J Comput Aided Mol Des (2011) 25:135–144
DOI 10.1007/s10822-010-9405-x

Quantitative structure–activity relationship analysis of β-amyloid aggregation inhibitors

Shiri Stempler · Michal Levy-Sakin · Anat Frydman-Marom · Yaniv Amir · Roni Scherzer-Attali · Ludmila Buzhansky · Ehud Gazit · Hanoch Senderowitz

Received: 10 October 2010 / Accepted: 3 December 2010 / Published online: 17 December 2010

Abstract Inhibiting the aggregation process of the β-amyloid peptide is a promising strategy in treating Alzheimer's disease. In this work, we have collected a dataset of 80 small molecules with known inhibition levels and utilized them to develop two comprehensive quantitative structure–activity relationship models: a Bayesian model and a decision tree model. These models have exhibited high predictive accuracy: 87% of the training and test sets using the Bayesian model and 89 and 93% of the training and test sets, respectively, by the decision tree model. Subsequently these models were used to predict the activities of several new potential β-amyloid aggregation inhibitors and these predictions were indeed validated by in vitro experiments. Key chemical features correlated with the inhibition ability were identified. These include the electro-topological state of carbonyl groups, AlogP and the number of hydrogen bond donor groups. The results demonstrate the feasibility of the developed models as tools for rapid screening, which could help in the design of novel potential drug candidates for Alzheimer's disease.

Keywords Amyloid · Structure–activity relationship · Bayesian classifier · Decision tree classifier · Alzheimer's disease

Introduction

The formation of amyloid fibrils is the major characteristics of various unrelated disorders [1–3]. Despite its fundamental clinical importance, the mechanism of amyloid formation is not fully understood. Alzheimer's disease (AD) is one of the key amyloid diseases. AD is a devastating neurodegenerative disorder, in which brain amyloid is composed almost entirely of β-amyloid (Aβ) peptide [4]. Many studies have pointed out a pathological role of Aβ aggregation in the development of AD [5–7]. Considerable evidence has indicated that Aβ fibrils are toxic [8, 9]. Several recent studies have suggested that low molecular weight species of aggregated Aβ may be more toxic than higher molecular weight fibrils [10–12].

At present, there is no cure or effective treatment for AD. Currently available drugs for the treatment of AD do not arrest the underlying disease process, but only treat its symptoms [13]. Therefore, there is an urgent need for a disease modifying therapy that would slow the neuronal loss and the cognitive decline.

Several therapeutic strategies are currently under development [14, 15]. We have focused on the prevention of the Aβ aggregation process. The prevention of amyloid formation is a potential goal in the therapy or the prevention of amyloid diseases, and similar strategies are possible for other related amyloid disorders [16].

Electronic supplementary material The online version of this article (doi:10.1007/s10822-010-9405-x) contains supplementary material, which is available to authorized users.

S. Stempler · M. Levy-Sakin · A. Frydman-Marom · Y. Amir · R. Scherzer-Attali · L. Buzhansky · E. Gazit (✉)
Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
e-mail: vpr@tauex.tau.ac.il

S. Stempler
e-mail: shirist2@post.tau.ac.il

H. Senderowitz (✉)
Department of Chemistry, Bar-Ilan University, 52900 Ramat Gan, Israel
e-mail: hanoch@mail.biu.ac.il

Various unrelated small molecules have been studied and shown to inhibit Aβ aggregation and Aβ toxicity, both in vitro and in vivo [17–20]. The aromatic moieties of Aβ have been shown to have a central role in the molecular recognition and the self-assembly processes, leading to the formation of Aβ fibrils [21]. Since the aromatic core of Aβ peptide was found important in the acceleration of the aggregation process, it can be used for designing aggregation inhibitors [22, 23]. Many researchers have reported the Aβ aggregation inhibitory ability of aromatic-rich small molecules [18–20, 24, 25]. The various aromatic inhibitors include different polyphenols, indole derivatives and small aromatic peptides. Although the importance of aromatic interactions in inhibiting Aβ aggregation was emphasized, [26, 27] in many of the cases variation of substituent on the aromatic ring resulted in changes in the inhibitory activity levels [28]. This may therefore suggest that the inhibitory activity is not exclusively dependent on the aromatic groups.

A large number of small molecules have been experimentally tested for their Aβ aggregation inhibition activity and have been reported in the literature. However, previous structure activity relationship (SAR) studies for the identification of the main characteristics that are significant for the inhibition ability included only limited data of small sets of similar compounds. These studies included SAR of *N*-phenyl anthranilic acid analogs, [29] SAR of amyloid aggregation inhibitors based on curcumin scaffold [30] and the study of polyphenols [27]. To the best of our knowledge, a comprehensive computational study for this purpose has not been reported yet.

A widely used in silico method in drug design is quantitative structure–activity relationship (QSAR). In the current research, we have developed classification QSAR models to find the correlation between the physicochemical properties of the various compounds and their inhibition activity. These models were successfully applied to the prediction of new potent inhibitors and could help gain further understanding on the amyloid self assembly mechanism.

Computational methods

Data set and molecular descriptors

A dataset of β-amyloid inhibitors composed of small aromatic compounds for which β-amyloid inhibition was measured experimentally was collected from the literature (see supplementary Tables S1, S2). 3D structures of the compounds were generated using ChemDraw version 7.0. For modeling purposes, experimental data were classified as active (1), for good inhibitors, or inactive (0) for the non-inhibitor molecules, using a threshold criterion. This binary classification was performed since quantitative inhibition data were not reported for all the molecules. In addition, it has allowed us to combine inhibition data from different sources. Inhibitors were defined as compounds exhibiting at least 60% aggregation inhibition when tested in a concentration ratio smaller than 1:20, protein to inhibitor. Compounds with mid-range inhibition activity were removed from the model, in order to reduce noise. In order to construct and validate the QSAR models, the dataset was divided into training and test sets comprising 80 and 20% of the total dataset, respectively. The test compounds were selected using the 'Find Diverse Molecules' module implemented in the Discovery Studio (DS) 2.0 software [31] which ensures that both the training and the test sets cover similar regions of property space. In addition, a similar ratio of inhibitors to non-inhibitors in the training and test sets was kept. QSAR models were developed based on the training set compounds. The test set compounds were used for external validation of the models generated. In order to further validate our approach, several random partitions of training and test sets were performed and independent QSAR models were generated for each partition. Similar results were obtained in all cases.

