Molecular docking and 3D-QSAR studies on the binding mechanism of statine-based peptidomimetics with β-secretase

Zhili Zuo, Xiaomin Luo,* Weiliang Zhu,* Jianhua Shen, Xu Shen,
Hualiang Jiang* and Kaixian Chen

*Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica,
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 555 Zu Chong Zhi Road, Shanghai 201203, China*

Abstract—β-Secretase is an important protease in the pathogenesis of Alzheimer's disease. Some statine-based peptidomimetics show inhibitory activities to the β-secretase. To explore the inhibitory mechanism, molecular docking and three-dimensional quantitative structure–activity relationship (3D-QSAR) studies on these analogues were performed. The Lamarckian Genetic Algorithm (LGA) was applied to locate the binding orientations and conformations of the peptidomimetics with the β-secretase. A good correlation between the calculated binding free energies and the experimental inhibitory activities suggests that the identified binding conformations of these potential inhibitors are reliable. Based on the binding conformations, highly predictive 3D-QSAR models were developed with q^2 values of 0.582 and 0.622 for CoMFA and CoMSIA, respectively. The predictive abilities of these models were validated by some compounds that were not included in the training set. Furthermore, the 3D-QSAR models were mapped back to the binding site of the β-secretase, to get a better understanding of vital interactions between the statine-based peptidomimetics and the protease. Both the CoMFA and the CoMSIA field distributions are in well agreement with the structural characteristics of the binding groove of the β-secretase. Therefore, the final 3D-QSAR models and the information of the inhibitor–enzyme interaction would be useful in developing new drug leads against Alzheimer's disease.
© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

Alzheimer's disease (AD) is the most common form of dementia and accounts for two thirds of all cases. It destroys brain cells and nerves, disrupting the transmitters, which carry messages in the brain, particularly those responsible for storing memories. As a result of an aging society, the prevalence of the disease is increasing dramatically. Nowadays, it has been a major social and financial burden for both the society and family.[1] Cure for this disease is currently unavailable, thus, investigations on new drug discovery and development for AD are of importance.

It has been found that the cerebral deposition of a 40–42-residue β-amyloid peptide (Aβ) is an early and critical feature in AD,[2] indicating a key role of Aβ in the pathogenesis of AD.[3] Accordingly, a hypothesis was proposed that overproduction of the 42-aminoacid form of Aβ might lead to the increased aggregation and deposition of Aβ as a senile plaques in the brain.[4–6] Aβ is generated from the endoproteolytic processing of the amyloid precursor protein (APP), which is a type I membrane protein with 770 amino acids in length. The endoproteolytic processing involves sequential actions of two proteases, the β-secretase and γ-secretase. Firstly, the β-secretase cleaves APP to generate the N-terminus of Aβ, and then, the γ-secretase cleaves APP to generate the C-terminus, leading to release of the Aβ.[7] It was followed that preventing the bioactivity of the β-secretase or γ-secretase could be therapeutically useful in the treatment of AD. But until 1999 the identity of neither the β-secretase nor γ-secretase was known. The β-secretase, also known as BACE (β-site APP-cleaving enzyme), which is the protease responding the β-site APP-cleaving, was independently identified by several laboratories using different approaches in 1999,[8–12] while the γ-secretase has not been conclusively identified. The identification of the β-secretase as BACE provided a target for novel therapies for Alzheimer's disease.[13]

Apparently, the interaction mechanism of the β-secretase with its inhibitors would be greatly helpful in discovering small molecule inhibitors for ceasing the function of the β-secretase. To investigate the mechanism, some three-dimensional structures of the β-secretase itself and its complex with different inhibitors (e.g., Lol-Alq based peptidomimetic inhibitors) have been experimentally determined.[14–17] These crystal structures provided not only insights into the interaction mechanisms of the β-secretase with the inhibitors, but also valuable clues for designing new inhibitors.[18–20] For example, crystal structures of the β-secretase complexed with Lol-Alq based peptidomimetic inhibitors OM00-3 and OM99-2 (PDB entries 1M4H and 1FKN) revealed eight interaction sub-domains of the β-secretase with the peptides analogues.[14,15] Based on the structural information, some new potential peptidomimetic inhibitors designed from structure-based modification have been reported.[18] Hom and co-workers designed and synthesized a series of statine-based peptidomimetic analogues, their inhibitory activities to the β-secretase were also measured.[20,21] However, as listed in Table 1, the IC_{50} of the most active inhibitor (**23**), a statine-based peptidomimetic inhibitor, is 0.3 μM, while the K_i of OM99-2 is as potent as 1.58 nM, calling for discovering new statine-based peptidomimetic inhibitors with higher bioactivities.[18] As shown in Table 1, most of the statine-based peptidomimetic inhibitors are very long structures and are too hydrophilic, resulting in difficulties in penetrating the blood–brain barrier (BBB). Thus, further structure based drug design study is needed to discover new statine-based peptidomimetic inhibitors that are more drug-like and more active. To the best of our knowledge, neither crystal structures of statine-based inhibitors in complex with the β-secretase nor a 3D-QSAR model are available. Thus, we carried out a molecular docking and 3D-QSAR study using comparative molecular field analysis (CoMFA)[22] and comparative molecular similarity indices analysis (CoMSIA)[23] methodologies. The results from this study should be useful in understanding the inhibitory mode of the β-secretase and in designing new drug leads against AD.

