Full length amylin oligomer aggregation: insights from molecular dynamics simulations and implications for design of aggregation inhibitors

Workalemahu Mikre Berhanu and Artem E. Masunov

aNanoScience Technology Center and Department of Chemistry, University of Central Florida, Orlando, FL 32826, USA; bDepartment of Chemistry, and Department of Physics, NanoScience Technology Center, University of Central Florida, Orlando, FL 32826, USA

Communicated by Ramaswamy H. Sarma

(Received 16 June 2013; accepted 5 August 2013)

Amyloid oligomers are considered to play essential roles in the pathogenesis of amyloid-related degenerative diseases including type 2 diabetes. Using an explicit solvent all atomic MD simulation, we explored the stability, conformational dynamics and association force of different single-layer models of the full-length wild-type and glycine mutants of amylin (pentamer) obtained from a recent high resolution fibril model. The RMSF profile shows enhanced flexibility in the disorder (Lys1-Cys7) and turn region (Ser19-Gly23), along with smallest fluctuation at the residues (Asn14-Phe15-Leu16-Val17-His18) of the $\beta_1$ region and (Ala25-Ile26-Leu27-Ser28-Ser29) of the $\beta_2$ region. We obtained a significant difference in backbone RMSD between the wild-type and the mutants, indicating that mutations affected the stability of the peptide. The RMSD and RMSF profiles indicate the edge and loop residues are the primary contributors to the overall conformational changes. The degree of structural similarity between the oligomers in the simulation and the fibril conformation is proposed as the possible explanation for experimentally observed shortening of the nucleation lag phase of amylin with oligomer seeding. On the basis of structure-stability findings, the $\beta_1$ and $\beta_2$ portions are optimal target for further anti-amyloid drug design. The MM-PBSA binding energy calculation reveals the binding of amylin: amylin strands in single layer is dominated by contributions from van der Waals interactions. The non-polar solvation term is also found to be favorable. While the electrostatic interactions and polar solvation energy was found to be favorable for the interaction for the larger aggregate and unfavorable for the smaller aggregates. A per-residue decomposition of the binding free energy has been performed to identify the residues contributing most to the self-association free energy. Residues found in the $\beta$-sheet regions were found to be key residue making the largest favorable contributions to the single-layer association. The result from our simulation could be used in rational design of new amylinomimetic agent, amylin aggregation inhibitors and amylin-specific biomarkers.

Keywords: critical nucleation; lag phase; cross seeding; oligomer; amylin; MM-PBSA; free energy; cluster analysis; secondary structure; RMSD; RMSF; pramlintide

1. Introduction

Amylin, also called islet amyloid polypeptide (IAPP), is the major peptide or protein component of the islet amyloid found in the pancreas of approximately 90% of type 2 diabetes patients (Cooper et al., 1987). It is a 37-amino acid peptide hormone that is co-secreted with insulin from islet $\beta$ cells in response to meals. This hormone helps improve blood glucose control after meal. It does this by promoting the following actions: (1) slowing digestion, (2) blocking glucagon secretion, and (3) enhancing satiety thereby reducing food consumption (Krentz, Patel, & Bailey, 2008). The formation of amyloid aggregates by association of peptides into ordered structures is a hallmark of many serious human disorders (Selkoe, 2004). Amyloid fibers can also have a functional role in bacteria, fungi, insects, invertebrates, and humans (Maury, 2009). Amyloid-forming proteins do not show any striking similarity, either in sequence or in their secondary or tertiary structure, yet all aggregate into a common core structure, the amyloid fibril.

Studies using X-ray fiber diffraction revealed that all amyloid fibrils share a cross-$\beta$-structure, composed of $\beta$-strands which run perpendicular to the axis of the fibril, and backbone hydrogen bonds running parallel to the fibril axis (Sawaya et al., 2007). High-resolution, atomic structures of full-length amyloid fibrils, obtained
through X-ray crystallography and solution nuclear magnetic resonance (NMR) spectroscopy, have not been solved as amyloid fibrils are insoluble and non-crystalline (Petkova et al., 2002). However, recent studies employing solid-state NMR (ss-NMR) have provided details on inter- and intra-molecular interactions within fibrils, and have shed light on mechanisms of aggregation (Antzutkin, Leapman, Balbach, & Tycko, 2002).

Amylin undergoes aggregation to form amyloid fibrils. Fibrillar amylin has been shown to be toxic to cultured β-cells and has been proposed to play a significant role in the pathogenesis of type 2 diabetes by contributing to the destruction of β-cells (Sasahara, Hall, & Hamada, 2010). It would be beneficial to improve the water solubility of amylin by means of protein engineering to prevent aggregation formation and thus avoid the interaction of the aggregate with cell membrane (Mack et al., 2010). Amylin-mimetic peptide pramlintide (a protein where Ala25, Ser28, and Ser29 of the wild amylin substituted with proline), approved for treatment of type 2 diabetes while another analog of amylin (davalintide), is under preclinical investigation (Mack & Hamada, 2010). It would be beneficial to improve the water solubility of amylin by means of protein engineering to prevent aggregation formation and thus avoid the interaction of the aggregate with cell membrane (Mack et al., 2010). Amylin-mimetic peptide pramlintide (a protein where Ala25, Ser28, and Ser29 of the wild amylin substituted with proline), approved for treatment of type 2 diabetes while another analog of amylin (davalintide), is under preclinical investigation (Mack et al., 2010). The disadvantage of pramlintide is that it has a short half-life and low bioavailability. Therefore, finding analogs that have improved pharmacokinetic and pharmacodynamic profiles is essential. Soluble oligomers of amylin, in common with other amyloidogenic proteins, have been implicated in the disruption of natural lipid membranes (Sasahara et al., 2010). It has been shown that aggregates of amylin can disrupt the integrity of the β-cell’s membrane, allowing entry of Ca²⁺ ions into the cell either through the formation of small pores or complete disruption of the membrane (Demuro et al., 2005). Wiltzius, Sievers, Sawaya, and Eisenberg (2009) have found that amylin can adopt a α-helical structure at residues 8–18 and 22–27. They suggested helical dimerization of amylin accelerates fibril formation and that insulin impedes fibrillation by blocking the amylin dimerization interface. Not all species form islet amyloid. In particular, human and other primates express amylin that can form amyloid fibrils, but rodents do not (Westerman, Engstrom, Johnson, Westerman, & Betsoltz, 1990). The primary sequences of rodent and human IAPP are very similar aside from the region of residues 20–29. Mouse amylin differ from the human protein at only 6 of 37 positions, 5 of which are located in the region of residues 20–29. There are three proline residues in this region of mouse amylin located at positions 25, 28, and 29, while human amylin has none. Other differences are replacement of His-18 in human amylin with Arg-18 in the rat sequence and substitution of Phe at position 23 and Ile at position 26 of the human sequence with Leu-23 and Val-26, respectively, in mouse IAPP (Wiltzius, Landau et al., 2009). The sequences of the human and mouse polypeptides are shown in Figure 1. The human amylin contains a disulfide bond between Cys2 and Cys7 and recent experimental studies suggest that the presence of an intra-molecular disulfide bond between Cys2 and Cys7 enhances the amylin fiber formation (Koo & Miranker, 2005; Yonemoto, Kroon, Dyson, Balch, & Kelly, 2008). The aggregation-prone region of the peptide (residues 20–29) has been identified through comparison of amylin variants from different species with variable amyloidogenic propensity. Similar to the full-length human Islet Amyloid Poly-Peptide (hIAPP), several fragments, including hIAPP8–20, hIAPP14–20, hIAPP20–29, hIAPP22–27, hIAPP23–27, and hIAPP30–37, form amyloid fibrils in vitro (Nilsson & Raleigh, 1999). Among these peptides, hIAPP20–29 has been suggested to be responsible for the amyloidogenic propensities of full-length human amylin.

Luca, Yau, Leapman, and Tycko (2007) published a single fibril structure derived from ss-NMR data. They report a structure in which a cross-β subunit consists of stacked β-sheets formed from the parallel in-registry assembly of a U-shaped β-strand-loop-β-strand motif. The loop is located at residues 18–27, straddled by two β-strands comprising residues 8–17 and 28–37, respectively. Recently, Wiltzius et al. (2008) published a fibril model of the full-length human amylin extrapolated from X-ray diffraction data of cross-β spine structures of two segments of human amylin (NNFGAIL and SSTNVG). The topology of this model is similar to that reported by Luca et al. (Wiltzius et al., 2008). The models revealed that the amyloid peptide adopts a β-strand-loop-β-strand motif, with each amylin peptide stacked in a parallel and staggered manner within β-sheets. Among the three aromatic residues, F15 is the only one that resides in the β-sheet core, while F23 is located in a bend and Y37 is exposed at the C-terminus. While atomistic characterization of the fibril form of this amyloid peptide has been emerging, (Luca et al., 2007; Wiltzius et al., 2008) the structures of the early aggregation species, including monomer and small oligomers, remain poorly understood. To date, atomic information for the aggregation of the amylin peptide in explicit water is limited.