150 descriptors were computed with the DS 2.0 software [31]. These descriptors included AlogP, molecular property counts, surface area, topological descriptors and E-state keys (see supplementary Table S3). Feature selection was performed using two different algorithms: the Bayesian model in the DS, in which five descriptors were chosen based on the bins statistics data, and the CFS attribute evaluator in WEKA 3.6 [32].

QSAR methodology

Two different QSAR models were built based on the training set: naïve Bayesian classifier as implemented in DS and recursive partitioning J48 decision tree (DT) model as implemented in WEKA.

Bayesian model

Bayesian analysis is a statistical categorization method. It is based on the mathematical theorem of Thomas Bayes [33]. Models are created through a straightforward learn by-example paradigm whereby the system learns to distinguish between the active sample data and the background data. The learning process generates a large set of Boolean features from the input descriptors, and then collects the frequency of the occurrence of each feature in the active subset and in all of the data samples. The Laplacian corrected estimator is used to adjust the uncorrected probability estimate of a feature to account for the different sampling frequencies of different features [34].

Decision tree (DT) model

The J48 algorithm [35] was used for generating a classification DT from a set of labeled training data. According to this method, each attribute of the data (i.e. descriptor) is considered for splitting the dataset into smaller subsets. The attribute with the highest normalized information gain is used for performing the actual split. Afterwards, the algorithm recurs on the smaller subsets. The splitting procedure stops if all instances in a subset belong to the same class, and then a leaf node is created in the DT with that class. In the present study, default parameters were used (see supplementary information).

Model evaluation

Leave-one-out (LOO) and tenfold Cross-Validation were used in the development of the Bayesian classifier and the DT, respectively. Furthermore, receiver operating characteristic (ROC) plots of the input data were computed. ROC graphs are 2D graphs in which the true positive rate (TPR) is plotted on the Y axis and the false positive rate (FPR) is plotted on the X axis [36]. TPR of a classifier is defined as: $\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$ and FPR is defined as: $\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$ (where TP, FP, TN, FN is true positive, false positive, true negative and false negative, respectively). The area under the ROC curve (AUC) indicates how often the model correctly identifies the TP and the TN and is commonly used to define model accuracy. The closer the area under the ROC curve is to 1, the greater is the predictive ability of the model. The ROC curve of the Bayesian model is presented in Fig. 1.

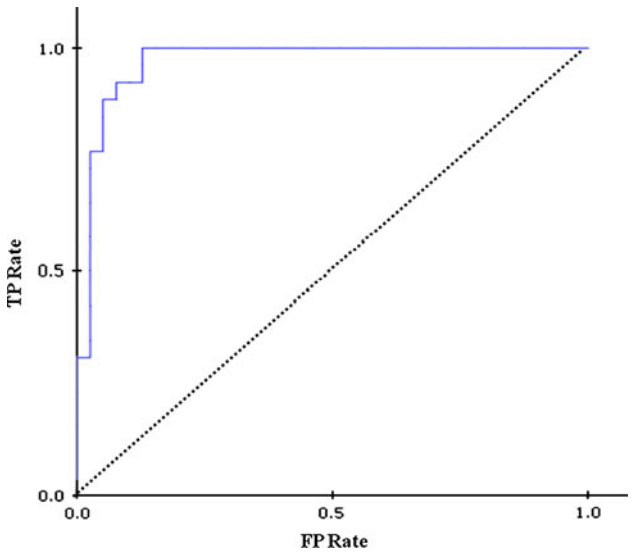

Fig. 1 ROC curve of the Bayesian model. The diagonal line represents a model with no discriminatory power in predicting the activity

The resulting models were further tested against chance correlation using Y-scrambling [37]. In this test, the dependent variable vector is randomly shuffled and a QSAR model is developed using the original independent variables matrix. Finally, the true predictive power of the models was evaluated on an external test set.

Virtual screening of chemical databases

For the mining of large molecular databases with the above QSAR models, we used the Zinc Database [38] and the Cambridge Structural Database (CSD) [39]. These databases were pre-filtered for small aromatic molecules, commercially available through several specific vendors. In addition, we have searched our internal database of small aromatic molecules. Predictions were made using a *consensus approach*, i.e. only compounds that were consensually predicted by both models (i.e. Bayesian and DT) to belong to the same inhibition class, were considered. All other predictions were considered unreliable and were excluded.

Results and discussion

A total set of 80 small aromatic molecules, published in the literature or assayed in our laboratory were collected (see supplementary Tables S1, S2). The inhibition levels of the selected compounds were all tested under various common protocols such as Thioflavin T (ThT) fluorescence assay and electron microscopy. For classification QSAR analysis, the inhibition activity was transformed into a binary format (see "Computational methods"). The dataset was divided into training and test sets. The training set contained 65 of the compounds, with a ratio of 40: 60% of inhibitors to non-inhibitors. The models were developed based on the training set. The test set contained 15 compounds with a similar ratio of inhibitors to non-inhibitor compounds.

In order to generate models with predictive ability over a wide range of structural classes, these models should be based on a heterogeneous dataset. The structural and the chemical diversity of the compounds in the training dataset are illustrated in Table S4. All compounds contain at least one aromatic ring.