2. Computational methods

2.1. Molecular structures

Thirty-two statine-based peptidomimetic inhibitors with a common skeleton of Sta-Val synthesized by Hom et al. were employed in this study (Table 1). The three-dimensional (3D) structure of each peptidomimetic analogue was constructed based on the structure of OM99-2 extracted from its β-secretase complex crystal structure (PDB entry 1FKN) from Protein Databank (PDB)[24] using SYBYL 6.8,[25] followed by an energy minimization to a convergence of 0.05 kcal/mol Å using the Tripos force field with Gasteiger–Hückel charges.[26]
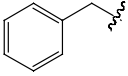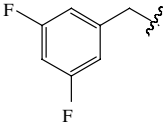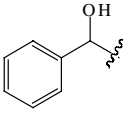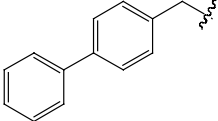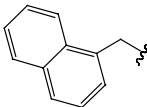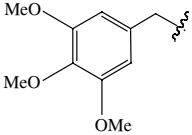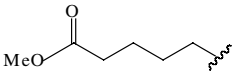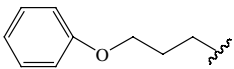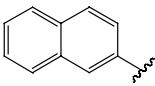
2.2. Molecular docking

To find the binding mode of the peptidomimetic inhibitors to the β-secretase, the advanced docking program

AUTODOCK 3.0.3[27] was used to automatically dock the ligands to the enzyme. The Lamarckian genetic algorithm (LGA)[28] was applied to deal with the protein–inhibitor interactions. Briefly, the LGA describes the relationship between the inhibitors and the β-secretase by translation, orientation, and conformation of the inhibitors. These so-called 'state variables' are the inhibitors' genotype, and the resulting atomic coordinates together w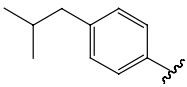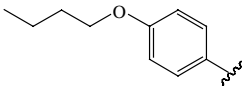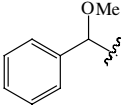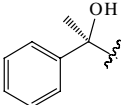ith the interaction and the intra-molecular energies are the inhibitors' phenotype. The environmental adaptation of the phenotype is reverse-transcribed into its genotype and become heritable traits. Each docking cycle, or generation, consists of a regimen of fitness evaluation, crossover, mutation, and selection. A Solis and Wets local search[29] is performed for the energy minimization on a user-specified proportion of the population. The docked structures of the inhibitors are generated after a reasonable number of evaluations.

The whole docking operation in this study could be stated as follows. First, the β-secretase was checked for polar hydrogen and assigned for partial atomic charges, a PDBQ file was then created. Atomic solvation parameters and fragmental volumes were assigned to the β-secretase using the ADDSOL module of the AUTODOCK program. Meanwhile, some of the torsion angles of the inhibitors that would be explored during molecular docking stage were defined, allowing the conformation search for the ligands during the docking process. Second, the grid map with $60 \times 80 \times 60$ points and a spacing of 0.375 Å was calculated using the AUTOGRID program to evaluate the binding energies between the inhibitors and the protease. The affinity and electrostatic potential grid were calculated for each type of atom in the inhibitors. The energetic configuration of a particular ligand was found by tri-linear interpolation of affinity values and electrostatic interaction of the eight grid points around each atom of the ligand. Third, some important parameters for LGA calculations were reasonably set up according to requirements of the Amber force field and our problem. The initial translation was set as initial coordinate, and the initial number of individuals in population is 300. The step size was set to 0.2 Å for translation and 5° for orientation and torsion. The maximum number of generations, energy evaluations, and docking runs were set to 5.0×10^5, 2.5×10^6, and 20, respectively. The elitism value is 1, which automatically survives into next generation. The mutation rate is 0.02, which is a probability that a gene would undergo a random change. The crossover rate, the probability of proportional selection, is 0.80. The pseudo-Solis and Wets local search method with a maximum of 300 iterations per local search was used. The probability of performing local search on an individual in the population is 0.06. The maximum number of consecutive successes before doubling or halving the local search step size is 4, and the same as failures. The lower bound on F, the termination criterion for the local search, is 0.01. Finally, the docked complexes of the inhibitor–enzyme for each inhibitor were selected according to the criterion of interaction energy combined with geometrical matching quality. These complexes were used as initial structures for further geometrical optimization

Table 1. The structure and in vitro inhibitory activities of the 32 statine-based peptidomimetic analogues

No	R_1	R_2	IC_{50} (μM)	pIC_{50}[a]	ΔG (kcal/mol)
R_1-CO-[Sta]-Val-R_2					
1*		-Ala-Glu-Phe	130	3.89	−15.26
2		-Ala-Glu-Phe	140	3.85	−15.41
3		-Ala-Glu-Phe	17	4.77	−16.65
4		-Ala-Glu-Phe	56	4.25	−15.30
5		-Ala-Glu-Phe	42	4.38	−17.26
6		-Ala-Glu-Phe	94	4.03	−17.08
7		-Ala-Glu-Phe	50	4.30	−15.38
8*		-Ala-Glu-Phe	25	4.60	−15.30
9		-Ala-Glu-Phe	48	4.32	−16.84
10		-Ala-Glu-Phe	28	4.55	−16.79
11		-Ala-Glu-Phe	27	4.57	−15.83
12		-Ala-Glu-Phe	28	4.55	−15.57
13		-Ala-Glu-Phe	41	4.39	−17.34

Table 1 (*continued*)
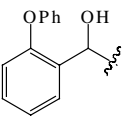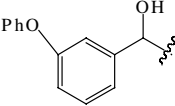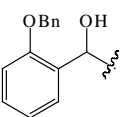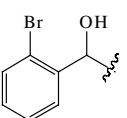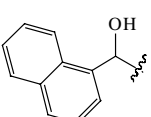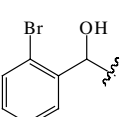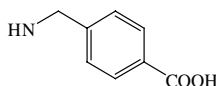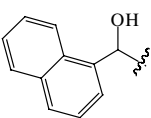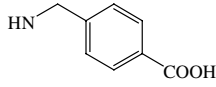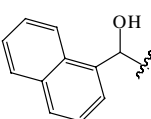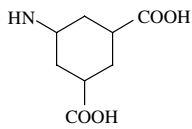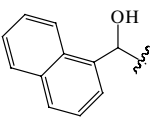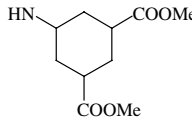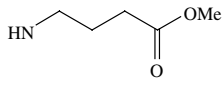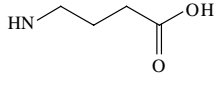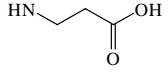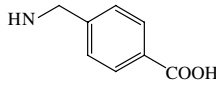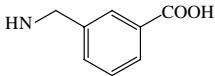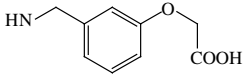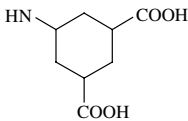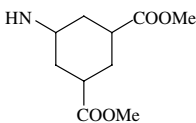
No	R_1	R_2	IC_{50} (µM)	pIC_{50}[a]	ΔG (kcal/mol)
14		-Ala-Glu-Phe	14	4.85	−16.61
15		-Ala-Glu-Phe	6	5.22	−17.37
16		-Ala-Glu-Phe	11	4.96	−15.58
17		-Ala-Glu-Phe	3	5.52	−17.56
18		-Ala-Glu-Phe	4	5.40	−18.20
19*			40	4.44	−17.24
20			50	4.30	−15.21
21			3	5.52	−18.09
22			140	3.85	−16.67

