J Mol Model (2012) 18:4567–4576
DOI 10.1007/s00894-012-1424-1

ORIGINAL PAPER

Response surface methodology in docking study of small molecule BACE-1 inhibitors

Nima Razzaghi-Asl · Ahmad Ebadi · Najmeh Edraki · Ahmadreza Mehdipour · Sara Shahabipour · Ramin Miri

Received: 19 January 2012 / Accepted: 2 April 2012 / Published online: 13 May 2012

Abstract Computational evaluation of ligand-receptor binding via docking strategy is a well established approach in structure-based drug design. This technique has been applied frequently in developing molecules of biological interest. However, any procedure would require an optimization set up to be more efficient, economic and time-saving. Advantages of modern statistical optimization methods over conventional one-factor-at-a-time studies have been well revealed. The optimization by experimental design provides a combination of factor levels simultaneously satisfying the requirements considered for each of the responses and factors. In this study, response surface method was applied to optimize the prominent factors (number of genetic algorithm runs, population size, maximum number of evaluations, torsion degrees for ligand and number of rotatable bonds in ligand) in AutoDock4.2-based binding study of small molecule β-secretase inhibitors as anti-alzheimer agents. Results revealed that a number of rotatable bonds in ligand and maximum number of docking evaluations were determinant variables affecting docking outputs. The interference between torsion degrees for ligand and number of genetic algorithm runs for docking procedure was found to be the significant interaction term in our model. Optimized docking outputs exhibited a high correlation with experimental fluorescence resonance energy transfer-based IC_{50}s for β-secretase inhibitors (R^2=0.9133).

Keywords Alzheimer · BACE-1 · Box-Behnken · Docking · Optimization

N. Razzaghi-Asl · A. Ebadi · N. Edraki · A. Mehdipour · S. Shahabipour · R. Miri
Medicinal and Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran PO Box 3288–71345

N. Razzaghi-Asl · A. Ebadi · R. Miri (✉)
Department of Medicinal Chemistry, Faculty of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran PO Box 1583–71345
e-mail: mirir@sums.ac.ir

A. Mehdipour
Computational Structural Biology Group, Max-Planck Institute of Biophysics, Max von Laue Strasse 3, Frankfurt am Main 60438, Germany

Introduction

Alzheimer disease (AD) affects a significant number of people all over the world [1]. Amyloid cascade hypothesis has received much attention in development of pharmacologically active anti-alzheimer agents [2]. The hypothesis suggests that aggregation of $A\beta_{40}$ and $A\beta_{42}$ oligopeptides and subsequent formation of neurotoxic polymeric plaques in the brain are significant features common in most patients involved with AD. Proteolytic action on a large transmembrane protein; amyloid precursor protein (APP), by two enzymes namely β and γ-secretases results in secretion of $A\beta_{40}$ and $A\beta_{42}$ peptides [3, 4]. β-secretase (Beta-site APP cleaving enzyme or BACE-1) is a type I membrane-associated aspartyl protease [2], which has been considered to be an attractive therapeutic target in AD due to the following reasons:

- The enzyme catalyzes the first step of Aβ production.
- The enzyme is mainly expressed in the brain.
- The enzyme is selective in its proteolytic activity.

Development of specific inhibitors of this key protease has been regarded as a major therapeutic challenge in AD treatment and many research groups have focused on development of BACE-1 inhibitors [5, 6]. However; the process

of drug development is challenging. It is time consuming, expensive, and requires tedious steps. To overcome these problems to some extent, virtual drug design approaches seem to be trying to provide cost- and time-efficient procedures [7].

In this regard, availability of a significant amount of crystallographic data on Protein Data Bank facilitated the performance of structure based drug discovery projects aiming at BACE-1 as a molecular target for Alzheimer disease (Brookhaven protein databank website. http://www.pdb.org). Docking is an important in silico structure based drug design technique which has been used to recognize correct spatial poses of ligands in the active site of receptors while at the same time predicting the affinity between ligand and protein [8]. In other words, docking describes a process to determine steric fit of ligand and receptor in terms of their geometric complementarily and optimum interaction profile [9].

Molecular docking has contributed much to drug discovery efforts and several drugs have been developed currently on the market [10]. AutoDock is a type of docking programs which has offered several fruitful advantages in drug design field (The AutoDock website. http://autodock.scripps.edu) [11, 12]. Moreover; AutoDock has been the most cited docking program in the literature up to the year 2005 [13].

Docking procedure like any other process would require an optimized condition to be more efficient, economic and time-fluent. There are several practical parameters that should be considered in a typical AutoDock procedure affecting the docking results. Optimization of a process can be done via several ways. Classic optimization method can be performed by varying any one of the process parameters and keeping the other parameters constant. When multiple variables are involved in a system, this technique becomes unproductive and time consuming. The modern statistical designs, namely design of experiments (DOE) consider all factors simultaneously and hence provide the possibility for evaluation of the whole effect all at once.

Response surface methods (RSM) have been designed for factors with more than three levels in which quadratic models can be established [14, 15]. Several prosperous reports of drug related RSM applications can be found in the literature [16–21]. Quantifiable response is one of the most important steps in a typical DOE. The most popular RSMs are Central composite, Box-Behnken and Doehlert designs [22, 23].

In the present study, we aimed to estimate the factor effects on Autodock 4.2-based docking output and optimize the effective factors on docking procedure of BACE-1 inhibitors. For this purpose, Box-Behnken method was applied to investigate the effect of the number of genetic algorithm runs, population size, maximum number of evaluations, torsion degrees for ligand and number of ligand active torsions on docking binding energy. Optimum docking conditions were interpreted. Subsequent regression analysis with known biological data for small molecule BACE-1 inhibitors was performed to validate the optimized docking procedure.

