

SCIENTIFIC REPORTS



OPEN

Hybrid approach to sieve out natural compounds against dual targets in Alzheimer's Disease

Sucharita Das¹, Sandipan Chakraborty² & Soumalee Basu¹

Excess A β production by the key protease BACE1, results in A β aggregation, forming amyloid plaques, all of which contribute to the pathogenesis of Alzheimer's disease. Besides the multi-factorial nature of the disease, the diversity in the size and shape of known ligands that bind to the active site of BACE1, that is the flexibility of the enzyme, pose a serious challenge for the identification of drug candidates. To address the issue of receptor flexibility we have carried out ensemble docking with multiple receptor conformations. Therein, two representative structures each from closed and semi-open BACE1 conformations were selected for virtual screening to identify compounds that bind to the active site of both the conformations. These outperformed compounds were ranked using pharmacophore models generated by a ligand-based approach, for the identification of BACE1 inhibitors. The inhibitors were further predicted for anti-amyloidogenic activity using a QSAR model already established by our group thus enlisting compounds with dual potency. BACE1 inhibitory and anti-amyloidogenic activity for the commercially available compounds were validated using *in vitro* studies. Thus, incorporation of receptor flexibility in BACE1 through ensemble docking in conjunction with structure and ligand-based approach for screening might act as an effective protocol for obtaining promising scaffolds against AD.

Alzheimer's disease (AD) is a debilitating disorder that has become one of the most common forms of dementia. It is characterised by loss of functional neurons and synapses leading to disrupted communication amongst nerve cells effecting in irreversible decline in intellectual abilities. Neuropathological symptoms of the disease include the presence of extracellular amyloid plaques and intracellular neurofibrillary tangles. Other pathophysiological aspects of AD include oxidative damage, mitochondrial dysfunction, failure of molecular transport mechanisms, inflammation and cell-cycle dysregulation. However, out of these, the amyloid hypothesis, according to which accumulation of A β in the brain drives AD pathogenesis, wins the spotlight¹. The rest of the disease process, including formation of neurofibrillary tangles, is believed to be a result of imbalance between A β production and A β clearance. A β is known to be generated upon sequential cleavage of the amyloid precursor protein (APP) by β -secretase (BACE1) followed by γ -secretase. Thus inhibition of BACE1 has emerged as a prime treatment strategy for the disease not only because cleavage by BACE1 is the rate limiting step of A β production more because BACE1 knock-out mice have been found to be healthy² unlike γ -secretase knock-out mice.

Several innovative methods to develop inhibitors against BACE1 have been adopted accordingly. The first generation inhibitors consisted of peptide based transition state analogs³⁻⁵. They were highly potent since they could occupy several sub-pockets of BACE1. But their undesirable pharmacological properties made them unfit as drugs. Thereafter the second generation small molecule inhibitors covering varied classes of molecules like imidazolidinone analogs, hydroxymethylcarbonyl isosteres, pyridinium-based derivatives, flavonoids and acylguanidines⁶⁻¹⁰ were developed which were more drug-like. These compounds emerged from advanced techniques like iterative X-ray crystallography, NMR, Surface Plasmon Resonance (SPR) to name a few. Several of them displayed improved CNS (central nervous system) penetration^{11,12}. Despite this, many of these compounds could not make it to the Phase 1 clinical trials due to severe side effects^{13,14}. It may be iterated that parallel to these efforts, novel methods of computer-based drug designing was customarily pursued¹⁵⁻¹⁸.

Ongoing treatment for the disease improves or stabilizes the symptoms of dementia by increasing the communication between the nerve cells. Currently approved medications for the disease include the cholinesterase inhibitors and NMDA receptor antagonists like Donepezil, Rivastigmine, Galantamine and Memantine that provide

¹Department of Microbiology, University of Calcutta, 35 Ballygunge Circular Road, Kolkata, 700 019, India.

²School of Chemical Sciences, Indian Association for the Cultivation of Science, Jadavpur, Kolkata, 700032, India. Correspondence and requests for materials should be addressed to S.B. (email: soumalee@gmail.com)

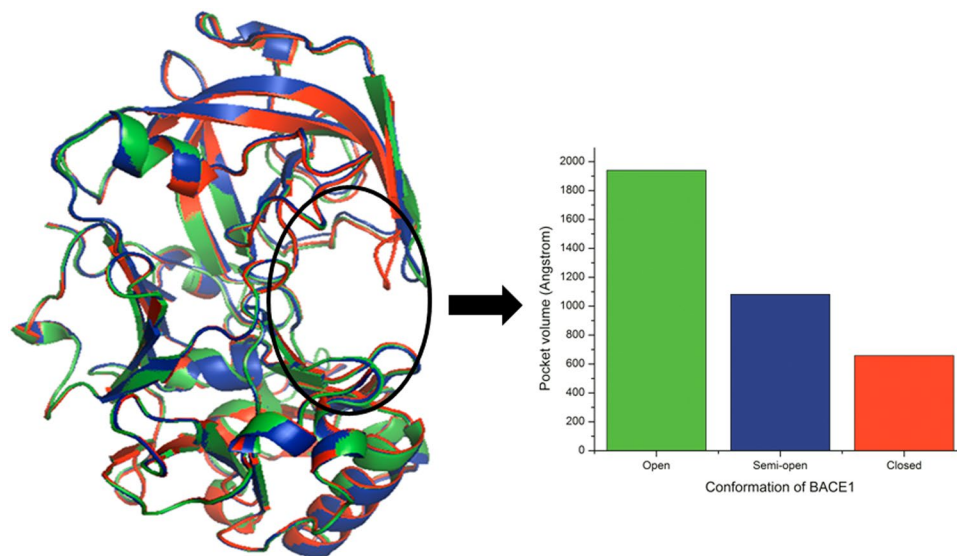


Figure 1. Three superimposed crystal structures of BACE1 representing the three different conformations: Flexible loops of the active site of open (green: 1SGZ), semi-open (blue: 4DJX) and closed (red: 3CIC) BACE1 structures show decreasing pocket volume (inset bar graph).

symptomatic relief without addressing the neuropathology^{19–21}. Hence the search for BACE1 inhibitors as drugs against AD still continues to be a hot pursuit.

As a consequence of this ongoing hunt, we find hundreds of BACE1 crystal structures in complex with inhibitors of different size and shape in the public structural repository, PDB. Comparison of the different crystal structures of BACE1-inhibitor complexes and their molecular dynamics studies show BACE1 to be quite flexible. The conformational flexibility of BACE1 enables it to switch from open to closed form upon ligand binding^{22,23} (Fig. 1). It has been observed that in the presence of peptide-like bound inhibitors, the flap moves closer to the catalytic aspartate dyad thus representing a closed conformation. Recent studies have revealed that upon binding with some special classes of inhibitors, the flap, due to steric clashes, adopts a semi-open form moving away from the dyads than in the closed form²⁴. Also an extensive study by Xu *et al.*²⁵ proved the existence of semi-open conformation of BACE1 exhibiting space group specificity with P6122. The space group specificity was however ruled out by a series of fragment screening approaches that have identified complexes of other space groups with a semi-open form too^{26–30}. Thus, to summarize, different classes of inhibitors may interact with the flap residues to a greater or lesser extent generating varied flap conformations of the protein. It is worth mentioning that lack of protein flexibility in structure-based drug designing severely limits the identification of true ligands as flexibility leads to structural rearrangement of the binding site and the adjoining structures, thus influencing ligand binding. Previous efforts of inhibitor identification have used a single BACE1 structure for virtual screening without consideration of the flap flexibility^{31,32}. So, BACE1 structure in which ligands of different inhibitor classes can dock needs to be determined³³.

One of the major problems related to design of inhibitors of enzymes is the consideration of the intrinsic flexibility of the binding site of the receptor, during docking. Scientists have been trying endlessly to incorporate flexibility using either combination of crystal and NMR structures or using computational methods³⁴ by combination of rotamers for receptor side chain conformations or sampling side chain torsion angle space³⁵ or even considering backbone flexibility through ensemble docking^{36–38}. Recently, Hou *et al.* have shown the success of appropriately using multiple structures for the open and closed conformation of the WPD loop of the lymphoid-specific tyrosine phosphatase (LYP) to virtually screen LYP inhibitors³⁹. Here, we have considered multiple crystal structures of closed and semi-open conformation of BACE1 taken from complexes with ligands of wide range of size and shape, for the fortification of receptor flexibility. It needs special mention at this point that out of 334 BACE1-inhibitor crystal structures in the PDB, none has a ligand attached to the open conformation of BACE1. Therefore, in this study, open conformation of BACE1 was excluded as an option while considering multiple crystal structures. Using multiple representative conformations of the enzyme, a database of natural compounds was screened virtually as inhibitors. These screened compounds that were capable of binding to both semi-open and closed conformation of BACE1 were further assessed for their anti-amyloidogenic activities using an already established QSAR model of our group. The flow of work is schematically depicted in Fig. 2 and discussed in details in the Methods section.