R_1-Val-Met-[Sta]-Val-R_2

No	R_1	R_2	IC_{50} (µM)	pIC_{50}[a]	ΔG (kcal/mol)
23	Ac–	-Ala-Glu-Phe	0.3	6.52	−20.83
24	$(CH_3)_3COOC–$	OH	94	4.03	−15.47
25	$(CH_3)_3COOC–$		47	4.33	−17.16
26	$(CH_3)_3COOC–$		17	4.77	−16.89
27	$(CH_3)_3COOC–$		30	4.52	−17.17
28*	$(CH_3)_3COOC–$		4	5.40	−19.87

Table 1 (*continued*)

No	R$_1$	R$_2$	IC$_{50}$ (μM)	pIC$_{50}$[a]	ΔG (kcal/mol)
29	(CH$_3$)$_3$COOC–		4	5.40	−18.64
30	(CH$_3$)$_3$COOC–		5	5.30	−18.79
31	(CH$_3$)$_3$COOC–		0.3	6.52	−19.90
32	(CH$_3$)$_3$COOC–		10	5.00	−19.13

*Inhibitors of testing set.
[a] pIC$_{50}$ = −log IC$_{50}$.

to discover the best binding conformation for each inhibitor.

The geometrical optimizations were carried out using molecular mechanics method encoded in SYBYL 6.8[24] with following parameters: the Amber all atom charge of the Amber 4.0 force field[30] and Powell method with the root-mean square (RMS) energy gradient of 0.05 kcal/(mol Å). The whole system was minimized to convergence. Although the solvation energies could not be explicitly considered during the minimization, the energy calculations were performed with a distance-dependent dielectric constant of 5 to mimic the solvation effect of the inhibitors in the protein environment.[31]

2.3. Binding free energy prediction

The new scoring function for predicting binding free energy encoded in AUTODOCK 3.0.3[26–28] was applied to evaluate the binding affinities between the enzyme and the 32 inhibitors. This scoring function was developed based on the traditional molecular force field model of interaction energy. Not only the restriction of internal rotors, the global rotation, and the translation were modeled depending on the number of torsion angles of the ligand but also the desolvation upon binding and the hydrophobic effect (solvent entropy changes at solute–solvent interfaces) were calculated. The total binding free energy was estimated based on the above-stated terms and a set of coefficient factors. Simulations on several systems in our laboratory demonstrated that this scoring function was sufficient to rank the inhibitors in the different levels of binding affinities.[32–38]

2.4. Structural alignment

Twenty conformations were obtained through AUTO-DOCK 3.0.3 for each ligand. The conformation with the strongest predicted binding affinity to the β-secretase was extracted from the optimized inhibitor–β-secretase

complex. These conformations were aligned together inside the binding pocket of β-secretase and used directly for CoMFA and CoMSIA analyses to explore the specific contributions of electrostatic, steric, and hydrophobic effects to the molecular bioactivities.

2.5. CoMFA

Steric and electrostatic interactions were calculated using the Tripos force field[39] with a distance-dependent dielectric constant at all intersections in a regularly spaced (2 Å) grid taking an sp^3 carbon atom as steric probe and a +1 charge as electrostatic probe. The cutoff was set to 30 kcal/mol. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated partial least squares (PLS)[40] method (leave one out). The minimum-sigma (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation was below this threshold. The final model, noncross-validated conventional analysis, was developed with the optimum number of components to yield a noncross-validated r^2 value.

2.6. CoMSIA

In CoMSIA, a distance-dependent Gaussian-type physicochemical property has been adopted to avoid singularities at the atomic positions and dramatic changes of potential energy for those grids in the proximity of the surface. With the standard parameters and no arbitrary cutoff limits, three physicochemical properties, that is, steric, electrostatic, and hydrophobic fields, were calculated. The steric contribution was reflected by the third power of the atomic radii of the atoms. Electrostatic properties were introduced as atomic charges resulted from molecular docking. An atom-based hydrophobicity was assigned according to the parameterization developed by Ghose et al.[41] The lattice dimensions were selected with a sufficiently large margin
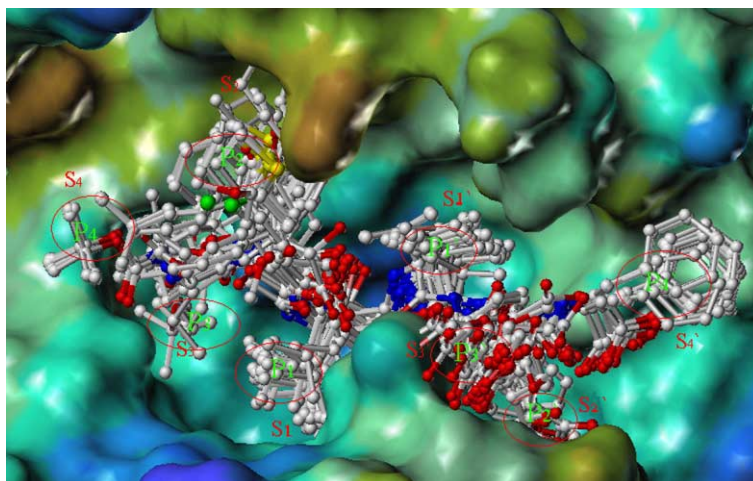
Figure 1. The binding conformations of the statine-based compounds displayed inside the active site of the β-secretase. This image was generated with the MOLCAD program in SYBYL 6.8,[24] with some residues removed for clear visualization. The β-secretase surface was rendered with electrostatic potential.