Materials and methods

Docking studies

Flexible-ligand docking studies were done by AutoDock4.2 program [24]. All the pre-processing steps for ligand and receptor crystallographic files were performed within WHAT IF server (European Molecular Laboratory Heidelberg, Germany) and AutoDock Tools 1.5.4 program (ADT) which has been released as an extension suite to the Python Molecular Viewer [24, 25]. All hydrogens were properly added to the receptor PDB file using What if server. ADT program was used to merge non-polar hydrogens into related carbon atoms of the receptor and Kollman charges were also assigned. For docked ligands, non-polar hydrogens were added; Gasteiger charges assigned and torsions degrees of freedom were also allocated by ADT program. Desirable independent genetic algorithm (GA) runs were considered for each ligand under study. For Lamarckian GA method; 27,000 maximum generations; a gene mutation rate of 0.02; and a crossover rate of 0.8 were used. A grid of $60\times60\times60$ points in x, y, and z direction was built centered on the center of mass of the catalytic site of β-secretase. Cluster analysis was performed on the docked results using an RMS tolerance of 2 Å. Chemical structures of the ligands under study with their relevant PDB codes are demonstrated in Fig. 1.

Subsequent validation of optimized docking condition was performed via docking of selected chemical scaffolds (Fig. 2) into active site of β-secretase enzyme.

Experimental design

All statistical analysis, modeling and numerical optimization was performed using Design-Expert software-v.7 (State-Ease, Corp., Minnesota).

Five experimental factors were varied at three levels: number of genetic algorithm runs, population size, maximum number of evaluations, torsions degrees for ligand during docking procedure, and number of ligand rotatable bonds (Table 1). These experimental factors were selected as they were considered to have the most significant effect on the efficiency of the method. The levels were chosen based on knowledge of the system acquired from initial experimental trials. The factors along with their assigned levels were defined as input for Box-Behnken method in DOE program. In our study, Box-Behnken matrix containing 46 solutions (docking protocols) were planned. Top ranked

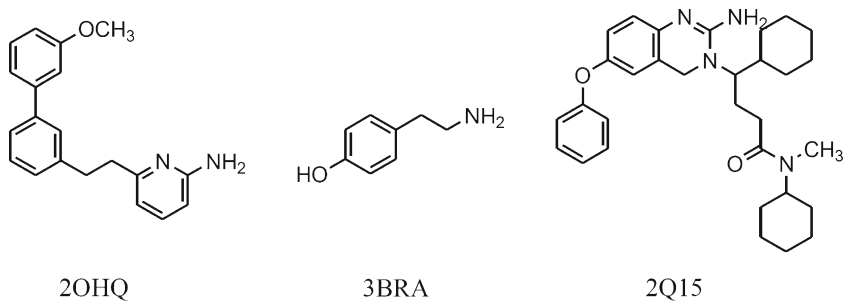

Fig. 1 Chemical structures of selected BACE-1 inhibitors and relevant PDB codes used for RSM based optimization of AutoDock4.2 procedure

ligand-receptor binding energies (kcal.mol^{-1}) in AutoDock dlg output file were considered as response to determine the best combination of docking parameters.

The introduction of central points (docking protocols including factors at their mid levels in the designed matrix) provided a more precise estimate of experimental error and enabled us to measure an adequacy of the model (lack of fit). Analysis of variance (ANOVA) was done on the data to determine the significant variables of the docking procedure.

For more detailed information on Box-Behnken design basis and methodology, readers are referred to the previous publication [26].

Results and discussion

Docking optimization

Knowledge on the optimized condition for docking procedure can be established in a typical validation (self-docking) study [27]. It is a commonly accepted practice to test docking performance on the studied system by re-docking the co-crystallized conformation of a ligand provided that the relevant PDB structure is present. Holo structures of the receptor which bear cognate ligand scaffolds are principal sources for this purpose [28]. Cosconati and coworkers have pointed out several advantages of self-docking techniques such as validation of the target preparation, establishment of docking parameters and validation of the method for prediction of the known binding pose [29].

In the present study, self-docking procedure between BACE-1 inhibitors (Fig. 1) and specified holo β-secretase enzyme structures was optimized by considering five variables under study (Table 1). The factor levels used in our experimental design were selected based on previous knowledge and default values [24, 30]. Each factor was considered in three levels. The assigned levels are given by their actual values in Table 1.

To construct an approximation model that can capture interactions between design variables, a Box-Behnken design matrix was planned to investigate all possible combinations of factors (Table 2) [31]. In our study, Box-Behnken matrix contained 46 solutions. Top ranked binding energies (kcal.mol^{-1}) in AutoDock dlg output file were considered as a response in each run. It should be noted that RMSD values below three could be considered as adoptable values [32].

Box-Behnken design provided fewer runs (46 runs) while similar 3-level full factorial design for five factors included 243 experiments. A comparison between these two statistical

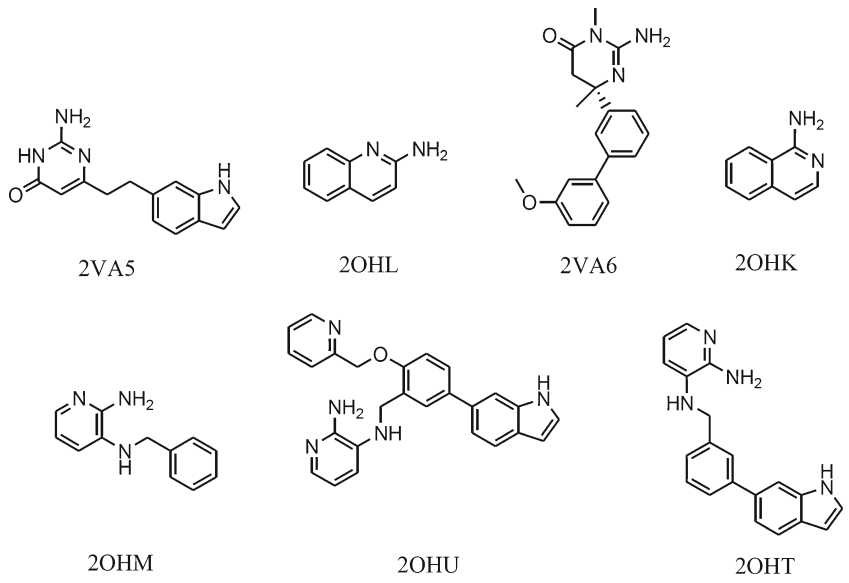

Fig. 2 Chemical structures of selected BACE-1 inhibitors and relevant PDB codes used for validation study of optimized AutoDock4.2 procedure

