Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

Original article

Receptor-dependent (RD) 3D-QSAR approach of a series of benzylpiperidine inhibitors of human acetylcholinesterase (HuAChE)

Jocley Queiroz Araújo [a,*], Monique Araújo de Brito [b], Lucas Villas Bôas Hoelz [a], Ricardo Bicca de Alencastro [a], Helena Carla Castro [c], Carlos Rangel Rodrigues [d], Magaly Girão Albuquerque [a,*]

[a] Universidade Federal do Rio de Janeiro (UFRJ), Instituto de Química (IQ), Programa de Pós-Graduação em Química (PPGQu), Laboratório de Modelagem Molecular (LabMMol), Rio de Janeiro, RJ, Brazil
[b] Universidade Federal Fluminense (UFF), Faculdade de Farmácia (FF), Santa Rosa, Niterói, RJ, Brazil
[c] Universidade Federal Fluminense (UFF), Instituto de Biologia (IB), Centro, Niterói, RJ, Brazil
[d] Universidade Federal do Rio de Janeiro (UFRJ), Faculdade de Farmácia (FF), Rio de Janeiro, RJ, Brazil

A R T I C L E   I N F O

A B S T R A C T

Acetylcholine inhibitors (AChEIs) are currently considered as potential drugs for treating Alzheimer disease. In this work, we developed a receptor-dependent 3D-QSAR (RD-3D-QSAR) models based on a series of 60 benzylpiperidine inhibitors of human acetylcholinesterase to support the design of new AChEIs. The best two models, A–F ($N = 47$, $q^2 = 0.736$, $r^2 = 0.860$) and C–F ($N = 47$, $q^2 = 0.753$, $r^2 = =0.900$) were developed and validated by a combined GA-PLS approach, available in WOLF. Residues of the aromatic gorge (Tyr341 and Trp439) and catalytic triad (His447) are related to both equations showing the consistency of these models with the SAR. Based on those models we have proposed four new benzylpiperidine derivatives and predicted the pIC_{50} for each molecule. The good predicted potency of benzylpiperidine derivative, **IIa**, indicates that it is a potential candidate as a new HuAChE inhibitor.

© 2010 Elsevier Masson SAS. All rights reserved.

1. Introduction

Alzheimer disease (AD) is a neurodegenerative disorder associated with a selective loss of cholinergic neurons in the brain and decreasing levels of the acetylcholine (ACh) [1]. This disease is related to the increase of beta-amyloid-rich senile plaques and neurofibrillary tangles in the brain [2], whereas it leads to a progressive decline of the cognitive function, executive function losses, memory deficits, and eventually to incapacitating dementia before death [3]. AD is associated to risk factors, such as parental age at time of birth [4], apolipoprotein-E [5], head injury [6], diabetes [7], hypertension [8], high cholesterol levels [9], strokes [10], and smoking [11].

The prevalence of AD increases dramatically with aging, and it doubles for every five-year interval after the age of 65 [12]. The disease affects more women than men, probably to the women higher life expectancy [13,14]. It has been estimated that 30%–50% of individuals in their 8th to 9th decades have some degree of AD [15]. Studies showed that the life expectancy for patients with AD can reach about ten years after diagnosis with considerable variability among patients [16]. In 2006, there were 26.6 million cases of AD in the world, and by 2050 it is predicted that AD prevalence will grow fourfold ($\sim$106.8 million) and one in 85 persons will be affected by this disease worldwide [17].

Currently there is no efficient treatment for AD but a drug-therapy that may temporarily relieve some of the symptoms. This drug-therapy is based on the "cholinergic hypothesis" [18], which states that the decrease of the cholinergic transmission plays a major role in the expression of cognitive, functional and, possibly, behavioral symptoms in AD [19–22]. Based on that, the neurotransmitters replacement is the current treatment option for AD patients.

In order to improve cholinergic neurotransmission, different strategies have been investigated including the increase of synthesis or pre-synaptic release of ACh, the stimulation of cholinergic post-

* Corresponding authors. Instituto de Química, UFRJ, Av. Athos da Silveira Ramos, 149, Centro de Tecnologia (CT), Bloco A, Sala 609, CEP 21941-909, Cidade Universitária, Rio de Janeiro, RJ, Brazil. Tel./fax: +55 21 2562 7132.
E-mail addresses: jocleymol@yahoo.com.br (J.Q. Araújo), moniquebrito@vm.uff.br (M.A. de Brito), lucashoelz@yahoo.com.br (L.V.B. Hoelz), bicca@iq.ufrj.br (R.B. de Alencastro), hcastrorangel@vm.uff.br (H.C. Castro), rangel@pharma.ufrj.br (C.R. Rodrigues), magaly@iq.ufrj.br (M.G. Albuquerque).

synaptic muscarinic and nicotinic receptors, and the reduction of ACh synaptic degradation using AChE inhibitors (e.g., AChEIs or anticholinesterase agents). However, except for the use of AChEIs, the current data do not support the use of ACh precursors, presynaptic releasing agents, or muscarinic agonists, due to its inefficacy and unacceptable side-effects [18].

Therefore, the main pharmacotherapy for AD are composed by compounds that inhibit AChE and consequently increase both ACh levels and cholinergic neurotransmission in the brain [23]. Currently, the use of AChEIs and NMDA (N-methyl-D-Aspartate) receptor antagonists is endorsed as standard first-line therapy in patients with mild-to-moderate AD [24,25].

Since AChE inhibition has beneficial effects on cognitive, functional, and behavioral symptoms of AD [26–28], five drugs have been approved by the U.S. Food and Drug Administration (FDA) for the treatment of AD. They include four AChEIs (Figure S1, Supporting Information), tacrine (Cognex™, approved in 1993), donepezil (Aricept™, 1996), rivastigmine (Exelon™, 2000), and galantamine (Reminyl™, 2001) - and one NMDA receptor antagonist - memantine (Namenda™, 2003) [24].

Among the AChEIs, tacrine is rarely used because of its hepatotoxicity (about 50% of patients) [23,29]. In contrast, donepezil and rivastigmine are commonly used in the early-to-moderate stages of AD patients to treat cognitive loss since they are not hepatoxic as tacrine. However, none of them are completely effective to cure AD patient and often present adverse effects [30]. Only galantamine shows clinically relevant neuroprotective effects with a promise profile in the treatment of AD patients, according to data from Phase II studies [31,32]. Therefore, due to the bioavailability problems and side-effects of the current pharmacotherapy, it is urgent to find more effective AChEIs to treat AD.

The available 3D-structures of AChE revealed its active site located at the bottom of a narrow cavity, about 20 Å deep, named as the aromatic gorge. This region is composed by highly conserved aromatic residues (e.g. Tyr72, Trp86, Phe123, Tyr124, Tyr133, Trp236, Trp286, Phe295, Phe297, Tyr337, Phe338, Tyr341, Trp439 and Tyr449) and plays different functional roles [33].

The AChE active site consists of a catalytic triad (CT, Ser203, Glu334, and His447), a catalytic anionic site (CAS, Trp86, Glu202, and Tyr337), an oxyanion hole (Gly121, Gly122, and Ala204), and an acyl pocket (Phe295 and Phe297). Additionally, there is a peripheral anionic site (PAS, Tyr72, Asp74, Tyr124, Trp286, and Phe297) located at the gorge entrance [34].

Therefore, in order to identify the relationship between the chemical structure and the biological activity of compounds that inhibit the human AChE, we have performed a receptor-dependent (RD) 3D-QSAR study of a series of 60 benzylpiperidine derivatives, analogues of donepezil (E2020, Figure S1, Supporting Information) described on the literature [35–37]. All the RD-3D-QSAR models were generated by WOLF program and validated. These models enabled the design and prediction of the biological activity values of four new benzylpiperidine derivatives (**Ia**, **Ib**, **IIa** and **IIb**) using molecular modification strategies widely employed in rational drug design. Our data may provide some guidelines to design new human AChE inhibitors for treatment of Alzheimer disease.

2. Computational approach

2.1. Biological data

The RD-3D-QSAR models were built using a series of 60 benzylpiperidine derivatives, analogues of donepezil (E2020, Figure S1, Supporting Information) [35–38]. The biological activity value of these compounds were obtained from literature that described it as the minimum concentration able to inhibit 50%

activity in vitro (IC_{50}) of human erythrocytes AChE (HuAChE) using Ellman pharmacological protocol [39]. Then, the IC_{50} values in nM were converted to M and turned into $-\log IC_{50}$ (pIC_{50}). The series of 60 benzylpiperidine derivatives was divided in two sets, a training (**1**–**47**) and test (**48**–**60**) sets (Table 1). The criteria of using 20% of the total of the compounds was choose to select the compounds of the test set as described in the literature [40]. In addition, we considered the structural diversity and biological activity of the database, which minimizes bias during the QSAR validation stage. Figure S2 (Supporting Information) shows the biological activity range of these benzylpiperidine derivatives in the training (pIC_{50} from 5.59 to 9.48 M) and test (pIC_{50} from 6.05 to 9.31 M) sets.