(>4 Å) to enclose all the binding conformations of the inhibitors determined by AUTODOCK 3.0.3. In general, similarity indices, $A_{F,K}$ between the compounds of interest were computed by placing a probe atom at the intersections of the lattice points using Eq. 1,

$$A_{F,K}^q(j) = -\sum_{i=1}^{n} W_{probe,k} W_{ik} e^{-a r_{iq}^2} \quad (1)$$

where q represents a grid point; i is the summation index over all atoms of the molecule j under computation; w_{ik} is the actual value of the physicochemical property k of atom i; and $w_{probe,k}$ is the value of the probe atom. In the present study, similarity indices were computed using a probe atom ($w_{probe,k}$) with charge +1, radius 1 Å, hydrophobicity +1, and attenuation factor α of 0.3 for the Gaussian type distance. The statistical evaluation for the CoMSIA analyses was performed in the same way as described for CoMFA.

3. Results and discussion

3.1. Binding modes between the statine-based inhibitors and the β-secretase

Figure 1 depicts the binding conformations of the statine-based peptidomimetics in the binding pocket of the β-secretase, which were derived from the docking simulations followed by energy minimizations. It is clearly that all the substructures (P_1, P_2, P_3, P_1', P_2', and P_3') of the peptidomimetics have very similar binding characteristic (Fig. 1). For example, substructure P_1 of all the inhibitors always binds to the same subsite S_1 of the β-secretase, interacting with residues of Leu30, Asp32, Phe108, Trp115, and Ile118 in the binding pocket.

3.2. Correlation between binding free energy and inhibitory activity

The predicted binding free energies (ΔGs) of the peptidomimetics to the β-secretase and corresponding experi-

mental $pIC_{50}(-\log IC_{50})$ values are listed in Table 1. A correlation was found between ΔGs and the pIC_{50}s via a linear regression analysis, as shown in Eq. 2 and Figure 2, the correlation coefficient (r^2) is 0.647. This rather good correlation demonstrates that the binding conformations and binding models of the peptidomimetic inhibitors to the β-secretase are reasonable.

$$\Delta G = -9.109 - 1.669 \times pIC_{50}$$
$$(N = 28, \ r^2 = 0.647, \ F = 47.627, \ S = 0.872) \quad (2)$$

where S is standard error, r^2 is correlation coefficient and F is testing factor of the reliability.

Furthermore, Eq. 2 indicates that there would be about 1.669 kcal/mol difference in binding free energy if there is one order of magnitude for numerical difference in the inhibitory potency (pIC_{50}). This value (1.669 kcal/mol) is close to the well-known value in theory (coefficient in $\Delta G^\circ = -RT \ln K$), 1.365 kcal/mol, which illuminates that Eq. 2 is accurate and reliable, indicating that we have successfully discovered the most probable binding conformation for each inhibitor. Therefore, they should be more accurate conformations
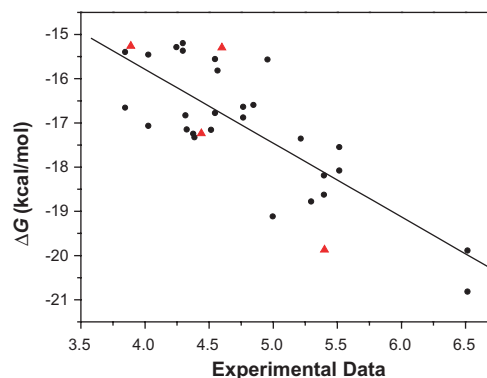

Figure 2. Correction between the predicted binding free energies (ΔGs, kcal/mol, $T = 298.15$ K) and the experimental activities (pIC_{50}). (●) Compounds of the training set; (▲) compounds of the testing set.

for QSAR study than usual ones by manually structural alignment.

3.3. CoMFA models

Twenty-eight of the 32 statine-based peptidomimetic inhibitors were randomly picked up as training set for constructing CoMFA models, the remaining four were used as testing set for the model validation.

PLS analysis was carried out for the 28 binding conformations, and the result is listed in Table 2, which shows that a CoMFA model with a cross-validated q^2 of 0.582 for five components was obtained. The noncross-validated PLS analysis with the optimum components of 5 revealed a conventional r^2 value of 0.986, $F = 307.229$, and an estimated standard error of 0.091. The steric field descriptors explain 48.4% of the variance, while the electrostatic descriptors explain 51.6%. The predicted activities for the 28 inhibitors versus their experimental activities with their residues (δ) are listed in Table 3, the correlation between the predicted activities and the experimental activities is depicted in Figure 3A. Table 3 and Figure 3A demonstrate that the predicted activities by the constructed CoMFA model are in good agreement with the experimental data, suggesting that the CoMFA model should have a satisfactory predictive ability.

