

Modulation of Binding Strength in Several Classes of Active Site Inhibitors of Acetylcholinesterase Studied by Comparative Binding Energy Analysis

Sonsoles Martín-Santamaría,^{†,‡} Jordi Muñoz-Muriedas,^{†,§} F. Javier Luque,^{*,§} and Federico Gago^{*,†}

Departament de Físicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain, and Departamento de Farmacología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain

Received February 10, 2004

The comparative binding energy (COMBINE) methodology has been used to identify the key residues that modulate the inhibitory potencies of three structurally different classes of acetylcholinesterase inhibitors (tacrine, huprine, and dihydroquinazoline) targeting the catalytic active site of this enzyme. The extended set of energy descriptors and the partial least-squares methodology used by COMBINE analysis on a unique training set containing all the compounds yielded an interpretable model that was able to fit and predict the activities of the whole series of inhibitors reasonably well ($r^2 = 0.91$ and $q^2 = 0.76$, 4 principal components). A more robust model ($q^2 = 0.81$ and SDEP = 0.25, 3 principal components) was obtained when the same chemometric analysis was applied to the huprine set alone, but the method was unable to provide predictive models for the other two families when they were treated separately from the rest. This finding appears to indicate that the enrichment in chemical information brought about by the inclusion of different classes of compounds into a single training set can be beneficial when an internally consistent set of pharmacological data can be derived. The COMBINE model was externally validated when it was shown to predict the activity of an additional set of compounds that were not employed in model construction. Remarkably, the differences in inhibitory potency within the whole series were found to be finely tuned by the electrostatic contribution to the desolvation of the binding site and a network of secondary interactions established between the inhibitor and several protein residues that are distinct from those directly involved in the anchoring of the ligand. This information can now be used to advantage in the design of more potent inhibitors.

Introduction

Ligand–receptor recognition is a complex problem to deal with theoretically because the binding affinity of a ligand results from the subtle interplay of forces that takes place within the receptor binding site, usually in competition with water molecules.¹ However, when a series of ligands display graded affinity toward a given receptor, as is often the case in medicinal chemistry projects, and structural information about one or more of their complexes with the target (or a highly homologous) receptor is available, molecular modeling techniques can be used to gain insight into the key features that modulate the strength of ligand–receptor binding within the series.² Quantitative structure–activity relationships (QSAR) can then be derived using molecular mechanics force fields, typically including van der Waals and Coulombic interactions, or alternative scoring functions that evaluate the goodness of the fit.^{3–5} These structure-based QSAR methods, however, are not devoid of hurdles as not always do the calculated values accurately reflect the differences in experimental activities or binding affinities. When discrepancies are found, likely causes are inadequate parametrization, insufficient structural refinement, inappropriate modeling of the dielectric environment, and/or omission of desolva-

tion and entropic contributions. Entropy can seldom be directly taken into consideration as it is usually one single conformation that is used to represent both the bound and unbound states of the binding partners. Multiple conformational states can, of course, be considered, also in the presence of explicit solvent molecules, as is done in free energy perturbation^{6,7} and linear response methods,⁸ but the high computational cost makes these two approaches impractical for the comparative study of more than a few ligand–receptor complexes. On the other hand, it has been shown that cancellation of errors and enthalpy–entropy compensations are not uncommon, especially within congeneric series.⁹

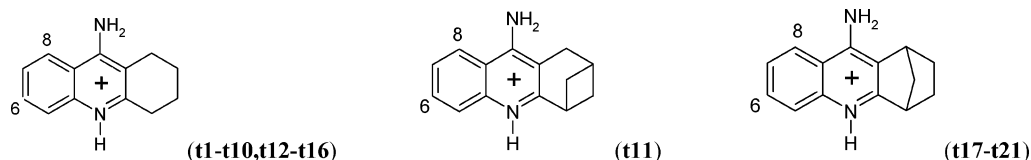
For QSAR studies of large series of ligand–receptor complexes, an alternative strategy has been proposed that is based on the partitioning of the calculated intermolecular binding energies into contributions from individual protein residues. A chemometric analysis is then performed that projects the resulting matrix of energy terms on to a small number of orthogonal “latent variables” (or principal components, PCs) that correlate with the activity differences using a partial least squares (PLS) analysis. At the end of the procedure, those pairwise interactions between the ligand and individual protein residues that are predictive of activity or binding free energy (i.e., the real “signal”) are selected and given weights, according to their importance in the model, in the form of PLS pseudocoefficients, whereas the “noise” present in the dataset is filtered out. This

* Address correspondence to either author. F.J.L.: tel, +34-934 024 557; fax, +34-934 035 987; e-mail, javier@far1.far.ub.es. F.G.: tel, +34-918 854 514; fax, +34-918 854 591; e-mail, federico.gago@uah.es.

[†] Both authors contributed equally to this work.

[‡] Universitat de Barcelona.

[§] Universidad de Alcalá.

Table 1. Chemical Structures and pIC₅₀ Data for Tacrine Derivatives

compound	substituents	pIC ₅₀		
		human ^a	human ^b	normalized ^c
t1 (tacrine)		7.47	6.60	6.69
t2	5-Cl	7.16		6.37
t3	6-Cl	8.74	8.00	8.09
t4 ^d	7-Cl	6.13	6.26	6.35
t5	8-Cl	8.03		7.24
t6	6-CH ₃		7.00	7.10
t7	7-CH ₃		5.09	5.18
t8	6-NO ₂		7.55	7.64
t9	7-NO ₂		5.52	5.60
t10	6-F		7.06	7.15
t11	8-F	7.54		6.76
t12	7-NH ₂		5.42	5.51
t13	6-OCH ₃		6.46	6.54
t14	6,7-diCl		6.33	6.41
t15	6,7-diOCH ₃		5.28	5.37
t16	7-OH	6.79		6.00
t17		7.50		6.72
t18	5-Cl	7.07		6.29
t19	6-Cl	8.35		7.56
t20	7-Cl	5.96		5.17
t21	8-Cl	7.37		6.59

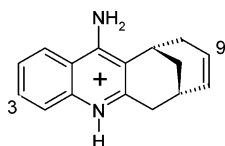
^a Reference 24. ^b Reference 25. ^c Experimental data were renormalized by adopting an IC₅₀ value for tacrine of 0.21 μM (ref 27). ^d Owing to the discrepancy between the IC₅₀ values observed for **t1**, **t3**, and **t4** (see text), the renormalized value was obtained by considering only the value determined spectrophotometrically (ref 25).

approach has been termed comparative binding energy (COMBINE) analysis, and, since its first implementation in 1995,¹⁰ several successful applications on different systems of medicinal chemistry^{11–15} and biochemical^{16–18} interest have been reported in the literature by a number of laboratories.

There are two clear advantages in using the COMBINE method for SAR studies: (i) it avoids the bias of a subjective, although expert, interpretation of structural data (often limited to one or a few complexes) that may, or may not, be sufficient to explain the differences in activity within a large series of ligands, and (ii) the computational cost is much lower than that required for *state-of-the-art* free energy calculations (on the other hand, only applicable to pairs of either compounds or targets). Nevertheless, the success of the COMBINE methodology, as well as of other SAR techniques, can be jeopardized by the quality of the experimental data, especially when they come from different sources. Ideally, *lead optimization* would have to rely on a homogeneous set of structural, chemical, and pharmacological data. In practice, however, medicinal chemists have to live with the uncertainties associated both with the use of different structural classes of compounds in the same project and with the fact that experimental activities may have been determined using different pharmacological assays in several laboratories. In an even less ideal scenario, they also may have to deal with 3D structures of drug–receptor complexes that are not exactly the same as those used in the experiments. In these circumstances, we asked ourselves whether it would still be possible to extract useful information to guide the optimization of lead compounds.

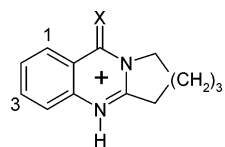
A case in point is acetylcholinesterase (AChE), whose blockade is currently the only clinical approach to the symptomatic treatment of Alzheimer's disease.¹⁹ The therapeutic impact of this enzyme has been reinforced by recent evidence suggesting an additional role in mediating the aggregation and deposition of the β-amyloid peptide.^{20,21} In AChE, the catalytic active site is at the end of a deep, narrow gorge (around 20 Å long) that penetrates into the enzyme and widens out close to its base.²² There is also a secondary binding locus, i.e., the peripheral site, which might act as an initial binding site for the incoming substrate. This structural complexity accounts for the large diversity of reversible AChE inhibitors,²³ which can interact with either the active site, the peripheral site, or both, making use of distinct sets of interactions.