A number of amyloid-based therapies including inhibitors of aggregation of the amyloid-β (Aβ) peptide have been examined for their potential application to the treatment of Alzheimer’s disease (Hawkes, Ng, & McLaurin, 2009). By contrast, only a limited number of investigations have been made into the similar design of compounds to inhibit aggregation of amylin (Potter et al., 2009). Mutagenesis has been used to probe secondary structure and inter-sheet side-chain packing. Single proline substitutions within the 20–29 fragments of amylin revealed that substitution of residues 22, 24, and 26–28 destabilizes fibrils and alters the kinetics of fibril formation (Moriarty & Raleigh, 1999). Abedini and
Raleigh (2006) designed a variant of the amyloidogenic 8–37 region of human amylin with proline substitutions at positions 17, 19, and 30. Compared to the wild-type, the mutant had dramatically greater solubility. Fox et al. (2010) recently have identified 22 single mutations, 4 double mutations, and 2 triple mutations of IAPP that appear to be less amyloidogenic than wild-type human amylin. A comparison of these sequences suggests residues 13 and 15–17 comprise an additional aggregation-prone region outside of the main amyloidogenic region (20–33 regions) of amylin.

The identification of amino acid important in stabilizing the amyloid aggregate using MD simulation based on wild-type and mutation could be a viable option in designing inhibitors. One approach that is used to design such inhibitors is to synthesize short peptides that correspond to a self-recognition element (SRE) of the native amyloid sequence but contain key modifications, so that the peptides bind to the parent protein at the SRE and prevent further aggregation. Such modifications include β-sheet-breaking amino acid substitutions, addition of N- and C-terminal blocking groups, replacement of

---

**Figure 1.** Comparison of sequence human and mouse amylin and structures of human amylin peptide. Primary sequences of human amylin (A) and mouse amylin (B). Both peptides have a disulfide bridge between Cys-2 and Cys-7. Residues that differ from that of the human peptide are colored red. The segments of human amylin important for fibril formation are underlined. The amino acids in the N terminal region (colored in green) are thought to be responsible for interaction with membrane. (C) The structures of an amylin peptide Eisenberg’s model monomer. Residues 1–8 are disordered. The (β1-strand)-loop-(β2-strand) motif of a monomeric subunit includes 8–37 with hydrophobic inter-sheet packing interactions of residues Q10 to H18 with residues I26 to S34 of the opposite side. Positively charged, polar, and hydrophobic residues are colored red, blue, green, and grey, respectively. (D) Pentameric stacks of amylin single layer representing the fibril conformations. The layers are arranged to form an inter-molecular parallel β-sheet. Color coding is as follows: chain A, green; chain B, yellow; chain C, red; chain D, blue; chain E, gray. The disulfide bonds are shown in cyan.
amide bonds with ester linkages, and introduction of R-disubstituted amino acids such as R-aminobutyric acid (Masman et al., 2009; Serio et al., 2000). Investigating the structural fluctuations of amylin may provide knowledge for designing drugs to inhibit amylin aggregation targeted at the less flexible portion. Understanding the dynamic behaviors of the amyloidogenic peptides is expected to provide insights into the possible mechanism of amyloid formation. Clearly, more detailed structural descriptions of the human amylin is very important for a molecular understanding of the oligomerization process and for the development of innovative therapeutic and diagnostic approaches in type 2 diabetics. Molecular dynamic (MD) simulation on amylin full-length model may provide valuable insight into the aggregation of the amylin peptide and the forces that drive peptide–peptide complexes. The study of the wild-type and mutants in an explicit solvent could provide valuable insight to future amyloid aggregation-guided inhibitor design efforts.

Computational studies have complemented experiments to provide insights into amyloid formation. Pervious MD studies have focused on the fibrilligenic properties of short human amylin peptides of 7–10 residues, aiming to identify regions likely to be responsible for the amyloidogenic properties of full-length amylin (Wu, Lei, & Duan, 2005; Zanuy, Porat, Gazit, & Nussinov, 2004). Xu, Ping, Li, and Mu (2009) investigated the β-sheets composed of seven anti-parallel decapptides, representing the 20–29 segments of hIAPP. The amyloid nucleus of hIAPP was mimicked with one β-sheet of different initial separation distances between the strands. Multiple all-atom MD simulations with explicit water solvent showed that the assembly occurs not only in the lateral direction but also along the longitudinal direction. Moreover, a possible oligomeric state with cyclic form was suggested based on one assembly model found in the simulations. The cyclic structures of amyloid oligomers have been reported to be the early intermediates in solution, capable to form ion-channel-like structures in the membrane that could be responsible for pathologic permeability and destabilization of the cellular ionic homeostasis (Quist et al., 2005).

The primary nuclei for fibrillar growth have been often described as oligomers (Serio et al., 2000). The distinctive feature of the amyloid oligomers is the development of cross-β-sheet structure (Morozova-Roche & Malisauskas, 2007). The structural similarities and differences between amylin oligomers and the mature fibril are not yet fully understood. Recent experimental observation indicates that oligomers play a role in both initiating fibril growth (Padrick & Miranker, 2002) and cytotoxicity (Zraika et al., 2010), it would be helpful to know whether the amylin conformation detected in the mature fibril is also the most prevalent topology in the oligomer aggregates. Simulation of the single-layer amylin aggregates based on the high-resolution models of amylin fibril is of interest for understanding the nature of self-interaction. There has been no systemic study on how mutation affects the stability of the amylin oligomer aggregates based on the recent high-resolution amylin fibril model. MD simulation of the different size of the amylin oligomer will be useful to obtain deeper understanding on possible nucleus seeds and conformation at the very early stage of fibril formation. Atomic knowledge of the structure of amylin smaller size oligomer (such as monomer to pentamer) is an essential step in understanding the aggregation process because they represent building blocks for larger oligomers. Theoretical study on designed interfacial mutation using MD simulation can be used to get an insight into the contribution of each interface residue to monomers assembly. MD simulations can be a complementary tool, as they show in atomic detail, to study how mutations influence the structure and dynamic relative to the wild-type. Specifically, we would like to answer the following questions: (1) which regions of the wild-type amylin oligomer aggregate are flexible. (2) The influence of three proline mutations at positions 25, 28, and 29, and single glycine mutation of the structure and flexibility of amylin oligomer? (3) The contribution of the different amino acids in the amylin: amylin assembly in the single-layer oligomer aggregate. (4) The possible minima nucleus seed for fibril formation and their similarity to the fibril conformation. (5) The influence of the number of amylin peptides on the structural stability and conformational dynamic of the amylin single layer assembly. Here, we investigated the conformational stability and association thermodynamics of different size of full-length wild-type amylin single layers and mutants (Table 1S). We conducted a total of ~0.5 μs of explicit-solvent MDs simulation on the single-layer amylin aggregate oligomer and mutants, using temperatures 330 K to emulate the experimental conditions of in vitro fibrillization (Sasahara, Naiki, & Goto, 2005).