Methods for Computational Studies

Grouping of the structures. Structures of BACE1-inhibitor complexes having resolution <2.5 Å, no missing residues, and known K_i were downloaded from the PDB database. The maximum number of semi-open conformations satisfying the above criteria was found to be five. We therefore formed the dataset with altogether ten members including five more receptor structures in the closed conformation. The receptor structures were

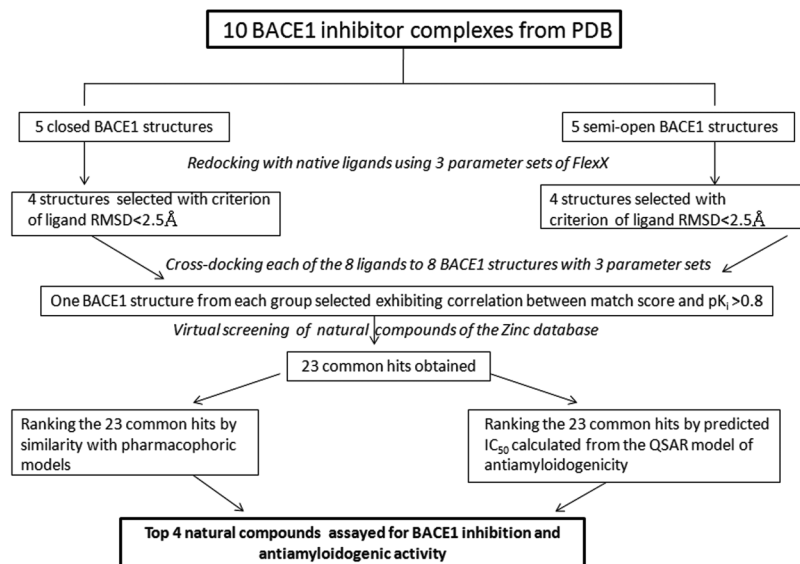


Figure 2. Flowchart showing the step by step progression of the work.

divided into two groups based on three distances - the depth of the active site cleft given by the distances between the flap tip and the two catalytic dyads individually (d_1 , d_2) and width of the mouth of the active site cleft (d_3). Here d_1 is the $C\beta$ D32- $C\alpha$ T72 distance, d_2 is the $C\alpha$ D228- $C\alpha$ T72 and d_3 is the $C\alpha$ S325- $C\alpha$ T72 distance. While one group consisted of semi-open conformations (PDB IDs: 4DJX, 4FS4, 4HA5, 4H3G, 4H3F) with d_1 , d_2 , d_3 ranging from 14.3 to 15 Å and the other group consisted of flap closed conformation (PDB IDs: 2P4J, 2G94, 2QMG, 3LPK and 3CIC) with d_1 , d_2 , d_3 ranging from 11.6 to 13.9 Å. Active site RMSD was calculated in Pymol⁴⁰ (version downloaded in 2014) considering the flap region (67–77), 10 s loop (9–14), A-loop (residues 158–167), F-loop (residues 311–318) and D-loop (residues 270–273)⁴¹.

Protein preparation. Water molecules from the crystal structures were removed. One of the aspartates (D32) of the aspartate dyad of BACE1 was protonated²³ for all the docking studies. Binding site was determined individually for each of the ten structures choosing the corresponding ligand as a reference ligand and specifying a radius of 13 Å.

Docking. All docking studies were carried out using the FlexX module⁴². The hybrid approach in which the ligand binding is driven by enthalpy and entropy was used as the docking strategy. During each run, default values for access scaling and all the stereo modes (E/Z, R/S, pseudo R/S) were considered. Most of the docking studies carried out previously using FlexX, have used the default values of maximum number of solutions per iterations and maximum number of solutions per fragmentations. However, we have varied the default values of these parameters to generate two more sets of parameters. This was done to check if changing the default parameters led to better docking results. So, three different parameter sets (200 and 200; 125 and 300; 400 and 400) with maximum number of solutions per iterations and maximum number of solutions per fragmentations were considered during re-docking and cross docking.

Re-docking. Each inhibitor was re-docked to its native BACE1 crystal structure using the three parameter sets mentioned above. RMSD between the crystal conformation of the ligand and the docked pose was calculated and a cut off of 2.5 Å was selected as a criterion for the subsequent cross-docking studies. During re-docking of inhibitors the common interactions were analysed and the presence of any two of the favourable interactions were used as a criterion to identify hits in virtual screening. The criterion was first determined manually and reaffirmed by the FlexX-Pharm docking.

Cross-docking. Cross-docking studies were carried out between inhibitors of varied sizes and receptors with varied volume of the active sites using the FlexX software. In FlexX, “Match score” is calculated from the contribution of the matched interacting groups. Therefore, a correlation between the Match score and pK_i of the BACE1-inhibitor complexes selected from re-docking was obtained to determine which receptor could dock to inhibitors of varied sizes with favourable interactions. $R^2 > 0.7$ was considered to be a criterion for good correlation. The p-value and F-value for the obtained correlation between Match scores and pK_i of the BACE1-inhibitor complexes were calculated for top ranked poses using the online server Wessa⁴³. P-value < 0.05 and F-Test value $\gg 1$ was considered to be significant. The BACE1 structure that resulted in highest squared correlation coefficient (R^2) was selected from each group (semi-open and closed flap conformations) for virtual screening. Using the correlation plot a hit selection criterion of Match score < -29 kJ/mol was set to screen hits from the database of compounds. Detailed analysis of the interactions was carried out using the Accelrys Discovery Studio Visualizer⁴⁴.

MM-GBSA calculations. The binding energies of the re-docked and cross-docked complexes were obtained by using the MM-GBSA option of the prime submenu in the Schrodinger software⁴⁵. It may be noted that the MM-GBSA energies for the common hits obtained from virtual screening of the compounds with 4H3G and 2P4J as receptors were also calculated using the same menu.

Determining squared correlation coefficient for compounds with similar scaffolds. We have collected the structure activity data reported for the ligands of 2P4J⁴⁶ and 4H3G⁴⁷ from Ghosh *et al.* and Mandal *et al.* The ligands were minimized using molecular mechanics force field in Hyperchem⁴⁸ to a convergence of 0.003 kcal Å⁻¹ mol⁻¹ and docked to their respective BACE1 structures. A correlation was drawn between pK_i and Match score.

Virtual Screening. Natural compounds in standard delay format (sdf) were downloaded from the natural product database, ZINC, and filtered by the Lipinski filter, using DruLiTo⁴⁹. Filtered compounds were screened by docking to the two selected semi-open and closed form of BACE1 structures using the FlexX docking program. Compounds showing interactions with frequently interacting BACE1 residues obtained from re-docking and exhibiting a Match score below -29 kcal/mol were considered as hits.

Generation of Pharmacophore Model and Pharmacophore Screening Using PharmaGist. A pharmacophore based ranking of the common hits from both forms of BACE1 was carried out. PharmaGist^{50,51} was used to generate candidate pharmacophore model using known BACE1 inhibitors⁵²⁻⁵⁹ that were obtained from BindingDB⁶⁰. Out of the input ligands one of the ligand is randomly selected as pivot by the algorithm. The method then identifies rigid groups in the remaining ligands and generates a set of transformations for superimposing the target rigid group onto the pivot. The superimpositions are reassembled into new conformations of the target ligand, aligned on the pivot and the alignment is scored by the weighted sum of the matched pivot features. A feature from pivot can be matched to a conformation of the target ligand, if they are of the same type and the distance between them is below a predefined threshold. The distance threshold for centres of aromatic rings is 1.8 Å and for the centres of the other feature types is 1.4 Å. The default weight values are 1 for each feature type except for hydrophobicity which is 0.3 and aromatic rings for which it is 3. The score of a pairwise alignment with a subset {*f_i*} of matched pivot features is $\sum_i s(f_i)$, where *s*(*f_i*) is the score for matching the pivot feature *f_i*. The high-scoring pairwise alignments are combined into multiple alignments in order to find subsets of pivot features that are matched by several pairwise alignments with different target ligands. The score of a multiple alignment is given by $(m + 1)^{1/2} \cdot \sum_i s(f_i)$, where *m* is the number of participating target ligands. Candidate pharmacophores are reported on the basis of this score. In our study, the ranking of the common hits was carried out based on the similarity of these molecules to the pharmacophore thus obtained. It may be mentioned that this similarity was measured using the same method of pairwise alignment used for pharmacophore generation. The default score values for matching the pivot features were used.

We used two different models to rank the common hits. Model 1 has been built using 12 known BACE1 inhibitors (IC₅₀ varying from 0.7–49 nM) and Model 2 was generated from 9 known BACE1 inhibitors (K_i varying from 0.3–270 nM).

Molecular overlay. Both the models were also validated by aligning to a reference pharmacophore corresponding to each set of input molecules. The reference pharmacophore was constructed by superposition of the BACE1-inhibitor complexes of each set of molecules, using Discovery Studio visualizer. The reference pharmacophore was then aligned to the PharmaGist generated pharmacophore using the molecular overlay tool of Discovery Studio visualizer.

Enrichment factor calculation. We investigated the effectiveness of the pharmacophore based ranking methodology using enrichment studies. Enrichment factor of the models was computed on a DUD (Directory of useful decoys) data set of BACE1, consisting of 10 actives and 250 decoys, using the following equation:

$$EF_{\%} = \frac{\frac{N_{active(\%)}}{N_{(\%)}}}{\frac{N_{active}}{N_{all}}}$$

where *N*_{active(%)} is the number of active compounds in a subset of the library, *N*_(%) is the number of compounds in the subset of the library, *N*_{active} and *N*_{all} are the total number of active compounds and the total number of compounds in the library.

