J Mol Model (2011) 17:2423–2442
DOI 10.1007/s00894-010-0912-4

ORIGINAL PAPER

# 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

**Workalemahu Mikre Berhanu · Artëm E. Masunov**

Received: 4 February 2010 / Accepted: 21 November 2010 / Published online: 21 December 2010


**Abstract** The aggregation modes of hexapeptide fragments of Tau, Insulin and Aβ peptide (VQIVYK, MVGGVV and LYQLEN) were found from their microcrystalline structures that had been recently resolved by X-ray analysis. The atomic structures reveal a dry self-complementary interface between the neighboring β-sheet layers, termed "steric zipper". In this study we perform several all-atom molecular dynamics simulations with explicit water to analyze stability of the crystalline fragments of 2-10 hexapeptides each and their analogs with single glycine replacement mutations to investigate the structural stability, aggregation behavior and thermodynamic of the amyloid oligomers. Upon comparing single and double layer models, our results reveal that additional strands contribute significantly to the structural stability of the peptide oligomers for double layer model, while in the case of single layer model the stability decreases (or remains the same in the case of LYQLEN). This is in agreement with the previous studies performed on different types of amyloid models. We also replaced the side-chains participating in the steric zipper interfaces with glycine. None of the mutants were structurally stable compared to the respective wild type model, except for mutants V2G and V6G in MVGGVV2 case. The exception can be explained by structural features of this particular polymorph. The double layer decamer and dodecamer aggregates of the wild type hexapeptides appear to be stable at 300K, which is confirmed by the conservation of high anti-parallel β-sheet content throughout the whole simulation time. Deletions of the side chains resulted in decline of secondary structure content compared to corresponding wild type indicating that the role of the replaced amino acid in stabilizing the structure. Detailed analysis of the binding energy reveals that stability of these peptide aggregates is determined mainly by the van der Waals and hydrophobic forces that can serve as quantitative measure of shape complementarities between the side chains. This observation implies that interactions among side chains forming the dehydrated steric zipper, rather than among those exposed to water, are the major structural determinant. The electrostatic repulsion destabilizes the studied double layer aggregates in two cases, while stabilizes the other two. Negative total binding free energy indicates that both wild type and mutants complex formation is favorable. However, the mutants complexation is less favorable than the wild type's. The present study provides the atomic level understanding of the aggregation behavior and the driving force for the amyloid aggregates, and could be useful for rational design of amyloid inhibitors and amyloid-specific biomarkers for diagnostic purposes.

**Keywords** Aggregation · Amyloid fibril · β sheet · Binding free energy · Cross-β structure · MM-PBSA · Molecular dynamic simulation · Oligomer · Secondary structure · Steric zipper

W. M. Berhanu
NanoScience Technology Center and Department of Chemistry,
University of Central Florida,
Orlando, FL 32826, USA

A. E. Masunov (✉)
NanoScience Technology Center, Department of Chemistry
and Department of Physics, University of Central Florida,
Orlando, FL 32826, USA
e-mail: amasunov@mail.ucf.edu

## Introduction

Aggregation of polypeptide chains and formation of amyloid fibrils are associated with the development of a number of disorders, including Alzheimer's, Parkinson's, type II diabetes, and Creutzfeldt-Jakob disease [1]. Amyloid deposits develop when proteins misfold out of their native conformations and aggregate into insoluble fibrils [2]. The amyloid fibrils share a sequence independent structure characterized by cross-$\beta$ spine structural motif in which protein $\beta$-strands run orthogonal to the fibril axis and repetitive hydrogen bonding extends parallel to the axis [3, 4]. This cross-$\beta$ spine may correspond to the global minimum energy conformation for a wide variety of proteins [3]. Identifying this structural motif in small model peptide systems and characterizing it under different conditions can yield valuable clues about the molecular-level details of amyloid formation. Recently, the micro-crystal structures of several amyloidogenic peptides have been determined by x-ray crystallography [5–7]. These high resolution structures provided researchers with a unique opportunity to understand the structural details and on the factors that destabilize/stabilize the amyloid fibrils. Molecular dynamics (MD) simulations, along with other theoretical approaches, based on these crystal structures, can often present significant contribution to this understanding [8–17]. By selecting an amyloid oligomer out of the crystal structure and evaluating its conformational stability in a crystal-free environment, these investigations have provided insights into the intrinsic propensities of peptide fragments to associate in amyloid-like states, the energetic factors stabilizing these aggregates, and the possible aggregation states of oligomeric precursors or larger assemblies up to 128 $\beta$-strands.

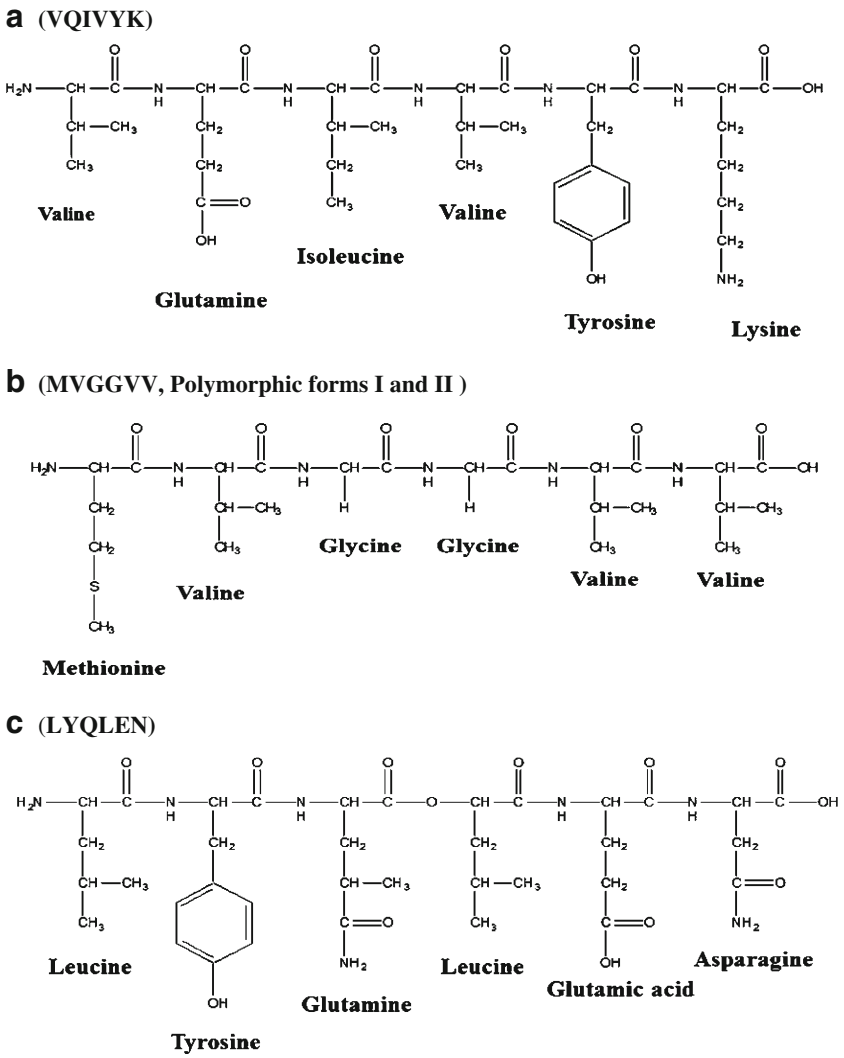
One of the common structural features, observed in many available X-ray structures of amyloidogenic polypeptides is pairing of the $\beta$-sheets by interdigitated side chains in a dry 'steric zipper'. It is worth noting, that most of the theoretical investigations have been conducted on the systems where steric zipper interface is composed of the large polar and/or aromatic side chains. In this contribution we focus on aggregates stabilized by steric zipper interfaces formed by small hydrophobic residues (VQIVYK, MVGGVV) (Scheme 1). A system with polar H-bonding side chains (LYQLEN) (Scheme 1) is also considered for comparison. We perform all-atom MD simulations with explicit solvent on both wild type and mutant polypeptides at various degrees of aggregation. The initial structure of the aggregates is based high resolution X-ray study [5]. The MVGGVV peptide represent the fragment (residues 35-40) from the C terminal of the $A\beta_{1\text{-}40}$ peptide, associated with Alzheimer's disease [18, 19]. The VQIVYK is a fragment (residues 306-311) of the Tau protein, which is also involved in the pathogenesis of Alzheimer's disease. LYQLEN peptide is a fragment of a chain (residues 13-18) of Insulin that had been shown to form amyloid-like fibrils [5].

In Alzheimer's disease (AD), the Tau protein forms intracellular amyloid tangles in neurons.[20, 21] The hexapeptide VQIVYK models the key amyloidogenic peptide sequence and forms amyloid-like fibrils with the same cross-$\beta$ structure found in full Tau amyloid fibrils [5]. The structural organization of VQIVYK is a parallel β-strand within the same $\beta$-sheet layer while maintaining anti-parallel organization between the adjacent $\beta$-sheet layers [5]. At the dry interface between the adjacent β-sheet layers, the shape complementarity is formed by the hydrpbobic steric zipper via the side chains of Val1, Ile3 and Tyr5 (Fig. 1a) [5], packing against each other forming the sheet-sheet interface.

Aggregation of Aβ peptides, which are the natural products of cellular proteolytic, is also linked to Alzheimer's disease (AD). The most abundant Aβ species are 40 residue peptides ($A\beta_{1\text{-}40}$). The MVGGVV peptide is a fragment (residue 35-40) from the C terminal of the $A\beta_{1\text{-}40}$ consists of parallel and anti-parallel $\beta$-strands within the same $\beta$-sheet layers. At the dry interface between the adjacent $\beta$-sheet layers, the shape complementarily is formed by the hydrophobic steric zipper via the side chains of Met1, Val2 and Val5 (Fig. 1b, c) [5].

Fibrils of Insulin are observed extracellular in the rare medical condition termed injection amyloidosis. These Insulin fibrils formed in vivo display the defining characteristics of amyloid aggregates such as binding the dye Congo red [22] and the cross-$\beta$ X-ray diffraction pattern [23]. Both A chain and B chain can form fibrils on their own [24, 25], and seeds of A chain or B chain can nucleate the fibrillation of full length Insulin [24]. The atomic-resolution picture of the interactions between segments of Insulin which may be part of fibrillar spine came from crystal structures of the fibril forming peptide segments LYQLEN (residues A13–A18) and VEALYL (residues B12–B17) [5]. The structural organization of LYQLEN is anti-parallel $\beta$-strands within the same $\beta$–sheet layer while maintaining parallel organization between the adjacent $\beta$-sheet layers [5]. At the dry interface between the adjacent β-sheet layers, the shape complementarity is formed by the polar side chain steric zipper (Tyr2, Gln3, Leu4 and Asn6) and side chain H-bonding (Fig. 1b) [5]. Recently, serum samples from patients with Parkinson's disease have been found to display an autoimmune response to Insulin oligomers and fibrils [26], possibly indicating the presence of Insulin aggregates in this disease as well. Insulin also reported to form amyloid-like fibrils in vitro under elevated temperatures, low pH, and increased ionic strength [27, 28]. This fibril formation has been a limiting factor in long-term storage of Insulin for treatment of diabetes. Thus, better

**Scheme 1** Structure of peptides studied in this work. **a** (VQIVYK). **b** (MVGGVV, Polymorphic forms I and II). **c** (LYQLEN)




understanding of Insulin fibrillation could lead to safer handling and more cost-effective storage of Insulin.

Previous theoretical study has demonstrated the significant role of steric zipper in the structural stability of the GNNQQNY and GGVVIA oligomers stabilized with polar side chain and H-bonding [16, 29]. Park et al. [8] address the structural selection mechanism of different double layer peptides including GNNQQNY, NNQQ, VEALYL, KLVFFAE and STVIIE, and find that the patterns with the lowest binding free energy correspond to X-Ray structures with high accuracy. The main contribution of the binding free energy of the double layer pattern is determined by the van der Waals and hydrophobic forces. These contributions can therefore serve as a quantitative measure of shape complementarity among side chains between the $\beta$-sheets. The steric self-complementary (known as steric zipper) selects the most stable packing modes. It also makes parallel $\beta$-sheets generally preferred over anti-parallel ones. The presence of charged side chains appears to give anti-parallel $\beta$-sheets kinetic preference at the early stages of assembly, while the double layer formation is likely to be thermodynamically controlled. Xu et al. [30] investigated the β-sheets composed of seven antiparallel decapeptides, representing the 20–29 segment of human Islet amyloid polypeptide (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. This provides a new insight into the assembly pathway at the early stage of fibril elongation. Based on the Poisson–Boltzmann free energy analysis and quasiharmonic configuration entropy estimation, the entropic contribution was found to play an important role in the longitudinal assembly. Moreover, a possible oligomeric state with cyclic form was suggested based on one assembly model found in the simulations. This evidenced the polymorphic nature of the amyloidogenic oligomerization and possible mechanism of its toxicity. The cyclic structures of amyloid oligomers have been reported to be the early intermediates in solution,

