Journal of Emerging Investigators

In silico comparative screening of anti-amyloidogenic properties of a targeted library of natural product polyphenols towards identification of potential leads in therapeutics for amyloid-linked neurodegenerative diseases

Manuscript Draft

Manuscript Number:

Full Title: In silico comparative screening of anti-amyloidogenic properties of a targeted library of natural product polyphenols towards identification of potential leads in therapeutics for amyloid-linked neurodegenerative diseases

Article Type: Research Article

Manuscript Classifications: Biology; Chemistry; Biophysics; Biochemistry; Organic Chemistry; Structural Biology

Keywords: Amyloid beta, polyphenols, molecular dynamics, molecular docking, Alzheimer's disease

Corresponding Author: Edward Njoo
Aspiring Scholars Directed Research Program
Fremont, CA UNITED STATES

Corresponding Author's Institution: Aspiring Scholars Directed Research Program

Corresponding Author E-Mail: edward.njoo@asdrp.org

Order of Authors: Aashi Shah
Shloka Raghavan
Shamita Bhattacharjee
Anushka Rajasekhar
Varsha Beldona
Pranjal Verma
Aishwarya Yuvaraj
Harsha Rajkumar
Edward Njoo

Manuscript Region of Origin: UNITED STATES

Abstract: The aggregation of the amyloid beta 42 (Aβ-42) peptide in neurons causes the formation of amyloid plaques, which are linked to the development and progression of Alzheimer’s disease. Inhibition of amyloid protein aggregation has been investigated as a potential therapeutic strategy in the treatment and prevention of Alzheimer’s disease. Polyphenols, which are natural product small molecules produced in plants with various bioactive properties, have previously demonstrated activity in inhibiting the aggregation of Aβ-42. Here, the Aβ-42 peptide inhibition potential of each polyphenol is determined using the compound’s individual stability and its computationally-predicted binding affinity to the peptide.

Additional Information:

Question Response

Author Comments:
In silico comparative screening of anti-amyloidogenic properties of a targeted library of natural product polyphenols towards identification of potential leads in therapeutics for amyloid-linked neurodegenerative diseases


¹Department of Chemistry, Biochemistry, & Physical Science, Aspiring Scholars Directed Research Program, Fremont, CA
²Amador Valley High School, Pleasanton, CA
³Irvington High School, Fremont, CA
⁴Quarry Lane School, Dublin, CA
⁵Dougherty Valley High School, San Ramon, CA
⁶American High School, Fremont, CA
⁷Fremont Christian High School, Fremont, CA

*Denotes equal contribution.

Correspondence: edward.njoo@asdrp.org
Abstract

The aggregation of the amyloid beta 42 (Aβ-42) peptide in neurons causes the formation of amyloid plaques, which are linked to the development and progression of Alzheimer’s disease. Inhibition of amyloid protein aggregation has been investigated as a potential therapeutic strategy in the treatment and prevention of Alzheimer’s disease. Polyphenols, which are natural product small molecules produced in plants with various bioactive properties, have previously demonstrated activity in inhibiting the aggregation of Aβ-42. Here, the Aβ-42 peptide inhibition potential of each polyphenol is determined using the compound’s individual stability and its computationally-predicted binding affinity to the peptide.
Introduction

Alzheimer’s disease (AD) is a neurodegenerative disease characterized by the deposition of β-amyloid plaques and Neurofibrillary tangles (NFTs) of the hyperphosphorylated tau. It is the leading cause of dementia worldwide and results in the cognitive decline of multiple domains. As of 2020, Alzheimer’s disease affected an estimated 5.8 million people in the United States and is projected to afflict nearly 14 million by the year 2060. Increased levels of Aβ-42 plaques have been found in the early stages of AD pathogenesis, resulting in neurotoxicity, the induction oxidative stress, and eventual cell death of neurons.  

Increased extracellular production of Aβ-42 is typically attributed to mutations observed in the amyloid precursor protein (APP) by β-secretase and γ-secretase subsequently. In a brain affected by AD, an inability to clear and regulate surplus Aβ-42 can be discerned, leading to acute dementia in mouse models, cell cultures, and synthetic simulations. Senile extracellular amyloid plaques are observed in many late onset AD cases, suggesting that Aβ-42 aggregation is in fact linked to the pathogenesis of the disease. Therefore, this provides a potential opportunity for the therapeutic treatment of Alzheimer’s disease, via inhibition of the AB-42 peptide.

