



# HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2018 April 21.

Published in final edited form as:

ACS Chem Biol. 2017 April 21; 12(4): 1113–1120. doi:10.1021/acscchembio.7b00045.

## Ion Mobility-Mass Spectrometry Reveals a Dipeptide That Acts as a Molecular Chaperone for Amyloid $\beta$

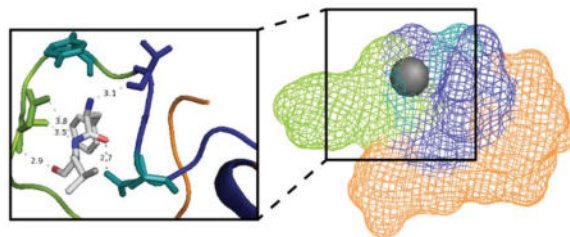
Molly T. Soper-Hopper, Joseph D. Eschweiler, and Brandon T. Ruotolo\*

Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, United States

### Abstract

Previously, we discovered and structurally characterized a complex between amyloid  $\beta$  1–40 and the neuropeptide leucine enkephalin. This work identified leucine enkephalin as a potentially useful starting point for the discovery of peptide-related biotherapeutics for Alzheimer's disease. In order to better understand such complexes that are formed *in vitro*, we describe here the analysis of a series of site-directed amino acid substitution variants of both peptides, covering the leucine enkephalin sequence in its entirety and a large number of selected residues of amyloid  $\beta$  1–40 (residues: D1, E3, F4, R5, H6, Y10, E11, H13, H14, Q15, K16, E22, K28, and V40). Ion mobility-mass spectrometry measurements and molecular dynamics simulations reveal that the hydrophobic C-terminus of leucine enkephalin (Phe-Leu, FL) is crucial for the formation of peptide complexes. As such, we explore here the interaction of the dipeptide FL with both wildtype and variant forms of amyloid  $\beta$  in order to structurally characterize the complexes formed. We find that FL binds preferentially to amyloid  $\beta$  oligomers, and attaches to amyloid  $\beta$  within the region between its N-terminus and its hydrophobic core, most specifically at residues Y10 and Q15. We further show that FL is able to prevent fibril formation.

### Graphical abstract



\*Corresponding Author. bruotolo@umich.edu.

#### ASSOCIATED CONTENT

Supporting Information. Detailed annotations of 3D IM-MS data, 2D MS data, expected and M/Z for each complex, CD data, as well as details of leucine enkephalin interactions with A $\beta$  variants can be found in the supporting information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### Author Contributions

Experiments were conceived of by MTS-H and BTR. IM-MS A $\beta$  experiments, docking studies, MD simulations, detailed modeling analysis, K<sub>D</sub> measurements, and TEM analysis were performed by MTS-H. JDE was responsible for development of the hierarchical clustering method. The manuscript was written by MTS-H and BTR. All authors have given approval to the final version of the manuscript.

## Introduction

The utility of ion mobility-mass spectrometry (IM-MS) as a tool for drug discovery and development, especially in the context of protein misfolding diseases such as Alzheimer's Disease (AD) and Type II Diabetes, has recently emerged.<sup>1-10</sup> While many hypotheses surrounding the causation of AD are proposed, a prominent theory is that the uncontrolled aggregation of Amyloid  $\beta$  ( $A\beta$ ) peptides, which form small toxic oligomers (dimer to dodecamer) leads ultimately to neuronal cell death.<sup>11-15</sup> As  $A\beta$  aggregation allows for many oligomeric states simultaneously, studying  $A\beta$  oligomers for the purposes of drug discovery is difficult with classical biophysical and structural biology techniques.<sup>16</sup> IM-MS allows for the use of low concentrations, small sample volumes, and separates protein oligomers based on mass and collision cross section (CCS)<sup>17-19</sup> making it an ideal technology for the analysis of  $A\beta$ -small molecule complexes.

We previously reported a set of experiments that identified  $A\beta$ -Neuropeptide complexes, with the ultimate goal of identifying a peptide or peptide fragment which might inhibit  $A\beta$  fibril formation and provide a scaffold for future potential biotherapeutics<sup>20</sup>. In these datasets, we found evidence of interactions between  $A\beta_{1-40}$  and both leucine enkephalin and galanin neuropeptides. We characterized the former in detail, providing CCSs, peptide-ligand dissociation constant ( $K_D$ ) values, and molecular dynamics (MD) simulations. In addition, we studied the ability of leucine enkephalin to inhibit  $A\beta$  fibril formation using both TEM and gel electrophoresis. Complexes containing  $A\beta$  and leucine enkephalin were detected in stoichiometric ratios ranging from 1:1 to 1:3, and from 2:1 to 2:3  $A\beta$ :leucine enkephalin. At increased concentrations of leucine enkephalin, the data indicated that attachment to  $A\beta$  dimers is favored over binding to the monomeric form.  $K_D$  values were measured in the  $\mu\text{M}$  range for these complexes, and MD simulations led the determination of the likely binding location of leucine enkephalin along the N-terminus of the  $A\beta$  sequence, similar to other natural products.<sup>5</sup> Transmission Electron Microscopy (TEM) and gel data reported decreased  $A\beta$  fibrillation and an increased amount of lower molecular weight  $A\beta$  oligomers when incubated in excess with leucine enkephalin, compared with control.

Here we report IM-MS measurements of complexes formed between  $A\beta$  and the site directed amino acid substitutions of the entire sequence of leucine enkephalin, which indicate that the hydrophobic C-terminal residues FL are critical for the formation of  $A\beta$ :leucine enkephalin interactions. Furthermore, we study the dipeptide FL and its complexes formed with  $A\beta$  sequence variants. For 1:1  $A\beta$ :FL complexes, we identify residues Y10 and Q15 as possessing key interactions with FL, necessary for strong complex formation. This observation matches well with the predicted location of FL binding from MD simulations of  $A\beta$ :leucine enkephalin.<sup>20</sup> MD simulations of  $A\beta$ :FL were conducted and models filtered using IM-MS data. From these, we predict that FL interacts with  $A\beta$  primarily through Y10, and that these interactions occur mostly through a series of specific interactions with the C-terminus of FL. Beyond monomeric  $A\beta$  complexes, we also study higher-order  $A\beta$  complexes (e.g. dimers and trimers) and find that FL interacts with  $A\beta$  dimers preferentially. Finally, we explored the downstream effects of FL binding on  $A\beta$  fibrillation through TEM analysis, finding strong evidence of the inhibition of fibrilization. We conclude by discussing the potential impact of FL as a scaffold for the future

development of AD therapies in the context of previously described small peptide-based amyloid inhibitors.

## Results and Discussion

To investigate the specific interactions responsible for A $\beta$ :leucine enkephalin binding, we first undertook a series of IM-MS experiments using a comprehensive set of site-directed amino acid substituted forms of leucine enkephalin (sequence: YGGFL). Each of these five leucine enkephalin variants was first evaluated using MS intensity values under controlled conditions as a quantitative measure of A $\beta$ <sub>1-40</sub> complex formation. (Figure 1A) Alanine substitution of the first three residues in the leucine enkephalin sequence (tyrosine and both glycine residues) yield complex ion signals similar to control experiments under our experimental conditions. However, when the two hydrophobic C-terminal residues F and L were substituted, neither of these variants were observed to generate signals associated with complex formation, indicating their importance in A $\beta$  binding overall.