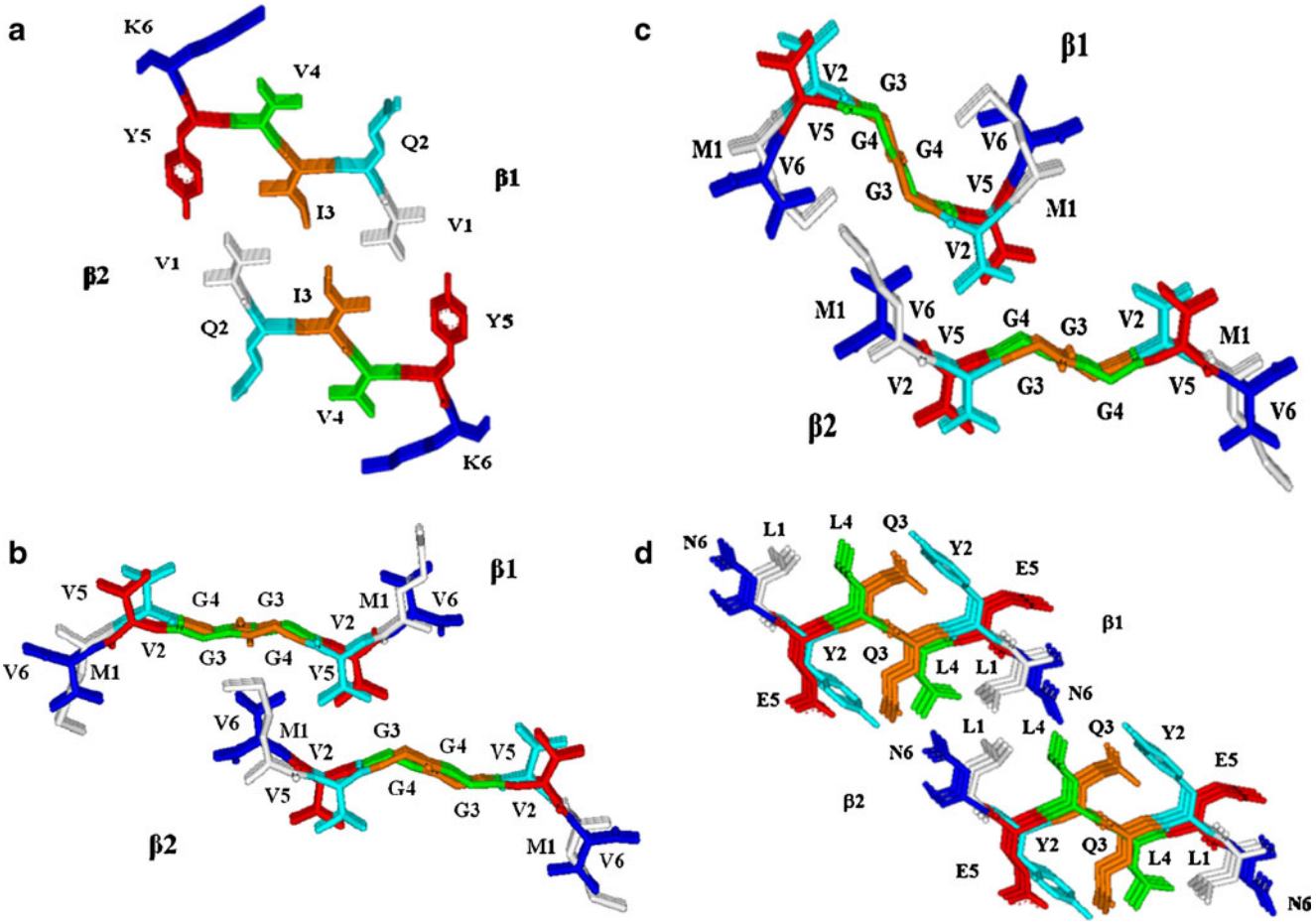



**Fig. 1** **a** The atomic representation of the VQIVYK oligomer; sandwich of 2 beta sheets formed by 5 hexapeptides each. Inter-sheet steric zipper is formed between the side chain of the residues Val1, Ile3 and Tyr5 of the β1 sheet and those of the same residue of β2 sheet. **b** The atomic representation of the MVGGVV1 (polymorph form I) oligomer; sandwich of 2 beta sheets formed by 6 hexapeptides each. Inter-sheet steric zipper is formed between the side chain of the residues Met1, Val2 and Val5 of the β1 sheet and those of the same residue of β2 sheet. **c** The atomic representation of the MVGGVV2 (polymorph form II) oligomer; sandwich of 2 beta sheets formed by 6 hexapeptides each. Inter-sheet steric zipper is formed between the side chain of the residues Met1, Val2 and Val5 of the β1 sheet and those of the same residue of β2 sheet. **d** The atomic representation of the LYQLEN oligomer; sandwich of 2 beta sheets formed by 6 hexapeptides each. Intersheet steric zipper is formed between the side chain of the residues Leu1, Tyr2, Gln3, Leu4, and Asn6 of the β1 sheet and those of the same residue of β2 sheet

capable to form ion-channel-like structures in the membrane that could be responsible for pathologic membrane permeability and destabilization of the cellular ionic homeostasis [31, 32]

Vitagliano et al. [33] in their molecular dynamics simulation characterizing assemblies formed by steric zipper assemblies composed of a pair of 10-stranded β-sheets of the peptides SSTSAA and VQIVYK show high fluctuations and significant distortion. The analysis of the VQIVYK crystal packing reveals two different double layers with significant interface area and surface complementarity [5]. One is characterized by nonpolar dry interface made up essentially by the side chains of Va11 and I1e3 of the two layers, while the other is polar and involves Tyr5 and Gln2 side chains [5]. The nonpolar interface exhibits larger values of the surface area (113 vs. 89 $Å^2$), but slightly lower surface complementarity (0.76 *vs.* 0.82) [5]. The stability of the nonpolar hydrophobic interfaces was studied by Vitagliano et al. [33] in their MD simulations. They report high fluctuations and significant distortion (RMSD *c.a.* 10 Å within 40 ns simulation) when investigate three layer assemblies formed by steric zipper and composed of a pair of 10-stranded β-sheets of the peptides VQIVYK. In the contrast, they found RMSD below 6 Å within 40 ns simulation, when study the 10-stranded double layer with nonpolar interface. Hence, the stability of the nonpolar interface is system dependent.

However, the atomic information for the early stage of the aggregation mechanism of the VQIVYK, MVGGVV and LYQLEN peptide is still limited so far. Thus, understanding the structural stability and aggregation behavior of the VQIVYK, MVGGVV and LYQLEN

**Table 1** Summary of the **VQIVYK** oligomeric models and simulation system

| Model | Systems | Sheet/strand organization | Simulation box size (A˚) | Simulation time (ns) | T(K) |
|---|---|---|---|---|---|
| Wilde type | | | | | |
| A1 *(Sh1-St2)* | sheet1, strands2 | ——/parallel | 49.31×49.31×49.31 | 10 | 300 |
| A2 (*Sh1-St3*) | sheet1, strands3 | ——/parallel | 50.84×50.84×50.84 | 10 | 300 |
| A3 (*Sh1-St4*) | sheet1, strands4 | ——/parallel | 51.67×51.67×51.67 | 10 | 300 |
| A4 (*Sh1-St5*) | sheet1, strands5 | ——/parallel | 54.06×54.06×54.06 | 10 | 300 |
| A5 (*Sh2-St2*) | sheet2, strands2 | Antiparallel / Parallel | 60.97×60.97×60.97 | 10 | 300 |
| A6 (*Sh2-St3*) | sheet2, strands3 | Antiparallel / Parallel | 63.36×63.36×63.36 | 10 | 300 |
| A7 (*Sh2-St4*) | sheet2, strands4 | Antiparallel / Parallel | 64.84×64.84×64.84 | 10 | 300 |
| A8 (*Sh2-St5)* | sheet2, strands5 | Antiparallel / Parallel | 65.88×65.88×65.88 | 10 | 300 |
| VQIVYK (*Sh2-St5*) | sheet2, strands5 | Antiparallel / Parallel | 65.88×65.88×65.88 | 10 | 300 |
| Mutants | | | | | |
| B1 | sheet2, strands5, V1G | Antiparallel / Parallel | 65.74×65.74×65.74 | 10 | 300 |
| B2 | sheet2, strands5, I3G | Antiparallel / Parallel | 65.71×65.71×65.71 | 10 | 300 |
| B3 | sheet2, strands5, Y5G | Antiparallel / Parallel | 65.26×65.26×65.26 | 10 | 300 |

peptide is expected to provide knowledge for designing an inhibitor aimed to decrease the self-aggregation into fibrils.

In this study, several all-atom MD simulations with explicit water at 300 K were conducted to investigate the structural stability, aggregation behavior and thermodynamics of the VQIVYK, MVGGVV and LYQLEN peptides with various sizes and its single glycine replacement mutations. Our aim is to elucidate: (i) the influence of the number of the peptides on the structural stability and conformational dynamics of the oligomers; (ii) the possible minimal nucleus seed for the fibril formation of the peptides; (iii) the principle driving force for the association of the peptides; and (iv) the effects of single glycine replacement mutations on the structural stability of the oligomers. The results of this study may provide insight into the possible mechanism of fibrillogenesis of the amyloid peptides. It may also be helpful for designing new or modified capping peptides capable of breaking the driving force for aggregations and preventing the fibril formation of the peptides.

## Computational details

The crystal structure of the VQIVYK, MVGGVV and LYQLEN had been determined by Sawaya et al. [5]. The atomic coordinates of the multiple unit cells were taken from the website [34], and the water molecules from the crystal structure were removed. The MVGGVV have two different polymorphic forms (form 1 and 2 with resolution of 2.0 Å

**Table 2** Summary of the **MVGGVV1, polymorphic form I,** oligomer models and simulation system

| Model | Systems | Sheet/strand organization | Simulation box size (A˚) | Simulation time (ns) | T(K) |
|---|---|---|---|---|---|
| Wilde type | | | | | |
| C1 (*Sh1-St2*) | sheet1, strands2 | ——/antiparallel | 52.59×52.59×52.59 | 10 | 300 |
| C2 (*Sh1-St3*) | sheet1, strands3 | ——/antiparallel | 53.46×53.46×53.46 | 10 | 300 |
| C3 (*Sh1-St4*) | sheet1, strands4 | ——/antiparallel | 54.14×54.14×54.14 | 10 | 300 |
| C4 (*Sh1-St5*) | sheet1, strands 5 | ——/antiparallel | 56.05×56.05×56.05 | 10 | 300 |
| C5 (*Sh2-St2*) | sheet2, strands2 | Antiparallel / Antiparallel | 63.02×63.02×63.02 | 10 | 300 |
| C6 (*Sh2-St3*) | sheet2, strands3 | Antiparallel /Antiparallel | 63.17×63.17×63.17 | 10 | 300 |
| C7 (*Sh2-St4*) | sheet2, strands4 | Antiparallel /Antiparallel | 66.36×66.36×66.36 | 10 | 300 |
| C8 (*Sh2-St5*) | sheet2, strands5 | Antiparallel / Antiparallel | 67.13×67.13×67.13 | 10 | 300 |
| MVGGVV1 (*Sh2-St6*) | sheet2, strands6 | Antiparallel / Antiparallel | 69.17×69.17×69.17 | 10 | 300 |
| Mutants | | | | | |
| D1 | sheet2, strands6, M1G | Antiparallel / Antiparallel | 69.16×69.16×69.16 | 10 | 300 |
| D2 | sheet2, strands6, V2G | Antiparallel / Antiparallel | 69.02×69.02×69.02 | 10 | 300 |
| D3 | sheet2, strands6, V5G | Antiparallel / Antiparallel | 69.45×69.45×69.45 | 10 | 300 |
| D4 | sheet2, strands6, V6G | Antiparallel / Antiparallel | 68.78×68.78×68.78 | 10 | 300 |

**Table 3** Summary of the **MVGGVV polymorphic form II** models (wild type and mutant) and simulation system

| Model | Systems | Sheet/strand organization | Simulation box size (A˚ ) | Simulation time (ns) | T(K) |
|---|---|---|---|---|---|
| Wilde type | | | | | |
| E1 (*Sh2-St2*) | sheet2, strands2 | Antiparallel /Antiparallel | 60.58×60.58×60.58 | 10 | 300 |
| E2 (*Sh2-St3*) | sheet2, strands3 | Antiparallel /Antiparallel | 61.70×61.70×61.70 | 10 | 300 |
| E3 (*Sh2-St3*) | sheet2, strands4 | Antiparallel /Antiparallel | 63.24×63.24×63.24 | 10 | 300 |
| E4 (*Sh2-St5*) | sheet2, strands5 | Antiparallel /Antiparallel | 64.50×64.50×64.50 | 10 | 300 |
| MVGGVV2 (*Sh2-St6*) | sheet2, strands6 | Antiparallel /Antiparallel | 66.83×66.83×68.83 | 10 | 300 |
| Mutants | | | | | |
| F1 | sheet2, strands6, M1G | Antiparallel /Antiparallel | 64.15×64.15×64.15 | 10 | 300 |
| F2 | sheet2, strands6, V2G | Antiparallel /Antiparallel | 68.86×68.86×68.86 | 10 | 300 |
| F3 | sheet2, strands6, V5G | Antiparallel /Antiparallel | 66.40×66.40×66.40 | 10 | 300 |
| F4 | sheet2, strands6, V6G | Antiparallel /Antiparallel | 66.25×66.25×66.25 | 10 | 300 |

and 1.8 Å) both of which were used in the simulation.The Sirius visualization program from San Diego Supercomputer Center (http://sirius.sdsc.edu) was used to construct the aggregates of various sizes. The initial geometry of the largest aggregate was taken as a pair of $\beta$-sheets composed of 6 strands (5 strands for VQIVYK), it is shown in Fig. 1. In the following we denote the aggregates ShN-StM, where N is the number of $\beta$-sheets, and M is the number of strands per $\beta$-sheet. The initial geometry of the largest wild type aggregate was taken as a pair of $\beta$-sheets composed of 6 strands (MVGGVV and LYQLEN) and 5 strands (VQIVYK), as shown in Fig. 1. For the smaller size wild type systems, the initial structures of oligomers were obtained by removing the β-strands one by one from the Sh2-St5 (VQIVYK) or Sh2-St6 (MVGGVV and LYQLEN) models. To construct the mutant systems, several glycine replacements were made in the wild type aggregate. The mutants are denoted as XnG, where X is the replaced residue, G is the glycine replacing this residue, and n is its position in the peptide sequence. Three or four mutants were designed for each peptide (V1G, I3G, and Y6G for VQIVYK; M1G, V2G, V5G, and V6G for MVGGVV; V2G, Q3G, L4G and V6G for LYQLEN). The simulation details for each model are summarized in Tables 1, 2, 3, and 4.

