Virtual quasi-2D intermediates as building blocks for plausible structural models of amyloid fibrils from proteins with complex topologies: A case study of insulin

Wojciech Puławski[1,2*] and Wojciech Dzwolak[1,3*]

[1]*Institute of High Pressure Physics, Polish Academy of Sciences, 29/37 Sokołowska Str., 01-142 Warsaw, Poland*
[2]*Bioinformatics Laboratory, Mossakowski Medical Research Institute, Polish Academy of Sciences, 5 Pawinski Str., 02-106 Warsaw, Poland (Present address)*
[3]*Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, 1 Pasteur Str., 02-093 Warsaw, Poland.*

[*] Corresponding authors:
W. Puławski: Phone: (+48) 22 849 93 58, E-mail: wpulawski@imdik.pan.pl
W. Dzwolak: Phone: (+48) 22 552 6567, E-mail: wdzwolak@chem.uw.edu.pl

Abstract

Conformational transitions of globular proteins into amyloid fibrils are complex multistage processes exceedingly challenging to simulate using molecular dynamics (MD). Slow monomer diffusion rates and rugged free energy landscapes disfavor swift self-assembly of orderly amyloid architectures within timescales accessible to all-atom MD. Here, we conduct a multiscale MD study of the amyloidogenic self-assembly of insulin: a small protein with complex topology defined by two polypeptide chains interlinked by three disulfide bonds. In order to avoid kinetic traps, unconventional pre-planarized insulin conformations are used as amyloid building blocks. These starting conformers generated through uniaxial compression of the native monomer in various spatial directions represent 6 distinct (out of 16 conceivable) 2D topological classes varying in N- / C-terminal segments of insulin's A- and B-chains being placed inside, or outside of the central loop constituted by the middle sections of both chains and Cys7A-Cys7B / Cys19B-Cys20A disulfide bonds. Simulations of the fibrillar self-assembly are initiated through a biased in-register alignment of 2, 3 or 4 layers of flat conformers belonging to a single topological class. The various starting topologies are conserved throughout the self-assembly process resulting in polymorphic amyloid fibrils varying in structural features such as helical twist, presence of cavities, as well as the overall stability. Some of the protofilament structures obtained in this work are highly compatible with the earlier biophysical studies on insulin amyloid and high resolution studies on insulin-derived amyloidogenic peptide models postulating presence of steric zippers. Our approach provides *in silico* means to study amyloidogenic tendencies and viable amyloid architectures of larger disulfide-constrained proteins with complex topologies.

Keywords: amyloid; self-assembly; molecular dynamics; in-register stacking; kinetic trap

Introduction

Formation of amyloid fibrils, the highly ordered β-sheet-rich aggregates, is a generic structural transition accessible to various proteins and polypeptides [1-3]. This phenomenon is of paramount biomedical importance. Historically, abnormal amyloid deposits in various tissues have been linked to a wide spectrum of degenerative maladies including Alzheimer's and Parkinson's diseases, cardiac amyloidosis, and type II diabetes mellitus [3-5]. Later studies have shown that early aggregates rather than mature amyloid fibrils may be involved in the etiology of some of these conditions (e.g. [6-7]). On the other hand, there are many recognized examples of biologically functional amyloid fibrils [8-10]. Formation of amyloid aggregates is often thermodynamically favored under close-to-physiological conditions [11]. However, *de novo* amyloidogenesis encompasses several discrete stages which may be associated with significant barriers of both enthalpic and entropic natures (for example: partial unfolding of the native state, local association of aggregation-prone intermediates and formation of amyloid nuclei) [1,3]. The downhill thermodynamics of amyloidogenesis becomes accessible only to proteins acquiring certain levels of instability and at, or above, certain critical concentrations [12-13]. The past two decades have witnessed significant progress in understanding the molecular and physicochemical mechanisms of amyloidogenesis and its role in biology. While mostly experimental efforts (including detailed structural studies on amyloid fibrils from typically moderate in length polypeptide chains [14-17]) have contributed to the present state of the field, for a long time those were paralleled and complimented by various computational approaches (e.g. [18-24]). Challenges faced by *in silico* studies on amyloid formation are multiple and significant (see [25-28]). Simplistic amyloid models based on the applications of coarse-grained methods lack realistic structural detail and do not address protein-protein and protein-solvent interactions to a satisfying degree. On the other hand, the sheer complexity of the simulated systems exacerbated by high numbers of polypeptide chains (monomers) that need to be considered for physically realistic simulations require enormous computational resources in the case of all-atom molecular dynamics. This holds true even for

implicit solvent simulations which are unlikely to adequately address solvation phenomena known to play an important role in the aggregation process. Therefore, there is a strong argument for undertaking even more costly explicit solvent simulations to probe conformational dynamics of amyloidogenic and transiently hydrated states ([29]). Furthermore, the force fields often used to simulate formation of fibrils may be suboptimally parametrized for this purpose. Especially if a simulation were to cover all the stages of aggregation of an initially folded and subsequently unfolded (and highly hydrated) protein. The problem of unrealistically high concentrations of proteins subjected to *in silico* aggregation is recognized, as well [26]. Unsurprisingly, the majority of computational amyloid studies are focused on short peptides lacking complex topological features (e.g. linear Aβ fragments [19, 30]). On the other hand, many proteins with complex topologies defined by the presence of multiple disulfide bonds such as insulin, lysozyme or α-lactalbumin are also amyloidogenic [31-35] and there is evidence that the native disulfide bonding is preserved in the fibrillar form (e.g. [36]). Amyloid fibrils of such proteins are even more challenging to study both experimentally and computationally. As a consequence, there are fewer developed ideas as to how these complex backbone topologies adopt the amyloidal β-sheet motive.

The initial goal of this study was to assess various plausible structures of amyloid fibrils of a model disulfide-constrained amyloidogenic proteins: insulin. Our work focuses on a subset of amyloidogenic structures consisting of layers of planarized insulin monomers stacked in an in-register parallel manner. This approach is based on the following symmetry argument: For aggregates of disulfide-constrained proteins with complex topologies saturation of intermolecular contacts may be achieved through stacking of planarized monomers. However, the low symmetry of such building blocks may require an in-register alignment (and a quasi-translational assembly mode) to facilitate saturation of van der Waals interactions and hydrogen bonds between flattened monomers forming interchain parallel β-sheet structure. As planarization of a protein molecule with complex disulfide-constrained structure produces discrete pseudo-2D-topologies we compare

stability and properties of various types of insulin amyloid fibrils assembled upon in-register stacking of these distinct 2D-topologies.

Methods

There are three main stages of the multiscale approach to the modeling of insulin amyloid fibril proposed in this study. In the first part, flat (quasi-2D) insulin conformers were simulated using an implicit-solvent method. Next, these structures were docked to each other to form layered amyloid-like structures. Finally, all the resulting amyloid-like structures were validated during long all-atom MD simulations. Substantial parts of modeling (points 1, 2, 3) were carried out using CHARMM program [37] and united atom representation of protein chains with the Gaussian solvent-exclusion model, EEF1 [38]. All parameters, unless stated otherwise, had default values consistent with the parametrization of the force field: switching function for electrostatic and van der Waals terms between 7.0 Å and 9.0 Å , distance-dependent dielectric screening. The solvent frictional forces were included by performing Langevin dynamic with 5.0 fs^{-1} friction coefficient.

1. The planarization protocol

The initial coordinates of an insulin monomer conformer were taken from the hexamer structure of bovine insulin, PDB entry 2A3G [39]. The conformer was placed over a flat surface composed of 625 attractive dummy beads with fixed positions on the rectangular 125 x 125 Å grid. As the initial orientation of the insulin monomer toward the surface is likely to affect the course of planarization, we used six distinct starting spatial orientations obtained by stepwise 90-degree rotations of the monomer around X, Y, and Z axes (see Fig. S1 in Supporting Information). The dummy beads of the surface were modeled with the attractive part of Lennard-Jones potential with r_{min} = 5.0 Å and ε modified every 10 ns (ε = -0.5, -1.0, -3.0, -6.0, -9.0 kcal/mol) in a step-wise manner. The simulation's temperature was set to 400 K causing unfolding and, eventually, flattening of the monomer. We have carried out 100 such simulations for each initial spatial orientation of the

insulin monomer (Fig. S1 in Supporting Information). Finally, 102 properly flattened conformers were selected and clustered into six topological groups depending on the placement of the N/C-termini of both A- and B-chains inside or outside of the central topological loop. Subsequently, each flat conformer was subjected to optimization procedure: minimization (50 steps of steepest descent method, SD and 50 steps of Adopted Basis Newton-Raphson, ABNR) and short molecular dynamics (10 ns) between two parallel (15 Å apart) surfaces in a manner similar to the flattening process except that only the repulsive part of Lennard-Jones potential was applied. In this way, the conformers became physically relaxed while their planarity was maintained. Finally, within each topological category, a single conformation was selected based on the Procheck score of Ramachandran dihedral angles compatible with the extended structure [40].