Assessing the applicability domain of the in-house build QSAR models. In a previous study, we have established a QSAR model for anti-amyloidogenicity⁶¹. We checked the applicability of the model on the 23 common hits. The applicability of the QSAR was checked using the PCA Bounding Box method. This method defines the applicability domain by the projection of molecules in the principal component space considering the maximum and minimum values of the principal components⁶². The molecules that were found to lie within the applicability domain of the QSAR were considered for prediction of their activities by the QSAR.

Prediction of anti-amyloidogenic activity. The common hits thus retrieved were subjected to further screening and ranking by an in-house built QSAR model of anti-amyloidogenic activity. The training set comprised of 18 polyphenols 5 of which are known to inhibit BACE1⁶¹. All the molecules were minimized using HyperChem with AM1 semi-empirical method to a gradient of 0.003 kcal/Å mol. Descriptor calculations were then carried out using the Codessa software⁶³. Prior to the calculation of quantum-chemical and thermodynamic

properties, the optimized structures were subjected to wave-function optimization using AM1 semi-empirical method of the AMPAC software⁶⁴. The model with the highest predicted R^2 , generated using heuristic approach was considered. The high predictive potential of the model was ascertained using the statistical parameters such as R^2 and F-Test values of 0.9135 and 34.31. Principal component analysis was used to characterize the outliers. Additionally cross validation technique was used to validate the QSAR model with R^2_{CV} of 0.844. The model has been previously used for screening natural compounds such as flavonoids, flavonoid derivatives, alkaloids, stilbenes and terpenoids. These screened compounds have exhibited anti-amyloidogenicity and BACE1 inhibitory activities^{65,66}.

The QSAR equation used for generating a list of top 23 compounds is as follows:

$$\begin{aligned} \text{LogIC}_{50} = & -11.244 - 0.012 * (\text{HBSA}) + 5.762 * (\text{AICO}) \\ & + 0.893 * (\text{ESP} - \text{RNCS}) + 1.767 * (\text{VE}) \end{aligned} \quad (1)$$

where HBSA is hydrogen bonding surface area, AICO is the average information content that reflects the shape and symmetry, ESP-RNCS is the relative negative charged surface area and VE is the vibrational entropy (300 K)/no of atoms.

Materials and methods for *in vitro* studies. The commercially available screened compounds of ZINC database were procured from Interbioscreen. BACE1 assay kit and HFIP were procured from Sigma Aldrich while $A\beta_{25-35}$ was bought from Anaspec in powdered form.

BACE1 inhibition assay. BACE1 activity assay kit was used to determine the BACE1 inhibition ability of the screened compounds. The substrate is fluorescently labelled with donor on one side and quencher group on the other side. In the absence of BACE1, due to FRET (Fluorescence resonance energy transfer) the donor fluorescence of the substrate is quenched. Upon cleavage of the labelled substrate by BACE1 the donor and quenching groups are separated, hence fluorescence increases. Therefore inhibition of BACE1 by an inhibitor would result in lesser substrate cleavage and hence lead to a decrease in fluorescence. In our study, concentrated stock solutions of the screened compounds were prepared by dissolving the compounds in DMSO. Final concentration of the screened compounds was 1 μM . A known inhibitor (myricetin) and a weak inhibitor (gossypetin) were also included in the assay. The assay was carried out according to the manufacturer's instructions. The reaction mixture was incubated at 37 °C for 2 hours. Fluorescence was then measured by excitation at 320 nm and emission at 405 nm. The inhibition percentage was obtained by the following equation:

$$\text{Single point inhibition \% (SPI)}_{1\mu\text{M}} = \left(1 - \frac{S - S_0}{C - C_0} \right) \times 100$$

where C is the fluorescence of the control (containing enzyme, assay buffer and substrate) after two hours of incubation, C_0 is the fluorescence of the control at the start, S is the fluorescence of the compound solution (containing enzyme, assay buffer, compound solution and substrate) after two hours of incubation and S_0 is the fluorescence of the compound solution at start.

Thioflavin T assay. The dye Thioflavin T (ThT), is used to quantify amyloid aggregates. On binding to the beta sheets of amyloid aggregates, the rotation of the ThT rings is hampered thus preserving the excited states leading to fluorescence enhancement. The $A\beta_{25-35}$ sample was prepared by dissolving it in HFIP to remove any preformed aggregates and keeping aliquots at -80 °C. On the day of assay, aliquot of the HFIP dissolved $A\beta_{25-35}$ (final concentration 100 μM) was lyophilized. Thereafter the lyophilized sample was dissolved in phosphate buffer. A stock solution of the compounds with concentration of 1 mM was prepared in DMSO. The compounds with final concentration of 100 μM were co-incubated with $A\beta_{25-35}$ for 72 hours at 37 °C. The co-incubated samples were diluted five times. A stock solution of Thioflavin T was prepared in DMSO and final concentration of 15 μM was added. The fluorescence was measured by excitation at 440 nm and emission at 460–550 nm.

Results and Discussion

Classification of the BACE1 conformations from crystal structures of complexes. We have excluded open BACE1 conformation from our study as none from about 300 ligand bound crystal structures from PDB were in the open conformation. Ten BACE1-ligand complexes with no missing residues were chosen and grouped into semi-open and closed structures on the basis of the depth of the active site judged by the flap tip-catalytic dyad distances (d1 and d2) and width of the mouth of active site given by the distance d3 respectively. Table 1 shows the RMSD of the flexible regions near the active site (as mentioned in the Methods section) between structures within and across the groups. PDB IDs from row or column 2 to 6 represent the semi-open conformations and comprise of Group1 while those from row or column 7 to 11 represent closed conformations and comprise of Group2. Within Group 1, the RMSD varied between 0.2–0.37 Å and surprisingly a wider range of 0.28–1.48 Å was observed within Group 2. Within Group 2, the range widened as a result of the differences in the conformation of the flap residues and also other loops near the active site of these closed structures. Thus it may be noted that the active site conformation of the closed structures are more diverse than those of the semi-open structures. The RMSD across groups (bold faced values in Table 1) varied from 0.6 to 1.63 Å. Out of all the pairs the most deviated active site conformations were found between 4H3G, 4HA5, 4FS4 (Group 1) and 2P4J (Group 2).

Unlike the RMSD values which varied within the same group of BACE1 conformation, we observed that the distances d1, d2 and d3 followed almost uniform values within the same group. Hence these distances would

PDB ID	4H3G	4DJX	4HA5	4FS4	4H3F	3CIC	3LPK	2P4J	2QMG	2G94
4H3G	0	0.24	0.34	0.36	0.33	0.65	0.61	1.62	0.69	1.1
4DJX		0	0.27	0.25	0.27	0.66	0.6	1.6	0.68	1.09
4HA5			0	0.2	0.3	0.74	0.72	1.63	0.76	1.13
4FS4				0	0.27	0.72	0.69	1.62	0.73	1.1
4H3F					0	0.75	0.72	1.61	0.75	1.12
3CIC						0	0.28	1.47	0.27	0.81
3LPK							0	1.49	0.35	0.84
2P4J								0	1.47	1.42
2QMG									0	0.75
2G94										0

Table 1. RMSD in Å between the flexible regions near the active site of the BACE1 structures.

be better to classify the open and closed conformations of BACE1 (Supplementary Fig. S1 and Supplementary Table S1).

Re-docking. Re-docking each ligand (Supplementary Table S2) to their respective BACE1 crystal structure was carried out to test the performance of the docking algorithm on the dataset. Using varied parameter sets, root mean square deviation (RMSD) was calculated between the re-docked ligand and the crystallographic conformation of the ligand. Ligands with an RMSD < 2.5 Å on re-docking were assumed to generate a docked pose relevant to the crystal pose and hence were considered for cross-docking. However in cases of failed re-docking (36.6% of the total re-docking studies), the corresponding BACE1 structures were excluded from further cross-docking studies. Thus the ligand of each of the ten BACE1 structures was re-docked using three different parameter sets (described in the docking sub-section of methods section). Re-docking studies with the parameter sets resulted in different RMSD of the ligands tabulated in Supplementary Table S3. Exception was the ligand of 4FS4 which showed no change in RMSD across all the three parameter sets. Two ligand structures 4H3F_L and 2G94_L, upon re-docking showed RMSD > 2.5 Å hence further studies using these receptor-ligand complexes were not carried out. Consistent high RMSD in case of 2G94 for all the parameter sets might have resulted from high number of rotatable bonds in the corresponding ligand (Supplementary Table S2). In case of the complexes 4H3G and 2P4J, only parameter set 3 could generate an RMSD < 2.5 Å. Also it is noteworthy that ligand RMSD with parameter set 3 has gone above 2.5 Å for two receptor structures in comparison to four and five for the other two parameter sets thus marking parameter set 3 superior over the two.

