trained in this AChE inhibition set, it is not completely fair to compare them directly with our specially tailored energy decomposition method. Therefore, we built a consensus score by training (over the 53-molecule set) a weighted sum over the above five commercially available scoring functions as follows:

$$pIC_{50} = 0.05682 \text{ Chemscore} - 0.00499 \text{ Drugscore} - 0.03582 \text{ Flexscore} - 0.01232 \text{ Gscore} + 0.01847 \text{ PMFscore} + 7.62820 \quad (2)$$

where Chemscore, Drugscore, Flexscore, Gscore, and PMFscore refer to computed affinities from the Chem-Score, DrugScore, FlexX score, G Score, and PMF score methods, respectively. Although an improvement was found for this consensus score within the training set itself ($R^2 = 0.26$), its predictive potency is poor, judging by no evident correlation within the 16-molecule testing set.

To help verify the physical sensibility of our model, we have mapped out the residues that contribute significantly to the scoring function. The coefficients of the 20 most important residues in terms of electrostatic and VDW contributions are shown in Figure 2. In those residues, Trp86, Ile451, Gly448, Tyr449, and Ser229 are the most important residues in the active site for VDW interactions. Trp86 functions by forming $\pi-\pi$ interaction with ligand aryl groups (when available), while the other residues define the shape of the gorge base, serving to discriminate according to ligand shape. In the upper gorge area and the acyl binding pocket, residues Tyr124, Phe295, Phe338, and Phe297 are responsible for providing hydrophobic contacts. The ring of Tyr72
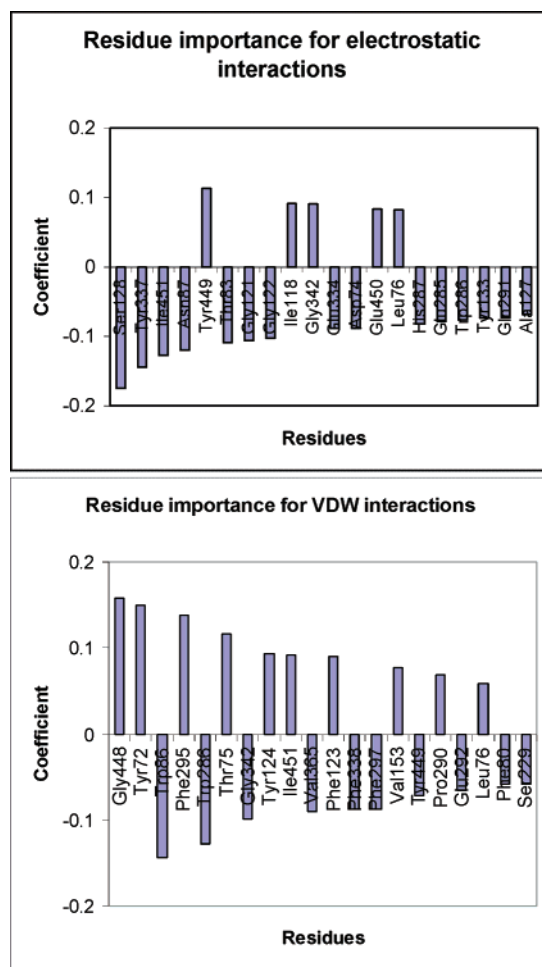

Figure 2. (Top) Coefficients of the 20 most important residues for electrostatic interactions. (Bottom) Coefficients of the 20 most important residues for VDW interactions.

is almost perpendicular to the Trp286 ring and forms a blocking wall to prevent the ligand ring from moving away from the position where it forms a $\pi-\pi$ interaction with the Trp286 ring. Phe295, Phe297, Val365, and Glu292 form another wall on the other side of the gorge, stretching from the acyl pocket toward the PAS.