2. Protofilament models

The flattened and energetically optimized insulin conformers were then used as building blocks for the molecular docking process. The first layer (flat conformer of insulin) was placed at the origin while its identical copy was placed 10 Å apart along the direction perpendicular to the monomer plane. During short dynamics (100 ps) the second docking fragment was pulled toward the first one by applying harmonic distance constraints between corresponding C_α carbon atoms of both layers, as in regular, in-register β-sheet structure. Constraints were removed when adjacent C_α carbons were less than 5 Å apart. The secondary structure of individual fragments was constrained during this process by holding the dihedrals near their initial values. The initial layer (docked-to) had all C_α coordinates fixed whereas the docking layer had its overall shape constrained (rigid docking). Repeating this procedure with other monomers allows us to build the fibril of a desired length.

3. All atom MD

Molecular dynamic simulations were performed using the AMBER CUDA version and FF15IPQ force-field [41-42]. All initial structures were neutralized with Na^+ / Cl^- ions and solvated

with SPC/E water molecules in a rectangular box with the smallest distance between any protein atom and the box's boundary of 16 Å. Size of the generated systems due to solvation varied between 30053 to 53024 atoms (9754 to 16406 water molecules, respectively). There were several stages of the initial equilibration process. First, each system was optimized with 200 steps of steepest descent algorithm followed by 800 steps of conjugate gradient method then heated gradually to 300 K during the first ns with backbone atoms constrained to their starting positions in the NPT ensemble followed by 5-ns-long equilibration with only C_α atoms constrained. All further production runs were carried out without any constraints imposed on heavy atoms. All bond lengths with hydrogen atoms were constrained with the SHAKE algorithm [43] which allows for integration with 2 fs step. We employed GPU code of *pmemd* with real-space cutoff of 8.0 Å and Langevin dynamics with collision frequency of 2 ps^{-1}. For each FC_1 .. FC_6 topology three independent 200-ns-long simulation runs were conducted. The most conserved structure (i.e. with the lowest RMSD) was subsequently used as a starting structure for high temperature unfolding simulations (at rate 50 K/ns) control runs and long (up to 2 μs) simulations.

4. Implicit solvent MD of long protofilaments

The docking procedure was applied to all the considered pseudo-2D topologies (FC_1..FC_6) to construct protofilaments consisting of thirty layers. For adequate solvation of such long structures in very large periodic boxes filled with water molecules the computation cost are prohibitive. Hence, we chose to simulate the long fibrils using an implicit solvation approach (EEF1 [38], details described above). All structures were optimized with 500 steps of SD minimization algorithm followed by 500 steps of ABNR method. The MD simulations were carried out for 100 ns, with three independent runs for every topology. The last ten nanoseconds of the simulation were used to acquire structural parameters of long protofilaments.

5. Definitions

β-Sheet content was defined as a fraction of residues involved in this structure as detected by the Stride algorithm [44]. RMSD was measured only for C_α atoms between the interchain disulfide bridges: residues 7-20 of A-chain and residues 7-19 of B-chain. We used the previously described approach to determine twist of fibrils [45] in which the angle of rotation is measured as dihedral angle between two vectors of two consecutive layers. Here, we use sulfur atoms in residues Cys7A and Cys20A, and Cys7B and Cys19B as two pairs of the anchor points. Width and height were measured as average parameters of the bounding box around the monomer. As many fragments of our protofilaments are highly mobile we chose to conduct the measurements in two regimes: [**i**] with all the atoms of the single layer taken into account, and [**ii**] measurements limited to the atoms within the central topological loop. In either case, conformers were aligned to the principal axes, so that the orientation would not influence the results. The average number of water molecules per layer was calculated in two stages. First, we identified atoms forming eventual cavities within the all final structures with CAVER software [46-47]. In the next stage, we searched for and counted water molecules that were within 2.5 Å distance from atoms forming the cavity's wall. This procedure is accurate only for small clusters of interior water molecules, thus each case was also visually inspected. The FC_4 aggregate contains large cluster of water molecules hence, in this case, the estimation was based on the direct counting.

Results and Discussion

The typical structural and thermodynamic features of amyloid fibrils such as high β-sheet content, high stability, and suppressed backbone fluctuations require that attractive intermolecular interactions within a fibril are effectively saturated. For a macromolecular system consisting of low symmetry building blocks one intuitive strategy to achieve this goal in a way that addresses both the short distance nature of van der Waals forces and the directionality of hydrogen bonds is to align individual structural units in a parallel in-register fashion. Following this symmetry argument we

have considered an *in silico* pathway in which planarized protein monomers are initially stacked in this manner. Structure of thus obtained oligomers is subsequently optimized using molecular dynamics. Fig. 1 depicts the basic idea of this process and the *in-silico* pathway leading from native protein to amyloid fibrils is juxtaposed to the one representing the 'real' physical process taking place either *in vivo* or *in vitro*.

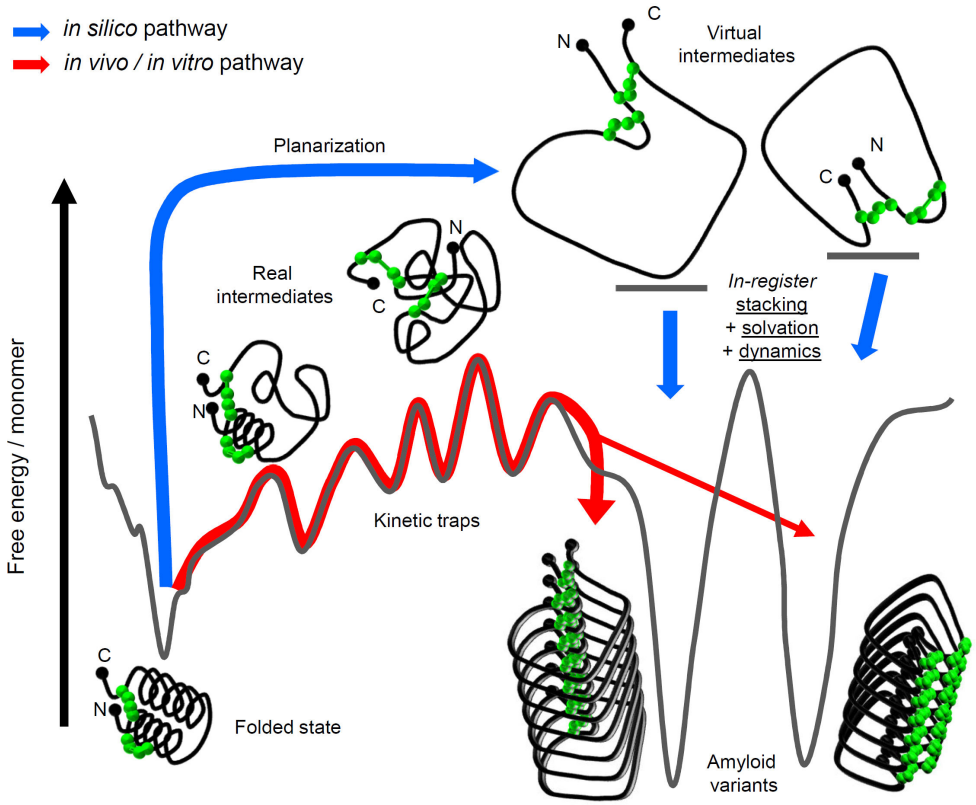

Figure 1. Relationship between the physical and the *in silico* conformational transition pathways leading form a folded protein to amyloid fibrils. The actual transition occurring via partly folded intermediate states is prone to kinetic traps (red pathway). The proposed *in silico* pathway involves initial planarization of the protein molecule to produce highly frustrated virtual intermediates which are subsequently stacked in register to form amyloid fibrils (blue pathway). Two amyloid structural variants distinct in terms of quasi-2D topology of building blocks are considered.

While an actual conformational transition of a native protein toward amyloid fibrils is likely to progress through a series of transient partly unfolded intermediates which form the layered

structure only at the final stages of aggregation the *in silico* pathway proceeds through a virtual planarized (a pseudo 2D object) intermediate. Certainly, such an intermediate would not form as a separate entity in solution due to prohibitive energetic and entropic barriers. The *in silico* pathway leading through quasi-2D intermediates mitigates the problems of kinetic traps related to residual native structure and strong local interactions within actual intermediate states. The initial in-register stacking of planar conformations is a biased process of negligible computational cost (in contrast to actual diffusion-limited aggregation of real intermediates in an explicit solvent). The hypothetical protein shown in Fig. 1 as a model contains two disulfide bonds. As a result, the flattened virtual intermediates would fall into one of four quasi-2D topological categories depending on whether N- and C-termini are placed inside or outside of the central loop determined by the backbone and the disulfide bonds (for the sake of clarity only two of these intermediates are shown in Fig. 1). Initial assemblies of in-register stacked identical quasi-2D intermediates subsequently undergo structural refinement and optimization through MD. Bovine insulin monomer extracted from the X-ray diffraction structure of the native hexamer (PDB entry: 2A3G [39]) was employed as the starting conformation for the planarization process carried out using CHARMM and implicit solvent environment [37-38]. The monomer was placed over a flat surface formed by dummy beads fixed on a rectangular grid. The planarization was induced by strong stepwise attraction between the surface and the protein molecule via modified Lennard-Jones potential (see Methods). The procedure was carried out 100 times for each of six different initial spatial orientations of the insulin monomer which could bias the outcome of the planarization process. Indeed, we found that the initial orientation impacts the success rate of the planarization (Fig. S1 – Supporting Information). We also note that certain initial orientations of the monomer appear to favor particular outcomes of the planarization procedure (i.e. quasi-2D conformations).