Table 1 Levels of experimental factors used for AutoDock4.2 based study of BACE-1 inhibitors

Factors under study	Low level	Medium level	High level
GA runs (A)	50	100	150
Population size (B)	50	150	300
Maximum number of evaluations (C)	2.5×10^5	2.5×10^6	1.0×10^7
Torsion degrees for ligand (D)	5	25	50
Ligand RTBs[a] (E)	4 (3BRA)	6 (2OHQ)	10 (2Q15)

[a] Rotatable bonds

methods reveals the Box-Behnken method to be more economic, efficient and time fluent [33]. With the Box-Behnken design methodology, major and interaction effects can be easily evaluated. The major effect refers to the effect caused by the varied factor, while the interaction effect is related to the case in which the effect of one factor is dependent on the value of another [34]. Analysis of variance (ANOVA) was applied to realize the significant factors in the regression model. Analysis of variance for response surface quadratic model is shown in Table 3.

The result of experimentation should be a model which will adequately predict the response within the design space. Regression analysis of the experimental data showed that a quadratic model (values of Prob > F less than 0.0001) could best fit the relationship between the dependent variable (binding energy of ligand to receptor) and independent variable terms. The model F-value of 55.75 implied that the model was significant. In the present study the "lack of fit F-value" of 0.80 implied that the lack of fit was not significant relative to the pure error. Non-significant lack of fit is good and the result of model fitness in terms of regression coefficients are shown in Table 4. The predicted R-squared of 0.9256 was in reasonable agreement with the adjusted R-square of 0.9605. "Adeq precision" measures the signal to noise ratio. A ratio greater than 4 is desirable. Ratio of 27.099 indicated an adequate signal therefore this model could probably be applied to navigate the design space.

In the ANOVA analysis (Table 4), "Prob > F" values less than 0.0500 indicate that the model terms are significant. Values greater than 0.1000 indicate that the model terms cannot be significant. According to the data, three main effects [GA runs (A), maximum number of evaluations (C) and number of rotatable bonds in ligand (E)] along with two second-order main effects (AD and E^2) were significant model terms. However AC and CD interaction terms were not significant but also not insignificant.

Number of active torsions in ligand which is defined by number of rotatable single connections in the molecular framework (factor E) was found to have the largest effect on response. It is worth noting that the defaults given for AutoDock are sufficient for docking systems with ten or less rotatable bonds (The AutoDock website; http://autodock.scripps.edu/). We set levels of RTBs (rotatable bonds) based on this rationalization, while probably many significant effects for factor E would be expected considering higher factor levels. Torsion degree of ligand per step (factor D) showed nearly no meaningful effect on binding process. The order of effective factors could be shown on the priority order of E > C > A > B > D. Significant effects of "maximum number of evaluations" and "GA runs" may be attributed to the opportunity to optimize docking procedure provided by increasing these terms. However in our opinion there should be a logic balance between a docking precision and docking period especially in the conditions where no automated and supercomputer facilities are available. A much larger effect of factor E (number of ligand rotatable bonds) on docking outputs implied that AutoDock program may be extensively sensitive to the number of active torsions in the docked ligand and interpretation of the results for ligands possessing more rotatable bonds should be done with care regarding this posed limitation. Lack of significant sensitivity for factor D (torsion degrees for ligand) may be interpreted by the fact that validation of docking studies are performed using the cognate inhibitors from crystallographic structures, where fitted pose of ligands will initiate the docking procedure (unlike the cross-docking procedure in which an external non-cognate ligand is docked). It is worth noting that ligand torsion degrees per each step in cross-docking procedure may probably be more determinant compared to the case of self-docking study.

An interaction between factors is likely to occur if different responses based on the settings of two factors are generated. These two factors will appear with two non-parallel lines in interaction plot, indicating that the effect of one factor depends on the level of the other. Referring to Fig. 3, AD interaction term versus binding energy can be well interpreted .

A cross point in the interaction plot can be interpreted in two ways; if this point belongs to the unique inhibitor structure processed under 105 GA runs (x: 0.1), different levels of factor D (5 and 50 torsion degrees/step) would produce docking outputs with the same binding energies. In the case of supposing different chemical scaffolds, two levels of factor D would lead to the interaction patterns with the same binding energies.

There is one outlier design point which is located outside the range of interaction lines. This point is related to the 6th run of the designed matrix (GA run: 100, maximum number of evaluation: 2.5×10^6, population size: 150, torsion degrees/step: 25 and RTBs: 6). The starting point for ligand conformation state in 6th docking run by 25 torsion degrees/step directed ligand in a way to produce false positive binding energy (ΔG: -8.71 $Kcal.mol^{-1}$). 3D surface interaction plot expressing significant AD interaction is also depicted for more clarity (Fig. 4).