Polyphenols are a diverse class of small molecule natural products that are biosynthesized by various plants. As demonstrated in previous studies, polyphenols serve as an effective antioxidant agent, serving as a line of defense against free radical damage and a variety of pathogens. Moreover, these compounds have been previously demonstrated to inhibit the aggregation of several amyloidogenic peptides including amyloid beta 42 (Aβ-42), alpha-synuclein, and the islet amyloid polypeptide (IAPP), which have a proclivity to aggregate and have been linked to human diseases including Alzheimer’s disease, Parkinson’s disease, and Type II diabetes, respectively.

Amyloidogenicity of such peptides is greatly influenced by the amino acid residue sequence and an abundance of pi-stacking interactions and hydrophobic interactions among aromatic and nonpolar side chains has been shown to be a significant contributor to amyloidogenic behavior. The amyloid fibril formation inhibition abilities of polyphenols were thought to be the result of their antioxidant capabilities. The most effective polyphenols for inhibiting fibril formation include catechin, curcumin, epigallocatechin gallate, and reverstarol.
Aβ aggregation is reportedly reduced by polyphenols through the disruption of key noncovalent interactions that drive amyloid formation. Compounds such as epigallocatechin gallate (EGCG) have been shown to directly inhibit aggregation by binding to unfolded polypeptides of the protein, stopping the formation of toxic beta sheet fibrils. However, due to the disorderly nature of amyloid fibrils, very few crystal structures have been arranged to date, and the exact structural basis for this mechanism of action of polyphenols such as EGCG remains unknown. Therefore, the use of computational tools to probe molecular-level details of polyphenolic inhibition of amyloidogenic peptides is a crucial path of drug discovery research.

Here, we screened a targeted library of natural product polyphenols for binding affinity to monomers and oligomers of the Aβ-42 peptide by molecular docking simulations. These results were used to identify the most effective potential inhibitors of Aβ-42 amyloid formation. These compounds were also screened for polyphenol-induced oligomer destabilization through molecular dynamics simulations. GROMACS, a molecular dynamics software, was used to perform these simulations and analyze Aβ-42 protein stability over time against certain polyphenols that had higher binding affinities to the peptide. We hypothesize that epigallocatechin gallate (EGCG) would exhibit the highest binding affinity to amyloid fibrils and potentially cause the greatest amyloid oligomer destabilization.

Results

TANGO
The Aβ-42 peptide, shown in Figure 1, had two major regions of predicted aggregation: residues 17-21 (LVFFA) and 29-41 (GAIIGLMVGGVVI) whereas the 10-35 residues, shown in figure 1b, only had one amino acid with high predictions: 17-21. This segment, which appears on both proteins, consists solely of amino acids with hydrophobic side chains, which may contribute to its high amyloidogenicity. The prediction values for the 10-35 residues were slightly lower than that of the Aβ-42 protein. In the 10 to 35 sequence, the values for the LVFFA segment are 77.514; however, in the Aβ-42 sequence the values for that same sequence are 81.431. This suggests that the LVFAA region might be particularly involved in aggregation. Hydrophobic interactions between the side chains may contribute to the amyloidogenicity of these protein sequences. Overall, residues 10-35 exhibited high predicted aggregation values. In comparison, the Aβ-42 protein had two amino acid sequences with high predicted aggregation values.

Docking
Three targets were chosen for docking simulations: the Aβ-42 oligomer (PDB:2MXU), monomer (PDB:1Z0Q) and a close-in sequence of residues 10-35, which were identified through TANGO as being the principally amyloidogenic region on the peptide. Solution state NMR structures of each target were obtained from the Protein Data Bank.\(^{15}\)