2D descriptors were calculated using the DS software (see "Methods"). Two different algorithms were applied for the selection of the statistically significant descriptors. The list of the descriptors selected by each algorithm is provided in Table 1.

Table 1 Significant descriptors identified by the two different algorithms

Descriptor name	Bayesian classifier	WEKA attribute selection
AlogP*	+	+
ES_sum_dO*	+	+
HBD_count	−	+
IC	+	−
Molecular_PolarSurfaceArea	+	−
Num_AromaticBonds	+	−

* Descriptors that were found to be significant by both algorithms

Bayesian model

The first classification model was developed by using the Bayesian model algorithm as implemented in DS. LOO cross validation was employed in the development of the model as internal validation. The model developed classified correctly a total percentage of 87.7% of the training set compounds (Table 2). A ROC curve was computed in order to check sensitivity and specificity of the model. The AUC calculated was 0.864. For external validation, model prediction of the test set compounds was applied (Table 2). Out of the 15 test set compounds, the model predicted correctly all active compounds and 8 out of 10 inactive compounds. To further evaluate the robustness of the model, we used Y scrambling method. Models with low prediction scores were generated, implying that the model was not based on a chance correlation.

DT model

A classification tree model was built using the WEKA software and is depicted in Fig. 2. This figure shows that the initial training set compounds were split into two nodes. Whereas the first one, corresponding to the inactive compounds, was not split further, the second node was subdivided to two nodes on the basis of the HBD_Count feature, representing the hydrogen bond donors count. Molecules with high HBD_Count are predicted by the model as inhibitors, while molecules with no hydrogen bond donors are predicted as non-inhibitors. These findings emphasize the importance of hydrogen bonds in the interaction of the inhibitor molecules with Aβ peptide. Several experimental studies have proposed the importance of hydrogen bond interactions between the inhibitors and the Aβ peptide [18, 40, 41]. These interactions are also known as the primarily interactions stabilizing the amyloid fibrils core structure. [42] Therefore, hydrogen bond interactions between the inhibitors and the peptide can disrupt the bonding between Aβ monomers and induce the inhibition of the aggregation process.

In the mid-range of HBD_Count values (1–4), the AlogP descriptor determines the prediction. Molecules with high values of AlogP, which account for more hydrophobicity, are predicted as inhibitors (Fig. 2), while molecules that are less hydrophobic predicted as non-inhibitors. These descriptors clearly corroborate the empirical findings according to which the hydrophobicity of the molecule plays a central role in the inhibition process. [18, 20, 41] Furthermore, these findings suggest that hydrogen bonds and hydrophobic interactions are complementary in terms of inhibition ability.

The prediction ability of the model was estimated with tenfold cross-validation (Table 2). The AUC score was 0.824. For additional validation, prediction of the 15 test set compounds was performed: Five out of five active compounds and nine out of ten inactive compounds were predicted correctly. Similar to the Bayesian model, low prediction scores were generated for the Y scrambled models.

Models analysis

For further analysis of the models, principal component analysis (PCA) graphs were calculated for both models (Fig. 3). PCA allows the visualization of the dataset in the property space. The descriptors that were used for developing each model were used in the PCA calculation. The PCA of both models shows good separation between inhibitors and non-inhibitors in the descriptors space for both training set and test set compounds.

Significant descriptors in the models

Among the significant descriptors identified by both algorithms, was the AlogP descriptor, which accounts for the hydrophobicity of the molecule. Low values of AlogP, corresponding to low hydrophobicity, were highly correlated with inactive molecules. The importance of the hydrophobicity of the small molecules for their binding to the hydrophobic region of Aβ and for their inhibition

Table 2 Statistical parameters on the prediction of the two QSAR models

Model	M	AUC	Accuracy (training)	FP (training)	FN (training)	Accuracy (test)	FP (test)	FN (test)
Bayesian	5	0.864	0.877	7	1	0.866	2	0
DT	3	0.824	0.892	5	2	0.933	1	0

Data set was split into 65 compounds for training set and 15 for test set. *M* number of descriptors used

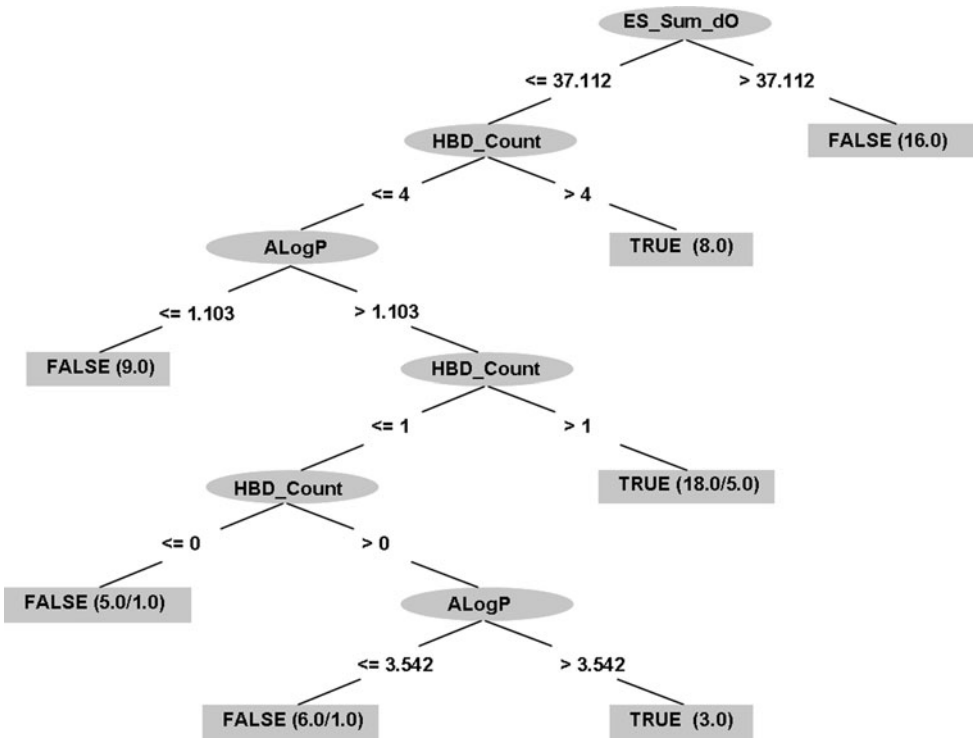

Fig. 2 Decision tree model developed for the prediction of Aβ aggregation inhibitors by the WEKA software

