Identification of Catechins Binding Pockets in Monomeric A β ₄₂ Through Ensemble Docking and MD Simulations

Rohoullah Firouzi^{a*}, Shahin Sowlati-Hashjin^b, Cecilia Chávez-García^(c,d), Mitra Ashouri^e, Mohammad Hossein Karimi-Jafari^f, Mikko Karttunen^{(c,d,g)*}

^(a) Department of Physical Chemistry, Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran. ^(b) Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, M5S 3G9, Canada. ^(c) Department of Chemistry, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 5B7, Canada. ^(d) The Centre of Advanced Materials and Biomaterials Research, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 5B7, Canada. ^(e) Department of Physical Chemistry, School of Chemistry, College of Science, University of Tehran, Tehran, Iran. ^(f) Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran. ^(g) Department of Physics and Astronomy, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 3K7, Canada.

* Corresponding Authors: rfirouzi@ccerci.ac.ir and mkarttu@uwo.ca

Abstract

The assembly of the Amyloid- β peptide (A β) into toxic oligomers and fibrils is associated with Alzheimer's disease and dementia. Therefore, disrupting amyloid assembly by direct targeting of the A β monomeric form with small molecules or antibodies is a promising therapeutic strategy. However, given the dynamic nature of A β , standard computational tools cannot be easily applied for high-throughput structure-based virtual screening in drug discovery projects. In the current study, we propose a computational pipeline – in the framework of the ensemble docking strategy – to identify catechins' binding pockets in monomeric A β ₄₂. It is shown that both hydrophobic aromatic interactions and hydrogen bonding are crucial for the binding of catechins to A β ₄₂. Also, it has been found that all the studied ligands, especially the EGCG, can act as potent inhibitors against amyloid aggregation by blocking the central hydrophobic region of the A β . Our findings are evaluated and confirmed with multi-microsecond MD simulations. Finally, it is suggested that our proposed pipeline, with low computational cost in comparison with MD simulations, is a suitable approach for the virtual screening of ligand libraries against A β .

Keywords: Catechins, Amyloid- β , Binding Pockets, Ensemble Docking, MD Simulations.

1. Introduction

Intrinsically disordered proteins (IDPs) are very flexible biomolecules without a well-defined folded structure and typically have important roles in biological processes, in particular in cellular signaling and gene regulation.¹⁻⁴ Under certain conditions, some IDPs may aggregate into highly toxic oligomers. These oligomers are associated with a wide range of serious human diseases such as cancer, neurodegenerative diseases, autoimmune disorders, cardiovascular disease, and type II diabetes.⁴⁻⁹ Thus, preventing or reducing aggregation of the IDPs involved in such diseases appears as an effective therapeutic strategy.

In recent years, there have been efforts to design and synthesize small molecules and short peptides to block IDPs aggregation at different stages along the aggregation pathway, in particular nucleation and oligomer formation.^{4,10-17} Several studies have revealed that a large number of natural compounds derived from plants, animals and microorganisms, have the potential to inhibit oligomerization.^{11-12, 18-20} For example, several computational and experimental observations have shown that polyphenolic plant compounds which occur naturally in fruit, vegetables, chocolate, and tea are capable of inhibiting IDP aggregation.^{12-13, 20-34}

Finding aggregation inhibitors and direct targeting of monomeric IDPs via small molecules is a very active area of research and a wide variety of computational techniques have been applied, yet there are many inherent difficulties.^{12-14, 27, 35-41} For example, most docking algorithms employ the flexible ligand and rigid receptor paradigm⁴²⁻⁴⁴ but IDPs display high conformational heterogeneity, and ligand binding causes large structural changes in the IDP conformations. To circumvent this problem, heterogeneous conformational ensembles of IDPs have been used for docking studies.^{27, 38, 45-46} Nevertheless, challenges associated with generating and choosing a set of suitable conformations for docking still remain. It will also be interesting to see how and if

Alphafold⁴⁷ will change the situation as it has already been applied to non-IDP related binding problems,⁴⁸ but IDPs appear to be a challenge even for Alphafold.⁴⁹

Several methods have been proposed for efficient sampling of IDP's conformational space and constructing a representative conformational ensemble. Examples include replica-exchange-⁵⁰⁻⁵³ and metadynamics-based methods,⁵⁴⁻⁵⁶ diffusion map approaches⁵⁷⁻⁵⁸ and Markov state modeling.⁵⁹⁻⁶² For comparisons between the methods, see for example Refs. ³⁹ and ⁶³⁻⁶⁵. One alternative for effective exploration of the conformational space is the use of multiple conventional MD trajectories (replicas) with different initial conditions (different velocities or/and different starting configurations). The strategy of choosing the initial conditions controls the effectiveness of this approach and its ability to enhance conformational sampling performance.⁶⁶⁻⁷⁰ We have recently proposed a new efficient algorithm for comprehensive exploring of the conformational space of IDPs, called Blockwise Excursion Sampling (BES).³⁹ It uses simulated annealing (SA) to find different low energy states of various regions of conformational space as optimal starting configurations for short conventional MD simulations. In BES, conformational sampling is based on many uncorrelated short MD simulations starting from different points of the accessible phase space. It has been shown that the protocol is successful in generating a diverse conformational library for IDP conformations in agreement with experimental data.⁷⁰

In this work, we applied the BES protocol to generate a reliable structural library for the full- length human Amyloid- β ($A\beta_{42}$) monomer involved in Alzheimer's disease.^{2,5,11-12,27,35,37,39} Through ensemble docking approach, catechins, that are important natural polyphenolic compounds, were docked onto the surfaces to identify the binding "hot spots" on the $A\beta_{42}$ monomers. It is noteworthy to point out that possible effects of catechins on $A\beta$ aggregation have been the subject of numerous experimental and computational investigations. Those studies have

proposed that catechins are able to disturb the $A\beta$ aggregation and induce its aggregation into nontoxic forms oligomers.^{21, 28, 71-80} In order to further evaluate the docking results and the stability of complexes, the obtained structures with the largest binding energies for each $A\beta_{42}$ -catechin complex were used as the starting structures for long multi-microsecond MD simulations (total of 15 μ s).

2. Methods

2.1. Protein and Ligand Library Preparation

A library of $A\beta_{42}$ (see **Figure 1** for a snapshot and the amino acid sequence) structures was generated using the BES protocol. Briefly, 2,000 excursion chains were performed such that all excursion chains were started from a fully extended structure; excursion chain refers to a sequence of MD and SA blocks in the BES protocol, for details and terminology please see Ref.⁷⁰.

Each excursion chain included five successive SA and MD blocks with maximum temperatures of 700, 600, 500, 400, and 350 K for the SA blocks. The relaxation time for each MD block was set to 120 ps and the last 100 ps were used to generate representative structures. In the next step, an average (mean) structure over the MD trajectory was obtained, and the root mean deviation (RMSD) was used as a criterion to identify the configuration in the MD trajectory that is structurally closest to the average structure. The selected structure was then energy minimized using the conjugate gradient method and used as a representative structure. As a result, for each MD block (five blocks in each excursion chain) one representative structure was derived. The final structural library included a total of 10,000 representative structures. **Scheme 1** summarizes the procedure. For more technical details and a complete description of the BES protocol see Ref.⁷⁰.

Catechins (or Flavan-3-ols) are dietary polyphenolic compounds that many experimental results have suggested as potentially useful in targeting $A\beta$.^{21, 28, 71-87} For this study, we selected

five different types of catechins: 1) (+)-catechin (*C*), 2) (-)-epicatechin (*EC*), 3) (-)-epigallocatechin (*EGC*), 4) (-)-epicatechin-3-gallate (*ECG*), and 5) (-)-epigallocatechin-3-gallate (*EGCG*). Their chemical structures are shown in **Figure 2** and their initial structures were taken from the ZINC database.⁸⁸ This was followed by optimization of the molecular geometries for all of them using the B3LYP exchange and correlation functional⁸⁹ and the 6-31+G(d,p) basis set. The GAMESS package⁹⁰ was used for geometry optimization.

2.2. Docking Setup

Docking simulations were performed with AutoDock Vina (version 1.1.2) software.⁹¹ The docking search space for exploring ligand binding conformations around each representative $A\beta_{42}$ structure was defined using a rectangular box centered at the center of mass of $A\beta_{42}$ with a minimal distance of 12 Å from the $A\beta_{42}$ to the edges of the box. Therefore, depending on size and shape of the $A\beta_{42}$ configuration, an optimized docking box was determined individually for each $A\beta_{42}$. Each docking run generates nine optimal $A\beta_{42}$ -ligand bound conformations and overall, a total of 90,000 (10,000 $\times$ 9) poses were generated for each catechin compound. The different poses in each run are rank-ordered by the Vina score, a quantity that correlates with the binding free energy. The top-scoring pose in each run achieves the lowest free energy of binding in the complex.

Very recently, it has been shown that the correct pose (the pose with the lowest RMSD from the corresponding experimental pose) is usually predicted by Vina but sometimes, does not get the top score in the Vina ranking.⁹²⁻⁹³ To avoid the problem and to capture the correct poses, it is recommended that except for the top-ranked pose, some important lower ranked poses for each docking run should be identified and selected for post-docking analysis. For more discussions about ranking, see also Refs. ⁹⁴⁻⁹⁷. For this purpose, the differences in the binding free energies

between the top-ranked pose and lower ranked poses were calculated for selected high-ranked modelled complexes for each docking run, $\Delta\Delta G_{binding}$ ($= \Delta G_{top-pose} - \Delta G_{lower ranked pose}$). Different cutoff values for the $\Delta\Delta G_{binding}$ threshold (0.1, 0.2 and 0.3 kcal/mol) were used for selecting docking complexes, since the optimal selection of complexes is not known. For larger $\Delta\Delta G_{binding}$ cutoff values, more complexes were selected. For example, with the cutoff value of 0.1 kcal/mol, 17,431 ($10,000 \times 1$ (top-ranked poses from each run) + 7,431 (lower ranked poses near top-poses with the cutoff)) complexes were selected for docking of *EGCG*, while 29,923 and 42,833 complexes were selected with the cutoff values of 0.2 and 0.3 kcal/mol, respectively. The number of selected docking complexes for all ligands and the corresponding cutoff values of $\Delta\Delta G_{binding}$ are provided and labelled as “set-*n*” in **Table 1**. Since the results for different sets are very similar, we only present the results of set-1 (with the cutoff value of 0.1) and the results for the other two sets can be found in Supporting Information.