Using the information provided in Figure 1A as well as CCS information from our previous study,<sup>20</sup> we filtered an ensemble of putative A $\beta$ :leucine enkephalin structures generated by MD in order to enrich for models where A $\beta$  constituents are within 4 Å of either the F or L residues of the bound leucine enkephalin peptide. This refined population of structures was used to calculate standardized probability values for each residue of A $\beta$ <sub>1-40</sub> in terms of their F or L contacts. Our analysis highlights several regions of importance, which are shown in Figures 1B and C, where a standard score greater than 1 indicates an increased significance of the contact observed between the indicated A $\beta$  residue and either the F or L within the leucine enkephalin sequence, respectively. The A $\beta$ <sub>1-40</sub> residues which most frequently interact with F within leucine enkephalin are Y10, H14, Q15, K16, E22 and D23, and those interacting most frequently with the L residue within leucine enkephalin are R5, Y10, E11, H14, Q16, K16, and K28. Those interacting significantly with both residues are Y10, H14, Q15, K16. This analysis agrees with our previous models of A $\beta$ :leucine enkephalin complexes, from which we concluded that leucine enkephalin binding likely occurs in the N-terminal hydrophobic core region of A $\beta$ , exhibiting similar behavior to other natural products that act as A $\beta$  amyloid inhibitors.<sup>5, 20</sup> As such, R5, Y10, E11, H14, Q15, K16, E22, and K28 alanine substitution variants of A $\beta$  were targeted for downstream experiments. Of the remaining alanine variants, D1A and V40A, the A $\beta$  termini, were used as negative controls. Furthermore, H6A and H13A were selected as additional controls for the H14A while E3 served as an additional control for the E11A and E22A substitutions. Each of these selections were designed to ensure that any detected change in complex  $K_D$  could be interpreted as related to the disruption of a specific binding motif observed by MD, rather than due to subtle changes in A $\beta$  structure. In addition, the F4A A $\beta$  variant has been identified as a potential binding site in the context of other AD therapeutics<sup>6</sup>, and so was also used here as a control.

As our data strongly indicates that it is primarily the hydrophobic C-terminus of leucine enkephalin that is responsible for its binding to A $\beta$ , we moved to focus our experimental efforts on the dipeptide FL. IM-MS data acquired from solutions containing 20  $\mu$ M A $\beta$ <sub>1-40</sub> and FL ranging from 10 – 80  $\mu$ M, reveals the formation of A $\beta$ <sub>1-40</sub>:FL complexes having

stoichiometries of 1:1 – 1:4, 2:1 – 2:5, 3:1 – 3:4, and 4:1 (Figure 2A). Detailed arrival time distributions can be found in the supplemental material. Companion MD simulations of  $A\beta_{1-40}$ :FL 1:1 complex yields evidence of significant interactions between the dipeptide and R5, Y10, K16, D23, S26, N27, and K28 within the  $A\beta$  sequence (Figure 2B). This result largely agrees with our analysis of  $A\beta_{1-40}$ : leucine enkephalin models discussed previously (Figure 1C), which predicted R5, Y10, K16, D23, and K28 as likely FL binding sites. While  $A\beta_{1-40}$  is found in greater quantities *in vivo*,  $A\beta_{1-42}$  is theorized to be the more pathogenic variant and aggregates more rapidly, making it more difficult to study *in vitro*. IM-MS data shown in Figure 2C indicates that  $A\beta_{1-42}$  also binds FL, exhibiting  $K_D$  and complex stoichiometry values similar to that of  $A\beta_{1-40}$ . Low energy models generated for  $A\beta_{1-42}$ :FL 1:1 complex exhibit a strong preference toward binding within the hydrophobic core of  $A\beta_{1-42}$ , specifically with residues E11, V12, H13, Q15, and K16 (Figure 2D), an observation that is generally retained from our  $A\beta_{1-40}$  MD and IM-MS data as well.

In addition, the measured  $K_D$  values for  $A\beta_{1-40}$  binding with leucine enkephalin (supplemental) and  $A\beta_{1-40}$  binding with FL are within an order of magnitude, a difference we consider to be negligible due to the variability of our  $K_D$  measurements. This observation further confirms the importance of phenylalanine and leucine in  $A\beta$  binding. If other residues were also critical for leucine enkephalin binding, we would expect that the additional inter-molecular interactions would produce an aggregate  $K_D$  value for FL binding significantly greater than the  $\sim 60 \mu\text{M}$  observed here.

To more-clearly elucidate the binding location of FL within the  $A\beta_{1-40}$  sequence,  $K_D$  values were extracted from the MS data acquired for all alanine substituted  $A\beta$  variants listed above, in addition to control variants: D1A, E3A, F4A, H6A, H13A, and V40A (Figure 3B). All control residues represent regions of the  $A\beta$  sequence that IM-MS filtered computational models did not highlight as critical for leucine enkephalin binding. Unavoidable signal variations in our data resulting from the complex mixture of ions analyzed in our experiments bring about relatively large standard deviations for some of our  $K_D$  measurements. Here, we conservatively assign significance to values that differ by greater than one standard deviation from the mean.

Of the fourteen alanine variants studied, only three produced  $K_D$  values significantly greater than those recorded for the WT  $A\beta_{1-40}$ :FL 1:1 complex ( $K_D = 58.1 \pm 25.9 \mu\text{M}$ ). These variants include Y10A, Q15A, and E22A ( $K_D = 135 \pm 71.3 \mu\text{M}$ ,  $126 \pm 86.6 \mu\text{M}$ ,  $112 \pm 53 \mu\text{M}$  respectively). Neighboring variants E11A, H14A and K16A do not deviate significantly from WT. In fact, the mutations R5A, K16A, and K28A all fall within error of the WT value, demonstrating the lysine and arginine residues of  $A\beta$  have no impact on FL binding. Of the control cases, D1A, E3A and H13A behaved similarly to the WT as expected. The 1:1 complex of F4A:FL was only observed twice in seven separate samples, having a measured  $K_D$  value of  $117 \mu\text{M}$  (data was insufficient to project experimental error). In contrast to these results, two variants, H6A and V40A, produced significantly reduced  $K_D$  values relative to WT,  $27.5 \pm 14.8 \mu\text{M}$  and  $17.2 \pm 8.5 \mu\text{M}$  respectively. (Figure 3B). Taken together, these data indicate that residues Y10, Q15, and E22 likely provide key molecular contacts that stabilize the  $A\beta_{1-40}$ :FL 1:1 complex, while H6 and V40 provide a destabilizing

or screening effect that, when removed, improves the strength of the resulting complex significantly.

Multiple FL binding events to monomeric A $\beta$  (1:2 and 1:3 A $\beta$ :FL) yielded weaker  $K_D$  values, generally, than those associated with the 1:1 interactions observed for all variants where 1:2 or 1:3 complexes were detected. Additionally, all  $K_D$  measurements for 1:2 complexes were within error of each other, providing no clear evidence of a preferred FL binding site in the higher-order complexes observed here.  $K_D$  values for recorded for binding between alanine variants and leucine enkephalin, as well as the values acquired for higher order A $\beta$ : leucine enkephalin complexes, are available in Supporting Information.

In order to validate that the structures of the variant forms of A $\beta_{1-40}$  are not significantly altered from that of the WT sequence, CD spectroscopy data was acquired for all alanine variants and WT A $\beta$  peptides (Supporting Information). While Y10A and Q15A peptides have CD signatures similar to that of A $\beta_{1-40}$ , the CD spectrum acquired for E22A exhibits a pronounced  $\beta$ -sheet structure similar to that of A $\beta_{1-42}$ . The known familial mutation E22G (Arctic mutation) of A $\beta_{1-40}$  has shown an increased propensity for aggregation and behavior mimicking A $\beta_{1-42}$ .<sup>21</sup> The alanine variant E22A differs from this mutation by only a methyl group and was suspected to mimic E22G structurally, making the CD results unsurprising. As such, CD data supports an analysis of our IM-MS  $K_D$  data interpreted on the basis of only the altered molecular contacts within the FL:A $\beta$  binding surface for Y10A and Q15A, but strongly suggests that E22A  $K_D$  data is a product of a significantly altered peptide conformation.