Our first aim was to examine the benefits and disadvantages of including three different chemical classes of AChE inhibitors that target the active site (21 tacrine [Table 1],^{24,25} 7 huprine [Table 2],^{26,27} and 7 dihydroquinazoline [Table 3]²⁸ derivatives) into a single training set versus examining each family in turn. This sort of comparison is further complicated in this particular case by the fact that the experimental data (i.e., 50% inhibitory concentrations, IC₅₀) have been compiled from independent studies in which enzyme sources, assay conditions and techniques differed from one laboratory to another. Subject to this caveat, the obvious enrichment in chemical information upon inclusion of different compound families into a unique training set for SAR studies can be negatively counterbalanced by the “noise” inherent in the mixture of experimental data. Thus, we decided to check whether the COMBINE method is well suited to identify key

Table 2. Chemical Structures and pIC₅₀ Data for Huprine Derivatives

compound	substituents	pIC ₅₀		
		bovine ^a	human ^b	normalized ^b
x1	9-CH ₃	7.33		7.88
x2	3-F,9-CH ₃	8.46		9.01
x3	3-Cl,9-CH ₃	8.94	9.49	9.49
x4	9-C ₂ H ₅	7.56		8.12
x5	3-F,9-C ₂ H ₅	8.17	8.68	8.68
x6	3-CH ₃ ,9-C ₂ H ₅	8.35		8.89
x7	3-Cl,9-C ₂ H ₅	8.89	9.49	9.49

^a Reference 26. ^b Reference 27. ^c Experimental data were renormalized by using a factor derived from the comparison of inhibitory activities in bovine and human AChE for **x3**, **x5**, and **x7**.

Table 3. Chemical Structures and pIC₅₀ Data for Dihydroquinazoline Derivatives

compound	substituents	pIC ₅₀	
		human ^a	normalized ^b
j1	X = H ₂	5.43	4.59
j2	X = H ₂ ; 4-Cl	5.82	4.99
j3	X = H ₂ ; 3-Cl	6.80	6.01
j4	X = H ₂ ; 2-Cl	6.30	5.47
j5	X = H ₂ ; 1-Cl	5.77	4.94
j6	X = O; 3-Cl	<4.00	<3.17
j7	X = S; 3-Cl	<4.00	<3.17

^a Reference 28. ^b Experimental data were renormalized by adopting an IC₅₀ value for tacrine of 0.21 μM (ref 27).

ligand–receptor interactions also in a clearly unfavorable case involving a nonhomogeneous dataset. For the PLS analysis, we have assessed the effect of incorporating as regressor variables not only the residue-based ligand interactions but also the interaction energies between the inhibitors and some individual water molecules used in the modeling that are common to all the complexes. We have tested as well the effect of replacing the force field Coulombic electrostatic interactions with those calculated using a continuum method, as reported previously.^{9,11} In addition, and as an innovation over previously reported COMBINE analyses, some selected residue-based desolvation contributions were calculated and used to replace the global term that describes the desolvation of the whole receptor binding site upon ligand binding. Finally, a test set of another 7 tacrine-related compounds, taken from an independent study (Table 6)²⁹ and not included in model derivation, was used to assess the predictive ability of the best resulting COMBINE model.

Results and Discussion

Dataset Description. Since all AChE inhibitors included in this study possess the same tricyclic tetrahydroacridine-like structure, it is safe to assume that they share with tacrine a common mode of binding to

AChE.³⁰ In fact, the binding mode for huprine derivatives that was anticipated from molecular modeling studies^{7,26,27} and later confirmed by the X-ray crystal structure of the AChE–huprine X complex³¹ shows that the tetrahydro-9-aminoacridine rings of both tacrine and huprine X appear virtually superimposed when the Cα traces of the enzyme in both complexes are overlaid. Furthermore, chlorination at position 6 in tacrine and 1,4-methylenetacrine, or at the equivalent position 3 in dihydroquinazoline and huprine, gives rise in all cases to a significant increase in the inhibitory potencies (Tables 1–3).²³ This finding can be easily explained by accepting that, given the similarity in their binding modes, the chlorine atom occupies a common pocket, which mainly involves Trp432 and Met436, according to the crystallographic data for the AChE–huprine X complex.³¹ These considerations support our contention that the differences in inhibitory activity can be interpreted in terms of differential interactions of the ligands with specific residues in the binding pocket of our modeled complexes. Importantly, this choice of inhibitors is particularly well suited for our purposes, because the COMBINE results can be directly contrasted with high-quality information provided by both X-ray crystallographic structures^{30,31} and previous free energy calculations.^{7,26,27,32}

Experimental Data Normalization. IC₅₀ values for tacrine (Table 1) and dihydroquinazoline (Table 3) derivatives, as well as for the test set of tacrine-related compounds (Table 6), were derived from inhibition of human erythrocyte AChE. On the other hand, for huprines, the inhibitory activity was determined using enzyme from either bovine or human (**x3**, **x5**, and **x7**) erythrocytes. Moreover, the experiments were performed by using either radiometric³³ (for some tacrine derivatives²⁴ and the dihydroquinazolines²⁸) or spectrophotometric³⁴ (for some tacrine derivatives,²⁵ the huprines,^{26,27} and the test set²⁹) techniques. Nevertheless, the consistency of the data was supported by inspection of the inhibitory activity for tacrine, which was used as the reference compound in all of these studies. Thus, IC₅₀ values of 33.5²⁴ and 30²⁸ nM and of 0.25,²⁵ 0.21,²⁷ and 0.14²⁹ μM were reported for tacrine by using either the radiometric or the spectrophotometric assay, respectively. For our purposes, a reference IC₅₀ value of 0.21 μM was adopted for tacrine, and the IC₅₀ data for the rest of tacrine (including the test set) and dihydroquinazoline derivatives were renormalized by using factors derived from the ratio between the IC₅₀ values reported for tacrine in each independent study and the reference value. Let us note that this implies adding corrections (in pIC₅₀ units) of 0.08 (ref 25), −0.18 (ref 29), −0.80 (ref 24) and −0.84 (ref 28) to the respective pIC₅₀ values given in the independent studies. For huprines **x1**, **x2**, **x4**, and **x6**, the IC₅₀ values on bovine AChE were scaled by a factor of ~3.6 (0.6 pIC₅₀ units) that corresponds to the roughly constant ratio observed between the IC₅₀ data determined on bovine and human AChE for **x3**, **x5**, and **x7**.

It is worth stressing that the general applicability of the renormalization procedure described above is not completely guaranteed for all the compounds, especially for the less active inhibitors. This is particularly noticeable, for example, in the fact that whereas the IC₅₀

values determined from spectrophotometric²⁵ and radiometric²⁴ assays for tacrine (**t1**) and 6-chlorotacrine (**t3**) have a similar ratio of ~6, a factor close to unity is found for the IC₅₀ values of 7-chlorotacrine (**t4**), which is one of the least potent inhibitors (see Table 1). This discrepancy illustrates the difficulties encountered when trying to derive a *homogeneous* set of pharmacological data for QSAR studies from values reported independently by several groups. Rather inevitably, a compromise has to be made between accuracy and generality, but the question remains of whether or not the larger content of chemical information gained from treating an expanded dataset of related compounds, rather than separate chemical families in turn, can be advantageous in our attempts to identify the molecular determinants responsible for the differences in biological activity.

COMBINE Analyses. The evolution of the figures of merit as a function of the number of PCs extracted for the different models encompassing the whole set of 35 inhibitors can be found in Table 4. The overall quality of all the models is fairly good (with $q^2 \geq 0.6$ for dimensionalities between 4 and 5), but clear improvements over the initial model are apparent upon incorporation of some of the refinements. Thus, when the Coulombic electrostatic interactions between ligand and individual residues calculated with the AMBER force field (model A) were replaced with those calculated using the continuum method (model B), the quality of the model improved, but only marginally. Likewise, the explicit inclusion of the two water molecules common to all the complexes did not lead to a very significant improvement (cf. model F vs model A, and model G vs model B). However, incorporation of the electrostatic contributions to the desolvation of both ligand and receptor upon complex formation (model H) significantly improved the quality of both the fitting and the predictions ($r^2 = 0.91$ and $q^2 = 0.72$, respectively, for a 5-PC model). Moreover, when each of these terms was considered independently, an almost negligible contribution was apparent for the desolvation of the ligand alone (model C) whereas the dimensionality of the model was reduced and its quality improved even further ($r^2 = 0.91$ and $q^2 = 0.76$ for a 4-PC model) when only the desolvation of the receptor binding site (model D) was taken into account. Model C behaved well for most of the compounds but overpredicted the less active inhibitors **j6** and **j7**. These outliers were fitted better in model D such that the standard deviation of error in predictions (SDEP) for the whole set attained a remarkable value of 0.78 (Figure 1), with the activities of only two compounds, **j4** and **t18**, significantly differing in more than 1 order of magnitude between predicted and experimental.