The perturbation plot of binding energy versus all modeled variables was applied to evaluate the contribution of each factor

Table 2 Box-Behnken design with actual and coded factor levels in AutoDock4.2 based study of BACE-1 inhibitors

Run	GA runs (A)	Population size (B)	Maximum number of evaluations (C)	Torsion degrees for ligand (D)	Ligand RTBs[a] (E)	Estimated binding energy (kcal.mol^{-1})	RMSD (Å)
1	100 (0)[b]	300 (+1)	250000 (−1)	25 (0)	6 (0)	−7.37	1.96
2	100 (0)	50 (−1)	250000 (−1)	25 (0)	6 (0)	−7.2	2.00
3	100 (0)	150 (0)	10000000 (+1)	25 (0)	4 (−1)	−6.27	1.91
4	100 (0)	50 (−1)	2500000 (0)	25 (0)	4 (−1)	−6.3	1.85
5	100 (0)	150 (0)	250000 (−1)	5 (−1)	6 (0)	−6.8	2.20
6	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−8.71	2.11
7	50 (−1)	150 (0)	10000000 (+1)	25 (0)	6 (0)	−8.8	1.85
8	100 (0)	150 (0)	10000000 (+1)	50 (+1)	6 (0)	−7.91	1.81
9	150 (+1)	150 (0)	2500000 (0)	50 (+1)	6 (0)	−7.75	1.86
10	50 (−1)	150 (0)	2500000 (0)	50 (+1)	6 (0)	−8.79	1.91
11	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−8.09	1.70
12	150 (+1)	150 (0)	2500000 (0)	25 (0)	4 (−1)	−6.2	1.83
13	50 (−1)	150 (0)	2500000 (0)	5 (−1)	6 (0)	−6.72	1.86
14	100 (0)	50 (−1)	2500000 (0)	5 (−1)	6 (0)	−8.83	0.48
15	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−7.7	1.54
16	100 (0)	300 (+1)	2500000 (0)	5 (−1)	6 (0)	−7.6	0.77
17	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−7.71	0.74
18	150 (+1)	50 (−1)	2500000 (0)	25 (0)	6 (0)	−8.68	0.66
19	100 (0)	150 (0)	10000000 (+1)	5 (−1)	6 (0)	−8.82	0.54
20	100 (0)	150 (0)	2500000 (0)	5 (−1)	1 0 (+1)	−11.89	2.76
21	100 (0)	50 (−1)	10000000 (+1)	25 (0)	1 0 (+1)	−12.1	0.54
22	100 (0)	300 (+1)	2500000 (0)	25 (0)	4 (−1)	−6.19	1.93
23	150 (+1)	150 (0)	2500000 (0)	25 (0)	1 0 (+1)	−11.9	0.64
24	150 (+1)	150 (0)	2500000 (0)	25 (0)	10 (+1)	−7.32	3.00
25	100 (0)	150 (0)	250000 (−1)	50 (+1)	6 (0)	−8.06	2.87
26	150 (+1)	150 (0)	250000 (−1)	25 (0)	6 (0)	−5.87	1.88
27	100 (0)	150 (0)	250000 (−1)	25 (0)	4 (−1)	−8.49	3.00
28	150 (+1)	150 (0)	2500000 (0)	5 (−1)	6 (0)	−8.55	2.51
29	100 (0)	50 (−1)	10000000 (+1)	25 (0)	6 (0)	−5.95	1.97
30	50 (−1)	150 (0)	2500000 (0)	25 (0)	4 (−1)	−5.97	1.97
31	100 (0)	150 (0)	2500000 (0)	5 (−1)	4 (−1)	−11.8	0.79
32	100 (0)	150(0)	250000 (−1)	25 (0)	1 0 (+1)	−8.44	1.68
33	150 (+1)	150(0)	10000000 (+1)	25 (0)	6 (0)	−7.48	1.13
34	50 (−1)	50 (−1)	2500000 (0)	25 (0)	6 (0)	−8.4	1.27
35	150 (+1)	300 (+1)	2500000 (0)	25 (0)	6 (0)	−12	0.62
36	50 (−1)	150 (0)	2500000 (0)	25 (0)	10 (+1)	−5.97	1.92
37	100 (0)	150 (0)	2500000 (0)	50 (+1)	4 (−1)	−8.55	2.12
38	100 (0)	300 (+1)	10000000 (+1)	25 (0)	6 (0)	−7.23	2.46
39	50 (−1)	300 (+1)	2500000 (0)	25 (0)	6 (0)	−8.5	3.00
40	100 (0)	50 (−1)	2500000 (0)	50 (+1)	6 (0)	−6.89	2.09
41	50 (−1)	150 (0)	10000000 (+1)	25 (0)	6 (0)	−11.88	2.69
42	100 (0)	150 (0)	2500000 (0)	50 (+1)	1 0 (+1)	−8.33	1.77
43	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−8.21	1.65
44	100 (0)	300 (+1)	2500000 (0)	50 (+1)	6 (0)	−8.44	2.33
45	100 (0)	150 (0)	2500000 (0)	25 (0)	6 (0)	−11.77	0.78
46	100 (0)	150 (0)	10000000 (+1)	25 (0)	1 0 (+1)	−12.04	0.64

[a] Rotatable bonds

[b] Numbers in parentheses indicate the coded levels of factors under study

Table 3 ANOVA for quadratic model in AutoDock4.2 based response surface study of BACE-1 inhibitors

Source	Sum of squares	DF*	Mean square	F value	p-value Prob>F
Model	155.65	20	7.78	55.75	<0.0001
Factor					
A	1.03	1	1.03	7.38	0.0118
B	0.34	1	0.34	2.41	0.1332
C	4.07	1	4.07	29.16	<0.0001
D	0.092	1	0.092	0.66	0.4258
E	136.07	1	136.07	974.69	<0.0001
AB	2.25E-004	1	2.25E-004	1.612E-003	0.9683
AC	0.59	1	0.59	4.19	0.0513
AD	1.97	1	1.97	14.14	0.0009
AE	0.031	1	0.031	0.22	0.6436
BC	7.225E-003	1	7.225E-003	0.052	0.8219
BD	0.22	1	0.22	1.58	0.2201
BE	0.012	1	0.012	0.087	0.7709
CD	0.51	1	0.51	3.66	0.0672
CE	6.4E-003	1	6.4E-003	0.046	0.8322
DE	2.5E −005	1	2.5E −005	1.791E-004	0.9894
A^2	0.07	1	0.07	0.5	0.4863
B^2	7.004E-004	1	7.004E-004	5.017E-003	0.9441
C^2	0.28	1	0.28	2.03	0.1666
D^2	0.12	1	0.12	0.85	0.3643
E^2	7.67	1	7.67	54.92	<0.0001
Residual	3.49	25	0.14		
Lack of fit	2.66	20	0.13	0.80	0.6747
Pure error	0.83	5	0.17		
Cur total	159.14	45			

* Degrees of freedom

to the docking output (Fig. 5). The perturbation plot illustrates the resulted binding energy as each variable moves from the chosen reference while all other factors are held constant at the middle of the design space (coded zero level) [35].