2. Computational details
2.1. System setup

Two experiment-based models of the three-dimensional structures of amylin oligomers were recently obtained: the first one is the Tycko’s model (Luca et al., 2007) and the second one is the Wiltzius et al.’s model (Wiltzius et al., 2008). The Tycko’s models consist of two different conformations (Luca et al., 2007). The main differences in peptide conformation in the two models (I and II) are in residues 19–21, 24, and 25, resulting in different sets of side-chain contacts between β-sheet layers (Luca et al., 2007). In the case of the Wiltzius et al.’s
model (2008), it is close to the Tycko’s model I. The coordinates of the pentamer models comprising of five chains of the 37 residue amylin were kindly provided by Tycko (Luca et al., 2007) and by Sawaya (Wiltzius et al., 2008). We used the Wiltzius et al.’s model to get the monomer, dimer, trimer, tetramer, and pentamer form of amylin. The initial structures of monomer, dimer, trimer, and tetramer of the wild-type were obtained by removing the monomers one by one from pentamer (Table 1S). The wild-type of the Wiltzius et al.’s single-layer model was mutated to examine the effect of the side-chain interactions of the amino acids involved in stabilizing the strand to strand association of amylin. The side chains of hydrophobic residues have been found to mediate specific interactions that direct the self-assembly of amylin. The mutants are obtained from wild-type termed 2G, 1034GG, 1626GG, and 252829PPP, respectively (Table 1S). The MD simulation was performed using AMBER11 amber99SB force field and explicit TIP3P water models (Case TAD TEC et al., 2010). Each of the amylin models and the corresponding mutants were solvated by explicit water molecules that extends 10 Å from any edge of the octahedral box to the protein atoms. Counter-ions were added to the box by randomly replacing water molecules to neutralize the system. Each system was initially energy minimized to remove bad contact by using conjugate gradient method with the peptide constrained and then to relax the atoms without position constrains (Case TAD TEC et al., 2010). The system was then subjected to 200 ps of heating procedure while constraining the backbone atoms of the protein to allow relaxation of water and ions, followed by 500 ps equilibration run without position constraints on the peptides. The production simulation time was 40 ns for both full amylin models and the corresponding mutants. Temperature was maintained with Langevin dynamics (Pastor, Brooks, & Szabo, 1988) protocol with a collision frequency of 1 ps⁻¹. The pressure was kept constant at 1 atm by weak coupling to a bath of constant pressure with a coupling time of 0.5 psec (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984). Electrostatic interactions were calculated by using the particle mesh Ewald method (Darden, York, & Pedersen, 1993). The cutoff radius for the Lennard-Jones interactions was also set to 12 Å. The SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1977) was used for bond constraints and the time step was 2 fs for all simulations. Each system was simulated for 40 ns and the trajectories were saved at 8.0 ps intervals for further analysis. Visual molecular dynamics (Humphrey et al., 1996) program was used for the visualization of trajectories.

2.3. Binding free energy calculation

The MM-GBSA single trajectory approach implemented as script (MMPBSA.py) in Amber11 (Case TAD TEC et al., 2010) was used to calculate the binding energy. To calculate binding free energies in MM-GBSA method, the explicit water simulations were used to generate the trajectory followed by the implicit Generalized Born-Surface Area method to calculate solvation energy terms. The gas phase and the solvation-free energies were calculated over the course of the 40 ns of the MD production trajectories. This approach was previously used to study the thermodynamics of amyloidogenic peptides (Berhanu & Masunov, 2011).

The amylin single-layer oligomer aggregates studied contain multiple protein–protein interfaces; the calculation of the free energy of the associations of monomers in single-layer oligomer aggregates requires a suitable interface. Because one of the aims of the present study is to assess the stability of the amylin oligomer with respect to the increase in the number of strand (the longitudinal growth), we measured the interaction energy...
between the edge (A) stand and the remaining strands (B) as shown in Figure S1. The free energy analysis was done using a single trajectory approach, where the complex (C), receptor (B) oligomer aggregate, and ligand (A) snapshots were taken from the snapshot of the performed MD trajectory. According to the MM-GBSA method, (Kollman et al., 2000) binding-free was calculated using Equation (1):

\[
\Delta G_{\text{bind}} = \langle G_C \rangle - \langle G_A \rangle - \langle G_B \rangle
\]  

(1)

The bracket \( \langle \rangle \) indicates an average of these energy terms over 2500 snapshot shots extracted from the MD simulation.

\[
\Delta G_{\text{bind}} = \langle \Delta E_{\text{MM}} \rangle + \langle \Delta G_{\text{solv}} \rangle - \langle T \Delta S \rangle
\]  

(2)

The free energy of each system \( X = A, B, \) or \( C \) was computed as a sum of the three terms:

\[
\langle \Delta G_X \rangle = \langle E_{\text{MM}} \rangle + \langle \Delta G_{\text{solv}} \rangle - T \langle S \rangle
\]  

(3)

where \( E_{\text{MM}} \) is the molecular mechanics energy of the molecule expressed as the sum of the internal energy (bonds, angles, and dihedrals) \( (E_{\text{int}}) \), electrostatic energy \( (E_{\text{ele}}) \), and van der waals term \( (E_{\text{vdw}}) \):

\[
E_{\text{MM}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdw}}
\]  

(4)

\( \Delta G_{\text{solv}} \) accounts for the solvation energy which can be divided into the polar and non-polar part:

\[
\Delta G_{\text{solv}} = \Delta G_{\text{GB}} + \Delta G_{\text{SA}}
\]  

(5A)

The polar part \( \Delta G_{\text{GB}} \) accounts for the electrostatic contribution to solvation and is obtained from Generalized Born (GB) calculations in a continuum model of the solvent. The second term \( \Delta G_{\text{SA}} \) is non-polar contribution to solvation-free energy that is linearly dependent on the solvent accessible surface area (SASA):

\[
\Delta G_{\text{SA}} = \gamma \text{SASA} + b
\]  

(5B)

The \( \Delta G_{\text{SA}} \) were calculated using AMBER11 default parameter for \( \gamma \) and \( b \) (5b). The entropic contribution was calculated in this study using the normal mode module in AMBER11 (Gohlke & Case, 2004). Since normal node calculation of the entropic contribution is computationally expensive, it was averaged only over 50 frames of the MD trajectory.

2.4. Mutant studies

It is still unclear what triggers the conversion of soluble monomeric IAPP into insoluble amyloid fibrils. Mostly, biophysical and mutagenesis approaches have been used to understand amylin fibril formation in vitro (Abedini & Raleigh, 2006; Jaikaran et al., 2001). A wealth of information about the role of the amino acid sequence in IAPP fibrillation has been obtained from rat amylin, which differs in only six amino acid residues but does not form amyloid fibrils. The non-aggregating nature of rat amylin was attributed to three proline residues in regions 20–29 in the peptide. Although the preventive properties of the proline residues in IAPP fibrillation have been well studied, biophysical studies that pinpoint the roles of specific residues are limited. Some experimental mutational studies performed on the small fragments of amylin in positions 20–29 have shown that proline substitution has led to the absence or reduced aggregation (Jiang, Xu, & Mu, 2009). Yan et al. show that the double N-methylated full-length IAPP analog [(N-Me) G24, (N-Me) I26]-IAPP (IAPP-GI) is a highly soluble, non-amyloidogenic, and non-cytotoxic amylin molecular mimic and an amylin receptor agonist (Yan, Tatarek-Nossol, Velkova, Kazantzis, & Kapurniotu, 2006). When proline residues from the non-amyloid mouse amylin sequence were substituted into the fibrilizing segments (region 20–29) of human amylin, the mutant did not form fibrils. The high-resolution amylin model has some residues which are in close proximity that forms inter-molecular side-chain to side-chain interactions which may stabilize the amylin oligomer. To investigate this hypothesis, we performed selected mutations of some of the residues involved in the side-chain–side-chain interaction in the opposite strands (such as residues Lys16, Ile26, Glu10, Ser34, and the disulfide bridge between residue 2 and 7) of single-layer amylin; the pentamer with the tiny \( \beta \)-breaker amino acid glycine (Kim & Hecht, 2005). The side chains that we examined are Q10S34 and 16L26I. In amylin, the disulfide bond between Cys2 and Cys7 has been identified as an important contributor of fiber formation. Fox et al. (2010) have identified two mutants (C2Y and C7R) as non-amyloidogenic. These mutations resulted in the loss of intra-molecular disulfide bond and, hence, would be expected to be less amyloidogenic. We performed a single point in silico mutation replacing Cys2 with glycine. The simulation of amylin pentamers mutants (2G, 1034GG, 1626GG and 252829PPP) was done using the simulation protocol mentioned above.

3. Result and discussion

3.1. Relative structure stability of amylin oligomers

The conformational change and the conservation of the oligomers were monitored by the time evolution of the backbone root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) through the simulation relative to their initial energy minimized structure as
shown in Figures 2 and 3. Higher RMSD and RMSF values indicate higher mobility and lower values indicated limited mobility.