Given the important role ascribed to the desolvation of the receptor binding site, we wondered whether it was a global property or whether it would be possible to pinpoint the most crucial residue(s) involved. To this end, the electrostatic contributions to the desolvation of a series of individual amino acids lining up the active site cavity were calculated ($\Delta G_{\text{desolv}}^{\text{residue}}$) and incorporated as additional descriptors in place of $\Delta G_{\text{desolv}}^{\text{R}}$. Strikingly, a marginal improvement or no improvement at all was observed when $\Delta G_{\text{desolv}}^{\text{residue}}$ values for residues Trp84, Gly117, Gly118, Glu199, Phe330, Tyr334, Trp432,

Table 4. Performance of Different COMBINE Models^a for the Whole Set of Inhibitors in Fitting and Prediction

data set ^b	no. of PCs	r^2	SDEC	q^2	SDEP
A	1	0.57	1.03	0.50	1.11
	2	0.70	0.86	0.59	1.00
	3	0.73	0.81	0.59	1.00
	4	0.76	0.76	0.60	0.99
	5	0.83	0.64	0.59	1.00
B	1	0.70	0.86	0.58	1.01
	2	0.79	0.71	0.59	1.00
	3	0.82	0.66	0.61	0.98
	4	0.86	0.58	0.64	0.93
	5	0.89	0.51	0.62	0.97
C	1	0.71	0.85	0.56	1.04
	2	0.79	0.72	0.55	1.05
	3	0.82	0.66	0.55	1.05
	4	0.84	0.62	0.59	1.00
	5	0.86	0.58	0.61	0.97
D	1	0.62	0.97	0.48	1.12
	2	0.72	0.82	0.55	1.05
	3	0.85	0.60	0.64	0.93
	4	0.91	0.47	0.76	0.78
	5	0.92	0.44	0.75	0.79
E	1	0.70	0.85	0.57	1.03
	2	0.81	0.68	0.61	0.98
	3	0.84	0.62	0.64	0.94
	4	0.88	0.54	0.68	0.89
	5	0.90	0.50	0.65	0.92
F	1	0.59	1.00	0.51	1.09
	2	0.73	0.82	0.63	0.95
	3	0.76	0.76	0.62	0.97
	4	0.82	0.67	0.62	0.96
	5	0.85	0.61	0.63	0.95
G	1	0.72	0.83	0.61	0.98
	2	0.77	0.74	0.60	0.99
	3	0.82	0.66	0.57	1.02
	4	0.86	0.58	0.56	1.04
	5	0.88	0.53	0.56	1.04
H	1	0.63	0.95	0.47	1.12
	2	0.73	0.82	0.51	1.06
	3	0.83	0.65	0.58	0.96
	4	0.87	0.56	0.66	0.87
	5	0.91	0.48	0.72	0.75

^a Abbreviations: PC, principal component; r^2 , correlation coefficient; SDEC, standard deviation of errors in correlation; q^2 , predictive correlation coefficient; SDEP, standard deviation of errors in prediction. ^b Models include the following variables: A, AMBER van der Waals and electrostatic interactions; B, AMBER van der Waals and DelPhi electrostatic interactions; C, AMBER van der Waals, DelPhi electrostatic interactions, and $\Delta G_{\text{desolv}}^{\text{L}}$; D, AMBER van der Waals, DelPhi electrostatic interactions and $\Delta G_{\text{desolv}}^{\text{R}}$; E, AMBER van der Waals, DelPhi electrostatic interactions, and $\Delta G_{\text{desolv}}^{\text{Tyr442}}$; F, AMBER van der Waals and electrostatic interactions including two water molecules;⁵⁸ G, AMBER van der Waals and DelPhi electrostatic interactions including two water molecules;⁵⁸ H, AMBER van der Waals, DelPhi electrostatic interactions, $\Delta G_{\text{desolv}}^{\text{L}}$, and $\Delta G_{\text{desolv}}^{\text{R}}$. The values in bold highlight the best quality model.

and Ile439 were considered (data not shown), and only when desolvation of Tyr442 was incorporated as an addition to model B ($\Delta G_{\text{desolv}}^{\text{Tyr442}}$, model E) did the SDEP value decrease from 0.93 to 0.89 (Table 4). In view of these results, it is clear that, at least in this particular case, no real benefit is obtained by increasing the computational cost of calculating and incorporating these individual contributions in place of the overall desolvation of the binding site (model D).

For the rest of the discussion we will concentrate on the best predictive model (model D), which contains the

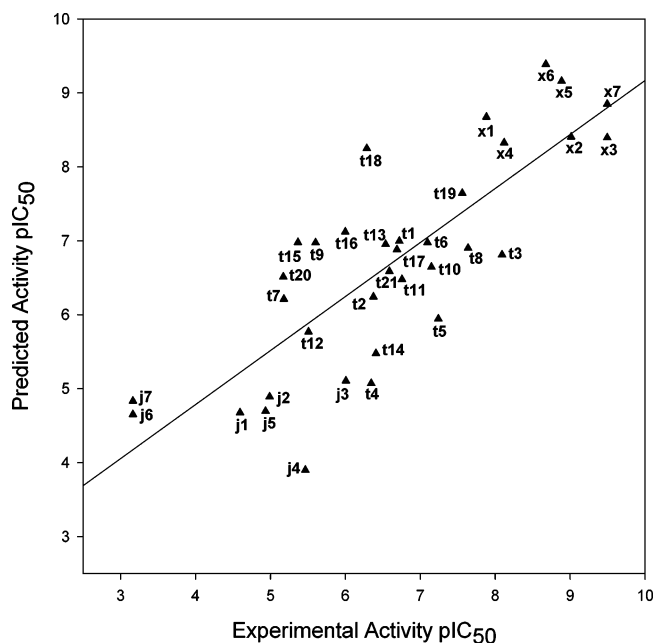


Figure 1. Scatterplot comparing experimental vs predicted activities in COMBINE model D for the 35 compounds of the training series belonging to the three families studied.

van der Waals interactions from AMBER, the electrostatic terms from DelPhi, and $\Delta G_{\text{desolv}}^{\text{R}}$.

To investigate the distribution of the 35 complexes in the space defined by their ligand–receptor interaction energies, it is useful to focus on the results of the principal component analysis (PCA). The essential data patterns can be easily visualized by plotting the complexes in the space defined by the first and second PC (*score plot*), whereas the relation between the original variables and the new orthogonal latent variables can be unveiled by plotting the contributions of the calculated energy descriptors to each of these PCs (*loading plot*). The first PC extracted, which is mostly made up by the electrostatic contribution to the desolvation of the receptor binding site and van der Waals interactions involving Gly118, Tyr121, Phe331 and His440, is enough to clearly distinguish the huprine family from the other two (Figure 2). This can be rationalized in terms of the extra space that is occupied by the carbobicyclic moiety of huprines, which fills a hydrophobic cavity³¹ lined by residues Tyr121, Phe290, Phe330, and Phe331 (Figure 3). The second PC, with major contributions from the electrostatic interactions with Glu199 and Asp72, and van der Waals interactions with Trp84, is able to separate dihydroquinazoline-based inhibitors (with the exception of **j2**) from the tacrines.

A quantitative and rapid assessment of the relevance of the different ligand–residue interactions to account for the differences in activity within the whole series of AChE inhibitors studied is provided by the normalized PLS coefficients shown in Figure 4. The signs of these coefficients express if favorable van der Waals and electrostatic interactions (negative energy values) correlate with an increase (negative coefficient) or a decrease (positive coefficient) in activity (the higher the pIC_{50} , the greater the potency). The negative PLS coefficients for Tyr121, Ser122, Phe290, Phe331, Tyr334, Trp432, Ile439, and Tyr442 indicate that favorable van der Waals interactions with these residues are beneficial

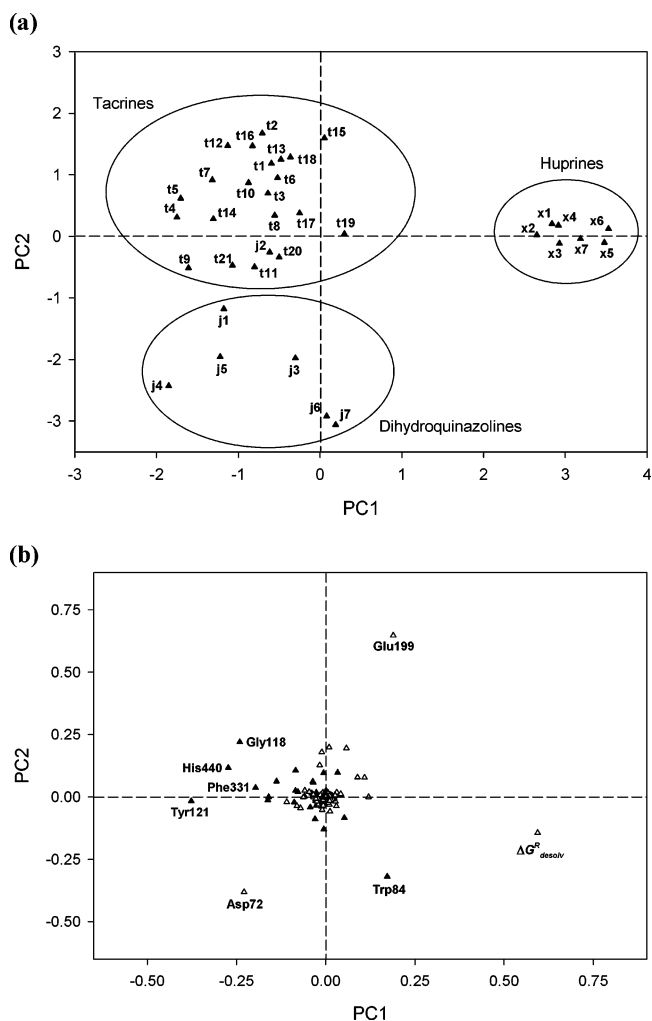


Figure 2. (a) Score plot and (b) loading plot (▲, van der Waals variables; △, electrostatic variables) of the first (PC1) and the second (PC2) principal components for COMBINE model D. Relevant energy descriptors have been labeled.