Snapshots of the insulin monomer undergoing the *in silico* planarization are shown in Fig. 2A. The molecule unfolds rapidly before becoming completely planar. The progress of the planarization is depicted in Fig. 2B by plotting the monomer's radius of gyration (R_g), the Z

coordinate of its center of mass, as well as the thickness (along the *Z* axis) as functions of time. The discrete points of stepwise increases in the surface-monomer attraction are marked as S_0, S_1, S_2, S_3 (corresponding to -1, -3, -6 and -9 kcal/mol energy). The abrupt increase in R_g correlates with flattening of the monomer. A successful planarization process would result in a flat conformation in which no segments of the main chains or disulfide bonds overlap each other, i.e. the thickness of thus obtained layer should be minimal and uniform and therefore compatible with in-register packing in a quasi-translational pattern. Results of planarization runs that did not meet these criteria were discarded. The filtered results of multiple planarization runs produced 102 flat structures which were clustered into 6 quasi-2D topological groups (marked as FC_1…FC_6, with the decreasing abundance of conformers belonging to given group) distinct in terms of the placement of A- and B-chains' N- and C-ends relative to the central topological loop of the insulin monomer restricted by the mid-sections of the main chains and the three disulfide bonds, see Fig. 2C. For a hypothetical molecule with zero van der Waals volume and the same as insulin's type of covalent structure, 16 distinct 2D topologies of the flatten conformers could be envisaged. Among the six observed in our simulations, the most abundant are those in which the N-termini of A-chain is placed inside of the central topological loop: FC_1 (68%), FC_2 (16%), FC_5 (2% of all cases) suggesting that with A-chain' N-terminal segment pointing inward flat conformations become more accessible. We noticed that a significant portion of the discarded conformers (which were refractory to the full planarization) had persistent 'bumps' around the Cys7A-Cys7B and Cys6A-Cys11A disulfide bonds and those correlated with A-chain's N-termini pointing outward. Overall, approximately 86 % of all flattened conformers have A-chain's N-terminal part placed inside the loop: significantly more that the total counts for A-chain's C-terminal part (21 % - see FC_6, FC_4, FC_2), and B-chain's N-terminal segment (3 % - see FC_5, FC_6). In none of these simulations, the largest free terminal segment, B-chain's C-termini, entered the loop of a planarized insulin conformer. It must be stressed that this outcome does not, in itself, prove that such a flat conformer is physically implausible. The trajectories reflecting time-dependent changes in monomer's R_g and thickness upon uniaxial compression

shown in Fig. 2B are quite similar regardless of the type of quasi-2D topology to which they ultimately lead.

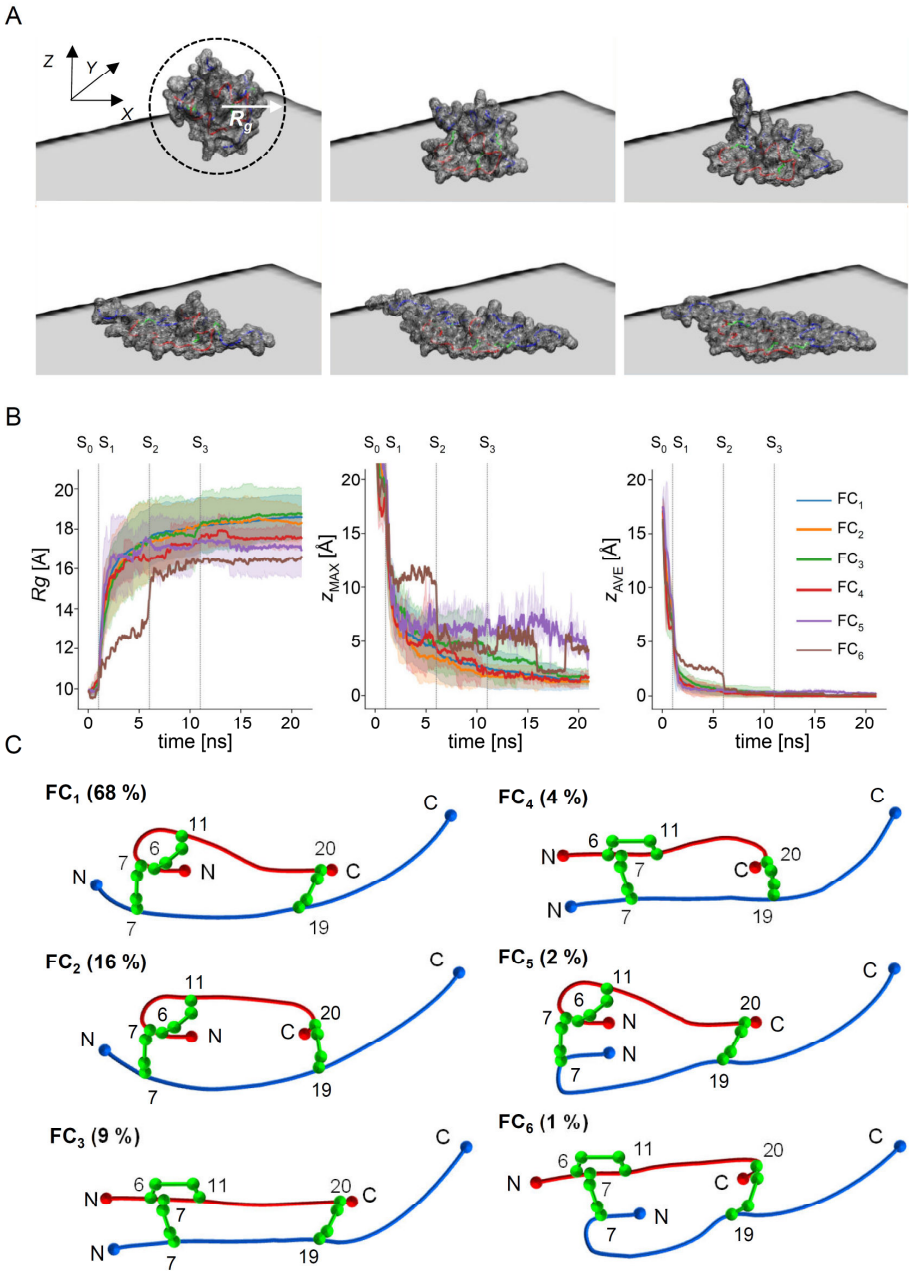

Figure 2. (A) Snapshots of an insulin monomer undergoing the *in silico* planarization process. The initial conformation is the native insulin (PDB: 2A3G); A- and B-chains are colored red and blue, respectively; disulfide bonds are represented as the green sticks and the molecular surface is shown as the gray wireframe (order: from left to right, from top to bottom). (B) the planarization process of the insulin monomer monitored by changes in radius of gyration (R_g), Z coordinate of the center of mass of insulin and maximal value of the protein's Z coordinate. Consecutive steps (S_0, S_1, S_2, S_3) of increasing attraction are marked with the vertical lines, ε = 1, 3, 6 and 9 kcal/mol. (C) The six pseudo-2D topologies of insulin monomer backbones obtained through the *in silico* planarization procedure.

The distinct appearances of the FC_6 and, to lesser degrees: FC_5 and FC_4 cases are most likely due to the low statistics and poor averaging.

Before thus obtained flat insulin monomers were used as building blocks for amyloidal self-assembly, all conformers were optimized through 10 ns MD dynamics between two parallel surfaces in which only the repulsive part of Lennard-Jones potential was employed. Selection of the best amyloid building blocks among the resulting unrestrained, yet maintaining planarity structures (in each topological category) was carried out according to the overall score of Ramachandran dihedral angles consistent with the extended structure (Methods). The docking of energetically optimized conformers was conducted by applying harmonic distance constraints between C_α atoms in aligned monomers. The process was repeated for the following layers resulting in energy optimized protofilament with two, three and four layers. The assemblies of flattened conformers representing the six observed topological classes obtained after the initial conformational optimization using all-atom MD and explicit solvent model (Methods) are shown in Fig. 3. The different packing regimes characteristic for the six stacked conformations limit the range of accessible inter- and intramolecular interactions which is reflected in the contact maps calculated for these ensembles (see Fig. S2 in Supporting Information). For example, the stretching of the B-chain found in all but FC_5 and FC_6 structures attenuates the intra-B-chain contacts resulting in the absence of off-diagonal features in the corresponding range of the contact maps (Fig. S2). The C-terminal segment of A-chain forms close contacts with the N-terminal sections of A-chain (FC_1, FC_2, FC_5) and B-chain (FC_5) in the most densely packed structures. The interactions in the FC_3 structure are dominated by interchain contacts and it is the only case where the parallel alignment of both A- and B-chains involves their whole lengths. The FC_4 structure contains a substantial void volume between separate midsections of A- and B-chains. Such a cavity, if preserved through the following stages of structural optimization, could cause either a significant van der Waals frustration, or become filled with water molecules. On the other hand the FC_5 structure exhibits

very dense packing with both N terminal segments filling the central loop, although the proximity of two charged groups could prove to be a destabilizing factor.