Residues Tyr449, Glu450, Ile451, Ala127, Ser128, Tyr133, Ile118 near Trp86, and the "oxyanion hole" residues Gly121 and Gly122 are important in providing electrostatic interactions in the active site. Tyr337, Asp74, Thr83, and Asn87 are the primary electrostatic contributors in the gorge area. Gly342, Leu76, Glu285, Trp286, His287, and Gln291 are probably helpful in enhancing the activity of ligands with polar groups oriented in this area, as is evidenced by reports that an AChE inhibitor tethering in the position of His287 can affect the binding affinity as much as 14-fold.[38] Site-directed mutagenesis in huAChE indicated that Asp74, Tyr337, Phe338, Phe295, Phe297, Tyr133, and Glu450 can affect the affinity although it has been difficult to determine experimentally whether these residues contribute mainly electrostatic or VDW interactions.[39,40]

The docked ligand structures generally support the above analysis regarding the identity of principle residues. In our current docking calculations, molecules 2−7 all share similar conformations in the PAS. The modified groups in those molecules are actually exposed to the solvent and do not contribute directly to the ligand−
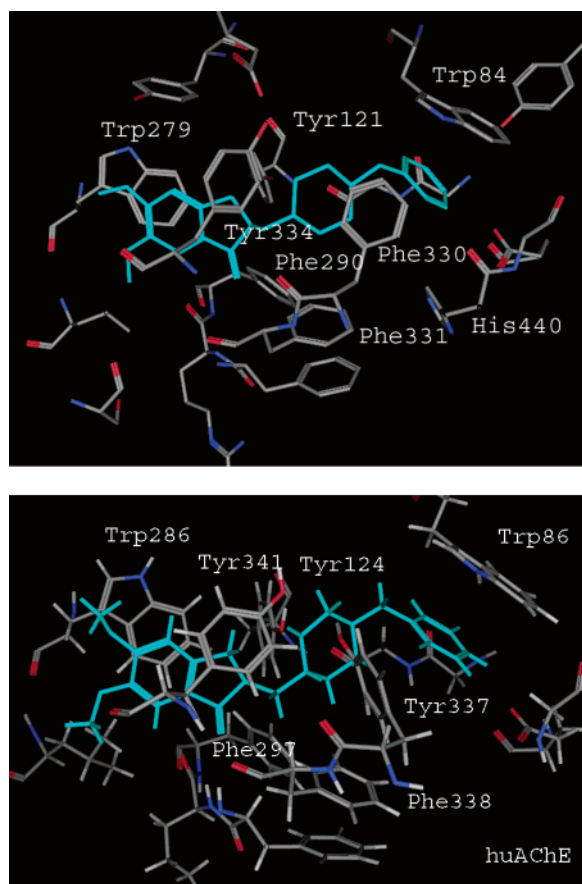
Figure 3. (Top) Structure of E2020 binding to tcAChE. E2020 is rendered in cyan. (Bottom) Structure of the E2020 binding in huAChE. E2020 is rendered in cyan.

receptor interaction. This is the same as observed in previous studies.[28] However, in the current work, molecule 8 takes a slightly different conformation with its morpholino group situating in the half-buried pocket by Trp286, His287, Ser293, Glu292, and Leu289. The morpholino oxygen has a distance of 3.65 Å from the backbone N of Glu292, and the morpholino nitrogen is 3.78 Å from the backbone O of Ser293. Such interactions could slightly pull the benzisoxazole ring away from Trp286 ring, leaving only a partial $\pi-\pi$ interaction. This particular conformation leads to an affinity increased by more than 10-fold to 0.8 nM relative to molecules 2−7. This particular region has been confirmed in an X-ray structure to be very important for the inhibitors binding to the PAS.[9]

As a final point of validation, we compared the calculated conformation for E2020 within the huAChE crystal structure relative to its original cocrystallized conformation in tcAChE. Our calculated structure suggests that E2020 has similar but not identical binding modes in tcAChE and huAChE (Figure 3). In the active site, its benzyl ring forms a $\pi-\pi$ interaction with the indole ring of Trp86 in huAChE and Trp84 in tcAChE. In the PAS, the indanone ring of E2020 forms a $\pi-\pi$ interaction with the indole ring of Trp286 in huAChE and Trp279 in tcAChE. The charged nitrogen of the E2020 piperidine ring undergoes a cation−π interaction with the phenyl ring of Phe330 in tcAChE. The corresponding residue Tyr337 in huAChE does not form a similar cation−π interaction due to the steric limitations