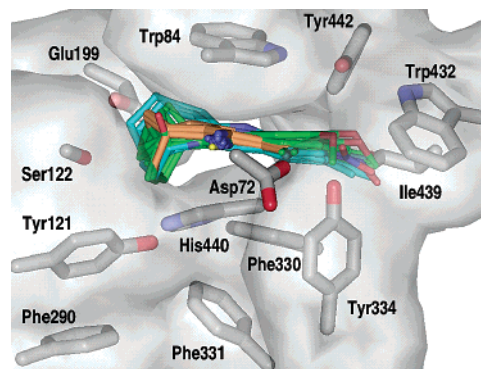


Figure 3. Superposition of the training set of inhibitors as found in their respective complexes with AChE. Relevant active site residues for one representative protein have been labeled. A semitransparent solvent-accessible surface envelops the side chains of the labeled residues (except Asp72) to delineate the active site cavity. Water molecules and all the hydrogen atoms have been omitted for clarity. Carbon atoms of tacrines, huprines, and dihydroquinazolines are shown in green, orange, and cyan, respectively.

for activity. These residues form hydrophobic pockets (Figure 3) that accommodate the substituents at position C6 of tacrine (Ile439, Trp432) and the methyl or ethyl group of huprines (Tyr121, Phe290, and Phe331),

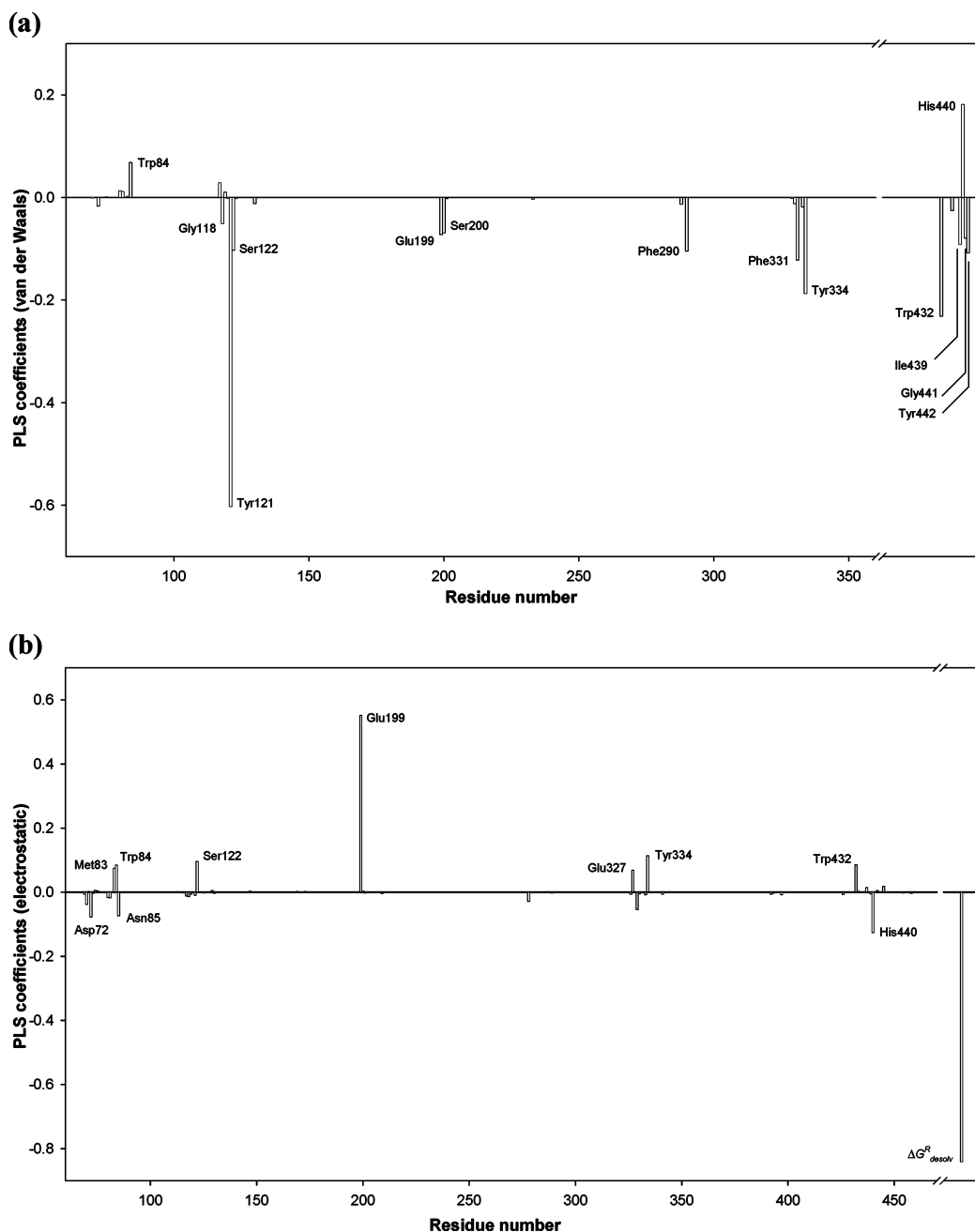


Figure 4. Normalized PLS coefficients of model D for each of the (a) van der Waals and (b) electrostatic interaction energies studied. Relevant residues for which the coefficients have absolute values ≥ 0.04 have been labeled.

or are involved in water-mediated contacts with the inhibitor (Tyr121 and Ser122).^{30,31} Tyr442, the desolvation of which was found to contribute significantly to model E (see above), is located close to the entrance of the hydrophobic pocket that accommodates the substituent attached to C6, and Tyr334 is fixed in position through a hydrogen bond between its hydroxyl group and the carboxylate of Asp72. Of all these ligand interactions, that involving Tyr121 appears to be the most discriminatory for activity when the three inhibitor families are considered together.

The positive PLS coefficients for His440 and Trp84, on the other hand, highlight the fact that favorable van der Waals interactions with these two residues are inversely correlated with activity. This can be striking at first sight, as the magnitude of the van der Waals interaction between the ligands and Trp84 is the great-

est of all, closely followed by Phe330 (Supporting Information, Figure 1a), but it must be borne in mind that the PLS analysis is seeking variables that can provide effective discrimination between weak and tight binders, and they do not need to be those with the greater absolute values. In fact, Phe330 (the second largest) is not even selected, meaning that differences in van der Waals interaction energies involving this residue cannot be used by the chemometric method to correlate with the differences in inhibitory potency. Thus, this interaction, although undoubtedly important for binding, makes up a relatively constant contribution within the series. Likewise, the counterintuitive sign assigned to the relatively small Trp84 coefficient simply means that the differences observed within the series for this interaction, its importance notwithstanding, tend to have an inverse relationship with activity.

With respect to the electrostatic block, it is dominated by the desolvation of the receptor binding site (the larger the cost of the desolvation, the more detrimental for activity) and by the interaction with Glu199 (Figure 4), precisely the two variables mainly used by PCA to discriminate between the three families and those in the electrostatic block showing the largest absolute magnitude and the greatest variation (Supporting Information, Figure 1b). It is noteworthy that the sign given here to the electrostatic interaction with His440 is negative whereas it was positive in the van der Waals block. Our interpretation, in agreement with findings in other systems,³⁵ is that both terms are inversely related in the AMBER force field when there exists a hydrogen bond, as occurs here between the carbonyl group of His440 and the NH group of the inhibitors (Figure 3). Therefore, the better this hydrogen bond between the ligand and His440, the better for activity. On the other hand, the positive coefficient assigned by the PLS model to the electrostatic interaction between the positively charged ligands and Glu199 is being used to effectively discriminate against the dihydroquinazolines, in agreement with the PCA loading plot (Figure 2b). Since the three classes of molecules are positioned in the field created by the negative charges of both Asp72 and Glu199 (each on a different side of the binding cavity, as shown in Figure 3), the electrostatic interactions with these two residues are balanced. The more potent huprines and tacrines interact more strongly with the carboxylate group of Asp72, through a water-mediated hydrogen bond, by virtue of the amino group attached to their quinoline ring. Lacking this amino group, the less potent dihydroquinazolines, which also present a different charge distribution on their quinazoline ring system, interact more weakly with Asp72 but strengthen their interaction with Glu199 at the expense of a larger desolvation penalty.