MD simulations were performed with the AMBER 10 software package [35]. The models were immersed in truncated octahedron filled with water molecules. The periodic water box was constructed in such a way that the solute was at least ~10 Å away from the box boundary and the minimum distance between the solute and its image was ~20 Å. When required, ions were added to neutralize the simulated system. The AMBER ff99SB force field and the

**Table 4** Summary of the **LYQLEN** oligomeric models and simulation system

| Model | Systems | Sheet/strand organization | Simulation box size (A˚ ) | Simulation time (ns) | T(K) |
|---|---|---|---|---|---|
| Wilde type | | | | | |
| G1(*Sh1-St2*) | sheet1, strands2 | ——/antiparallel | 50.43×50.43×50.43 | 10 | 300 |
| G2(*Sh1-St3*) | sheet1, strands3 | ——/antiparallel | 51.95×51.95×51.95 | 10 | 300 |
| G3(*Sh1-St42*) | sheet1,strands 4 | ——/antiparallel | 51.93×51.93×51.93 | 10 | 300 |
| G4(*Sh1-St5*) | sheet1, strands5 | ——/antiparallel | 55.75×55.75×55.75 | 10 | 300 |
| G5(*Sh2-St2*) | sheet2, strands2 | Antiparallel /Antiparallel | 65.67×65.67×65.67 | 10 | 300 |
| G6(*Sh2-St3*) | sheet2, strands3 | Antiparallel /Antiparallel | 66.97×66.97×66.97 | 10 | 300 |
| G7(*Sh2-St4*) | sheet2, strands4 | Antiparallel /Antiparallel | 68.59×68.59×68.59 | 10 | 300 |
| G8(*Sh2-St5*) | sheet2, strands5 | Antiparallel /Antiparallel | 69.82×69.82×69.82 | 10 | 300 |
| LYQLEN(*Sh2-St6*) | sheet2, strands6 | Antiparallel /Antiparallel | 70.46×70.46×70.46 | 10 | 300 |
| Mutants | | | | | |
| Y2G | sheet2, strands6, Y2G | Antiparallel /Antiparallel | 70.04×70.04×70.04 | 10 | 300 |
| Q3G | sheet2, strands6, Q3G | Antiparallel /Antiparallel | 70.35×70.35×70.35 | 10 | 300 |
| L4G | sheet2, strands6, L4G | Antiparallel /Antiparallel | 70.23×70.23×70.23 | 10 | 300 |
| N6G | sheet2, strands6, V6G | Antiparallel /Antiparallel | 69.35×69.35×69.35 | 10 | 300 |

TRIP 3 water model were used in the simulations. For all systems, in order to relax bond geometries, the potential energy of the system (peptides and water) was minimized by using the steepest-descent method until convergence was reached. The electrostatic interactions were calculated using Particle Mesh Ewald (PME) method [36] with a cutoff of 12 Å. The cutoff radius for the Lennard-Jones interactions was also set to 12 Å. At the end of the equilibration simulation, energy and temperature of the systems were stable. In order to constrain all bonds connecting hydrogen atoms, and to allow a larger time step of 2.0 fs, the SHAKE algorithm [37] was applied. SLeAP module from the AMBER package was used to finalize the compounds for calculations. Initially, each system in the explicit solvent model underwent a two stage minimization process. The first stage involved an initial 1000 iterations (500 steepest descent, 500 conjugate gradient) minimization of the solvent holding the protein constrained, followed by a 2500 iteration (1000 steepest descent, 1500 conjugate gradient) minimization of the entire system. After minimizations were complete, the systems were warmed up from 0 to 300 K at constant volume using weak solute restraints for 20 ps. The explicit solvent model systems were then subjected to constant pressure dynamics at 1 atm for 200 ps to adjust the solvent density to ~1 g/cm$^3$. The main system properties were analyzed to verify successful equilibration for all systems and then constant volume molecular dynamics simulations were performed for all systems. All the production simulations were carried out without restraints for 10 ns. The temperature of the system was regulated by the Langevin thermostat. The molecular dynamic trajectories were saved every 1000 steps (2 ps interval) for subsequent analysis. Hydrogen bond occupancies and structure RMSDs was calculated using PTRAJ module available within AMBER. Secondary structure analysis was performed using the Dictionary of Secondary Structure of Protein (DSSP) software [38]. The simulation results were visualized suing VMD 1.8.7 [39]. The MM-PBSA single trajectory approach implemented as script in AMBER 10 [35] was used to calculate the steric zippers binding energy for non-covalent association between the β-sheets within the double layer. In this approach an assumption is made that no significant conformational changes occur upon binding, i.e., structural adaptation is negligible and the snapshots for all three species can be obtained from the single trajectory of the complex by separating it into two components. To calculate binding free energies in MM-PBSA method, the explicit water simulations were used to generate the trajectory followed by the implicit Poisson-Boltzman/surface area method to calculate solvation energy terms. The gas phase and the solvation free energies were calculated over 500 snapshots taken at 20 ps intervals from the last 8 ns of the MD trajectories.

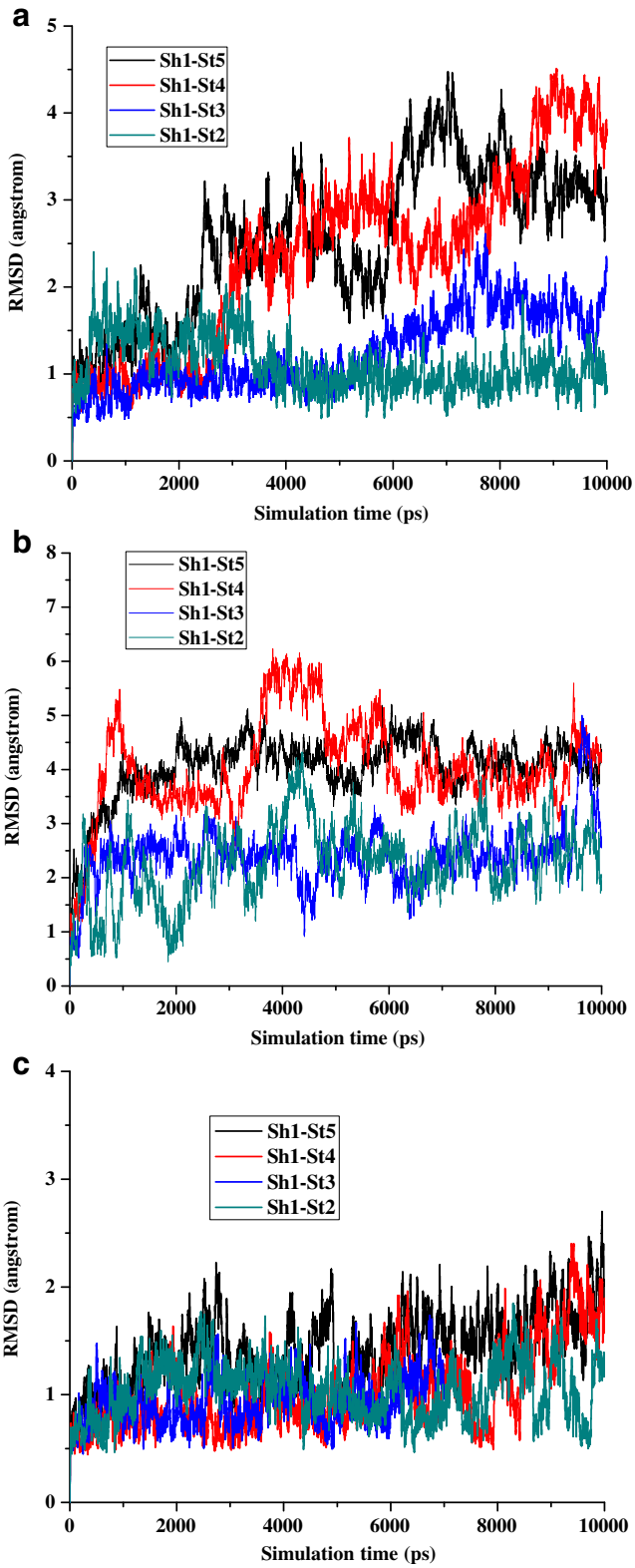



**Fig. 2** Time evolution of the $C_\alpha$-RMSD values with respect to initial structure of the wild type systems with one sheet and different strands within the sheet: (**a**) tau oligomer (VIQVYK), (**b**) Abeta amyloid I (MVGGVV1) polymorph form I and (**c**) Insulin amyloid (LYQLEN)

## Results and discussion

### Size dependent structural stability of the wild type peptides aggregates

Eight simulations of wild type VQIVYK were conducted for the aggregates build of one (models A1-A4) and two (models A5–A8) antiparallel β-sheets with parallel strands within each sheet. The relative stability of the model aggregates was measured by the backbone root mean-squared deviation (RMSD). The reference structure for calculating backbone RMSD was the energy-minimized structure. As one can see in Fig. 2a, for the model systems of A1 (Sh1-St2) and A2 (Sh1-St3), the RMSDs remained below 2.Å for 10 ns, while for A3 (Sh1-St4) and A4 (Sh1-St5) the RMSDs increased to 4.5 Å, indicating the lower relative instability of the one layer aggregate with larger number of strands. The larger two-layer model systems of A7 (Sh2-St4) and A8 (Sh2-St5), maintained RMSDs *c.a.* 4.0 Å within 10 ns, indicating relative stability of the structures compared to the smaller bilayer models A5 (Sh2-St2) and A6 (Sh2-St3), which showed large fluctuations up to 7.0 Å (Fig. 3a). The results of two-layer models suggested that the structural stability of the VQIVYK oligomers increases with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands, while for one-layer models the trend is opposite.

Our simulation for 5-stranded double layers of the wildtype VQIVYK oligomers was found to have a RMSD of 4 Å, in good agreement with the result reported by Vitagliano et al. [33]. The comparison of the RMSD values of the nonpolar interface models (VQIVYK and MVGGVV) with LYQLEN that has polar residues on the dry interface indicates the nonpolar are significantly less stable. The smaller RMSD values of the polar LYQLEN is in good agreement with the result of Zhang et al. [29], who found an RMSD of 2 Å by simulation of 4-stranded double layer GGNNQQNY, which has polar residues on the dry interlayer interface. Our results indicate that the polar dry interface significantly improves stability.

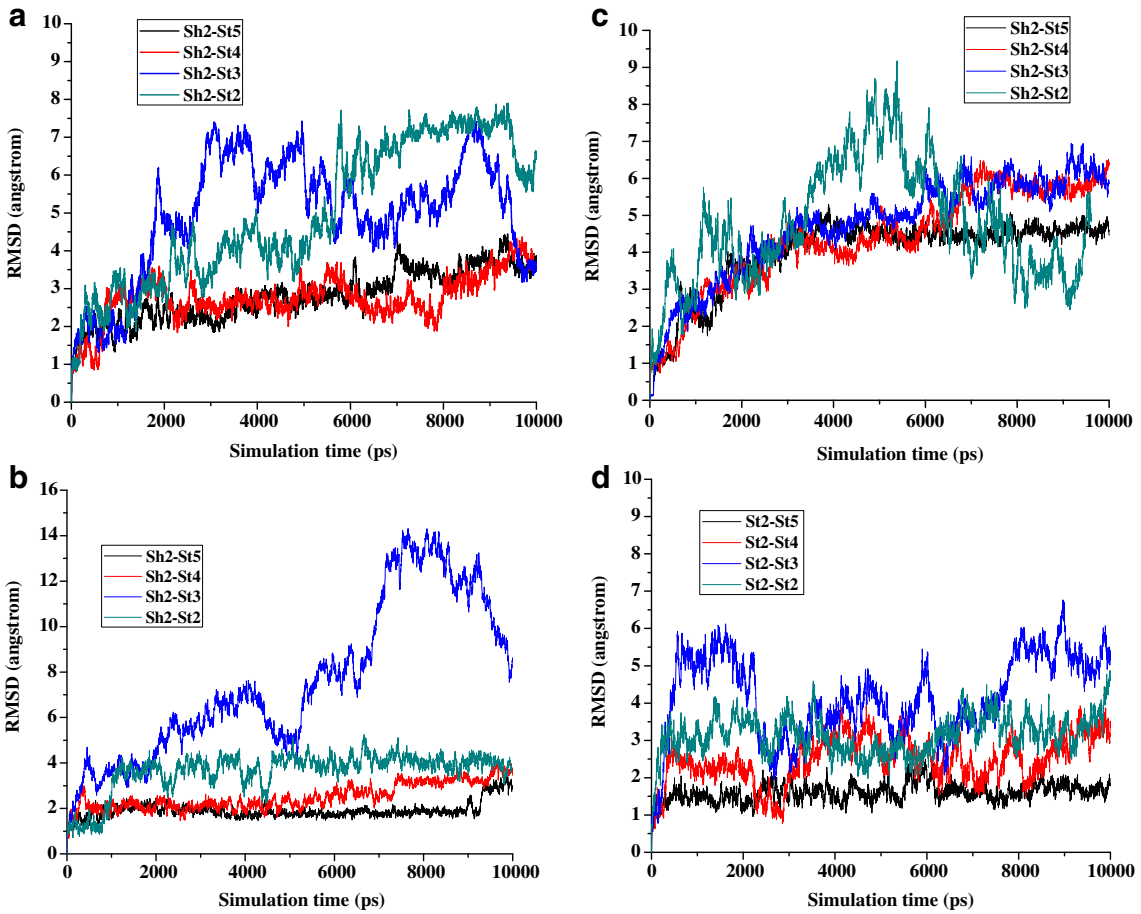



**Fig. 3** Time evolution of the $C_\alpha$-RMSD values with respect to initial structure of the wild type systems with two sheets and different strands within the sheets: (**a**) tau oligomer (VIQVYK), (**b**) Abeta amyloid I (MVGGVV1) polymorph form I, (**c**) Abeta amyloid II (MVGGVV2) polymorph form II and (**d**) Insulin amyloid (LYQLEN)

Another eight wild type simulations of MVGGVV1 (models C1–C8) were conducted for anti-parallel β-sheets with parallel strands within the sheets. As shown in Fig. 2b, for the model systems of C1 (Sh1-St2) and C2 (Sh1-St3), the RMSDs were below 4.50 Å within 10 ns. For C3 (Sh1-St4) and C4 (Sh1-St5) the RMSDs were maintained below 4.5 Å and 6 Å within 10 ns respectively, the two layer model systems of C7 (Sh2-St4) and C8 (Sh2-St5), the RMSDs were below 3.0 Å within 10 ns as shown in Fig. 3b. Aggregate C6 (Sh2-St2) maintained RMSDs below 4.0 and C5 (Sh2-St3) showed large fluctuations RMSD within the first 5 ns and then increased to 12 Å after 7 ns. Our results for one-layer models suggest that the structural stability of the MVGGVV1 oligomers increases as the number of stands decreases, while the results of two-layer models suggest that the structural stability of the MVGGVV1 oligomers increases remarkably with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands.