in this area associated with an inauspicious orientation of the Tyr337 ring. As a result, the nitrogen of the piperidine ring of the E2020 instead interacts with the hydroxyl groups of the Tyr337, Tyr124, and Ser125 within distances of 3.41, 3.12, and 4.18 Å relative to the oxygen atom of the respective hydroxyl groups. To probe the role of Tyr337, we examined the potential energy curve for Tyr337 side chain rotation relative to the other residues (energy evaluation according to the MMFF94s force field with appropriate charges) but found only one minimum in the potential curve. Closer analysis reveals that that the Tyr337 ring is trapped in a local pocket formed by Phe338, Tyr341, Trp439, Trp449, and His447. In previous structural studies and molecular dynamics simulations, it has been found that Phe330 in tcAChE can adopt a wide range of conformations in the complex structure and may function as a swinging gate[15,23,39] that structurally couples the anionic subsite of the active site and the PAS. It is natural to expect similar behavior of Tyr337 in the huAChE compared with the analogous Phe330 in tcAChE. Such a gate swing movement of the Tyr337 ring would certainly induce a shift in attached and neighboring residues; thus, one function of this PAS−active site coupling could be to effect proper residue alignment within the anionic subsite. The fact that the huAChE crystal structure used in this study to train our scoring function did not have a cocrystallized active site inhibitor (having only a PAS-bound fasciculin molecule[22]), and thus did not reflect such a conformational shift in the Tyr337, may constitute a subtle flaw in our model. Given our model's fairly strong predictive capacity, we expect the flaw to be of only minor consequence, however.

The accord between results derived from our scoring method and those of X-ray structures and mutagenesis observations indicates the effectiveness of the current analysis and the scoring function's predictive power. Since this method requires a set of compounds with known activity data to fit the score function, it is most applicable to the task of lead optimization as opposed to de novo discovery. In contrast to the comparative molecular field analysis (CoMFA)[41] method, which uses probe atoms such as nitrogen, carbon, etc. to calculate possible interaction fields between ligands and a putative receptor (and does not require explicit atomic-level representation of the receptor structure), our method uses a "real" interaction field between the ligands and receptor, thus requiring advance knowledge of the receptor structure. The benefit of using the current method is that the important residues can be directly identified and verified by site-directed mutagenesis, whereas CoMFA merely generates a hypothetical map of possible favorable and unfavorable interaction regions based on statistically postulated correlations between activity and orientations of specific functional groups within the ligand training set. Such a hypothetical map may correspond to the real interaction field if the underlying postulated statistical correlation is valid but is unlikely to be as accurate a representation as a real interaction field. In cases where crystal structures of the relevant receptor are available, CoMFA-generated hypotheses may be adjusted in order to directly correspond to the real field. In this vein, there have been successful studies that used receptor-based docking

methods to align the ligands and then used CoMFA on the resulting alignments to ascertain and analyze the resulting grid for drug design purposes.[13,14,16,17] This method is similar in principle to the COMBINE-like technique that we have applied herein to this huAChE system except that COMBINE methods are easier to compute and to understand because they only require treatment of ligand−residue interactions (hundreds of terms or less) rather than the thousands of grid points in CoMFA models. Another benefit relative to CoMFA models is that the relative simplicity of our scoring function also allows it to be used directly in any subsequent docking calculations, thus providing a means to accurately score novel inhibitor candidates and predict their bound conformations rather than having to do so in a stepwise manner.

Conclusions

A new AChE-specific scoring function has been developed herein and used in predicting binding affinity of AChE inhibitors. The method is based on a COMBINE-type approach, incorporating the electrostatic and VDW interaction fields between the ligands and receptor. A 53-compound training set was used to construct the scoring function, and a further 16 compounds were used to test the resulting model. Strong statistical correlations were found between predicted and observed affinities for both the training and testing set. Analysis of the scoring function has permitted identification of those receptor residues making the most important contributions to ligand binding. These analyses are consistent with the X-ray structure and mutagenesis studies. A comparison with other scoring methods and consensus scoring indicated the high effectiveness and predictability of this method.

Supporting Information Available: Scientific vector language (SVL) script for extracting nonbonding energy terms to compute the ligand−receptor interaction energy. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Taylor, P.; Radic, Z. The cholinesterases: from genes to proteins. *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 281−320.
(2) Massoulie, J.; Pezzementi, L.; Bon, S.; Krejci, E.; Velette, F. M. Molecular and cellular biology of the cholinesterases. *Prog. Neurobiol.* **1993**, *41*, 31−39.
(3) Crismon, M. L. Tacrine: first drug approved for Alzheimer's disease. *Ann. Pharmacother.* **1994**, *28*, 744−751.
(4) Barner, E. L.; Gray, S. L. Donepezil use in Alzheimer disease. *Ann. Pharmacother.* **1998**, *32*, p 70−77.
(5) Sidell, F. R.; Borak, J. Chemical warfare agents: II. Nerve agents. *Ann. Emerg. Med.* **1992**, *21*, p 865−871.
(6) Marrs, T. C. Organophosphate poisoning. *Pharmacol. Ther.* **1993**, *58*, 51−66.