It is interesting to note that the study of the tacrine and the dihydroquinazoline families on their own led in both cases to models with no predictive ability (negative q^2 values) whereas utilization of only the huprines subset yielded a relatively modest initial model ($q^2 = 0.4$), the predictive ability of which was very much improved ($q^2 = 0.81$, SDEP = 0.25 for a 3-PC model) upon incorporation of the term representing desolvation of the receptor binding site (Figure 5). Comparison of the most significant PLS coefficients for this subset (Table 5) with those derived for the whole set is interesting as it reveals that both COMBINE models have in common $\Delta G_{\text{desolv}}^R$, the van der Waals interaction with Trp432 and Ile439, and the electrostatic interaction with Glu199, albeit the relative importance of each of these contributions varies. For the huprines, additional interactions, especially with Phe330 but also with Gly119, are used to fine-tune the binding affinity. Remarkably, a role for both Phe330 and Glu199 in the binding of huprine derivatives has already been proposed on the basis of free energy calculations.⁷ Although the biochemical role of Glu199 in catalysis is well established,^{36–38} interaction with this residue does not appear to have been exploited yet in the design of more potent tacrine-like reversible inhibitors. In this respect, it is worth recalling that the carboxylate of Glu199 establishes a direct hydrogen bond with a hydroxyl

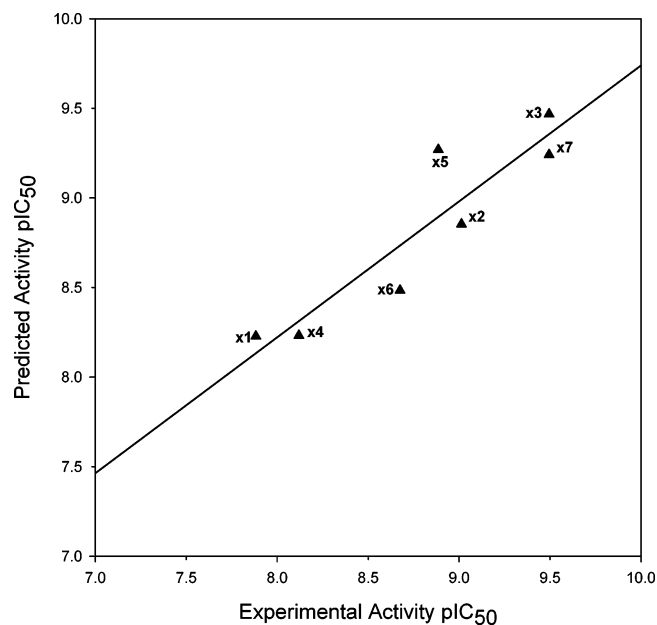


Figure 5. Scatter plot comparing experimental vs predicted activities in the COMBINE model derived for the huprine series alone.

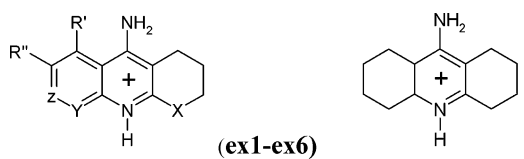
Table 5. Normalized PLS Coefficients^a for the COMBINE Model Derived for the Huprine Family of AChE Inhibitors

residue	van der Waals	electrostatic
Gly119	0.056	
Glu199		-0.067
Phe330	-0.118	
Trp432	-0.261	0.042
Ile439	-0.040	
$\Delta G_{\text{desolv}}^R$		-0.250

^a Only PLS coefficients with absolute values ≥ 0.04 are shown.

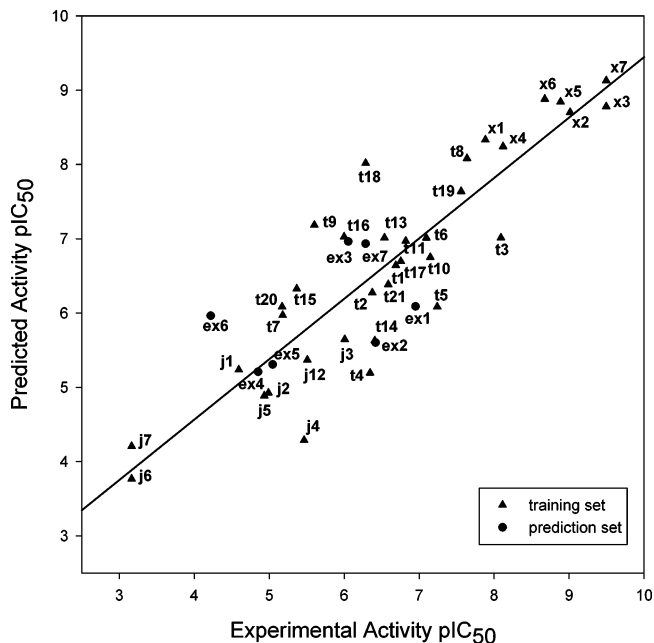
group of the reversible AChE inhibitor galanthamine,³⁹ as well as a water-mediated hydrogen bond with the pyridine nitrogens of (-)-huperzine A,⁴⁰ (-)-huperzine B,⁴¹ and two alkylene-linked dimers of hupryridone.⁴²

The enormous differences in model quality found for the three families studied most likely reflect the greater reliability of the experimental data for huprines, which are the most active inhibitors, with activity data taken from a single source, as well as the larger uncertainties associated with the normalized inhibition data for the other compounds, particularly the tacrine derivatives. In fact, when the tacrine family was split into two subsets, according to the experimental method employed to evaluate their inhibitory activity,^{24,25} the q^2 value of each subset increased ($q^2 \sim 0.3$), which we take as a clear indication of improvement in the signal-to-noise ratio. Also, when huprines and dihydroquinazolines were considered as the training set, a good model was obtained ($q^2 = 0.80$) that could be used to predict the activity of the tacrines with reasonable accuracy. In this case, the external SDEP of 1.03 compared quite well with the cross-validated SDEP of 0.92 obtained for the other two families in the training set. All of these results support the assumed binding mode for the dihydroquinazolines and reveal the advantages of including as large a description as possible of both the chemical space and the energy landscape around the binding site. Thus, the extended set of energy descriptors provided by the three structurally diverse families help to build a

Table 6. Chemical Structures and pIC₅₀ Data for the AChE Inhibitors in the External Validation Set


compound	substituents	pIC ₅₀	
		human ^a	normalized ^b
ex1	X = CH ₂ , Y = Z = CH, R' = F, R'' = H	7.10	6.95
ex2	X = O, Y = Z = CH, R' = R'' = H	6.58	6.42
ex3	X = CH ₂ , Y = Z = CH, R' = OCH ₃ , R'' = H	6.21	6.05
ex4	X = O, Y = Z = CH, R' = H, R'' = Cl	5.00	4.85
ex5	X = O, Y = N, Z = CH, R' = R'' = H	5.20	5.05
ex6	X = O, Y = CH, Z = N, R' = R'' = H	4.38	4.22
ex7		6.44	6.29

^aReference 29. ^bExperimental data were renormalized by adopting an IC₅₀ value for tacrine of 0.205 μM (ref 27).

**Figure 6.** Plot of recalculated versus experimental activities for the training set (▲) showing the predictions for the compounds in the test set (●, series ex, Table 6) using COMBINE model D.

comprehensive model that is able to predict the activity differences among the whole series of compounds with relatively high accuracy.

As a final test of the goodness of the derived COMBINE model with 4 PCs, we tried to predict the activity of an external set consisting of seven 9-amino-1,2,3,4-tetrahydroacridine derivatives²⁹ that were originally designed to explore brain uptake (Table 6). The quality of the predictions was reasonably good (Figure 6), as assessed by an SDEP of 0.93 log unit, which is only slightly larger than the value of 0.78 obtained for the training set (Table 4). The fact that inhibitory potencies can be predicted with an error of less than 1 order of magnitude highlights the possibility of extending the

validity of the model to other classes of ligands provided they bind in the same region that has been explored by the training series.¹³

Conclusions

Exploitation of pharmacological data generated independently by various research groups is a desired goal in medicinal chemistry projects but is often hampered by the logical uncertainties associated with the use of different assays and/or experimental conditions. This is a particularly worrying limitation when experimental activities have to be assigned in a reliable and consistent way to several chemical series that need to be included in a unique training set for chemometric analysis. On these premises, the results presented here for three different classes of AChE inhibitors demonstrate that the signal/noise ratio can be increased relative to the individual series provided the experimental data are “normalized” and a suitable chemometric tool is employed. Indeed, the extended set of energy descriptors and the PLS analysis implemented in a COMBINE framework enabled us to build a comprehensive QSAR model that was able to predict the inhibitory potency not only of a whole series of AChE inhibitors but also of an external validation set with reasonable accuracy. Of the three families of AChE inhibitors studied, the more active huprines were found to provide the more consistent results as a COMBINE model derived for this family alone and consisting of only 3 PC was highly predictive in cross-validation ($q^2 = 0.81$, SDEP = 0.25).

Although other existing methods (e.g., free energy perturbation [FEP] or thermodynamics integration)⁷ are more useful for determining highly accurate differences in binding free energy, there are some clear advantages in using a method such as COMBINE analysis to exploit the information contained in a large number of crystallographic or modeled 3D structures of related drug–receptor complexes: (i) the global picture it provides encompasses the whole of the binding site explored by the different ligands; (ii) the computational cost is much lower than that required for state-of-the-art free energy calculations, on the other hand only applicable to pairs of either compounds or targets; and (iii) it supplements the subjective interpretation of structural data,³⁵ allowing researchers to gain quantitative or semiquantitative insight into the key role played by specific ligand–receptor interactions and/or desolvation components. As a result, those residue-based van der Waals and electrostatic contributions that are endowed with a higher discriminatory ability can be identified, which provides clues for further chemical modification throughout the series.