Excluding 4H3F and 2G94, the rest of the eight complexes were analyzed for the identification of the enzyme-inhibitor interactions. For complexes in the closed BACE1 conformations, the common hydrogen bond interactions included T232, G230, Q73, T72, G34 and either of the aspartate dyads D228 and D32. In case of complexes with semi-open BACE1 conformations, van der Waals interaction was much more prominent and hydrogen bonding interaction was lesser. However in these complexes hydrogen bonds with both D228 and D32 were observed, strengthening the interaction in the absence of the proximity of the flap residues. Stacking interactions with F108 and Y71 were observed both in semi-open and closed conformations. A study by Liu *et al.*⁶⁷ corroborated these findings as they showed that van der Waals interactions with Y71 and H-bonding with T72 stabilize the BACE1-inhibitor complex. They also showed that H-bond interactions with G230 of S3 subpocket and T232 of S2 subpocket are favourable interactions that occur in several BACE1 inhibitors. In the present study, H-bond interactions with BACE1 residues, T72, G230 and T232 and π -stacking interaction with Y71 were considered to be important and therefore presence of any two of these favourable interactions were used as a criterion for the selection of hits during virtual screening. It is worth mentioning that since even a weak inhibitor forms hydrogen bond with D228 and/or D32 therefore favourable interaction with either of them was kept as a default criterion.

Cross-docking. Cross-docking studies were carried out with receptors and inhibitors taken from eight complexes selected from re-docking. For each receptor, a correlation was obtained using eight Match scores (from a self-docking and seven cross docking studies) and eight corresponding inhibition constants. The purpose of the study was to identify BACE1 structures that could show reasonable agreement between Match scores and activity data (pK_i) for varied inhibitor sizes and also to set a threshold Match score for virtual screening.

The cross-docked complexes within a group share interaction profiles similar to the cognate pair indicated by the bold-faced residues in Table 2. However it was found that the original pose could not be recovered in several cases as evident from the corresponding interaction profiles (Supplementary Table S4). It was observed that docking with closed BACE1 conformations were less correlated ($R^2 < 0.8$) to the inhibition constants except for the structure of 2P4J (Table 3). Docking of inhibitors with few rotatable bonds to the small cavity of the closed structures (Supplementary Tables S1 and S2) probably resulted in steric clashes leading to poor scores. On the other hand, large inhibitors with high number of rotatable bonds occupied several subpockets and thus gave better scores. Moreover, lesser steric clashes due to larger cavity volume could be the reason for better correlation as observed in case of semi-open conformation over the closed one (Table 3). Hence receptors 4H3G from the semi-open and 2P4J from the closed conformation have been selected that could favourably interact with BACE1 inhibitors of varied sizes. Best R^2 along with permissible F-test and p-values were considered for the selection. It may be emphasized that the pair showed considerable deviation of the flexible regions of the active site (Table 1) implying distinct representation of the two conformations of the receptor, BACE1. Incidentally, cross-docking

Ligands								
BACE1	23I	SC6	316	Z76	10Q	0KQ	13W	H24
	T72, Q73, Y198, D228, G230, T231, T232, N23, S325.	D32, G34, T72, R128, Y198.	Q73, G230, R235, S325.	D32, G34, T72, Q73, R128, Y198.	S35, T72, Q73, R128, S325.	D32, T72, R128.	D32, S36, N37, T72, Q73.	D32, T72, G230, R235, S325.
	-44	-41	-40.52	-39.6	-27.33	-23.6	-20.2	-21.07
	-90.31	-87.22	-70.89	-89.49	-49.7	-40.69	-59.03	-33.95
2QMG	T72, Q73, G230, N233, R235, K321, S325.	G34, Q73, Y198, D228, G230, T231, T232.	D32, G34, T72, Q73, D228, T231, T232.	G34, Q73, Y198, D228, G230, T231, T232.	Q73, Y198, T232.	G34, T72, D228, T232.	S36, Q73, D228.	D32, T72, Q73, T232.
	-37.61	-47.9	-44	-52.8	-23.96	-24.37	=-30	-22.8
	-65.82	-115.8	-111.4	-97.4	-59.63	-56.45	-80.95	-37.12
3CIC	G34, P70, T72, Q73, Y198, G230, T232.	D32, T72, D228, G230, N233.	G34, T72, Q73, D228, G230, T232.	G34, T72, Q73, D228, G230, T232.	P70, R128, Y198, R235, S328.	T72, D228.	T72, Q73, R128, D228.	D32, T72, Q73, D228, T232.
	-37.57	-38.25	-41.51	-41.77	-26.22	-23.81	-32.18	-22.09
	-92.1	-91.2	-99.27	-100.9	-42.48	-59.72	-73.95	-23.58
3LPK	T72, R128, R235, Q326, T329.	D32, Q73, D228, G230, T232.	G34, Q73, D228, G230, T231, T232.	G34, Q73, Y198, D228, G230, T231, T232.	D32, T72, D228, R235.	G34, T72, D228, T232.	S36, T72, Q73, D228.	D32, T72, Q73, T232.
	-33.64	-42.18	-44.35	-44.98	-26.69	-26.69	-32.71	-26.78
	-52.68	-110.1	-103.8	-118.9	-72.79	-66.55	-71.73	-46.95
4H3G	T72, Q73, Y198, R235, G230, T232, N233.	D32, D228, G230, T232, N233.	D32, D228, G230, T232, N233.	D32, G34, Y198, D228, G230, T232.	D32, W76, D228.	D32, D228.	D32, D228.	I126, Y198.
	-36.38	-34.94	-29.16	-36.3	-26.31	-26.66	-26.99	-21.27
	-60.31	-79.57	-66.56	-89.19	-84.06	-68.44	-62.24	-48.83
4DJX	P70, T72, Y198, R235, Q326, S328, T329.	D32, G34, R128, D228.	D32, G34, D228, G230, T232.	T72, Q73, N233, R235, S328, T329.	G34, Q73, D228.	D32, D228.	D32, D228, T232.	D32, D228, T232.
	-36.3	-41.9	-33.02	-36.79	-22.5	-22.73	-23.28	-20.85
	-65.28	-75.61	-105.4	-63.1	-88.46	-59.29	-72.3	-25.17
4HA5	S10, G11, T232, N233, R307, P308.	D32, G34, Y198, D228, T232.	D32, G34, S36, D228.	Y71, T72, Q73, Y198, R235, T329.	D32, Y71, D228.	D32, D228.	D32, W115, D228.	D32, D228.
	-36.84	-38.43	-34	-34.27	-27.03	-25.89	-22.65	-20.25
	-58.4	-90.8	-78.69	-64.77	-88.46	-54.4	-61.16	-20.72
4FS4	T72, Q73, R235, S328, T329.	Q73, K107, R235.	D32, D228, G230, T232.	G34, Y198, D228, G230, T232.	D32, Y198, D228.	D32, D228.	D32, D228.	D32, G34, T232.
	-40.4	-28.57	-33.42	-39.56	-25.77	-23.76	-21.14	-22
	-55.09	-58.95	-72.79	-84.79	-86.42	-60.97	-67.32	-25.95

Table 2. Comparison of the interacting residues involved in H-bonds, Match scores and MM-GBSA energies for the re-docked and cross-docked complexes (rank 1 poses for parameter set 3) used in the study. For more

details refer to the note[§]. [§]The rows represent the 8 receptors used in cross-docking and the columns indicate the 8 ligands used therein. For each receptor, the H-bonding residues, Match score and MM-GBSA energy are indicated in the first, second and third row respectively. The diagonal cells correspond to re-docked data, entries of which have been italicized. The bold-faced residues indicate interacting residues similar to the cognate pair.

PDB ID	R ² , F-test and p-value of the model obtained from Parameter1	R ² , F-test and p-value of the model obtained from Parameter2	R ² , F-test and p-value of the model obtained from Parameter3
4H3G	Not done	Not done	R ² = 0.943 F-test = 55.11 p-value = 8.89e-10
4DJX	R ² = 0.685 F-test = 17.326 p-value = 4.463e-08	R ² = 0.765762 F-test = 26.05 p-value = 9.238e-09	R ² = 0.783 F-test = 21.573 p-value = 3.275e-08
4HA5	R ² = 0.869 F-test = 26.429 p-value = 2.15e-09	Not done	R ² = 0.9 F-test = 27.846 p-value = 2.059e-09
4FS4	R ² = 0.51 F-test = 8.557 p-value = 2.067e-7	R ² = 0.69 F-test = 17.964 p-value = 7.796e-9	R ² = 0.55 F-test = 9.79 p-value = 1.625e-7
3CIC	R ² = 0.28 F-test = 0.45 p-value = 2.065e-5	R ² = 0.58 F-test = 8.449 p-value = 8.146e-8	R ² = 0.66 F-test = 10.4 p-value = 8.66e-7
3LPK	R ² = 0.503 F-test = 4.679 p-value = 1.669e-6	R ² = 0.593 F-test = 4.460 p-value = 4.873e-7	R ² = 0.619 F-test = 3.12 p-value = 1.629e-6
2P4J	Not done	Not done	R ² = 0.85 F-test = 32.08 p-value = 1.512e-8
2QMG	R ² = 0.746 F-test = 11.13 p-value = 1.06e-6	R ² = 0.64 F-test = 5.389 p-value = 1.07e-5	R ² = 0.654 F-test = 12.69 p-value = 6.52e-7

Table 3. Statistical details of the linear correlation models obtained by cross-docking studies.