2.2. Ligands building

The 3D-structures of the entire set of 60 benzylpiperidine derivatives were built using the Spartan'06 software [41]. E2020 complexed with AChE from Torpedo californica - TcAChE (PDB ID: 1EVE [42]) - was used as a template. All structures were submitted to a single point calculation, using the AM1 semiempirical method [43]. Figure S3 (Supporting Information) shows the superposition of E2020 and compound **33**, the most potent of this series.
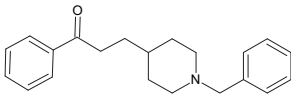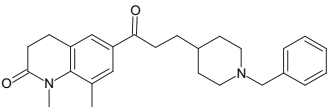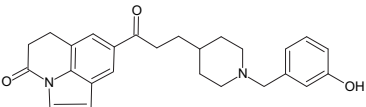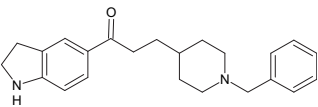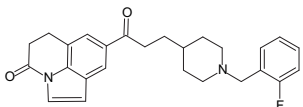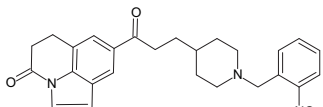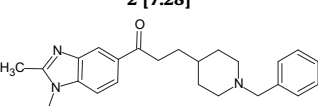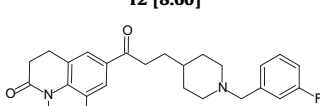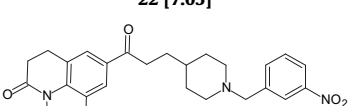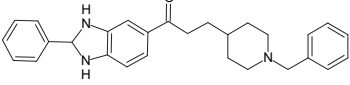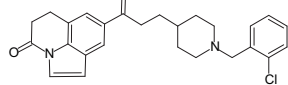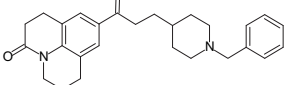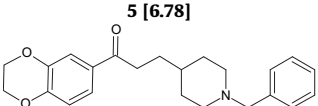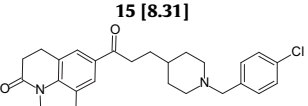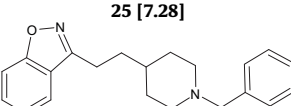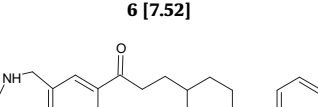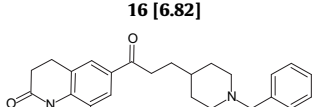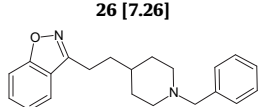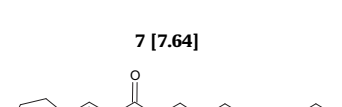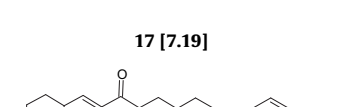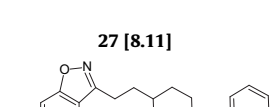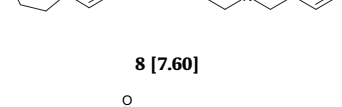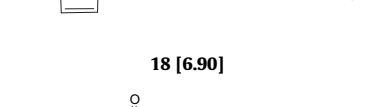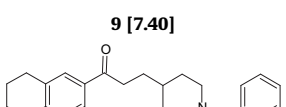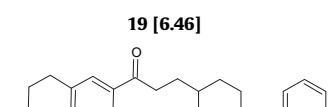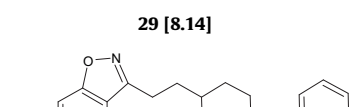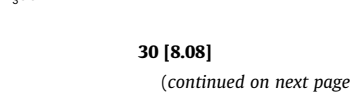
2.3. Human AChE (HuAChE) model building

Since the pharmacological assays of the 60 benzylpiperidine derivatives were performed with the HuAChE and because there is not an E2020 structure bound to the HuAChE available in the PDB, we built a E2020/HuAChE model, which is necessary for the construction of the benzylpiperidines/HuAChE complexes. The E2020/HuAChE model was generated, using as reference the X-ray structures of fasciculin/HuAChE (PDB ID: 1B41) [44] and E2020/TcAChE (PDB ID: 1EVE) [42]. Both complexes were corrected by removal of water, ligands, cofactors, and ions, except the E2020 ligand from TcAChE that would be transferred to the new model. Since the 1B41 structure was incomplete, e.g. there were two missing coils (Pro259-Gly264 and Arg493-Asp494), these two sequences in FASTA format were generated, using the MODELLER program [45], and then, they were aligned with the T-COFFEE server [46]. The alignment thus obtained and the corrected structures of 1B41 (HuAChE) and 1EVE (TcAChE) enabled the building of the E2020/HuAChE model, showing full coils (Figure S4, Supporting Information). Then, we submitted the model to geometry optimization using the MODELLER program, and to subsequent validation by the Ramachandran plot [47] using the PROCHECK program [48] (Figure S5, Supporting Information).

2.4. Complexes building and optimization
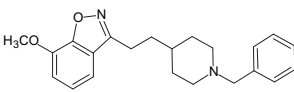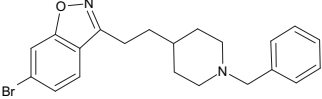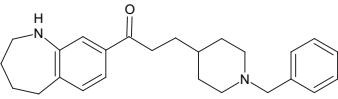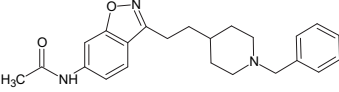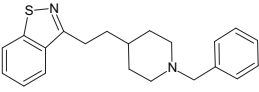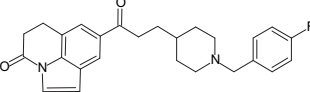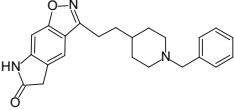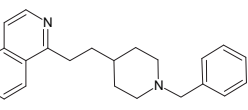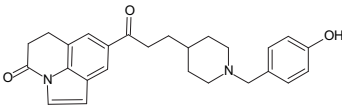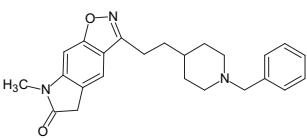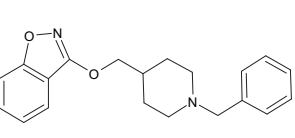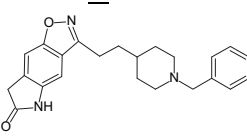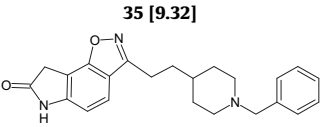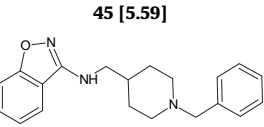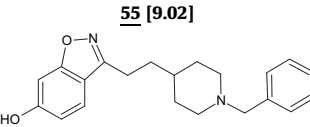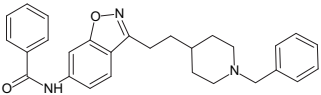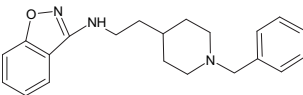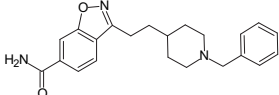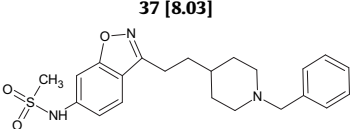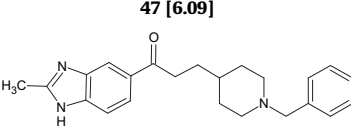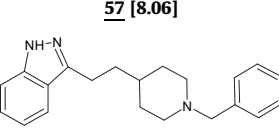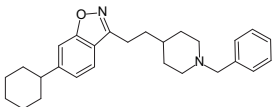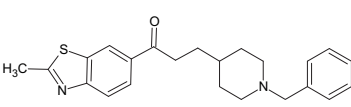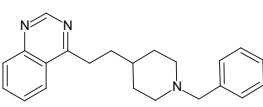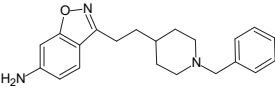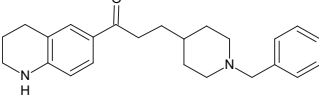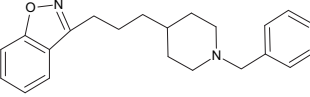
Each structure of the 60 benzylpiperidine derivatives was docked into the active site of HuAChE model through superposition with E2020, by alignment of selected atoms (see Figure S6, Supporting Information), using the Fit-Atoms module of Sybyl v.8.0 package [49]. Then, the E2020 reference structure in each ligand/HuAChE complex was excluded. Hydrogen atoms were added to the enzyme, the acid and basic groups were ionized, and the ligand/enzyme complexes were optimized in two steps: first, only the ligand and then the entire complex. All optimizations were carried out using Tripos force field (Sybyl) with an energy criterion less than 1.0 kcal/mol.Å, using the steepest descent algorithm, followed by the conjugate gradient algorithm with an energy criterion less than 0.5 kcal/mol.Å.

Table 1
Structure of the 60 benzylpiperidine inhibitors of human AChE with its respective pIC$_{50}$ (M) values in brackets. The numeration of the test set is underlined.

1 [6.52] **11 [8.01]** **21 [8.06]**

2 [7.28] **12 [8.60]** **22 [7.05]**

3 [8.37] **13 [8.89]** **23 [8.54]**

4 [7.48] **14 [8.29]** **24 [7.80]**

5 [6.78] **15 [8.31]** **25 [7.28]**

6 [7.52] **16 [6.82]** **26 [7.26]**

7 [7.64] **17 [7.19]** **27 [8.11]**

8 [7.60] **18 [6.90]** **28 [8.24]**

9 [7.40] **19 [6.46]** **29 [8.14]**

10 [8.44] **20 [8.96]** **30 [8.08]**

Table 1 (*continued*)

31 [8.15]	**41 [7.30]**	**51 [6.71]**
32 [8.55]	**42 [7.00]**	**52 [8.34]**
33 [9.48]	**43 [6.66]**	**53 [9.31]**
34 [9.24]	**44 [6.68]**	**54 [7.37]**
35 [9.32]	**45 [5.59]**	**55 [9.02]**
36 [8.44]	**46 [6.49]**	**56 [7.59]**
37 [8.03]	**47 [6.09]**	**57 [8.06]**
38 [7.85]	**48 [7.92]**	**58 [6.92]**
39 [9.10]	**49 [8.17]**	**59 [6.47]**
40 [7.70]	**50 [7.19]**	**60 [6.05]**

2.5. Molecular dynamics simulations (MDS) of the complexes

Each ligand/HuAChE complex was submitted to a preliminary optimization, with the steepest descent and conjugate gradient algorithms, carried out in the GROMACS 4.0 program [50] using the GROMOS96 force field [51]. In the next step, the molecular topologies of the ligands were generated by the PRODRG server [52]. Then, the MDS were performed in the GROMACS program using the GROMOS96 force field. The ligand/protein complexes were submitted to these simulations by 1.0 ns under constant conditions of temperature (310 K) and pressure (1.0 bar) using the Berendsen thermostat [53]. All bonds that involved hydrogen atoms for ligands and protein were frozen by the LINCS algorithm [54], enhancing the integration step to 1.0 ns by the Verlet algorithm [55]. Finally, the long-range electrostatic interactions were treated using the PME algorithm (Particle-Mesh Ewald) [56].

2.6. RD-3D-QSAR models construction

The steric and electrostatic interaction energies were calculated between each ligand and the HuAChE residues within radii of 10 Å around the ligand, according to the pruning approach [57,58]. They were considered as independent variables in this RD-3D-QSAR study. These steric (Lennard-Jones) and electrostatic (Coulomb) interaction energies were calculated as an average value from 500 ps to 1 ns of the molecular dynamics simulation using the GROMACS program.

In order to estimate the influence of its descriptors (independent variables) on the predictive capability of the RD-3D-QSAR models, we constructed and tested 12 databases clustered in three groups (each database is composed by one set of dependent variables and one set of independent variable). Then, each database was submitted to the WOLF program [59] to construct QSAR equations with six to ten terms using the combined Genetic Function Approximation (GFA) and Partial Least Squares (PLS) technique [60], three principal components, population of 700–800 equations, crossover operations of 10000–500000, mutation rate of 100%, and smoothing factor of 0.1–1.0.