2.3. MD simulation setup

All-atom MD simulations were performed on the peptide with five different ligands. The force field parameters for each ligand were created using the Antechamber program in the Ambertools19 package⁹⁸ and described by the General Amber Force Field (GAFF)⁹⁹ using AM1-BCC charges.¹⁰⁰ The Amberff99SB*-ILDNP force field¹⁰¹ and the TIP3P water model¹⁰² were adopted for the protein and water, respectively. Each protein was placed in a dodecahedral box such that the distance from the edges of the box to every atom in the protein was at least 1 nm and 150 mM of KCl was added to reproduce physiological conditions. Overall charge neutrality was preserved by adding 3 K⁺ counterions. The GROMACS 2016.3¹⁰³ package was used for all simulations. Each system was energy minimized using the method of steepest descents. This was followed by a pre-

equilibration in the canonical ensemble, i.e., at constant particle number, volume, and temperature, for 100 ps. The Lennard-Jones potential was truncated using a shift function between 1.0 and 1.2 nm. Electrostatic interactions were calculated using the particle-mesh Ewald method (PME)¹⁰⁴⁻¹⁰⁵ with a real space cut-off of 1.2 nm. The temperature was set to 310 K with the V-rescale algorithm¹⁰⁶ and pressure was kept at 1 atm using the Parrinello-Rahman barostat.¹⁰⁷ Bonds involving hydrogens were constrained using the Linear Constraint Solver (P-LINCS) algorithm.¹⁰⁸ Pre-equilibration was followed by a production run of 3 μ s with a time step of 2 fs for each of the five peptide-ligand systems.

3. Results and Discussion

3.1. Docking Analysis

Structural Analysis and Identifying Ligand Binding Site

Determination of the important residues that are in close contact with the ligand is very important for the identification of potential inhibitors of A β ₄₂ aggregation. The distances between the heavy atoms of the ligand and the residues of A β ₄₂ were used to define the binding sites of A β ₄₂. Based on this assumption, a list of binding residues for each selected complex was generated for which at least one heavy atom of the residues falls within the cutoff distance of any ligand heavy atom. To assess the effect of the cutoff distance, five distances were evaluated: 3.0, 3.5, 4.0, 4.5, and 5.0 Å; cutoff distances from 3.0 to 5.0 Å are commonly used to study ligand-protein binding interactions, such as hydrogen bonds, hydrophobic contacts, and aromatic interactions.^{22, 109-114}

The number of contacts between each ligand and A β ₄₂ residues for the set-1 (with $\Delta G_{binding}$ = 0.1 kcal/mol) was counted, see **Tables 2-6**. The first observation from the tables is that all the ligands have the most contacts with residues Tyr10, Phe19, and Phe20. Moreover, based on the

ranking (**Tables 1-6**), other aromatic residues of (His13, His14, His6, and Phe4) also contribute to stabilizing the interactions with catechins: These polyphenolic compounds tend to interact with the aromatic residues through stacking and/or T-shaped interactions, see the snapshots in **Figures 3-7**. Another important observation is that the tendency of ligands to interact with Tyr10 correlates with the number of hydroxyl groups on the ligands. This is seen in the data in **Tables 2-6**: *EGCG* possesses the largest number of hydroxyl groups (8 OH's) and has a greater tendency to interact with Tyr10 than with Phe19 or Phe20, i.e., the number of contacts between *EGCG* and Tyr10 is larger than those between *EGCG* and Phe19 or Phe20. Thus, the data appears to imply that, hydrophobic aromatic interactions and hydrogen bonding are both crucial for the binding process. Finally, a comparative look at the tabulated values immediately shows that all ligands have a tendency to associate with the hydrophobic region of A β ₄₂ spanning residues from Tyr10 to Phe20. This region contains most of the aromatic residues found in full-length A β . This region also encompasses the central hydrophobic region (¹⁶KL VFF²⁰) that, based on many experimental and computational studies, is involved in the initiation of amyloid aggregation.^{11-13, 22, 36, 115-119} Therefore, our docking results show that all studied ligands, especially the *EGCG*, can act as potent inhibitors against amyloid aggregation through blocking the central hydrophobic region. These findings are in agreement with experimental studies.^{28, 72-74}

The distributions of binding affinities (kcal/mol) of all ligands in set-1 (cutoff of 0.1 kcal/mol) are shown in **Figures 3-7**. In addition, the figures show a couple of complexes with the highest binding affinities and for each complex the different types of aromatic interactions, such as π - π , XH- π (X = C, N, O), and lone pair- π interactions between the aromatic rings of the ligands, and the sidechain of aromatic residues have been highlighted by colored dashed-lines for clarity. In all cases, at least one aromatic residue was in contact with the ligand and in most cases, catechin

compounds possessing multiple aromatic rings were capable of interacting with several aromatic residues simultaneously. Therefore, it seems that these aromatic interactions play an important role in binding to aggregation-prone regions of $A\beta_{42}$ and are essential for high affinity and binding specificity.

3.2. Analysis of MD Simulations

In order to evaluate the docking results, microsecond-scale MD simulations were performed on the $A\beta_{42}$ -L (L = C, EC, ECG, EGC, EGCG) structures for which the largest binding energies were obtained by the docking procedure. The initial and final MD structures are provided in **Figure S1** which show that the ligands maintain their interactions with $A\beta_{42}$ throughout the simulation. Root mean square deviation (RMSD) of the $A\beta_{42}$ indicates that in all cases the system has reached steady state (**Table S1** and **Figure S2**). The stability of the $A\beta_{42}$ -L complexes is also reflected in the relatively low radius of gyration (~ 1.0 nm) of the $A\beta_{42}$ (**Table S1** and **Figure S3**). Moreover, the steady solvent accessible surface area (SASA; with $< 7\%$ fluctuations) observed for the chains provides further support for the stability and compactness of the system (**Table S1** and **Figure S4**).

The docking results suggest that Tyr10, Phe19, and to some extent Phe20, are the main residues involved in ligand binding. Average distances between these residues and the five ligands from the MD simulations are provided in **Table 7**. In agreement with the results from our docking procedure, with the exception of C for which the distance is ~ 1.0 nm, the average distance between the ligands and all three residues is ~ 0.5 nm (**Table 7** and **Figures S5 – S9**). The time evolution of these distances and the end-to-end distance for $A\beta_{42}$ are shown in **Figures S5 – S9**.

Inspection of hydrogen bonds between the ligands and $A\beta_{42}$ reveals that hydrogen bonding plays a more important role in binding of *EGCG* and *EGC* than with the other three ligands (**Table 8**). Binding of *C* seems to have the least dependence on the hydrogen bonding among the ligands and relies on the π - π , XH - π ($X = C, O$) interactions. Average number of $A\beta_{42}$...ligand, $A\beta_{42}$...solvent, and ligand...solvent hydrogen bonds are collected in **Table 8**.

Figures S10 – S14 depict the contributions of $A\beta_{42}$ residues that involve hydrogen bonding with the ligands. A comparison among the five $A\beta_{42}$...ligand systems reveals that there are a few residues in the $A\beta_{42}$ sequence that frequently form hydrogen bonds with the five different catechins (e.g., Glu22, Asp23, Ala42, etc.). These intransient hydrogen bonds alternate between the ligand and the residues, which may suggest that hydrogen bonds have a less important role in the binding of the ligands to $A\beta_{42}$ compared to stacking interactions. The average number of hydrogen bonds between all the amino acids of the peptide and ligands are depicted in **Figure S15**. Similar conclusion has been previously made for $A\beta_{42}$ protofibrils and *EGCG*, where *EGCG* was shown to bind to $A\beta_{42}$ monomer through hydrophobic, π - π stacking, and hydrogen bonds.¹²⁰ Moreover, same study identifies Asp1, Glu22, and Ala42 residues to form the most hydrogen bonds with *EGCG*, which except for Asp1, agrees with our observation (**Figure S14**). It should be noted, however, that while we have examined disordered structures in this study, Li *et al.*¹²⁰ considered fibrils and 25 *EGCG* molecules.

4. Conclusions

In this work, binding of various well-known catechins present in green tea to the amyloid- β peptide ($A\beta$) has been predicted and analyzed. For this purpose, a computational pipeline in the framework of the ensemble docking strategy has been proposed in which a structurally heterogeneous

ensemble of conformations of $A\beta_{42}$ is used. The ensemble is generated by the Blockwise Excursion Sampling (BES) protocol⁷⁰ in which the conformational sampling is performed on the basis of many uncorrelated short-time MD simulations starting from different reasonable points of the accessible phase space.

It was observed that all green tea catechins compounds tended to interact with the aromatic residues through stacking and/or T-shaped interactions and, because of this, all compounds show a high tendency to interact with the hydrophobic region of $A\beta_{42}$ spanning residues from Tyr10 to Phe20, the region with the highest number of the aromatic residues in full-length $A\beta_{42}$. This region also encompasses the central hydrophobic core (CHC, residues 16 – 20) that, based on many experimental and computational studies, plays a key role in the aggregation process of $A\beta_{42}$. Therefore, the docking results indicate that all studied ligands, especially the *EGCG*, can act as potent inhibitors against amyloid aggregation by blocking the central hydrophobic core. Additionally, it has been suggested that both hydrophobic aromatic interactions and hydrogen bonding are crucial for the binding of catechins to $A\beta_{42}$.

To evaluate the obtained findings in binding of catechin compounds to $A\beta_{42}$, long multi-microsecond MD simulations were performed. It was shown that the present docking protocol is highly successful in identifying catechins' binding pockets in monomeric $A\beta_{42}$, in agreement with previous MD simulations and some recent experimental observations for similar $A\beta_{42}$ -catechin complexes.^{28, 71-73, 77-79} Finally, we suggested that our proposed pipeline with low computational cost in comparison with MD simulations is a suitable approach for high-throughput structure-based virtual screening of ligand libraries against the intrinsically disordered proteins (IDPs), such as the $A\beta$.

Conflicts of Interest Statement

There are no conflicts of interest to declare.

Data and Software Availability

GAMESS package and AutoDock Vina (version 1.1.2) were used under a free academic license for ligands preparation and docking simulations. Produced and analyzed data are available upon request. MD simulations were performed using GROMACS 2016.3.

Supporting Information

RMSD, R_g , and SASA of $A\beta_{42}$ in the different systems: Table S1 (standard deviations in parentheses). Initial and final structures of $A\beta_{42}$ in the presence of ligands: Figure S1. RMSD, radius of gyration and solvent accessible surface area of the $A\beta_{42}$ backbone in the presence of different ligands, Figures S2, S3 and S4, respectively. Figures S5-S9 show the average $A\beta_{42}$ end-to-end distance and distances between residues Y10, F19 and F20, and C, EC, ECG, EGC and EGCG, respectively. Figures S10-S14: Hydrogen bonds between selected residues of $A\beta_{42}$ and C, EC, ECG, EGC and EGCG, respectively. Figure S15: Average number of hydrogen bonds between $A\beta_{42}$ amino acids and the ligands.