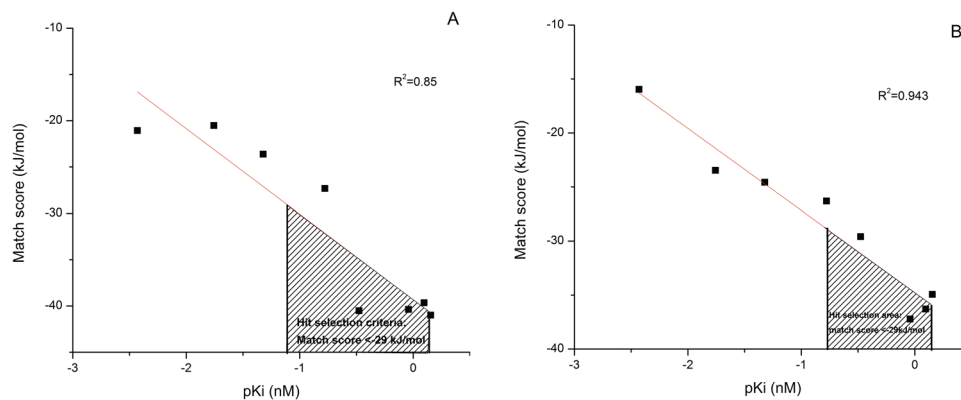


Figure 3. Linear correlation plots for 2P4J and 4H3G. pK_i versus Match Score for the receptors (A) 2P4J and (B) 4H3G. The shaded region of the graphs corresponds to the highly active inhibitors of BACE1 with a Match score < -29 kJ/mol which has been used as a selection criterion during virtual screening.

with parameter set 3 showed the highest R² value which was in corroboration with the re-docking studies where too parameter set 3 outperformed. Figure 3A,B shows the dependence of Match score, calculated from parameter set 3, on the pK_i for the receptors 4H3G and 2P4J respectively. It is clear for both the receptors that a score higher than -29 kJ/mol corresponds to high pK_i or low inhibition constant. Thus, a Match score less than -29 kJ/mol was regarded as a threshold during the screening of ligands as inhibitors.

Correlation of Match score with experimental activity. Before using the two structures (2P4J and 4H3G) for virtual screening of compounds using docking study, the correlation between Match score and binding affinity was validated. Therefore optimized structures obtained for 8 ligands of 2P4J (compound no-22, 5a, 5b, 5c, 5d, 5e, 5f and 5g as in Ghosh *et al.*⁴⁶) and 13 ligands of 4H3G (compound no-26, 27, 28, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41 as in Mandal *et al.*⁴⁷) were docked to the respective receptors and the Match score was noted down. pK_i was calculated from the K_i values reported by Ghosh *et al.* for 2P4J and Mandal *et al.* for 4H3G. The pK_i versus

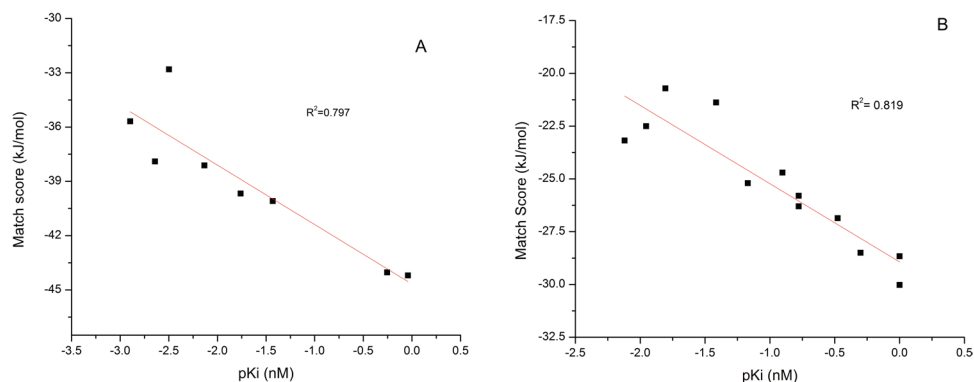


Figure 4. Plot of pK_i and Match score for known ligands of (A) 2P4J (B) 4H3G.

Match score plot for both (Fig. 4A,B) indicate good correlation. The K_i , Match score, MM-GBSA energy and receptor ligand interactions of all are tabulated in Tables S5 and S6.

Virtual screening. A library of 12000 compounds was simultaneously docked to the selected semi-open (4H3G) and closed (2P4J) conformation of BACE1. The compounds were thus declared hits if (i) they showed two out of the four interactions analyzed from re-docking studies and (ii) exhibited a Match score < -29 kJ/mol, a threshold selected from cross-docking studies. 4H3G yielded 30 compounds while 2P4J yielded 57 compounds, out of which 23 compounds were found common (Table S7).

MM-GBSA calculations of the screened hits. Binding energies were calculated for the complexes with the screened compounds (Table S7) and also the known BACE1 inhibitors (Table 2) to assess the stability of the docked complexes. The ΔG values varied from -100.92 kcal/mol to -21.93 kcal/mol for the complexes with closed BACE1 structure. In case of the complexes with semi-open BACE1 structures the ΔG values varied from -89.8 kcal/mol to -39.69 kcal/mol. Interestingly ZINC31167296 and ZINC01758814 were the top two compounds implying high affinity for both the conformations (bold-faced in Table S7). As seen in Fig. 5, ZINC01758814 forms H-bond with T72, T232 and G230 for both the forms and so does ZINC31167296 with I126, F108, D228 and R235.

Arranging the common compounds with the pharmacophore models. Two pharmacophore models (Model 1 and Model 2) were generated using known active inhibitors (Supplementary Tables S2 and S8) of BACE1 (Fig. 6). Both Model 1 and Model 2 gave the highest score of alignment with its pivot (Table 4). As evident from the figure, some of the donor and acceptor features overlapped in both the models. As a result the total spatial features of donor and acceptor for each of the two models remained same i.e. 5. Analysis of the corresponding reference pharmacophore for Model 1 and Model 2 indicate that the donors and acceptors are involved in H-bond interactions with active site residues of BACE1 such as D32, T72 Q73, G230, T232 and N233. Additionally the presence of aromatic feature involved in the stacking interactions with Y71 has been observed in both the models. The two models were validated by overlaying to their respective reference pharmacophore obtained by superposition of the X-ray crystal structures of the input molecules (Table 5). The generated models were found to be similar to their reference pharmacophore.

The 23 common compounds were arranged in descending order according to the pairwise alignment score obtained by aligning each compound with the two pharmacophore models generated. For the pairwise alignment, the default criterion of a minimum match of 3 features between the model and the common compound was set. The arrangement of the common compounds selected through virtual screening was assessed using the enrichment factor. The maximum value of enrichment factor that could be attained using the top 10 and 15 compounds were 26 and 17.34 respectively. We found that the pharmacophore models [Model 1 (using 12 inhibitors with IC_{50} values) and Model 2 (using 9 inhibitors with K_i values)] showed comparable performance (Table 6) and therefore both were used for arranging the compounds.

Activity prediction of BACE1 inhibitors as multi-target agents. AD being a multi-factorial neurodegenerative disease, a promising designing approach should be one that would target more than one molecule responsible for the pathology. It is to be noted that the disease is characterized by elevated levels of amyloid β in the brain as an underlying pathogenesis of the disease. Therefore, 23 virtually screened and arranged BACE1 inhibitors were predicted for anti-amyloidogenic activity to finally declare them as predicted BACE1 inhibitors with dual potency.

Antiamyloidogenic activity prediction of BACE1 inhibitors. All of the 23 screened compounds lay within the applicability domain of the in-house developed QSAR model⁶¹ that was used to predict the anti-amyloidogenic activity ($\log IC_{50}$). Using the predicted IC_{50} values, a list of the 23 compounds was prepared. IC_{50} of the compounds have been calculated considering the ligand-based descriptors like hydrogen bonding surface area, average information content reflecting the shape and symmetry, the relative negative charged surface