Docking results from AutoDock Vina are listed in Table 1, and indicate that tannic acid, naringin, rutin, curcumin, and epigallocatechin gallate (EGCG), in that order, demonstrate the most promising binding affinities to the Aβ-42 peptide. Table 1 shows the predicted binding affinities of each compound against Aβ-42 oligomers, monomers, and the amyloidogenic segment based on molecular docking. Tannic acid exhibited the greatest predicted binding affinity to the amyloid beta oligomer (-12 kcal/mol), followed by naringin and rutin (-7.7 kcal/mol and -7.5 kcal/mol, respectively), which are both glycosylated polyphenols. This trend is possibly a product of the greater availability of hydrogen bond donors and acceptors of tannic acid, which might also explain the greater calculated binding affinities of glycosylated polyphenols rutin and naringin over their non-glycosylated counterparts.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Oligomer (kcal/mol)</th>
<th>Monomer (kcal/mol)</th>
<th>Residues 10 to 35 (kcal/mol)</th>
<th>Average (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tannic Acid</td>
<td><img src="image" alt="Tannic Acid" /></td>
<td>-12</td>
<td>-7.2</td>
<td>-4.3</td>
</tr>
<tr>
<td>Naringin</td>
<td><img src="image" alt="Naringin" /></td>
<td>-7.7</td>
<td>-6</td>
<td>-6.3</td>
</tr>
<tr>
<td>Rutin</td>
<td><img src="image" alt="Rutin" /></td>
<td>-7.5</td>
<td>-6.5</td>
<td>-4.8</td>
</tr>
<tr>
<td>Genistein</td>
<td><img src="image" alt="Genistein" /></td>
<td>-6.8</td>
<td>-6.5</td>
<td>-4.9</td>
</tr>
<tr>
<td>EGCG</td>
<td><img src="image" alt="EGCG" /></td>
<td>-5.9</td>
<td>-6.4</td>
<td>-5.6</td>
</tr>
<tr>
<td>Compound</td>
<td>(-\infty) to (-10)</td>
<td>(-10.999) to (-9.999)</td>
<td>(-9.999) to (-8.999)</td>
<td>(-8.999) to (-7.999)</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Apigenin</td>
<td>-6.8</td>
<td>-5.9</td>
<td>-5.1</td>
<td>-5.933</td>
</tr>
<tr>
<td>Quercetin</td>
<td>-6.8</td>
<td>-5.7</td>
<td>-5.1</td>
<td>-5.867</td>
</tr>
<tr>
<td>Kaempferol</td>
<td>-6.7</td>
<td>-5.0</td>
<td>-5.3</td>
<td>-5.667</td>
</tr>
<tr>
<td>Chrysin</td>
<td>-6.1</td>
<td>-5.3</td>
<td>-4.7</td>
<td>-5.367</td>
</tr>
<tr>
<td>Curcumin</td>
<td>-6.9</td>
<td>-5.1</td>
<td>-3.6</td>
<td>-5.2</td>
</tr>
<tr>
<td>Resveratrol</td>
<td>-4.7</td>
<td>-4.4</td>
<td>-5.3</td>
<td>-4.8</td>
</tr>
<tr>
<td>Gallic Acid</td>
<td>-4.8</td>
<td>-4.1</td>
<td>-3.7</td>
<td>-4.2</td>
</tr>
<tr>
<td>Caffeic Acid</td>
<td>-5.1</td>
<td>-4.3</td>
<td>-3.2</td>
<td>-4.2</td>
</tr>
</tbody>
</table>

**Key:**
- \(\leq -10\)
- \(-9.999\) to \(-9.0\)
- \(-8.999\) to \(-8.0\)
- \(-7.999\) to \(-7.5\)
- \(-7.499\) to \(-7.0\)
- \(-6.999\) to \(-6.5\)
- \(-6.499\) to \(-6\)
Table 1: Predicted binding affinities (ΔG) of screened polyphenols against Aβ-42 via molecular docking in three trials. Every square is color-coded by increasing binding affinity.

Molecular Dynamics
Molecular dynamics (MD) simulations were performed using GROningen MAchine for Chemical Simulations (GROMACS) to find which polyphenols best destabilize the Aβ-42 protein. Root-mean-square deviation (RMSD) calculations were conducted to find similarity between various atomic positions of the polyphenols. RMSD shows how much the atoms of Aβ-42 change over time, which can explain whether the polyphenols destabilize or stabilize the protein. Over time, gallic acid, naringin, EGCG, curcumin, and apigenin cause fluctuations in the RMSD of Aβ-42. Fluctuation in RMSD indicate destabilization of the Aβ-42 oligomer. The polyphenols with the highest average RMSD values, listed in Table 1, were found to be apigenin at 4.62 Å, curcumin at 4.182 Å, and rutin at 4.2615 Å, while unliganded Aβ42 was observed to have a much lower RMSD at 0.881. High average RMSDs indicate more deviation of the atomic positions in the AB-42 protein, which suggests relative destabilization of the aggregated state protein.