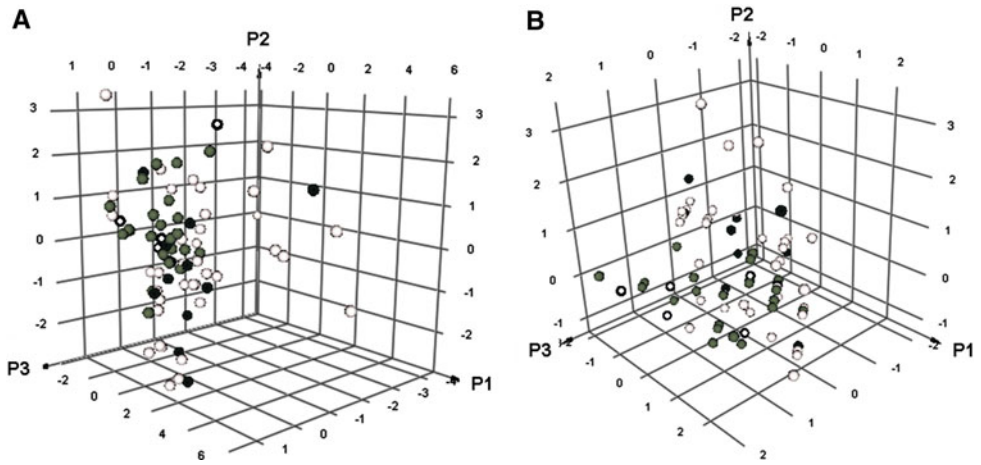

Fig. 3 PCA of the training and test sets compounds. **a** Compound representation in the descriptors space of the Bayesian model. **b** Compound representation in the descriptors space of the DT. *White circles* represent non-inhibitor molecules in the training set, *gray circles* represent inhibitors in the training set, *black circles* represent non-inhibitor molecules in the test set and *black circles* with *white dots* represent inhibitor molecules in the test set. The total variance explained is 0.92 and 0.99 in **a** and **b**, respectively

mechanism was suggested by several experimental studies [17, 43].

The ES_sum_dO [44, 45] descriptor was also selected by both models. It represents the E-state for oxygen atoms of carbonyl groups in the molecule. According to the models, molecules containing high values of E-state values have a higher probability to be predicted as inactive compounds. Carbonyl groups can interact with the peptide backbone amides in order to block the peptide aggregation process, as was shown in several works [18, 46]. However, our results demonstrate that a large number of carbonyl groups could disrupt the inhibition ability of the small molecules. This may be due to repulsive electrostatic effects, since the Aβ peptide itself contains negatively charged residues. Results of a recent study of indole derivatives as inhibitors of lysozyme amyloid fibrilization are in agreement with our models, and have shown a decreased activity of molecules containing carbonyl groups [47].