Author Contributions

All authors contributed to the design, implementation and analysis of the research, interpretation of the results, and writing of the manuscript. All authors have given their approval for the final version of the manuscript.

Acknowledgments

C.C.G is recipient of an Ontario Trillium Scholarship (OTS). M.K. thanks NSERC Canada Research Chairs Program the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support. Computational resources were provided by SharcNet and Compute Canada.

References

1. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **2005**, *6*, 197-208.
2. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M., Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews* **2014**, *114*, 6589-6631.
3. Davey, N. E., The functional importance of structure in unstructured protein regions. *Current Opinion in Structural Biology* **2019**, *56*, 155-163.
4. Dobson, C. M.; Knowles, T. P. J.; Vendruscolo, M., The Amyloid Phenomenon and Its Significance in Biology and Medicine. *Cold Spring Harbor Perspectives in Biology* **2020**, *12*, a033878.
5. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annual Review of Biophysics* **2008**, *37*, 215-246.
6. Uversky, V. N., Chapter Four - Intrinsic Disorder, Protein-Protein Interactions, and Disease. In *Advances in Protein Chemistry and Structural Biology*, Donev, R., Ed. Academic Press: 2018; Vol. 110, pp 85-121.
7. Owen, M. C.; Gnutt, D.; Gao, M.; Wärmländer, S. K. T. S.; Jarvet, J.; Gräslund, A.; Winter, R.; Ebbinghaus, S.; Strodel, B., Effects of in vivo conditions on amyloid aggregation. *Chemical Society Reviews* **2019**, *48*, 3946-3996.
8. La Rosa, C.; Condorelli, M.; Compagnini, G.; Lolicato, F.; Milardi, D.; Do, T. N.; Karttunen, M.; Pannuzzo, M.; Ramamoorthy, A.; Fraternali, F.; Collu, F.; Rezaei, H.; Strodel, B.; Raudino, A., Symmetry-breaking transitions in the early steps of protein self-assembly. *European Biophysics Journal* **2020**, *49*, 175-191.
9. Strodel, B., Energy Landscapes of Protein Aggregation and Conformation Switching in Intrinsically Disordered Proteins. *Journal of Molecular Biology* **2021**, *433*, 167182.
10. Metallo, S. J., Intrinsically disordered proteins are potential drug targets. *Current Opinion in Chemical Biology* **2010**, *14*, 481-488.
11. Nie, Q.; Du, X.-g.; Geng, M.-y., Small molecule inhibitors of amyloid β peptide aggregation as a potential therapeutic strategy for Alzheimer's disease. *Acta Pharmacologica Sinica* **2011**, *32*, 545-551.
12. Lemkul, J. A.; Bevan, D. R., The Role of Molecular Simulations in the Development of Inhibitors of Amyloid β -Peptide Aggregation for the Treatment of Alzheimer's Disease. *ACS Chemical Neuroscience* **2012**, *3*, 845-856.
13. Härd, T.; Lendel, C., Inhibition of Amyloid Formation. *Journal of Molecular Biology* **2012**, *421*, 441-465.
14. Heller, G. T.; Sormanni, P.; Vendruscolo, M., Targeting disordered proteins with small molecules using entropy. *Trends in Biochemical Sciences* **2015**, *40*, 491-496.
15. Heller, G. T.; Aprile, F. A.; Vendruscolo, M., Methods of probing the interactions between small molecules and disordered proteins. *Cellular and Molecular Life Sciences* **2017**, *74*, 3225-3243.
16. Ruan, H.; Sun, Q.; Zhang, W.; Liu, Y.; Lai, L., Targeting intrinsically disordered proteins at the edge of chaos. *Drug Discovery Today* **2019**, *24*, 217-227.
17. Firouzi, R.; Noohi, B., Identification of key stabilizing interactions of amyloid- β oligomers based on fragment molecular orbital calculations on macrocyclic β -hairpin peptides. *Proteins: Structure, Function, and Bioinformatics* **2022**, *90*, 229-238.
18. Tang, C.; Ye, Y.; Feng, Y.; Quinn, R. J., TCM, brain function and drug space. *Natural Product Reports* **2016**, *33*, 6-25.
19. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G., Counting on natural products for drug design. *Nature Chemistry* **2016**, *8*, 531-541.
20. Habtemariam, S., Natural Products in Alzheimer's Disease Therapy: Would Old Therapeutic Approaches Fix the Broken Promise of Modern Medicines? *Molecules* **2019**, *24*, 1519.

21. Porat, Y.; Abramowitz, A.; Gazit, E., Inhibition of Amyloid Fibril Formation by Polyphenols: Structural Similarity and Aromatic Interactions as a Common Inhibition Mechanism. *Chemical Biology & Drug Design* **2006**, *67*, 27-37.
22. Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E., EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nature Structural & Molecular Biology* **2008**, *15*, 558-566.
23. Jiang, L.; Liu, C.; Leibly, D.; Landau, M.; Zhao, M.; Hughes, M. P.; Eisenberg, D. S., Structure-based discovery of fiber-binding compounds that reduce the cytotoxicity of amyloid beta. *eLife* **2013**, *2*, e00857.
24. Stefani, M.; Rigacci, S., Protein Folding and Aggregation into Amyloid: The Interference by Natural Phenolic Compounds. *International Journal of Molecular Sciences* **2013**, *14*, 12411.
25. Zheng, Q.; Kebede, M. T.; Kemeh, M. M.; Islam, S.; Lee, B.; Bleck, S. D.; Wurfl, L. A.; Lazo, N. D., Inhibition of the Self-Assembly of A β and of Tau by Polyphenols: Mechanistic Studies. *Molecules* **2019**, *24*, 2316.
26. Murakami, K.; Irie, K., Three Structural Features of Functional Food Components and Herbal Medicine with Amyloid β 42 Anti-Aggregation Properties. *Molecules* **2019**, *24*, 2125.
27. Zhu, M.; De Simone, A.; Schenk, D.; Toth, G.; Dobson, C. M.; Vendruscolo, M., Identification of small-molecule binding pockets in the soluble monomeric form of the A β 42 peptide. *The Journal of Chemical Physics* **2013**, *139*, 035101.
28. Fusco, G.; Sanz-Hernandez, M.; Ruggeri, F. S.; Vendruscolo, M.; Dobson, C. M.; De Simone, A., Molecular determinants of the interaction of EGCG with ordered and disordered proteins. *Biopolymers* **2018**, *109*, e23117.
29. Jia, L.; Zhao, W.; Sang, J.; Wang, W.; Wei, W.; Wang, Y.; Zhao, F.; Lu, F.; Liu, F., Inhibitory Effect of a Flavonoid Dihydromyricetin against A β 40 Amyloidogenesis and Its Associated Cytotoxicity. *ACS Chemical Neuroscience* **2019**, *10*, 4696-4703.
30. Tomaselli, S.; La Vitola, P.; Pagano, K.; Brandi, E.; Santamaria, G.; Galante, D.; D'Arrigo, C.; Moni, L.; Lambruschini, C.; Banfi, L.; Lucchetti, J.; Fracasso, C.; Molinari, H.; Forloni, G.; Balducci, C.; Ragona, L., Biophysical and in Vivo Studies Identify a New Natural-Based Polyphenol, Counteracting A β Oligomerization in Vitro and A β Oligomer-Mediated Memory Impairment and Neuroinflammation in an Acute Mouse Model of Alzheimer's Disease. *ACS Chemical Neuroscience* **2019**, *10*, 4462-4475.
31. Windsor, P. K.; Plassmeyer, S. P.; Mattock, D. S.; Bradfield, J. C.; Choi, E. Y.; Miller, B. R.; Han, B. H., Biflavonoid-Induced Disruption of Hydrogen Bonds Leads to Amyloid- β Disaggregation. *International Journal of Molecular Sciences* **2021**, *22*, 2888.
32. Henríquez, G.; Gomez, A.; Guerrero, E.; Narayan, M., Potential Role of Natural Polyphenols against Protein Aggregation Toxicity: In Vitro, In Vivo, and Clinical Studies. *ACS Chemical Neuroscience* **2020**, *11*, 2915-2934.
33. Nam, G.; Hong, M.; Lee, J.; Lee, H. J.; Ji, Y.; Kang, J.; Baik, M.-H.; Lim, M. H., Multiple reactivities of flavonoids towards pathological elements in Alzheimer's disease: structure–activity relationship. *Chemical Science* **2020**, *11*, 10243-10254.
34. Pandey, S. N.; Rangra, N. K.; Singh, S.; Arora, S.; Gupta, V., Evolving Role of Natural Products from Traditional Medicinal Herbs in the Treatment of Alzheimer's Disease. *ACS Chemical Neuroscience* **2021**, *12*, 2718-2728.
35. Rezaei-Ghaleh, N.; Blackledge, M.; Zweckstetter, M., Intrinsically Disordered Proteins: From Sequence and Conformational Properties toward Drug Discovery. *ChemBioChem* **2012**, *13*, 930-950.
36. Attanasio, F.; Convertino, M.; Magno, A.; Caffisch, A.; Corazza, A.; Haridas, H.; Esposito, G.; Cataldo, S.; Pignataro, B.; Milardi, D.; Rizzarelli, E., Carnosine Inhibits A β 42 Aggregation by Perturbing the H-Bond Network in and around the Central Hydrophobic Cluster. *ChemBioChem* **2013**, *14*, 583-592.
37. Nguyen, P.; Derreumaux, P., Understanding Amyloid Fibril Nucleation and A β Oligomer/Drug Interactions from Computer Simulations. *Accounts of Chemical Research* **2014**, *47*, 603-611.
38. Heller, G. T.; Bonomi, M.; Vendruscolo, M., Structural Ensemble Modulation upon Small-Molecule Binding to Disordered Proteins. *Journal of Molecular Biology* **2018**, *430*, 2288-2292.