3.1.1. Root-mean-square deviation

The RMSD provides an overall measure of the departure of the structures from the initial X-ray Cartesian coordinates, at each frame of the MD trajectory. The RMSDs provide useful information on relative stability of the oligomers, and were previously used in stability analyses of oligomers with β-sheet structure (Berhanu & Masunov, 2011; Masman et al., 2009). The conformation change and conservation of oligomers’ stability of the three different double-layer models of U-shaped disordered-β-strand-loop-β-strand motif of amylin (Tycko’s model I, Tycko’s model II and Wiltzius et al.’s model) were monitored by the time evaluation of the RMSD. The recent high-resolution model of amylin (Wiltzius et al.’s model) has a lower RMSD than Tycko’s et al. (Luca et al., 2007) models I and II (data not shown). Even though all the three models are similar, they differ in the details of the packing of side chains (Wiltzius et al., 2008) and this is the reason for the difference in the RMSD of the models. We used Wiltzius et al.’s model as it has the lowest RMSD for our studies. The RMSD variation of the wild-type and mutants along each simulation is shown in Figure 2. The structures of the monomer and dimer significantly depart from the initial conformations as shown in the RMSDs (Figure 2(A)). The greatest changes of RMSD occur for the monomer (∼11 Å), followed by those for the dimer (∼9 Å). The RMSD of the trimer was about 4 Å and the RMSD of the bigger aggregates (tetramer and pentamer) are ∼3 Å (see Figure 2(A)) over the 40 ns simulation. The RMSD of the monomer quickly increased to 12 Å within after 2 ns. The dimer has an RMSD of 6 Å within the first 10 ns and then it increase to 8 Å and remains within an
RMSD of 8 Å for the last 15 ns of the simulation. These suggest the monomer and dimer are not stable. However, the RMSD of the trimer, tetramer, and pentamer are small indicating the structural stability of the larger complex. The RMSD of the wild-type pentamer and glycine mutants are shown in Figure 2(B). The result indicates double glycine mutation at L16 and I26 leads to significant instability with an RMSD value of about 7 Å double the value of the RMSD of the wild-type. The next unstable double glycine mutants are 2G, 1034GG, and 252829PPP with an RMSD of about 6 Å. None of the mutants were found to be structurally stable as the wild-type. Comparison of the dynamic of wild-type with the double mutants suggests the replacement of the residues L16 and I26 with glycine has a more pronounced destabilizing effect indicating that this side chain is important for the stability of amylin. The amino acids Leu16 and Ile26 contain hydrophobic side chains and large surface area. Their substitution with the uncharged small side chain Gly which is found in flexible regions of proteins perturbs the stability of amylin. The change in hydrophobicity and the surface area resulting from mutation is the major factor responsible for instability of this mutation. In other mutants, the major change from the mutation is a reduction in the surface area.

3.1.2. Root-mean-square fluctuation

The residue-based RMSF of the backbones was used to assess the local dynamics and flexibility of each residue
using ptraj tool in amber. Figure 3 shows the RMSF profiles of the different oligomers of the wild-type and mutants of amylin single-layer models. The RMSF of the monomer, dimer, trimer, tetramer, and pentamer are shown in Figure 3(A). Analysis of the flexibility shows the residue base fluctuation decreases with increase in the number of strand stacking. In agreement with others previous studies on the β-strand-loop-β-strand (U) fold of amyloid oligomers (Zheng, Jang, Ma, Tsai, & Nussinov, 2007), the N- and C-termini are the most flexible parts of the peptides for all the systems. The enhanced flexibility is also observed in the disorder (Lys1-Cys7), turn region (Ser19-Gly23), and at the glycine residues (Gly33). The N-terminal part of the peptides demonstrates reduced flexibility despite it being located in the highly disordered region compared to the C-terminus. The fluctuation difference between the edge and central residues becomes smaller for the larger oligomers (Figure 3(A)). The smallest and central residues becomes smaller for the larger oligopeptides demonstrates reduced glycine residues (Gly33). The N-terminal part of the (Lys1-Cys7), turn region (Ser19-Gly23), and at the C-terminus. The fluctuation difference between the edge and central residues becomes smaller for the larger oligomers (Figure 3(A)). The smallest fluctuation occurred at the residues Asn14-Phe15-Leu16-Val17-His18 of β1 region and Ile26-Leu27-Ser28-Ser29 of the β2 region. The shorter peptide sequences derived from fibrillogenic regions of the human amylin corresponding to residues 8–20, 10–19, and 20–29 have been identified as “hot spots” of amyloid formation (Kajava, Aebi, & Steven, 2005). Our simulation also shows these regions (14–18 and 26–29 above) are crucial stabilizing elements in the process of aggregation and possibly in the elongation of the amylin aggregates. These regions could be a promising target for further drug design of sequence-specific β-sheet breaker peptides capable of interacting specifically with these portions of the aggregates. Recently, disulfide bond between Cys2 and Cys7 has been implicated with the stabilization of N-terminal and the initialization of amylin amyloid formation (Milton & Harris, 2010). Our simulation of the single point of glycine mutation of the amino acid cystine at position 2 shows that the flexibility of the N-terminal increases indicating the significant role of the disulfide bond between cys2 and cys7 in stabilizing the N-terminal region (Figure 3). The RMSF values of the mutant pentamers of amylin are shown in Figure 3(B). The RMSF profiles of all the mutants except 1626GG show that the residues at the β1 region and β1 region have relatively lower fluctuations than residues at the edges. The results indicate that the relatively large structural fluctuation comes mainly from the edges of the oligomer. In the case of 1626GG mutant, the lowest fluctuation occurs in the residues 10–14 and it has the greatest residue fluctuation of all the mutants, showing important role of the Lys16 to Ile26 side-chain interaction in stabilizing the amylin. The other mutants were also found to have a larger RMSF value compared to the wild-type pentamer. Both the RMSD and RMSF consistently show that the addition of extra β hair-pin of the wild-type increases the stability of the amylin. The difference in RMSD and RMSF between the wild-type and the corresponding mutants was significant indicating that the mutations affect the stability of the aggregate.

3.2. Hydrogen bond analysis

The amyloid configuration and properties primarily depend on the density of hydrogen bonds involving the backbone of the polypeptides (Fox et al., 2010), while the side-chain hydrogen bonds are involved in the geometrical details and extension of the disordered parts of the structure (Kim & Hecht, 2006). To further characterize the structural stability of the studied amylin wild-type and mutant models, the hydrogen bonds between backbones and side chains were examined. The hydrogen bonds between peptides play a major role in stabilizing the structures of proteins. If a sufficient number of backbone hydrogen bonds are deleted, the fibril formation and growth can be inhibited. Introduction of the β sheet-breaking residue within the β strands resulted in a loss of fibril formation and growth. The side-chain–side-chain hydrogen bonds as well as backbone–backbone hydrogen bonds were found to be affected by the mutations (Figure 4). Hydrogen bonding plays crucial roles in the formation of protofilaments and in the packing in the protofilaments. The average number of hydrogen bonds per layer for monomer and dimers is lower than those of the other larger oligomer aggregate (trimer, tetramer, and pentamer). The increase in hydrogen bonds improves the formation of ordered β-sheet layer (Figure 4(A)). The mutant structures are less organized, with smaller average number of hydrogen bonds (Figure 4(B)). The average number of hydrogen bonds of the wild-type is the largest, compared to the hydrogen bond contents of the similar size aggregates of the mutants. This helps the wild-type to retain fibril topology with β-cross sheets organization. The destabilizing effect of a mutation was further found to correlate with its influence on the number of hydrogen bonds per monomer. When proline replaces a residue of the β-strand, the kink formed at the position of the proline residue will considerably weaken the hydrogen bonding between the neighboring strands. Triple backbone modifications (252829PPP) at the positions Ala 25, Ser28, and Ser29 with proline blocking hydrogen bonding between β-sheets disrupting β-sheet extension and packing. This approach has been a viable strategy for arresting fibril growth. Another approach has been the incorporation of N-methyl amino acids into peptide analogs of the aggregating peptide (Yan, Velкова, Tatarek-Nossol, Andreetto, & Kapurniotu, 2007). The addition of an N-methyl group reduces hydrogen-bonding capabilities due to the replacement of the amide hydrogen, and so disrupts the network of hydrogen bonds required for formation of a stable β-sheet-based fibril.
3.3. Secondary structure analysis