From the protein sequence based RMSDs, depicted in Figure 6, the AB-42 peptide seems to have the most fluctuations in RMSD when simulated with EGCG, genistein, tannic acid, and caffeic acid. These polyphenols have high spikes in RMSD at Lys102, Val221, and Glu321, meaning that these amino acids are likely being targeted specifically by the polyphenols. At Lys102, Genestein with AB-42 has the highest RMSD value of 0.409 Å, while at Val221 Tannic Acid with AB-42 has a value of 0.512 Å.

Discussion
In the search for naturally-occurring polyphenols which demonstrate high potential to inhibit the activity of the Aβ-42 peptide, we found via computational analysis that some polyphenols possess this ability. With the docking software AutoDock Vina, it was observed that some polyphenols, such as tannic acid, naringin, rutin, and epigallocatechin gallate, exhibited high binding affinities to Aβ-42. This suggests that these compounds may be more effective in binding to the amyloid peptide. Contrary to our initial hypothesis, EGCG was not found to have the highest binding affinity to either the Aβ-42 monomer or the oligomer, and was outperformed by tannic acid and by two glycosylated polyphenols, naringin and rutin.
These results suggest that glycosylation of polyphenols may improve binding affinity to Aβ-42 and other amyloidogenic peptides.

As expected, the predictions obtained from TANGO showed that residues 10-35 contained a number of the regions most prone to amyloidogenicity, and in particular residues 10-15 were predicted to have the greatest amyloidogenic propensity. Several of the polyphenols screened in this study were found to dock to this particular region of the 10-35 region. Namely, the glycosylated polyphenol naringin, which exhibited the highest binding affinity to the 10-35 region of the amyloid beta subsequence, and this high binding affinity (-6.3 kcal/mol) is attributed to hydrogen bonding interactions between the sugar fragment of the ligand and the tyrosine, glutamate, and histidine residues of the YEVHH subsequence. This further reinforces the significance of the glycoside in the performance of certain polyphenols in binding to the amyloid protein.

Time-dependent RMSDs from the molecular dynamics simulations, depicted in Figure 5, suggest that gallic acid, naringin, EGCG, curcumin, and apigenin are most effective in attenuating the aggregation of Aβ-42, evidenced by an increase in RMSD fluxions, which suggests destabilization of the aggregated state. The average RMSDs, depicted in Figure 6 from the simulation show that apigenin, curcumin, and rutin with Aβ-42 have the highest RMSD values per residue. These polyphenols result in the most deviation in atomic positions within the Aβ-42, thus rendering it the least stable and making them seem the most effective as Aβ-42 peptide inhibitors. As expected, a simulation of native Aβ-42 in water has the lowest average RMSD, suggesting a comparatively stabilized aggregate state and thus less inhibition of the Aβ-42 peptide. While further study must be conducted to understand how the interaction between polyphenols and the amino acids of Aβ-42, this initial study indicates that Lys102, Val221, and Glu321 play a role in hindering the aggregation of the protein. However, further mechanistic studies would need to be conducted in order to determine whether the polyphenols successfully destabilize AB-42 over an extended period versus our 1 nanosecond simulation. From our results, we find that apigenin and EGCG have the most consistent results in destabilizing Aβ-42 from the analyses of molecular dynamics simulations. Future work in evaluating these polyphenols in vitro will provide insight into whether these trends determined computationally do indeed play a role in inhibition of amyloidogenesis of Aβ-42.

Methods

Density Functional Theory (DFT)
Thermodynamically minimized structures of each polyphenol was found via an initial Molecular Mechanics (MM) pre-optimization using the Universal Force Field (UFF) at 10,000 steps. Subsequently, rigorous quantum mechanical structural optimizations were performed using ORCA, an \textit{ab initio} quantum mechanical software package.\textsuperscript{13} Density functional theory optimizations were completed with the B3LYP functional and def2-SVP basis set, and an implicit conductor-like solvation model (CPCM) simulating the dielectric and permittivity of water. Density functional theory calculations were performed on a Dell PowerEdge 710 server with a 24 core Intel Xeon X5660 processor @ 2.80GHz and 32GB RAM.