2.7. RD-3D-QSAR models validation

The best models from each database were ranked according to its lack-of-fit (LOF) values (Eq. (1)), which prevent the models overfitting as it penalizes additional terms when identical LSE values are observed [59].

$$LOF = 1 - \frac{LSE}{1 - \frac{(c+dp)^2}{M}} \qquad (1)$$

In Equation (1), **LSE** is the least squares error calculated by difference between the experimental and predicted activity values; **c** is the number of equation terms; **d** is the smoothing factor, e.g. the adjusted parameter that controls the number of equation terms (independent variables); **p** is the number of characteristics in the equation terms (default = 1); and **M** is the number of training set compounds.

Thus, these models were submitted to the leave-one-out cross-validation method (LOO$_{cv}$) to test the internal predictive capability of models [61] represented by the q^2 (r$^2_{cv}$) parameter (Eqs. (2) and (3)).

$$r^2_{cv} = 1 - \frac{PRESS}{\sum (Y - Y_{mean})^2} \qquad (2)$$

$$PRESS = \sum \left(Y - Y_{pred} \right)^2 \qquad (3)$$

According to Equations (2) and (3), **PRESS** is the predictive sum of squares of LOO$_{cv}$; **Y** is the observed potency value; **Y$_{mean}$** is the mean value of potency; and **Y$_{pred}$** is the predicted potency value. After the validation by LOO$_{cv}$, the models were submitted to non-cross validation procedure to evaluate the fitting quality of the models, represented by r^2 and SEE (standard error of estimate).

In addition, an external validation procedure was performed using the test set compounds, where the predictivity level of the RD-3D-QSAR models were evaluated in terms **S$_{PRESS}$** (standard deviation of cross-validation) [40] (Eq. (4)).

$$S_{PRESS} = \sqrt{\frac{PRESS}{n - k - 1}} \qquad (4)$$

According to Equation (4), **n** is the number of compounds; and **k** is the number of terms in the model, where the "**n-k**-1" expression corresponds to model degrees of freedom.

The cross-correlation matrix of the residuals of fit (pIC$_{50}$ experimental − pIC$_{50}$ predicted) was computed to determine if there are highly correlated models [57,62]. Also, the cross-correlation matrix of descriptors was computed to determine if there are highly correlated descriptors ($R > 0.7$) in the same equation, in order to avoid redundant information [63].

In this work we considered as outliers all compounds of the training and test sets with a residue value higher than twofold of the model Standard Error of Estimate (SEE) [64].

Finally, the best RD-3D-QSAR models were those with larger values of r^2, q^2 and q$^2_{adj}$ (Eq. (5)), associated to smaller errors (SE$_{cv}$ e SEE) and number of outliers.

$$q^2_{adj} = 1 - \left(1 - q^2\right)\frac{n - 1}{n - p} \qquad (5)$$

In Equation (5), **q^2** is the cross-validation squared correlation coefficient; **n** is the number of training set compounds; and **p** is the number of equation terms.

3. Results and discussion

In this work, we built 12 databases (DBs) derived from three main groups: A (DB-A01, DB-A02, DB-A03 and DB-A04), B (DB-B01, DB-B02, DB-B03 and DB-B04) and C (DB-C01, DB-C02, DB-C03 and DB-C04). In the DBs from Group A, the independent variables are the steric and electrostatic interaction energies calculated by residue; in DBs from Group B, these steric and electrostatic energies were summed for each residue; and in the DBs from Group C, the independent variables from the DBs of Groups A and B were combined.

From the pool of equations generated by the WOLF program, we selected the five best QSAR equations with 6–10 terms for each DB. These equations were evaluated relative to its statistics parameters, as well as the number of outlier compounds to select the best equation for each DB. We analyzed the predicted potency and the respective residual values (training and test sets), the cross-correlation matrix of residuals and descriptors, and the outliers of each QSAR equation.

Table 2 shows the best QSAR equation for each DB, except for the DBs from Group B whose equations showed reduced predictivity. In this work, we will not analyze all the equations from Table 2 since some of them are statistically similar. Instead, only the equation with the greatest q$^2_{adj}$ value for each group of DB will be analyzed. Therefore, as two groups were selected (A and C), two equations will be analyzed: Eq. A–F and Eq. C–F.

3.1. Analysis of Equation A–F

This equation shows 10 terms (independent variables) where each is represented by a 3-letter code of the selected residue and a code, LJ or C, relative to the steric (Lennard-Jones) or electrostatic (Coulomb)

Table 2
Statistics parameters of the best RD-3D-QSAR equations for Groups A and C.

Equation	Terms	q^2_{adj}	q^2	SEcv	r^2	SEE	Crossover	Outliers
A–C	8	0.650	0.703	0.482	0.794	0.648	450000	3
A–F	10	0.672	0.736	0.454	0.860	0.857	440000	4
A–K	10	0.531	0.623	0.540	0.832	0.865	250000	4
A–P	10	0.514	0.609	0.547	0.805	0.888	210000	3
C–B	9	0.565	0.641	0.529	0.814	1.134	250000	2
C–F	10	0.693	0.753	0.431	0.900	1.301	390000	2
C–K	10	0.538	0.628	0.537	0.823	0.894	410000	3
C–P	10	0.533	0.624	0.537	0.792	0.640	410000	3

contributions, respectively. Fig. 1 shows the graphic representation of Eq. A–F, using derivatives **33** (Fig. 1A) and **45** (Fig. 1B), the most and the least potent compounds of this series, respectively.

As seen, only one term of Eq. A–F (**Trp439C**) represents the Coulomb contribution whereas the other nine terms (**Gln71LJ**, **Pro78LJ**, **Gly126LJ**, **Asp131LJ**, **Val153LJ**, **Val282LJ**, **Tyr341LJ**, **Arg364LJ**, and **His447LJ**) correspond to the Lennard-Jones contribution. This result points to feasible predominance of the steric over the electrostatic contribution in the HuAChE active site. Possibly, any conformational change that allows favorable or avoid unfavorable steric contacts of a ligand of this series with neighbor HuAChE residues should be more important than its interaction with long-range charged residues, which is peculiar of electrostatic interaction in this environment.

$$pIC_{50} = 4.862 + 0.124\mathbf{Gln71LJ} - 25.745\mathbf{Pro78LJ} - 0.271\mathbf{Gly126LJ}$$

$$- 2.751\mathbf{Asp131LJ} + 7.481\mathbf{Val153LJ} - 3.215\mathbf{Val282LJ} - 0.068\mathbf{Tyr341LJ}$$

$$- 4.641\mathbf{Arg364LJ} + 0.249\mathbf{Trp439C} - 0.037\mathbf{His447LJ}$$

Fig. 1. Graphic representation of Eq. **A–F** showing compounds **33** (A) and **45** (B) (stick model and color by element), the most and the least potent derivatives of this series, respectively. The residues are represented in stick models and colored according the Lennard-Jones (green: Gln71, Pro78, Gly126, Asp131, Val153, Val282, Tyr341, Arg364 and **His447**) and Coulomb (red: Trp439) contributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The equation coefficients (Fig. 1) also may provide useful information to confirm this supposition. The terms with positive coefficients signs (**Gln71LJ**, **Val153LJ**, and **Trp439C**) must show positive interaction energy to increase the predicted potency of a compound. The same concept can be applied to terms with negative coefficient signs (**Pro78LJ**, **Gly126LJ**, **Asp131LJ**, **Val282LJ**, **Tyr341LJ**, **Arg364LJ**, and **His447LJ**), but the corresponding interaction energy of these terms must have a negative sign to increase the predicted potency. In both cases, the predicted potency values are directly proportional to the magnitude of each term coefficient of this equation.

For instance, the analysis of the compound **33** according Eq. A–F (Fig. 1) reveals the **Trp439C** term with a positive coefficient (+0.249) whereas the corresponding interaction energy also shows a positive sign (+3.853 kcal/mol), which contributes for increasing the potency of this compound. On the other hand, the **Tyr341LJ** term has a negative coefficient (- 0.068) analogously to the corresponding interaction energy that presents a negative sign (- 14.109 kcal/mol), also contributing to the increase in the potency of this compound.

The same analysis may be performed for the remaining compounds considering the average interaction energies among the HuAChE selected residues of Eq. A–F with the 60 compounds and the normalized interaction energies (e.g., the average interaction energies multiplied by the correspondent coefficients) (Fig. 2).

As seen in Fig. 2, the average interaction energy values can be clustered in two groups: i) terms with average interaction energies lower than |0.10| kcal/mol (**Pro78LJ**, **Asp131LJ**, **Val153LJ**, **Val282LJ**, and **Arg364LJ**); and ii) terms with average interaction energies between |1.00| and |14.00| kcal/mol (**Gln71LJ**, **Gly126LJ**, **His447LJ**, **Tyr341LJ**, and **Trp439C**). However, the normalized interaction energy (Fig. 2) shows a compensation of the importance of each term in Eq. A–F, increasing the weight of the terms in the first group and decreasing it in the second group. Therefore, considering the normalized interaction energies, **Gly126LJ**, **Tyr341LJ**, **Trp439C**, and **His447LJ** are the most important terms in Eq. A–F and only **Gln71LJ** and **Val153LJ** contribute to decrease the predicted potency, since both have negative average interaction energies and positive coefficients in Eq. A–F.

Among all terms of Eq. A–F, **Trp439C** is the only one that showed a positive average interaction energy value (+1.47 kcal/mol), corresponding to a repulsive electrostatic interaction. However, it contributes to the increase of the predicted potency because its coefficient has a positive sign (+0.249).

In addition, three terms selected in Eq. A–F (**Tyr341LJ**, **Trp439C**, and **His447LJ**) are related to two residues from the aromatic gorge (Tyr341 and Trp439) and one from the HuAChE catalytic triad (His447), showing the consistency of this model with the structure-activity relationship.

3.2. Analysis of Equation C–F

This equation shows 10 terms (independent variables) where each one is represented by the same codes used in Eq. A–F. The terms are described only by the 3-letter code of the selected residue corresponding to the sum of Lennard-Jones and Coulomb contributions, since Eq. C–F is derived from a DB from group C, which is resultant of DBs from groups A and B. Fig. 3 shows the graphic representation of Eq. C–F using derivative **33** (Fig. 3A) and **45** (Fig. 3B), the most and the least potent compounds of this series, respectively.

Our study showed that one term of Eq. C–F (**Gln279C**) represents the individual Coulomb contribution, four terms (**Gly126LJ**, **Val153LJ**, **Val282LJ**, and **Tyr341LJ**) represent the individual Lennard-Jones contribution and other four terms (**Pro78**, **Val340**, **Trp439**, and **Met443**) represent the combined Lennard-Jones and Coulomb contribution. These data pointed that not only the steric contribution, but also the sum of Lennard-Jones and Coulomb

Fig. 2. Average interaction energies between the HuAChE selected residues of Eq. A—F with the 60 compounds and the normalized interaction energies.

contributions are important in the HuAChE active site. So, this result updates what was shown by Eq. A—F where the steric contribution had the larger importance.