Validation of optimized docking parameters

The optimum docking condition in terms of studied effective factors was applied to further correlate the docking results with experimental biological data of an external data set (Fig. 2). For GA run, the optimum value was found to be 100 (Table 2) which is minimally required for generation of all the possible conformational clusters. It should be noted that for apo crystallographic files, higher number of GA runs could be considered. Population size would not make any significant variation in response and the default value of the software (150) which was in accordance with the optimized value could be applied. In the case of maximum number of evaluations, it was found that for ligands bearing four or less RTBs (2OHL, 2OHK, and 2OHM), 2.5×10^6 evaluations could meet the requirements for achieving optimized binding mode. Ligands containing five or more RTBs were subjected to 1.0×10^7 evaluations (2VA5, 2OHT, and 2OHU). Torsion degrees for ligand were found to be five in the case of ligands possessing four RTBs and this value

Table 4 Statistic values for the full quadratic model from ANOVA analysis in AutoDock4.2 based study of BACE-1 inhibitors

Standard deviation	Mean	Coefficient of variation (%)	PRESS	R-squared	R-adjusted	Predicted R-squared	Adeq precision*
0.37	−8.36	4.47	11.84	0.9781	0.9605	0.9256	27.099

* Adeq precision: signal to noise ratio

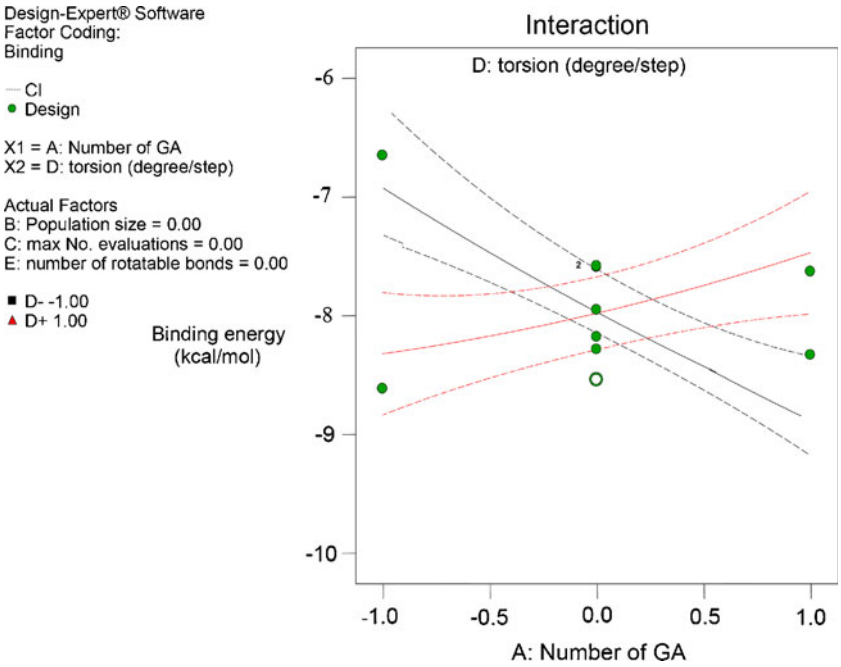

Fig. 3 Interaction plot for GA runs × torsion degrees/step term in RSM model for AutoDock4.2 based study of BACE-1 inhibitors, assigned coded factor levels for population size (0), maximum number of evaluations (0) and number of ligand rotatable bonds (0) are 150, 2500000 and 6, respectively

was also used for our regression study. As was mentioned above, ligands under study are representatives of cognate ones; hence the effect of this factor is of less importance compared to non-cognate ligands.

Selection criteria for BACE-1 inhibiting external data set were based on the following items:

- Availability of a holo crystallographic PDB structure with reliable resolution.
- Availability of a relevant experimental biological data.
- RTBs in a range of studied levels.

Top ranked docking binding energies in each trial were chosen for further correlation with in-vitro FRET-based experimental biological data (Fig. 6) [36, 37]. All the docking results produced RMSD values below 2. After correlation, Eq. 1 was obtained for BACE-1 inhibiting activities represented as pIC_{50} values. Our results revealed a high correlation between docking outputs and in vitro anti-Alzheimer activities. Binding energy (ΔG_b) was the sole descriptor variable.

$$pIC50 = -0.7476\Delta G_b - 1.1807 \quad \left(R^2 = 0.9131\right) \qquad (1)$$

The predicted binding energies of these inhibitors into the active site of receptor are listed in Table 5.