Table 2. CoMFA and CoMSIA results

	CoMFA	CoMSIA
PLS statistics		
q^2	0.582	0.622
r^2	0.986	0.982
S	0.091	0.105
F	307.229	191.762
Optimal comp.	5	6
Field distribution (%)		
Steric	48.4	19.2
Electrostatic	51.6	47.5
Hydrophobic		33.3
Testing set		
r^2	0.756	0.853
S	0.237	0.225

Table 3. Experimental activities (pIC$_{50}$s) and predicted activities (PAs) and by CoMFA and CoMSIA

No	pIC$_{50}$	CoMFA		CoMSIA	
		PA	δ	PA	δ
1*	3.89	4.38	−0.49	4.21	−0.32
2	3.85	3.90	−0.05	3.90	−0.05
3	4.77	4.66	0.11	4.60	0.17
4	4.25	4.37	−0.12	4.29	−0.04
5	4.38	4.34	0.04	4.28	0.10
6	4.03	3.99	0.04	4.09	−0.06
7	4.30	4.27	0.03	4.31	−0.01
8*	4.60	4.35	0.25	4.37	0.23
9	4.32	4.46	−0.14	4.41	−0.09
10	4.55	4.56	−0.01	4.58	−0.03
11	4.57	4.53	0.04	4.54	0.03
12	4.55	4.53	0.02	4.61	−0.06
13	4.39	4.46	−0.07	4.51	−0.12
14	4.85	4.78	0.07	4.86	−0.01
15	5.22	5.23	−0.01	5.21	0.01
16	4.96	4.98	−0.02	4.87	0.09
17	5.52	5.45	0.07	5.47	0.05
18	5.40	5.43	−0.03	5.43	−0.03
19*	4.44	4.47	−0.03	4.65	−0.21
20	4.30	4.30	0.00	4.23	0.07
21	5.52	5.40	0.12	5.39	0.13
22	3.85	3.92	−0.07	3.99	−0.14
23	6.52	6.50	0.02	4.06	2.46
24	4.03	4.05	−0.02	4.47	−0.44
25	4.33	4.47	−0.14	4.67	−0.34
26	4.77	4.57	0.20	4.47	0.30
27	4.52	4.50	0.02	4.06	0.46
28*	5.40	5.18	0.22	5.30	0.10
29	5.40	5.44	−0.04	5.34	0.06
30	5.30	5.38	−0.08	5.33	−0.03
31	6.52	6.63	−0.11	6.74	−0.22
32	5.00	4.89	0.11	4.86	0.14

The CoMFA contour plots of steric and electrostatic interactions are shown in Figure 4A. To aid in visualization, the most active compound (**23**) is displayed with green color in the maps. The colored polyhedra in the map show these areas in 3D space where changes in the field values for the peptidomimetics correlate strongly with concomitant changes in inhibitory activities. Detrimental and beneficial steric interactions are respectively displayed in yellow and green contours,
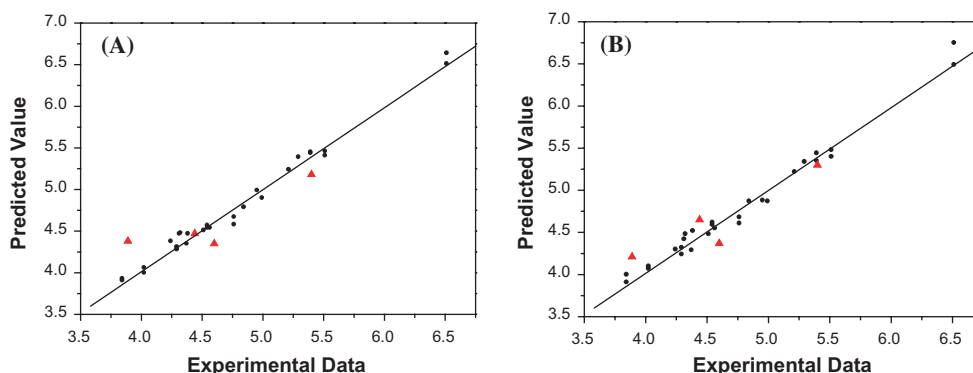

Figure 3. The predicted activities versus the experimental activities (pIC$_{50}$) of the statine-based compounds. (A): CoMFA; (B): CoMSIA. (●) Compounds of the training set; (▲) compounds of the testing set.
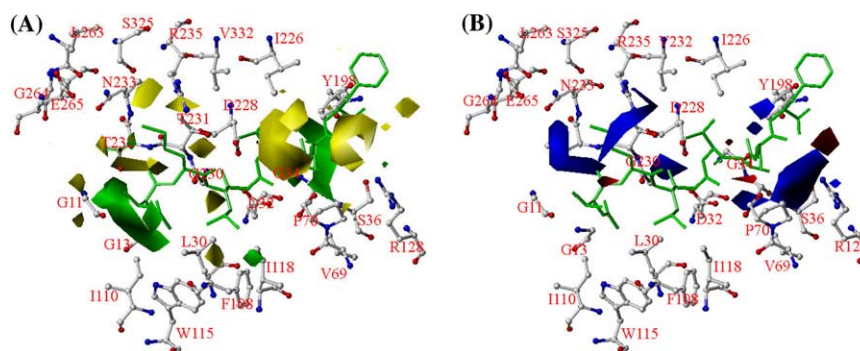
Figure 4. CoMFA contour maps displayed with compound **23** and the key residues in the binding site of the β-secretase. (A) Steric field distribution; and (B) electrostatic field distribution. The residues are represented as ball-and-stick, and the compound is shown in green sticks. Sterically favored areas are in green; sterically unfavored areas are in yellow; positive potential favored areas are in blue; positive potential unfavored areas are in red.

while blue and red contours illustrate the regions of desirable positive and negative electrostatic interactions.

Two big regions of green contour near the P_3 substructure of inhibitor **23** suggest that more bulky substituents in these positions could significantly improve the molecular biological activities. Exactly, compounds **14**–**16** exhibit considerable increases in their binding affinities due to the increased bulky substituents around the P_3 substructure (Table 1). There is also a big region of green contour near the P'_2 substructure of compound **23**, indicating that more bulky substituents are needed in this position to improve the inhibitory activity. This is in agreement with the fact that the inhibitory activities of compounds **28** and **29** with more bulky substituents in P'_2 are higher than those of compounds **26** and **27**. The prediction was confirmed by the molecular surface of the binding pocket of the β-secretase, which was shown in Figure 1 that S_3 and S'_2 subsites of the β-secretase are big enough to accommodate more bulky substituents. Some yellow polyhedra near P_2, P_4, P'_3 groups of inhibitor **23** indicate that increased steric bulk in these areas is unfavorable for the inhibitory activities. Figure 4A shows that the residues Thr232, Asn233, Ser325, Arg235, and Val332 in the binding pocket of the β-secretase are in the distance of less than 3.0 Å to P_2 substructure of the inhibitor. Therefore, any larger substitute may lead to collision with these residues in the pocket.