In the present example, most of the residues that are more strongly involved in the binding of these ligands (e.g., Trp84, Phe330, Asp72) are found not to be the most relevant to explain the differences in activity, thus indicating that their contribution to the binding strength of the ligands is relatively constant within the series. In contrast, differences in the cost of desolvating the receptor binding site, as well as in the electrostatic interaction with Glu199 and van der Waals interactions with Tyr121 and Trp432, account for most of the variation in potency. The model then points to a network of secondary interactions between the inhibitor and

several well-defined residues, different from those usually considered for ligand anchoring and not immediately obvious from visual inspection alone, as important for fine-tuning the inhibitory potency. This information can now be used to advantage in the design of new AChE inhibitors.

Methods

Building and Refinement of the Inhibitors. All 42 AChE inhibitors, including the 21 tacrine derivatives (Table 1), 7 huprine X analogues (Table 2), and 7 dihydroquinazoline compounds (Table 3) used for model derivation, as well as the set of 7 related molecules (Table 6) used for external validation, were model-built in Insight-II⁴³ utilizing standard bond lengths and angles. The geometry of each inhibitor was then fully optimized at the Hartree–Fock level with the 6-31G(d) basis set⁴⁴ using the Gaussian-98 program.⁴⁵ According to the basicity of the aminoquinoline ring,⁴⁶ the protonated species was considered in all cases. Charge distributions for the inhibitors were calculated by fitting the HF/6-31G(d) electrostatic potential to atom centers using the RESP procedure,⁴⁷ and van der Waals parameters for inhibitor atoms were transferred from those defined for related atoms in the AMBER-95 all-atom force field⁴⁸ (parm98), which was also used for protein atoms.

Choice of X-ray Structure and Modeling of the Complexes. Although inhibitory potencies have been determined using human erythrocyte AChE for most of the compounds studied (the exceptions being huprines **x1**, **x2**, **x4**, and **x6**; see text), no crystal structure is available of human AChE (hAChE) in complex with a tacrine-related analogue. The Protein Data Bank does contain a crystal structure of a hAChE-fasciculin complex (1B41),⁴⁹ with this latter toxin bound at the outermost region of the gorge, and several structures of mouse AChE complexed to decamethonium (1MAA) or other inhibitors that interact at the peripheral binding site (1J07, 1N5M, 1N5R, 1MAH), as well as an apo form (1J06).⁵⁰ This means that, at least in principle, none of these mammalian AChE crystal structures appears to be the best option to model the binding of inhibitors that target the catalytic binding pocket which is located at the end of the gorge. On the other hand, the homologous enzyme from *Torpedo californica* (TcAChE) has been cocrystallized and solved with tacrine and huprine X bound at the catalytic site (PDB entries 1ACJ³⁰ and 1E66,³¹ respectively). TcAChE is highly homologous to the mammalian enzymes (Supporting Information, Figure 2). In fact, since the only difference in the active site region is a Phe330 → Tyr substitution, most molecular modeling studies of ligand–AChE complexes have focused on TcAChE itself (even in cases where experimental activities have been determined using AChE from organisms in which this residue is actually a Tyr, as in references 51 and 52) or have simply replaced Phe330 with Tyr (e.g., references 25 and 53). In the absence of direct crystallographic information, it does indeed seem reasonable to assume that, in the complexes of the mammalian enzymes with tacrine or huprine X, Tyr337 will be arranged similarly to Phe330 of TcAChE for the series of tacrine-like inhibitors examined here. This assumption is supported by the following pieces of evidence:

(a) The side chain conformation of Phe330 in TcAChE largely depends on the specific inhibitor being present in the complex. Thus, the χ_1 torsion angle is around -170° in the apo enzyme or in the complexes with either huperzine A or edrophonium, -120° in the complexes with either E2020 or decamethonium, -140° in the complexes with bivalent ligands related to huperzine A, and finally $+160^\circ$ in the complexes with either tacrine or huprine X.^{23,30,31,54}

(b) The side chain of Tyr337 is endowed with similar flexibility showing χ_1 angles of $+167^\circ$ in the apo form, $+178^\circ$ in its complex with decamethonium, and -170° in the complexes with inhibitors bound at the peripheral binding site.⁵⁵ It has also been noted that, as a result of side chain rotations, the hydroxyl group experiences positional deviations of up to

2 Å, and does not appear to be involved in specific, strong hydrogen bonds with other residues, as the minimal distance separating this phenolic oxygen and that of Tyr341 is between 3.5 Å (1MAA) and 3.8 Å (1MAH),⁵⁵ surely longer than that expected for an optimal O···H–O interaction.

(c) Since tacrine, 1,4-methylenetacrine, dihydroquinazoline, and huprine derivatives are structurally similar, it seems safe to assume that they adopt a similar binding mode (experimentally confirmed for tacrine and huprine X, as discussed above) and also that the conformation of Tyr337 will be similar in all cases. By adopting a conformation similar to that of Phe330 in the TcAChE complexes, 1ACJ and 1E66, the aromatic ring of Tyr337 might give rise to a cation– π interaction with the positive charge of the protonated inhibitors. Such an interaction has indeed been shown to be completely stable in previous nanosecond-long molecular dynamics (MD) simulations.^{7,27} Moreover, extension of these MD trajectories up to 5 ns (F.J.L., unpublished work) has revealed no relevant changes in the stacking interactions of both tacrine and huprine X with Trp and Phe (or Tyr).

(f) With the side chain in this conformation, the hydroxyl group does not form specific interactions with either tacrine or huprine X derivatives, as noted in several MD simulations.^{7,27} Thus, Tyr and Phe at position 330 can be safely assumed to behave very similarly in all the complexes studied. A completely different situation can be expected, of course, for the complexes with other inhibitors, such as (–)-huperzine A, since not only does Tyr337 adopt a different conformation⁴⁰ but also its hydroxyl group probably gets engaged in a hydrogen bond with the ammonium group of the inhibitor.³¹

(g) Free energy calculations performed for inhibitors bound to the TcAChE enzyme (with either Phe or Tyr at position 330) satisfactorily predicted the changes in binding affinities for mammalian AChE due to chemical modifications on huprine derivatives.^{7,27}

For the above reasons, the TcAChE structure was chosen for the modeling work as it provided the finest details of both ligand–receptor and ligand–water interactions that were thought most relevant for the series of compounds studied. TcAChE was modeled in its physiologically active form, as described before,⁷ with neutral His440 and deprotonated Glu327 which, together with Ser200, form the catalytic triad. The standard ionization state at neutral pH was considered for the rest of the ionizable residues with the exceptions of Asp392 and Glu443, which were neutral, and His471, which was protonated, according to previous numerical titration studies.³²

Tacrine and dihydroquinazoline derivatives were positioned in the active site on the basis of the crystallographic structures of the complexes of TcAChE with tacrine (PDB entry 1ACJ) whereas the complex with huprine X (PDB entry 1E66) provided the docking orientation for the huprine derivatives. The missing atoms in the original PDB files were built in using the program Insight II.⁴³ Each modeled complex was then hydrated by centering a sphere of 26 Å radius of TIP3P water molecules⁵⁶ around the inhibitor.

Energy Refinement of the Complexes. Each TcAChE–inhibitor complex was energy-minimized in a sequential way. First, the hydrogen positions were refined for 1000 steps of steepest descent energy minimization. Then, the water molecules were allowed to reorientate in the electric field of the complex for a further 5000 steps of steepest descent. Next, those residues in the active site (i.e., having an atom at less than 4 Å from any atom of the inhibitor) were relaxed for 2000 steps of steepest descent and 3000 steps of conjugate gradient. Finally, the whole system was optimized for 2000 steps of steepest descent and 3000 steps of conjugate gradient energy minimization. This rather conservative minimization protocol was deemed sufficient to take into account the minor conformational adjustments reported on complex formation without causing artifactual distortions in the protein.

Breakdown of the Intermolecular Interaction Energy and Pretreatment of the Resulting Energy Matrix. Individual residue contributions to the calculated ligand–

receptor interaction energies in the refined complexes were obtained by using the ANAL module in AMBER and including all atom pairs in the calculation. A uniform dielectric of 1 was used in the evaluation of electrostatic interactions. Each inhibitor was regarded as a single fragment, and no intramolecular energy terms were considered. Since there are 537 amino acids in the protein and 2 energy contributions (van der Waals and electrostatic) are considered for each residue, 1074 variables were used to characterize each complex. These energy descriptors made up the matrix for the chemometrics program Q².⁵⁷ No scaling or variable selection was carried out except for a mild pretreatment that consisted of zeroing all the variables with absolute values lower than 0.01 kcal mol⁻¹ and by removing those variables with a standard deviation below 0.01 kcal mol⁻¹. This procedure reduced the number of energy descriptors that entered the PLS analysis to around 200. For the calculations including two water molecules,⁵⁸ these were considered as two more receptor residues thereby yielding 4 additional variables.