39. Bhattacharya, S.; Xu, L.; Thompson, D., Revisiting the earliest signatures of amyloidogenesis: Roadmaps emerging from computational modeling and experiment. *WIREs Computational Molecular Science* **2018**, *8*, e1359.
40. Nguyen, P. H.; Ramamoorthy, A.; Sahoo, B. R.; Zheng, J.; Faller, P.; Straub, J. E.; Dominguez, L.; Shea, J.-E.; Dokholyan, N. V.; De Simone, A.; Ma, B.; Nussinov, R.; Najafi, S.; Ngo, S. T.; Loquet, A.; Chiricotto, M.; Ganguly, P.; McCarty, J.; Li, M. S.; Hall, C.; Wang, Y.; Miller, Y.; Melchionna, S.; Habenstein, B.; Timr, S.; Chen, J.; Hnath, B.; Strodel, B.; Kayed, R.; Lesné, S.; Wei, G.; Sterpone, F.; Doig, A. J.; Derreumaux, P., Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chemical Reviews* **2021**, *121*, 2545-2647.
41. Boopathi, S.; Poma, A. B.; Garduño-Juárez, R., An Overview of Several Inhibitors for Alzheimer's Disease: Characterization and Failure. *International Journal of Molecular Sciences* **2021**, *22*, 10798.
42. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M., Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer-Aided Drug Design* **2011**, *7*, 146-157.
43. 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. *Journal of Chemical Information and Modeling* **2010**, *50*, 511-524.
44. Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J., Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols* **2016**, *11*, 905-919.
45. Limongelli, V.; Marinelli, L.; Cosconati, S.; Braun, H. A.; Schmidt, B.; Novellino, E., Ensemble-Docking Approach on BACE-1: Pharmacophore Perception and Guidelines for Drug Design. *ChemMedChem* **2007**, *2*, 667-678.
46. Tóth, G.; Gardai, S. J.; Zago, W.; Bertoncini, C. W.; Cremades, N.; Roy, S. L.; Tambe, M. A.; Rochet, J.-C.; Galvagnion, C.; Skibinski, G.; Finkbeiner, S.; Bova, M.; Regnstrom, K.; Chiou, S.-S.; Johnston, J.; Callaway, K.; Anderson, J. P.; Jobling, M. F.; Buell, A. K.; Yednock, T. A.; Knowles, T. P. J.; Vendruscolo, M.; Christodoulou, J.; Dobson, C. M.; Schenk, D.; McConlogue, L., Targeting the Intrinsically Disordered Structural Ensemble of α -Synuclein by Small Molecules as a Potential Therapeutic Strategy for Parkinson's Disease. *PLOS ONE* **2014**, *9*, e87133.
47. Ruff, K. M.; Pappu, R. V., AlphaFold and Implications for Intrinsically Disordered Proteins. *Journal of Molecular Biology* **2021**, *433*, 167208.
48. Yuan, Q.; Chen, S.; Rao, J.; Zheng, S.; Zhao, H.; Yang, Y., AlphaFold2-aware protein-DNA binding site prediction using graph transformer. *bioRxiv* **2021**, 2021.08.25.457661.
49. Wilson, C. J.; Choy, W.-Y.; Karttunen, M., AlphaFold2: A role for disordered protein prediction? *bioRxiv* **2021**, 2021.09.27.461910.
50. Ball, K. A.; Phillips, Aaron H.; Wemmer, David E.; Head-Gordon, T., Differences in β -strand Populations of Monomeric A β 40 and A β 42. *Biophysical Journal* **2013**, *104*, 2714-2724.
51. Rosenman, D. J.; Wang, C.; García, A. E., Characterization of A β Monomers through the Convergence of Ensemble Properties among Simulations with Multiple Force Fields. *The Journal of Physical Chemistry B* **2016**, *120*, 259-277.
52. Ziegler, Z.; Schmidt, M.; Gurry, T.; Burger, V.; Stultz, C. M., Mollack: a web server for the automated creation of conformational ensembles for intrinsically disordered proteins. *Bioinformatics* **2016**, *32*, 2545-2547.
53. Ball, K. A.; Phillips, A. H.; Nerenberg, P. S.; Fawzi, N. L.; Wemmer, D. E.; Head-Gordon, T., Homogeneous and Heterogeneous Tertiary Structure Ensembles of Amyloid- β Peptides. *Biochemistry* **2011**, *50*, 7612-7628.
54. Do, T. N.; Choy, W.-Y.; Karttunen, M., Accelerating the Conformational Sampling of Intrinsically Disordered Proteins. *Journal of Chemical Theory and Computation* **2014**, *10*, 5081-5094.
55. Camilloni, C.; Vendruscolo, M., Statistical Mechanics of the Denatured State of a Protein Using Replica-Averaged Metadynamics. *Journal of the American Chemical Society* **2014**, *136*, 8982-8991.

56. Zerze, G. H.; Miller, C. M.; Granata, D.; Mittal, J., Free Energy Surface of an Intrinsically Disordered Protein: Comparison between Temperature Replica Exchange Molecular Dynamics and Bias-Exchange Metadynamics. *Journal of Chemical Theory and Computation* **2015**, *11*, 2776-2782.
57. Zheng, W.; Rohrdanz, M. A.; Clementi, C., Rapid Exploration of Configuration Space with Diffusion-Map-Directed Molecular Dynamics. *The Journal of Physical Chemistry B* **2013**, *117*, 12769-12776.
58. Fujisaki, H.; Moritsugu, K.; Mitsutake, A.; Suetani, H., Conformational change of a biomolecule studied by the weighted ensemble method: Use of the diffusion map method to extract reaction coordinates. *The Journal of Chemical Physics* **2018**, *149*, 134112.
59. Noé, F.; Clementi, C., Kinetic Distance and Kinetic Maps from Molecular Dynamics Simulation. *Journal of Chemical Theory and Computation* **2015**, *11*, 5002-5011.
60. Paul, A.; Samantray, S.; Anteghini, M.; Khaled, M.; Strodel, B., Thermodynamics and kinetics of the amyloid- β peptide revealed by Markov state models based on MD data in agreement with experiment. *Chemical Science* **2021**, *12*, 6652-6669.
61. Cukier, R. I., Generating Intrinsically Disordered Protein Conformational Ensembles from a Database of Ramachandran Space Pair Residue Probabilities Using a Markov Chain. *The Journal of Physical Chemistry B* **2018**, *122*, 9087-9101.
62. Lin, Y.-S.; Bowman, Gregory R.; Beauchamp, Kyle A.; Pande, Vijay S., Investigating How Peptide Length and a Pathogenic Mutation Modify the Structural Ensemble of Amyloid Beta Monomer. *Biophysical Journal* **2012**, *102*, 315-324.
63. Barducci, A.; Pfandtner, J.; Bonomi, M., Tackling sampling challenges in biomolecular simulations. *Methods Mol Biol* **2015**, *1215*, 151-71.
64. Ilie, I. M.; Caflisch, A., Simulation Studies of Amyloidogenic Polypeptides and Their Aggregates. *Chemical Reviews* **2019**, *119*, 6956-6993.
65. Nasica-Labouze, J.; Nguyen, P. H.; Sterpone, F.; Berthoumieu, O.; Buchete, N.-V.; Coté, S.; De Simone, A.; Doig, A. J.; Faller, P.; Garcia, A.; Laio, A.; Li, M. S.; Melchionna, S.; Mousseau, N.; Mu, Y.; Paravastu, A.; Pasquali, S.; Rosenman, D. J.; Strodel, B.; Tarus, B.; Viles, J. H.; Zhang, T.; Wang, C.; Derreumaux, P., Amyloid β Protein and Alzheimer's Disease: When Computer Simulations Complement Experimental Studies. *Chemical Reviews* **2015**, *115*, 3518-3563.
66. Nakajima, N.; Nakamura, H.; Kidera, A., Multicanonical Ensemble Generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *The Journal of Physical Chemistry B* **1997**, *101*, 817-824.
67. Caves, L. S.; Evanseck, J. D.; Karplus, M., Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. *Protein Sci* **1998**, *7*, 649-666.
68. Harada, R.; Takano, Y.; Baba, T.; Shigeta, Y., Simple, yet powerful methodologies for conformational sampling of proteins. *Physical Chemistry Chemical Physics* **2015**, *17*, 6155-6173.
69. Harada, R.; Shigeta, Y., Efficient Conformational Search Based on Structural Dissimilarity Sampling: Applications for Reproducing Structural Transitions of Proteins. *Journal of Chemical Theory and Computation* **2017**, *13*, 1411-1423.
70. Salehi, N.; Amininasab, M.; Firouzi, R.; Karimi-Jafari, M. H., Efficient construction of a diverse conformational library for amyloid- β as an intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **2019**, *88*, 183-193.
71. Ahmed, R.; VanSchouwen, B.; Jafari, N.; Ni, X.; Ortega, J.; Melacini, G., Molecular Mechanism for the (-)-Epigallocatechin Gallate-Induced Toxic to Nontoxic Remodeling of A β Oligomers. *Journal of the American Chemical Society* **2017**, *139*, 13720-13734.
72. Lopez del Amo, J. M.; Fink, U.; Dasari, M.; Grelle, G.; Wanker, E. E.; Bieschke, J.; Reif, B., Structural Properties of EGCG-Induced, Nontoxic Alzheimer's Disease A β Oligomers. *Journal of Molecular Biology* **2012**, *421*, 517-524.
73. Wang, S.-H.; Liu, F.-F.; Dong, X.-Y.; Sun, Y., Thermodynamic Analysis of the Molecular Interactions between Amyloid β -Peptide 42 and (-)-Epigallocatechin-3-gallate. *The Journal of Physical Chemistry B* **2010**, *114*, 11576-11583.