The CoMFA electrostatic contours, in which show blue-colored regions where an increased positive charge is favorable for inhibitory activity and red-colored regions where an increased negative charge is favorable for the activity, are shown in Figure 4B. Three big blue polyhedra near the P'_3 substructure of inhibitor **23** indicate that the positively charged substituents around these areas may increase molecular inhibitory activity. Figure 4B shows that the S'_3 subsite of the β-secretase is composed of the residues Tyr198, Tyr71, Ser36, and Ser35 that have a hydroxyl on each sidechain. Thus, the positive substituents should strengthen the binding of the inhibitors to the S'_3 subsite of the β-secretase. Another region of blue contour between the P_1 and P'_1 substructure of inhibitor **23** suggests another area for more positively charged substituents. As indicated in Figure 4B, the key catalytic negative residues of Asp32 and Asp228

are rather near this region, favoring a strong electrostatic interaction with the positively charge substitute group of the inhibitor. One more blue polyhedron was found between the P_3 and P_4 substructures of inhibitor **23**, which was supported by the crystal structure that there are negatively charged polar residues between the subsites of S_3 and S_4, namely, Asn233, Thr232, and Thr231. A red polyhedron around P'_2 and P'_4 of inhibitor **23** (Fig. 4B) indicates that more negative substituents are favorable in this region to improving the inhibitor's activity. This prediction is in agreement with the appearance of a positively charged residue of Arg128 at the corresponding position of the β-secretase. There is another red region of contour between the P_2 and P_4 substructures, suggesting that the inhibitor with a more negative substituent in this area would increase its inhibitory potency. An example is inhibitor **3**, which has an additional hydroxyl there, thus, much more active than inhibitor **1** (Table 1).

3.4. CoMSIA

The CoMSIA study revealed a cross-validated q^2 of 0.622 with an optimum component number of 6, a conventional r^2 of 0.982 with a standard error of 0.105 and $F = 191.762$ (Table 2). The steric field descriptor explains 19.2% of the variance; and the electrostatic descriptor is 47.5%, while the hydrophobic field explains the rest 33.3%. The total field distribution of the steric and hydrophobic fields is 52.5%, which is approximately equal to the steric field contribution in CoMFA (48.4%). This indicated that CoMSIA and CoMFA models are in consistent with each other. The predicted inhibitory activities are also listed in Tables 2 and 3. The correlation between the experimental and predicted bioactivities is shown in Figure 3B. All the results demonstrate that the CoMSIA model is also fairly predictive.

The electrostatic and steric contour maps from the CoMSIA analysis (Fig. 5A and B) are generally in accordance with the field distributions of CoMFA maps (Fig. 4A and B). Thus, the discussion of the CoMSIA result would be focused on the hydrophobic interactions only.

Magenta contours refer to areas where the hydrophobic substituents are favorable to improve inhibitor
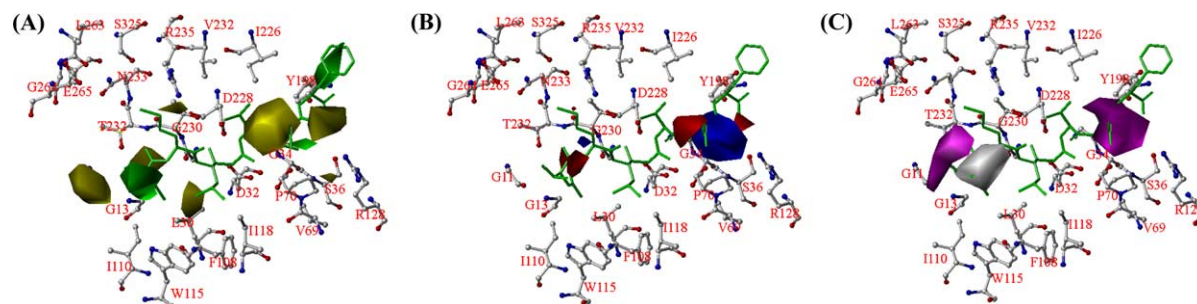
Figure 5. CoMSIA contour maps displayed with the compound **23** and the key residues in the active site of the β-secretase. (A) steric field distribution; (B) electrostatic field distribution; and (C) hydrophobic field distribution. The residues are represented as sticks, and the ligand is shown in ball-and-stick and colored in green. Sterically favored areas are in green; sterically unfavored areas are in yellow; positive potential favored areas are in blue; positive potential unfavored areas are in red; hydrophobic favored areas are in magenta; hydrophilic favored areas are in white.

activities; while the white contours indicate the area where hydrophilic substituents are favorable to increase the inhibitory activity of these peptidomimetics. As shown in Figure 5C, there is one white-colored polyhedron that is in between the P_2 and P_3 substructures, which means that more hydrophilic substituents in the area is beneficial to increase molecular bioactivity. There are two magenta polyhedra near the P_4 and P'_3 substructures of inhibitor **23**, respectively. The magenta-colored regions indicate that the hydrophobic substituents are needed near the substructures to improve molecular bioactivity. It is noticed that the residues Gly11, Leu263, Gly264, Glu265, and Ala323 are near the S_4 subsite, and that Pro70, Val69, and Tyr71 are close to the S'_3 subsite of the β-secretase, leading to strong hydrophobic interaction with the inhibitors in these regions.

3.5. Validation of the 3D-QSAR models

The four randomly selected compounds (compounds **1***, **8***, **19***, and **28*** in Table 1) were used as testing set to verify the constructed CoMFA and CoMSIA models. The calculated results are also listed in Table 3 (labeled with asterisk), and displayed in Figure 3 (in triangle). The predicted pIC_{50} with the QSAR models are in good agreement with the experimental data within a statistically tolerable error range, with a correlation coefficient of $r^2 = 0.756$ and 0.853 for CoMFA and CoMSIA models, respectively (Table 2). The testing results indicate that the CoMFA and CoMSIA models would be reliably used in new statine-based peptidomimetic inhibitor design for developing drug leads against AD.