The comparison of the identical residues of equations C—F and A—F showed that the terms relative to residues Gly126, Val153, Val282, and Tyr341 are the same, in contrast to the terms related to residues Pro78 and Trp439. The analysis of the cross-correlation matrix of descriptors (Table S1, Supplemental Material) is pertinent only for different terms and shows a perfect correlation ($r = 1.00$) between **Pro78LJ** (Eq. A—F) and **Pro78** (Eq. C—F), and a high correlation ($r = 0.94$) between **Trp439C** (Eq. A—F) and **Trp439** (Eq. C—F). The high correlations in both cases are probably because these residues are contributing with the same kind of information (*e.g.* Pro78 and Trp439 contribute mainly with Lennard-Jones and Coulomb interactions, respectively). In addition, there is a medium correlation ($R = 0.62$) between **Arg364LJ** (Eq. A—F) and **Pro368LJ** (Eq. C—F) (Table S1, Supplemental Material), probably because Arg364 and Pro368 residues are close in space and contribute with the same kind of interaction (Lennard-Jonnes).

Analyzing the coefficients of the terms of Eq. C—F, four have positive signs (**Val153LJ**, **Val340**, **Trp439**, and **Met443**) and six have negative signs (**Pro78**, **Gly126LJ**, **Gln279C**, **Val282LJ**, **Tyr341LJ**, and **Pro368LJ**). To increase the predicted potency, the same analysis performed in Eq. A—F can be performed in Eq. C—F; *i.e.*, terms with positive sign should have positive interaction energy whereas terms with negative coefficient should have negative sign.

For example, in the analysis of compound **33** according Eq. C—F (Fig. 3), the **Val340** term has a positive coefficient ($+0.063$) and the correspondent interaction energy has a positive sign ($+7.066$ kcal/mol), contributing to increase the potency. On the other hand, the **Gly126LJ** term has a negative coefficient ($- 0.240$) and the corresponding interaction energy has a negative sign ($- 3.347$ kcal/mol), also contributing to increase the potency.

Similar to Eq. A—F, the same analysis may be performed with the remaining compounds of Eq. C—F, considering the average interaction energies among the HuAChE selected residues of Eq. C—F with the 60 compounds and the normalized interaction energies (*i.e.*, the average interaction energies multiplied by the correspondent coefficients) (Fig. 4).

As seen in Fig. 4, just like in Eq. A—F, the average interaction energy values of Eq. C—F can be clustered in two groups: i) terms with average interaction energies lower than |0.10| kcal/mol (**Pro78**, **Val153LJ**, **Gln279C**, **Val282LJ**, **Pro368LJ** and **Met443**); and ii) terms with average interaction energies between |1.00| and |14.00| kcal/mol (**Gly126LJ**, **Val340**, **Tyr341LJ** and **Trp439**). However, the

normalized interaction energy (Fig. 4) shows a compensation of the importance of each term in Eq. C—F, increasing the weight of the terms in the first group and decreasing it in the second group. Therefore, considering the normalized interaction energies,

$$pIC_{50} = 5.464 - 43.168\textbf{Pro78} - 0.240\textbf{Gly126LJ} + 7.715\textbf{Val153LJ}$$

$$- 10.240\textbf{Gln279C} - 1.908\textbf{Val282LJ} + 0.063\textbf{Val340} - 0.045\textbf{Tyr341LJ}$$

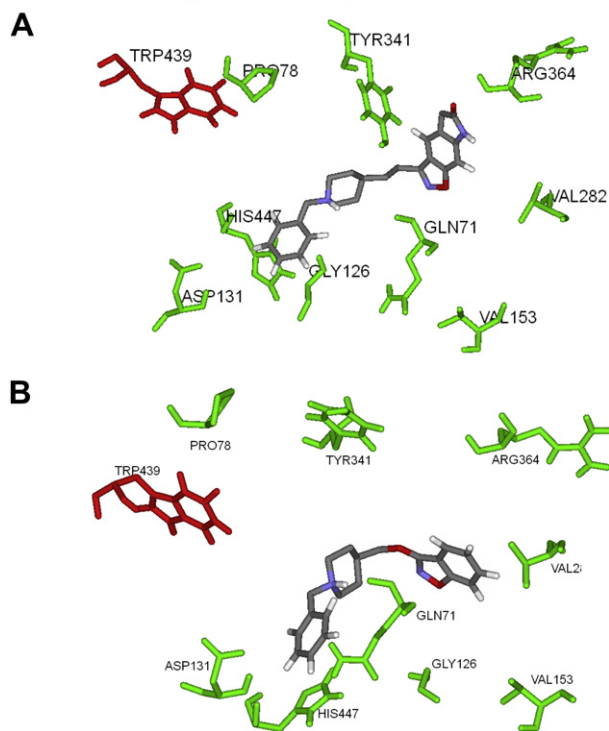$$- 3.845\textbf{Pro368LJ} + 0.416\textbf{Trp439} + 0.560\textbf{Met443}$$
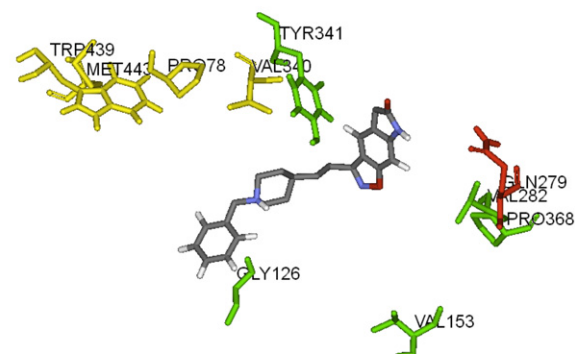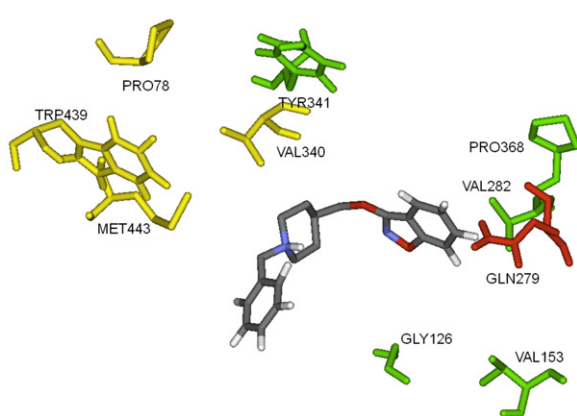
Fig. 3. Graphic representation of Eq. C—F showing compounds 33 (A) and 45 (B) (stick model and color by element), the most and the least potent derivatives of this series, respectively. The residues are represented in stick model and colored according the Lennard-Jones (green: Gly126, Val153, Val282, Tyr341), Coulomb (red: Gln279) and Lennard-Jones/Coulomb (yellow: Pro78, Val340, Trp439, Met443) contributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
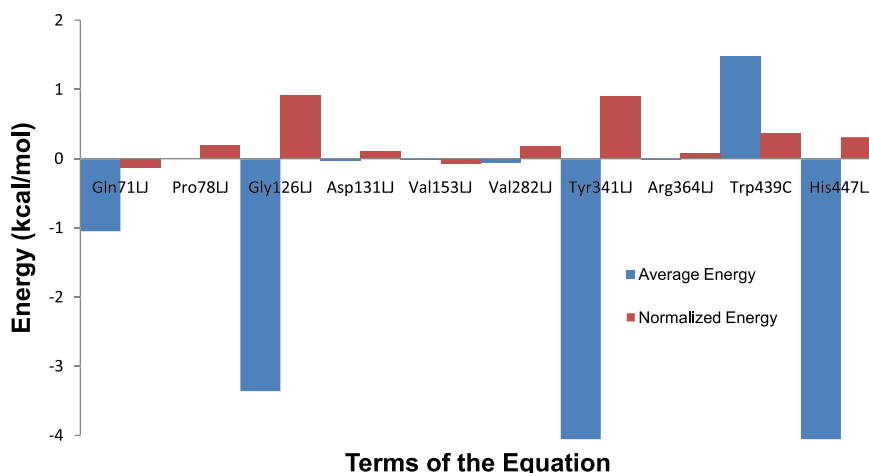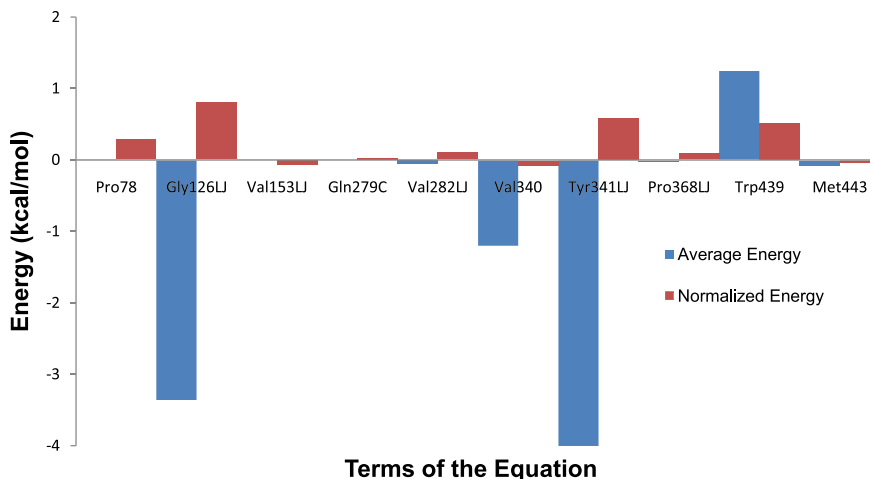
Fig. 4. Average interaction energies between the HuAChE selected residues of Eq. C–F with the 60 compounds and the normalized interaction energies.

Gly126LJ, **Tyr341LJ**, **Trp439**, and **Pro78** are the most important terms in Eq. C–F, while **Val153LJ**, **Val340**, and **Met443** contribute to the decrease of the predict potency, since all have negative average interaction energies and positive coefficients in Eq. C–F.

In addition, differently from Eq. A–F, only two terms selected in Eq. C–F (**Tyr341LJ** and **Trp439C**) are related to two residues from the aromatic gorge (Tyr341 and Trp439) of HuAChE.

3.3. Residues analysis of Equations A–F and C–F

Table 3 shows the experimental (pIC_{50Exp}) and predicted (pIC_{50Pred}) activity and the residual ($pIC_{50Exp} - pIC_{50Pred}$) values for

the training (**1–47**) and test (**48–60**) set compounds, according to equations A–F and C–F.

Almost all compounds in the training set (Eq. A–F, 89% and Eq. C–F, 91%) have residues lower than |0.5| (Table 3), and there are no outliers, since none of them had residual values greater than twice the standard error of the estimate (Eq. A–F, 2SEE = 1.58 and Eq. C–F, 2SEE = 2.39). These results revealed a good internal predictive capability for both models.