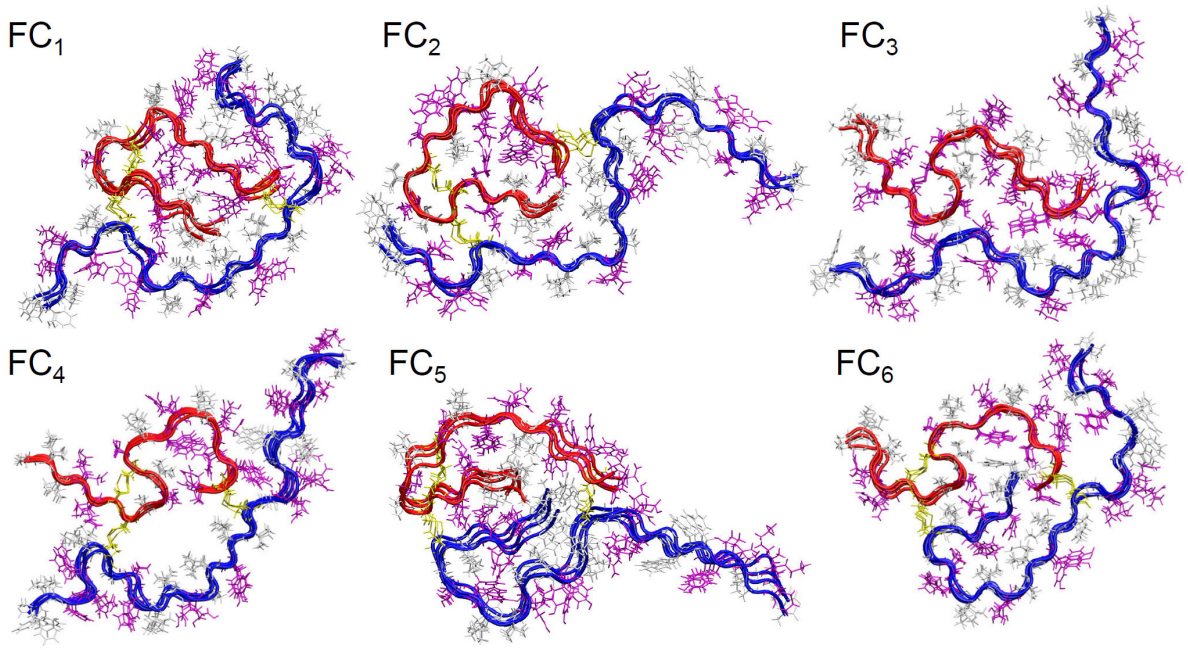

Figure 3. Initial models of insulin amyloidal assemblies obtained by in-register stacking of planarized conformers. Here, each model consists of four flat monomers. Hydrophobic side chains are colored in white; polar and charged residues are presented in magenta. The A- and B-chains are shown in red and blue, respectively; cysteine disulfide bridges are marked in yellow.

Removal of the biasing restraints imposed on these assemblies at the early stages of simulations did not cause their dissociation and unfolding during prolonged MD simulations in explicit solvent (Fig. S3, Supporting Information). While the in-register alignment of parallel β-sheets was typically maintained for all the six topological classes, removal of the constraints triggered various changes in fibril's morphology. Polymorphism of these assemblies arising from the distinct quasi-2D topologies of the ingredient insulin monomers has several consequential manifestations including the protofilament's twist (Table 1), β-sheet content, presence of void volumes, and the overall stability. The latter property is not only a function of the topological class of the building block but first and foremost of the number of stacked monomer layers. In Fig. 4, the influence of the number of stacked insulin monomers belonging to the six quasi-2D-topological classes on the stability of assemblies is probed by measuring RMSD values of C_α carbons during

isothermal all-atom MD simulations. For the sake of clarity only C_α atoms confined to the disulfide-bond-restricted core amyloid region (residues Cys7A to Cys19A and Cys7B to Cys20B) were selected.

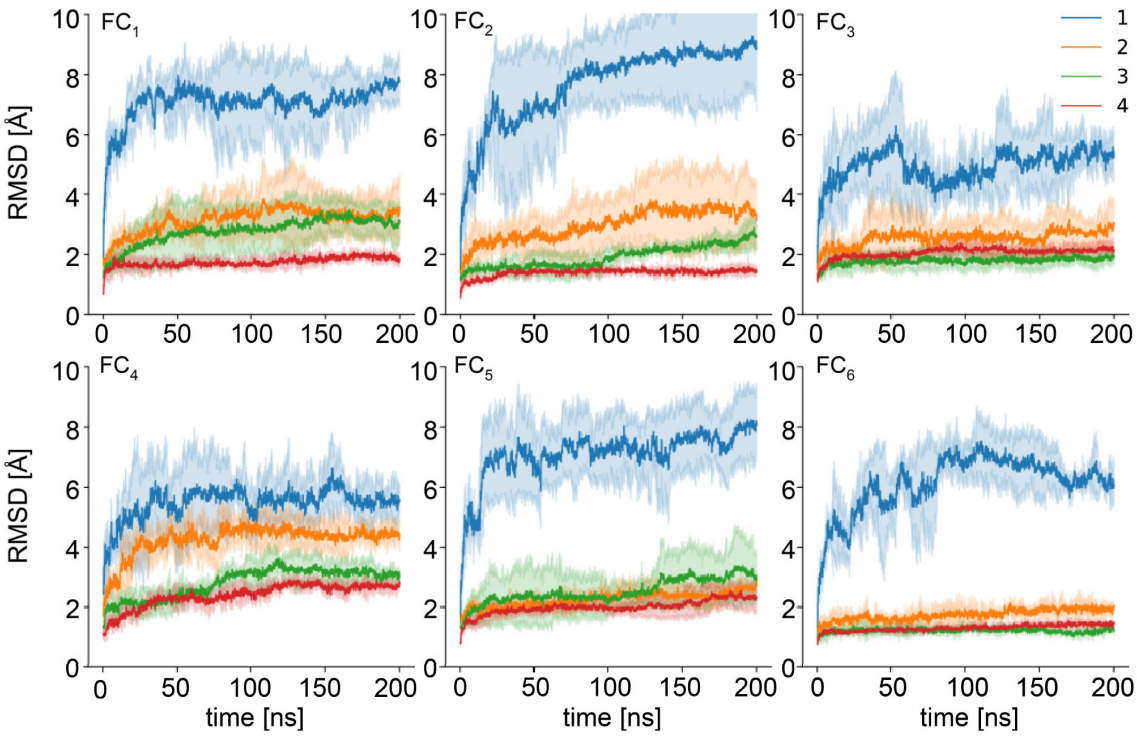

Figure 4. Stability of amyloidal stacks of the planarized insulin monomers in various 2D-topologies probed by MD simulations at 300 K: temporal changes in backbone C_α root mean square deviation are impacted by the number of layers. Only C_α carbons between the interchain disulfide bonds are included in the calculation of RMSD (single layer – blue line, two layers - orange, three layers – green, four layers – red). Thick lines represent average values of the three independent runs. Topologies FC_3, FC_5, FC_6 exhibit similar constant stability for two, three, and four layers. A single layer of FC_1 transiently adopts conformations similar to the starting structure.

In the absence of any constraints biasing the system to maintain the extended conformation, the single layers are entirely unstable: rapid increases in RMSD above 4 Å occur within the first 10 ns of the simulations irrespective of the topologic class. However, the single layer of FC_3, transiently forms conformation similar to the one present in the fibril consisting of multiple layers. The stabilizing effect of the layer multiplication is pronounced and universal becoming very clear

already for pairs of layers, especially for FC_2, FC_5, and FC_6. Although the stretches of B-chain outside the core region (residues Cys1B to Cys6B and Cys22B to Cys31B) are rather unstable, the core holds its initial state. With the addition of subsequent layers the stability builds up, however the magnitude of these gains appear to be different for some structural variants (e.g. FC_3, FC_5, FC_6, and FC_4 fibrils are already quite stable as trimers and the effect of the forth monomer is less pronounced). Importantly, the four-layered structures consistently maintain moderate RMSD values (< 2 Å) throughout the simulation (in particular: FC_1, FC_2, FC_6). The presence of FC_6 structure among them is interesting given that only one planarization trajectory led to this topology.

Steadily low RMSD values point to an elevated structural order which would be consistent with a high β-sheet content. The calculated ratios of backbone atoms involved in the canonical β-sheet conformation (as defined by the Stride algorithm [44]) during the final 100 ns of the MD simulations are reported in Fig. 5.