To further validate the CoMFA and CoMSIA models, the software LIGPLOT[42] was employed to study the hydrophobic and hydrogen bonding interactions between the inhibitors and the β-secretase. As indicated in two-dimensional representative of the interacting model of inhibitor **23** with the β-secretase (Fig. 6), there are hydrophobic interactions between the residue Pro70 and the P'_3 substructure of inhibitor **23**, which is in agreement with the CoMSIA model in Figure 5C that there is a big magenta-colored polyhedron near the P'_3 substructure, suggesting again that more hydrophobic interactions around this area should improve molecular inhibitory activities. There are four hydrogen bonds formed between inhibitor **23** and the residues Gly11
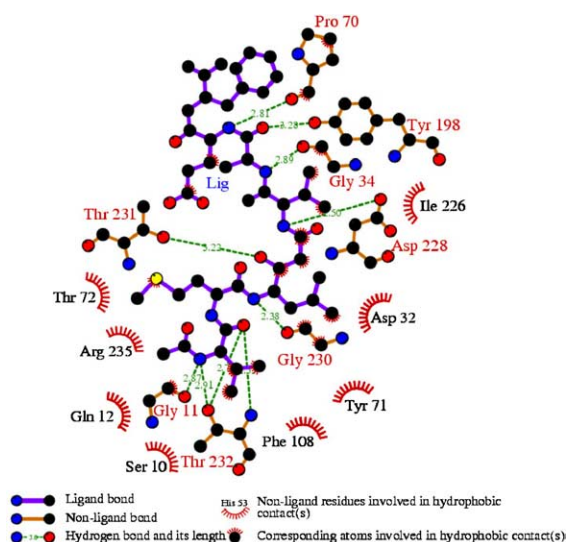

Figure 6. Two-dimensional representative of the interaction model of inhibitor **23** with the β-secretase, drawn by LIGPLOT.[42] The distance between the donor and acceptor of less than 3.4 Å is considered as a hydrogen bond, and a 4.1 Å distance between two hydrophobic atoms is considered to be a hydrophobic interaction.

and Thr232 (Fig. 6), supporting the CoMSIA conclusion that more hydrophilic substituents to inhibitors in this area should increase molecular bioactivity. The hydrogen bonds between inhibitor **23** and residues Gly34, Tyr198, Pro70, Asp228, and Thr231 (Fig. 6) are a good evidence to support the CoMFA conclusion in Figure 4B that the more positively charged substituents near the P'_3 substructure and the area between the P_1 and P'_1 substructures of inhibitor **23** may increase the molecular bioactivity. Thus, the 2D representative of the interaction model of inhibitor **23** with the β-secretase, drawn by LIGPLOT, supports the findings from the 3D-QSAR studies.

4. Conclusions

In this work, molecular docking and 3D-QSAR studies were carried out to explore the binding mechanism of statine-based peptidomimetic inhibitors to the β-secretase, and to construct highly predictive 3D-QSAR

models for designing new β-secretase inhibitors for the treatment of Alzheimer's disease. Both the binding conformations of 32 statine-based peptidomimetic molecules and their binding free energies were determined and predicted by molecular docking. The binding models of the inhibitors show clearly the mechanism of how the peptidomimetic compounds bind to the β-secretase. The binding free energies of these compounds to the β-secretase, estimated by LGA algorithm, were found to have a good correlation with the experimental inhibitory potencies. Based on the binding conformations from molecular docking, highly predictive CoMFA and CoMSIA models were developed. These models match well the 3D topology of the binding site of the β-secretase. The reliability of the models was verified by the compounds in the testing set. The 3D-QSAR results revealed some important sites, where steric, electrostatic and hydrophobic modifications should significantly affect compounds' bioactivities, thus, useful clues to designing novel inhibitors of β-secretase with high capabilities to penetrate the blood–brain barrier for the treatment of Alzheimer's disease.