Four wild type MVGGVV2 peptide aggregates (two layer with different number of strands), simulations were conducted for antiparallel β-sheets with parallel strands within the sheets (models E1–E4). We did not do single layer simulation, assuming the result will be the same as for polymorphic form I. As shown in Fig. 2c, E2 (Sh2-St3) and E3 (Sh2-St4), the RMSDs were almost identical and remained below 6.0 Å within 10 ns. While E1 (Sh2-St2) showed a large fluctuations within 1 ns and remained c.a. 7.0 Å within 10 ns. E4 (Sh2-St5) shows the same RMSD change as E2 and E3 for the first 4 ns increasing to ~5.0 Å and maintained an RMSD ≤5.0 Å during the 10 ns simulation. The results of two-layer models suggested that the structural stability of the MVGGVV2 oligomers increases with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands.

Finally eight wild type LYQLEN peptide aggregates (models G1–G8) were considered for antiparallel β-sheets, parallel strands within the sheets. As shown in

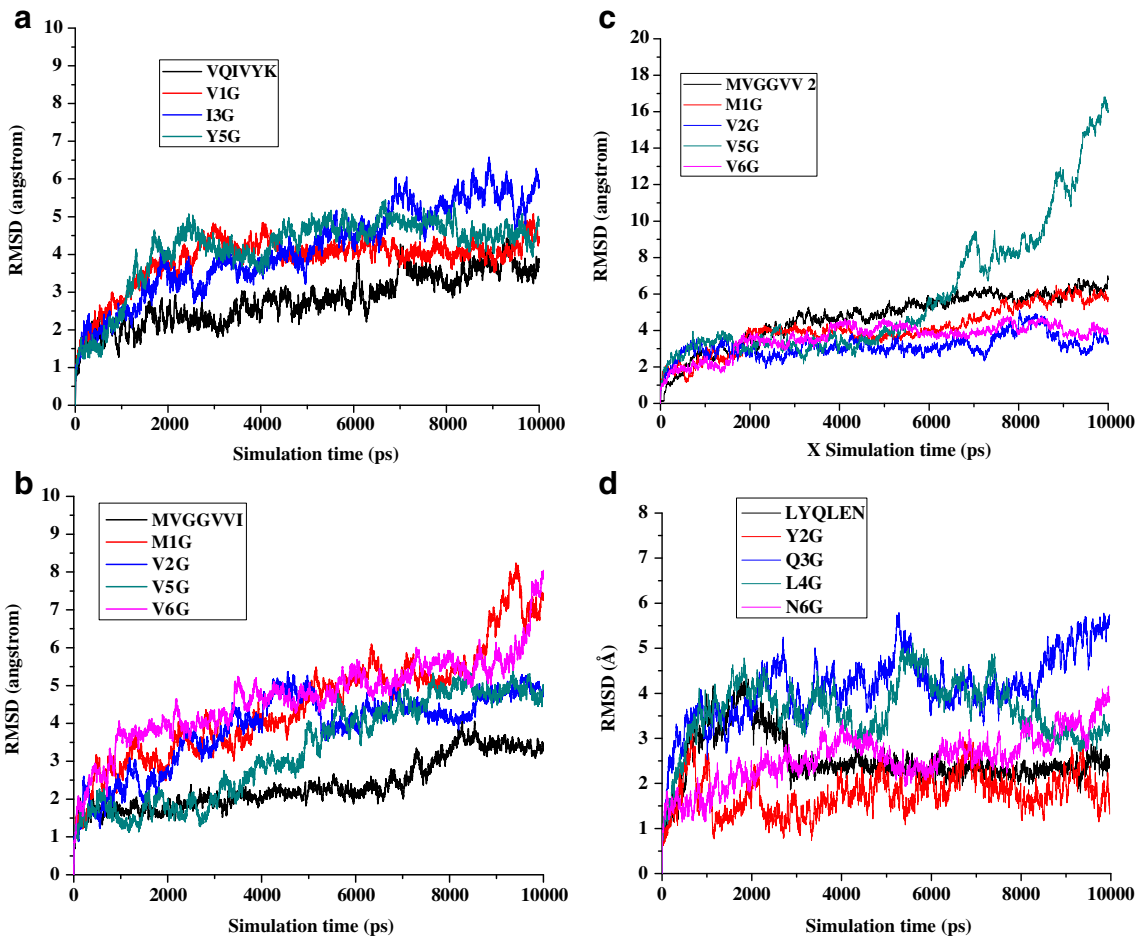



**Fig. 4** Time evolution of the $C_\alpha$-RMSD of the wild type systems of the studied amyloids oligomers and their mutants. (**a**) Tau oligomer **(Sh2-St5)** and its glycine mutants, (**b**) Aβ amyloid polymorph form I **(Sh2-St6)** and its glycine mutants, (**c**) Aβ amyloid polymorph form I **(Sh2-St6)** and its glycine mutants and (**d**) Insulin amyloid **(Sh2-St6)** and its glycine mutants

Fig. 2d, for the model systems G1 (Sh1-St2), G2 (Sh1-St3), G3 (Sh1-St4) and G4 (Sh1-St5) consisting of one layer and different number of strands, the RMSDs remained at 2.0 Å within 10 ns, indicating exceptional stability of these structures. Figure 3d shows that for the model systems G5 (Sh2-St2) and G7 (Sh2-St4), the RMSDs shown were maintained at ~4.50 Å, for G8 (Sh2-St5) RMSDs is c.a. 2.50 Å, and for G6 (Sh2-St3) RMSD demonstrates a large fluctuation within the first 4 ns and then stabilized at 7 Å after 8 ns, which indicated that they lost their original structural organization. The results of two-layer models suggested that the structural stability of the LYQLEN oligomers increases remarkably with increasing the numbers of β-strands, with four and five strands being the most stable. Our results for one-layer models suggested that the structural stability of the LYQLEN oligomers is the same irrespective of the number of strands. One layer with two, three, four and five stands (model G1-G4) are structurally stable with RMSD *c.a* 2.0 Å. The LYQLEN oligomers is stabilized with backbone to backbone and side chain hydrogen bonding while in the case of the VQIVYK , MVGGVV1 and MVGGVV 2 are stabilized with back bone to back to bone hydrogen bonding.

Comparing single and double layer models, our results reveal that the extra β-sheet strand contributes significantly to the structural stability of the VQIVYK, MVGGVV1 and MVGGVV2 oligomers for double layer model while in the case of single layer model it decreases. In the case of the LYQLEN our results also shows that an extra β-sheet strands contributes significantly to the structural stability of the LYQLEN oligomers for double layer models while in the case of single layer model it is almost the same irrespective of the number of strands. This is in agreement with previous studies done on different types of amyloid models. The above results all together suggest that the structural stability of the oligomers increases significantly with increasing the number of β-strands for double layer models implying extra sheet-sheet interactions are necessary for the formation of steric zipper to associate the

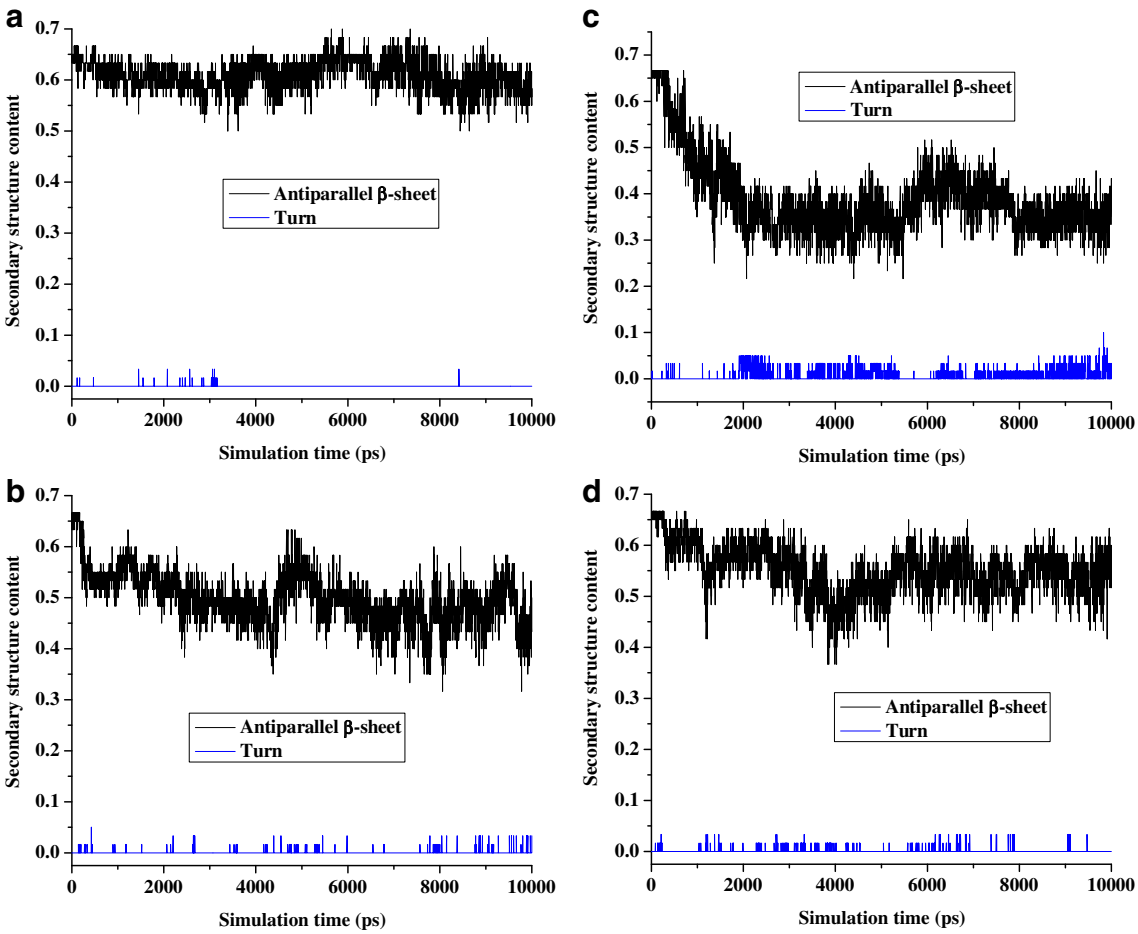



**Fig. 5** Time evolutions of the secondary structure contents for Tau amyloid oligomer Sh2-St5 amyloid aggregate formed by Tau amyloid fragment **VQIVYK** and its mutants (**a**) Wildtype, (**b**) Y5G, (**c**) I3G, and (**d**) V1G during the 10-ns MD simulations

strands, resulting in more stable oligomeric organizations. Our findings are in agreements with previous observations [16, 29] which indicates that the minimal nucleus seed for the amyloid fibril formation could be as small as three or four peptides.

The evolution of the root mean square deviations (RMSDs) between initial and the current trajectory structures indicates that the system undergoes certain rearrangement. The initial structures are taken from X-ray and may be stabilized by the crystal environment. However, the simulations are performed in the solution state. Due to this different environment, relatively large RMSDs may not always correspond to the unstable structures. To further analyze the structural stability we also performed secondary structure analysis and binding free energy calculations.