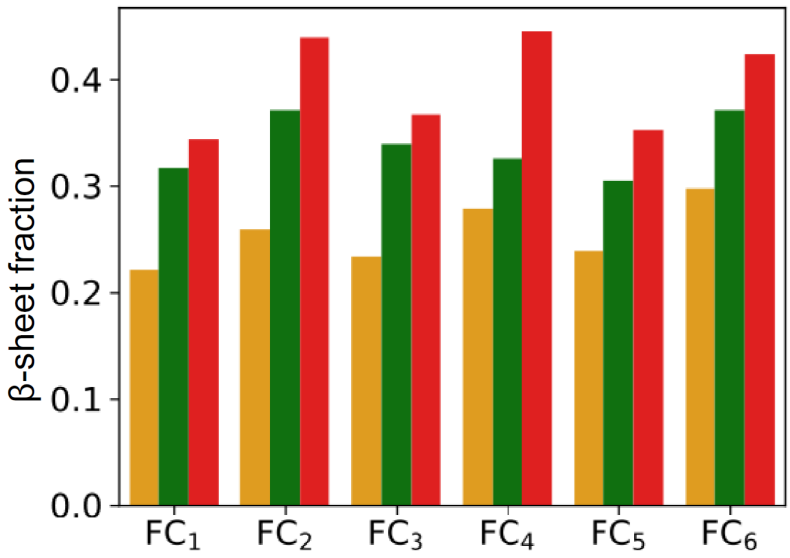

Figure 5. Average β-sheet content during the second part (100 – 200 ns) of simulations for various stacked quasi-2D topologies and different numbers of layers. β-Sheet content is defined as a ratio of the number of the C_α carbons involved in β-sheet structure (as recognized by the Stride algorithm) to all the C_α carbons in the molecule (two layers – orange bars, three layers – green bar, four layers – red bars).

Unsurprisingly, in all these cases, β-sheet becomes more abundant with the increasing number of stacked monomers. The three top scoring variants are FC_2, FC_4, and FC_6. Hence, along with the two more stable structures, the relatively fluctuation-prone FC_4 type reveals a significant presence of β-sheet, as well conformation. We have conducted an additional survey of stability of four-layered amyloidal assemblies of all six topological variants by subjecting them to temperature ramp MD simulations summarized in Fig. 6.

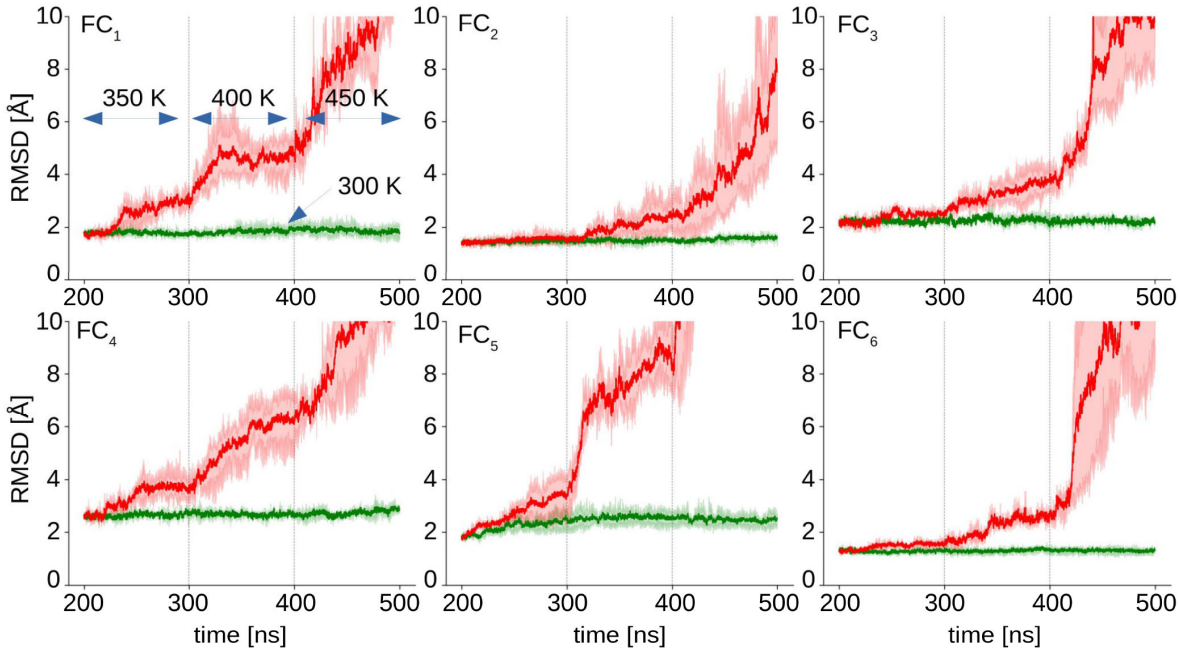

Figure 6. Relative stability of quadruple layer stacks of various topologies probed by changes in RMSD during high temperature unfolding. These simulations correspond to prolonged MD runs of the simulations shown in Fig. 4. Only C_α carbons between the interchain disulfide bonds are included in the calculation of RMSD. The temperature was maintained at 350 K for the initial 100 ns followed by jumps to 400 K, and finally 450 K as indicated with vertical dashed lines. Thick red lines denote average RMSD values from the three independent simulations. Green lines represent average RMSD from the control runs at 300 K.

The simulations were carried out at 350 K for the first 100 ns followed by the temperature jumps to 400 K and 450 K (at 400 ns) while control simulations were run at 300 K. Based on the behavior at 350 K, the six amyloid variants can be placed in the following approximate order of decreasing stability: FC_2, FC_6> FC_3, FC_5> FC_4, FC_1. Furthermore, assemblies of FC_2, FC_3 and FC_6 appear to be rather resistant to temperature increasing to 400 K contrasting with the fast structural

disruption caused in the other three isomers. The unfolding at 450 K is fastest for the FC_5 aggregate. The data presented so far suggests that the FC_2 and FC_6 assemblies are particularly stable and that this stability coincides with increased β-sheet content. What distinguishes both these isomers on the structural level is the inward placement of N- and C-termini (belonging to a single A-chain in the case of FC_2, and B-, and A-chains in FC_6) in the core region (Figs. 2C and 3). Hypothetically, the proximity of N- and C-termini could lead to a salt-bridge-like interaction between them which could contribute strongly to the overall stability of the aggregate [48-49].

In order to assess the structural features of protofilaments consisting of different insulin monomers obtained in this study we also built longer filaments composed of 30 monomer layers. Cross-sections and side views of such protofilament structures obtained for the six different quasi-2D topologies are presented in Fig. 7 along with the visualizations of internal cavities that could be filled with water molecules. In general, these multilayer structures proved to be remarkably stable during the simulations with the exception of FC_5 where the N-terminal fragments of both chains are repulsed from each other causing a significant disruption. There are marked differences in terms of fibril's twist and the number of water molecules that could be possibly sequestered within the central topological loop (Fig. 7, Table 1). The FC_3 and FC_4 fibrils reveal negative (left-handed) twist of approximately ~ 1° per monomer layer. For other structures, the obtained twist values are either even lower (FC_1, FC_2, FC_6), or impossible to determine due to the fibril's instability (FC_5). These values translate into very long helical pitches (Table 1) which contrast with those observed experimentally for mature insulin amyloid fibrils (~ 240 nm [32]). However, we have shown earlier that the thinnest fibrillar specimen of insulin amyloid (which correspond most likely to individual protofilaments, as simulated in this study) may be essentially flat, according to atomic force microscopy (AFM) [32], while the superstructural twist (which is most accessible to AFM) of several intertwined protofilaments is a an outcome of the complex interplay of various factors including twist of component protofilaments [50].