Negative control

We performed a negative control to further validate the suggested optimization model. For this purpose, two

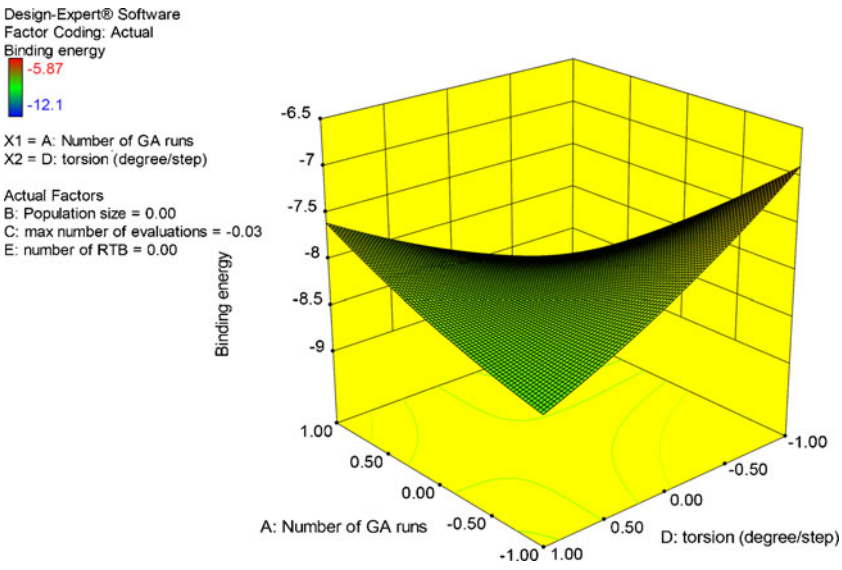

Fig. 4 3D surface plot for genetic algorithm runs × ligand torsion degrees/step RSM model term in AutoDock4.2 based study of BACE-1 inhibitors, assigned coded factor levels for population size (0), maximum number of evaluations (−0.03) and number of ligand rotatable bonds (0) are 150, 2432500 and 6, respectively

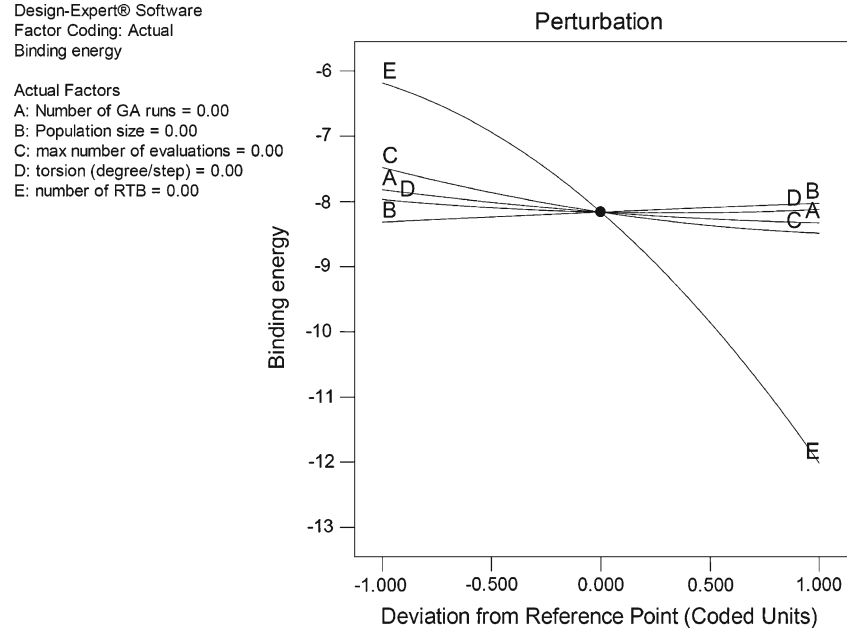

Fig. 5 Perturbation plot of docking binding energy versus number of genetic algorithm runs (**a**), population size (**b**), maximum number of evaluations (**c**), torsion degrees for ligand (**d**) and number of rotatable bonds in ligand (**e**)

platforms were utilized. First, self-docking of ligands bearing more than four rotatable bonds (2VA6, 2OHU, and 2OHT; Fig. 6) were carried out in a defined condition for ligands possessing four or fewer RTBs (Table 6). RMSD was selected as response for this assay due to its robustness as an appropriate criterion in self-docking studies.

RMSD values obtained for various self-docking procedures (Table 6) indicate that the optimized condition described for docking of a typical ligand with definite number of active torsions (RTBs≤4) may not support desirable results (accurately oriented binding poses) for more flexible BACE-1 inhibiting scaffolds (RTBs>4). However this negative control demonstrates the lack of any extraneous significant confounding effects on response.

One expectable point is that the amount of departure from reference structure (RMSD) is more pronounced for ligands designated by PDB codes 2VA6 (RTB: 7), 2OHU (RTB: 8), and 2OHT (RTB: 5) in the external data set. These results confirmed that flexibility of docked ligands is a major limitation in a simulation process. Higher number of energy evaluations along with higher number of genetic algorithm runs (significant model terms in AutoDock) would be required to support more accurate docking results in more flexible scaffolds. Our negative control analysis showed that self-docking of BACE-1 inhibitors bearing more than four RTBs could not be properly validated using 2.5×10^6 energy evaluations in Autodock4 program and to get validated results, energy evaluations need to be increased.

In the second strategy for negative control, we performed a self-docking simulation of selected ligands (Fig. 6) on the grid box including the whole BACE-1 structure (blind docking) ignoring the position of active site. The results are shown in Table 7. In a typical blind docking procedure, accurate re-orientation of a cognate ligand in the BACE-1 active site could be affected to a significant amount using pre-optimized levels of docking factors. An expectable feature of this study is the less populated top conformation clusters in the case of blind docking. Furthermore; RMSD values were also affected. Our results showed that accurate prediction of the active site in blind docking is a major bottleneck which may be unraveled at least to some extent by increasing the number of GA runs and number of energy evaluations in a self-docking process.