Using all the data described above, we filtered the output of MD simulations to produce a range of models that describe the molecular details of the A $\beta_{1-40}$ :FL 1:1 complex observed in our experiments. Our workflow begins by filtering all models that have CCS values outside of our experimental error ( $652.6 \pm 22.1 \text{ \AA}^2$  for the A $\beta_{1-40}$ :FL complex). Next, we remove all models that lack clear interactions (contact within 4  $\text{\AA}$ ) between FL and Y10 and Q15, which our experimental data suggest are key elements for the efficient formation of a stable A $\beta_{1-40}$ :FL complex.

Hierarchical clustering of this filtered model pool yields five structural families. Figure 4A shows the output of our clustering algorithm, along with low-resolution averaged peptide structures that depict an averaged location for FL binding on the A $\beta$  surface. In most cases, FL is found between the N-terminus and the hydrophobic core while the c-terminus is further away. A detailed, manual analysis of these five structural families reveals common themes in the molecular contacts created between FL and the Q15 residue within the A $\beta$  sequence. Two examples of these details are shown in Figure 4B and 4C, where we highlight bonding of FL and A $\beta$  residue Q15.

An interaction distance of 2.8  $\text{\AA}$  between the carbonyl oxygen of L within FL, and an  $\epsilon$  nitrogen (atom name NE2 in pdb format) of Q15 is observed in the lowest energy structure, matching our experimental restraints (data not shown). The ligand configuration is stabilized by R5, H6, and D7 with hydrogen bonding distances of 3.1  $\text{\AA}$  and 3.2  $\text{\AA}$  between R5 and D7 with the carbonyl oxygen of F, and a 3.8  $\text{\AA}$  interaction distance between the F nitrogen and

H6. In this particular model, Y10 is too far away from the carbonyl oxygen on L to hydrogen bond. In another low energy structure, FL is wrapped tightly within the A $\beta$  structure (Figure 4B). Here, hydrogen bonding occurs between the NE2 position of Q15 and the carbonyl oxygen on F (2.7 Å). While there are no contacts less than 4 Å between FL and Y10 in this model, the ring structure surrounds the FL pocket. The interaction is stabilized through an interaction having a 3.1 Å distance between carboxylic acid oxygen on V12 and the backbone nitrogen on F, in addition to 3 apparent hydrogen bonds between D7 and L.

Despite the Y10 not being close enough for hydrogen bonding in the two models discussed thus far, another low energy structure from a separate cluster does demonstrate such interactions (Figure 4C). Here, Y10 displays evidence of a hydrogen bond between the backbone nitrogen of F (distance of 3.0 Å). In addition, Q15 is within 3.0 Å of the oxygen on Q15 and the alpha carbon of L. The interaction is further stabilized through hydrogen bonding between an oxygen of L and N of L17 (2.8 Å), as well as a  $\delta$  nitrogen (atom name ND1 in pdb format) of H14 and the oxygen of F (3.1 Å). We note that all FL hydrogen bonding captured within our models occur along the dipeptide backbone. As such hydrogen bonds are relatively weak interactions, we believe our MD data only captures a subset of all potential interactions. Side-chain specific interactions are thus also possible, and perhaps likely, in these complexes, but remain undocumented in the three lowest energy models detailed here.

Discussed thus far, all K<sub>D</sub> and computational analysis has focused on the analysis of FL binding to monomeric A $\beta$ . IM-MS data circled in Figure 2A however, clearly shows the interaction of FL with small soluble oligomers of A $\beta$ <sub>1-40</sub>. We calculated the percentage of oligomers found in both unbound and bound states. Of the total monomeric population detected, 34 ± 12 % is found bound to at least 1 FL. In contrast, 69 ± 4.5% of the dimer population is bound to FL and 50 ± 20% of the trimers are bound. This suggests preference for binding to the dimer and trimer populations over the monomer in WT A $\beta$ <sub>1-40</sub>, despite the monomer being the dominant species in our samples at time of measurement. Binding with the tetramer was not detected in enough replicates to calculate a precise value for the average bound fraction.

In order to evaluate the ability of FL to alter A $\beta$  fibril formation, TEM images were acquired for A $\beta$ <sub>1-40</sub> samples dosed with varying amounts of FL. The features observed in the TEM data shown in Figure 5 have a marked dependence upon FL concentration, with long mature fibril structures observed for A $\beta$  samples where no FL is added, transitioning to truncated fibril structures in samples where equimolar FL is added, and finally a complete abrogation of A $\beta$  fibril formation when excess FL is present. Truncated fibrils observed by TEM upon equimolar FL addition to A $\beta$  samples are less dense when compared to control samples, further indicating that mature fibril formation is inhibited under these conditions. The weakly-stained forms observed under conditions of FL excess appear universally amorphous, indicating that while A $\beta$  aggregation still occurs, it likely does so through an altered route that does not share a morphology connected with the production of toxic species, which have been observed in previous studies and similarly interpreted.<sup>22</sup>

## Conclusions

Through a comprehensive set of IM-MS experiments and MD simulations for an exhaustive series of amino acid variants, we have identified the two C-terminal residues of leucine enkephalin, FL, as critical for its previously-reported interaction with A $\beta$ . A subsequent IM-MS analysis of A $\beta$  1:1 complex formed with the isolated FL dipeptide from leucine enkephalin revealed both equivalent binding strengths to the full length construct. Through detailed IM-MS and CD spectroscopy covering 14 alanine variants of A $\beta$ <sub>1-40</sub>, carefully selected through an analysis of MD simulation output, we have narrowed the potential sites of FL contact on the A $\beta$ <sub>1-40</sub> surface to the Y10 and Q15 residues. Models extensively filtered through a range of structural biology data indicate that FL primarily interacts with Q15 through hydrogen bonding to the FL backbone, and that the dipeptide is positioned between the N-terminus and the hydro-phobic core in a manner similar to previous A $\beta$  binders of therapeutic note.<sup>5, 11, 20</sup> With regard to small soluble oligomers we note improved attachment over the monomers. Critically, TEM data shows that FL inhibits A $\beta$ <sub>1-40</sub> fibril formation, which further points to its potential as a foundation for future AD therapeutics. Additional computational modeling and further experiments are required to give details concerning the binding of FL to A $\beta$  at the dimer level and above, although we speculate it to access a similar mode as found for the monomer. To move forward, we look towards the chemical modification FL in the creation of a therapeutic library for IM-MS screening. Modifications should reduce potential cellular toxicity and increase binding affinity. Additionally, amidated and acetylated FL variants are of interest, as such modifications in the A $\beta$ <sub>1-42</sub> C-terminal fragment VVIA have been shown to decrease the toxicity of resultant A $\beta$  oligomers.<sup>23</sup> It is worth noting that the binding location of FL is somewhat rare among previously described peptidic A $\beta$  binders, with most accessing the C-terminus and hydrophobic core rather than the N-terminal region of the peptide.<sup>23</sup> It is likely that further IM-MS screens, in combination with FL optimization, will lead to a more detailed mechanism of A $\beta$  fibril formation inhibition in general.