74. Sinha, S.; Du, Z.; Maiti, P.; Klärner, F.-G.; Schrader, T.; Wang, C.; Bitan, G., Comparison of Three Amyloid Assembly Inhibitors: The Sugar scyllo-Inositol, the Polyphenol Epigallocatechin Gallate, and the Molecular Tweezer CLR01. *ACS Chemical Neuroscience* **2012**, *3*, 451-458.
75. Mori, T.; Koyama, N.; Tan, J.; Segawa, T.; Maeda, M.; Town, T., Combined treatment with the phenolics (-)-epigallocatechin-3-gallate and ferulic acid improves cognition and reduces Alzheimer-like pathology in mice. *Journal of Biological Chemistry* **2019**, *294*, 2714-5444.
76. Lantz, R.; Busbee, B.; Wojcikiewicz, E. P.; Du, D., Flavonoids with Vicinal Hydroxyl Groups Inhibit Human Calcitonin Amyloid Formation. *Chemistry – A European Journal* **2020**, *26*, 13063-13071.
77. Minh Hung, H.; Nguyen, M. T.; Tran, P.-T.; Truong, V. K.; Chapman, J.; Quynh Anh, L. H.; Derreumaux, P.; Vu, V. V.; Ngo, S. T., Impact of the Astaxanthin, Betanin, and EGCG Compounds on Small Oligomers of Amyloid A β 40 Peptide. *Journal of Chemical Information and Modeling* **2020**, *60*, 1399-1408.
78. Zhan, C.; Chen, Y.; Tang, Y.; Wei, G., Green Tea Extracts EGCG and EGC Display Distinct Mechanisms in Disrupting A β 42 Protofibril. *ACS Chemical Neuroscience* **2020**, *11*, 1841-1851.
79. Acharya, A.; Stockmann, J.; Beyer, L.; Rudack, T.; Nabers, A.; Gumbart, J. C.; Gerwert, K.; Batista, V. S., The Effect of (-)-Epigallocatechin-3-Gallate on the Amyloid- β Secondary Structure. *Biophysical Journal* **2020**, *119*, 349-359.
80. Chen, T.; Yang, Y.; Zhu, S.; Lu, Y.; Zhu, L.; Wang, Y.; Wang, X., Inhibition of A β aggregates in Alzheimer's disease by epigallocatechin and epicatechin-3-gallate from green tea. *Bioorg Chem* **2020**, *105*, 104382.
81. Ono, K.; Hamaguchi, T.; Naiki, H.; Yamada, M., Anti-amyloidogenic effects of antioxidants: Implications for the prevention and therapeutics of Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2006**, *1762*, 575-586.
82. Ono, K.; Yoshiike, Y.; Takashima, A.; Hasegawa, K.; Naiki, H.; Yamada, M., Potent anti-amyloidogenic and fibril-destabilizing effects of polyphenols in vitro: implications for the prevention and therapeutics of Alzheimer's disease. *Journal of Neurochemistry* **2003**, *87*, 172-181.
83. Ahmed, R.; Akcan, M.; Khondker, A.; Rheinstädter, M. C.; Bozelli, J. C.; Epand, R. M.; Huynh, V.; Wylie, R. G.; Boulton, S.; Huang, J.; Verschoor, C. P.; Melacini, G., Atomic resolution map of the soluble amyloid beta assembly toxic surfaces. *Chemical Science* **2019**, *10*, 6072-6082.
84. Takai, E.; Uda, K.; Yoshida, T.; Zako, T.; Maeda, M.; Shiraki, K., Cysteine inhibits the fibrillisation and cytotoxicity of amyloid- β 40 and 42: implications for the contribution of the thiophilic interaction. *Physical Chemistry Chemical Physics* **2014**, *16*, 3566-3572.
85. Gonçalves, P. B.; Sodero, A. C.; Cordeiro, Y., Green Tea Epigallocatechin-3-gallate (EGCG) Targeting Protein Misfolding in Drug Discovery for Neurodegenerative Diseases. *Biomolecules* **2021**, *11*.
86. Pervin, M.; Unno, K.; Ohishi, T.; Tanabe, H.; Miyoshi, N.; Nakamura, Y., Beneficial Effects of Green Tea Catechins on Neurodegenerative Diseases. *Molecules* **2018**, *23*, 1297.
87. Martinez Pomier, K.; Ahmed, R.; Melacini, G., Catechins as Tools to Understand the Molecular Basis of Neurodegeneration. *Molecules* **2020**, *25*, 3571.
88. Irwin, J. J.; Shoichet, B. K., ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling* **2005**, *45*, 177-182.
89. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994**, *98*, 11623-11627.
90. Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery Jr, J. A., General atomic and molecular electronic structure system. *Journal of Computational Chemistry* **1993**, *14*, 1347-1363.
91. Trott, O.; Olson, A. J., AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2010**, *31*, 455-461.

92. Nguyen, N. T.; Nguyen, T. H.; Pham, T. N. H.; Huy, N. T.; Bay, M. V.; Pham, M. Q.; Nam, P. C.; Vu, V. V.; Ngo, S. T., Autodock Vina Adopts More Accurate Binding Poses but Autodock4 Forms Better Binding Affinity. *Journal of Chemical Information and Modeling* **2020**, *60*, 204-211.
93. Bagheri, S.; Behnejad, H.; Firouzi, R.; Karimi-Jafari, M. H., Using the Semiempirical Quantum Mechanics in Improving the Molecular Docking: A Case Study with CDK2. *Molecular Informatics* **2020**, *39*, 2000036.
94. Cole, J.; Davis, E.; Jones, G.; Sage, C. R., 3.12 - Molecular Docking—A Solved Problem? In *Comprehensive Medicinal Chemistry III*, Chackalamannil, S.; Rotella, D.; Ward, S. E., Eds. Elsevier: Oxford, 2017; pp 297-318.
95. Liu, J.; Wang, R., Classification of Current Scoring Functions. *Journal of Chemical Information and Modeling* **2015**, *55*, 475-482.
96. Leach, A. R.; Shoichet, B. K.; Peishoff, C. E., Prediction of Protein–Ligand Interactions. Docking and Scoring: Successes and Gaps. *Journal of Medicinal Chemistry* **2006**, *49*, 5851-5855.
97. Scoring Functions for Protein-Ligand Docking. *Current Protein & Peptide Science* **2006**, *7*, 407-420.
98. Case, D. A.; Cheatham Iii, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005**, *26*, 1668-1688.
99. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157-1174.
100. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002**, *23*, 1623-1641.
101. Aliev, A. E.; Kulke, M.; Khaneja, H. S.; Chudasama, V.; Sheppard, T. D.; Lanigan, R. M., Motional timescale predictions by molecular dynamics simulations: Case study using proline and hydroxyproline sidechain dynamics. *Proteins: Structure, Function, and Bioinformatics* **2014**, *82*, 195-215.
102. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926-935.
103. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25.
104. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089-10092.
105. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995**, *103*, 8577-8593.
106. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101.
107. Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, *52*, 7182-7190.
108. Hess, B., P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 116-122.
109. Pérez-Nueno, V. I.; Rabal, O.; Borrell, J. I.; Teixidó, J., APIF: A New Interaction Fingerprint Based on Atom Pairs and Its Application to Virtual Screening. *Journal of Chemical Information and Modeling* **2009**, *49*, 1245-1260.
110. Tan, L.; Batista, J.; Bajorath, J., Rationalization of the Performance and Target Dependence of Similarity Searching Incorporating Protein–Ligand Interaction Information. *Journal of Chemical Information and Modeling* **2010**, *50*, 1042-1052.
111. Batista, J.; Tan, L.; Bajorath, J., Atom-Centered Interacting Fragments and Similarity Search Applications. *Journal of Chemical Information and Modeling* **2010**, *50*, 79-86.

112. Nishio, M.; Umezawa, Y.; Fantini, J.; Weiss, M. S.; Chakrabarti, P., CH- π hydrogen bonds in biological macromolecules. *Physical Chemistry Chemical Physics* **2014**, *16*, 12648-12683.
113. Weisel, M.; Bitter, H.-M.; Diederich, F.; So, W. V.; Kondru, R., PROLIX: Rapid Mining of Protein-Ligand Interactions in Large Crystal Structure Databases. *Journal of Chemical Information and Modeling* **2012**, *52*, 1450-1461.
114. Li, G.-B.; Yang, L.-L.; Wang, W.-J.; Li, L.-L.; Yang, S.-Y., ID-Score: A New Empirical Scoring Function Based on a Comprehensive Set of Descriptors Related to Protein-Ligand Interactions. *Journal of Chemical Information and Modeling* **2013**, *53*, 592-600.
115. Chimon, S.; Shaibat, M. A.; Jones, C. R.; Calero, D. C.; Aizezi, B.; Ishii, Y., Evidence of fibril-like β -sheet structures in a neurotoxic amyloid intermediate of Alzheimer's β -amyloid. *Nature Structural & Molecular Biology* **2007**, *14*, 1157-1164.
116. Berhanu, W. M.; Hansmann, U. H. E., Side-chain hydrophobicity and the stability of A β 16-22 aggregates. *Protein Science* **2012**, *21*, 1837-1848.
117. Crespi, G. A. N.; Hermans, S. J.; Parker, M. W.; Miles, L. A., Molecular basis for mid-region amyloid- β capture by leading Alzheimer's disease immunotherapies. *Scientific Reports* **2015**, *5*, 9649.
118. Riek, R.; Eisenberg, D. S., The activities of amyloids from a structural perspective. *Nature* **2016**, *539*, 227-235.
119. Liu, H.; Morris, C.; Lantz, R.; Kent, T. W.; Elbassal, E. A.; Wojcikiewicz, E. P.; Du, D., Residue-Specific Dynamics and Local Environmental Changes in A β 40 Oligomer and Fibril Formation. *Angewandte Chemie International Edition* **2018**, *57*, 8017-8021.
120. Li, F.; Zhan, C.; Dong, X.; Wei, G., Molecular mechanisms of resveratrol and EGCG in the inhibition of A β 42 aggregation and disruption of A β 42 protofibril: similarities and differences. *Physical Chemistry Chemical Physics* **2021**, *23*, 18843-18854.

Table 1. The number of selected docking complexes for each ligand with the different $\Delta\Delta G_{binding}$ cutoffs.

	$\Delta\Delta G_{binding}$ (kcal/mol) cutoff	<i>C</i>	<i>EC</i>	<i>EGC</i>	<i>ECG</i>	<i>EGCG</i>
Set-1	0.1	13495	13478	15646	15027	17431
Set-2	0.2	20768	20722	24811	23327	29923
Set-3	0.3	30544	30712	35758	34105	42833

Table 2. The number of contacts between the amino acid residues (the amino acid sequence is provided in the caption of **Figure 1**) of $A\beta_{42}$ and the *C* ligand of set-1 (13,495 complexes, **Table 1**) with five different distance cutoffs as indicated. The aromatic residues are in bold typeface and three most favorable aromatic residue hotspots are highlighted in light teal.