In the test set (Table 3), few compounds (Eq. A–F, 23% and Eq. C–F, 15%) have residues lower than |0.5|, while four (**48, 51, 53,** and **59**) and two (**49** and **51**) compounds are outliers in Eq. A–F and Eq. C–F, respectively. As the training and test set residual values show random

Table 3
Experimental (pIC_{50Exp}) and predicted (pIC_{50Pred}) activity and residual ($pIC_{50Exp} - pIC_{50Pred}$) values for the training (**1–47**) and test (**48–60**) set compounds, according to equations A–F and C–F.

#	pIC_{50Exp}	pIC_{50Pred}		Residue[a]		#	pIC_{50Exp}	pIC_{50Pred}		Residue[a]	
		A–F	C–F	A–F	C–F			A–F	C–F	A–F	C–F
1	6.52	6.60	6.45	−0.08	0.07	**31**	8.15	7.99	8.09	0.16	0.06
2	7.28	7.33	7.58	−0.05	−0.30	**32**	8.55	8.59	8.87	−0.04	−0.32
3	8.37	8.20	8.01	0.17	0.36	**33**	9.48	8.67	9.15	0.81	0.33
4	7.48	7.40	7.99	0.08	−0.51	**34**	9.24	8.99	9.40	0.25	−0.16
5	6.78	6.96	7.01	−0.18	−0.23	**35**	9.32	9.01	9.44	0.31	−0.12
6	7.52	7.20	7.50	0.32	0.02	**36**	8.44	8.36	7.98	0.08	0.46
7	7.64	8.12	7.92	−0.48	−0.28	**37**	8.03	7.38	7.40	0.65	0.63
8	7.60	6.84	7.60	0.76	0.00	**38**	7.85	7.61	7.80	0.24	0.05
9	7.40	7.41	7.65	−0.01	−0.25	**39**	9.10	9.39	9.22	−0.29	−0.12
10	8.44	8.41	8.51	0.03	−0.07	**40**	7.70	7.85	7.61	−0.15	0.09
11	8.01	7.98	7.62	0.03	0.39	**41**	7.30	7.26	7.30	0.04	0.00
12	8.60	8.53	8.31	0.07	0.29	**42**	7.00	7.19	7.02	−0.19	−0.02
13	8.89	8.69	8.70	0.20	0.19	**43**	6.66	6.33	6.58	0.33	0.08
14	8.29	8.22	8.32	0.07	−0.03	**44**	6.68	7.40	7.01	−0.72	−0.33
15	8.31	8.59	8.75	−0.28	−0.44	**45**	5.59	6.08	5.49	−0.49	0.10
16	6.82	7.48	7.16	−0.66	−0.34	**46**	6.49	6.09	6.84	0.40	−0.35
17	7.19	7.18	7.00	0.01	0.19	**47**	6.09	6.27	6.34	−0.18	−0.25
18	6.90	7.33	6.64	−0.43	0.26	**48**	7.92	6.19	6.02	**1.73**	1.90
19	6.46	6.86	6.64	−0.40	−0.18	**49**	8.17	7.16	4.79	1.01	**3.38**
20	8.96	9.42	8.76	−0.46	0.20	**50**	7.19	6.87	6.49	0.32	0.70
21	8.06	8.33	7.87	−0.27	0.19	**51**	6.71	10.55	14.08	**−3.84**	**−7.37**
22	7.05	7.14	7.01	−0.09	0.04	**52**	8.34	8.16	9.13	0.18	−0.79
23	8.54	8.75	8.41	−0.21	0.13	**53**	9.31	7.43	7.17	**1.88**	2.14
24	7.80	7.77	7.75	0.03	0.05	**54**	7.37	7.64	7.37	−0.27	0.00
25	7.28	7.34	7.88	−0.06	−0.60	**55**	9.02	8.41	8.51	0.61	0.51
26	7.26	7.30	7.61	−0.04	−0.35	**56**	7.59	8.75	8.10	−1.16	−0.51
27	8.11	7.81	7.55	0.30	0.56	**57**	8.06	6.92	6.73	1.14	1.33
28	8.24	7.75	7.86	0.49	0.38	**58**	6.92	7.82	7.92	−0.90	−1.00
29	8.14	8.28	7.95	−0.14	0.19	**59**	6.47	4.56	5.48	**1.91**	0.99
30	8.08	8.01	8.11	0.07	−0.03	**60**	6.05	6.98	6.18	−0.93	−0.13

[a] Residual values in bold are related to outliers.

variations in relation to the potency variation, we may conclude that there is not tendentious behavior, which is very important for a QSAR study.

3.4. Outliers analysis of Equations A–F and C–F

Eq. A–F has four outliers (Table 3): **48** (Residue = 1.53), **51** (Residue = −3.84), **53** (Residue = 1.88), and **59** (Residue = 1.91). Compound **51** was predicted to be more potent than its experimental potency, the worst prediction, while compounds **48**, **53**, and **59** were predicted to be less potent than their experimental potency. Therefore, we should analyze the descriptors in Eq. A–F that contribute to increase the potency in **51**, and those that contribute to decrease the predict potency in **48**, **53**, and **59**. The analysis of the descriptors of these compounds and of the normalized interaction energies (Fig. 5) for explaining the outlier behavior revealed **Pro78LJ** and **Gln71LJ** as the most significant for outliers **51** for **59**, respectively, according to Eq. A–F (Fig. 1). On the other hand, we could not identify a most significant one for outliers **48** and **53**.

In the case of the **Pro78LJ** descriptor for outlier **51**, the product (*e.g.* normalized interaction energy) of its interaction energy value (−0.208 kcal/mol) and the correspondent coefficient (−25.745) is the largest one for this compound (Fig. 5). Thus, the largest predicted potency value for **51** (pIC_{50Pred} = 10.55), in relation to the experimental one (pIC_{50Exp} = 6.71), can be explained by the large magnitude of its negative coefficient, associated to the negative interaction energy of this descriptor. However, in terms of structural analysis, there is no comprehensive explanation for the outlier behavior of **51**, since the close related analogs in the training set (**5**, **7**, and **8**) were well predicted.

In the case of the **Gln71LJ** descriptor for outlier **59**, the product of its interaction energy value (−7.117 kcal/mol) and the correspondent coefficient (+0.124) in Eq. A–F is the largest one for this compound (Fig. 5). Therefore, its behavior is inverse to that of the **Pro78LJ** descriptor (*e.g.*, the more negative the interaction energy value, the lower will be the predicted potency). In terms of structural analysis, there is no close related analog of **59** (test set) in the training set to which it may be compared.

Since **Gln71LJ** and **Val153LJ** are the unique descriptors that contribute to decrease the potency in Eq. A–F but are irrelevant for **48** and **53**, we could not identify a most significant descriptor to explain the outlier behavior of these derivatives (Fig. 5). In this case, the sum of the products of the interaction energy values of the descriptors that increase the potency and the correspondent coefficients is not larger enough to reach the experimental potency value. In terms of structural analysis, there is no close related analog of **48** (test set) in the training set to which it may be compared, while there is no comprehensive explanation for the outlier behavior of **53**, since the close related analogs in training set (**19**, **20**, and **21**) were well predicted.

Eq. C–F has only two outliers (Table 3): **49** (Residue = 3.38) and **51** (Residue = −7.37). The potency of **51** was overestimated and this was the worst prediction, as in Eq. A–F, while the potency of **49** was underestimated. Therefore, we should analyze the descriptors in Eq. C–F that contribute to increase the potency in **51**, and those that contribute to decrease the predicted potency for **49**. Analyzing the descriptors of these compounds and considering the normalized interaction energies (Fig. 4), according to Eq. C–F (Fig. 3), **Pro78** and **Val340** were the most significant in explaining the outlier behavior of **51** and **49**, respectively.

Similar to the analysis of **Pro78** for outlier **51** in Eq. A–F, the product of its interaction energy value (−0.183 kcal/mol) and the correspondent coefficient (−43.168) in Eq. C–F is the largest one for this compound (Fig. 6). Thus, the larger predicted potency value for **51** ($pIC_{50Pred.}$ = 14.08), in relation to the experimental one (pIC_{50Exp} = 6.71), can be explained by the large magnitude of its negative coefficient, associated to the negative interaction energy of this descriptor.

In the case of **Val340** for outlier **49** in Eq. C–F, the product of its interaction energy value (−31.575 kcal/mol) and the correspondent coefficient (+0.063) in Eq. C–F is the largest one for this compound (Fig. 6). Thus, the lower predicted potency value for **49** ($pIC_{50Pred.}$ = 4.79), in relation to the experimental one (pIC_{50Exp} = 8.17), can be explained by the large magnitude of its negative interaction energy, associated to the small positive coefficient of this descriptor. In terms of structural analysis, there is no close related analog of **49** (test set) in the training set to which it may be compared.

3.5. Analysis of the cross-correlation matrix of the descriptors of Equations A–F and C–F

The cross-correlation matrix of the descriptors of Eq. A–F and Eq. C–F (Table S1, Supplementary Material) shows that there is no significant correlation among the descriptors of each equation, since no R values were greater than |0.7|. The largest correlations were shown to be between **Val153LJ** and **Trp439C** ($R = −0.43$) of Eq. A–F and **Val153LJ** and **Trp439** ($r = −0.44$) of Eq. C–F. This result means that each equation term is independent (*i.e.* there is not redundant information), which is suitable for a QSAR study [63].
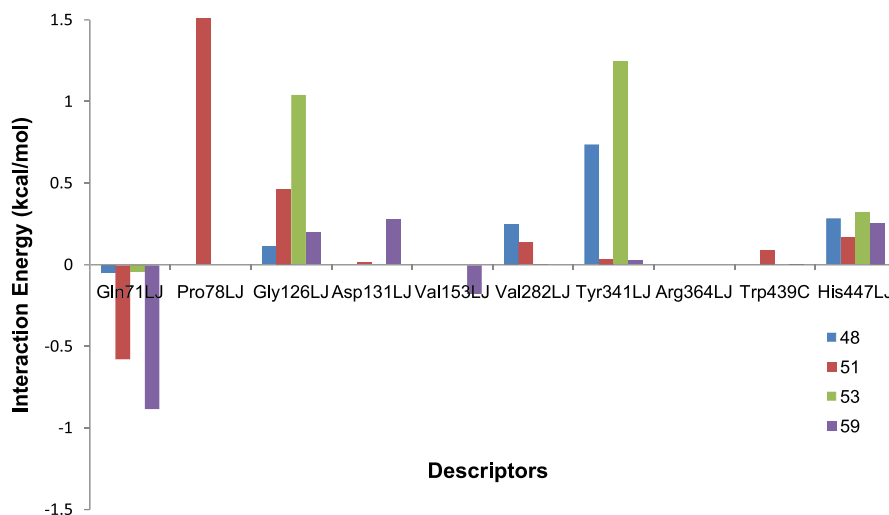

Fig. 5. Product of the descriptors interaction energies and coefficients for outliers of Eq. A–F.
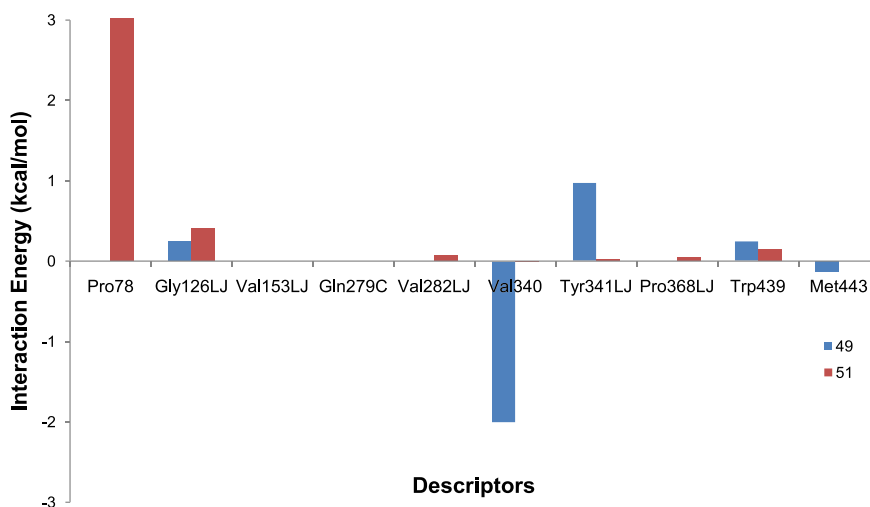
Fig. 6. Product of the descriptors interaction energies and coefficients for outliers of Eq. C–F.