## Methods

Reagents were purchased from commercial sources and used without further purification unless otherwise specified. A $\beta$ <sub>1-40</sub> was purchased from Anaspec. Alanine mutations were custom synthesized and received at a purity of at least 95% (GenScript). A $\beta$  variants (1 mg) were dissolved in 200  $\mu$ L 1% (v/v) ammonium hydroxide and diluted with pH 6.9 100 mM ammonium acetate to a total volume of 1500  $\mu$ L. Peptides were then injected into a Slide-A-Lyzer Dialysis Cassette (2K MWCO, 3 mL volume). 100mM ammonium acetate was added to fill volume of cassette capacity, and dialysis was performed over a 24 h period. Samples were then removed from the cassette, flash frozen with liquid nitrogen and lyophilized overnight. Peptides were reconstituted using 200  $\mu$ L 1% (v/v) ammonium hydroxide and diluted with pH 6.9 100 mM ammonium acetate to a total volume of 1500  $\mu$ L. Peptide concentration for the stock solution was calculated from absorbance at 280nm ( $\epsilon = 1490 \text{ M}^{-1} \text{ cm}^{-1}$ ). Peptides were aliquoted and stored at  $-80 \text{ }^\circ\text{C}$ . Leucine enkephalin acetate hydrate (YGGFL) and prepared in pH 6.9 100mM ammonium acetate. FL was synthesized by AnaSpec, and prepared in dimethylsulfoxide.

## IM-MS

Mass spectra were collected on a quadrupole-ion mobility-time-of-flight (TOF) mass spectrometer (Synapt G1 HDMS, Waters) with a nano-electrospray ionization (nESI) source. Protein ions were generated using a nESI source and optimized to allow transmission of noncovalent protein complexes using electrospray capillaries prepared as described previously.<sup>23</sup> Protein complex ions were generated using an aliquot of the sample (ca. 7  $\mu$ L) sprayed from the nESI emitter at a capillary voltage of 1.2kV. The source was operated in positive ion mode with the sample cone at 35 V and extraction cone at 1.0V. The bias voltage was 21 V, with backing pressure at 2.32 mbar and TOF pressure at  $1.81 \times 10^{-6}$  mbar. The travelling-wave IM separator was operated at a pressure of approximately 3.0 mbar of nitrogen. Mass spectra were calibrated externally using a solution of cesium iodide (100 mg mL<sup>-1</sup>) and analyzed using MassLynx 4.1 and Driftscope 2.0 software. CCS ( $\Omega$ ) measurements were externally calibrated using a database of known values in helium, using values for peptides and proteins that bracket the likely CCS and ion mobility values of the unknown ions.<sup>24, 25</sup> Samples were prepared by mixing stock solutions (as prepared above) of neuropeptide and A $\beta$ <sub>1-40</sub> variants at pH 6.9, to generate a final A $\beta$ <sub>1-40</sub> concentration of 20  $\mu$ M, with 2.88% v/v DMSO present in all samples. Samples were incubated on ice for 1 hour prior to IM-MS analysis.

## Docking Studies

Flexible ligand docking studies were performed using AutoDock Vina<sup>26</sup>. FL was created using PyMOL and studies were conducted against the A $\beta$ <sub>1-40</sub> monomer (aqueous solution NMR structure PDB 2LFM)<sup>27</sup> and A $\beta$ <sub>1-42</sub> (aqueous solution NMR structure PDB 1Z0Q).<sup>28</sup> Twenty docking simulations with A $\beta$ <sub>1-40</sub> and thirty docking simulations with A $\beta$ <sub>1-42</sub>, one with each conformation within the ensemble uploaded to the PDB, were performed. The structures for FL and A $\beta$  were prepared for use with Auto Dock Vina using AutoDock Tools.<sup>26</sup> All hydrogens were added to A $\beta$  and torsions were kept as the default selected in AutoDock Tools. The exhaustiveness for the docking studies was set at 8, resulting in 9 output models for each A $\beta$  conformation. Docked models of FL were visualized with A $\beta$  using PyMOL.

## MD Simulations

Simulations were started using the minimized A $\beta$  solution NMR structures (PDB 2LFM or PDB 1Z0Q).<sup>27, 28</sup> Charges were applied to R5, K16, and K28 to mimic the 3<sup>+</sup> charge state observed in the gas phase. The simulations were performed using periodic boundary conditions in a dodecahedron with the minimum distance between the simulated molecules and the box wall being 1.0 nm. GROMACS ligand topology was prepared using the GlycoBioChem PRODRG2 server.<sup>29</sup> The MD simulations were carried out using the GROMACS software package<sup>30</sup> and GROMOS96 force field.<sup>31</sup> To constrain the bond length in the A $\beta$  and FL, the LINCS algorithm was used, allowing an integration time step of 2 fs. Long-range electrostatic interactions were treated with the particle mesh Ewald method. Temperature was maintained using the method of Berendsen et al.<sup>32</sup> The FL and the A $\beta$  were separately coupled to external temperature bath with a temperature-coupling constant of 0.1 ps.



The system was energy-minimized by steepest decent for 500 steps. After equilibration, simulated annealing was performed for the A $\beta$ :FL complex in the gas-phase. The three most basic A $\beta$  side chains (R5, K16, and K28) were charged. The system was heated from 300 K to 500 K over 100 ps, then cooled down to 300 K over the next 100 ps. The cycle was repeated over 20 ns in order to allow for escape from local minima and enhance equilibration. For the A $\beta$ <sub>1-40</sub>:FL complex, 20 independent simulated annealing runs, each running for a total of 20 ns, were performed from the lowest energy complexes generated by AutoDock Vina. For the A $\beta$ <sub>1-42</sub>:FL complex, 30 independent simulated annealing runs, each running for a total of 20 ns, were performed from the lowest energy complexes generated by AutoDock Vina. From the MD trajectory generated, 100 structures were sampled at 300 K and the CCS was calculated with Mobcal using the trajectory method algorithm.<sup>33, 34</sup> Models of the A $\beta$ :FL complexes were visualized in PyMOL. In total 2000 structures were generated for A $\beta$ <sub>1-40</sub> and 3000 structures for A $\beta$ <sub>1-42</sub>. Of the structures which were within  $\pm 3\%$  of the experimentally determined CCS, the 200 structures with lowest energy were analyzed to determine the A $\beta$  residues within 4 Å of FL. Standardized values (Z-scores) were calculated for each residue of A $\beta$ <sub>1-40</sub> and A $\beta$ <sub>1-42</sub> and plotted in standard deviation ( $\sigma$ ) space in order to determine the relative likelihood of FL binding within a given region of A $\beta$ .

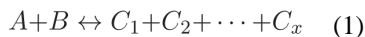
### Hierarchical Clustering Analysis

The CCS-filtered structures of A $\beta$ <sub>1-40</sub>:FL, generated from MD simulations in the previous step were subjected to another filter for agreement with site-directed amino acid substitution experiments (FL was within 4 Å of A $\beta$  residue Y10 and Q15). This reduced the set of possible structures from 963 to 60. We then characterized the structural families present in the dataset using hierarchical clustering. A $\beta$  candidate structures were temporarily stripped of their ligands while being aligned using the Kabsch algorithm,<sup>35</sup> and pairwise RMSD values were calculated for each structure in the dataset. These pairwise RMSD values were then used as a distance matrix for hierarchical clustering using the average linkage clustering module from Scipy.<sup>36</sup> Intra and inter-cluster RMSD values were analyzed and it was found that the dataset could reasonably be broken into 5 clusters, allowing for facile interpretation of the dataset. The results of the clustering are represented as a dendrogram where each leaf represents an individual structure, and the height of the branches represents the distance between structures or clusters in RMSD space. For each structural cluster, an average structure was calculated to represent the structural family, and all ligands were then superimposed on the representative structure, revealing strong structure-specific ligand clustering, in addition to an averaged ligand location. Structures were visualized in PyMOL.