We carried out secondary structure analysis using DSSP (Kabsch & Sander, 1983). Allam et al. (Sasahara et al., 2005), based on molecular simulations and infrared experiments, have found the monomer amylin protein exhibits a helical secondary structure in the hydrophobic region spanning residues 11–19 and 23–32. The hydrophobic residues LANFLVH (residues 12–18) and AILSSTNV (residues 25–32) show high propensity for β structures in amylin fibril model structure used in this study. The hydrophobic region (residues 14–18 and 25–29) is considered to be the core part of β-sheets amylin aggregation based on aggregation propensity experiments (Bhak, Choe, & Paik, 2009). The C terminal residues 1–8 are generally unstructured, as their preferences for any of the three sorts of secondary structures are very low expect in the case of the monomer which shows high propensity for turn/bend. The β1 and β2 regions (residues 10–18 and 25–32, respectively) show a rather high propensity for β structures for all amylin models except the amylin monomer. Figures 4(C) shows the average secondary structure content of amylin single-layer oligomers. The amylin monomer exhibits lower content of β-sheets, and greater number of residues in helices and in turn-like conformation. In the case of the larger size, oligomers (such as SH1-ST2 to SH1-ST5) exhibit higher content of β-sheets, and smaller number of residues in helices and in turn conformation. The larger aggregates retain the fibril conformation mainly due to an increase in the number of backbone hydrogen bonds (W. M. Berhanu & Masunov, 2011). In the case of the mutants, the β sheet contents declined compared to the wild-type (Figure 4(D)). Our results indicate that the wild-type exhibits a higher degree of the original secondary structures compared to the mutants. The secondary structure analysis shows the role of the mutated side chains in stabilizing the amylin oligomers. The results of β sheet content are in agreement with the RMSD and RMSF result in that those with highest RMSD and RMSF fluctuation are found to be lacking the ability to preserve secondary structure and conformational stability.

Figure 4. Comparison of the number of hydrogen bonds per layer and secondary structure content of wild-type and mutants of amylin. Number of hydrogen bonds per layer of the wild-type models (A) and the 5 strands per of the hairpin of amylin wild-type and mutants (B). Secondary structure content for the wild-type models (C) and the pentamer wild-type and its corresponding mutants (D).
3.4. Clustering analysis

If one is interested to organize and conceptualize the enormous structural data-set from MDs simulations in order to understand how the molecule moves within this space, and how these parts and movements eventually connect to particular macroscopic and microscopic properties, structural clustering is important (Keller, Daura, & van Gunsteren, 2010). Conventional clustering algorithms are used for reducing any large MD trajectory to obtain reduced representations of these trajectories. Cluster analysis puts similar samples of data into groups called clusters, such that an ensemble of data, for example, the different structures obtained from a MD trajectory, is partitioned into groups of similar objects. To identify the most populated conformations sampled, clustering of all snapshots from the trajectories was performed using the ptraj program of Amber11. The standard approach, which has been used with considerable success, is to cluster the configurations in terms of an RMSD. For clustering, we utilized the average linkage algorithm implemented in ptraj (Shao, Tanner, Thompson, & Cheatham, 2007). The uniqueness or equivalence of different clusters was assessed based on visual comparison of representative structures. The clustering was performed on a 5000-frame reference set (8 ps sampling rate). Tables 2S and 3S show the values of the different clustering metrics. Figure 5(A–E) shows the superposition of the initial structure and the most populated cluster structure of the various sizes of amylin single-layer aggregate. The analysis of the structures indicated the most populated clusters detected from the smaller size oligomers (single and double strand) indicating larger structural rearrangements compared to the initial conformation taken from the fibril model. While in the case of the larger aggregate (SH1-ST3 and SH1-ST5), the conformation is preserved. The trimer to pentamer shows an overall conservation of the U-shaped topology. Due to the high stability of the C-terminal β-strand and the U-shaped topology, the trimer to pentamer adopts a confirmation that allows longitudinal elongation along the fibril axis, suggesting that these conformations might serve as nuclei for protofibril growth. The smaller fluctuations observed in the C-terminal strand suggest that the interface involved in lateral fibril association (Figure 3) is fully preformed in these small oligomers. The different stability of the secondary structure elements suggests that the trimer to pentamer should exist in conformations competent for both longitudinal elongation and lateral association of the oligomers. The initial topology is already lost in the

![Figure 5](https://example.com/figure5.png)

Figure 5. Superposition of the initial fibril structure onto the respective most populated clusters representative structures of the (A) Monomer, (B) Dimer, (C) Trimer, (D) Tetramer and (E) Pentamer. The initial fibril conformations are shown in blue and the most populated cluster with the corresponding percentage (refers to cluster occupancy) are shown in magenta.
early steps of monomer simulation due to the lack of stabilizing interactions between adjacent layers of the parallel β-sheet (Figure 5(A)). In contrast to the folds of monomer, trimer through pentamer on the other hand retains the β-hairpin conformation (Figure 5(C–E)). The initial U-shaped structure is not stable in the dimer and the hydrophobic contacts between the C- and C-terminal stack of β-sheets are lost, resulting in an overall enhanced flexibility. Most of the mutants showed the initial U-shaped structure was lost especially by the mutants with a sequence similar to pramlintide, 1626GG, and 1034GG with the loss of C-terminal contact (Figure S2). In the case of the 2G mutants, the deletion of the disulfide bridge between res2 and 7 enhanced the flexibility of the N terminal region (Figures S2 and 3) with loss of the initial U-shaped conformation.

3.5. Free energy calculation

Detailed characterization of individual energy terms of the calculated binding-free energy of the studied amylin oligomer aggregates is shown in Table 1. An inspection of the free energy components for the wild-types reveals that the electrostatic component of the free energy of binding ($\Delta G_{\text{ele}}$) contributes favorably to binding ($\Delta G < 0$). The non-polar component contributes favorably ($\Delta G_{\text{n-polar}} < 0$) as expected, since formation of complexes reduces solvent-accessible surface area. In most cases, the electrostatic component of the solvation-free energy $\Delta G_{\text{solv}}$ is consistently favorable. The MM-GBSA binding energy calculation reveals the binding of amylin: amylin strands in single layer is dominated by contributions from van der Waals interactions. The non-polar solvation term is also found to be favorable. The electrostatic interactions are unfavorable for the larger aggregate, while for the same system polar solvation energy was found to be favorable. The result of the binding-free energy calculation indicates the structurally stable models have the lowest binding-free energy (i.e. more negative in value), while the models which are structurally unstable were found to have the largest binding-free energy. The trend in the calculated binding-free energy is in agreement with the observed instability based on RMSD, RMSF. Those aggregate oligomer models which show structural instability were found to have unfavorable binding energy compared to the stable ones. Results from this work provided valuable insight into the forces that drive the stability of the peptide–peptide complexes of the single-layer aggregate oligomer models of amylin and those that lead to unstable complexes. The study of the wild-type and mutants in an explicit solvent will provide valuable insight to guided future short- and long-acting amylin analogs’ design efforts.

3.6. Decomposition free energy on a per-residue basis

The free energy decomposition not only identifies the binding energy hot spots but also gives insight into the nature of key interaction (Zoete, Meuwly, & Karplus, 2005). To provide the basic information on the intermolecular interactions contributed from the individual residues in the amylin, single-layer aggregate interaction decomposition of free energy (the per residue total, side chain and backbone binding-free energy) was evaluated using the decomposition energy module in AMBER. The calculation was performed over the 2500 MD snapshots taken from the 40 ns simulation. The summations of per residue interaction-free energies were separated into the residue backbone ($\Delta G_{\text{backbone}}$) and the side chain ($\Delta G_{\text{side chain bind}}$). The energy contributions from the selected residues are summarized in Figure 6.