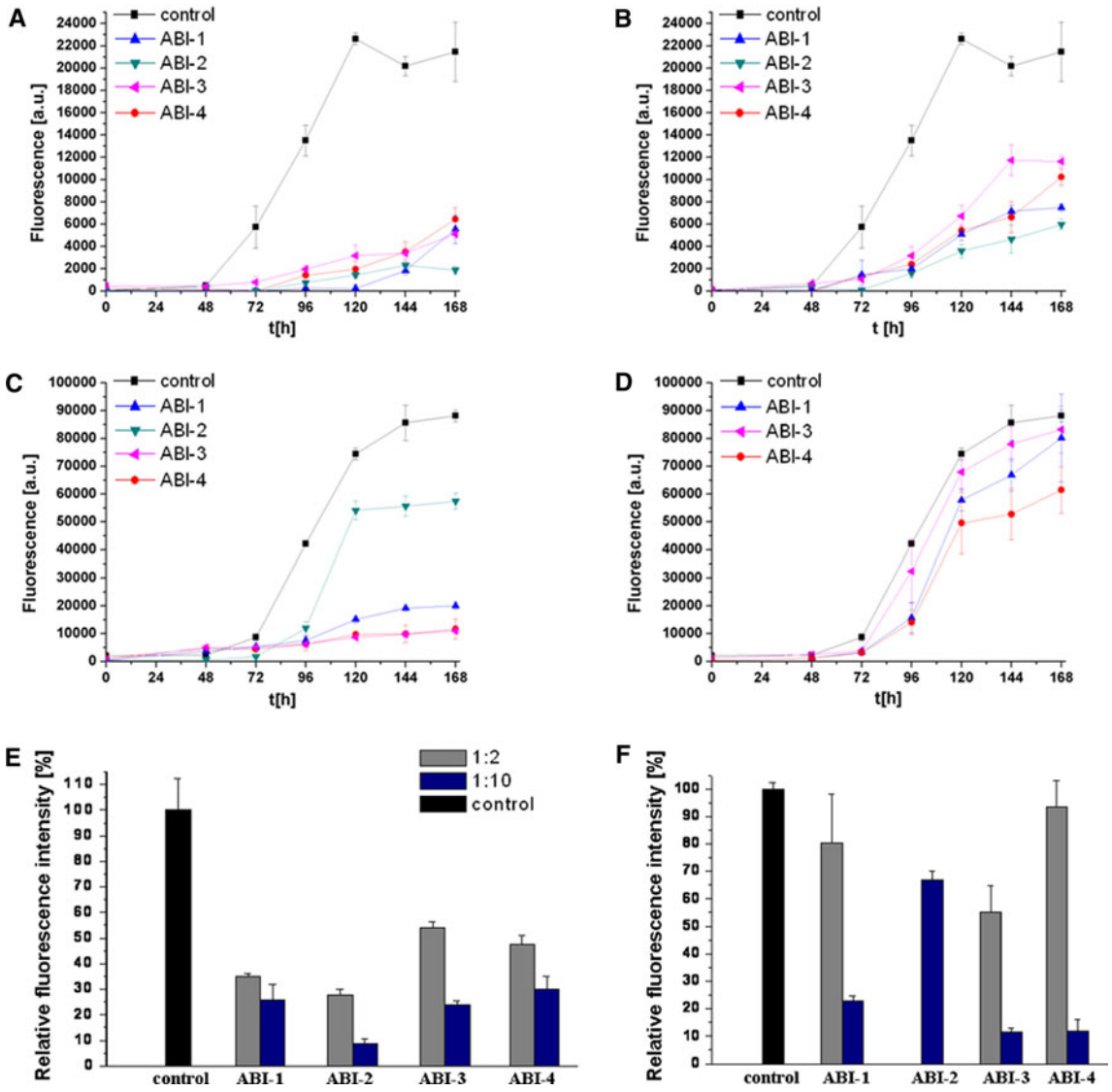

Fig. 4 Inhibition of Aβ aggregation by the four inhibitors. a 50 μM and **b** 10 μM of each inhibitor were added to a fixed amount of 5 μM Aβ (1–40). **c** 50 μM and **d** 10 μM of each inhibitor were added to a fixed amount of 5 μM Aβ (1–42). The kinetics of Aβ (5 μM) fibril formation with and without the inhibitors as assessed by the ThT fluorescence assay is shown. **e** and **f** Relative inhibition intensity for the different inhibitors in the ThT fluorescence experiments of Aβ (1–40) and of Aβ (1–42), respectively, after incubation for 7 days

Virtual screening

In the second phase of this study, the QSAR models were applied to the mining of available chemical databases in search for new inhibitors. Two databases have been explored: CSD [39] and Zinc [38]. The search included ~3,000 small aromatic molecules (see "Computational methods"). The screening was performed in a consensus fashion, in which only compounds that were consensually predicted by both models to belong to the same inhibition class were considered. 68 compounds were selected by the consensus model as potential Aβ aggregation inhibitors as a result of the computational screening. Three compounds predicted as inhibitors, were chosen for further in vitro experiments in order to experimentally validate the QSAR models: *N*,*N*′-bis(4-methylphenyl) ethanediamide (ABI-1), 5,12-dihydroquinolino [2,3-b]acridine-7,14-dione (ABI-2), *N*-(3-Indolylacetyl)-*L*-phenylalanine (ABI-3). (Figure S1). The choice was based on high prediction confidence and availability. The fourth compound has been selected for experimental validation based on the screening of an internal database of aromatic compounds. The compound selected was a tryptophan-derived synthetic peptide (ABI-4) and was also predicted as an aggregation inhibitor by both models.