3.0 Å			3.5 Å			4.0 Å			4.5 Å			5.0 Å		
Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population
PHE	19	4455	PHE	19	5160	PHE	19	5510	PHE	19	5853	PHE	19	6290
TYR	10	4263	TYR	10	4916	PHE	20	5297	PHE	20	5739	PHE	20	6242
PHE	20	4138	PHE	20	4884	TYR	10	5267	TYR	10	5603	TYR	10	6014
GLN	15	3963	GLN	15	4602	GLN	15	4987	GLN	15	5419	GLN	15	5911
LYS	16	3745	LYS	16	4220	LYS	16	4611	LYS	16	5040	LYS	16	5586
LEU	17	3472	LEU	17	4053	LEU	17	4515	LEU	17	4999	LEU	17	5581
HIS	13	3402	HIS	13	4030	HIS	13	4441	HIS	13	4885	VAL	18	5423
HIS	14	3338	HIS	14	4000	HIS	14	4412	HIS	14	4850	HIS	13	5386
VAL	18	3246	VAL	18	3938	VAL	18	4371	VAL	18	4838	HIS	14	5381
VAL	12	3027	VAL	12	3713	VAL	12	4171	VAL	12	4661	VAL	12	5279
GLU	11	2926	GLU	11	3418	GLU	11	3797	GLU	11	4265	GLU	11	4785
ARG	5	2642	ALA	21	3237	ALA	21	3688	ALA	21	4169	ALA	21	4721
PHE	4	2628	GLU	22	3061	GLU	22	3411	GLU	22	3776	GLY	9	4225
ALA	21	2617	ARG	5	3054	ARG	5	3360	GLY	9	3718	GLU	22	4223
GLU	22	2587	HIS	6	3038	HIS	6	3321	ARG	5	3712	SER	8	4122
HIS	6	2530	PHE	4	3026	GLY	9	3262	SER	8	3632	ARG	5	4075
SER	8	2337	SER	8	2869	SER	8	3258	HIS	6	3615	HIS	6	3934
GLY	9	2265	GLY	9	2866	PHE	4	3255	PHE	4	3452	ASP	7	3715
VAL	24	2154	ASP	7	2587	VAL	24	2931	VAL	24	3289	VAL	24	3710
ASP	7	2140	VAL	24	2587	ASP	7	2904	ASP	7	3265	PHE	4	3686
ASP	23	1905	ASP	23	2327	ASP	23	2602	ASP	23	2985	ASP	23	3430
ASN	27	1895	ASN	27	2190	ASN	27	2393	SER	26	2721	SER	26	3110
SER	26	1738	SER	26	2076	SER	26	2374	ASN	27	2667	ASN	27	2975
GLU	3	1705	GLU	3	2037	GLU	3	2256	GLY	25	2514	GLY	25	2927
LYS	28	1684	LYS	28	1957	GLY	25	2186	GLU	3	2463	LYS	28	2766
GLY	25	1459	GLY	25	1879	LYS	28	2172	LYS	28	2433	GLU	3	2685
ILE	31	1452	ILE	31	1711	ILE	31	1916	ILE	31	2172	ILE	31	2473
ALA	30	1327	ALA	30	1613	ALA	30	1844	ALA	30	2101	ALA	30	2427
ILE	32	1291	ILE	32	1538	ILE	32	1749	GLY	29	1981	GLY	29	2310
LEU	34	1225	GLY	29	1489	GLY	29	1719	ILE	32	1979	ILE	32	2254
ALA	2	1169	LEU	34	1461	ALA	2	1641	ALA	2	1843	LEU	34	2110
GLY	29	1151	ALA	2	1434	LEU	34	1637	LEU	34	1842	ALA	2	2040
VAL	36	1049	VAL	36	1264	MET	35	1449	MET	35	1652	MET	35	1875
MET	35	966	MET	35	1245	VAL	36	1415	VAL	36	1582	VAL	36	1822
GLY	33	914	GLY	33	1148	GLY	33	1297	GLY	33	1471	GLY	33	1672
VAL	39	854	VAL	39	1048	VAL	39	1193	VAL	39	1342	VAL	39	1559
VAL	40	832	VAL	40	982	ASP	1	1116	ASP	1	1265	ASP	1	1416
ASP	1	806	ASP	1	977	VAL	40	1084	VAL	40	1225	VAL	40	1392
ILE	41	771	ILE	41	948	ILE	41	1074	ILE	41	1213	GLY	37	1373
GLY	37	648	GLY	37	815	GLY	37	984	GLY	37	1135	ILE	41	1372
GLY	38	643	GLY	38	814	GLY	38	953	GLY	38	1092	GLY	38	1305
ALA	42	468	ALA	42	580	ALA	42	674	ALA	42	764	ALA	42	870

Table 3. The number of contacts between the amino acid residues (the amino acid sequence is provided in the caption of **Figure 1**) of $A\beta_{42}$ and the *EC* ligand for set-1 (13,478 complexes, **Table 1**) with five different distance cutoffs as indicated. The aromatic residues are in bold typeface and three most favorable aromatic residue hotspots are highlighted in light teal.

3.0 Å			3.5 Å			4.0 Å			4.5 Å			5.0 Å		
Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population
PHE	19	4264	PHE	19	4971	PHE	19	5355	PHE	19	5741	PHE	19	6194
TYR	10	4181	TYR	10	4899	TYR	10	5273	TYR	10	5620	PHE	20	6083
PHE	20	4008	PHE	20	4776	PHE	20	5190	PHE	20	5617	TYR	10	5967
GLN	15	3947	GLN	15	4527	GLN	15	4907	GLN	15	5308	GLN	15	5800
LYS	16	3643	LYS	16	4103	LYS	16	4483	LYS	16	4954	LEU	17	5497
HIS	13	3433	HIS	13	4070	HIS	13	4459	HIS	13	4892	LYS	16	5456
LEU	17	3355	LEU	17	3963	LEU	17	4403	LEU	17	4890	VAL	18	5418
HIS	14	3324	HIS	14	3959	HIS	14	4377	HIS	14	4834	HIS	14	5403
VAL	18	3255	VAL	18	3878	VAL	18	4322	VAL	18	4820	HIS	13	5375
VAL	12	3113	VAL	12	3740	VAL	12	4186	VAL	12	4660	VAL	12	5214
GLU	11	2890	GLU	11	3391	GLU	11	3789	GLU	11	4250	GLU	11	4776
ARG	5	2802	ALA	21	3299	ALA	21	3718	ALA	21	4169	ALA	21	4671
ALA	21	2678	ARG	5	3237	ARG	5	3500	ARG	5	3817	GLY	9	4282
GLU	22	2648	HIS	6	3112	GLU	22	3432	GLU	22	3812	GLU	22	4244
HIS	6	2624	GLU	22	3099	HIS	6	3364	GLY	9	3772	SER	8	4177
PHE	4	2551	PHE	4	2991	GLY	9	3325	SER	8	3719	ARG	5	4119
SER	8	2461	GLY	9	2896	SER	8	3275	HIS	6	3671	HIS	6	3999
GLY	9	2279	SER	8	2895	PHE	4	3242	PHE	4	3467	ASP	7	3829
ASP	7	2218	ASP	7	2659	ASP	7	2982	ASP	7	3355	PHE	4	3694
VAL	24	2094	VAL	24	2520	VAL	24	2862	VAL	24	3214	VAL	24	3688
ASN	27	1894	ASP	23	2283	ASP	23	2582	ASP	23	2934	ASP	23	3414
ASP	23	1871	ASN	27	2182	SER	26	2442	SER	26	2769	SER	26	3150
SER	26	1831	SER	26	2158	ASN	27	2419	ASN	27	2695	ASN	27	3001
GLU	3	1702	GLU	3	2024	GLY	25	2259	GLY	25	2579	GLY	25	2952
LYS	28	1617	GLY	25	1972	GLU	3	2238	GLU	3	2431	LYS	28	2717
GLY	25	1571	LYS	28	1891	LYS	28	2116	LYS	28	2395	GLU	3	2676
ILE	31	1473	ILE	31	1753	ILE	31	1939	ILE	31	2201	ILE	31	2501
ALA	30	1373	ALA	30	1666	ALA	30	1871	ALA	30	2123	ALA	30	2458
ILE	32	1307	ILE	32	1582	ILE	32	1785	GLY	29	2049	GLY	29	2336
LEU	34	1250	GLY	29	1505	GLY	29	1742	ILE	32	2010	ILE	32	2298
GLY	29	1179	LEU	34	1451	ALA	2	1634	ALA	2	1837	LEU	34	2095
ALA	2	1128	ALA	2	1427	LEU	34	1602	LEU	34	1831	ALA	2	2068
VAL	36	1014	VAL	36	1227	MET	35	1414	MET	35	1630	MET	35	1889
MET	35	971	MET	35	1207	VAL	36	1371	VAL	36	1575	VAL	36	1810
GLY	33	967	GLY	33	1185	GLY	33	1332	GLY	33	1522	GLY	33	1729
VAL	39	856	VAL	39	1043	VAL	39	1183	VAL	39	1344	VAL	39	1569
ASP	1	793	ASP	1	981	ASP	1	1125	ASP	1	1260	ASP	1	1414
VAL	40	764	VAL	40	918	VAL	40	1043	VAL	40	1191	GLY	37	1412
ILE	41	724	ILE	41	890	GLY	37	1030	GLY	37	1191	VAL	40	1353
GLY	37	686	GLY	37	868	ILE	41	1023	ILE	41	1170	GLY	38	1346
GLY	38	608	GLY	38	803	GLY	38	926	GLY	38	1112	ILE	41	1310
ALA	42	446	ALA	42	534	ALA	42	631	ALA	42	734	ALA	42	837

Table 4. The number of contacts between the amino acid residues (the amino acid sequence is provided in the caption of **Figure 1**) of $A\beta_{42}$ and the *ECG* ligand for set-1 (15,027 complexes, **Table 1**) with five different distance cutoffs as indicated. The aromatic residues are in bold typeface and three most favorable aromatic residue hotspots are highlighted in light teal.