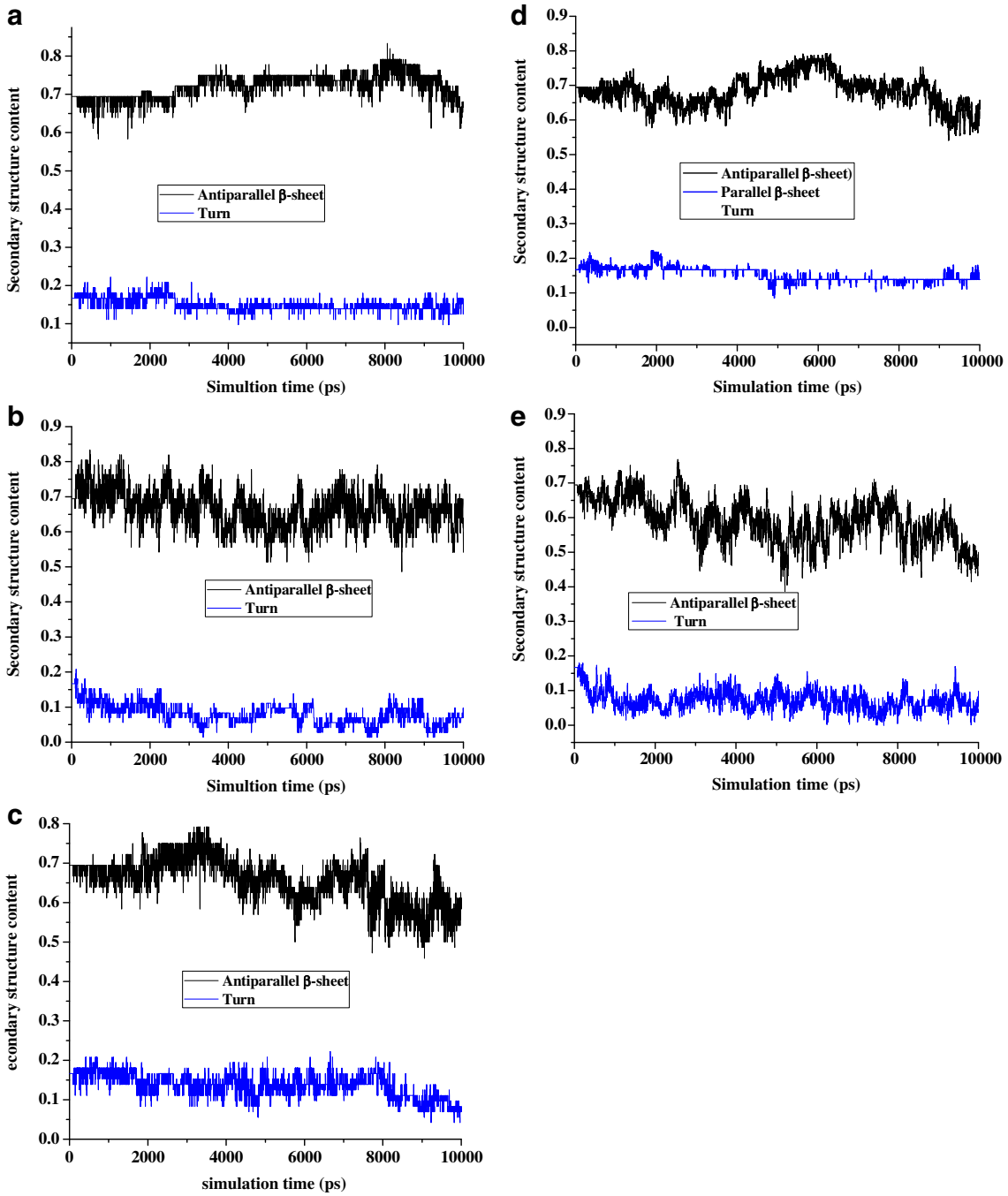



**Fig. 6** Time evolutions of the secondary structure contents for Sh2-St6 amyloid aggregate formed by Aβ peptide fragment (**MVGGVV1**) polymorph form I and its mutants (**a**) wild type, (**b**) V6G, (**c**) V5G, (**d**) V2G, and (**e**) M1G during the 10-ns MD simulations

The effect of single-glycine mutations on structural stabilities of the aggregates

To investigate how the steric zipper interfaces influence the structural stability of the double-sheet aggregates of VQIVYK, MVGGVV and LYQLEN peptides, the side-chains participating in these interfaces were replaced by glycine (Tables 1–4). As one can see from Fig. 4a, the largest aggregates composed of these mutants were less stable, compared to the respective wild type aggregates. The RMSD of I3G and Y5G are somewhat higher (c.a 5.5 Å), than those of V1G (c.a. 4.5 Å), indicating that I3G and Y5G exhibit higher potential to destabilize the structure of the VQIVYK aggregate. Mutations of the nonpolar side chain Ile3 or Tyr5 to Gly were found to result in destabilization of the oligomeric structures. Figure 4b indicates that none of MVGGVV1 mutants were structurally stable compared to the value of the respective wild type

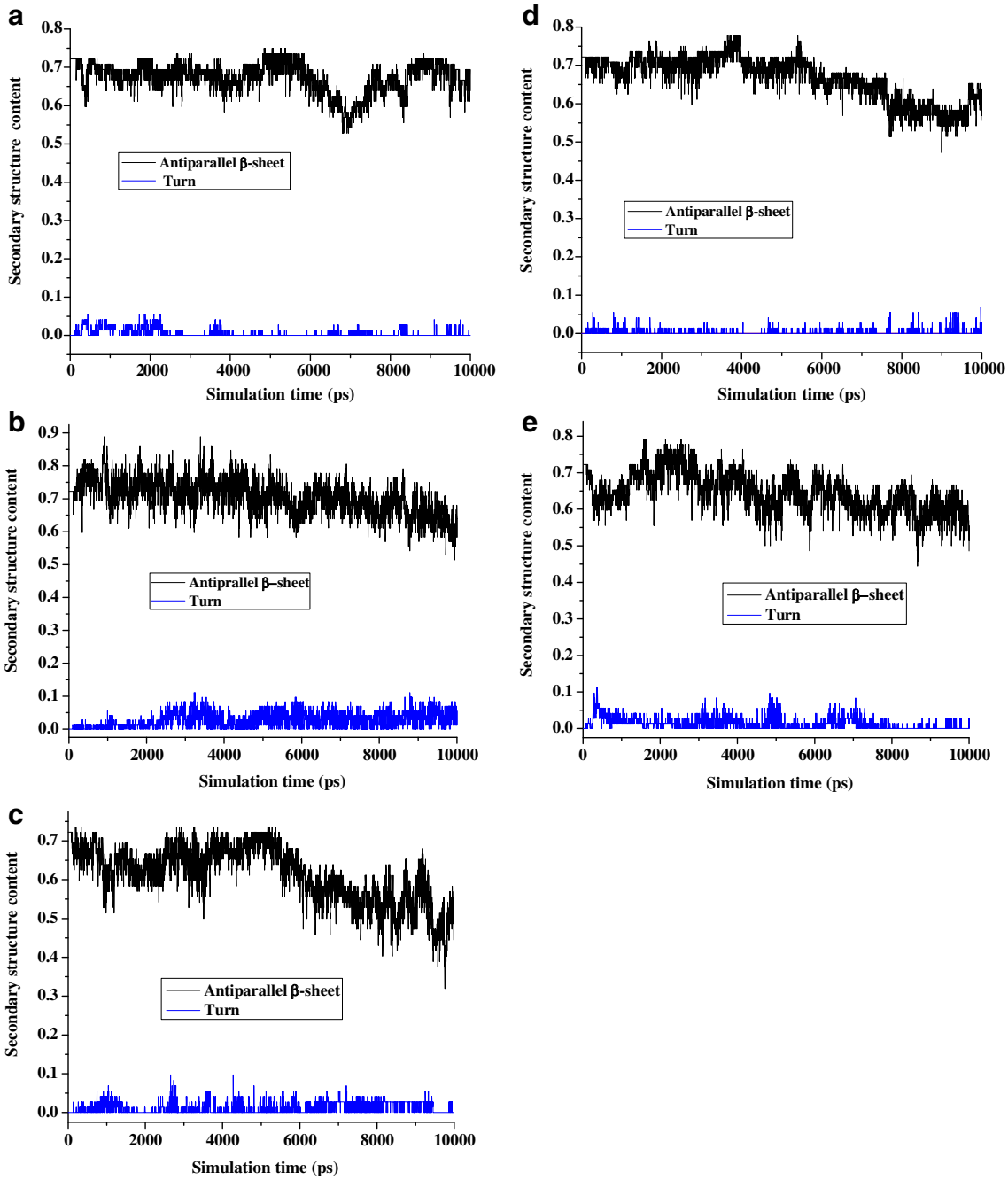



**Fig. 7** Time evolutions of the secondary structure contents for Sh2-St6 amyloid aggregate formed by Aβ peptide fragment (**MVGGVV1**) polymorph form I and its mutants (**a**) wildtype, (**b**) V6G, (**c**) V5G, (**d**) V2G, and (**e**) M1G during the 10-ns MD simulations

model. It shows that the RMSD of M1G and V6G are significantly higher (<8.5 Å) than those of V2G and V5G (<5.5.Å), indicating that V2G and V5G exhibit higher potential to destabilize the structural integrity of the MVGGVV1 oligomer. The result also (Fig. 4b), shows that mutation of the nonpolar sidechain Met1 or Val6 to Gly negatively affects the intersheet steric zipper destabilizing the structural integrity of the MVGGVV1 oligomers to a greater extent than the V2G and V5G mutants and the wild type. Figure 4c shows that some of MVGGVV2 mutants (V2G and V6G) were structurally stable compared to the wild type model. It shows that the RMSD of the mutant V2G and V6G are lower (nearly 4.0 Å) than those of M1G and wild type (nearly 6.0 Å), indicating that V2G and V6G exhibit higher potential to stabilize the structural integrity of the MVGGVV2 oligomer. The result also shows

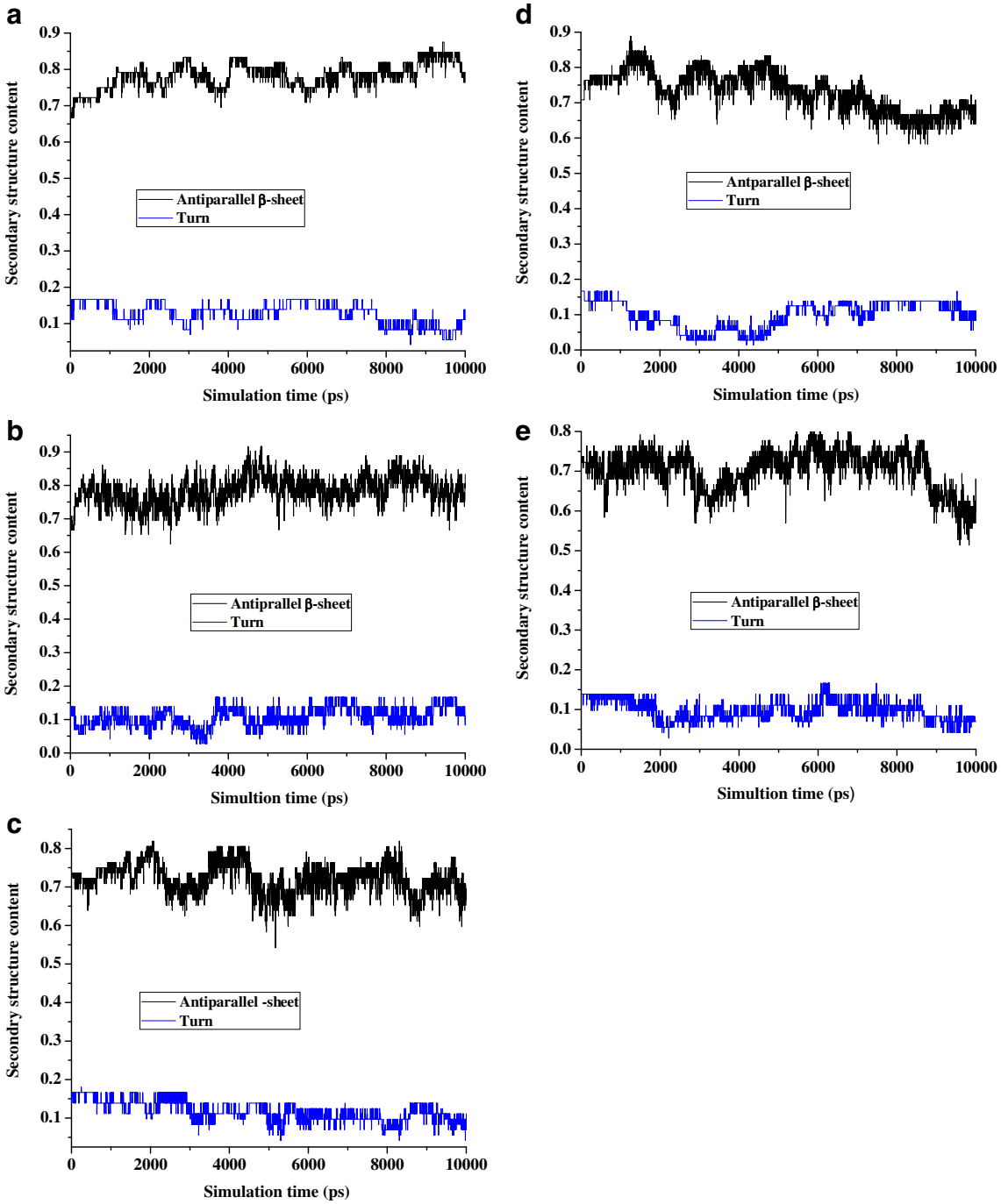



**Fig. 8** Time evolutions of the secondary structure contents Sh2-St6 amyloid aggregate formed by Insulin fragment **LYQLEN** and four of its mutants (**a**) wild type, (**b**) N6G, (**c**) L4G, (**d**) Q3G, and (**e**) Y2G during the 10-ns MD simulations

(Fig. 4c), that Y5G (RMSD >15.0 Å) destabilize the structural integrity of the MVGGVV2 oligomers to a greater extent than the other mutants and the wild type. Our finding that the MVGGVV2 wild type aggregates are less stable compared to certain mutants is in contrast to other oligomers in our study and to conclusions of the previous work done on the hexapeptide amyloid [16, 29]. This apparent contradiction could be explained based on the structural difference between this particular polymorph and other amyloid X–ray structure in that there is 90° bending in the upper sheet of MVGGVV form 2 [5]. As can be seen in Fig. 4d, none of the mutants of LYQLEN are as structurally stable as the wild type, indicating that the side chain interactions play an important role in determining the stability of the LYQLEN oligomers. However, the N6G mutant have small RMSD values (~2.5 Å), whereas the Y2G, Q3G and L4G mutants have large RMSD values (>4.0 Å). Comparison between the dynamics of the wild type and its mutants suggests that mutations N6G have little effect on the structural stability of the LYQLEN (low RMSD), whereas mutations Y2G, Q3G and L4G destabilize the oligomeric structures. The destabilization of the Q3G and L4G mutants is even more pronounced. As seen in Fig. 4d mutations of the polar side chain Gln3, or Leu4 to Gly would affect the intersheet steric zipper, leading to greater destabilization of the oligomeric structures.