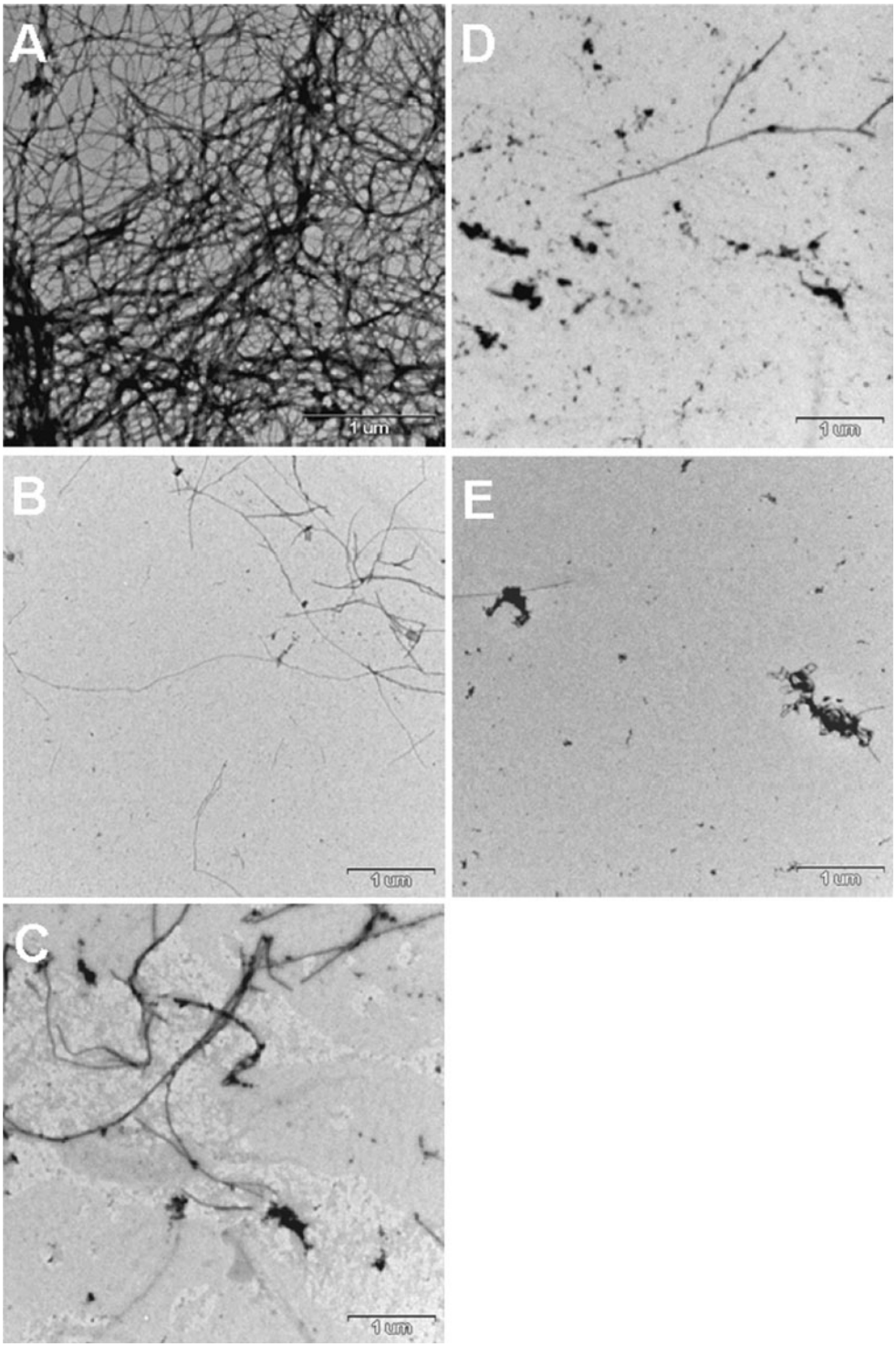

Fig. 5 **a** TEM micrograph of a control sample containing Aβ (1–40), after incubation for 7 days. **b–e** TEM micrographs of samples containing Aβ (1–40) in the presence of the inhibitor molecules tested: ABI-1, ABI-2, ABI-3, ABI-4 (1:10 M ratio). *Scale bar* 1 μm

PCA calculations were carried out in order to determine the relative positioning of the four molecules in the selected descriptor space in order to test whether they reside within the model's applicability domain. All 4 compounds were shown to be in close vicinity to the training and test sets molecules. It could therefore be assumed that the models can reliably predict their activity.

Three indole derivatives were predicted by the consensus model as non-inhibitors: Indole, 3-metyl-indole and indole-3-carboxylic acid. These compounds were tested

experimentally in the past in our laboratory [25]. Results of the in vitro inhibition experiments indeed have shown none or poor inhibitory activity.

Experimental validation

We have investigated the ability of the four chosen small aromatic molecules to inhibit both Aβ (1–40) and Aβ (1–42) aggregation. For this purpose, we used the ThT binding assay. This method is used to measure the growth of fibrils in a solution. The method provides us with a dynamic view of the amyloid formation process. We allowed 5 μM of Aβ to form amyloid fibrils in 10 and 50 μM inhibitor solutions or in the absence of inhibitor solutions as a positive control. ABI-2 was tested in a concentration ratio of 1:5 and not 1:10 due to low solubility in this concentration. ThT was added to aliquots of the solutions after every 24 h, and the fluorescence was measured (see "Experimental section"). An obvious difference was observed between the fibrilization of both Aβ peptides with and without the inhibitors (Fig. 4). Aβ solutions without the inhibitors have revealed a typical sigmoid curve of a nucleation-dependent polymerization and displayed high level of fluorescence in the different experiments. All of the four compounds tested have exhibited good inhibitory ability. In a concentration of tenfold ratio of inhibitors to Aβ (1–40), the inhibition percentage was more than 70% for all molecules tested (Fig. 4e), exhibiting a higher inhibition ability than some of the training set inhibitors. The lag-times of the solutions containing the inhibitors were longer than Aβ peptide itself, ranging from 72 to 120 h. Relative inhibition intensity of all the experiments is shown in Fig. 4.

The quantitative fluorescence assays were followed by qualitative transmission electron microscopy (TEM) analysis. Samples were taken from the Aβ inhibition experiment after 7 days of incubation. Fibrilar structures were observed in the control samples. The specimen containing the Aβ alone contained large broad branching fibrils (Fig. 5), while the specimen with the inhibitors showed shorter fibrils and a significantly decreased amount of fibrils as compared to the fibrils formed by Aβ alone. These results are highly correlated with the ThT assay results.

Conclusions

We have developed computational models that identified a relationship between several structural features and inhibitory activity of known β-amyloid aggregation inhibitors. The hydrophobicity of the inhibitor molecules characterized by the AlogP descriptor, the E-state of carbonyl groups and the hydrogen bonding ability were found important by both models. These results emphasize the high significance of these attributes. Internal and external validations were used and have demonstrated the good performances of the new models and their predictive power. Overall, the models performed very well in the prediction of external validation compounds. To the best of our knowledge, this is the first time that a comprehensive computational QSAR model of Aβ aggregation inhibitors has been reported.

The consensus model successfully predicted the activity of several small molecules that have never been tested for their ability to inhibit amyloid formation. The four molecules that were chosen for the experimental validation showed a very good Aβ aggregation inhibition activity, indicating 100% accuracy of prediction. Thus, the use of 2D descriptors for QSAR modeling was proven to be efficient. It enabled virtual screening in a rapid fashion and the selection of only small number of molecules as potential inhibitors for the experimental validation. Furthermore, consensus QSAR modeling, as used in this study, permits more reliable predictions, as opposed to predictions based on a single model only.

These findings will hopefully lead to additional rational and computational investigations on the chemistry of specific Aβ oligomerization inhibitors, which could help in the development of novel potential drug candidates for AD. Extending the experimental dataset and using a quantitative inhibition data (e.g. IC_{50}), will greatly help in developing robust and reliable models.