3.6. Prediction of the biological potency (pIC$_{50}$) of an external benzylpiperidine analogue

The theoretical prediction of the biological potency (pIC$_{50}$) for E2020 (numbered as compound **61**, Table 4), which was not included into the training or test sets, was performed using the best RD-3D-QSAR equations (Eq. A–F and Eq. C–F). Table 4 shows the experimental (pIC$_{50Exp}$) and predicted (pIC$_{50Pred}$) potency values for compound **61** according to equations A–F and C–F, compared to the values of the most (**33**) and the least (**45**) potent benzylpiperidine analogues from the training set. The E2020 inhibitor was modeled as the (R) enantiomer, because according to the X-ray diffraction experiment it was soaked as a racemate, but only the (R) form seems to bind to the enzyme (TcAChE) [42].

The experimental biological potency for this benzylpiperidine derivative assayed in HuAChE is reported with different values such as pIC$_{50}$ = 7.74 M (IC$_{50}$ = 18 nM) [65] and pIC$_{50}$ = 8.10 M (IC$_{50}$ = 8 nM) [35], but both are in the range from 5.59 to 9.48 M of biological potency of training and test sets. We should note that the experimental pIC$_{50}$ values reported for E2020 do not mention its stereochemistry (absolute configuration), and probably are derived from the racemate (rac) (i.e. 50:50 mixture of enantiomers R and S). Moreover, the literature does not report the experimental potency of the pure (R) and (S) enantiomers of E2020, probably because the enantiomers are interconverted in the biological medium [66]. Therefore, we calculated hypothetical pIC$_{50Exp}$ values for the R enantiomer, using the pIC$_{50Exp}$ from the racemate [35,65], considering two hypotheses: the R and S enantiomers are equipotent and the R enantiomer is the eutomer.

As seen in Table 4, the predicted potency values for compound **61** were underestimated in all equations, with low residual values, where **61** does not show an outlier behavior. Comparing each of the residual values for compound **61** with the residual values relative to compounds **33** and **45**, we may conclude that the E2020 was correctly predicted, even for the largest residue of compound **61** (residue = 1.44), according to Eq. C–F. It is worth highlight that the good results of prediction for E2020 were not expected, because this compound is the only one in the entire series that shows the indenone ring, a feature that models were not trained to recognize.

3.7. Design of new benzylpiperidine derivatives

In this work we proposed four new compounds with structural profile similar to those of the 60 benzylpiperidine derivatives. The

design of these new structures was based on the results for the two benzylpiperidine analogues **33** and **61**. Compound **33**, evaluated as the most potent in this series, was well predicted by all RD-3D-QSAR models. Likewise, compound **61**, an anti-Alzheimer drug well established in market and a benzylpiperidine inhibitor of AChE, also shows good prediction results. The main strategies of molecular modification in design of these new benzylpiperidine analogues were molecular hybridization, bioisosterism, molecular simplification and introduction of substituent groups (Fig. 7).

The benzylpiperidine group present in both analogues (**33** and **61**) was maintained in the new compounds (**Ia, Ib, IIa,** and **IIb**) since it characterizes the class of benzylpiperidine derivatives. In addition, we also maintained the methylene spacer in position 4 of

Table 4
Experimental (pIC$_{50Exp}$) and predicted (pIC$_{50Pred}$) activity and residual (pIC$_{50Exp}$ − pIC$_{50Pred}$) values for compounds **61** (E2020), **33** and **45**, according to equations A–F and C–F.

#	pIC$_{50Exp}$	pIC$_{50Pred}$ (residue)	
		Eq. A–F	Eq. C–F
33	9.48	8.67 (0.81)	9.15 (0.33)
45	5.59	6.08 (−0.49)	5.49 (0.10)
61 (rac) [a]	7.74[b]/8.10[c]	nd [d]	nd
	7.44 [f]	7.24 (0.20)	6.96 (0.48)
61 (R) [e]	8.05 [g]	7.24 (0.81)	6.96 (1.09)
	7.80 [h]	7.24 (0.56)	6.96 (0.84)
	8.40 [i]	7.24 (1.16)	6.96 (1.44)

[a] Experimental pIC$_{50}$ values reported for E2020 do not mention its stereochemistry (absolute configuration), and probably are derived from the racemate (rac).
[b] Clark, J. K. et al. Quaternary salts of E2020 analogues as acetylcholinesterase inhibitors for the reversal of neuromuscular block. *Bioorganic & Medicinal Chemistry Letters* **2002** 12, 2565–2568.
[c] Villalobos, A. et al. Novel Benzisoxazole Derivatives as Potent and Selective Inhibitors of Acetylcholinesterase. *Journal of Medicinal Chemistry* **1994** 37, 2721–2734.
[d] Not determined.
[e] E2020 modeled as R enantiomer according to reference 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.
[f] Hypothetical pIC$_{50Exp}$ value considering R and S enantiomers as equipotent, calculated from pIC$_{50Exp}$ of reference b.
[g] Hypothetical pIC$_{50Exp}$ value considering R enantiomer as the eutomer, calculated from pIC$_{50Exp}$ of reference b.
[h] Hypothetical pIC$_{50Exp}$ value considering R and S enantiomers as equipotent, calculated from pIC$_{50Exp}$ of reference c.
[i] Hypothetical pIC$_{50Exp}$ value considering R enantiomer as the eutomer, calculated from pIC$_{50Exp}$ of reference c.
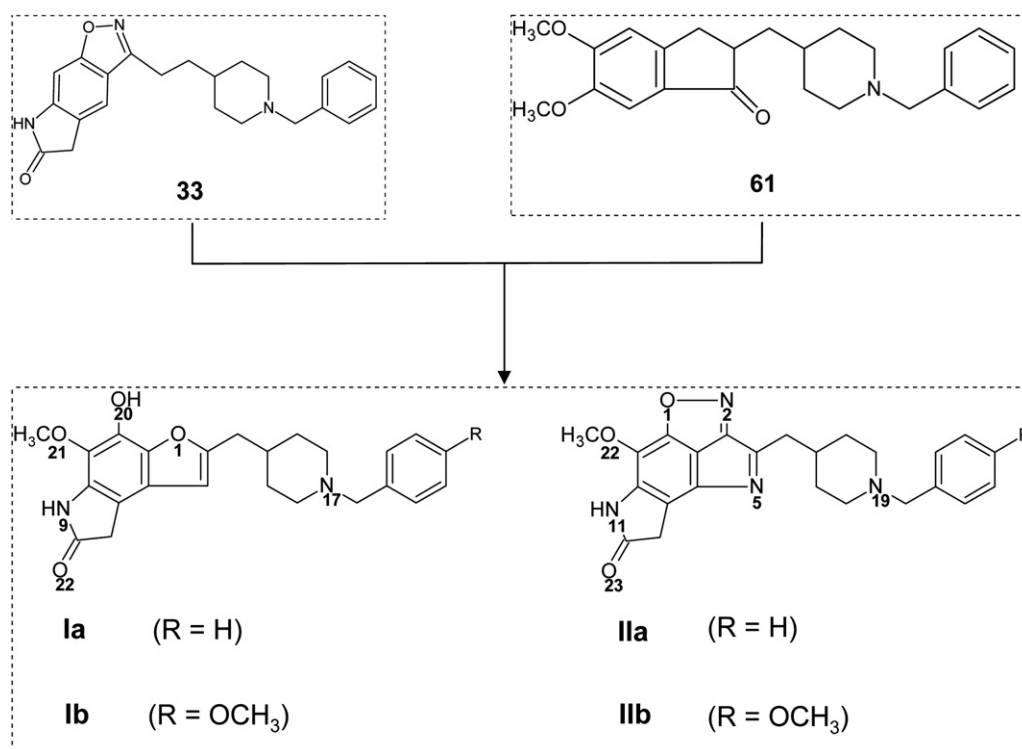
Fig. 7. Design of the new benzylpiperidine analogues by molecular modification strategies.

piperidine ring. In contrast, the indenone ring of **61** was replaced by benzofuran (**Ia** and **Ib**) and indol (**IIa** and **IIb**) bioisosteres, abolishing the stereogenic center of **61**, and consequently avoiding racemic mixtures on the synthesis of these new compounds.

The 5-methoxyl substituent of the indenone ring (**61**) was maintained in all compounds to maximize the hydrophobic interactions with aromatic residues (*i.e.*, Tyr72, Tyr124, and Trp286), since there is an available space (~5.0 Å) between these residues and C8 of **33** (the corresponding position of 5-OMe group of **61**).

The pyrrolidinone ring of **33** was also preserved in all compounds, while both the isoxazol and pyrrolidinone rings of **33** were maintained in compounds **IIa** and **IIb**. Compounds **Ib** and **IIb**, corresponds to compounds **Ia** and **IIa**, respectively, by the introduction of a methoxyl group as the *para*-benzyl group substituent (Fig. 7). This strategy should enable the formation of a potential hydrogen bond or hydrophobic interaction with the catalytic histidine (His447) that is distant about 3.8 Å from the benzyl ring of **33**.

3.8. Prediction of the biological potency (pIC$_{50}$) of the new benzylpiperidine derivatives

Each of the proposed compounds was evaluated by the best RD-3D-QSAR models in this work (Eq. A–F and Eq. C–F), and its biological potencies were predicted (Table 5). Considering both equations simultaneously, the overall order of predicted potency is the same: **IIa** > **Ib** > **IIb** > **Ia**. Comparing series **I** and **II**, the pattern of methoxylation of the benzyl ring *para*-position is different, since in series **I**, the 4-OMe derivative **Ib** is more potent than the non-methoxylated derivative **Ia**, while in series **II**, the non-methoxylated derivative **IIa** is more potent than the 4-OMe derivative **Ia**.