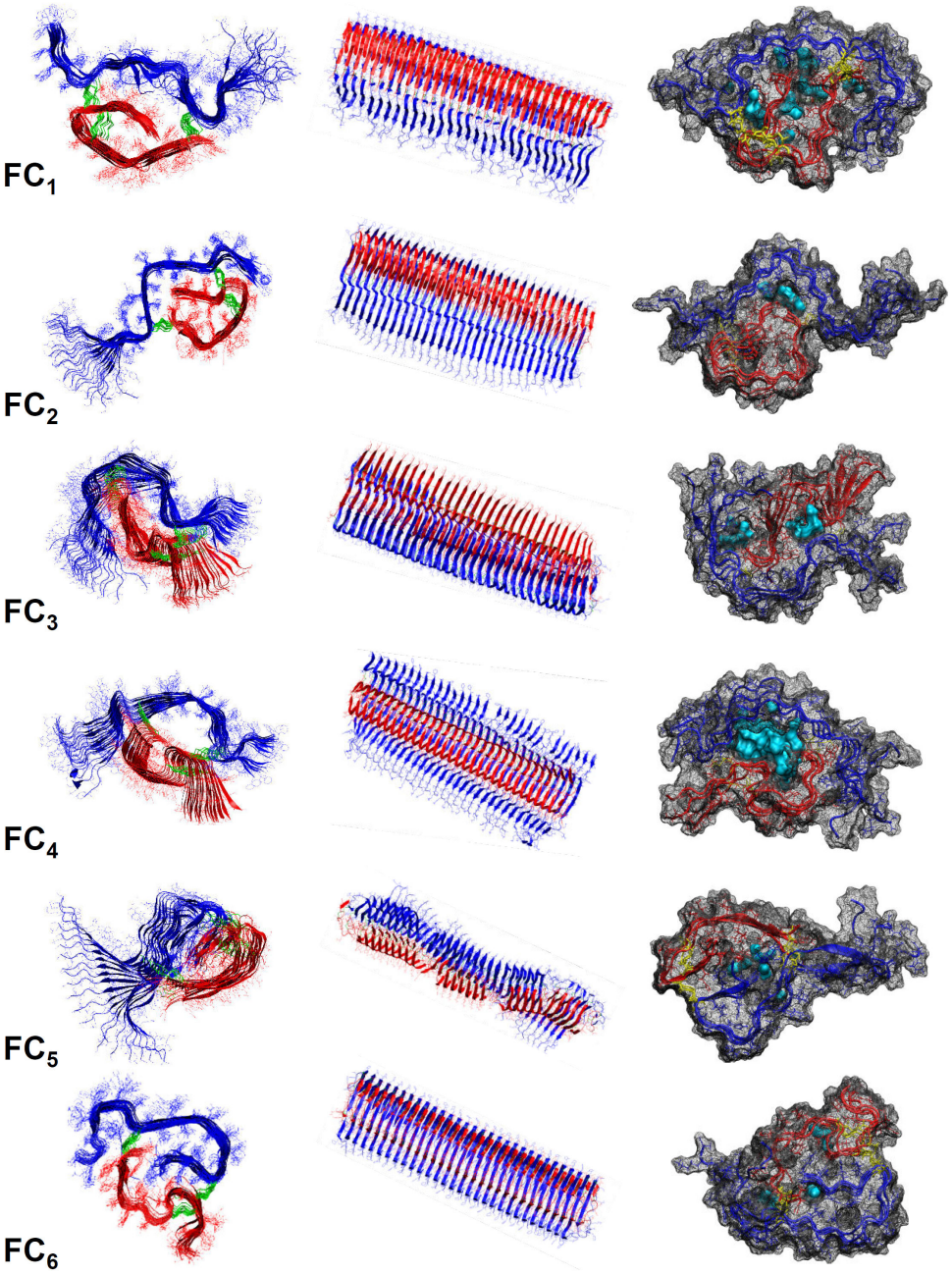

Figure 7. Insulin amyloid structures obtained by in-register stacking (followed by molecular dynamics in implicit solvent) of 30 planarized monomers of indicated quasi-2D topologies represented by final models after 100 ns simulations. A- and B-chains are marked in red and blue, respectively. In the right column, the final structures from the explicit solvent MD simulations are shown: the internal cavities within fibrils are filled with water molecules (light blue). The amounts of the sequestered water in the internal canals of FC_1..FC_6 fibrils vary significantly (Table 1).

In the case of FC_1, FC_2, and FC_6 fibrils, the negligible twist correlates with either A- or B-chain's N-termini being placed inside of the central loop. Hypothetically, fully extended (as in FC_3 and FC_4 fibrils) A- and B-chains could adopt more easily the conformational twist that is intrinsic to the canonical parallel β-sheet structure. However, should either chain become bent with its N-terminal fragment directed inward to the loop's center (FC_1, FC_2, FC_5, FC_6), the conflicting torque forces acting on the sections of A- and B-chains could diminish the effective twist of the whole protofilament. Additional studies are needed to assess plausibility of this scenario and its involvement in non-twisted insulin fibrils [32, 51].

All the mature structures examined in this study contained core void volumes which could be filled with at least a few water molecules (per monomer – Fig. 7 right column and Table 1). There have been conflicting reports on the presence of water within amyloid fibrils. Some studies postulated that amyloid fibrils are densely packed interfaces that are generally dry [15], while others suggested that the amyloid core is loosely packed [52-53] and therefore permissive to being penetrated by and filled with water [54-55]. Our previous experimental studies suggested that bovine insulin amyloid fibrils grown under typical conditions are devoid of sequestered water [56-57]. By applying the criterion of hydration of the internal canal, the six amyloid structures examined here could be divided into three groups: [i] *dry fibrils* (FC_5 and FC_6) containing 1-2 sequestered water molecules per monomer, [ii] *moderately wet fibrils* (3-5 water molecules per monomer in FC_1, FC_2, FC_3), and [iii] *water-soaked amyloid* (more than 20 H_2O molecules per insulin monomer in the case of FC_4 fibrils) which is clearly least consistent with the aforementioned experimental studies.

Previous studies pointed to the importance of certain regions of the insulin's covalent structure for the transition to amyloid fibrils. The LVEALYL section of B-chain is likely to be involved in formation of steric zipper stabilizing the fibrils [58] whereas the disulfide-bond N-terminal fragment of A-chain encompassing the first 11-13 amino acid residues has a very strong propensity to form amyloid [59-63]. In the fibrillar state, both these segments are expected to form

stable β-sheet conformations. From this perspective, it becomes crucial to look at the stability and β-sheet content of the FC_1..FC_6 four-layer assemblies at the resolution of individual amino acid residues. Fig. 8 provides an insight into local propensity to form β-sheet structure along the amino acid sequence of insulin in respect to the four-layer insulin monomer assemblies. We note that while there is not a single structure saturated in β-sheet content at both these segments simultaneously, some topologies (FC_2, FC_4, FC_6) perform better than others (FC_1, FC_3, FC_5). In Fig. S4 (Supporting Information), the parallel data on molecular fluctuations is shown.

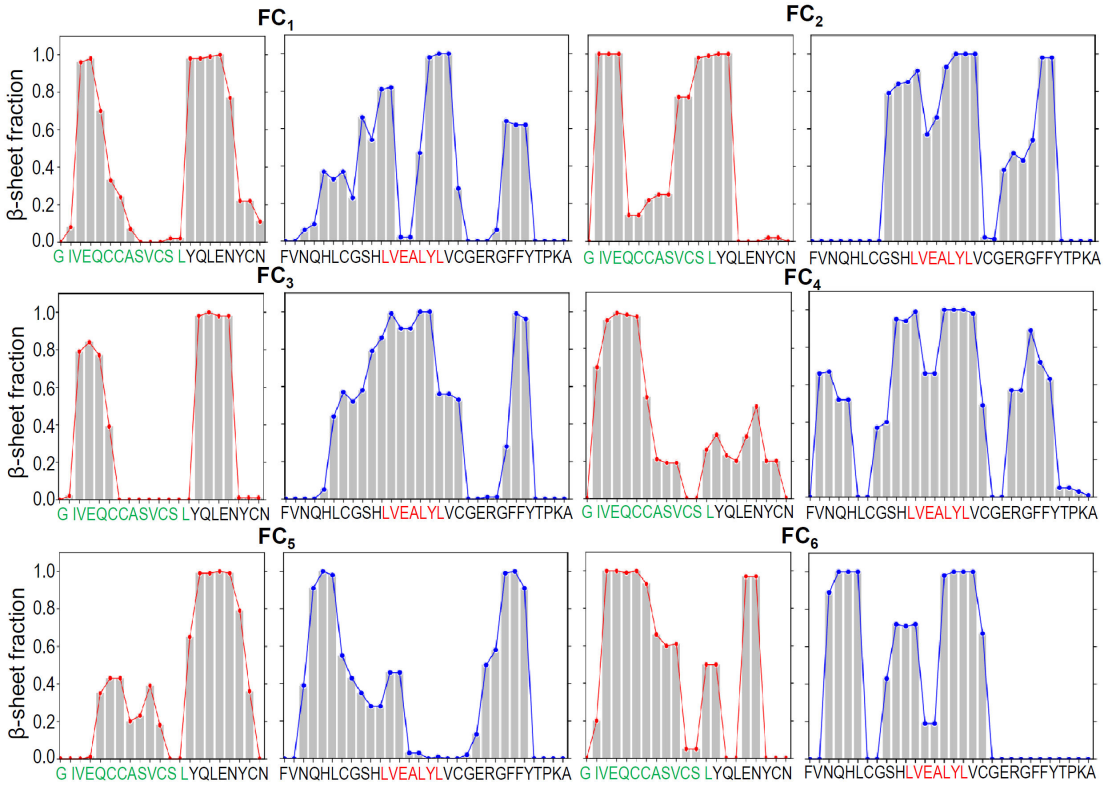

Figure 8. Local β-sheet content for amyloidal quadruple insulin monomer layers in various quasi-2D topologies along the amino acid sequence (β-sheet content is defined as a fraction of MD simulation time in which particular residue is assigned to this conformation). The calculations were based on averaging of the local conformational state over three independent MD runs (final 100 ns of a 200-ns-long simulation). The LVEALYL segment implicated in insulin amyloid steric zipper is blue-marked whereas strongly amyloidogenic N-terminal fragment of A-chain is marked in green.