## Secondary structure assessment

We analyzed the secondary structure of the oligomers using the DSSP algorithm written by Kabsch and Sander [38]. This algorithm is based on identification of hydrogen-bonding (H-bonding) patterns and recognizes seven types of secondary structures which can be grouped into three classes: helix (α-helix, $3_{10}$-helix, π-helix), β-strand (isolated β-bridge, extended β-sheet) and loop (turn, bend). The result of the secondary structure analysis for two layer models of the amyloid peptides shows that the wild type of Tau (St2-St5 VQIVYK), Insulin (St2-St6 LYQLEN) and Aβ amyloids Polymorphic form I and II (Sht2-St6 MVGGVV) appear to be stable at 300 K, which is confirmed by the conservation of high anti-parallel β-sheet content throughout the whole simulation time (as shown in Fig. 5a-d). In all systems the contents of antiparallel β sheets are much greater than the turn or parallel β-sheet indicating the conservation of the original structure. The results also indicate that antiparallel β-sheet, parallel β-sheet and turn content were preserved

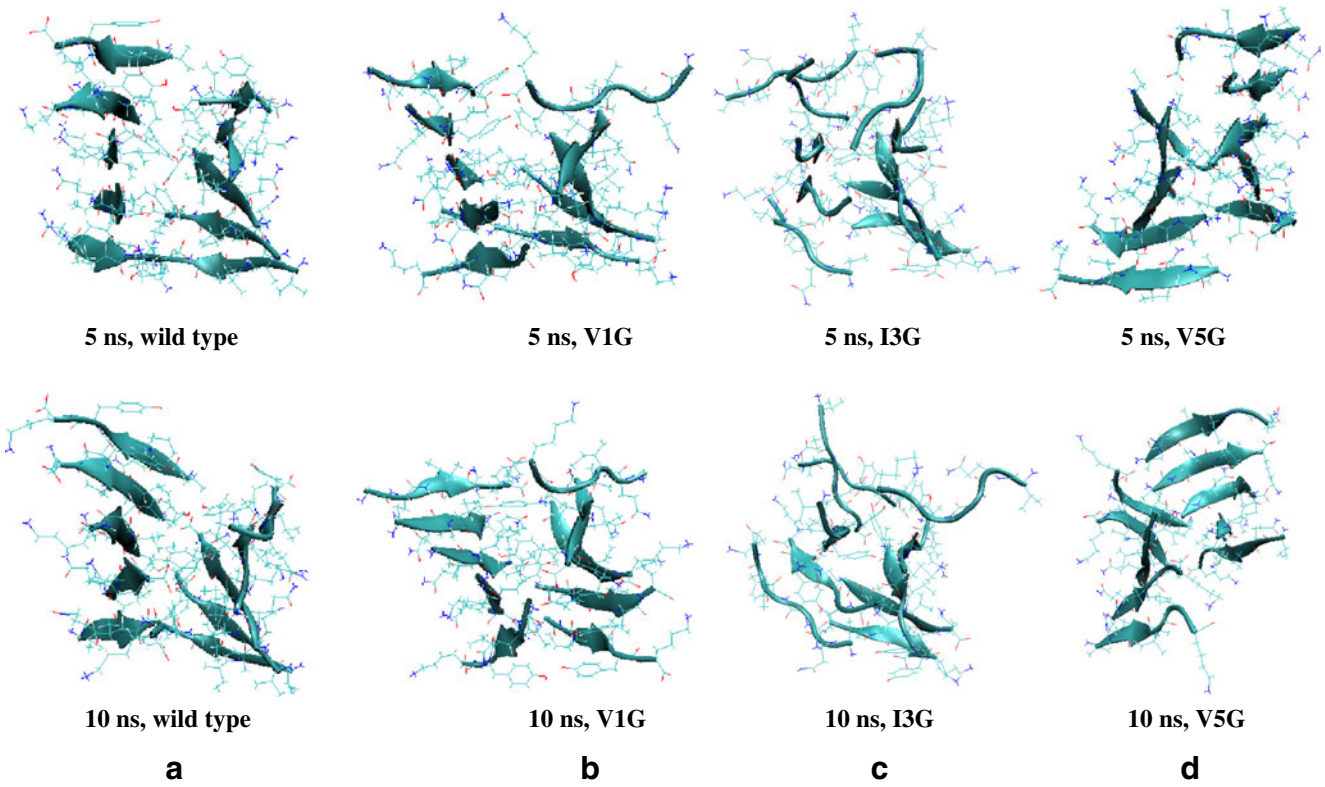



**Fig. 9** The snapshots of Sh2-St5 amyloid aggregate formed by Tau amyloid fragment **VQIVYK** and three of its mutants at the 5 ns midpoint (top) and at the end (bottom) of 10-ns MD trajectory for (**a**) wild type, (**b**) V1G, (**c**) I3G and (**d**) I5G. While the wild type is the most stable, the mutant I3G is the least stable

throughout the 10 ns simulation. The analyses of the secondary structure evolution throughout the simulation for the mutant form of the amyloids are shown in Figs. 5, 6, 7 and 8.

The results for Tau mutant and the wild type secondary structure is shown in Fig. 5a. In the case of the Tau mutants the content of the secondary structure declined, specially in the last 3 ns of the simulation. This result is in agreement with the RMSD results discussed above. The decline is largest in I3G which has the largest RMSD value. Figure 6b shows the results for Aβ peptide polymorph form I for both the wild type and mutants. The result for the wild type shows that the secondary structure is preserved as indicated with its high content of antiparallel β-sheets (~0.7). Among the mutants the content of the secondary structure was more unstable for M1G indicating the greater destabilizing effects of replacing methione with glycine. This is in agreement with the RMSD result, the highest RMSD among the mutants was from the M1G mutant (see Fig. 4b). Figure 7c shows the results for Aβ peptide polymorph form II for both the wild type and mutants. The result for the wild type shows that the secondary structure is preserved as indicated with its high content of antiparallel β-sheets (~0.7). Among the mutants the content of the secondary structure was more unstable for M1G indicating the greater destabilizing effects of replacing methione with glycine. This is in agreement with the RMSD result, the highest RMSD among the mutants was from the M1G mutant (see Fig. 4c).

The results for Insulin amyloid mutants (Fig. 8d) and the wild type shows that the secondary structure is preserved as indicated by its high content antiparallel β–sheets (~0.7). In the case of the Q3G and Y2G mutants the content of antiparallel β-sheets declined, specially in the last 2 ns of the simulation. This result is in agreement with the RMSD results. The RMSD for both Q3G and Y2G mutants was the largest ~5 Å.

Two trajectory snapshots (at 5 ns and 10 ns) are shown in Figs. 9, 10, 11 and 12 for each of the two layer oligomer aggregates. As the structure evolves, some of the terminal strands break the β-sheet ordering and twist relative to the remaining strands although they do not dissociate from the aggregate completely. Degree of this disorder correlates with the RMSD values reported in Figs. 2 and 3. Among the most disordered structures are mutant I3G of the Tau fragment (VQIVYK), mutants V6G and M1G of the Aβ fragment polymorph 1 (MVGGVV1) mutant V5G of the Aβ fragment polymorph 2 (MVGGVV2) and the mutants Y2G and Q3G of the

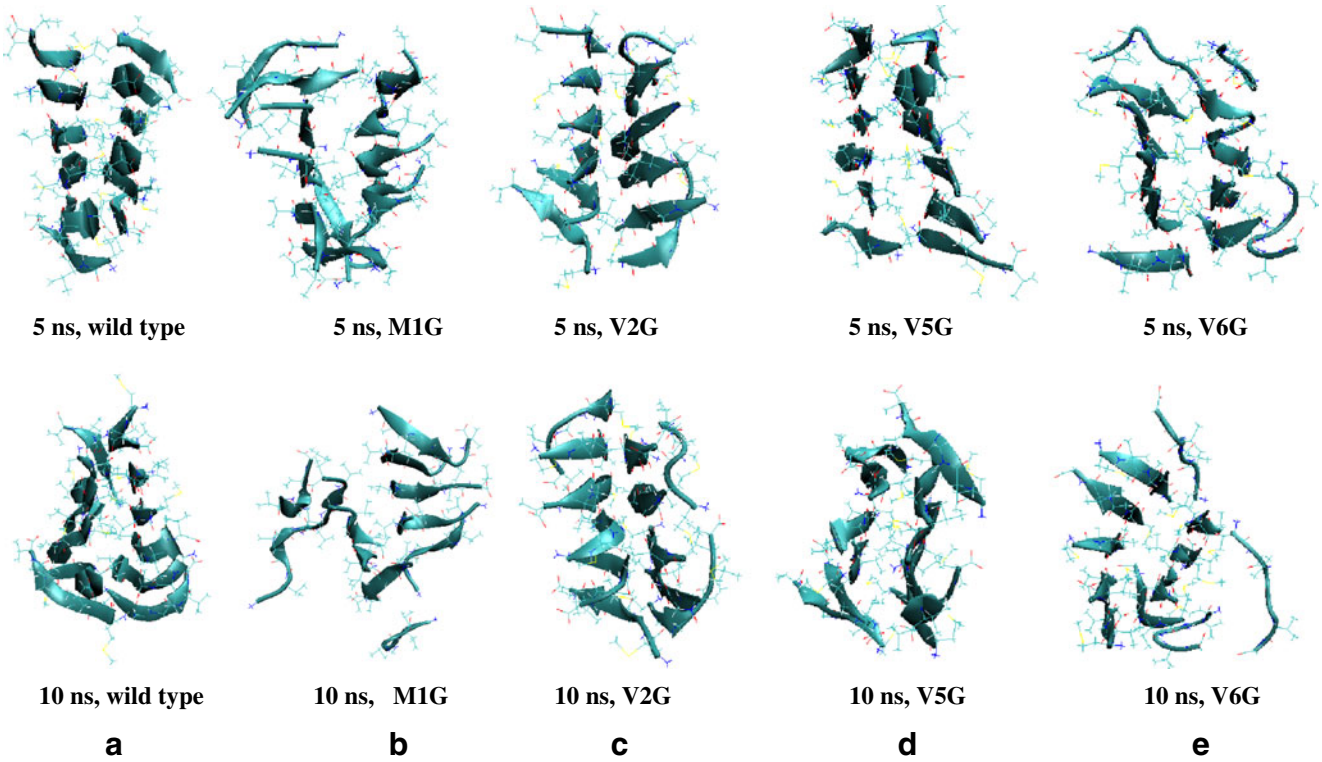



**Fig. 10** The snapshots of Sh2-St6 amyloid aggregate formed by Aβ peptide fragment (**MVGGVV1**) and four of its mutants. The initial structure is selected from the X-Ray structure of the polymorph I. The snapshots are shown at the mid-point (top) and at the end (bottom) of 10-ns MD trajectory for (**a**) wild type, (**b**) M1G, (**c**) V2G, (**d**) V5G, and (**e**) N6G. The wild type is the most stable, while the V6G and M1G mutants are the least stable