The optimal dimensionality of the PLS models was determined by monitoring the cross-validation indexes as a function of the number of principal components (PCs) extracted. For cross-validation, the compounds were assigned randomly to any of five groups of approximately the same size, and the whole procedure was repeated 20 times. The predictive ability of the resulting models is reported by both the cross-validated correlation coefficient (q^2) and the standard deviation of error in predictions (SDEP). This cross-validation procedure is more rigorous than the more widely employed leave-one-out method and also provides more conservative results: a smaller q^2 and a higher SDEP.

For external validation, the PLS models obtained were used to predict the biological activity of seven related analogues (prediction set) not included in the initial training set. The complexes were modeled and refined as described above.

Continuum Electrostatics Calculations. Finite difference solutions to the linearized Poisson–Boltzmann equation,⁵⁹ as implemented in the DelPhi module of Insight II,⁴³ were used to describe the electrostatic effects of ligand binding in a format appropriate for use in COMBINE analysis, as reported previously.¹¹ In brief, following the classical approach,⁹ the change in electrostatic free energy on molecular association (ΔG_{ele}) was split into three separate components: (i) the ligand–receptor interaction energy in the presence of the surrounding solvent ($E_{\text{ele}}^{\text{LR}}$), (ii) the change in solvation energy of the ligand upon binding ($\Delta G_{\text{desolv}}^{\text{L}}$), and (iii) the change in solvation energy of the receptor upon binding ($\Delta G_{\text{desolv}}^{\text{R}}$):

$$\Delta G_{\text{ele}} = E_{\text{ele}}^{\text{LR}} + (\Delta G_{\text{desolv}}^{\text{L}} + \Delta G_{\text{desolv}}^{\text{R}})$$

The latter two terms were calculated by considering the effects on the respective electrostatic free energies of replacing the high dielectric medium of the solvent with the low dielectric medium of the other molecule in those regions that are occupied by the binding partner in the complex. The first term was calculated for each complex by computing the solvent-corrected potential generated by the charges on the ligand at the positions of each of the uncharged atoms of the receptor. To calculate the electrostatic contribution to the desolvation of just one protein residue, charges were assigned only to the atoms belonging to that residue in both the protein and the protein–ligand complex, as reported previously.⁶⁰ Selected residues for which this contribution was calculated were Trp84, Gly117, Gly118, Glu199, Phe330, Tyr334, Trp432, Ile439, and Tyr442.

The atomic coordinates employed were those of the AMBER-optimized complexes once the solvent molecules used in the energy refinement were removed (except when the two water molecules were explicitly considered in the analysis). A dielectric of 4 was chosen for the interior of the protein, the ligands, and the complexes whereas the surrounding solvent was assigned a dielectric of 80 with ionic strength of 0.145 M. Cubic grids with a resolution of 0.75 Å were centered on the molecular systems considered, and the charges were distrib-

uted onto the grid points.⁶¹ Solvent-accessible surfaces, calculated with a spherical probe of 1.4 Å radius, defined the solute boundaries, and a minimum separation of 5 Å was left between any solute atom and the borders of the box. The potentials at the grid points delimiting the box were calculated analytically by treating each charge atom as a Debye–Hückel sphere.

Acknowledgment. J.M.-M. is the recipient of a research fellowship from the Catalan Departament d'Universitats, Recerca i Societat de la Informació. Financial support from Neuropharma S.A. (Tres Cantos, Madrid) and the Spanish Ministerio de Ciencia y Tecnología (Grants SAF2002-04282 and GEN2001-4758 to F.J.L.) is gratefully acknowledged. S.M.-S. and F.G. also thank the National Foundation for Cancer Research for partial support of this research.

Supporting Information Available: Two figures containing the distribution of energy variables (van der Waals and electrostatic) for all the complexes in the training set (Figure 1) and the AChE sequence alignment (Figure 2) (PDF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- Ajay; Murcko, M. A. Computational methods to predict binding free energy in ligand–receptor complexes. *J. Med. Chem.* **1995**, *38*, 4953–4967.
- Apostolakis, J.; Caflisch, A. Computational ligand design. *Comb. Chem. High Throughput Screening* **1999**, *2*, 91–104.
- Marrone, T. J.; Luty, B. A.; Rose, P. W. Discovering high-affinity ligands from the computationally predicted structures and affinities of small molecules bound to a target: A virtual screening approach. *Perspect. Drug Discovery Des.* **2000**, *20*, 209–230.
- Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein–ligand interactions. *J. Mol. Biol.* **2000**, *295*, 337–356.
- Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11–26.
- Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
- Barril, X.; Orozco, M.; Luque, F. J. Predicting relative binding free energies of tacrine–huperzine A hybrids as inhibitors of acetylcholinesterase. *J. Med. Chem.* **1999**, *42*, 5110–5119.
- Hansson, T.; Marelus, J.; Åqvist, J. Ligand binding affinity prediction by linear interaction energy methods. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27–35.
- Checa, A.; Ortiz, A. R.; de Pascual-Teresa, B.; Gago, F. Assessment of solvation effects on calculated binding affinity differences: Trypsin inhibition by flavonoids as a model system for congeneric series. *J. Med. Chem.* **1997**, *40*, 4136–4145.
- Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. Prediction of drug binding affinities by comparative binding energy analysis. *J. Med. Chem.* **1995**, *38*, 2681–2691.
- Pérez, C.; Pastor, M.; Ortiz, A. R.; Gago, F. Comparative binding energy analysis of HIV-1 protease inhibitors: Incorporation of solvent effects and validation as a powerful tool in receptor-based drug design. *J. Med. Chem.* **1998**, *41*, 836–852.
- Lozano, J. J.; Pastor, M.; Cruciani, G.; Gaedt, K.; Centeno, N. B.; Gago, F.; Sanz, F. 3D-QSAR methods on the basis of ligand–receptor complexes. Application of COMBINE and GRID/GOLPE methodologies to a series of CYP1A2 ligands. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 341–353.
- Cuevas, C.; Pastor, M.; Pérez, C.; Gago, F. Comparative binding energy (COMBINE) analysis of human neutrophil elastase inhibition by pyridone-containing trifluoromethylketones. *Comb. Chem. High Throughput Screening* **2001**, *4*, 627–642.
- Wang, T.; Wade, R. C. Comparative binding energy (COMBINE) analysis of influenza neuraminidase–inhibitor complexes. *J. Med. Chem.* **2001**, *44*, 961–971.
- Murcia, M.; Ortiz, A. R. Virtual screening with flexible docking and COMBINE-based models. Application to a series of factor Xa inhibitors. *J. Med. Chem.* **2004**, *47*, 805–820.
- Tomic, S.; Nilsson, L.; Wade, R. C. Nuclear receptor–DNA binding specificity: A COMBINE and Free–Wilson QSAR analysis. *J. Med. Chem.* **2000**, *43*, 1780–1792.