The energy decomposition shows the major contribution to the binding energy of amylin oligomer aggregate is gained from the key amino acid residues (those with a $\Delta G_{\text{binding}} < -0.50$ kcal/mol) occurring mainly in the β-sheet region (Table 2, Figures 6 and S4). These residues are in β1 (res 14–16) and in chain β2 (res 26–30 and res 35–36). The result of the per residue decomposition indicates the importance of the particular residues in the β-sheet region with regards to the formation and stabilization of amylin and this is in agreement with experimental observation that this region is important for amyloidogenesis (Wiltzius, Landau et al., 2009). Due to the electrostatic repulsion of adjacent amylin layers, the contribution of the two positively charged residues (Lys1 and Arg11) to binding energy is unfavorable, the unfavorable electrostatic term is counteracted by favorable solvation energy. The polar arginine residues buried in the hydrophobic regions have a reduced effect on electrostatic contribution (Arg 14, Arg 21, Arg 22, Arg

| Table 1. Binding-free energy components (kcal mol$^{-1}$) calculated with MM-GBSA for different size aggregates of the single-layer wild-type full amylin. |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| $\Delta E_{\text{vdw}}$     | $\Delta E_{\text{ele}}$     | $\Delta G_{\text{GB}}$      | $\Delta G_{\text{solv}}$    | $\Delta G_{\text{subtotal}}$| $\Delta S$                  | $\Delta G_{\text{binding}}$| $\Delta G_{\text{model}}$  |
| Dimer                      | $-133.7 \pm 18.9$          | $-86.8 \pm 72.0$            | $121.3 \pm 70.7$           | $-18.8 \pm 2.6$             | $102.5 \pm 70.8$            | $-117.9 \pm 20.9$           | $-54.1 \pm 7.7$             |
| Trimer                     | $-168.4 \pm 10.1$          | $6.0 \pm 56.7$              | $34.0 \pm 55.2$            | $-20.7 \pm 9.1$             | $133.5 \pm 55.2$            | $-149.0 \pm 12.4$           | $-61.9 \pm 5.1$             |
| Tetramer                   | $-176.8 \pm 5.5$           | $146.2 \pm 30.5$            | $-113.3 \pm 40.3$          | $-20.7 \pm 9$               | $134.3 \pm 40.3$            | $-164.8 \pm 8.5$            | $-60.8 \pm 6.3$             |
| Pentamer                   | $-177.4 \pm 9.4$           | $167.9 \pm 59.6$            | $-129.1 \pm 57.5$          | $-21.0 \pm 1.0$             | $151.9 \pm 57.5$            | $-161.3 \pm 13.38$          | $-61.96 \pm 6.43$           |

Note: $E_{\text{vdw}}$ and $E_{\text{ele}}$ are the van der Waals and electrostatic binding terms, $\Delta G_{\text{GB}}, \Delta G_{\text{solv}}, \Delta G_{\text{ele}}$ are the polar, non-polar and total solvation energies. $\Delta G_{\text{subtotal}} = \Delta E_{\text{vdw}} + \Delta E_{\text{ele}} + \Delta G_{\text{ele}} + \Delta G_{\text{sol}} = \Delta G_{\text{GB}} + \Delta G_{\text{solv}}$. Data are shown as mean ± SD.
and Arg 35). Our study provides the atomic level understanding of the aggregation behavior and the driving force for the amylin aggregates could be used to guide experimental investigations to design amylin aggregation inhibitors.

3.7. Fibril nucleation and the structure of amylin oligomers

Understanding the process of amyloid fibril formation is an important goal of protein aggregation studies. Amyloids grow in a nucleation-dependent manner (Bhak et al., 2009; Harper & Lansbury, 1997). Fibrillation kinetics is typically characterized by an initial apparent lag phase related to the formation of oligomers, protofibrils, and aggregation nuclei (Soto, Estrada, & Castilla, 2006). The typical fibril formation process is characterized by a lag phase in which no detectable fibers are formed. This is then followed by an elongation phase in which fiber is formed over a time period often shorter than the lag phase. Eventually, the process reaches equilibrium when most soluble proteins are converted into fibrils (Padrick & Miranker, 2002). On the other hand, if fibers (oligomers, protofibrils, and fibrils) are already formed, they grow extremely fast with very short lag times. The seeding phenomenon, i.e. lag time being reduced or eliminated by adding preformed fibrils, has been regarded as the key evidence of the nucleation mechanism for amyloid (Han, Weinreb, & Lansbury, 1995).

A recent experimental work on oligomers’ capability to stabilize fibril nucleation activity by Ono et al. (Ono, Condron and Teplow, 2009) on Aβ amyloid and on amylin (Padrick & Miranker, 2002) has indicated that the oligomers and the fibril showed different capability to act as seeds. Anselm et al. (Horn & Sticht, 2010) used the degree of structural similarity to the fibril conformation detected for the oligomers in their simulation as a possible reason for difference among various sizes of oligomer with respect to the efficiency as nucleation seeds for Aβ amyloid. Amylin shares amino acid sequence similarity with Aβ in the presumably ordered region and shows a similar secondary structure in the fibrillar state (Jayasinghe & Langen, 2004). The results from our simulation show the single-layer amylin oligomer as small as the trimer is capable of preserving the conformation present in the fibril (See Figure 5). The

![Figure 6. M-GBSA energy decomposition by residue inter-molecular interaction energies for amylin dimer (A), trimer (B), tetramer (C), and pentamer (D). Negative values indicate that the energy term is stabilizing. Black bars correspond to backbone energies while grey bars correspond to side-chain energies.](image-url)
### Table 2. Decomposition of the free energy on a per-residue basis of key residues of the full-length amylin pentamer association. (A) Total binding-free energy decomposition, (B) Decomposition of the side chain contributions to the binding free energy, and (C) Decomposition of the backbone contributions to the binding free energy.

| Residue | \(\langle E_{\text{vdw}} \rangle\) | \(\langle E_{\text{elec}} \rangle\) | \(\langle G_{\text{elec,solv}} \rangle\) | \(\langle G_{\text{pp,solv}} \rangle\) | \(\langle G_{\text{binding}} \rangle\) |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **A**   |                 |                 |                 |                 |                 |
| LYS1    | -.185           | 60.383          | -59.275         | -.058           | .865            |
| ALA5    | -.872           | -.036           | .258            | -.115           | -.765           |
| THR6    | -1.095          | -2.889          | 3.373           | -.149           | -.76            |
| CYX7    | -.988           | -.774           | 1.03            | -.098           | -.83            |
| GLN10   | -.997           | -3.777          | 4.097           | -.137           | -.815           |
| LEU12   | -1.181          | -2.017          | 2.301           | -.173           | -.107           |
| ASN14   | -1.593          | -3.102          | 3.181           | -.192           | -.105           |
| PHE15   | -1.184          | .022            | .402            | -.105           | -.865           |
| LEU16   | -1.962          | -1.391          | 1.704           | -.221           | -.187           |
| VAL17   | -.692           | -.512           | .466            | -.075           | -.812           |
| HIS18   | -1.497          | -1.051          | 1.589           | -.163           | -.123           |
| ASN21   | -1.18           | -1.395          | 1.873           | -.164           | -.866           |
| ASN22   | -1.511          | -3.897          | 3.291           | -.2             | -.231           |
| PHE23   | -.113           | -.23            | .819            | -.133           | -.675           |
| ILE26   | -1.096          | -1.363          | 1.388           | -.098           | -.116           |
| LEU27   | -.834           | -.423           | .417            | -.104           | -.943           |
| SER28   | -.722           | -1.279          | 1.023           | -.044           | -.1023          |
| THR30   | -.891           | -1.786          | 1.866           | -.101           | -.912           |
| VAL32   | -1.962          | -1.532          | 1.456           | -.138           | -.1176          |
| SER34   | -.809           | -2.565          | 2.791           | -.166           | -.749           |
| ASN35   | -1.713          | -4.401          | 4.621           | -.213           | -.706           |
| THR36   | -1.612          | -1.42           | 1.667           | -.286           | -.651           |

The energy terms correspond to the contribution of the entire considered residue (backbone and side-chain). A negative value of \(\Delta G_{\text{binding}}\) indicates that the corresponding residue makes a favorable contribution. Only residues making a significant favorable or unfavorable contribution are given (\(|G_{\text{binding}}|\geq 0.75\) kcal/mol). Energies are in kcal/mol.

| **B**   |                 |                 |                 |                 |                 |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| CYX2    | -.396           | -1.036          | .945            | -.059           | -.546           |
| ALA5    | -.404           | .108            | -.123           | -.101           | -.519           |
| CYX7    | -456            | -.626           | .564            | -.058           | -.577           |
| LEU12   | -.586           | 1.898           | -.1827          | -.093           | -.608           |
| PHE15   | -.641           | .356            | -.122           | -.089           | -.495           |
| LEU16   | -1.142          | 1.097           | -.986           | -.139           | -.117           |
| HIS18   | -1.497          | -1.051          | 1.589           | -.163           | -.123           |
| ASN22   | -.886           | -1.137          | 1.211           | -.119           | -.1164          |
| PHE23   | -.578           | .053            | .158            | -.084           | -.451           |
| ALA25   | -.132           | -.002           | -.002           | -.033           | -.169           |
| ILE26   | -.57            | 1.115           | -1.086          | -.032           | -.573           |
| LEU27   | -.502           | 1.007           | -.964           | -.096           | -.554           |
| VAL32   | -.495           | 2.35            | -.272           | -.074           | -.492           |
| ASN35   | -.964           | -3.22           | 3.202           | -.178           | -.116           |
| THR36   | -.927           | -.386           | .553            | -.221           | -.981           |