Table 5 Docking results for studied β-secretase inhibitors with BACE-1 target

PDB code	No. rotatable bonds in ligand	Estimated binding Energy (kcal.mol^{-1})	Experimental IC_{50}* (μM)	Experimental pIC_{50}
2VA5	5	−7.52	130	3.89
2VA6	4	−10.54	0.38	6.42
2OHL	1	−5.35	2000	2.70
2OHK	1	−5.47	2000	2.70
2OHM	4	−6.17	310	3.51
2OHT	5	−7.47	9.1	5.04
2OHU	8	−8.18	4.2	5.38

* All the reported IC_{50}s are on the basis of FRET assay protocol

Conclusions

When multiple variables are involved in a typical docking study, optimization of the system using a conventional

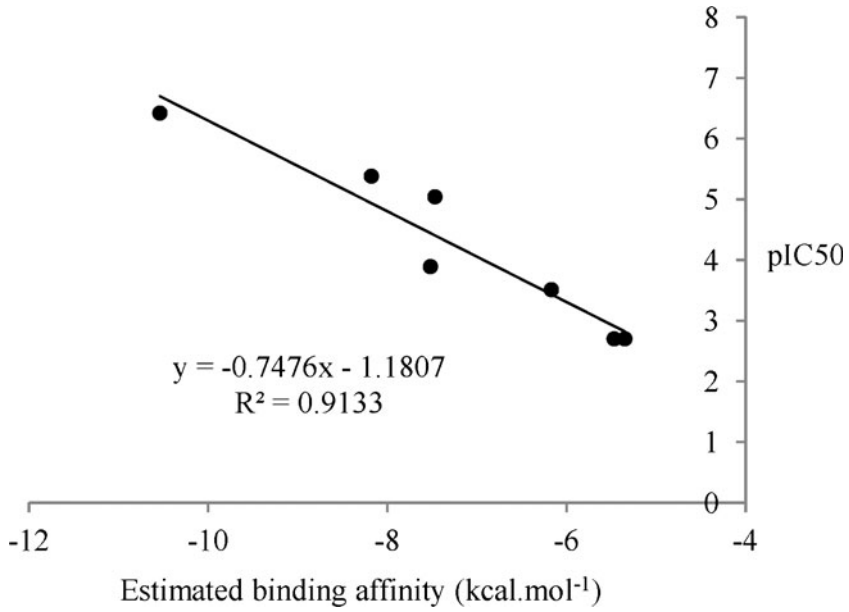

Fig. 6 Correlation plot of AutoDock4.2 binding energies versus experimental pIC_{50}s (FRET assay) for studied BACE-1 inhibitors

approach of maintaining other variables involved at constant level would not determine the order of factor effects on response and combined interaction effects on docking process may not be identified. Also the method would be time consuming and probably requires a costly route. Therefore, developing mathematical models describing the relationship between the response and independent variables, in which the significance of individual factors and multifactor interactions can be determined would be desirable. Box-Behnken method could be applied to optimize the docking simulation of BACE-1 inhibitors into receptor active site while at the same time taking all of the evaluated variables and their interactions into consideration simultaneously. The optimized docking condition showed that the number of ligand active torsions is a critical factor in final docking results while maximum number of evaluations may be the most determinant AutoDock parameter in docking free energy of binding and RMSD from reference structure. Obtained optimization levels were validated via regression of docking outputs with an external data set possessing

Table 6 Negative control and positive control in validation of optimized docking model for BACE-1 inhibitors

Docked PDB code	No. RTBs	RMSD from reference structure (Å) (positive control)[a]	RMSD from reference structure (Å) (negative control)[b]
2VA5	4	1.29	1.29
2VA6	7	1.45	3.00
2OHL	1	0.71	0.71
2OHK	1	0.65	0.65
2OHM	4	1.12	1.12
2OHT	5	1.87	2.16
2OHU	8	2.00	2.87

[a] For a positive control, suggested optimized levels of significant factors were used (1.0×10^7 energy evaluations and 100 GA runs for ligands with RTBs>4 and 2.5×10^6 energy evaluations and 100 GA runs for ligands with RTBs≤4)

[b] For a negative control, suggested optimized levels of significant factors for ligands with RTBs≤4 were used

Table 7 Negative control (blind docking) and positive control (active site-oriented docking) in validation of optimized docking model for BACE-1 inhibitors

Docked PDB code	No. RTBs	No. of conformation clusters with RMSD<2 Å out of 100 (positive control)	No. of conformation clusters with RMSD<2 Å out of 100 (negative control)	RMSD (Å) (positive control)	RMSD (Å) (negative control)
2VA5	4	95	50	1.29	1.48
2VA6	7	88	27	1.45	3.42
2OHL	1	100	55	0.71	1.61
2OHK	1	100	60	0.65	0.99
2OHM	4	91	55	1.12	1.73
2OHT	5	85	46	1.87	2.54
2OHU	8	63	29	2.00	3.70

[a] For a positive control, suggested optimized levels of significant factors were used (1.0×10^7 energy evaluations and 100 GA runs for ligands with RTBs>4 and 2.5×10^6 energy evaluations and 100 GA runs for ligands with RTBs≤4)

[b] For a negative control, suggested optimized levels of significant factors for active site-oriented docking were used

known biological data. Docking outputs exhibited high correlation coefficient with in vitro anti-Alzheimer activities (R^2=0.9133). Future perspectives may be directed toward other enzymes or considering more diverse receptor-based, docking-based or ligand-based factors in optimization procedures.

Acknowledgments Financial supports of this project by Vice-chancellor Research of Shiraz University of Medical Sciences are acknowledged.