### K<sub>D</sub> Measurements by MS

K<sub>D</sub> values for peptide:A $\beta$  complexes were calculated using the relative intensity of each species from the mass spectra, as described previously.<sup>37</sup> In our previous work,<sup>20</sup> we modified this method to accommodate multiple ligand binding events with the following assumptions: (1) the spray and detection efficiency of all species are similar, (2) the ligand concentration is sufficiently high so that [L]<sub>eq</sub> remains constant and (3) the ligand binds to the complex one at a time in a stepwise fashion. Here we present further modification to better represent multiple ligand binding stoichiometry.

For the equilibrium binding of A (unbound protein) and B (a ligand/binding partner):



$$R_x = \frac{[C_x]_{\text{eq}}}{[A]_{\text{eq}}} \quad (2)$$

Where  $R_x$  is an equilibrium quotient between the bound form of the protein ( $C_x$ , having  $x$  ligands attached) when interacting with ligand B, and is unbound for (A).

$$[C_x]_{\text{eq}} = \frac{R_x([A]_0 - (\sum_{i=1}^{x-1} y[C_i]_{\text{eq}}))}{1+yR_x} \quad (3)$$

Eqn (3) above defines  $R_x$  for all bound forms of the protein ( $C_i$ ), where  $y$  is the number of species A in the complex and:

$$K_{D_x} = \frac{[C_{(x-1)}]_{\text{eq}}[B]_{\text{eq}}}{[C_x]_{\text{eq}}} \quad (4)$$

Allows for the determination of  $K_D$  for any given step in the sequential equilibrium described in eqn (1), where  $C_0 = A$ . Standard deviation values ( $\sigma$ ) for the  $K_D$  measurements reported here are shown from at least three replicate measurements.

## TEM

Negatively stained specimens for transmission electron microscopy were prepared by applying 5  $\mu\text{L}$  of peptide solution to hydrophilic 400 mesh carbon-coated formvar support films mounted on copper grids (Ted Pella, Inc.). The samples were allowed to adhere for 3 min, rinsed twice with distilled water, and stained for 60–90 s with 1% uranyl acetate (Ted Pella, Inc.). All samples were imaged at an accelerating voltage of 60 kV in a CM-100 electron microscope (Philips, Inc.). Samples were prepared in 100 mM ammonium acetate, with 2.88% DMSO at a concentration of 25  $\mu\text{M}$   $A\beta$  and allowed to incubate at room temperature for 24 h.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Funding Sources

This research was supported jointly by the UM protein folding disease (PFD) initiative, part of the University of Michigan Health System (UMHS) Strategic Research Initiative and Fast Forward programs, and National Institutes of Health (National Institute of General Medical Sciences, GM095832).

The authors wish to thank M. Ivanova (UM) for the collection of TEM images. Additionally J. Clark (UM) for his assistance and advice in determining modes of binding. R. Kerr (UM) for his discussion of A $\beta$  variants and preparation ideas. K. Korshaven and M. Wolf (UM) for assistance in troubleshooting CD.

## ABBREVIATIONS

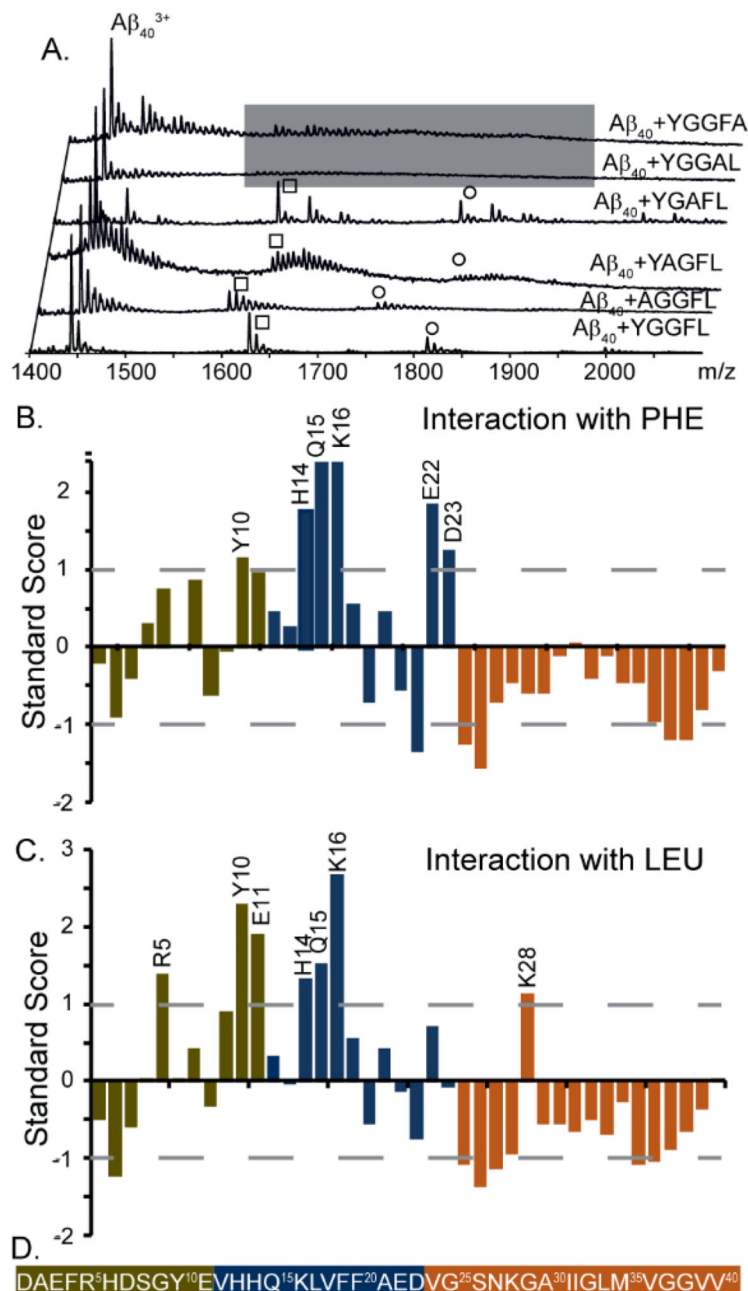
<b>IM-MS</b>	Ion mobility-mass spectrometry
<b>AD</b>	Alzheimer's Disease
<b>TEM</b>	Transmission Electron Microscopy
<b>CD</b>	Circular Dichroism
<b>KD</b>	Dissociation Constant
<b>LE</b>	Leucine Enkephalin
<b>A<math>\beta</math></b>	Amyloid $\beta$ peptide