We also looked at structural similarities between LVEALYL segments: as a part of the steric zipper (PDB: 3HYD, [58]) and as part of simulated fibrillar structure (Fig. 9). The requirement to

form the steric zipper appears to be best satisfied by the topologies of FC_1 (C_α RMSD < 2.0 Å), FC_3 and FC_4, followed by FC_6, FC_5 and finally FC_2 in which the turn around the residue Glu3B causes the main dissimilarity (C_α RMSD ~ 3.2 Å). Interestingly, the structures become nearly identical (C_α RMSD ~ 0.5 Å) when the leucine residue of the LVEALYL segment is neglected. Overall, the assemblies most consistent with the expected β-sheet presence in the N-terminal region of A-chain's are those of FC_6, FC_4 followed FC_2, FC_1, FC_3, and FC_5 (although molecular fluctuations in these segments are strongly dumped in FC_2 and FC_1).

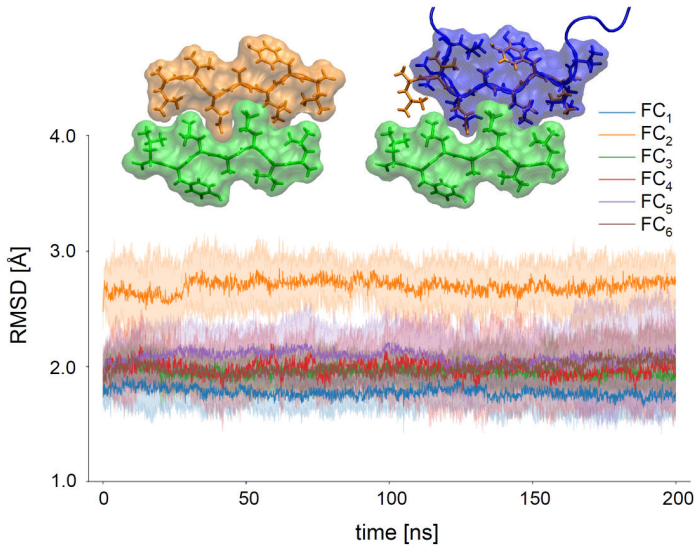

Figure 9. Similarity between LVEALYL segments: as a part of the insulin's steric zipper (PDB: 3HYD) and a part of FC_1....FC_6 fibrils (C_α RMSD profiles for all topologies). Thick lines represent the average of the three independent, 200 ns simulations. Inset: in the left panel, the crystal structure of LVEALYL fibrils (3HYD) is presented, the orange sticks and surface represent one sheet whereas green surface and green sticks depict the other sheet, together forming steric zipper; in the right panel, an example of structural alignment between 3HYD peptide and LVEALYL segment in the FC_3 topology.

In conclusion, here we have put forward and examined an alternative approach to the exceedingly challenging problem of *in silico* simulations of amyloid structures formed by proteins with complex topologies. In our approach, monomers of insulin, selected as a model amyloidogenic protein with the topology constrained by disulfide bonds, were first 'planarized' and subsequently

used as building blocks for amyloid aggregates. Through this protocol, six different amyloid structures varying in pseudo-2D topologies of flat insulin monomers were accessed. These structural variants were analyzed in terms of overall stability and compatibility with the existing low resolution experimental data on insulin amyloid fibrils, as well as high resolution structural data on fibrils derived from short insulin fragments. One of the here-presented plausible structural variants of insulin fibrils, marked as FC_6, is distinct in terms of the N-terminal part of insulin's B-chain being pointed inward to the center of the main topological loop of flattened insulin monomer with B-chain's C-terminus and A-chain's both termini protruding outside the fibril. The FC_6 variant is remarkably stable and presents a significant β-sheet content especially in the regions expected to acquire this fold upon fibrillization, namely: the steric zipper LVEALYL region, and the N-terminal fragment of A-chain. The variant is also largely devoid of internally trapped water molecules which is also in accordance with the earlier experimental studies. Remarkably, the building block topology for the FC_6 variant is the rarest among the all planarized insulin conformers. Different stages of our multiscale approach may be easily modified and supplemented by various computational methods – for example by coupling to molecular-docking tools such as CABS-dock [64]. As the phenomenon of amyloidogenesis is not restricted to simple single-chain protein and peptides further efforts are urgently needed to create tools to develop viable molecular structures of amyloid fibrils formed by proteins with complex topologies.

Table 1: Key characteristics of long protofilaments assembled from planarized insulin monomers belonging to various pseudo-2D-topological categories. Due to the instability of FC_5 fibrils the twist and helical pitch could not be determined (N.D.). The approximate number of potentially sequestered (within the central topological loop) water molecules is given per insulin monomer.

Category	*Inward directed termini*	*Twist [deg.]*	*Pitch [nm]*	*Overall stability*	*H_2O mol. per monomer*
FC_1	N- (A-chain)	0.1	1782	+	4.5
FC_2	N-, C- (A-chain)	0.0	-	++	3.5
FC_3	None	-0.9	198	++	3.5
FC_4	C- (A-chain)	-1.1	162	+	>20
FC_5	N- (A-chain), N- (B-chain)	N.D.	N.D.	-	2
FC_6	C- (A-chain), N- (B-chain)	-0.1	1782	+++	1.5

Acknowledgments:

This work was supported by the National Science Centre of Poland, grant no. 2017/25/B/ST5/02599. We are grateful for discussions to Dr. Grzegorz Łach (Faculty of Physics, University of Warsaw) and Dr. Szymon Niewieczerzał (Centre of New technologies, University of Warsaw).

References

[1] Ke, P. C., Zhou, R., Serpell, L. C., Riek, R., Knowles, T. P., Lashuel, H. A., ... & Mezzenga, R. (2020). Half a century of amyloids: past, present and future. Chemical Society Reviews, 49(15), 5473-5509.

[2] Fändrich, M., Fletcher, M. A., & Dobson, C. M. (2001). Amyloid fibrils from muscle myoglobin. Nature, 410(6825), 165-166.

[3] Chiti, F., & Dobson, C. M. (2017). Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. Annual review of biochemistry, 86, 27-68.

[4] Ruberg, F. L., Grogan, M., Hanna, M., Kelly, J. W., & Maurer, M. S. (2019). Transthyretin amyloid cardiomyopathy: JACC state-of-the-art review. Journal of the American College of Cardiology, 73(22), 2872-2891.

[5] Sacchettini, J. C., & Kelly, J. W. (2002). Therapeutic strategies for human amyloid diseases. Nature Reviews Drug Discovery, 1(4), 267-275.

[6] Cline, E. N., Bicca, M. A., Viola, K. L., & Klein, W. L. (2018). The amyloid-β oligomer hypothesis: beginning of the third decade. Journal of Alzheimer's Disease, 64(s1), S567-S610.

[7] Janson, J., Ashley, R. H., Harrison, D., McIntyre, S., & Butler, P. C. (1999). The mechanism of islet amyloid polypeptide toxicity is membrane disruption by intermediate-sized toxic amyloid particles. Diabetes, 48(3), 491-498.

[8] Fowler, D. M., Koulov, A. V., Balch, W. E., & Kelly, J. W. (2007). Functional amyloid–from bacteria to humans. Trends in biochemical sciences, 32(5), 217-224.

[9] Dueholm, M. S., Petersen, S. V., Sønderkær, M., Larsen, P., Christiansen, G., Hein, K. L., ... & Otzen, D. E. (2010). Functional amyloid in Pseudomonas. Molecular microbiology, 77(4), 1009-1020.

[10] Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., ... & Hultgren, S. J. (2002). Role of Escherichia coli curli operons in directing amyloid fiber formation. Science, 295(5556), 851-855.

[11] Gazit, E. (2002). The “correctly folded” state of proteins: is it a metastable state?. Angewandte Chemie International Edition, 41(2), 257-259.

[12] Lomakin, A., Teplow, D. B., Kirschner, D. A., & Benedek, G. B. (1997). Kinetic theory of fibrillogenesis of amyloid β-protein. Proceedings of the National Academy of Sciences, 94(15), 7942-7947.Eis

[13] Chiti, F., & Dobson, C. M. (2009). Amyloid formation by globular proteins under native conditions. Nature chemical biology, 5(1), 15-22.