(7) Barril, X.; Orozco, M.; Luque, F. J. Towards improved acetylcholinesterase Inhibitors: A structural and computational approach. *Mini-Rev. Med. Chem.* **2001**, *1*, 255−266.
(8) Rachinsky, T. L; Camp, S.; Li, Y.; Ekstrom, J.; Newton, M.; Taylor, P. Molecular cloning of mouse acetylcholinesterase: tissue distribution of alternatively spliced mRNA species. *Neuron* **1990**, *5*, 317−327.
(9) Bourne, Y.; Taylor, P.; Radic, Z.; Marchot, P. Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site. *EMBO J.* **2003**, *22*, 1−12.
(10) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of aetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. *Science* **1991**, *253*, 872−879.
(11) Harel, M.; Quinn, D. M.; Nair, H. K.; Silman, I.; Sussman, J. L. 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.* **1996**, *118* (10), 2340−2346.
(12) Kojima, J.; Nakajima, K.; Ochiai, M.; Nakayama, K. Effects of NIK-247 on cholinesterase and scopolamine-induced amnesia. *Methods. Find. Exp Clin. Pharmacol.* **1997**, *19*, 245−251.
(13) Sippl, W.; Contreras, J.-M.; Parrot, I.; Rival, Y. M.; Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 395−410.
(14) Sippl, W. Development of biologically active compounds by combining 3D QSAR and structure-based design methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 825−830.
(15) Kua, J.; Zhang, Y.; Mccammon, J. A. Studying enzyme binding specificity in acetylcholinesterase using a combined molecular dynamics and multiple dockign approach. *J. Am. Chem. Soc.* **2002**, *124*, 8260−8267.
(16) Bernard, P.; Kireev, D. B.; Chretien, J. R.; Fortier, P.-L.; Coppet, L. Automated docking of 82 *N*-benzylpiperidine derivatives to mouse acetylcholinesterase and comparative molecular field analysis with natural alignment. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 355−371.
(17) Cho, S. J.; Garsia, M. L. S.; Bier, J.; Tropsha, A. Strucutre-Based Alignment and Comparative Molecular Field Analysis of Acetylcholinesterase Inhibitors. *J. Med. Chem.* **1996**, *39*, 5064−5071.
(18) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule−ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269−288.
(19) Jones, G.; Willett, P.; Glen, R.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(20) Rarey M.; Kramer, B.; Lengauer, T.; Klebe, G. J. A fast flexible docking method using an imcremental constructuion algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.
(21) Haq, Z.-U.; Wellenzohn, B.; Liedl, K. r.; Rode, B. M. Molecular docking studies of Natural Cholinesterase Inhibiting Steroidal Alkaloids from *Sarcococca saligna*. *J. Med. Chem.* **2003**, *46*, 4087−4090.
(22) Kryger, G.; Harel, M.; Giles, K.; Toker, L.; Velan, B.; Lazar, A.; Kronman, C.; Barak, D.; Ariel, N.; Shafferman, A.; Silman, I.; Sussman, J. L. Structure of Recombinant Native and E202Q Mutant Human Acetylcholinesterase Complexed with the Snake-Venom Toxin Fasciculin-II. *Acta Crystallogr., Sect. D* **2000**, *56*, 1385−1394.
(23) 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*, 297−307.
(24) *Sybyl/Unity*, version 6.9; Tripos Inc.: St. Louis, MO, 2002.
(25) *Molecular Operating Enviroment (MOE) 2002.03*; Chemical Computing Group Inc., Montreal, Quebec, Canada, 2002.
(26) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**. *20*, 720−729.
(27) Palin, R.; Clark, J. K.; Cowley, P.; Muir, A. W.; Pow, E.; Prosser, A. B.; Taylor, P.; Zhang, M.-Q. Novel Piperidinium and Pyridinium Agents as Water-Soluble Acetylcholinesterase Inhibitors for the Reversal of Neutomuscular Blockade. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2569−2572.
(28) Villalobos, A.; Blake, J. F.; Biggers, C. K.; Butler, T. W.; Chapin, D. S.; Chen, Y. L.; Ives, J. L.; Jones, S. B.; Liston, D. R.; Nagel, A. A.; Nason, D. M.; Nielsen, J. A.; Shalaby, I. A.; White, W. F. Novel Benzisoxazole Derivatives as Potent and Selective Inhibitors of Acetylcholinesterase. *J. Med. Chem.* **1994**, *37*, 2721−2734.
(29) Villalobos, A.; Butler, T. W.; Chapin, D. S.; Chen, Y. L.; Demattos, S. B.; Ives, J. L.; Jones, S. B.; Liston, D. R.; Nagel, A. A.; Nason, D. M.; Nielsen, J. A.; Ramirez, A. D.; Shalaby, I. A.; White, W. F. 5,7-Dihydro-3-[2-[1[(phenylmehtyl)-4-piperidinyl]-ethyl]-6*H*-pyrrolo[3,2-*f*]-1,2-benzisoxazol-6-one: A Potent and Centrally-Selective Inhibitor of Acetylcholinesterase with an Improved Margin of Safety. *J. Med. Chem.* **1995**, *38*, 2802−2808.