## References

1. Young LM, Saunders JC, Mahood RA, Revill CH, Foster RJ, Ashcroft A, Radford S. ESI-IMS-MS: A method for rapid analysis of protein aggregation and its inhibition by small molecules. *Methods*. 2015
2. Young LM, Saunders JC, Mahood RA, Revill CH, Foster RJ, Tu L-H, Raleigh DP, Radford SE, Ashcroft AE. Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nature chemistry*. 2015; 7:73-81.
3. Woods LA, Platt GW, Hellewell AL, Hewitt EW, Homans SW, Ashcroft AE, Radford SE. Ligand binding to distinct states diverts aggregation of an amyloid-forming protein. *Nature chemical biology*. 2011; 7:730-739. [PubMed: 21873994]
4. Young LM, Cao P, Raleigh DP, Ashcroft AE, Radford SE. Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. *Journal of the American Chemical Society*. 2013; 136:660-670. [PubMed: 24372466]
5. Hyung S-J, DeToma AS, Brender JR, Lee S, Vivekanandan S, Kochi A, Choi J-S, Ramamoorthy A, Ruotolo BT, Lim MH. Insights into antiamyloidogenic properties of the green tea extract (-)-epigallocatechin-3-gallate toward metal-associated amyloid- $\beta$  species. *Proceedings of the National Academy of Sciences*. 2013; 110:3743-3748.
6. Beck MW, Oh SB, Kerr RA, Lee HJ, Kim SH, Kim S, Jang M, Ruotolo BT, Lee J-Y, Lim MH. A rationally designed small molecule for identifying an in vivo link between metal-amyloid- $\beta$  complexes and the pathogenesis of Alzheimer's disease. *Chemical Science*. 2015; 6:1879-1886. [PubMed: 28706643]
7. Pithadia AS, Kochi A, Soper MT, Beck MW, Liu Y, Lee S, DeToma AS, Ruotolo BT, Lim MH. Reactivity of diphenylpropynone derivatives toward metal-associated amyloid- $\beta$  species. *Inorganic chemistry*. 2012; 51:12959-12967. [PubMed: 23153071]
8. Bleiholder C, Do TD, Wu C, Economou NJ, Bernstein SS, Buratto SK, Shea J-E, Bowers MT. Ion Mobility Spectrometry Reveals the Mechanism of Amyloid Formation of A $\beta$  (25-35) and Its Modulation by Inhibitors at the Molecular Level: Epigallocatechin Gallate and Scyllo-inositol. *Journal of the American Chemical Society*. 2013; 135:16926-16937. [PubMed: 24131107]
9. Gessel MM, Wu C, Li H, Bitan G, Shea J-E, Bowers MT. A $\beta$  (39-42) modulates A $\beta$  oligomerization but not fibril formation. *Biochemistry*. 2011; 51:108-117. [PubMed: 22129303]

10. Zheng X, Gessel MM, Wisniewski ML, Viswanathan K, Wright DL, Bahr BA, Bowers MT. Z-Phe-Ala-diazomethylketone (PADK) disrupts and remodels early oligomer states of the Alzheimer disease A $\beta$ 42 protein. *Journal of Biological Chemistry*. 2012; 287:6084–6088. [PubMed: 22253440]
11. Kirkitadze MD, Bitan G, Teplow DB. Paradigm shifts in Alzheimer's disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies. *Journal of neuroscience research*. 2002; 69:567–577. [PubMed: 12210822]
12. Lesné S, Koh MT, Kotilinek L, Kaye R, Glabe CG, Yang A, Gallagher M, Ashe KH. A specific amyloid- $\beta$  protein assembly in the brain impairs memory. *Nature*. 2006; 440:352–357. [PubMed: 16541076]
13. Necula M, Kaye R, Milton S, Glabe CG. Small molecule inhibitors of aggregation indicate that amyloid  $\beta$  oligomerization and fibrillization pathways are independent and distinct. *Journal of Biological Chemistry*. 2007; 282:10311–10324. [PubMed: 17284452]
14. Ono K, Condrón MM, Teplow DB. Structure–neurotoxicity relationships of amyloid  $\beta$ -protein oligomers. *Proceedings of the National Academy of Sciences*. 2009; 106:14745–14750.
15. Bernstein SL, Dupuis NF, Lazo ND, Wyttenbach T, Condrón MM, Bitan G, Teplow DB, Shea J-E, Ruotolo BT, Robinson CV. Amyloid- $\beta$  protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nature chemistry*. 2009; 1:326–331.
16. Bitan G, Fradinger EA, Spring SM, Teplow DB. Neurotoxic protein oligomers—what you see is not always what you get. *Amyloid*. 2005; 12:88–95. [PubMed: 16011984]
17. Clemmer DE, Jarrold MF. Ion mobility measurements and their applications to clusters and biomolecules. *Journal of Mass Spectrometry*. 1997; 32:577–592.
18. Bohrer BC, Merenbloom SI, Koeniger SL, Hilderbrand AE, Clemmer DE. Biomolecule analysis by ion mobility spectrometry. *Annual review of analytical chemistry (Palo Alto, Calif.)*. 2008; 1:293.
19. Benesch JL, Ruotolo BT, Simmons DA, Robinson CV. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chemical reviews*. 2007; 107:3544–3567. [PubMed: 17649985]
20. Soper MT, DeToma AS, Hyung S-J, Lim MH, Ruotolo BT. Amyloid- $\beta$ –neuropeptide interactions assessed by ion mobility-mass spectrometry. *Physical Chemistry Chemical Physics*. 2013; 15:8952–8961. [PubMed: 23612608]
21. Paivio A, Jarvet J, Graslund A, Lannfelt L, Westlind-Danielsson A. Unique physicochemical profile of  $\beta$ -amyloid peptide variant A $\beta$ 1–40E22G protofibrils: Conceivable neuropathogen in arctic mutant carriers. *Journal of molecular biology*. 2004; 1:145–159.
22. Walsh D, Klyubin I, Fadeeva J, Rowan M, Selkoe D. Amyloid-beta oligomers: their production, toxicity and therapeutic inhibition. *Biochemical Society Transactions*. 2002; 30:552–557. [PubMed: 12196135]
23. Hernández H, Robinson CV. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols*. 2007; 2:715–726. [PubMed: 17406634]
24. Ruotolo BT, Hyung SJ, Robinson PM, Giles K, Bateman RH, Robinson CV. Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. *Angewandte Chemie International Edition*. 2007; 46:8001–8004. [PubMed: 17854106]
25. Bush MF, Hall Z, Giles K, Hoyes J, Robinson CV, Ruotolo BT. Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Analytical chemistry*. 2010; 82:9557–9565. [PubMed: 20979392]
26. Trott O, Olson AJ. 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. [PubMed: 19499576]
27. Vivekanandan S, Brender JR, Lee SY, Ramamoorthy A. A partially folded structure of amyloid-beta (1–40) in an aqueous environment. *Biochemical and biophysical research communications*. 2011; 411:312–316. [PubMed: 21726530]
28. Tomaselli S, Esposito V, Vangone P, van Nuland NA, Bonvin AM, Guerrini R, Tancredi T, Temussi PA, Picone D. The  $\alpha$ -to- $\beta$  Conformational Transition of Alzheimer's A $\beta$ -(1–42) Peptide in

- Aqueous Media is Reversible: A Step by Step Conformational Analysis Suggests the Location of  $\beta$  Conformation Seeding. *ChemBioChem*. 2006; 7:257–267. [PubMed: 16444756]
29. SchuËttelkopf AW, Van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*. 2004; 60:1355–1363. [PubMed: 15272157]
  30. Lindahl E, Hess B, Van Der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling*. 2001; 7:306–317.
  31. van Gunsteren WF, Billeter S, Eising A, Hünenberger PH, Krüger P, Mark AE, Scott W, Tironi IG. Biomolecular simulation: The {GROMOS96} manual and user guide. 1996
  32. Berendsen HJ, Postma JPM, van Gunsteren WF, DiNola A, Haak J. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*. 1984; 81:3684–3690.
  33. Mesleh M, Hunter J, Shvartsburg A, Schatz G, Jarrold M. Structural information from ion mobility measurements: effects of the long-range potential. *The Journal of Physical Chemistry*. 1996; 100:16082–16086.
  34. Shvartsburg AA, Jarrold MF. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chemical Physics Letters*. 1996; 261:86–91.
  35. Kabsch W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*. 1976; 32:922–923.
  36. Jones E, Oliphant T, Peterson P. {SciPy}: Open source scientific tools for {Python}. 2014
  37. Wang W, Kitova EN, Klassen JS. Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Analytical chemistry*. 2003; 75:4945–4955. [PubMed: 14708765]