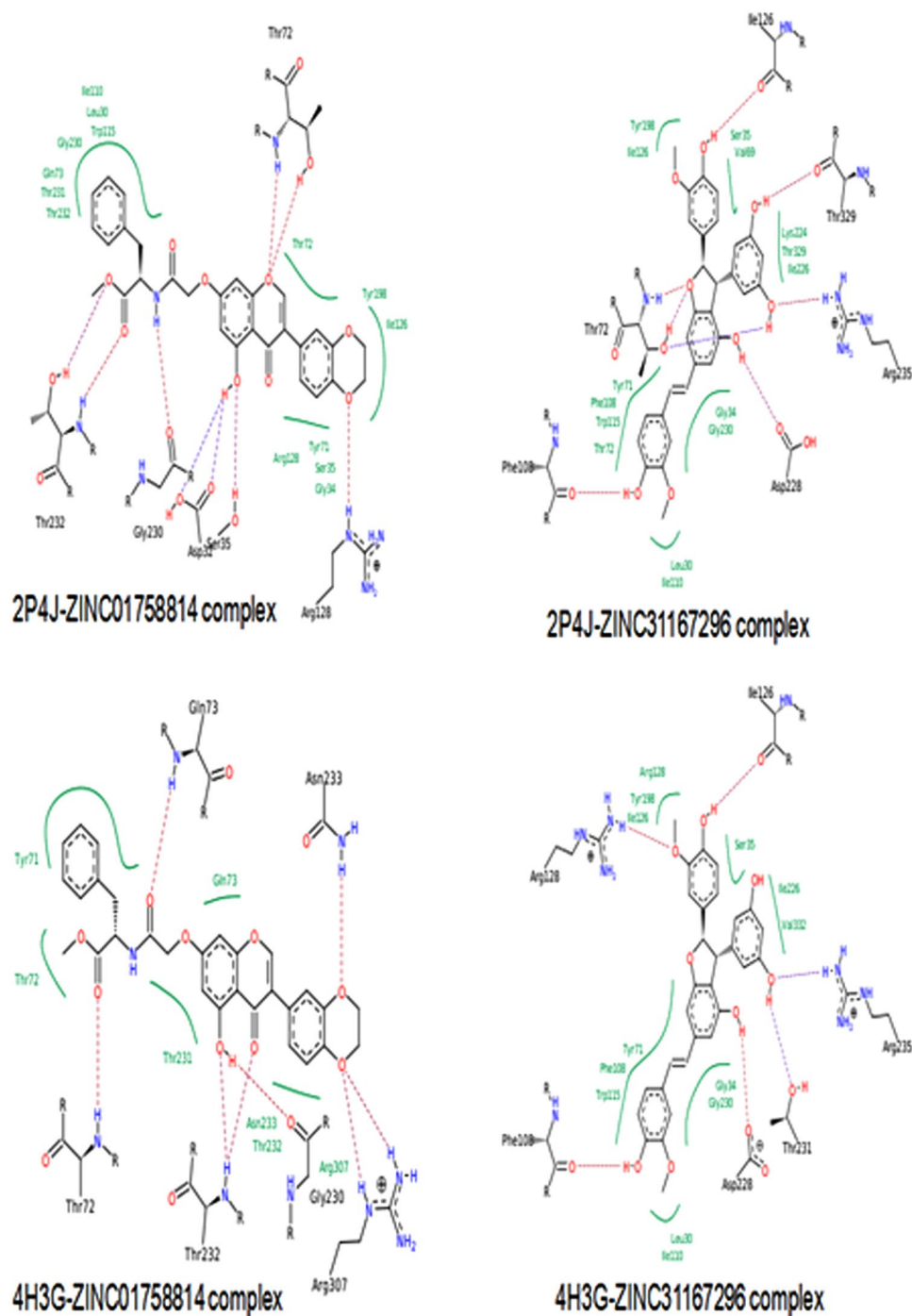


Figure 5. Interactions exhibited by the ligands (ZINC31167296 and ZINC01758814) with lowest binding energies. The interaction map has been generated by FlexX.

area and vibrational energy of the atoms as in Eqn. 1 of Methods section. An important descriptor that might have contributed significantly to its antiamyloidogenic activity is hydrogen bonding surface area (HBSA) which was high for ZINC72319974, ZINC03175470, ZINC31167296, ZINC16027834 and ZINC01758814 signifying less ability to form hydrogen bonds with bulk solvent which in turn would increase the anti-amyloidogenic activity. The molecules thus bind to the hydrophobic region of the fibril exhibiting anti-amyloidogenic activity. Additionally ZINC01758814 showed highest value of vibrational entropy which elucidates flexibility of the molecule and the associated conformational and rotational degrees of freedom.

In order to identify compounds with dual potency a shortened list was prepared using a criterion of PharmaGist score greater than 20 and predicted antiamyloidogenicity (IC_{50}) less than $550 \mu\text{M}$. Therefore, ZINC31167296, ZINC16027834, ZINC01758814, ZINC02119155 and ZINC53276039 may be considered as compounds with dual potency to be validated using *in vitro* studies.

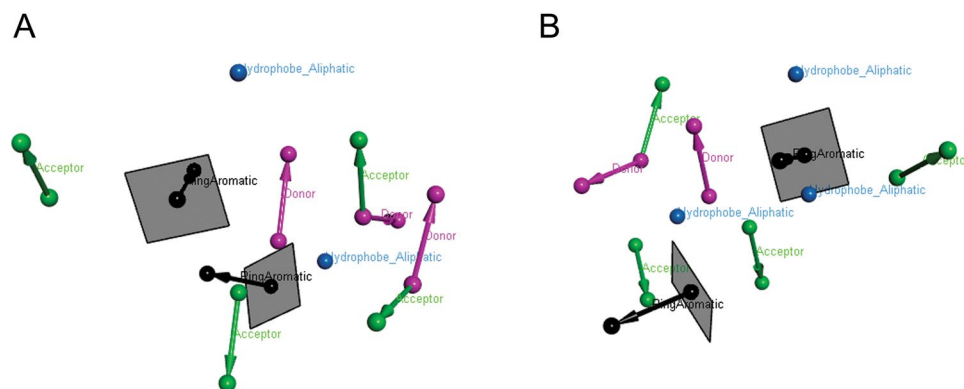


Figure 6. Pharmacophore models obtained from PharmaGist. (A) Model 1- denotes the model generated from 12 BACE1 inhibitors (with IC_{50}) and (B) Model 2- denotes the model generated from 9 BACE1 inhibitors (with K_i).

Pharmacophore	Score	No of spatial features	No. of aromatic feature	No of Hydrophobic features	No. of donor features	No. of acceptor features	No of positives	No of negatives
Model 1	41	9	2	2	3	4	0	0
Model 2	35.3	10	2	3	2	4	0	0

Table 4. The two pharmacophore models obtained from PharmaGist using active inhibitors of BACE1.

Pharmacophore model	Score
Model 1	0.7154
Model 2	0.6022

Table 5. Overlay score between the pharmacophore model generated by PharmaGist and reference pharmacophore.

PharmaGist models	Enrichment factor in top 10	Enrichment factor in top 15
Model 1	10.52	12.28
Model 2	13.157	13.157

Table 6. Performance of the pharmacophore models.

In vitro validation. *Probing the BACE1 inhibition ability of the screened compounds.* Four (bold-faced) out of the five screened compounds enlisted in Table 7 are commercially available. Therefore four compounds were procured for validation. All of them showed moderate BACE1 inhibition ability at a concentration of $1\ \mu\text{M}$ (Table 8) with the $SPI_{1\mu\text{M}}$ (single point inhibition percentage) of the compounds varying across the range of 16.98 to 21.6. Interestingly, $SPI_{1\mu\text{M}}$ of myricetin (20.6), a known BACE1 inhibitor falls within this range. $SPI_{1\mu\text{M}}$ of gossypetin, another flavonoid, was found to be low with a value of 4.6.

Probing the antiamyloidogenicity of the screened compounds. On adding Thioflavin T (ThT) to the 3-day aged $A\beta_{25-35}$ samples the ThT fluorescence was enhanced due to restricted rotation of the ThT rings in the hydrophobic environment of the amyloid fibrils. Interestingly, all the four compounds procured, showed a reduction in fluorescence when co-incubated with $A\beta_{25-35}$ for 3 days. Our results (Fig. 7) indicate that the compounds show an ability to interfere with the formation of the amyloid fibrils thus reflecting a reduction in the degree of fibril aggregates. One of the compounds, ZINC01758814 showed highest anti-amyloidogenicity with more than 50% reduction in the fibril formation.

Conclusions

Current treatment of AD is limited to temporary symptomatic support to the cognitive abilities of the patient and is therefore not curative. Since $A\beta$ is believed to play an early role which is crucial in the etiology of AD, therapeutics targeting production of $A\beta$ or its aggregation would lead to more effective disease-modifying drugs. Scientists are working hard to discover/design inhibitor molecules for the protease BACE1 that is known to cleave APP. Although further cleavage of APP by γ -secretase produces $A\beta$, cleavage by BACE1 forms the rate limiting step. Development of potent BACE1 inhibitors is still being persuaded with much less success than expected.

Compounds	Selection criterion for		
	BACE1 inhibitor based on pharmacophore model 1	BACE1 inhibitor based on pharmacophore model 2	Antiamyloidogenic compound
ZINC31167296	Score >20	Score >20	Predicted IC ₅₀ < 550 μM
ZINC16027834			
ZINC01758814			
ZINC02119155			
ZINC53276039			

Table 7. Common compounds predicted to be potent BACE1 inhibitor as well as potent antiamyloidogenic agent.

Compound	SPI ₁ μM
Known inhibitor (Myricetin)	20.6
ZINC01758814	16.98
ZINC16027834	20.7
ZINC02119155	19.62
ZINC53276039	21.6
Weak inhibitor (Gossypetin)	4.6

Table 8. BACE1 inhibition ability.