3.0 Å			3.5 Å			4.0 Å			4.5 Å			5.0 Å		
Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population
PHE	19	5225	PHE	19	6050	PHE	19	6491	PHE	19	6953	PHE	19	7425
TYR	10	5192	TYR	10	5937	TYR	10	6338	PHE	20	6751	PHE	20	7304
PHE	20	4974	PHE	20	5792	PHE	20	6266	TYR	10	6743	TYR	10	7187
GLN	15	4954	GLN	15	5629	GLN	15	6025	GLN	15	6506	GLN	15	7058
LYS	16	4384	HIS	13	5062	HIS	13	5589	HIS	13	6053	HIS	14	6641
HIS	13	4322	LYS	16	4985	LYS	16	5467	HIS	14	6017	HIS	13	6620
HIS	14	4218	HIS	14	4944	HIS	14	5459	LYS	16	5968	LEU	17	6578
LEU	17	4134	LEU	17	4838	LEU	17	5332	LEU	17	5901	VAL	18	6545
VAL	18	4081	VAL	18	4820	VAL	18	5322	VAL	18	5877	LYS	16	6545
VAL	12	4029	VAL	12	4770	VAL	12	5305	VAL	12	5857	VAL	12	6495
GLU	11	3786	GLU	11	4406	GLU	11	4862	GLU	11	5393	GLU	11	5988
ARG	5	3652	ALA	21	4112	ALA	21	4620	ALA	21	5151	ALA	21	5763
ALA	21	3416	ARG	5	4101	ARG	5	4460	GLY	9	4876	GLY	9	5452
GLU	22	3332	GLU	22	3894	GLY	9	4328	ARG	5	4796	SER	8	5246
HIS	6	3297	SER	8	3889	SER	8	4312	SER	8	4712	ARG	5	5178
SER	8	3275	HIS	6	3882	GLU	22	4260	GLU	22	4690	GLU	22	5176
PHE	4	3272	GLY	9	3830	HIS	6	4245	HIS	6	4610	HIS	6	4978
GLY	9	3080	PHE	4	3787	PHE	4	4088	PHE	4	4357	ASP	7	4792
ASP	7	2965	ASP	7	3563	ASP	7	3917	ASP	7	4315	PHE	4	4623
VAL	24	2712	VAL	24	3259	VAL	24	3632	VAL	24	4062	VAL	24	4612
ASP	23	2550	ASP	23	3042	ASP	23	3404	ASP	23	3805	ASP	23	4334
ASN	27	2459	ASN	27	2851	SER	26	3202	SER	26	3560	SER	26	3998
SER	26	2413	SER	26	2848	ASN	27	3116	ASN	27	3406	ASN	27	3808
GLU	3	2320	GLU	3	2752	GLU	3	3025	GLY	25	3273	GLY	25	3751
LYS	28	2143	GLY	25	2527	GLY	25	2890	GLU	3	3259	GLU	3	3548
GLY	25	2052	LYS	28	2468	LYS	28	2772	LYS	28	3153	LYS	28	3512
ILE	31	1780	ILE	31	2114	ILE	31	2365	ILE	31	2642	ALA	30	3044
ALA	30	1659	ALA	30	2066	ALA	30	2325	ALA	30	2634	ILE	31	3014
ILE	32	1640	GLY	29	1937	GLY	29	2248	GLY	29	2533	GLY	29	2903
LEU	34	1568	ILE	32	1932	ALA	2	2194	ILE	32	2476	ILE	32	2808
ALA	2	1529	ALA	2	1919	ILE	32	2191	ALA	2	2470	ALA	2	2753
GLY	29	1507	LEU	34	1841	LEU	34	2027	LEU	34	2252	LEU	34	2549
VAL	36	1310	VAL	36	1589	MET	35	1803	MET	35	2060	MET	35	2415
MET	35	1251	MET	35	1553	VAL	36	1786	VAL	36	2018	VAL	36	2284
ASP	1	1192	GLY	33	1436	GLY	33	1646	GLY	33	1854	GLY	33	2139
GLY	33	1163	ASP	1	1435	ASP	1	1618	ASP	1	1805	VAL	39	2038
VAL	39	1115	VAL	39	1347	VAL	39	1572	VAL	39	1785	ASP	1	2008
ILE	41	1046	VAL	40	1267	VAL	40	1424	ILE	41	1622	VAL	40	1824
VAL	40	1043	ILE	41	1234	ILE	41	1420	VAL	40	1604	ILE	41	1805
GLY	37	848	GLY	37	1094	GLY	37	1300	GLY	37	1505	GLY	37	1792
GLY	38	836	GLY	38	1073	GLY	38	1272	GLY	38	1477	GLY	38	1761
ALA	42	656	ALA	42	829	ALA	42	948	ALA	42	1082	ALA	42	1202

Table 5. The number of contacts between the amino acid residues (the amino acid sequence is provided in the caption of **Figure 1**) of $A\beta_{42}$ and the *EGC* ligand for set-1 (15,646 complexes, **Table 1**) with five different distance cutoffs as indicated. The aromatic residues are in bold typeface and the three most favorable aromatic residue hotspots are highlighted in light teal.

3.0 Å			3.5 Å			4.0 Å			4.5 Å			5.0 Å		
Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population
TYR	10	5045	TYR	10	5821	TYR	10	6228	TYR	10	6643	TYR	10	7107
PHE	19	4812	PHE	19	5579	PHE	19	5989	PHE	19	6443	PHE	19	7010
GLN	15	4686	PHE	20	5397	PHE	20	5873	PHE	20	6376	PHE	20	6931
HIS	13	4194	GLN	15	5325	GLN	15	5777	GLN	15	6236	GLN	15	6792
LYS	16	4094	HIS	13	4938	HIS	13	5403	HIS	13	5875	HIS	13	6456
HIS	14	4085	HIS	14	4787	HIS	14	5262	HIS	14	5824	HIS	14	6438
LEU	17	3818	LYS	16	4649	LYS	16	5108	LYS	16	5665	LYS	16	6300
VAL	12	3747	LEU	17	4491	LEU	17	5014	VAL	12	5588	VAL	12	6254
VAL	18	3739	VAL	12	4463	VAL	12	5000	LEU	17	5571	LEU	17	6251
GLU	11	3570	VAL	18	4417	VAL	18	4910	VAL	18	5514	VAL	18	6212
ARG	5	3505	GLU	11	4095	GLU	11	4554	GLU	11	5115	GLU	11	5730
HIS	6	3346	ARG	5	4024	ARG	5	4408	ALA	21	4895	ALA	21	5459
GLU	22	3201	HIS	6	3899	ALA	21	4339	ARG	5	4776	GLY	9	5250
ALA	21	3132	ALA	21	3839	HIS	6	4232	GLY	9	4640	SER	8	5198
PHE	4	3114	GLU	22	3712	SER	8	4130	HIS	6	4601	ARG	5	5144
SER	8	3062	SER	8	3654	GLU	22	4117	SER	8	4588	GLU	22	5051
GLY	9	2901	PHE	4	3601	GLY	9	4093	GLU	22	4553	HIS	6	4989
ASP	7	2810	GLY	9	3593	PHE	4	3903	PHE	4	4202	ASP	7	4670
VAL	24	2480	ASP	7	3312	ASP	7	3698	ASP	7	4123	PHE	4	4522
ASN	27	2348	VAL	24	2955	VAL	24	3331	VAL	24	3806	VAL	24	4373
ASP	23	2341	ASP	23	2796	ASP	23	3113	ASP	23	3558	ASP	23	4127
SER	26	2283	SER	26	2710	SER	26	3041	SER	26	3430	SER	26	3871
GLU	3	2180	ASN	27	2683	ASN	27	2937	ASN	27	3260	ASN	27	3660
GLY	25	1904	GLU	3	2599	GLU	3	2828	GLY	25	3151	GLY	25	3598
LYS	28	1903	GLY	25	2383	GLY	25	2753	GLU	3	3094	GLU	3	3387
ILE	31	1689	LYS	28	2235	LYS	28	2489	LYS	28	2860	LYS	28	3259
ALA	30	1618	ILE	31	1987	ILE	31	2239	ILE	31	2544	ILE	31	2928
GLY	29	1510	ALA	30	1961	ALA	30	2203	ALA	30	2516	ALA	30	2910
ILE	32	1502	GLY	29	1867	GLY	29	2139	GLY	29	2443	GLY	29	2808
ALA	2	1433	ILE	32	1826	ALA	2	2086	ILE	32	2347	ILE	32	2689
LEU	34	1412	ALA	2	1826	ILE	32	2057	ALA	2	2336	ALA	2	2595
VAL	36	1166	LEU	34	1638	LEU	34	1851	LEU	34	2110	LEU	34	2397
MET	35	1135	VAL	36	1412	MET	35	1608	MET	35	1849	MET	35	2185
GLY	33	1115	GLY	33	1385	VAL	36	1608	VAL	36	1840	VAL	36	2104
ASP	1	1059	MET	35	1379	GLY	33	1547	GLY	33	1767	GLY	33	2004
VAL	39	1027	ASP	1	1274	ASP	1	1457	ASP	1	1664	ASP	1	1872
VAL	40	901	VAL	39	1236	VAL	39	1419	VAL	39	1618	VAL	39	1867
ILE	41	885	VAL	40	1077	VAL	40	1221	VAL	40	1407	GLY	37	1635
GLY	37	785	ILE	41	1060	ILE	41	1218	ILE	41	1403	VAL	40	1620
GLY	38	770	GLY	37	1014	GLY	37	1170	GLY	38	1365	GLY	38	1591
ALA	42	547	GLY	38	1009	GLY	38	1160	GLY	37	1360	ILE	41	1557
			ALA	42	669	ALA	42	790	ALA	42	921	ALA	42	1060

Table 6. The number of contacts between the amino acid residues (the amino acid sequence is provided in the caption of **Figure 1**) of $A\beta_{42}$ and the *EGCG* ligand for set-1 (17,431 complexes, **Table 1**) with five different distance cutoffs as indicated. The aromatic residues are in bold typeface and three most favorable aromatic residue hotspots are highlighted in light teal.

3.0 Å			3.5 Å			4.0 Å			4.5 Å			5.0 Å		
Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population	Residue	ID	Population
TYR	10	6305	TYR	10	7109	TYR	10	7572	TYR	10	8059	TYR	10	8585
PHE	19	5892	PHE	19	6734	PHE	19	7225	PHE	19	7758	PHE	19	8302
GLN	15	5757	GLN	15	6495	PHE	20	7012	PHE	20	7648	PHE	20	8282
PHE	20	5580	PHE	20	6450	GLN	15	6960	GLN	15	7545	GLN	15	8159
HIS	13	5184	HIS	13	6020	HIS	13	6604	HIS	13	7189	HIS	13	7861
LYS	16	5071	HIS	14	5811	HIS	14	6400	HIS	14	7066	HIS	14	7760
HIS	14	4997	LYS	16	5723	VAL	12	6323	VAL	12	6989	VAL	12	7712
VAL	12	4802	VAL	12	5694	LYS	16	6253	LYS	16	6840	LYS	16	7546
LEU	17	4699	VAL	18	5496	VAL	18	6093	VAL	18	6733	LEU	17	7538
VAL	18	4685	LEU	17	5460	LEU	17	6017	LEU	17	6728	VAL	18	7463
GLU	11	4579	GLU	11	5311	GLU	11	5860	GLU	11	6492	GLU	11	7189
ARG	5	4483	ARG	5	5048	ARG	5	5437	ALA	21	5911	ALA	21	6624
HIS	6	4115	HIS	6	4797	ALA	21	5303	ARG	5	5886	GLY	9	6552
SER	8	4048	ALA	21	4779	GLY	9	5224	GLY	9	5856	SER	8	6389
PHE	4	3974	SER	8	4685	SER	8	5218	SER	8	5737	ARG	5	6294
ALA	21	3972	GLY	9	4605	HIS	6	5212	HIS	6	5660	HIS	6	6105
GLU	22	3883	PHE	4	4525	GLU	22	4907	GLU	22	5434	GLU	22	5944
GLY	9	3787	GLU	22	4479	PHE	4	4901	ASP	7	5281	ASP	7	5906
ASP	7	3646	ASP	7	4351	ASP	7	4781	PHE	4	5269	PHE	4	5621
VAL	24	3142	VAL	24	3745	VAL	24	4162	VAL	24	4662	VAL	24	5316
ASP	23	2959	ASP	23	3502	ASP	23	3880	ASP	23	4398	ASP	23	4988
SER	26	2873	SER	26	3372	SER	26	3766	SER	26	4208	SER	26	4651
ASN	27	2856	GLU	3	3328	GLU	3	3658	GLU	3	3988	ASN	27	4429
GLU	3	2819	ASN	27	3274	ASN	27	3584	ASN	27	3957	GLY	25	4363
LYS	28	2470	GLY	25	2920	GLY	25	3358	GLY	25	3820	GLU	3	4314
GLY	25	2373	LYS	28	2842	LYS	28	3226	LYS	28	3618	LYS	28	4067
ILE	31	2113	ILE	31	2453	ALA	2	2712	ALA	30	3080	ALA	30	3506
ALA	2	1937	ALA	2	2382	ILE	31	2712	ILE	31	3029	ILE	31	3471
ALA	30	1898	ALA	30	2371	ALA	30	2696	ALA	2	3005	ALA	2	3358
ILE	32	1863	GLY	29	2255	GLY	29	2597	GLY	29	2921	GLY	29	3320
GLY	29	1809	ILE	32	2239	ILE	32	2524	ILE	32	2866	ILE	32	3261
LEU	34	1780	LEU	34	2104	LEU	34	2330	LEU	34	2582	LEU	34	2977
ASP	1	1531	MET	35	1837	MET	35	2104	MET	35	2408	MET	35	2792
VAL	36	1516	VAL	36	1818	VAL	36	2047	VAL	36	2296	VAL	36	2595
MET	35	1469	ASP	1	1815	ASP	1	2027	ASP	1	2251	ASP	1	2452
VAL	39	1350	GLY	33	1661	GLY	33	1883	GLY	33	2135	GLY	33	2451
GLY	33	1348	VAL	39	1629	VAL	39	1872	VAL	39	2115	VAL	39	2423
ILE	41	1208	VAL	40	1455	ILE	41	1675	ILE	41	1885	GLY	38	2126
VAL	40	1205	ILE	41	1454	VAL	40	1649	VAL	40	1848	ILE	41	2103
GLY	38	1060	GLY	37	1340	GLY	37	1571	GLY	38	1787	VAL	40	2094
GLY	37	1040	GLY	38	1336	GLY	38	1541	GLY	37	1781	GLY	37	2090
ALA	42	818	ALA	42	1006	ALA	42	1147	ALA	42	1299	ALA	42	1466