The energy terms correspond to the contribution of side-chain of key residue. A negative value of \(\Delta G_{\text{binding}}\) indicates that the corresponding residue makes a favorable contribution. Only residues making a significant favorable or unfavorable contribution are given (\(|G_{\text{binding}}|\geq 0.5\) kcal/mol). Energies are in kcal/mol.

| **C**   |                 |                 |                 |                 |                 |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| LYS1    | -.124           | 33.223          | -32.331         | -.048           | .72             |
| ASN3    | -.313           | -1.949          | 1.836           | -.042           | -.468           |
| THR6    | -.669           | -3.941          | 4.197           | -.072           | -.485           |
| GLN10   | -.546           | -3.549          | 3.63            | -.075           | -.54            |
| ARG11   | -.386           | -1.339          | 1.267           | -.02            | -.478           |
| LEU12   | -.595           | -3.915          | 4.128           | -.08            | -.461           |
| ASN14   | -.776           | -1.832          | 1.963           | -.08            | -.725           |
| LEU16   | -.82            | -2.488          | 2.69            | -.082           | -.7             |

(Continued)
monomer is a very poor seed for amylin, but dimers can already act as seeds. Trimers and tetramers exhibit very similar properties and are almost as active as fibrils themselves as seeds. The secondary structure content and clustering analysis on the trajectories from the various size single-layer amylin oligomer shows the larger aggregates retain the fibril conformation and the smaller ones (SH1-ST1 and SH1-ST2) lose the conformation. The observation could be used to explain the observed shortening of the nucleation lag phase of amylin aggregation with oligomer seeds observed experimentally (Padrick & Miranker, 2002). Based on the result of the secondary structure and cluster analysis, we proposed the SH1-ST3 to be a critical nucleus for the single amylin fibril oligomer growth. To characterize the critical nucleation, we computed the association energy which is different between our proposed minimum nuclei and the other oligomers (SH1-ST2, SH1-ST4, and SH1-ST5) we used in the equation.

$$
\Delta \Delta G(n) = G(n) - G(3); \quad n = 2, 3, 4, 5 \quad (6)
$$

The results are shown in Table 1 and it is plotted in Figure S5. Our calculation shows that at high numbers of chains, the oligomer is stable and its free energy is favorable for the addition of the new chains. The free energy terms correspond to the contribution of key residues backbone. A negative value of $\Delta G_{binding}$ indicates that the corresponding residue makes a favorable contribution. Only residues making a significant favorable or unfavorable contribution are given ($|G_{binding}| \geq .5 \text{kcal/mol}$). Energies are in kcal/mol.

| Residue | $\langle F_{vdw} \rangle$ | $\langle F_{elec} \rangle$ | $\langle G_{elec,solv} \rangle$ | $\langle G_{np,solv} \rangle$ | $\langle G_{binding} \rangle$ |
|---------|------------------|------------------|------------------|------------------|------------------|
| VAL17   | -.394            | -2.063           | 1.976            | -.013            | -.493            |
| ASN21   | -.532            | -.177            | -.129            | -.036            | -.52             |
| ASN22   | -.625            | -2.528           | 2.08             | -.081            | -1.153           |
| ILE26   | -.526            | -2.478           | 2.473            | -.065            | -.596            |
| LEU27   | -.332            | -1.431           | 1.381            | -.007            | -.389            |
| SER28   | -.501            | -2.019           | 1.571            | -.039            | -.989            |
| THR30   | -.512            | -2.425           | 2.31             | -.074            | -.702            |
| VAL32   | -.467            | -3.882           | 3.728            | -.063            | -.684            |
| ASN35   | -.749            | -1.18            | 1.419            | -.035            | -.546            |
| THR36   | -.685            | -1.034           | 1.115            | -.065            | -.669            |

The energy terms correspond to the contribution of key residues backbone. A negative value of $\Delta G_{binding}$ indicates that the corresponding residue makes a favorable contribution. Only residues making a significant favorable or unfavorable contribution are given ($|G_{binding}| \geq .5 \text{kcal/mol}$). Energies are in kcal/mol.

Figure 7. Snapshots of amylin (A and B) and Aβ pentamers (C and D) at the beginning and end of 40 ns simulation.
energy calculations demonstrate the increased stabilization of the aggregate with the number of monomers in this aggregate (dynamic cooperativity effect). As one can see from the Table 1, in the amylin the binding energy grows until it approaches the asymptotic value at the tetramer. Further size increase keeps the binding energy nearly constant (Berhanu & Masunov, 2012).

Studies have shown protein mis-folding can be enhanced by a cross-seeding mechanism, indicating that one protein mis-folding process may be an important risk factor for the development of a second (Yan et al., 2007). The structure of the amyloid aggregates are similar, they are characterized with a cross-β motif. Thus, a direct interaction between amyloid aggregate and dissimilar peptide may result in acceleration of protein aggregation through cross-seeding mechanism, leading to faster accumulation of amyloid aggregates (Morales et al., 2010). Even though human amylin displays sequence differences with Aβ peptide, the two peptides are able to efficiently cross-seed (Andreetto et al., 2010; Yan et al., 2007). Amylin was shown to have about 65% sequence similarity with Aβ peptide when overlapped in the central core region (residue 15–37) of Aβ peptide. Amyloid cross-seeding between 37-residue human amylin and 40- and 42-residue Aβ amyloid peptide occurs in spite of the difference in the primary sequence (Andreetto et al., 2010). To explore the conformational similarity, we performed a 40 ns simulation of the single-layer amylin pentamer and Aβ amyloid pentamer in order to explain why their cross-seeding occurs. The initial coordinates of the Aβ17-42 pentamer were extracted from the averaged NMR-based structure (PDB code: 2BEG) derived from NMR (Luhrs et al., 2005).

Our simulation indicates (Figure 7) both amylin and Aβ17-42 pentamer retained the initial β-hairpin conformation. The visual structural analysis of the snapshots of pentamer of amylin and Aβ in which the initial U-shape structure in both peptides is retained, suggested either of the peptides can serve as a template for the congruent incorporation of the other one. This might explain the experimentally observed cross-seeding between amylin and Aβ (O’Nuallain, Williams, Westermark, & Wetzel, 2004). The structural similarity between amylin and Aβ peptide and the cross-seeding between these two peptides indicate the similar approach can be utilized in designing inhibitors (Dessalew & Mikre, 2008). Thus, a number of amyloid aggregation inhibitors designed or screening for inhibition of Aβ amyloid could be effective in inhibiting the aggregation of human amylin.

4. Conclusions

The major findings of this study can be summarized as follows:

1. The stability of the amylin single-layer β-hairpin peptides oligomers increases with increasing the number of strands. The RMSF profile shows enhanced flexibility in the disordered and turn regions (Lys1-Cys7 and Ser19-Gly23), along with the smallest RMSF fluctuations in β1 and β2 regions (residues Asn14-Phe15-Leu16-Val17-His18 of the β1 region and Ala25-Ile26-Leu27-Ser28-Ser29 of the β2 region). On the basis of these structure-stability findings, the β2 and β1 portion is an optimal target for further anti-amyloid drug design.

2. The glycine substitution of the amino acids in pentamer mutants in which the side chains are involved in a face-to-face hydrophobic zipper that maximize intra-sheet packing interactions was found to be unstable oligomers. These indicate those residues are important for the stability of the wild-type and can be a potential target for structure-based design of aggregation inhibitor. The three proline substitutions at positions Ala25, Ser28, and Ser29 of amylin lead into unstable aggregation, which is in agreement with the enhanced water solubility and non-amyloidogenic property of pramlintide (where the Ala25, Ser28 and Ser29 residues of amylin is replace with three proline residues).

3. The secondary structure content and clustering analysis on the trajectories from the various size single-layer amylin oligomer show the larger aggregates retain the fibril conformation and the smaller ones (SH1-ST1 and SH1-ST2) lose the conformation. The observation could explain the observed shortening of the nucleation lag phase of amylin aggregation with oligomer seeds.

4. The binding energy calculated by MM-GBSA method shows the hydrophobic interactions play an important role in stabilizing the structural organizations of the single-layer amylin. Per residue decomposition shows the key amino acid residues (whose with a $E_{\text{binding}} \leq -1.00 \text{ kcal/mol}$) occur mainly in the β-sheet regions (β1 and β2). Due to the electrostatic repulsion of adjacent amylin layers, the contribution of the two positively charged residues (Lys1 and Arg11) to binding energy is unfavorable, the unfavorable electrostatic term counteracted by favorable sovation energy. The polar Asparagine resides buried in the hydrophobic regions have a reduced effect on electrostatic contribution (Asn 14, Asn 21, Asn 22, Asn 31 and Asn 35).