The analysis of the non-methoxylated derivatives showed the presence of indol and isoxazol rings in **IIa** as crucial to increase the predicted potency compared to **Ia** that do not present any of these structural features. On the other hand, these features lose their importance for the methoxylated derivatives as the methoxyl

group seems to induce a rearrangement of these ligands in the AChE active site.

3.9. The analysis of the proposed benzylpiperidine derivative IIa into the HuAChE binding site

Despite all derivatives had potential structural and pharmaceutical drug features according to the Molinspiration [67] and Osiris [68] servers, we selected derivative **IIa** as the most suitable candidate to be a new HuAChE inhibitor. Fig. 8 shows the potential interactions (*e.g.*, hydrogen bond, hydrophobic, and van der Waals) that **IIa** makes with the residues of the catalytic anionic site (Tyr337), peripheral anionic site (Tyr72, Asp74, Trp286, and Glu292), and catalytic triad (His447), together with the potential hydrogen bond interactions with other residues (Asn283 and Tyr341) of the AChE active site.

Four hydrogen bond interactions were detected between **IIa** and HuAChE: (i) **IIa**-O23 and Tyr72-NH (O···N distance = 2.70 Å); (ii) **IIa**-O22 and Asn283-NH$_2$ (O···N distance = 3.53 Å); (iii) **IIa**-N2 and Glu292-CO$_2$ (N77···O distance = 2.85 Å); and (iv) **IIa**-O1 and Tyr341-OH (O···O distance = 2.86 Å). Three hydrophobic interactions were detected between the **IIa** and Trp286 (PAS), Tyr337 (PAS) and His447 (catalytic triad) residues besides a van der Waals interaction between **IIa** and the Asp74 (PAS). Among the residues that make direct contact with **IIa**, only Tyr341 and His447 are

Table 5
Predicted values of pIC$_{50}$ (M) for the benzylpiperidine derivatives **Ia**, **Ib**, **IIa** and **IIb** according to the best 3D-QSAR equations.

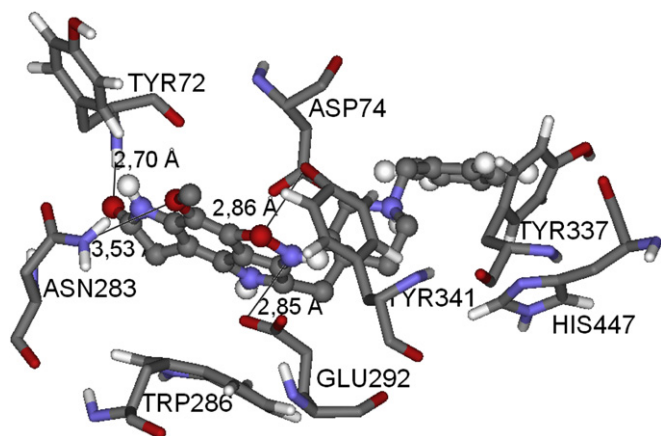| #	Eq. A–F	pIC$_{50Pred}$	Eq. C–F
Ia	6.16		7.33
Ib	9.03		8.01
IIa	16.36		10.86
IIb	7.13		7.73

Fig. 8. Graphic representation of potential interactions (hydrogen bond, hydrophobic, and van der Waals) between IIa (ball-and-stick model and color by element) and residues (stick model and color by element) of catalytic anionic site (Tyr337), peripheral anionic site (Tyr72, Asp74, Trp286, and Glu292), and catalytic triad (His447). The other AChE active site residues showed are relative to potential hydrogen bond interactions.

related to terms **Tyr341LJ** and **His447LJ** of Eq. A–F and **Tyr341LJ** of Eq. C–F. Hence, Eq. A–F should be more relevant than Eq. C–F to explain the predicted activity of derivative **IIa** ($pIC_{50Pred} = 16.36$), since there are two terms selected in this equation that make interaction with **IIa**.

Thus, we pointed the derivative **IIa** for further organic synthesis and pharmacological evaluations due to the larger potency predicted by equations A–F and C–F, good calculated structural and pharmaceutical features by Osiris and Molinpiration servers and lack of stereogenic center.

4. Conclusions

In this work we built and evaluated receptor-dependent 3D-QSAR models of a series of 60 benzylpiperidine inhibitors of HuAChE (compounds **1**–**47** as the training set and **48**–**60** as the test set). These models were obtained from 12 databases (DB) clustered in three groups, which were submitted to a combined genetic algorithm and partial least square approach, available in the WOLF program, to obtain the QSAR equations. The independent variables in this new RD-3D-QSAR study are the steric (Lennard-Jones) and electrostatic (Coulomb) interaction energies, calculated between each ligand and the HuAChE residues within radii of 10 Å around the ligand.

The best two models are named Eq. A–F ($q^2_{adj} = 0.672$, $q^2 = 0.736$, $SE_{cv} = 0.454$, $r^2 = 0.860$, SEE = 0.788, outliers = **48, 51, 53**, and **59**) and Eq. C–F ($q^2_{adj} = 0.693$, $q^2 = 0.753$, $SE_{cv} = 0.431$, $r^2 = 0.900$, SEE = 1.196, outliers = **49** and **51**). According to both models, the Lennard-Jones and also the sum of Lennard-Jones and Coulomb contributions are more important than the Coulomb ones to the structure–activity relationships. These data suggest that the hydrophobic residues of the active site of HuAChE are more important in the interaction with this series of inhibitors when compared to polar residues. Comparing both models, two residues from the aromatic gorge (Tyr341 and Trp439) are related to terms **Tyr341LJ** and **Trp439C** of Eq. A–F and terms **Tyr341LJ** and **Trp439C** of Eq. C–F, while one residue from the catalytic triad (His447) is related to term **His447LJ** of Eq. A–F, showing the consistency of these models with the structure–activity relationships. As an additional external validation test, the theoretical prediction of the biological potency (pIC_{50}) for E2020 (**61**), a benzylpiperidine analog modeled as the (R) enantiomer, was performed using the best RD-3D-QSAR models.

Finally, four new inhibitors of HuAChE were proposed using the data from this RD-3D-QSAR study, and drug design strategies of molecular modification. These new benzylpiperidine derivatives (**Ia**, **Ib**, **IIa**, and **IIb**) preserve crucial structural features of two analogues, **33** and **61**, that had good results in this study. The prediction of biological potency (pIC_{50}) for each of these derivatives was performed using the best two RD-3D-QSAR models. The good results of predicted potency for derivative **IIa**, together with the calculated structural and pharmaceutical features by Osiris and Molinspiration servers pointed this benzylpiperidine derivative as a potential new HuAChE inhibitor candidate.

Appendix. Supporting information

Supplementary information associated with this article can be found in online version at doi:10.1016/j.ejmech.2010.10.009.

References

[1] M.D.M. Alcala, N.M. Vivas, S. Hospital, P. Camps, D. Munoz-Torrero, A. Badia, Characterisation of the anticholinesterase activity of two new tacrine-huperzine A hybrids, Neuropharmacology 44 (2003) 749–755.
[2] N.R. Cutler, J.J. Sramek, Review of the next generation of Alzheimer's disease therapeutics: Challenges for drug development, Prog. Neuro-Psychopharmacol. Biol. Psychiatry 25 (2001) 27–57.
[3] S.D. Ginsberg, S. Che, S.E. Counts, E.J. Mufson, Single Cell gene expression profiling in Alzheimer's disease, Neurorx 3 (2006) 302–318.
[4] L. Bertram, R. Busch, M. Spiegl, N.T. Lautenschlager, U. Muller, A. Kurz, Paternal age is a risk factor for Alzheimer disease in the absence of a major gene, Neurogenetics 1 (1998) 277–280.
[5] M.S. Tsai, E.G. Tangalos, R.C. Petersen, G.E. Smith, D.J. Schaid, E. Kokmen, R.J. Ivnik, S.N. Thibodeau, Apolipoprotein-E - Risk factor for Alzheimer-disease, Am. J. Hum. Genet. 54 (1994) 643–649.
[6] S. Fleminger, Head injury as a risk factor for Alzheimer's disease, J. Neurol. Neurosurg. Psychiatry 74 (2003) 832.
[7] I.J. Martins, E. Hone, J.K. Foster, S.I. Sunram-Lea, A. Gnjec, S.J. Fuller, D. Nolan, S.E. Gandy, R.N. Martins, Apolipoprotein E, cholesterol metabolism, diabetes, and the convergence of risk factors for Alzheimer's disease and cardiovascular disease, Mol. Psychiatry 11 (2006) 721–736.
[8] M. Kivipelto, M.P. Laakso, J. Tuomilehto, A. Nissinen, H. Soininen, Hypertension and hypercholesterolaemia as risk factors for Alzheimer's disease - Potential for pharmacological intervention, CNS Drugs 16 (2002) 435–444.
[9] M. Kivipelto, A. Solomon, Cholesterol as a risk factor for Alzheimer's disease - epidemiological evidence, Acta. Neurol. Scand. 114 (2006) 50–57.
[10] J.C. De la Torre, How do heart disease and stroke become risk factors for Alzheimer's disease, Neurol. Res. 28 (2006) 637–644.
[11] K.J. Anstey, C. von Sanden, A. Salim, R. O'Kearney, Smoking as a risk factor for dementia and cognitive decline: a meta-analysis of prospective studies, Am. J. Epidemiol. 166 (2007) 367–378.
[12] I. Mebane-Sims, 2009 Alzheimer's disease facts and figures, Alzheimers Demen. 5 (2009) 234–270.
[13] R. Schmidt, E. Assem-Hilger, T. Benke, P. Dal-Bianco, M. Delazer, G. Ladurner, K. Jellinger, J. Marksteiner, G. Ransmayr, H. Schmidt, E. Stogmann, J. Wancata, C. Wehringer, Sex differences in Alzheimer disease, Neuropsychiatrie 22 (2008) 1–15.
[14] L.E. Hebert, P.A. Scherr, J.J. McCann, L.A. Beckett, D.A. Evans, Is the risk of developing Alzheimer's disease greater for women than for men? Am. J. Epidemiol. 153 (2001) 132–136.
[15] M.B. Patwardhan, D.C. McCrory, D.B. Matchar, G.P. Samsa, O.T. Rutschmann, Alzheimer disease: operating characteristics of PET– A meta-Analysis, Radiology 273 (2004) 73–80.
[16] R.C. Mohs, The clinical syndrome of Alzheimer's disease: aspects particularly relevant to clinical trials, Genes Brain Behav. 4 (2005) 129–133.
[17] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, H.M. Arrighi, Forecasting the global burden of Alzheimer's disease, Alzheimer's Demen. 3 (2007) 186–191.
[18] E. Scarpini, P. Scheltens, H. Feldman, Treatment of Alzheimer's disease: current status and new perspectives, Lancet Neurol. 2 (2003) 539–547.
[19] R.T. Bartus, R.L. Dean, B. Beer, A.S. Lippa, The cholinergic hypothesis of Geriatric memory Dysfunction, Science 217 (1982) 408–417.
[20] P.T. Francis, A.M. Palmer, M. Snape, G.K. Wilcock, The cholinergic hypothesis of Alzheimer's disease: a review of progress, J. Neurol. Neurosurg. Psychiatry 66 (1999) 137–147.