**Figure 1.** Interactions between  $A\beta_{1-40}$  and leucine enkephalin variants. A) IM-MS analysis of alanine substituted variants of leucine enkephalin reveals complexes formed between  $A\beta_{1-40}$  and AGGFL, YAGFL, and YGAFI in a manner similar to WT leucine enkephalin. Samples containing YGGAL and YGGFA produced no measurable complex signals, indicating F and L to be critical for the interaction with  $A\beta$ . Square annotations represent the 1:1 interaction, and circles the 1:2 interaction. B) and C) Plots of the standard score (Z-score) for  $A\beta$  residues that are within 4Å of the F and L in leucine enkephalin peptide structure. D) Sequence of  $A\beta_{1-40}$ , where the N-terminus (green) is defined as residues 1–11, the

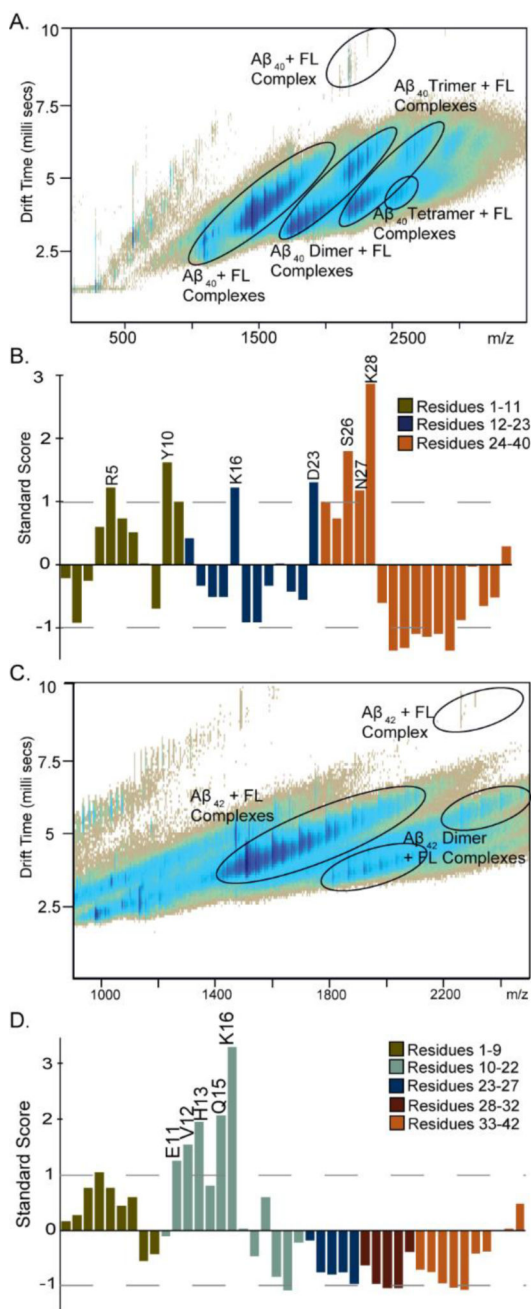
hydrophobic core (blue) is residues 12–23, and the C-terminus is residues 24–40. This color-coding for A $\beta$ <sub>1–40</sub> applies to all figures in this report.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2.**

Interaction between Aβ with the dipeptide FL. A) IM-MS data shows the presence of Aβ<sub>1-40</sub> in complex with FL as a monomer, dimer, trimer, and tetramer at 20 μM Aβ: 60 μM FL. Multiple copies of FL bind to each oligomer under these concentrations, while at lower FL concentrations (10 and 20 μM) only 1:1 complexes are observed. B and D) Plots of the standard score (Z-score) for Aβ residues within 4 Å of the bound FL. Larger values denote contacts of greater significance on the standard deviation ( $\sigma$ ) scale. Negative values denote contacts of reduced significance. C) IM-MS data shows the presence of Aβ<sub>1-42</sub> in complex with FL as both monomer and dimer, with up to 4 copies of FL binding to each oligomer at



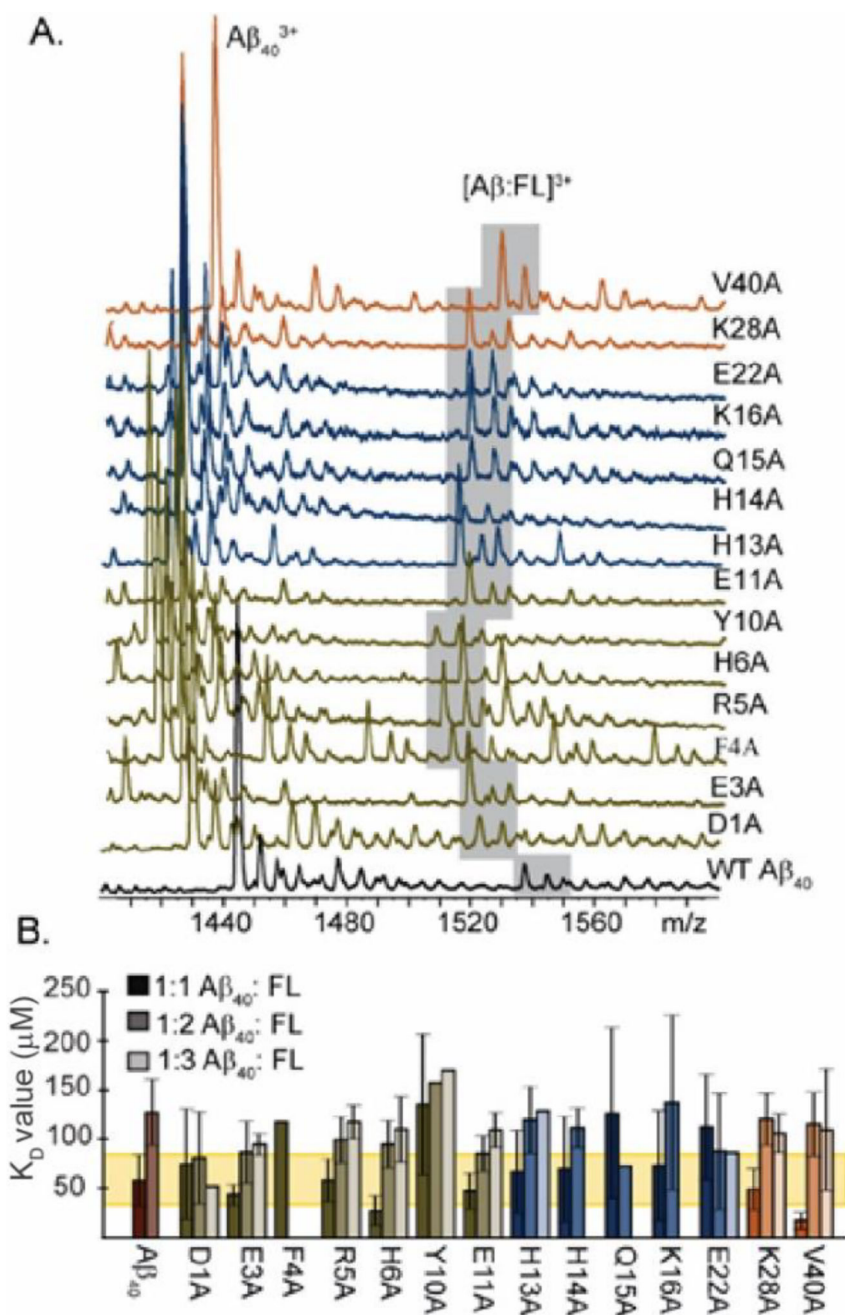
20  $\mu\text{M}$   $\text{A}\beta$ : 60  $\mu\text{M}$  FL. As with  $\text{A}\beta_{1-40}$ , with lower concentrations of FL, fewer FL copies are bound.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Figure 3.**