5. The structural similarity between amylin and Aβ17-42 pentamer has been suggested as possible explanation for the experimental observation of the cross-seeding between amylin and Aβ oligomer. Our study provides the atomic level
understanding of the aggregation behavior and the driving force for the amylin aggregates, and could be useful for rational design of amylinomimetic agent, amylin aggregation inhibitors, and amylin-specific biomarkers for diagnostic purposes.

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2013.832635.

Acknowledgements
This work was supported in part by the National Science Foundation (CHE-0832622). Research was performed in part using Stokes HPCC facility at UCF Institute for Simulation and Training (IST), and the National Energy Research Scientific Computing Center (NERSC) a DOE Office of Science user facility at Lawrence Berkeley National Laboratory.

References
Abedini, A., & Raleigh, D. P. (2006). Destabilization of human IAPP amyloid fibrils by proline mutations outside of the putative amyloidogenic domain: Is there a critical amyloidogenic domain in human IAPP? Journal of Molecular Biology, 355, 274–281. doi: 10.1016/j.jmb.2005.10.052
Andreetto, E., Yan, L. M., Taterek-Nossol, M., Velkova, A., Frank, R., & Kapurniotu, A. (2010). Identification of hot regions of the A beta-IAPP interaction interface as high-affinity binding sites in both cross- and self-association. Angewandte Chemie-International Edition, 49, 3081–3085. doi: 10.1002/anie.200904902
Antzutkin, O. N., Leapman, R. D., Balbach, J. J., & Tycko, R. (2002). Supramolecular structural constraints on Alzheimer’s beta-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. Biochemistry, 41, 15436–15450. doi: 10.1021/bi0204185
Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & J. R., Haak (1984). Molecular dynamics with coupling to an external bath. Journal of Chemical Physics, 81, 3684–3690.
Berhanu, W. M., & Masunov, A. E. (2011). Can molecular dynamics simulations assist in design of specific inhibitors and imaging agents of amyloid aggregation? Structure, stability and free energy predictions for amyloid oligomers of VQIVYK, MVGGVV and LYQLEN. Journal of Molecular Modeling, 17, 2423–2442.
Berhanu, W. M., & Masunov, A. E. (2012). Controlling the aggregation and rate of release in order to improve insulin formulation: molecular dynamics study of full-length insulin amyloid oligomer models. Journal of Molecular Modeling, 18, 1129–1142.
Bhak, G., Choe, Y. J., & Paik, S. R. (2009). Mechanism of amyloidogenesis: Nucleation-dependent fibrillation versus double-concerted fibrillation. Bmb Reports, 42, 541–551.
Cao, P., Marek, P., Noor, H., Patsalo, V., Tu, L. H., Wang, H., ... Raleigh, D. P. (2013). Islet amyloid: From fundamental biophysics to mechanisms of cytotoxicity. FEBS Letters, 587, 1106–1118. doi: 10.1016/j.febslet.2013.01.046

Full length amylin oligomer aggregation

Case TAD, D. A., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., ... Kollman, P. A. (2010). Amber 10. San Francisco, CA: University of California.
Cheng, P. N., Pham, J. D., & Nowick, J. S. (2013). The Supramolecular chemistry of beta-sheets. Journal of the American Chemical Society, 135, 5477–5492. doi: 10.1021/ja3088407
Cooper, G. J. S., Willis, A. C., Clark, A., Turner, R. C., Sim, R. B., & Reid, K. B. M. (1987). Purification and characterization of a peptide from amyloid-rich pancreases of type-2 diabetic patients. Proceedings of the National Academy of Sciences of the United States of America, 84, 8628–8632.
Darden, T., York, D., & Pedersen, L. (1993). Particle mesh ewald – an n.log(n) method for ewald sums in large systems. Journal of Chemical Physics, 98, 10089–10092.
Demurn, A., Mina, E., Kayed, R., Milton, S. C., Parker, L., & Glabe, C. G. (2005). Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers. Journal of Biological Chemistry, 280, 17294–17300. doi: 10.1074/jbc.M500997200
Dessalew, N., & Mikre, W. (2008). On the paradigm shift towards multitarget selective drug design. Current Computer-Aided Drug Design, 4, 76–90.
Fox, A., Snolwaerts, T., Casanova, C. E., Calciano, A., Nogaj, L. A., & Moffet, D. A. (2010). Selection for nonamyloidogenic mutants of islet amyloid polypeptide (IAPP) identifies an expanded region for amyloidogenicity. Biochemistry, 49, 7783–7789. doi: 10.1021/bi100337p
Gohlke, H., & Case, D. A. (2004). Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. Journal of Computational Chemistry, 25, 238–250. doi: 10.1002/jcc.10379
Han, H. Y., Weinreb, P. H., & Lansbury, P. T. (1995). The core alzheimers peptide nac forms amyloid fibrils which seed and are seeded by beta-amyloid – is nac a common trigger or target in neurodegenerative disease. Chemistry & Biology, 2, 163–169.
Harper, J. D., & Lansbury, P. T. (1997). Models of amyloid seeding in Alzheimer’s disease and scrapie: Mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. Annual Review of Biochemistry, 66, 385–407.
Hawkes, C. A., Ng, V., & McLaurin, J. (2009). Small molecule inhibitors of a beta-aggregation and neurotoxicity. Drug Development Research, 70, 111–124. doi: 10.1002/ddr.20290
Horn, A. H. C., & Sticht, H. (2010). Amyloid-beta 42 Oligomer Structures from Fibils: A Systematic Molecular Dynamics Study. Journal of Physical Chemistry B, 114, 2219–2226. doi: 10.1021/jp100023q
Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. Journal of Molecular Graphics, 14, 33.
Jayasinghe, S. A., & Langen, R. (2004). Identifying structural features of fibrillar islet amyloid polypeptide using site-directed spin labeling. Journal of Biological Chemistry, 279, 48420–48425. doi: 10.1074/jbc.M406853200
Sasahara, K., Hall, D., & Hamada, D. (2010). Effect of lipid type on the binding of lipid vesicles to islet amyloid polypeptide amyloid fibrils. [Article]. Biochemistry, 49, 3040–3048. doi: 10.1021/bi0919252

Sasahara, K., Naiki, H., & Goto, Y. (2005). Kinetically controlled thermal response of beta(2)-microglobulin amyloid fibrils. [Article]. Journal of Molecular Biology, 352, 700–711. doi: 10.1016/j.jmb.2005.07.033

Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., … Eisenberg, D. (2007). Atomic structures of amyloid cross-beta spines reveal varied steric zippers. [Article]. Nature, 447, 453–457. doi: 10.1038/nature05695

Selkoe, D. J. (2004). Folding proteins in fatal ways (Vol. 426, Article). Proceedings of the National Academy of Sciences of the United States of America, 101, 15075–15080. doi: 10.1073/pnas.0408388101

Sotomayor, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Serio, T. R., & Betsholtz, C. (1990). Islet amyloid polypeptide fragment and its analogs. [Article]. Protein Science, 17, 1467–1474. doi: 10.1002/pro.145

Staunton, D., Porat, Y., Gazit, E., & Nussinov, R. (2004). Peptide sequence and amyloid formation: Molecular simulations and experimental study of a human islet amyloid polypeptide fragment and its analogs. [Article]. Structure, 12, 439–455. doi: 10.1016/j.str.2004.02.002

Zheng, J., Jang, H., Ma, B., Tsai, C. J., & Nussinov, R. (2007). Modeling the Alzheimer A beta(17–42) fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities. [Article]. Biophysical Journal, 93, 3046–3057. doi: 10.1529/biophysj.107.110700

Zoete, V., Meuwly, M., & Karplus, M. (2005). Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. [Article]. Protein Structure Function and Bioinformatics, 61, 79–93. doi: 10.1002/prot.20528

Zraika, S., Hull, R. L., Verchere, C. B., Clark, A., Potter, K. J., Fraser, P. E., … Kahn, S. E. (2010). Toxic oligomers and islet beta cell death: guilty by association or convicted by circumstantial evidence?. Diabetologia, 53, 1046–1056. doi: 10.1007/s00125-010-1671-6