Experimental section

Materials

Aβ(1–40) and Aβ(1–42) were purchased from Bachem. ABI-1 was purchased from AKos, ABI-2 was purchased from TCI, ABI-3 was purchased from Sigma–Aldrich, and ABI-4 was synthesized in our laboratory. The purity of the tested compounds was greater than 95% as determined by analytical HPLC (DIONEX ultimate 3000).

Peptides preparation

To avoid preaggregation, synthetic lyophilized Aβ (1–42) was pretreated with hexafluoroisopropanol (HFIP). Aβ (1–42) was dissolved in 100% HFIP, incubated for an hour at 37 °C under shaking at 100 RPM and evaporated in a SpeedVac. Synthetic lyophilized Aβ (1–40) and Aβ (1–42) were further dissolved in dimethylsulfoxide (DMSO) to a concentration of 100 μM and sonicated for 20 s to avoid pre-aggregation. Aβ solutions were prepared by immediate dilution with 10 mM phosphate-buffered saline (PBS) (100 mM NaCl, 0.5 mM EDTA, pH 7.4) to a final concentration of 10 μM.

Inhibitors preparation

The different small molecules were dissolved in DMSO to a concentration of 1 mM and then diluted with 10 mM PBS buffer, to final concentration of 20 and 100 μM.

Thioflavin T binding fluorescence

Aβ solutions were immediately mixed with the inhibitors stock solutions to the final concentration. The samples were incubated at 37 °C and the fibrilogenesis rates were followed by ThT fluorescence assay (excitation at 450 nm, 2.5 nm slit, and emission at 480 nm, 5 nm slit). ThT was added to a ten-fold diluted sample and measured using a Jobin–Yvon Horiba Fluoromax 3 fluorimeter.

Transmission electron microscopy

10 μl samples from the different Aβ inhibition assays were placed on 400 mesh copper grids covered by carbon-stabilized Formvar film (SPI Supplies, West Chester, PA). After 2 min, excess fluid was removed, and the grids were negatively stained with 10 μl of 2% uranyl acetate solution for 2 min. Finally, excess fluid was removed and the samples were viewed in a JEOL 1200EX electron microscope operating at 80 kV.

Supporting information

A complete list of compounds used in the training and test sets along with the bibliographic references, observed inhibition and predicted classification, a list of the descriptors calculated, molecular structures of compounds tested experimentally and other supplementary data indicated in the text.

Acknowledgments We thank Nir Ben-Tal and members of his laboratory for helpful discussions regarding the models and members of the Gazit laboratory for helpful discussions. The authors acknowledge the support of the DIP German-Israel Cooperation Program and the support of MERZ cooperation for this research.

References

1. Cohen AS, Calkins E (1959) Electron microscopic observation on a fibrous component in amyloid of diverse origins. Nature 183: 1202
2. Sunde M, Blake CCF (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. Q Rev Biophys 31:1–39
3. Gazit E (2002) Mechanistic studies of the process of amyloid fibrils formation by the use of peptide fragments and analogues: implications of the design of fibrilization inhibitors. Curr Med Chem 9:1725–1735
4. Gandy S (2005) The role of cerebral amyloid β accumulation in common forms of Alzheimer disease. J Clin Invest 115: 1121–1129
5. Mann D (1989) Cerebral amyloidosis aging and Alzheimer's disease; a contribution from studies on Down's syndrome. Neurobiol Aging 10:397–399
6. Price D, Tanzi R, Borchelt D, Sisodia S (1998) Alzheimer's disease: genetic studies and transgenic models. Annu Rev Genet 32:461–493
7. Van Leuven F (2000) Single and multiple transgenic mice as models for Alzheimer's disease. Prog Neurobiol 61:305–312
8. Pike CJ, Burdick D, Walencewicz AJ, Glabe CG, Cotman CW (1993) Neurodegeneration induced by beta-amyloid peptides in vitro: the role of peptide assembly state. J Neurosci 13: 1676–1687
9. Grace EA, Rabiner CA, Busciglio J (2002) Characterization of neuronal dystrophy induced by fibrilar amyloid beta: implications for Alzheimer's disease. J Neurosci 114:265–273
10. Walsh DM, Townsend M, Podlisny MB, Shankar GM, Fadeeva JV, El Agnaf O, Hartley DM, Selkoe DJ (2005) Certain inhibitors of synthetic amyloid-β peptide (Aβ) fibrillogenesis block oligomerization of natural Aβ and thereby rescue long-term potentiation. J Neurosci 25:2455–2462
11. Ono K, Condron MM, Teplow DB (2009) Structure-neurotoxicity relationships of amyloid {beta}-protein oligomers. Proc Natl Acad Sci USA 106:14745–14750
12. Lesne S, Koh MT, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallagher M, Ashe KH (2006) A specific amyloid beta protein assembly in the brain impairs memory. Nature 16:352–357
13. Van Marum RJ (2008) Current and future therapy in Alzheimer's disease. Fundam Clin Pharmacol 22:265–274
14. Soto C (1999) Plaque busters: strategies to inhibit amyloid formation in Alzheimer's disease. Mol Med Today 5:343–350
15. Rafii MS, Aisen PS (2009) Recent developments in Alzheimer's disease therapeutics. BMC Med 7:7
16. Gilead S, Gazit E (2004) Inhibition of amyloid fibril formation by peptide analogues modified with alpha-aminoisobutyric acid. Angew Chem Int Ed Engl 43:4041–4044
17. Findeis MA (2000) Approaches to discovery and characterization of inhibitors of amyloid β-peptide polymerization. Biochim Biophys Acta 1502:76–84
18. Frydman-Marom A, Rechter M, Shefler I, Bram Y, Shalev DE, Gazit E (2009) Cognitive-performance recovery of Alzheimer's disease model mice by modulation of early soluble amyloidal assemblies. Angew Chem Int Ed 48:1981–1986
19. Ono K, Hasegawa K, Naiki H, Yamada M (2004) Curcumin has potent anti-amyloidogenic effects for Alzheimer's beta-amyloid fibrils in vitro. J Neurosci Res 75:742–750
20. Riviere C, Richard T, Vitrac X, Merillon JM, Vlls J, Monti JP (2008) New polyphenols active on β-amyloid aggregation. Bioorg Med Chem Lett 18:828–831
21. Soto C, Sigurdsson EM, Morelli L, Kumar RA, Castano EM, Frangione B (1998) Beta-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. Nat Med 4:822–826
22. Gazit E (2002) A possible role for π-stacking in self-assembly of amyloid fibrils. FASEB J 16:77–83
23. Tjernberg LO, Näslund J, Lindquist F, Iohansson J, Karlström AR, Thyberg J, Terenius L, Nordstedt C (1996) Arrest of beta-amyloid fibril formation by a pentapeptide ligand. J Biol Chem 271:8545–8548
24. Howlett DR, Perry AE, Godfrey F, Swatton JE, Jennings KH, Spitzfaden C, Wadsworth H, Wood SJ, Markwell RE (1999)