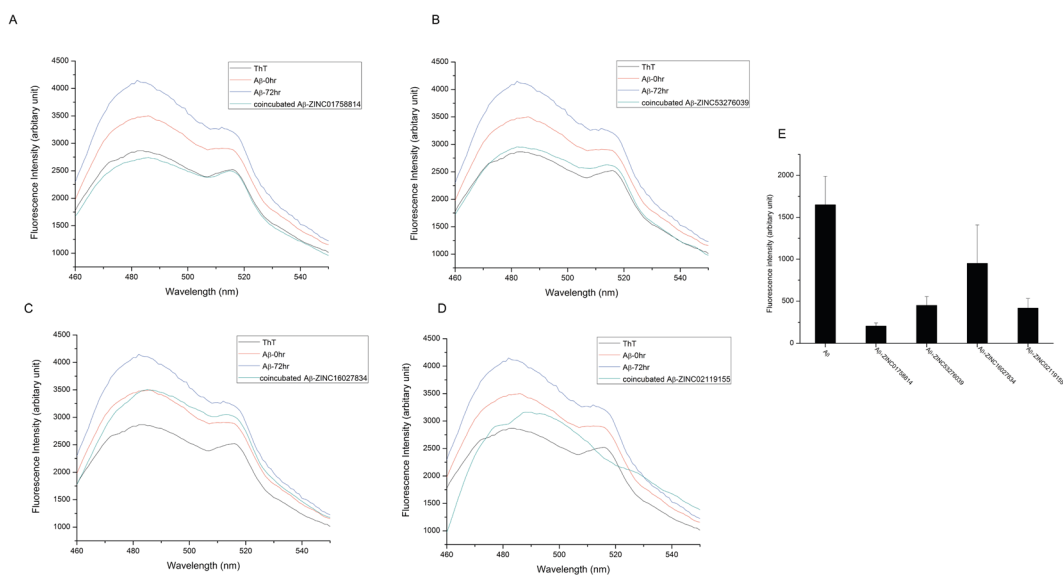


Figure 7. Thioflavin T assay for compounds (A) ZINC01758814 (B) ZINC53276039 (C) ZINC16027834 (D) ZINC02119155 (E) Comparison of ThT assay for A β_{25-35} aggregation for all four compounds.

One of the reasons for the lack of success is probably the flexibility of BACE1. The enzyme is known to undergo a conformational transition from open to closed through a half open half closed (HOHC) conformation⁶⁸. Thus, for such an enzyme, use of a static crystal structure during docking would fail to accommodate diverse range of ligand conformations. On the contrary, taking into account receptor flexibility considering multiple crystal structures with slight movement of amino acid residues, loops or turns in different directions gives the ligand a better chance to accommodate itself in the active site of the enzyme. The other reason for the low success rate is the multi-factorial aspect of the disease. For a compound to be used as a therapeutic agent for AD it is thus not enough to target a single molecule. Rather a compound acting on more than one target would fit better.

In the present study, we have considered multiple crystal structures for closed and semi-open conformations of BACE1 with bound ligands of varied size and shape in order to account for the flexibility of the enzyme. By re-docking and cross-docking we could select two most suitable and diverged representative structures from each group, that could adopt ligands of varied size. The closed and semi-open conformations were subsequently used to virtually screen the natural product database of ZINC and from the two lists of screened compounds, 23 common compounds were selected. The common compounds were believed to outperform the other ligands by their versatility to bind to different conformations of BACE1. The common compounds were further arranged

in the descending order of their similarity to known BACE1 inhibitors using two pharmacophore models. For the multi-factorial aspect of the disease, these virtually screened BACE1 inhibitors were further predicted for anti-amyloidogenic activity. Finally, we validated our studies through *in vitro* experimentations of four of the screened compounds. ZINC16027834, ZINC02119155, ZINC01758814 and ZINC53276039 showed moderate BACE1 inhibition activity. Additionally, all the compounds showed the ability to interfere with amyloid fibril formation. Hence these compounds might act as potential starting compounds to target BACE1 inhibition as well as anti-amyloidogenicity and therefore be used in AD therapeutics. The incorporation of flexibility considering multiple structures for different conformations of BACE1 with subsequent use of ligand-based and structure-based approaches for selection of compounds against dual targets makes this effort different from the ones pursued before. Thus we have successfully developed a bilayered (structure and ligand-based approach), dual-target directed screening methodology imbibing receptor flexibility to identify natural compounds that might serve as scaffolds for leads to be used in therapeutics for AD.

References

- Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
- Roberds, S. L., Anderson, J., Basu, G. & Bienkowski, M. J. BACE knockout mice are healthy despite lacking the primary beta-secretase activity in brain: implications for Alzheimer's disease therapeutics. *Hum. Mol. Genet.* **1**, 1317–1324 (2001).
- Ghosh, A. K. *et al.* Design of Potent Inhibitors for Human Brain Memapsin 2 (β -Secretase). *J. Am. Chem. Soc.* **122**, 3522–3523 (2000).
- Hong, L. *et al.* Structure of the Protease Domain of Memapsin 2 (β -Secretase) Complexed with Inhibitor. *Science* **290**, 150–153 (2000).
- Asai, M. *et al.* The novel β -secretase inhibitor KMI-429 reduces amyloid β peptide production in amyloid precursor protein transgenic and wild-type mice. *J. Neurochem.* **96**, 533–540 (2006).
- Jain, P. & Jadhav, H. R. Quantitative structure activity relationship analysis of aminoimidazoles as BACE-I inhibitors. *Med. Chem. Res.* **22**, 1740–1746 (2013).
- Li, Z., Zhou, M., Wu, F., Li, R. & Ding, Z. Self-organizing molecular field analysis on human β -secretase nonpeptide inhibitors: 5, 5-disubstituted aminohydantoins. *Eur. J. Med. Chem.* **46**, 58–64 (2011).
- Al-Nadaf, A., Sheikha, G. A. & Taha, M. O. Elaborate ligand-based pharmacophore exploration and QSAR analysis guide the synthesis of novel pyridinium-based potent β -secretase inhibitory leads. *Bioorganic Med. Chem.* **18**, 3088–3115 (2010).
- Jeon, S. Y., Bae, K., Seong, Y. H. & Song, K. S. Green tea catechins as a BACE1 (β -secretase) inhibitor. *Bioorganic Med. Chem. Lett.* **13**, 3905–3908 (2003).
- Jia-kuo, L. *et al.* Design and synthesis of cyclic acylguanidines as BACE1 inhibitors. *Chinese Chem. Lett.* **26**, 1–4 (2015).
- Forman, M. *et al.* The novel bace inhibitor MK-8931 dramatically lowers cerebrospinal fluid ab peptides in healthy subjects following single- and multiple-dose administration. *Alzheimer's Dement.* **8**, P704 (2012).
- Bernier, F. *et al.* Clinical study of E2609, a novel BACE1 inhibitor, demonstrates target engagement and inhibition of BACE1 activity in CSF. *Alzheimer's Dement.* **9**, P886 (2013).
- Vassar, R. BACE1 inhibitor drugs in clinical trials for Alzheimer's disease. *Alzheimer's Res. Ther.* **6**, 1–14 (2014).
- Miller, B. W., Willett, K. C. & Desilets, A. R. Rosiglitazone and Pioglitazone for the Treatment of Alzheimer's Disease. *Ann. Pharmacother.* **45**, 1416–1424 (2011).
- Liu, S., Fu, R., Zhou, L. H. & Chen, S. P. Application of Consensus Scoring and Principal Component Analysis for Virtual Screening against β -Secretase (BACE-1). *PLoS ONE* **7**(6), e38086 (2012).
- Liu, H. *et al.* AlzPlatform: An Alzheimer's Disease Domain-Specific Chemogenomics Knowledgebase for Polypharmacology and Target Identification Research. *J. Chem. Inf. Model.* **54**(4), 1050–1060 (2014).
- Zeng, H. & Wu, X. Alzheimer's disease drug development based on Computer-Aided Drug Design. *Eur. J. Med. Chem.* **121**, 851–863 (2016).
- Manoharan, P., Chennou, K. & Ghoshal, N. Computational analysis of BACE1-ligand complex crystal structures and linear discriminant analysis for identification of BACE1 inhibitors with anti P-glycoprotein binding property. *J. Biomol. Struct. Dyn.* **36**(1), 1–15 (2018).
- Clark, C. M. & Karlawish, J. H. Alzheimer Disease: Current Concepts and Emerging Diagnostic and Therapeutic Strategies. *Ann. Intern. Med.* **138**, 400–10 (2003).
- Cummings, J. L. Alzheimer's disease. *N. Engl. J. Med.* **351**, 56–67 (2004).
- Scarpini, E., Schelterns, P. & Feldman, H. Treatment of Alzheimer's disease; current status and new perspectives. *Lancet Neurol.* **2**, 539–547 (2003).
- Hong, L. & Tang, J. Flap Position of Free Memapsin 2 (β -Secretase), a Model for Flap Opening in Aspartic Protease Catalysis. *Biochemistry* **43**, 4689–4695 (2004).
- Shimizu, H. *et al.* Crystal Structure of an Active Form of BACE1, an Enzyme Responsible for Amyloid β Protein Production. *Mol. Cell. Biol.* **28**, 3663–3671 (2008).
- Yuan, J. *et al.* Structure-based design of beta-site APP cleaving enzyme 1 (BACE1) inhibitors for the treatment of Alzheimer's disease. *J. Med. Chem.* **56**, 4156–4180 (2013).
- Xu, Y. *et al.* Flexibility of the flap in the active site of BACE1 as revealed by crystal structures and molecular dynamics simulations. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 13–25 (2012).
- Malamas, M. S. *et al.* New pyrazolyl and thienyl aminohydantoins as potent BACE1 inhibitors: Exploring the S2' region. *Bioorganic Med. Chem. Lett.* **21**, 5164–5170 (2011).
- Zhu, Z. *et al.* Discovery of cyclic acylguanidines as highly potent and selective β -site amyloid cleaving enzyme (BACE) inhibitors: Part I - Inhibitor design and validation. *J. Med. Chem.* **53**, 951–965 (2010).
- Cumming, J. N. *et al.* Structure based design of iminohydantoin BACE1 inhibitors: Identification of an orally available, centrally active BACE1 inhibitor. *Bioorganic Med. Chem. Lett.* **22**, 2444–2449 (2012).
- Gravenfors, Y. *et al.* New Aminoimidazoles as beta-Secretase (BACE-1) Inhibitors Showing Amyloid-beta (A β) Lowering in Brain. *J. Med. Chem.* **55**, 10316 (2012).
- Efremov, I. V. *et al.* Discovery and Optimization of a Novel Spiropyrrolidine Inhibitor of Beta-Secretase (BACE1) through Fragment-Based Drug Design. *J. Med. Chem.* **55**, 9069–9088 (2012).
- Ju, Y. *et al.* Identification of Novel BACE1 Inhibitors by Combination Pharmacophore Modeling, Structure-Based Design and *In Vitro* Assay. *Curr Comput Aided Drug Des.* **12**(1), 73–82 (2016).
- Polgar, T., Magyar, C., Simon, I. & Keseru, G. M. Impact of Ligand Protonation on Virtual Screening against β -Secretase (BACE1). *J. Chem. Inf. Model.* **47**(6), 2366–2373 (2007).
- Bajda, M., Jonczyk, J., Malawska, B. & Filipek, S. Application of computational methods for the design of BACE-1 inhibitors: Validation of *in silico* modelling. *Int. J. Mol. Sci.* **15**, 5128–5139 (2014).