Acknowledgements

References and notes

1. Roggo, S. *Curr. Top. Med. Chem.* **2002**, *2*, 359.
2. Wolfe, M. S.; Xia, W.; Ostaszewski, B. L.; Diehl, T. S.; Kimberly, W. T.; Selkoe, D. J. *Nature* **1999**, *398*, 513.
3. Hardy, J.; Allsop, D. *Trends Pharmacol. Sci.* **1991**, *12*, 383.
4. Sherrington, R.; Rogaev, E. I.; Liang, Y.; Rogaeva, E. A.; Levesque, G.; Ikeda, M.; Chi, H.; Lin, C.; Li, G.; Holman, K. *Nature* **1995**, *375*, 754.
5. Levy-Lahad, E.; Wasco, W.; Poorkaj, P.; Romano, D. M.; Oshima, J.; Pettingell, W. H.; Yu, C. E.; Jondro, P. D.; Schmidt, S. D.; Wang, K. *Science* **1995**, *269*, 973.
6. Rogaev, E. I.; Sherrington, R.; Rogaeva, E. A.; Levesque, G.; Ikeda, M.; Liang, Y.; Chi, H.; Lin, C.; Holman, K.; Tsuda, T. *Nature* **1995**, *376*, 775.
7. Walter, J.; Kaether, C.; Steiner, H.; Haass, C. *Curr. Opin. Neurobiol.* **2001**, *11*, 585.
8. Lin, X.; Koelsch, G.; Wu, S.; Downs, D.; Dashti, A.; Tang, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1456.
9. Hussain, I.; Powell, D.; Howlett, D. R.; Tew, D. G.; Meek, T. D.; Chapman, C.; Gloger, I. S.; Murphy, K. E.; Southan, C. D.; Ryan, D. M.; Smith, T. S.; Simmons, D. L.; Walsh, F. S.; Dingwall, C.; Christie, G. *Mol. Cell. Neurosci.* **1999**, *14*, 419.
10. Yan, R.; Bienkowski, M. J.; Shuck, M. E.; Miao, H.; Tory, M. C.; Pauley, A. M.; Brashier, J. R.; Stratman, N. C.; Mathews, W. R.; Buhl, A. E.; Carter, D. B.; Tomasselli, A. G.; Parodi, L. A.; Heinrikson, R. L.; Gurney, M. E. *Nature* **1999**, *402*, 533.
11. Vassar, R.; Bennett, B. D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E. A.; Denis, P.; Teplow, D. B.; Ross, S.; Amarante, P.; Loeloff, R.; Luo, Y.; Fisher, S.; Fuller, J.; Edenson, S.; Lile, J.; Jarosinski, M. A.; Biere, A. L.; Curran, E.; Burgess, T.; Louis, J. C.; Collins, F.; Treanor, J.; Rogers, G.; Citron, M. *Science* **1999**, *286*, 735.
12. Sinha, S.; Anderson, J. P.; Barbour, R.; Basi, G. S.; Caccavello, R.; Davis, D.; Doan, M.; Dovey, H. F.; Frigon, N.; Hong, J.; Jacobson-Croak, K.; Jewett, N.; Keim, P.; Knops, J.; Lieberburg, I.; Power, M.; Tan, H.; Tatsuno, G.; Tung, J.; Schenk, D.; Seubert, P.; Suomensaari, S. M.; Wang, S.; Walker, D.; John, V. *Nature* **1999**, *402*, 537.
13. Skovronsky, D. M.; Lee, V. M. *Trends Pharmacol. Sci.* **2000**, *21*, 161.
14. Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A. K.; Zhang, X. C.; Tang, J. *Science* **2000**, *290*, 150.
15. Hong, L.; Turner, R. T., III; Koelsch, G.; Shin, D.; Ghosh, A. K.; Tang, J. *Biochemistry* **2002**, *41*, 10963.
16. Hong, L.; Tang, J. *Biochemistry* **2004**, *43*, 4689.
17. Patel, S.; Vuillard, L.; Cleasby, A.; Murray, C. W.; Yon, J. *J. Mol. Biol.* **2004**, *343*, 407.
18. Ghosh, A. K.; Bilcer, G.; Harwood, C.; Kawahama, R.; Shin, D.; Hussain, K. A.; Hong, L.; Loy, J. A.; Nguyen, C.; Koelsch, G.; Ermolieff, J.; Tang, J. *J. Med. Chem.* **2001**, *44*, 2865.
19. Tung, J. S.; Davis, D. L.; Anderson, J. P.; Walker, D. E.; Mamo, S.; Jewett, N.; Hom, R. K.; Sinha, S.; Thorsett, E. D.; John, V. *J. Med. Chem.* **2002**, *45*, 259.
20. Hom, R. K.; Gailunas, A. F.; Mamo, S.; Fang, L. Y.; Tung, J. S.; Walker, D. E.; Davis, D.; Thorsett, E. D.; Jewett, N. E.; Moon, J. B.; John, V. *J. Med. Chem.* **2004**, *47*, 158.
21. Hom, R. K.; Fang, L. Y.; Mamo, S.; Tung, J. S.; Guinn, A. C.; Walker, D. E.; Davis, D. L.; Gailunas, A. F.; Thorsett, E. D.; Sinha, S.; Knops, J. E.; Jewett, N. E.; Anderson, J. P.; John, V. *J. Med. Chem.* **2003**, *46*, 1799.
22. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
23. Klebe, G.; Abraham, U. *J. Comput. Aided Mol. Des.* **1999**, *13*, 1.
24. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
25. Sybyl Version 6.8; Tripos Associates: St. Louis, MO, 2000.
26. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
27. Morris, G. M.; Goodsell, D. S.; Huey, R.; Hart, W. E.; Halliday, R. S.; Belew, R. K.; Olson, A. J. Autodock Version 3.0.3. The Scripps Research Institute, Molecular Graphics Laboratory, Department of Molecular Biology; 1999.
28. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
29. Solis, F. J.; Wets, R. J. B. *Maths. Pera. Res.* **1981**, *9*, 19.
30. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
31. Mehler, E. L.; Solmajer, T. *Protein Eng.* **1991**, *4*, 903.
32. Xu, Y.; Liu, H.; Niu, C.; Luo, C.; Luo, X.; Shen, J.; Chen, K.; Jiang, H. *Bioorg. Med. Chem.* **2004**, *12*, 6193.

33. Liu, G.; Zhang, Z.; Luo, X.; Shen, J.; Liu, H.; Shen, X.; Chen, K.; Jiang, H. *Bioorg. Med. Chem.* **2004**, *12*, 4147.
34. Yu, C.; Chen, L.; Luo, H.; Chen, J.; Cheng, F.; Gui, C.; Zhang, R.; Shen, J.; Chen, K.; Jiang, H.; Shen, X. *Eur. J. Biochem.* **2004**, *271*, 386.
35. Liu, H.; Li, Y.; Song, M.; Tan, X.; Cheng, F.; Zheng, S.; Shen, J.; Luo, X.; Ji, R.; Yue, J.; Hu, G.; Jiang, H.; Chen, K. *Chem. Biol.* **2003**, *10*, 1103.
36. Liu, H.; Huang, X.; Shen, J.; Luo, X.; Li, M.; Xiong, B.; Chen, G.; Yang, Y.; Jiang, H.; Chen, K. *J. Med. Chem.* **2002**, *45*, 4816.
37. Cui, M.; Huang, X.; Luo, X.; Briggs, J. M.; Ji, R.; Chen, K.; Shen, J.; Jiang, H. *J. Med. Chem.* **2002**, *45*, 5249.
38. Huang, X.; Xu, L.; Luo, X.; Fan, K.; Ji, R.; Pei, G.; Chen, K.; Jiang, H. *J. Med. Chem.* **2002**, *45*, 333.
39. Clark, M. C.; Cramer, R. D., III; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.
40. Bush, B. L.; Nachbar, R. B., Jr. *J. Comput. Aided Mol. Des.* **1993**, *7*, 587.
41. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Combust. Chem.* **1999**, *1*, 55.
42. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127.