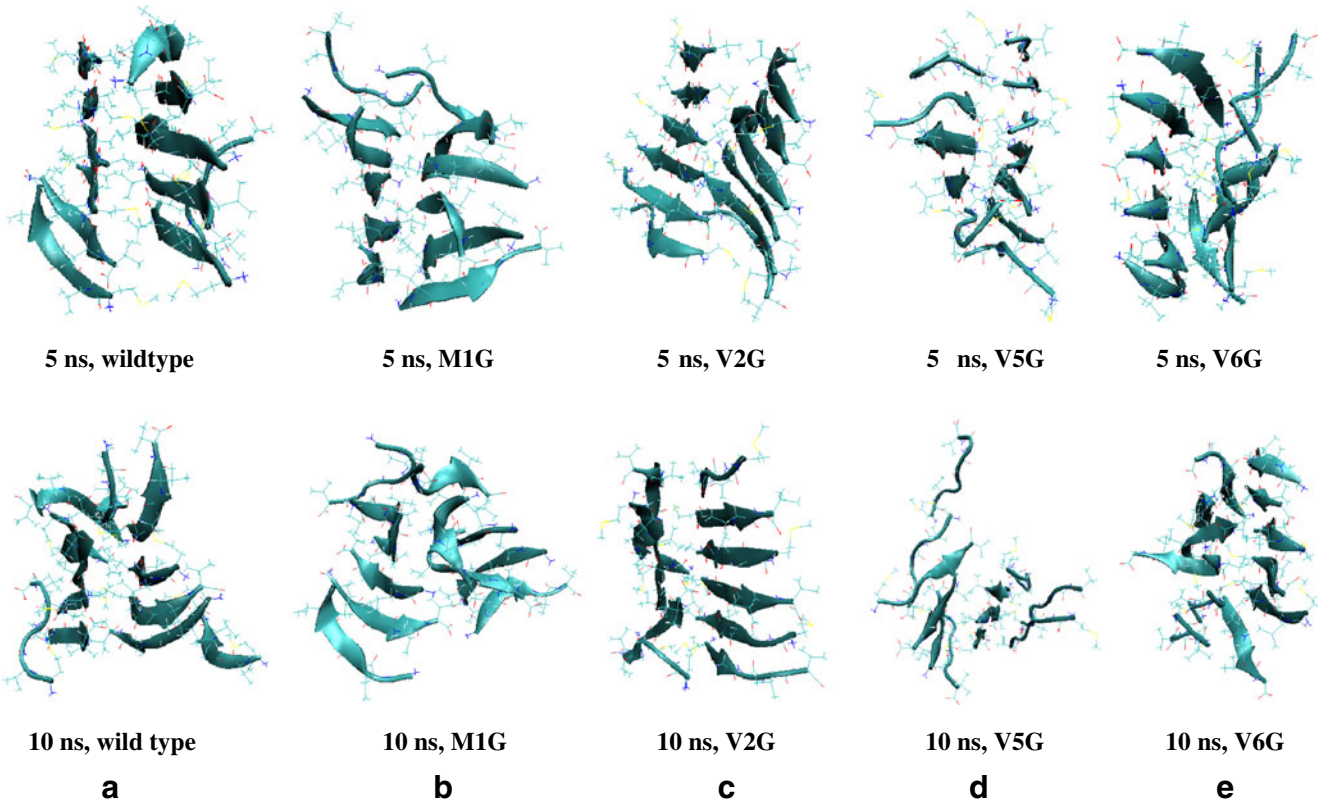



**Fig. 11** The snapshots of Sh2-St6 amyloid aggregate formed by Aβ peptide fragment (**MVGGVV2**) and four of its mutants. The initial structure is selected from polymorph II X-Ray structure. The snapshots at the mid-point (top) and at the end (bottom) of 10-ns MD trajectory for: (**a**) wildtype, (**b**) M1G, (**c**) V2G, (**d**) V5G, and (E) V6G. The wildtype, V2G, and V6G are more stable than the other mutants, presumably due to the bending of the strands in one of the layers

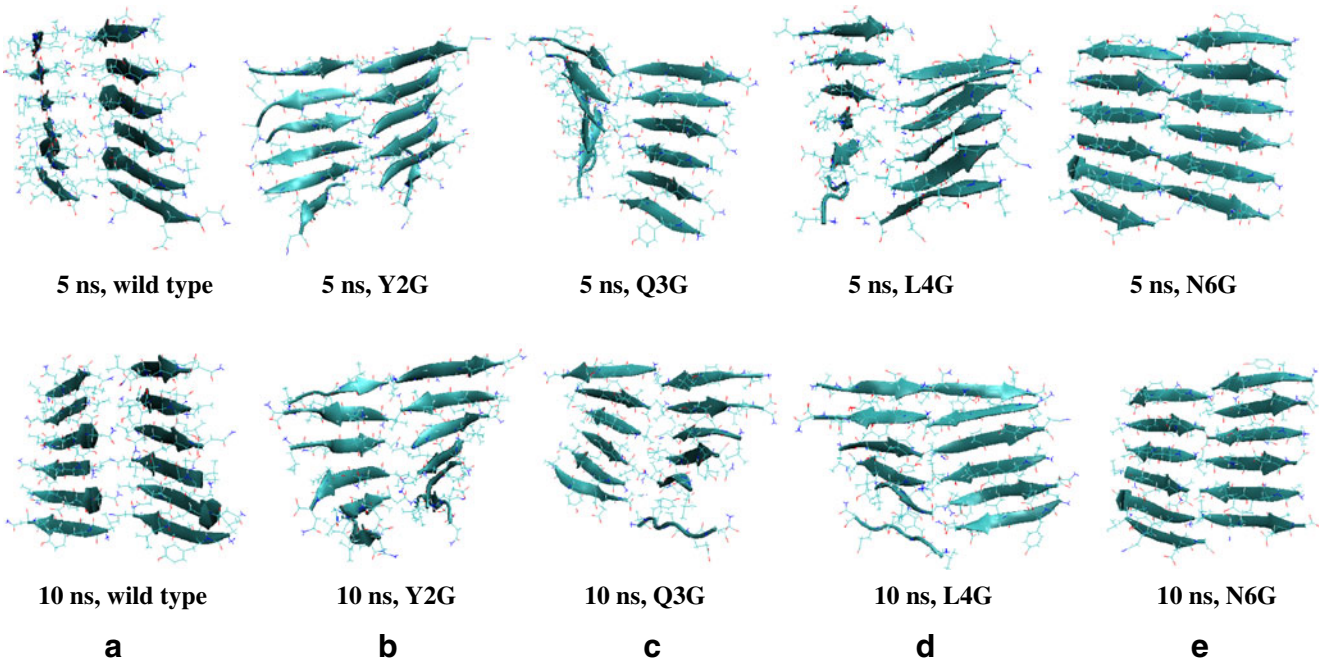



**Fig. 12** The snapshots of Sh2-St6 amyloid aggregate formed by Insulin fragment (**LYQLEN**) and four of its mutants at the 5 ns mid-point (top) and at the end (bottom) of 10-ns MD trajectory for: (**a**) wild type, (**b**) Y2G, (**c**) Q3G, (**d**) L4G, and (**e**) N6G. While the wild type is the most stable, the mutants Y2G and Q3G are the least stable

**Table 5** Binding free energy components (kcal $mol^{-1}$) and standard deviations calculated with MM-PBSA* for wild type and mutants of the tau (**VQIVYK**) oligomer bi-layers (**Sh2-St5** models)

| Energy (Kcal/mol) | Wild type Mean ±std | Mutant -V1G Mean ±std | Mutant -I3G Mean±std | Mutant –Y5G Mean±std |
|---|---|---|---|---|
| $\Delta E_{elec.}$ | 446.47±24.29 | 396.40±49.64 | 379.36±22.97 | 396.29±39.45 |
| $\Delta E_{vdw}$ | -101.84±4.74 | -85.05±4.29 | -89.53±4.29 | -80.01±4.47 |
| $\Delta E_{int}$ | 0 | 0 | 0 | 0 |
| $\Delta E_{gas}$ | 344.62±24.30 | 311.35±50.17 | 289.83±22.37 | 316.28±38.48 |
| $\Delta G_{PB}$ | -404.42±23.24 | -359.20±45.04 | -331.30±19.30 | -370.80±38.51 |
| $\Delta G_{sur}$ | -15.65±0.34 | -14.05±0.48 | -14.78±0.43 | -13.68±0.33 |
| $\Delta G_{polar}$ | 42.05±7.58 | 37.20±10.03 | 48.06±9.13 | 25.48±8.44 |
| $\Delta G_{non\text{-}polar}$ | -117.49±5.08 | -99.1±4.77 | -104.3±4.72 | -93.69±4.80 |
| $\Delta G_{TOT}$ | -75.44±6.22 | -61.91±9.29 | -56.25±8.34 | -68.20±7.04 |

* $E_{vdw}$, and $E_{elec}$ are the van der Waals and electrostatic binding terms. $\Delta G_{GB}$ and $\Delta G_{sur}$ are the solvation energies of polar and nonpolar residues, calculated by Amber 10 using the Generalized Born model. $\Delta G_{polar}$ and $\Delta G_{non\text{-}polar}$ are the sums of polar energy ($\Delta E_{elec} + \Delta G_{PB}$) and non-polar energy components ($\Delta E_{vdw} + \Delta G_{sur}$), respectively. $\Delta G_{TOT}$ is the sum of $\Delta G_{polar}$ and $\Delta G_{non\text{-}polar.}$ $\Delta G_{TOT}$ (the binding free energy can also be obtained using the equation, $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{PB} + \Delta G_{SUR}$

Insulin amyloid (LYQLEN). Apparently, the mutated aminoacids were involved in the steric zippers, which were not only holding the β-sheets together, but also preserving them from disaggregation. Inversely, the complementarity of the aminoacids sidechains would be essential for the formation of the ordered aggregate. On the other hand, disordered random aggregation may take place for any polypeptide studied in this work, as suggested by the negative values of association free energies, reported in the next section.

## Free energy calculations

The binding free energies were calculated with the molecular mechanics Poisson-Boltzmann solvent accessible surface area (MM-PBSA) model [40], as implemented in AMBER. In this method, the total binding free energy in water is approximated by $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{PB} + \Delta G_{SUR}$. The $\Delta E_{GAS}$ is the gas phase interaction energy. The $\Delta G_{PB/GB}$ is the polar part of the solvation free energy represented by Poisson-Boltzmann approaches. The $\Delta G_{SUR}$ is the surface area term, approximating the non-polar part of the solvation free energy. In this formula, the conformational entropy of the solute is not considered, while the solvent entropy is implicitly considered in the $\Delta G_{PB}$ and $\Delta G_{SUR}$. Although the MM-PBSA calculations may overestimate the absolute binding free energy due to the missing terms (e.g., conformational entropy change of the solute upon binding), they usually give a reasonable estimate on the relative binding free energy when the conformational entropy changes of two binding modes are comparable [41].

The binding energy was calculated by MM-PBSA method as specified in method section. The breakdown of binding energy components is listed in Tables 5, 6, 7 and 8. The MM-PBSA analysis allows us to separate the

**Table 6** Binding free energy components (kcal $mol^{-1}$) and standard deviations calculated with MM-PBSA (Amber PBSA) for wild type and mutant complexes of the A-beta polymorphic form I **(MVGGVV1)** oligomer bi-layer (**Sh2-St6** models)

| Energy (Kcal/mol) | Wild type Mean ±std | Mutant- M1G Mean ±std | Mutant -V2G Mean ±std | Mutant -V5G Mean±std | Mutant -V6G Mean±std |
|---|---|---|---|---|---|
| $\Delta E_{elec.}$ | -214.20±16.40 | -200.00±25.12 | -262.67±30.24 | -187.40±21.12 | -239.17±29.63 |
| $\Delta E_{vdw}$ | -94.88±5.58 | -74.93±6.31 | -76.93±4.85 | -78.13±4.56 | -81.06±5.58 |
| $\Delta E_{int}$ | 0 | 0 | 0 | 0 | 0 |
| $\Delta E_{gas}$ | -309.09±15.72 | -275.62±24.88 | -339.61±28.93 | -265.53±20.19 | -320.23±29.43 |
| $\Delta G_{PB}$ | 230.47±55.68 | 222.67±23.68 | 265.34±26.56 | 199.03±18.34 | 244.04±28.33 |
| $\Delta G_{sur}$ | -14.97±0.45 | -12.97±0.76 | -13.20±0.47 | -13.19±0.49 | -13.65±0.63 |
| $\Delta G_{polar}$ | 16.27±52.64 | 21.98±8.22 | 2.66±8.03 | 11.63±6.42 | 4.87±7.37 |
| $\Delta G_{non\text{-}polar}$ | -109.85±6.03 | -87.90±7.07 | -90.13±5.32 | -91.32±5.05 | -94.70±6.21 |
| $\Delta G_{TOT}$ | -93.58±53.51 | -65.91±6.36 | -87.47±6.04 | -79.69±5.17 | -89.83±7.18 |

**Table 7** Binding free energy components (kcal $mol^{-1}$) and standard deviations calculated with MM-PBSA (Amber PBSA) for wild type and mutant complexes of the **A-beta polymorphic form II (MVGGVV2) oligomer** bilayers (**Sh2-St6** models)

| Energy (Kcal/mol) | Wild type Mean ±std | Mutant -M1G Mean ±std | Mutant- V2G Mean ±std | Mutant -V5G Mean±std | Mutant- V6G Mean±std |
|---|---|---|---|---|---|
| **Δ E $_{elec}$** | -641.67±38.91 | -419.66±37.36 | -592.76±33.31 | -527.91±54.10 | -544.54±26.74 |
| **Δ E $_{vdw}$** | -132.12±7.26 | -102.91±5.88 | -98.90±6.73 | -106.59±6.44 | -117.32±6.79 |
| **Δ E $_{int}$** | 0 | 0 | 0 | 0 | 0 |
| **Δ E $_{gas}$** | -773.80±35.40 | -522.57±35.91 | -691.66±33.23 | -634.50±54.88 | -661.86±24.16 |
| **Δ G $_{PB}$** | 631.05±33.17 | 432.44±32.78 | 580.15±27.02 | 530.17±49.25 | 538.05±22.21 |
| **Δ G $_{sur}$** | -22.81±0.32 | -19.16±0.38 | -19.83±0.63 | -19.51±0.38 | -20.52±0.36 |
| **Δ G $_{polar}$** | -10.62±10.21 | 12.78±9.91 | -12.61±10. 51 | 2.26±10.51 | -6.50±10.27 |
| **Δ G $_{non-polar}$** | -154.93±7.58 | -122.07±6.26 | -118±7.36 | -126.10±6.82 | -137.84±7.15 |
| **Δ $G_{TOT}$** | -165.56±7.57 | -109.29±7.85 | -131.34±9.80 | -123.84±9.09 | -144.33±7.47 |

total free energy of binding into electrostatic, van der Waals interactions and solute-solvent interactions, and thereby gain additional insights into the monomer to monomer association process in the formation of the double layer of the amyloid oligomers. As shown in Tables 5 and 8, van der Waals interactions play a very important role in the simulation, contributing significantly more to the total interaction energy than the electrostatic interaction for the Tau (VQIVYK) and Insulin (LYQLEN) aggregates. In the case of Aβ oligomers, MVGGVV, the electrostatic interactions constitute the main component of the total interaction energy (Tables 6 and 7). Nonpolar solvation energies favor the binding and the polar solvation energies disfavor it in all the systems studied. Note that the internal component of the molecular-mechanical energy (bond, angle, and torsional energies) has zero contribution to the binding free energy, because the structures of the monomers in its unbound and bound states were assumed to be the same (data not shown). Tables 5, 6, 7 and 8 also reports the apolar/hydrophobic and polar/electrostatic contributions to the free energy. We found that the predicted binding free energy is dominated by the magnitudes of the apolar components ($\Delta E_{vdw} + \Delta E_{sur}$) in all the four oligomers and their mutants. In contrast, the polar interactions ($\Delta E_{elec} + \Delta G_{PB}$) show less contribution to the binding free energy. The result of the binding free energy calculation also indicated that the wild type is the most stable structure compared to the mutants. From the negative total binding free energy of the wild types we clearly see that this is a favorable protein-protein complex in pure water. The mutants also form a stable complex based on the negative total binding free energy. However, the mutant complex is less thermodynamically favorable than the wild type complex. The calculation indicated that the mutation of bulky polar side chain from the steric zipper structure leads to the less stable dimer (see example Tables 5, 6, 7 and 8) giving the mutant with smallest binding free energy, and indicating that these residues are important for stabilizing the structure.