Site directed amino acid substitution effects on FL binding. A) Shown is a representative spectra of 20  $\mu M$   $A\beta_{1-40}$  variant incubated with 80  $\mu M$  FL for 1 hour on ice. The grey boxed region indicates the 1:1 complex detected at a  $3^+$  charge state. Signal for the 1:1 complexes of Y10A: FL and Q15A: FL show decreases in intensity when compared to the WT interaction. B)  $K_D$  values calculated for each amino acid substitution. Error bars indicate 1 standard deviation ( $\sigma$ ). The yellow bar is  $\pm 1$  ( $\sigma$ ) of the 1:1 complex. Of the 1:1 interactions observed, only Y10A, Q15A, and E22A are above 1  $\sigma$  different than the WT. All 1:2

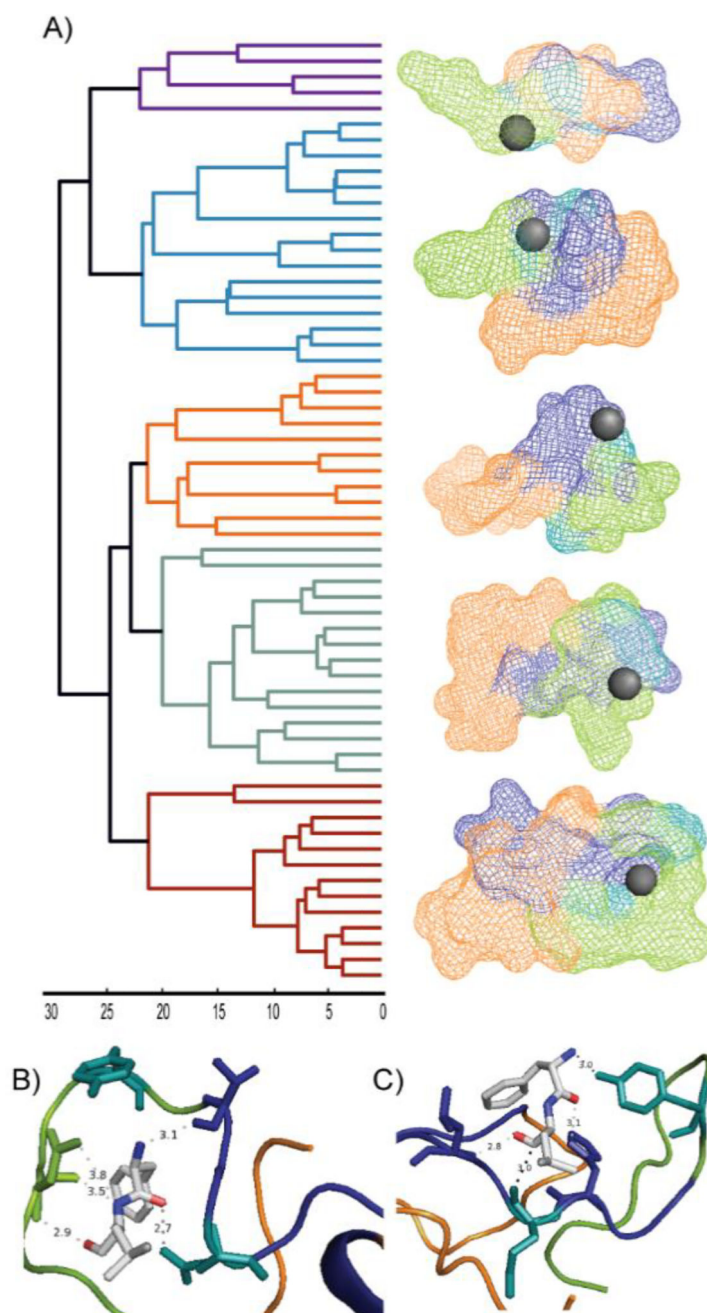
interactions fell within standard deviation of the WT 1:2 interaction and a WT 1:3 interaction could not be measured in triplicate due to signal-to-noise constraints.

Author Manuscript

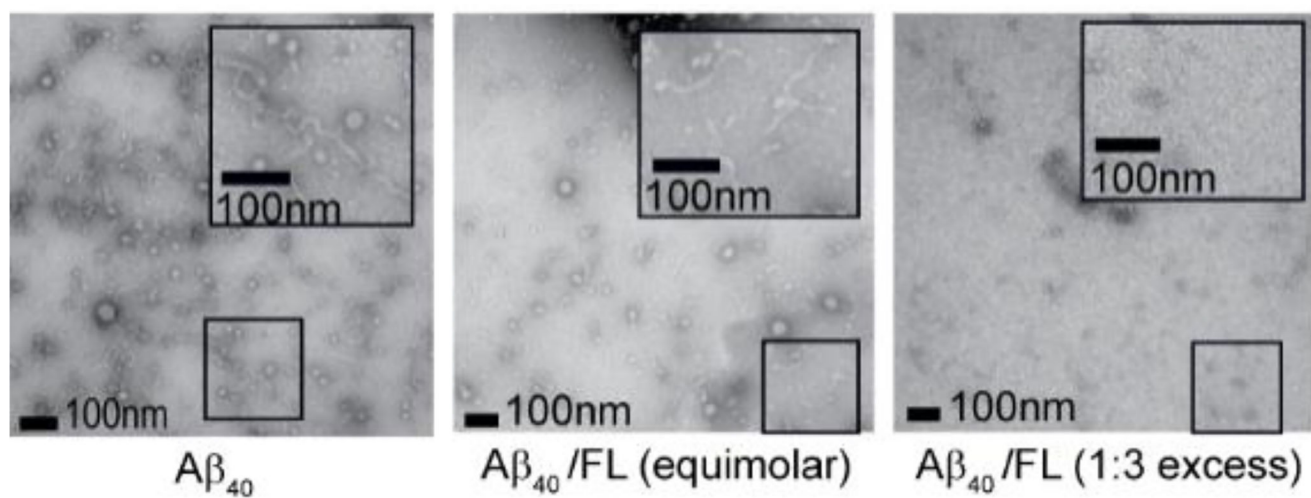
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 4.** Hierarchical clustering analysis and model visualization of A $\beta_{1-40}$ :FL complexes. A) Clustering hierarchy, and the averaged family structure for each cluster. Each leaf of the dendrogram represents an individual A $\beta_{1-40}$ :FL complex and the height of the branches represents the average linkage distance as measured by pairwise RMSD (Å). B and C) High resolution representative models of A $\beta$ :FL complex. Q15 and Y10 (teal) primarily interact with the leucine of FL. N-terminus (green), hydrophobic core (blue), C-terminus (orange), and FL (grey, oxygen in red, nitrogen in blue).



**Figure 5.** TEM images of  $A\beta_{1-40}$  incubated with FL (as labeled), displaying reduced fibril formation upon FL addition, in a dose-dependent fashion.