Table 7. Average $A\beta_{42}$ end-to-end distance (D1–A42) and distances between EGCG and Y10, F19, F20 residues (the amino acid sequence is provided in the caption of Figure 1). Standard deviations are provided in parentheses.

	Distance (nm)				
	C	EC	ECG	EGC	EGCG
end-to-end	1.10 (0.66)	1.69 (0.73)	1.30 (0.48)	1.77 (0.60)	1.54 (0.59)
Ligand...Y10	1.08 (0.45)	0.47 (0.40)	0.54 (0.18)	0.29 (0.08)	0.70 (0.22)
Ligand...F19	0.97 (0.54)	0.95 (0.51)	0.55 (0.25)	0.50 (0.20)	0.40 (0.23)
Ligand...F20	0.74 (0.53)	0.72 (0.39)	0.44 (0.24)	0.44 (0.20)	0.36 (0.26)

Table 8. Average number of hydrogen bonds (donor...acceptor ≤ 0.35 nm and α ($\angle$ (Hydrogen–Donor–Acceptor) $\leq 30^\circ$). Standard deviations are provided in parentheses.

	Number of hydrogen bonds (SD)				
	C	EC	ECG	EGC	EGCG
$A\beta$...Ligand	2 (2)	2 (1)	4 (2)	4 (1)	5 (2)
$A\beta$...Solvent	109 (17)	114 (20)	108 (21)	114 (21)	112 (8)
Ligand...Solvent	6 (2)	5 (3)	7 (4)	5 (3)	7 (3)

Scheme 1. Flowchart of steps taken in this study.

Figure 1. A random structure of Amyloid β_{42} . $A\beta_{42}$ sequence: ¹DAEFRHDSG ¹⁰YEVHHQKLVF ²⁰FAEDVGSNKG ³⁰AIIGLMVGGV ⁴⁰VI⁴²A. Asp1 (N-terminus) and Ala42 (C-terminus) are shown using van der Waals radii. Other residues are represented in licorice. Acidic, basic, polar, and non-polar amino acids are shown in red, blue, green, and white, respectively.

Figure 2. 2D and 3D representations of catechins considered in this study.

Figure 3. The histogram of binding affinity (kcal/mol) of the C ligand in set-1 (cutoff of 0.1 kcal/mol, Table 1). The binding affinities were divided to bins of 0.1 kcal/mol, and the numbers on top of the bins show the number of complexes in each bin. The three aromatic residue hotspots (*i.e.* Tyr10, Phe19, and Phe20, see Figure 1 for sequence) and interacting aromatic residues with the ligand for some complexes with the highest binding affinities are shown in licorice representation. Favorable aromatic interactions between the ligand and peptide are depicted with orange dashed lines.

Figure 4. The histogram of binding affinity (kcal/mol) of the *EC* ligand over the set-1(cutoff of 0.1 kcal/mol, Table 1). The binding affinities were divided to bins of 0.1 kcal/mol, and the numbers on top of the bins show the number of complexes in each bin. The three aromatic residue hotspots (*i.e.* Tyr10, Phe19, and Phe20, see Figure 1 for sequence) and interacting aromatic residues with the ligand for some complexes with the highest binding affinities are shown in licorice representation. Favorable aromatic interactions between the ligand and peptide are depicted with orange dashed lines.

Figure 5. The histogram of binding affinity (kcal/mol) of the *ECG* ligand over the set-1 (cutoff of 0.1 kcal/mol, Table 1). The binding affinities were divided to bins of 0.1 kcal/mol, and the numbers on top of the bins show the number of complexes in each bin. The three aromatic residue hotspots (*i.e.* Tyr10, Phe19, and Phe20, see Figure 1 for sequence) and interacting aromatic residues with the ligand for some complexes with the highest binding affinities are shown in licorice representation. Favorable aromatic interactions between the ligand and peptide are depicted with orange dashed lines.

Figure 6. The histogram of binding affinity (kcal/mol) of the *EGC* ligand over the set-1 (cutoff of 0.1 kcal/mol, Table 1). The binding affinities were divided to bins of 0.1 kcal/mol, and the numbers on top of the bins show the number of complexes in each bin. The three aromatic residue hotspots (*i.e.* Tyr10, Phe19, and Phe20, see Figure 1 for sequence) and interacting aromatic residues with the ligand for some complexes with the highest binding affinities are shown in licorice representation. Favorable aromatic interactions between the ligand and peptide are depicted with orange dashed lines.

Figure 7. The histogram of binding affinity (kcal/mol) of the *EGCG* ligand over the set-1 (cutoff of 0.1 kcal/mol, Table 1). The binding affinities were divided to bins of 0.1 kcal/mol, and the numbers on top of the bins show the number of complexes for each bin. The three aromatic residue hotspots (*i.e.* Tyr10, Phe19, and Phe20, see Figure 1 for sequence) and interacting aromatic residues with the ligand for some complexes with the highest binding affinities are shown in licorice representation. Favorable aromatic interactions between the ligand and peptide are depicted with orange dashed lines.

Supporting Information

Identification of Catechins Binding Pockets in Monomeric A β ₄₂ through Ensemble Docking and MD Simulations

Rohoullah Firouzi^{a*}, Shahin Sowlati-Hashjin^b, Cecilia Chávez-García^(c,d), Mitra Ashouri^e, Mohammad Hossein Karimi-Jafari^f, Mikko Karttunen^{(c,d,g)*}

^(a) Department of Physical Chemistry, Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran. ^(b) Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, M5S 3G9, Canada. ^(c) Department of Chemistry, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 5B7, Canada. ^(d) The Centre of Advanced Materials and Biomaterials Research, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 5B7, Canada. ^(e) Department of Physical Chemistry, School of Chemistry, College of Science, University of Tehran, Tehran, Iran. ^(f) Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran. ^(g) Department of Physics and Astronomy, The University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 3K7, Canada.

Table S1. RMSD, R_g, and SASA of A β ₄₂ in the different systems (standard deviations in parentheses).

	Systems				
	C	EC	ECG	EGC	EGCG
RMSD (nm)	0.85 (0.12)	1.26 (0.09)	0.7 (0.09)	1.26 (0.17)	0.99 (0.14)
R_g (nm)	0.96 (0.09)	0.98 (0.07)	0.96 (0.07)	0.99 (0.07)	1.01 (0.06)
SASA (nm²)	35.56 (3.19)	36.82 (2.53)	36.96 (2.03)	37.07 (1.91)	38.19 (1.89)

Figure S1. Initial (left) and final (right, at 3.0 μs) structures of Aβ₄₂ in the presence of ligands. Aβ₄₂ is shown using ribbons and the ligands using the van der Waals representation.

Figure S2. RMSD of $A\beta_{42}$ backbone in the presence of different ligands.

Figure S3. Radius of gyration of $A\beta_{42}$ backbone in the presence of different ligands.

Figure S4. Solvent accessible surface area of $A\beta_{42}$ backbone in the presence of different ligands. SASA was calculated based on the numerical Double Cubic Lattice Method (DCLM)¹ as implemented in the GROMACS package using 1 ns intervals.

Figure S5. (a) Average $A\beta_{42}$ end-to-end distance (D1–A42) and distances between C and (b) Y10, (c) F19, and (d) F20 residues.

Figure S6. (a) Average $A\beta_{42}$ end-to-end distance (D1–A42) and distances between EC and (b) Y10, (c) F19, and (d) F20 residues.

Figure S7. (a) Average $A\beta_{42}$ end-to-end distance (D1-A42) and distances between ECG and (b) Y10, (c) F19, and (d) F20 residues.

Figure S8. (a) Average $A\beta_{42}$ end-to-end distance (D1-A42) and distances between ECG and (b) Y10, (c) F19, and (d) F20 residues.

Figure S9. (a) Average $A\beta_{42}$ end-to-end distance (D1–A42) and distances between EGCG and (b) Y10, (c) F19, and (d) F20 residues.

Figure S10. Hydrogen bonds between C and $A\beta_{42}$ selected residues. Anionic, cationic, and neutral residues are shown in red, blue, and green, respectively.

Figure S11. Hydrogen bonds between EC and Aβ₄₂ selected residues. Anionic, cationic, and neutral residues are shown in red, blue, and green, respectively.

Figure S12. Hydrogen bonds between ECG and A β_{42} selected residues. Anionic, cationic, and neutral residues are shown in red, blue, and green, respectively.

Figure S13. Hydrogen bonds between EGC and A β ₄₂ selected residues. Anionic, cationic, and neutral residues are shown in red, blue, and green, respectively.

Figure S14. Hydrogen bonds between EGCG and Aβ₄₂ selected residues. Anionic, cationic, and neutral residues are shown in red, blue, and green, respectively.

Figure S15. Average number of hydrogen bonds between Aβ₄₂ amino acids and the ligands.

References

1. Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M., The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **1995**, *16*, 273-284.