References

1. Querfurth HWLF (2010) Alzheimer's disease. N Engl J Med 362 (4):329–344
2. Hardy J (2009) The amyloid hypothesis for Alzheimer's disease: a critical reappraisal. J Neurochem 110(4):1129–1134
3. Pietrzik C, Behl C (2005) Concepts for the treatment of Alzheimer's disease: molecular mechanisms and clinical application. Int J Exp Pathol 86(3):173–185
4. Selkoe DJ (2008) Soluble oligomers of the amyloid [beta]-protein impair synaptic plasticity and behavior. Behav Brain Res 192(1):106–113
5. Albert JS (2009) Progress in the development of β-secretase inhibitors for Alzheimer's disease. Prog Med Chem 48:133–161
6. Ziora Z, Kimura T, Kiso Y (2006) Small-sized BACE1 inhibitors. Drugs Future 31(1):53–63
7. Mandal S, Moudgil M, Mandal SK (2009) Rational drug design. Eur J Pharmacol 625(1–3):90–100
8. Krovat EM, Steindl T, Langer T (2005) Recent advances in docking and scoring. Curr Comput Aided Drug 1(1):93–102
9. Putta S, Beroza P (2007) Shapes of things: computer modeling of molecular shape in drug discovery. Curr Top Med Chem 7 (15):1514–1524
10. Laurie R, Alasdair T, Jackson RM (2006) Methods for the prediction of protein-ligand binding sites for structure-based drug design and virtual ligand screening. Curr Protein Pept Sci 7(5):395–406
11. Sousa SF, Fernandes PA, Ramos MJ (2006) Protein–ligand docking: current status and future challenges. Proteins Struct Funct Bioinf 65(1):15–26
12. Sellers RP, Alexander LD, Johnson VA, Lin CC, Savage J, Corral R, Moss J, Slugocki TS, Singh EK, Davis MR (2010) Design and synthesis of Hsp90 inhibitors: exploring the SAR of sansalvamide a derivatives. Bioorg Med Chem 18(18):6822–6856
13. Sergio Filipe Sousa PAF, Ramo MJ (2006) Protein–Ligand Docking: Current Status and Future Challenge. Proteins Struct Funct Bioinf 65:15–26
14. Khuri AI, Mukhopadhyay S (2010) Response surface methodology. WIREs Comp Stat 2(2):128–149
15. Hanrahan G, Lu K (2006) Application of factorial and response surface methodology in modern experimental design and optimization. Crit Rev Anal Chem 363(4):141–151
16. Minto CF, Schnider TW, Short TG, Gregg KM, Gentilini A, Shafer SL (2000) Response surface model for anesthetic drug interactions. Anesthesiology 92(6):1603–1616
17. Nazzal S, Khan MA (2002) Response surface methodology for the optimization of ubiquinone self-nanoemulsified drug delivery system. AAPS Pharm Sci Tech 3(1):23–31
18. Palamakula A, Nutan MTH, Khan MA (2004) Response surface methodology for optimization and characterization of limonene-based coenzyme Q10 self-nanoemulsified capsule dosage form. AAPS Pharm Sci Tech 5(4):114–121
19. Jessy S, Shital L (2008) Response surface methodology for the optimization of celecoxib self-microemulsifying drug delivery system. Indian J Pharm Sci 70(5):585–590
20. Khattab IS (2008) Studies on development of insoluble drugs as pharmaceutical suspensions by response surface methodology. Pharmazie 63(10):726–730
21. Cho BR, Shin S, Choi Y, Kovach J (2009) Development of a multidisciplinary optimization process for designing optimal pharmaceutical formulations with constrained experimental regions. Int J Adv Manuf Technol 44(9):841–853
22. Massart D, Vandeginste B, Buydens L, De Jong S, Lewi P, Smeyers-Verbeke J (1997) Handbook of chemometrics and qualimetrics: Part A. Elsevier Science Pub Co.
23. Elibol M (2002) Response surface methodological approach for inclusion of perfluorocarbon in actinorhodin fermentation medium. Process Biochem 38(5):667–673
24. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30(16):2785–2791
25. Vriend GHR, van Aalten T (1997) WHAT IF manual. European Molecular Biology Laboratory, Heidelberg
26. Edrissi M, Razzaghi-Asl N (2007) Complexation of iron with piroxicam: evaluation via response surface methodology. Acta Chim Slov 54(4):825–833
27. Hevener KE, Zhao W, Ball DM, Babaoglu K, Qi J, White SW, Lee RE (2009) Validation of molecular docking programs for virtual screening against Dihydropteroate synthase. J Chem Inf Model 49 (2):444–460
28. McGovern SL, Shoichet BK (2003) Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. J Med Chem 46(14):2895–2907
29. Cosconati S, Forli S, Perryman AL, Harris R, Goodsell DS, Olson AJ (2010) Virtual screening with AutoDock: theory and practice. Expert Opin Drug Discov 5(6):597–607
30. Forli S, Botta M (2007) Lennard-Jones potential and dummy atom settings to overcome the AutoDock limitation in treating flexible ring systems. J Chem Inf Model 47(4):1481–1492
31. Box GEP, Draper NR (1987) Empirical model-building and response surfaces. Wiley, New York
32. Ghorban Dadrass L, Madadkar SA, Shafiei A, Mahmoudian M (2004) Flexible ligand docking studies of matrix metalloproteinase inhibitors using lamarkian genetic algorithm. DARU 12(1):1–10
33. Ferreira SLC, Bruns RE, Ferreira HS, Matos GD, David JM, Brand˜ao GC, da Silva EGP, Portugal LA, dos PS Reis, Souza AS, dos Santos WNL (2007) Box-Behnken design: An alternative for the optimization of analytical methods. Anal Chim Acta 597:179–186
34. Box GEP, Hunter JS, Hunter WG (2005) Statistics for experimenters: design, innovation, and discovery, vol 13. Wiley-Interscience, New York
35. Myers RH, Montgomery DC, Anderson-Cook CM (2009) Response surface methodology: process and product optimization using designed experiments, vol 705. Wiley, New York
36. Edwards PD, Albert JS, Sylvester M, Aharony D, Andisik D, Callaghan O, Campbell JB, Carr RA, Chessari G, Congreve M (2007) Application of fragment-based lead generation to the discovery of novel, cyclic amidine -secretase inhibitors with nanomolar potency, cellular activity, and high ligand efficiency. J Med Chem 50(24):5912–5925
37. Congreve M, Aharony D, Albert J, Callaghan O, Campbell J, Carr RAE, Chessari G, Cowan S, Edwards PD, Frederickson M (2007) Application of fragment screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of β-secretase. J Med Chem 50(6):1124–1132