34. Källblad, P., Todorov, N. P., Willems, H. M. G. & Alberts, I. L. Receptor flexibility in the in Silico screening of reagents in the S1' pocket of human collagenase. *J. Med. Chem.* **47**, 2761–2767 (2004).
35. Alberts, I. L., Todorov, N. P. & Dean, P. M. Receptor flexibility in de novo ligand design and docking. *J. Med. Chem.* **48**, 6585–6596 (2005).
36. Tian, S. *et al.* Assessing ensemble docking-based virtual screening strategy for kinase targets by considering protein flexibility. Assessing ensemble docking-based virtual screening strategy for kinase targets by considering protein flexibility. *J. Chem. Inf. Model.* (2014).
37. Craig, I. R., Essex, J. W. & Spiegel, K. Ensemble docking into multiple crystallographically derived protein structures: An evaluation based on the statistical analysis of enrichments. *J. Chem. Inf. Model.* **50**, 511–524 (2010).
38. Huang, S. Y. & Zou, X. Q. Ensemble Docking of Multiple Protein Structures: Considering Protein Structural Variations in Molecular Docking. *Proteins: Struct. Funct. Genet.* **66**, 399–421 (2007).
39. Hou, X., Li, K., Yu, X., Sun, J. & Fang, H. Protein Flexibility in Docking-Based Virtual Screening: Discovery of Novel Lymphoid-Specific Tyrosine Phosphatase Inhibitors Using Multiple Crystal Structures. *J. Chem. Inf. Model.* **55**, 1973–1983 (2015).
40. DeLano, W. L. The PyMOL Molecular Graphics System, DeLano Scientific: San Carlos, CA, 2002 (2009).
41. Chakraborty, S., Kumar, S. & Basu, S. Conformational transition in the substrate binding domain of β -secretase exploited by NMA and its implication in inhibitor recognition: BACE1–myricetin a case study. *Neurochem. Int.* **58**(8), 914–23 (2011).
42. Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–89 (1996).
43. Wessa, P. Free Statistics Software, Office for Research Development and Education, version 1.1.23-r7, 2016, <http://www.wessa.net/>.
44. Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, San Diego: Dassault Systèmes, 2016 Release 2017.
45. Li, J. *et al.* The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* **79**, 2794–812 (2011).
46. Ghosh, A. K. *et al.* Design, Synthesis, and X-ray Structure of Potent Memapsin 2 (α -Secretase) Inhibitors with Isophthalamide Derivatives as the P2-P3-Ligands. *J. Med. Chem.* **50**, 2399–2407 (2007).
47. Mandal, M. *et al.* Design and Validation of Bicyclic Iminopyrimidinones As Beta Amyloid Cleaving Enzyme 1 (BACE1) Inhibitors: Conformational Constraint to Favor a Bioactive Conformation. *J. Med. Chem.* (2012).
48. Hypercube, Inc, USA. Hyperchem (2002).
49. http://www.niper.ac.in/pi_dev_tools/DruLiToWeb/DruLiTo_index.html.
50. Schneidman-Duhovny, D., Dror, O., Inbar, Y., Nussinov, R. & Wolfson, H. J. PharmaGist: a webserver for ligand-based pharmacophore detection. *Nucleic Acids Res.* **36**, 223–228 (2008).
51. Dror, O., Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. Novel approach for efficient pharmacophore-based virtual screening: Method and applications. *J. Chem. Inf. Model.* **49**, 2333–2343 (2009).
52. Clarke, B. *et al.* BACE-1 inhibitors part 2: Identification of hydroxy ethylamines (HEAs) with reduced peptidic character. *Bioorganic Med. Chem. Lett.* **18**, 1017–1021 (2008).
53. Charrier, N. *et al.* Second generation of BACE-1 inhibitors. Part 1: The need for improved pharmacokinetics. *Bioorganic Med. Chem. Lett.* **19**, 3664–3668 (2009).
54. Charrier, N. *et al.* Second generation of BACE-1 inhibitors part 3: Towards non hydroxyethylamine transition state mimetics. *Bioorganic Med. Chem. Lett.* **19**, 3674–3678 (2009).
55. Cumming, J. N. *et al.* Rational design of novel, potent piperazinone and imidazolidinone BACE1 inhibitors. *Bioorganic Med. Chem. Lett.* **18**, 3236–3241 (2008).
56. Cheng, Y. *et al.* From fragment screening to *in vivo* efficacy: Optimization of a series of 2-aminoquinolines as potent inhibitors of beta-site amyloid precursor protein cleaving enzyme 1 (BACE1). *J. Med. Chem.* **54**, 5836–5857 (2011).
57. Rueeger, H. *et al.* Discovery of cyclic sulfone hydroxyethylamines as potent and selective beta-site APP-cleaving enzyme 1 (BACE1) inhibitors: structure-based design and *in vivo* reduction of amyloid beta-peptides. *J. Med. Chem.* **55**, 3364–3386 (2012).
58. Sund, C. *et al.* Design and synthesis of potent hydroxyethylamine (HEA) BACE-1 inhibitors carrying prime side 4,5,6,7-tetrahydrobenzazole and 4,5,6,7-tetrahydropyridinoazole templates. *Bioorganic Med. Chem. Lett.* **22**, 6721–6727 (2012).
59. Pennington, L. D. *et al.* Hydroxyethylamine-based inhibitors of BACE1: P1-P3 macrocyclization can improve potency, selectivity, and cell activity. *Bioorganic Med. Chem. Lett.* **23**, 4459–4464 (2013).
60. Liu, T., Lin, Y., Wen, X., Jorissen, R. N. & Gilson, M. K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **35**, 198–201 (2007).
61. Chakraborty, S. & Basu, S. Insight into the anti-amyloidogenic activity of polyphenols and its application in virtual screening of phytochemical database. *Med. Chem. Res.* **23**, 5141–5148 (2014).
62. Sahigara, F. *et al.* Comparison of different approaches to define the applicability domain of QSAR models. *Molecules* **17**, 4791–4810 (2012).
63. CODESSA Version 2.7.16. University of Florida and institute of chemical physics, University of Tartu, Estonia and Semi Chem Inc (1995).
64. AMPAC 9.0 Ó Semichem, 7128 Summit, Shawnee, K. S. 66216 (1994).
65. Chakraborty, S., Bandyopadhyay, J., Chakraborty, S. & Basu, S. Multi-target screening mines hesperidin as a multi-potent inhibitor: Implication in Alzheimer's disease therapeutics. *Eur. J. Med. Chem.* **121**, 810–822 (2016).
66. Chakraborty, S. & Basu, S. Multi-functional activities of citrus flavonoid narirutin in Alzheimer's disease therapeutics: An integrated screening approach and *in vitro* validation. *Int. J. Biol. Macromol.* **103**, 733–743 (2017).
67. Liu, S., Fu, R., Cheng, X., Chen, S.-P. & Zhou, L.-H. Exploring the binding of BACE-1 inhibitors using comparative binding energy analysis (COMBINE). *BMC Struct. Biol.* **12**, 21 (2012).
68. Chakraborty S., Basu S. *Translating the knowledge of functional dynamics towards designing inhibitors of BACE1, a key aspartate protease in Alzheimer's disease in Proteases in Human Diseases* (eds Chakraborti, S., Chakraborti, T.) 171–184 (Springer, 2017).

Acknowledgements

SB and SD gratefully acknowledge the financial support from the Department of Science and Technology (Govt. of India, Ref No: SR/SO/BB-0105/2010, dated: 20/06/2012). We also acknowledge the Centre for High Performance Computing for Modern Biology and Department of Biophysics Molecular Biology & Bioinformatics, University of Calcutta for the computational resources and fluorimetric studies respectively.

Author Contributions

S.D., S.C. and S.B. conceived the work and S.D. carried it out. S.D. and S.B. analyzed the results. S.D. wrote the manuscript and S.B. and S.C. corrected it critically.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-40271-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019