- (17) Kmuniček, J.; Luengo, S.; Gago, F.; Ramírez Ortiz, A.; Wade, R.; Damborský, J. Comparative Binding Energy (COMBINE) analysis of the substrate specificity of haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10. *Biochemistry* **2001**, *40*, 8905–8917.
- (18) Kmuniček, J.; Boháč, M.; Luengo, S.; Gago, F.; Wade, R. C.; Damborský, J. Comparative Binding Energy analysis of haloalkane dehalogenase substrates: Modelling of enzyme substrate complexes by molecular docking and quantum mechanical calculations. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 299–311.
- (19) Lahiri, D. K.; Farlow, M. R.; Sambamurti, K.; Greig, N. H.; Giacobini, E.; Schneider, L. S. A critical analysis of new molecular targets and strategies for drug developments in Alzheimer's disease. *Curr. Drug Targets.* **2003**, *4*, 97–112.
- (20) Alvarez, A.; Opazo, C.; Alarcón, R.; Garrido, J.; Inestrosa, N. C. Acetylcholinesterase promotes the aggregation of amyloid-beta peptide fragments by forming a complex with the growing fibrils. *J. Mol. Biol.* **1997**, *272*, 348–361.
- (21) Castro, A.; Martínez, A. Peripheral and dual binding site acetylcholinesterase inhibitors: Implications in treatment of Alzheimer's disease. *Mini-Rev. Med. Chem.* **2001**, *1*, 267–272.
- (22) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of acetylcholinesterase from *Torpedo californica*: A prototypic acetylcholine-binding protein. *Science* **1991**, *253*, 872–879.
- (23) Barril, X.; Kalko, S. G.; Orozco, M.; Luque, F. J. Rational design of reversible acetylcholinesterase inhibitors. *Mini-Rev. Med. Chem.* **2002**, *2*, 27–36.
- (24) Gregor, V. E.; Emmerling, M. R.; Lee, C.; Moore, C. J. The synthesis and in vitro acetylcholinesterase and butyrylcholinesterase inhibitory activity of tacrine (Cognex) derivatives. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 861–864.
- (25) Recanatini, M.; Cavalli, A.; Belluti, F.; Piazzi, L.; Rampa, A.; Bisi, A.; Gobbi, S.; Valenti, P.; Andrisano, V.; Bartolini, M.; Cavrini, V. SAR of 9-amino-1,2,3,4-tetrahydroacridine-based acetylcholinesterase inhibitors: Synthesis, enzyme inhibitory activity, QSAR, and structure-based CoMFA of tacrine analogues. *J. Med. Chem.* **2000**, *43*, 2007–2018.
- (26) Camps, P.; El Achab, R.; Morral, J.; Muñoz-Torrero, D.; Badia, A.; Baños, J. E.; Vivas, N. M.; Barril, X.; Orozco, M.; Luque, F. J. Synthesis, in vitro pharmacology, and molecular modeling of very potent tacrine-huperzine A hybrids as acetylcholinesterase inhibitors of potential interest for the treatment of Alzheimer's disease. *J. Med. Chem.* **1999**, *42*, 3227–3242.
- (27) Camps, P.; El Achab, R.; Morral, J.; Muñoz-Torrero, D.; Badia, A.; Baños, J. E.; Vivas, N. M.; Barril, X.; Orozco, M.; Luque, F. J. New tacrine-huperzine A hybrids (huprines): Highly potent tight-binding acetylcholinesterase inhibitors of interest for the treatment of Alzheimer's disease. *J. Med. Chem.* **2000**, *43*, 4657–4666.
- (28) Jaén, J. C.; Gregor, V. E.; Lee, C.; Davis, R.; Emmerling, M. Acetylcholinesterase inhibition by fused dihydroquinazoline compounds. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 737–742.
- (29) Desai, M.; Thadeio, P. F.; Lipinski, C. A.; Liston, D. R.; Spencer, R. W.; Williams, I. H. Physical parameters for brain uptake: Optimizing logP, logD and pK_a of THA. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 411–414.
- (30) Harel, M.; Schalk, I.; Ehret-Sabatier, L.; Bouet, F.; Coeldner, M.; Hirth, C.; Axelsen, P. H.; Silman, I.; Sussman, J. L. Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9031–9035.
- (31) Dvir, H.; Wong, D. M.; Harel, M.; Barril, X.; Orozco, M.; Luque, F. J.; Muñoz-Torrero, D.; Camps, P.; Rosenberry, T. L.; Silman, I.; Sussman, J. L. 3D structure of *Torpedo californica* acetylcholinesterase complexed with huprine X at 2.1 Å: Kinetic and molecular dynamics correlates. *Biochemistry* **2002**, *41*, 2970–2981.
- (32) Wlodek, S. T.; Antosiewicz, J.; McCammon, J. A.; Straatsma, T. P.; Wilson, M. K.; Briggs, M. J.; Humblet, C.; Sussman, J. L. Binding of tacrine and 6-chlorotacrine by acetylcholinesterase. *Biopolymers* **1996**, *38*, 109–117.
- (33) Emmerling, M. R.; Sobkowitz, H. M. Differentiation and distribution of acetylcholinesterase molecular forms in the mouse cochlea. *Hearing Res.* **1988**, *32*, 137–145.
- (34) Ellman, G. L.; Courtney, K. D.; Andres, B., Jr.; Featherstone, R. M. A new and rapid colorimetric determination of acetylcholinesterase activity. *Biochem. Pharmacol.* **1961**, *7*, 88–95.
- (35) Rodríguez-Barrios, F.; Gago, F. Chemometrical identification of mutations in HIV-1 reverse transcriptase conferring resistance or enhanced sensitivity to arylsulfonfylbenzotriazoles. *J. Am. Chem. Soc.* **2004**, *126*, 2718–2719.
- (36) Radic, Z.; Gibney, G.; Kawamoto, S.; MacPhee-Quigley, K.; Bongiorno, C.; Taylor, P. Expression of recombinant acetylcholinesterase in a baculovirus system: Kinetic properties of glutamate 199 mutants. *Biochemistry* **1992**, *31*, 9760–9767.
- (37) Wlodek, S. T.; Antosiewicz, J.; Briggs, J. M. On the mechanism of action of acetylcholinesterase: The electrostatically induced acceleration of the catalytic acylation step. *J. Am. Chem. Soc.* **1997**, *119*, 8159–8165.
- (38) Millard, C. B.; Koellner, G.; Ordentlich, A.; Shafferman, A.; Silman, I.; Sussman, J. L. Reaction products of acetylcholinesterase and VX reveal a mobile histidine in the catalytic triad. *J. Am. Chem. Soc.* **1999**, *121*, 9883–9884.
- (39) Bartolucci, C.; Perola, E.; Pilger, C.; Fels, G.; Lamba, D. Three-dimensional structure of a complex of galanthamine (Nivalin) with acetylcholinesterase from *Torpedo californica*: Implications for the design of new anti-Alzheimer drugs. *Proteins* **2001**, *42*, 182–191.
- (40) Raves, M. L.; Harel, M.; Pang, Y.-P.; Silman, I.; Kozikowski, A. P.; Sussman, J. L. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (–)-huperzine A. *Nature Struct. Biol.* **1997**, *4*, 57–63.
- (41) Dvir, H.; Jiang, H. L.; Wong, D. M.; Harel, M.; Chetrit, M.; He, X. C.; Jin, G. Y.; Yu, G. L.; Tang, X. C.; Silman, I.; Bai, D. L.; Sussman, J. L. X-ray structures of *Torpedo californica* acetylcholinesterase complexed with (+)-huperzine A and (–)-huperzine B: Structural evidence for an active site rearrangement. *Biochemistry* **2002**, *41*, 10810–10818.
- (42) Wong, D. M.; Greenblatt, H. M.; Dvir, H.; Carlier, P. R.; Han, Y.-F.; Pang, Y.-P.; Silman, I.; Sussman, J. L. Acetylcholinesterase complexed with bivalent ligands related to huperzine A: Experimental evidence for species-dependent protein–ligand complementarity. *J. Am. Chem. Soc.* **2003**, *125*, 363–373.
- (43) Insight II, version 98.0, 1998; Molecular Simulations Inc., 9685 Scantion Road, San Diego, CA 92121-2777.
- (44) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–219.
- (45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, Jr., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11.2; Gaussian, Inc.; Pittsburgh, PA, 2001.
- (46) Foye, W. O.; Lemke, T. L.; Williams, D. A. *Principles of Medicinal Chemistry*, 4th ed.; Williams and Wilkins: Media, PA, 1995.
- (47) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (48) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Kollman, P. A. Second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (49) 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. Structures of recombinant native and E202Q mutant human acetylcholinesterase complexed with the snake-venom toxin fasciculin-II. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 1385–1394.
- (50) 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.
- (51) Barreiro, E. J.; Camara, C. A.; Verli, H.; Brazil-Mas, L.; Castro, N. G.; Cintra, W. M.; Aracava, Y.; Rodrigues, C. R.; Fraga, C. A. M. Design, synthesis, and pharmacological profile of novel fused pyrazolo[4,3-d]pyridine and pyrazolo[3,4-b][1,8]naphthyridine isozomers: A new class of potent and selective acetylcholinesterase inhibitors. *J. Med. Chem.* **2003**, *46*, 1144–1152.
- (52) Contreras, J. M.; Parrot, I.; Sippl, W.; Rival, Y. M.; Wermuth, C. G. Design, synthesis, and structure–activity relationships of a series of 3-[2-(1-benzylpiperidin-4-yl)ethylamino]pyridazine derivatives as acetylcholinesterase inhibitors. *J. Med. Chem.* **2001**, *44*, 2707–2718.
- (53) Tumiatti, V.; Rosini, M.; Bartolini, M.; Cavalli, A.; Marucci, G.; Andrisano, V.; Angeli, P.; Banzari, R.; Minarini, A.; Recanatini, M.; Melchiorre, C. Structure–activity relationships of acetylcholinesterase noncovalent inhibitors based on a polyamine backbone. 2. Role of the substituents on the phenyl ring and nitrogen atoms of caproactime. *J. Med. Chem.* **2003**, *46*, 954–966 (see also ref 10 herein).

- (54) Kryger, G.; Silman, I.; Sussman, J. L. Structure of acetylcholinesterase complexed with E2020 (Aricept): Implications for the design of new anti-Alzheimer drugs. *Structure* **1999**, *7*, 297–307.
- (55) 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.
- (56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (57) Q² version 4.5, 2000, Multivariate Infometric Analysis, S.r.l. Perugia, Italy.
- (58) For complexes with **t1** and dihydroquinazolines: WAT 589 and WAT 607. For complexes with tacrines (with the exception of **t1**): WAT 590 and WAT 607. For complexes with huprines: WAT 848 and WAT 891.
- (59) Nicholls, A.; Honig, B. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *J. Comput. Chem.* **1991**, *12*, 435–445.
- (60) Rodríguez-Barrios, F.; Pérez, C.; Lobatón, E.; Velázquez, S.; Chamorro, C.; San-Félix, A.; Pérez-Pérez, M. J.; Camarasa, M. J.; Pelemans, H.; Balzarini, J.; Gago, F. Identification of a putative binding site for [2',5'-bis-*O*-(*tert*-butyldimethylsilyl)- β -D-ribofuranosyl]-3'-spiro-5''-(4''-amino-1'',2''-oxathiole-2'',2''-dioxide) thymine (TSAO) derivatives at the p51–p66 interface of HIV-1 reverse transcriptase. *J. Med. Chem.* **2001**, *44*, 1853–1865.
- (61) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **1987**, *9*, 327–335.

JM049877P