[21] J.L. Cummings, D. Kaufer, Neuropsychiatric aspects of Alzheimer's disease: the cholinergic hypothesis revisited, Neurology 47 (1996) 876–883.
[22] E.K. Perry, The cholinergic hypothesis - 10 Years on, Br. Med. Bull. 42 (1986) 63–69.
[23] A. Lleo, S.M. Greenberg, J.H. Growdon, Current pharmacotherapy for Alzheimer's disease, Annu. Rev. Med. 57 (2006) 513–533.
[24] Alzheimer's disease: FDA's role in new product development. www.fda.gov/NewsEvents/Testimony/ucm110879.htm (Access date May 2010).
[25] M.R. Farlow, J.L. Cummings, Effective pharmacologic management of Alzheimer's disease, Am. J. Med. 120 (2007) 388–397.
[26] J.L. Cummings, Cholinesterase inhibitors: a new class of psychotropic compounds, Am. J. Psychiatry 157 (2000) 4–15.
[27] S.L. Rogers, M.R. Farlow, R.S. Doody, R. Mohs, L.T. Friedhoff, B. Albala, B. Baumel, G. Booker, J. Dexter, M. Farmer, J.P. Feighner, S. Ferris, B. Gordon, D.G. Gorman, G. Hanna, L.E. Harrell, R. Hubbard, J. Kennedy, J. McCarthy, D.W. Scharre, F. Schaerf, L. Schneider, B. Seltzer, A. Siegal, S.R. Stark, A. Strauss, T.M. Walshe, A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease, Neurology 50 (1998) 136–145.
[28] P.N. Tariot, P.R. Solomon, J.C. Morris, P. Kershaw, S. Lilienfeld, C. Ding, A 5-month, randomized, placebo-controlled trial of galantamine in AD, Neurology 54 (2000) 2269–2276.
[29] M.A. Findeis, The role of amyloid beta peptide 42 in Alzheimer's disease, Pharmacol. Ther. 116 (2007) 266–286.
[30] L.L. Shen, G.X. Liu, Y. Tang, Molecular docking and 3D-QSAR studies of 2-substituted 1-indanone derivatives as acetylcholinesterase inhibitors, Acta. Pharmacol. Sin 28 (2007) 2053–2063.
[31] J. Marco-Contelles, M.D. Carreiras, C. Rodriguez, M. Villarroya, A.G. Garcia, Synthesis and pharmacology of galantamine, Chem. Rev. 106 (2006) 116–133.
[32] H. Sugimoto, Y. Yamanishi, Y. Iimura, Y. Kawakami, Donepezil hydrochloride (E2020) and other acetylcholinesterase inhibitors, Curr. Med. Chem. 7 (2000) 303–339.
[33] Y.C. Xu, J.P. Colletier, M. Weik, H.L. Jiang, J. Moult, I. Silman, J.L. Sussman, Flexibility of aromatic residues in the active-site gorge of acetylcholinesterase: X-ray versus molecular dynamics, Biophys. J. 95 (2008) 2500–2511.
[34] G. Kryger, I. Silman, J.L. Sussman, Three-dimensional structure of a complex of E2020 with acetylcholinesterase from Torpedo californica, J. Physiol. Paris 92 (1998) 191–194.
[35] A. Villalobos, J.F. Blake, C.K. Biggers, T.W. Butler, D.S. Chapin, Y.P.L. Chen, J.L. Ives, S.B. Jones, D.R. Liston, A.A. Nagel, D.M. Nason, J.A. Nielsen, I.A. Shalaby, W.F. White, Novel Benzisoxazole derivatives as potent and selective inhibitors of acetylcholinesterase, J. Med. Chem. 37 (1994) 2721–2734.
[36] A.A. Nagel, D.R. Liston, S. Jung, M. Mahar, L.A. Vincent, D. Chapin, Y.L. Chen, S. Hubbard, J.L. Ives, S.B. Jones, J.A. Nielsen, A. Ramirez, I.A. Shalaby, A. Villalobos, W.F. White, Design and synthesis of l-Heteroaryl-3-(1-benzyl-4-piperidinyl)propan-1-one derivatives as potent, selective acetylcholinesterase inhibitors, J. Med. Chem. 38 (1995) 1084–1089.
[37] Y. Ishichi, M. Sasaki, M. Setoh, T. Tsukamoto, S. Miwatashi, H. Nagabukuro, S. Okanishi, S. Imai, R. Saikawa, T. Doi, Y. Ishihara, Novel acetylcholinesterase inhibitor as increasing agent on rhythmic bladder contractions: SAR of 8-{3-[1-(3-fluorobenzyl)piperidin-4-yl]propanoyl}-1,2,5,6-tetrahydro-4H- pyrrolo [3,2,1-ij]quinolin-4-one (TAK-802) and related compounds, Bioorg. Med. Chem. 13 (2005) 1901–1911.
[38] A. Villalobos, T.W. Butler, D.S. Chapin, Y.L. Chen, S.B. Demattos, J.L. Ives, S.B. Jones, D.R. Listen, A.A. Nagel, D.M. Nason, J.A. Nielsen, A.D. Ramirez, I.A. Shalaby, W.F. White, 5,7-Dihydro-3-[2-[1-(phenylmethyl)-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. 38 (1995) 2802–2808.
[39] G.L. Ellman, K.D. Courtney, V. Andres, R.M. Featherstone, New, Rapid Colorimetric, Determination of acetylcholinesterase activity, Biochem. Pharmacol. 7 (1961) 88.
[40] H. Kubinyi, Hansch Analysis and Related Approaches. VCH, New York, 1993.
[41] Spartan `06. Wavefunction, Inc, 18401 Von Karman Avenue, Suite 370 Irvine, CA 92612, USA, 2006.
[42] G. Kryger, I. Silman, J.L. Sussman, Structure of acetylcholinesterase complexed with E2020 (Aricept ®): implications for the design of new anti-Alzheimer drugs, Structure 7 (1999) 297–307.
[43] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, J. Am. Chem. Soc. 107 (1985) 3902–3909.
[44] G. Kryger, M. Harel, K. Giles, L. Toker, B. Velan, A. Lazar, C. Kronman, D. Barak, N. Ariel, A. Shafferman, I. Silman, J.L. Sussman, Structures of recombinant native and E202Q mutant human acetylcholinesterase complexed with the snake-venom toxin fasciculin-II, Acta Crystallogr. Sect. D: Biol. Crystallogr. 56 (2000) 1385–1394.
[45] A. Sali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[46] C. Notredame, D.G. Higgins, J. Heringa, T-Coffee: a novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 (2000) 205–217.
[47] G.N. Ramachandran, V. Sasisekharan, Conformation of Polypeptides and Proteins, Adv. Protein Chem. 23 (1968) 283–437.
[48] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck - A program to Check the Stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[49] Sybyl V.8.0. Tripos Inc, 1699 S. Hanley Rd., St. Louis, MO, USA, 2008.
[50] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (2008) 435–447.
[51] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: The GROMOS96 Manual and User Guide, VdF: Hochschulverlag AG an der ETH Zürich and BIOMOS b.v. Zürich, Groningen, 1996.
[52] D.M.F. vanAalten, R. Bywater, J.B.C. Findlay, M. Hendlich, R.W.W. Hooft, G. Vriend, PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules, J. Comput-Aided. Mol. Des 10 (1996) 255–262.
[53] H.J.C. Berendsen, J.P.M. Postma, W.F. Vangunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[54] B. Hess, H. Bekker, H.J.C. Berendsen, J. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[55] L. Verlet, Computer experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones molecules, Phys. Rev. 159 (1967) 98–103.
[56] T. Darden, D. York, L. Pedersen, Particle Mesh Ewald - an N.Log(N) method for Ewald sums in large Systems, J. Chem. Phys. 98 (1993) 10089–10092.
[57] A.J. Hopfinger, S. Wang, J.S. Tokarski, B.Q. Jin, M. Albuquerque, P.J. Madhav, C. Duraiswami, Construction of 3D-QSAR models using the 4D-QSAR analysis formalism, J. Am. Chem. Soc. 119 (1997) 10509–10524.
[58] N.C. Romeiro, M.G. Albuquerque, R.B. de Alencastro, M. Ravi, A.J. Hopfinger, Free-energy force-field three-dimensional quantitative structure-activity relationship analysis of a set of p38-mitogen activated protein kinase inhibitors, J. Mol. Model. 12 (2006) 855–868.
[59] D. Rogers, A.J. Hopfinger, Application of genetic function approximation to quantitative structure-activity-relationships and quantitative structure-property relationships, J. Chem. Inf. Comput. Sci. 34 (1994) 854–866.
[60] W.G. Glen, W.J. Dunn, D.R. Scott, Principal components analysis and partial least squares regression, Tetrahedron Comput. Methodol. 2 (1989) 349–376.
[61] H. Kubinyi, QSAR and 3D QSAR in drug design .1. methodology, Drug Discov. Today 2 (1997) 457–467.
[62] D. Rogers, Some Theory and examples of genetic function approximation with comparison to Evolutionary techniques. in: D. James (Ed.), Genetic Algorithms in Molecular Modeling. Academic Press, London, 1996, pp. 87–107.
[63] D. Livingstone, Data Analysis for Chemists. Oxford University Press, New York, 1995.
[64] D. Jaiswal, C. Karthikeyan, P. Trivedi, Rationalization of Physicochemical Properties of Alkanoic acid derivatives towards Histone Deacetylase inhibition, Internet Electron. J. Mol. Des. 5 (2006) 13–26.
[65] J.K. Clark, P. Cowley, A.W. Muir, R. Palin, E. Pow, A.B. Prosser, R. Taylor, M.Q. Zhang, Quaternary salts of E2020 analogues as acetylcholinesterase inhibitors for the reversal of neuromuscular block, Bioorg. Med. Chem. Lett. 12 (2002) 2565–2568.
[66] K. Matsui, Y. Oda, H. Ohe, S. Tanaka, N. Asakawa, Direct determination of E2020 enantiomers in plasma by liquid-chromatography mass-spectrometry and column-switching techniques, J. Chromatogr. A 694 (1995) 209–218.
[67] Molinspiration property calculation Service. http://www.molinspiration.com/ (Access date April 2010).
[68] Thomas Sander, Osiris Property Explorer. http://www.organic-chemistry.org/prog/peo/ (Access date April 2010).