Inhibition of fibril formation in β-amyloid peptide by a novel series of benzofurans. Biochem J 340:283–289
25. Cohen T, Frydman-Marom A, Rechter M, Gazit E (2006) Inhibition of amyloid fibril formation and cytotoxicity by hydroxyindole derivatives. Biochemistry 45:4727–4735
26. Gazit E (2002) Mechanistic studies of the process of amyloid fibrils formation by the use of peptide fragments and analogues: implications of the design of fibrilization inhibitors. Curr Med Chem 9:1725–1735
27. Porat Y, Abramowitz A, Gazit E (2006) Inhibiton of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibiton mechanism. Chem Biol Drug Des 67:7–37
28. Levy M, Porat Y, Bacharach E, Shalev DE, Gazit E (2008) Phenolsulfonphthalein, but not phenolphthalein, inhibits amyloid fibril formation: implications for the modulation of amyloid self-assembly. Biochemistry 47:5896–5904
29. Simons JL, Caprathea BW, Callahana M, Grahama JM, Kimurab T, Laia Y, LeVine H, Lipinskia IW, Sakkaba AT, Tasakib Y, Walkera LC, Yasunagab T, Yea Y, Zhuanga N, Augelli-Szafran CE (2009) The synthesis and structure–activity relationship of substituted N-phenyl anthranilic acid analogs as amyloid aggregation inhibitors. Bioorg Med Chem Lett 19:654–657
30. Reinke AA, Gestwicki JE (2007) Structure–activity relationships of amyloid beta-aggregation inhibitors based on curcumin: influence of linker length and flexibility. Chem Biol Drug Des 70:206–215
31. Discovery Studio, Accelrys, Inc. http://www.accelrys.com/dstudio/
32. http://www.cs.waikato.ac.nz/ml/weka
33. Feller W (1950) An introduction to probability theory and its applications, vol 1. Wiley, New York
34. Xia X, Maliski EG, Gallant P, Rogers D (2004) Classification of kinase inhibitors using a Bayesian Model. J Med Chem 47: 4463–4470
35. Quinlan RJ (1993) C4.5: programs for machine learning. Morgan Kaufmann Publishers, San Mateo
36. Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27:861–874
37. Rucker C, Rucker G, Meringer MY (2007) Randomization and its variants in QSPR/QSAR. J Chem Inf Model 47:2345–2357
38. Irwin JJ, Shoichet BK (2005) ZINC—A free database of commercially available compounds for virtual screening (http://zinc.docking.org). J Chem Inf Model 45:177–182
39. Allen FH (2002) The cambridge structural database: a quarter of a million crystal structures and rising. Acta Cryst B 58:380–388
40. Akaishi T, Morimotoa T, Shibaoa M, Watanabea S, Sakai-Katob K, Utsunomiya-Tateb N, Abe K (2008) Structural requirements for the flavonoid fisetin in inhibiting fibril formation of amyloid β protein. Neurosci Lett 444:280–285
41. Lashuel HA, Hartley DM, Balakhaneh D, Aggarwal A, Teichberg S, Callaway DJ (2002) New class of inhibitors of amyloid-beta fibril formation. Implications for the mechanism of pathogenesis in Alzheimer's disease. J Biol Chem 277:42881–44290
42. Dobson CM (1999) Protein misfolding, evolution and disease. Trends Biochem Sci 24:329–332
43. Riviere C, Richard T, Quentin L, Krisa S, Merillon JM, Mont JP (2007) Inhibitory activity of stilbenes on Alzheimer's b-amyloid fibrils in vitro. Bioorg Med Chem 15:1160–1167
44. Hall LH, Mohney B, Kier LB (1991) The electrotopological state: structure information at the atomic level for molecular graphs. J Chem Inf Comput Sci 31:76–82
45. Hall LH, Kier LB (2000) The E-state as the basis for molecular structure space definition and structure similarity. J Chem Inf Comput Sci 40:784–791
46. Convertino M, Pellarin R, Catto M, Carotti A, Caflish A (2009) 9, 10-anthraquinone hinders β-aggregation: how does a small molecule interfere with Aβ -peptide amyloid fibrillation? Protein Sci 18: 792–800
47. Morshedi D, Rezaei-Ghaleh N, Ebrahim-Habibi A, Ahmadian S, Nemat-Gorgani M (2007) Inhibition of amyloid fibrillation of lysozyme by indole derivatives—possible mechanism of action. FEBS J 274:6415–6425