(30) Clark, J. K.; Cowley, P.; Muir, A. W.; Palin, R.; Pow, E.; Prosser, A. B.; Taylor, R.; Zhang, M.-Q. Quaternary Salts of E2020 Analogues as Acetylchoinesterase Inhibitors fo rthe Reversal of Neuromuscular Block. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2565–2568.

(31) Eldridge, M. D.; Murray, C. W.; Auton, R. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.

(32) Muegge, I.; Martin, Y. C. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. *J. Med. Chem.* **1999**, *42*, 791–804.

(33) Gohlke, H.; Hendlich, M.; Klebe, G. Predicting Binding Modes, Binding Affinities and "Hot Spots" for Protein–Ligand Complexes Using a Knowledge-Based Scoring Function. *Perspect. Drug Discovery Des.* **2000**, *20*, 115–144.

(34) Wang, T.; Wade, R. C. Comparative binding energy (COMBINE) analysis of OppA–peptide complexes to relate structure to binding thermodynamics. *J. Med. Chem.* **2002**, *45* (22), 4828–4837.

(35) Kmunicek, J.; Bohac, M.; Luengo, S.; Gago, F.; Wade, R. C.; Damborsky, J. Comparative binding energy analysis of haloalkane dehalogenase substrates: modelling of enzyme–substrate complexes by modelcular docking and quantum mechanical calculations. *J. Comput.-Aided Mol. Des.* **2003**, *17* (5–6), 299–311.

(36) Murcia, M.; Ortiz, A. R. Virtual screening with flexible docking and COMBINE-based models. Appliction to a series of factor Xa inhibitors. *J. Med. Chem.* **2004**, *47* (4), 805–820.

(37) *SIMCA-P*; Umereics AB, Umea, Sweden, 2001.

(38) Shafferman, A.; Velan, B.; Ordentlich, A.; Kronman, C.; Grosfeld, H.; Leitner, M.; Flashner, Y.; Cohen, S.; Barak, D.; Ariel, N. Substrate inhibition of acetylcholinesterase: residues affecting signal transduction from teh surface to the catalytic center. *EMBO J.* **1992**. *11*, 3561–3568.

(39) Johnson, J. L.; Cusack, B.; Hughes, T. F.; Mccullough, E. H.; Fauq, A.; Romanovskis, P. R.; Spatola, A. F.; Rosenberry, T. L. Inhibitors tethered near the acetylcholinesterase active site serve as molecular rulers of the peripheral and acylation sites. *J. Biol. Chem.* **2003**, *278*, 38948–38955.

(40) Ordentlich, A.; Barak, D.; Kronman, C.; Ariel, N.; Segall, Y.; Valan, B.; Shafferman, A. The architecture of human acetylcholinesterase active center probed by interactions with selected organophosphate inhibitors. *J. Biol. Chem.* **1996**, *271*, 11953–11962.

(41) Cramer, R. D., III; Petterson, D. E.; Flume, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.