\textbf{TANGO}

In order to determine the segments of the amyloid peptide that contribute to the aggregation propensity of Aβ-42, we used TANGO, a computer software that is used to predict the aggregation propensity of regions in polypeptide chains.\textsuperscript{14} The algorithm that TANGO utilizes can predict aggregation in four possible structural states: Alpha Helix, Beta Sheet Aggregation, Alpha Helix Aggregation, and Beta Turn. Beta sheet potential aggregation values were examined for Aβ-42.

TANGO calculations were conducted at a pH of 5.37, while the temperature was 310.05 Kelvin. These were found to be the standard physiological conditions of the human brain. The ionic strength of the protein was calculated using the Calistry Ionic Strength Calculator.

\textbf{Molecular Docking}

The nature and thermodynamics of protein-ligand interactions were determined via molecular docking. As the binding sites of polyphenols to the Aβ-42 protein are currently unknown, docking software AutoDock Vina\textsuperscript{15} was used to blind-dock the selected polyphenols to the monomer (PDB: 1Z0Q), oligomer (PDB: 2MXU), and 10 to 35 (PDB: 1HZ3) peptides of the Aβ-42 protein. Blind docking allowed us to determine this binding site, which we found was consistent between different polyphenols.\textsuperscript{11, 16}

\textbf{Molecular Dynamics}

Molecular dynamics was simulated using GROMACS\textsuperscript{17} in a CHARMM36 force field with a 3-point dodecahedron water model. Prior to simulation, ligand parameters were prepared using AC-PYPE Server. Energy minimization was performed in order to remove clashes between the solvation model and the protein-ligand system. Equilibration was then performed for temperature and pressure coupling. The last MD simulation run was performed in 50000 steps using the LINCS algorithm.
Molecular dynamics simulations were performed on a Dell Poweredge 710 server with a 24 core Intel Xeon X5660 processors @ 2.80GHz and 32GB RAM. Analysis of the resulting trajectory files and subsequent data were performed using Visual Molecular Dynamics (VMD) and MDWeb Server.18

Acknowledgments

We would like to thank our advisor, Edward Njoo, for his guidance and support of our research. We would also like to acknowledge the Chemistry Department at the Aspiring Scholars Directed Research Program (ASDRP) and its community and corporate sponsors and supporters for providing the facilities and funding for us to conduct our research. The authors also gratefully acknowledge Prof. Robert Downing and the Computer Science & Engineering Department at ASDRP for graciously providing us with computational resources towards our research.

References


Figure Captions

Figure 1: A chart of the predicted aggregation propensity values for the Aβ42 Amino Acid Sequence in Beta Sheet structural state, with residues 10-15 (YEVHH) exhibiting the greatest amyloidogenic propensity.

Figure 2: a) The overlay of the oligomer (brown) and 10 to 35 peptide sequence (blue) with the docking results of the oligomer. In this, all of the polyphenols dock in the same part of both sequences. b) shows the overlay of the oligomer and 10 to 35 sequence with the docking results of the amyloid-forming segment. All the polyphenols dock against different parts of the 10 to 35 sequence, but generally dock in the same area with the oligomer. c) The docked structure of the polyphenol glycoside naringin bound to the 10-35 amyloidogenic subsequence, with key hydrogen bonding interactions predicted between the sugar fragment and the tyrosine, glutamate, and histidine residues in the 10-15 region.

Figure 3: Root mean square deviation (RMSD) values for the AB42 protein from each polyphenolic complex were derived from GROMACS and graphed as a function of time. RMSD as a function of time provides insight into whether destabilization of the AB42 oligomers occurs.

Figure 4: Root mean square deviation (RMSD) values over time from each simulation were averaged and graphed. An average RMSD shows the overall deviation of atoms of the protein within the 1ns time frame, which can help understand stability of the interaction between the polyphenols and the AB42 protein.

Figure 5: The Aβ42 sequence was graphed in relation to the average RMSD per amino acid residue. The relationship between the sequence and RMSD can help examine which amino acids are possibly affected by the polyphenols simulated.
Beta Sheet Aggregation in the Amyloid Beta 42 Sequence
Figure 4

Average RMSD of Polyphenols with AB-42

Polyphenols

RMSD

Water  EGC   Rutin  Naringin  Gallic Acid  Chrysin  Tannic Acid  Resveratrol  Quercetin  Kaempferol  Genistein  Curcumin  Caffeic Acid  Apigenin