**Table 8** Binding free energy components (kcal $mol^{-1}$) and standard deviations calculated with MM-PBSA (Amber PBSA) for wild type and mutant complexes of the insulin (**LQYLEN**) oligomer bi-layers (**Sh2-St6** models)

| Energy (Kcal/ mol) | Wild type Mean ±std | Mutant – Y2G Mean±std | Mutant – Q3G Mean±std | Mutant –L4G Mean±std | Mutant – N6G Mean±std |
|---|---|---|---|---|---|
| Δ $E_{elec}$ | 298.79±36.86 | 374.35±30.91 | 359.94±33.49 | 327.50±36.68 | 304.01±36.57 |
| Δ $E_{vdw}$ | -105.15±4.97 | -80.86±4.71 | -87.77±5.63 | -86.63±4.87 | -91.68±5.30 |
| Δ $E_{int}$ | 0 | 0 | 0 | 0 | 0 |
| Δ $E_{gas}$ | 193.65±35.14 | 293.49±30.22 | 272.16±33.19 | 240.87±36.64 | 212.32±34.90 |
| Δ $G_{PB}$ | -280.31±32.94 | -362.93±26.10 | -345.09±33.94 | -307.99±33.68 | -289.87±34.98 |
| Δ $G_{sur}$ | -16.58±0.32 | -14.11±0.48 | -15.08±0.43 | -15.27±0.42 | -15.36±0.48 |
| Δ G $_{polar}$ | 18.48±8.39 | 11.41±8.26 | 14.85±6.86 | 19.50±8.27 | 14.14±7.59 |
| Δ G $_{non-polar}$ | -121.73±5.29 | -94.97±5.19 | -102.85±6.06 | -101.90±5.29 | -107.85±5.78 |
| Δ $G_{TOT}$ | -103.24±6.30 | -83.56±7.25 | -88.00±6.07 | -82.40±7.42 | -92.91±6.26 |

The widely accepted hypothesis on the amyloid disease is that soluble protein oligomers are the source for toxicity and are the primary pathogenic factor in these diseases and thus small molecules that prevent or reverse protein oligomerization may provide a mechanism to target the actual cause of the disease [42, 43]. Peptidomimetics are a promising class of small molecules capable of inhibiting oligomerization. Most fibrillogenesis inhibitors of this type were designed rationally based on molecular recognition elements found in the site of aggregation [44]. Identification of this aggregation site is often based on the mutational data. Such data helps to pinpoint the small regions on the protein interaction interface that are responsible for a disproportionate contribution to the binding energy of the two proteins [45–47]. In this work we have shown that most mutations at the aggregation site reduce the binding free energy and weaken the aggregation. Therefore, the computational studies can serve the same purpose of the rational design, as experimental mutation studies.

Another potential application of the presented approach is design of imaging agents. The progress in therapeutic agents for treatment of neurodegenerative amyloid diseases calls for development of more specific biomarkers to detect early stages of amyloid diseases [48]. Design of peptidomimetics based on the data obtained in the molecular dynamic simulation may provide the starting point for design of specific aggregation inhibitors drugs and diagnostic agents. Both structural and thermodynamic results reported in this study illustrate the higher fluctuation in RMSD values and less negative binding free energies for the mutated peptides. These mutants, therefore, may serve as aggregation inhibitors pending the experimental confirmation.

## Conclusions

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

(i) the stability of the VQIVYK, MVGGVV and LYQLEN peptides oligomers increases with increasing the number of β-strand;
(ii) the Sh2-St4 model acts as a stable seed in prompting amyloid fibril formation for all the cases considered;
(iii) the binding energy calculated by MM-PBSA method and the analysis of individual contributions to the binding energy shows the hydrophobic interactions play an important role in stabilizing the structural organizations between β-sheet layers in the oligomers. The result of the binding free energy calculation also indicated that the wild type is the most stable structure compared to the mutants;
(iv) the hydrophobic steric zipper on the intersheet interface contributes significantly to the stability of the entire aggregate structures. Mutations of the side-chains participating in the steric zipper interfaces of the oligomeric (VQIVYK, MVGGVV1 and LYQLEN) peptides to Gly resulted in decline of secondary structure content compared to corresponding wild type indicating that the role of the replaced amino acid in stabilizing the structure;
(v) a single glycine substitution at the steric zipper interface disrupts the hydrophobic steric zipper remarkably, indicating that the hydrophobic attraction is a major driving force for stabilizing and aggregation of oligomers. Consequently, the substantial reduction in the van der Waals intersheet interactions leads to destabilization of the oligomers. Overall, aggregation of both wild type and mutant peptides is driven by nonpolar interaction.

Some evidence from the experimental work suggests that short peptides may share similar intermolecular interactions to their parent proteins while forming amyloid fibril [49]. Thus, exploring the structural stability and aggregation behavior of the short peptides may gain insights into the self-assembly process at the early stage of fibril formation and provide a clue to understand the possible aggregation mechanism of their parent proteins. The hexapeptide NFGAIL, a fragment truncated from human islet amyloid polypeptide (hIAPP, residues 22–27), is one of the shortest fragments that have been shown to form amyloid fibrils similar to those formed by the full polypeptide [49] and the fibrils are cytotoxic toward the pancreatic cell line. Therefore, this hIAPP "amyloid-core" peptide has been used as a simplified model system to facilitate the discovery of key factors underlying amyloid fibril formation and the development of anti-amyloid agents. Porat et al. [50] showed that whereas the NFGAIL was a minimal fibril forming fragment from hIAPP with Tyr substituted for Phe (*i.e.*, $^{22}$NFGAILSS$^{29}$ to $^{22}$NYGAILSS$^{29}$ did not form fibrils by itself and even inhibited fibril formation. Along these lines one can envision a possible strategy to inhibit the formation of early aggregates that includes the design of specific inhibitor, breaking the hydrophobic steric zipper observed in the structure of hydrophobic region of the amyloid aggregate. Proof of principle for such a strategy has been published recently [51, 52]. Thus, designing new peptidomimetic inhibitors able to prevent the fibril formation based on the steric zipper motif of the oligomers, similar to the ones examined in this study may become a viable therapeutic strategy. The peptidomimetic approach can also be implemented in designing specific biomarkers for early stage detection of aggregate formation [52, 53].

**Acknowledgments** This work was supported in part by the National Science Foundation (CCF/CHE 0832622). This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The author appreciates the anonymous reviewers for their insightful comments, which greatly helped improving this paper.

## References

1. Selkoe DJ (2003) Nature 426:900–904
2. Hamley IW (2007) Angew Chem Int Edit 46:8128–8147
3. Makin OS, Serpell LC (2005) Febs J 272:5950–5961
4. Nelson R, Eisenberg D (2006) Curr Opin Struc Biol 16:260–265
5. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJW, McFarlane HT, Madsen AO, Riekel C, Eisenberg D (2007) Nature 447:453–457
6. Wiltzius JJW, Sievers SA, Sawaya MR, Cascio D, Popov D, Riekel C, Eisenberg D (2008) Protein Sci 17:1467–1474
7. Ivanova MI, Sievers SA, Sawaya MR, Wall JS, Eisenberg D (2009) Proc Natl Acad Sci USA 106:18990–18995
8. Park J, Kahng B, Hwang W (2009) PLoS Comput Biol 5:17
9. De Simone A, Esposito L, Pedone C, Vitagliano L (2008) Biophys J 95:1965–1973
10. Gsponer J, Haberthur U, Caflisch A (2003) Proc Natl Acad Sci USA 100:5154–5159
11. Esposito L, Paladino A, Pedone C, Vitagliano L (2008) Biophys J 94:4031–4040. doi:10.1529/biophysj.107.118935
12. Esposito L, Pedone C, Vitagliano L (2006) Proc Natl Acad Sci USA 103:11533–11538
13. Wei GH, Song W, Derreumaux P, Mousseau N (2008) Front Biosci 13:5681–5692
14. Wu C, Wang ZX, Lei HX, Zhang W, Duan Y (2007) J Am Chem Soc 129:1225–1232
15. Berryman JT, Radford SE, Harris SA (2009) Biophys J 97:1–11
16. Chang LK, Zhao JH, Liu HL, Liu KT, Chen JT, Tsai WB, Ho Y (2009) J Biomol Struct Dyn 26:731–740
17. Nerelius C, Johansson J, Sandegren A (2009) Front Biosci 14:1716–U3856
18. Rauk A (2008) Dalt Transact 1273-1282
19. Rauk A (2009) Chem Soc Rev 38:2698–2715
20. Teplow DB, Lazo ND, Bitan G, Bernstein S, Wyttenbach T, Bowers MT, Baumketner A, Shea JE, Urbanc B, Cruz L, Borreguero J, Stanley HE (2006) Acc Chem Res 39:635–645
21. Berriman J, Serpell LC, Oberg KA, Fink AL, Goedert M, Crowther RA (2003) Proc Natl Acad Sci USA 100:9034–9038
22. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D (2005) Nature 435:773–778
23. Sunde M, Blake C (1997) Adv Prot Chem , Vol 50. Academic, San Diego, pp 123-159
24. Devlin GL, Knowles TPJ, Squires A, McCammon MG, Gras SL, Nilsson MR, Robinson CV, Dobson CM, MacPhee CE (2006) J Mol Biol 360:497–509
25. Hong DP, Fink AL (2005) Biochemistry 44:16701–16709
26. Wilhelm KR, Yanamandra K, Gruden MA, Zamotin V, Malisauskas M, Casaite V, Darinskas A, Forsgren L, Morozova-Roche LA (2007) Eur J Neurol 14:327–334
27. Brange J, Andersen L, Laursen ED, Meyn G, Rasmussen E (1997) J Pharm Sci 86:517–525
28. Ahmad A, Uversky VN, Hong D, Fink AL (2005) J Biol Chem 280:42669–42675
29. Zhang ZQ, Chen H, Bai HJ, Lai LH (2007) Biophys J 93:1484–1492
30. Xu WX, Ping J, Li WF, Mu YG (2009) J Chem Phys 130:8
31. Kayed R, Pensalfini A, Margol L, Sokolov Y, Sarsoza F, Head E, Hall J, Glabe C (2009) J Biol Chem 284:4230–4237
32. Quist A, Doudevski L, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R (2005) Proc Natl Acad Sci USA 102:10427–10432
33. Vitagliano L, Stanzione F, De Simone A, Esposito L (2009) Biopolymers 91:1161–1171
34. http://www.doe-mbi.ucla.edu/~sawaya/chime/xtalpept (2009)
35. D.A. Case TAD TEC, III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Ross C. Walker, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA (2008) University of California, San Francisco
36. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8593
37. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
38. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
39. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38
40. Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Acc Chem Res 33:889–897
41. Chong LT, Duan Y, Wang L, Massova I, Kollman PA (1999) Proc Natl Acad Sci USA 96:14330–14335
42. Hardy J, Selkoe DJ (2002) Science 297:353–356
43. Lesne S, Koh MT, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallagher M, Ashe KH (2006) Nature 440:352–357
44. Sciarretta KL, Gordon DJ, Meredith SC (2006) Amyloid, Prions, and Other Protein Aggregates. Pt C Elsevier, San Diego, pp 273–312
45. Bogan AA, Thorn KS (1998) J Mol Biol 280:1–9
46. Haydar SN, Yun HED, Staal RGW, Hirst WD (2009) Annu Rep Med Chem, Vol 44. Elsevier, San Diego, pp 51–69
47. Blazer LL, Neubig RR (2009) Neuropsychopharmacology 34:126–141
48. Otto M, Lewczuk P, Wiltfang J (2008) Methods 44:289–298
49. Tenidis K, Waldner M, Bernhagen J, Fischle W, Bergmann M, Weber M, Merkle ML, Voelter W, Brunner H, Kapurniotu A (2000) J Mol Biol 295:1055–1071
50. Porat Y, Mazor Y, Efrat S, Gazit E (2004) Biochemistry 43:14454–14462
51. Sato T, Kienlen-Campard P, Ahmed M, Liu W, Li HL, Elliott JI, Aimoto S, Constantinescu SN, Octave JN, Smith SO (2006) Biochemistry 45:5503–5516
52. Takahashi T, Ohta K, Mihara H. Prot Struct Func Bioinf 78:336-347
53. Kim YS, Lee JH, Ryu J, Kim DJ (2009) Curr Pharm Des 15:637–658