[14] Tycko, R. (2011). Solid-state NMR studies of amyloid fibril structure. Annual review of physical chemistry, 62, 279-299.
[15] Eisenberg, D. S., & Sawaya, M. R. (2017). Structural studies of amyloid proteins at the molecular level. Annual review of biochemistry, 86, 69-95.
[16] Moran, S. D., & Zanni, M. T. (2014). How to get insight into amyloid structure and formation from infrared spectroscopy. The journal of physical chemistry letters, 5(11), 1984-1993.
[17] Gremer, L., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R. B., ... & Schröder, G. F. (2017). Fibril structure of amyloid-β (1–42) by cryo–electron microscopy. Science, 358(6359), 116-119.
[18] Urbanc, B., Cruz, L., Yun, S., Buldyrev, S. V., Bitan, G., Teplow, D. B., & Stanley, H. E. (2004). In silico study of amyloid β-protein folding and oligomerization. Proceedings of the National Academy of Sciences, 101(50), 17345-17350.
[19] Buchete, N. V., Tycko, R., & Hummer, G. (2005). Molecular dynamics simulations of Alzheimer's β-amyloid protofilaments. Journal of molecular biology, 353(4), 804-821.
[20] Cecchini, M., Rao, F., Seeber, M., & Caflisch, A. (2004). Replica exchange molecular dynamics simulations of amyloid peptide aggregation. The Journal of chemical physics, 121(21), 10748-10756.
[21] Gsponer, J., Haberthür, U., & Caflisch, A. (2003). The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. Proceedings of the National Academy of Sciences, 100(9), 5154-5159.
[22] Chebaro, Y., Pasquali, S., & Derreumaux, P. (2012). The coarse-grained OPEP force field for non-amyloid and amyloid proteins. The Journal of Physical Chemistry B, 116(30), 8741-8752.
[23] Atsmon-Raz, Y.; Miller, Y. A Proposed Atomic Structure of the Self-Assembly of the Non-Amyloid-β Component of Human α-Synuclein as Derived by Computational Tools. J. Phys. Chem. B 2015, 119 (31), 10005–10015.
[24] Wineman-Fisher, V.; Atsmon-Raz, Y.; Miller, Y. Orientations of Residues along the β-Arch of Self-Assembled Amylin Fibril-like Structures Lead to Polymorphism. Biomacromolecules 2015, 16 (1), 156–165.
[25] Ilie, I. M., & Caflisch, A. (2019). Simulation studies of amyloidogenic polypeptides and their aggregates. Chemical reviews, 119(12), 6956-6993.
[26] Strodel, B. (2021). Amyloid aggregation simulations: challenges, advances and perspectives. Current opinion in structural biology, 67, 145-152.
[27] Nguyen, P. H., Ramamoorthy, A., Sahoo, B. R., Zheng, J., Faller, P., Straub, J. E., ... & Derreumaux, P. (2021). Amyloid oligomers: A joint experimental/computational perspective on

Alzheimer's disease, Parkinson's disease, Type II diabetes, and amyotrophic lateral sclerosis. Chemical Reviews, 121(4), 2545-2647.
[28] Grasso, G., & Danani, A. (2020). Molecular simulations of amyloid beta assemblies. Advances in Physics: X, 5(1), 1770627.
[29] Aggarwal, L., & Biswas, P. (2021). Hydration Thermodynamics of the N-Terminal FAD Mutants of Amyloid-β. Journal of Chemical Information and Modeling, 61(1), 298-310.
[30] Luttmann, E., & Fels, G. (2006). All-atom molecular dynamics studies of the full-length β-amyloid peptides. Chemical physics, 323(1), 138-147.
[31] Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M., & Robinson, C. V. (2000). Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. Protein Science, 9(10), 1960-1967.
[32] Jansen, R., Dzwolak, W., & Winter, R. (2005). Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy. Biophysical journal, 88(2), 1344-1353.
[33] Jimenez, J. L., Nettleton, E. J., Bouchard, M., Robinson, C. V., Dobson, C. M., & Saibil, H. R. (2002). The protofilament structure of insulin amyloid fibrils. Proceedings of the National Academy of Sciences, 99(14), 9196-9201.
[34] Frare, E., Mossuto, M. F., de Laureto, P. P., Dumoulin, M., Dobson, C. M., & Fontana, A. (2006). Identification of the core structure of lysozyme amyloid fibrils by proteolysis. Journal of molecular biology, 361(3), 551-561.
[35] Goers, J., Permyakov, S. E., Permyakov, E. A., Uversky, V. N., & Fink, A. L. (2002). Conformational prerequisites for α-lactalbumin fibrillation. Biochemistry, 41(41), 12546-12551.
[36] Kurouski, D., Washington, J., Ozbil, M., Prabhakar, R., Shekhtman, A., & Lednev, I. K. (2012). Disulfide bridges remain intact while native insulin converts into amyloid fibrils. PloS one, 7(6), e36989.
[37] B. R. Brooks, C. L. Brooks III, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus: CHARMM: The Biomolecular simulation Program, J. Comp. Chem. 30, 1545-1615 (2009)
[38] T. Lazaridis and M. Karplus, Effective energy function for proteins in solution, Proteins, 35:133-152 (1999)
[39] Smith, G. David, Walter A. Pangborn, and Robert H. Blessing. "The structure of T6 bovine insulin."Acta Crystallographica Section D: Biological Crystallography 61.11 (2005): 1476-1482.

[40] Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst., 26, 283-291

[41] R. Salomon-Ferrer, A.W. Goetz, D. Poole; S. Le Grand, and R.C. Walker. (2013) "Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald."J. Chem. Theory Comput. 9, 3878-3888

[42] Debiec, Karl T., et al. "Further along the road less traveled: AMBER ff15ipq, an original protein force field built on a self-consistent physical model." Journal of chemical theory and computation 12.8 (2016): 3926-3947.

[43] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. J. Comput. Phys. 1977, 23 (3), 327–341.

[44] Frishman D, Argos P. Knowledge-based protein secondary structure assignment. Proteins. 1995 Dec;23(4):566-79. doi: 10.1002/prot.340230412. PMID: 8749853.

[45] Periole, Xavier, et al. "Energetics underlying twist polymorphisms in amyloid fibrils." The Journal of Physical Chemistry B 122.3 (2018): 1081-1091.

[46] Brezovský J., Kozlíková B., Damborský J.: Computational Analysis of Protein Tunnels and Channels, Protein Engineering: Methods and Protocols, Methods in Molecular Biology, Uwe T. Bornscheuer and Matthias Hohne (eds.),, vol. 1685, pp 25-42, 2018, 978-1-4939-7364-4.

[47] Jurcik, A., Bednar, D., Byska, J., Marques, S. M., Furmanova, K., Daniel, L., Kokkonen, P., Brezovsky, J., Strnad, O., Stourac, J., Pavelka, A., Manak, M., Damborsky, J., Kozlikova, B. : CAVER Analyst 2.0: Analysis and Visualization of Channels and Tunnels in Protein Structures and Molecular Dynamics Trajectories., Bioinformatics, bty386, 2018.

[48] Kumar, S., & Nussinov, R. (1999). Salt bridge stability in monomeric proteins. Journal of molecular biology, 293(5), 1241-1255.

[49] Tarus, B., Straub, J. E., & Thirumalai, D. (2006). Dynamics of Asp23− Lys28 salt-bridge formation in Aβ10-35 monomers. Journal of the American Chemical Society, 128(50), 16159-16168.

[50] Gruziel, M., Dzwolakc, W., & Szymczak, P. (2013). Chirality inversions in self-assembly of fibrillar superstructures: a computational study. Soft Matter, 9, 8005-8013.

[51] Devlin, Glyn L., et al. "The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner." Journal of molecular biology 360.2 (2006): 497-509.

[52] Perutz, M. F., Finch, J. T., Berriman, J. & Lesk, A. (2002). Amyloid fibers are water-filled nanotubes. Proc. Natl Acad. Sci. USA, 99, 5591–5595.

[53] Kishimoto, A., Hasegawa, K., Suzuki, H., Taguchi, H.,Namba, K. & Yoshida, M. (2004). b-Helix is a likely core structure of yeast prion Sup35 amyloid fiber. Biochem. Biophys. Res. Commun. 315, 739–745

[54] Thirumalai, D., Govardhan Reddy, and John E. Straub. "Role of water in protein aggregation and amyloid polymorphism." Accounts of chemical research 45.1 (2012): 83-92.

[55] Krone, Mary Griffin, et al. "Role of water in mediating the assembly of Alzheimer amyloid-β Aβ16− 22 protofilaments." Journal of the American Chemical Society 130.33 (2008): 11066-11072.

[56] Dzwolak, W., Loksztejn, A., & Smirnovas, V. (2006). New Insights into the Self-Assembly of Insulin Amyloid Fibrils: An HD Exchange FT-IR Study. Biochemistry, 45, 8143-8151.

[57] Surmacz-Chwedoruk, W., Malka, I., Bożycki, Ł., Nieznańska, H., & Dzwolak, W. (2014). On the Heat Stability of Amyloid-Based Biological Activity: Insights from Thermal Degradation of Insulin Fibrils. PLoS ONE, 9(1).

[58] Ivanova, M. I., Sievers, S. A., Sawaya, M. R., Wall, J. S., & Eisenberg, D. (2009). Molecular basis for insulin fibril assembly. Proceedings of the National Academy of Sciences, 106(45), 18990-18995.

[59] Piejko, M., Dec, R., Babenko, V., Hoang, A., Szewczyk, M., Mak, P., & Dzwolak, W. (2015). Highly Amyloidogenic Two-chain Peptide Fragments Are Released upon Partial Digestion of Insulin with Pepsin*♦. Journal of Biological Chemistry, 290(10), 5947-5958.

[60] Dec, R., Koliński, M., & Dzwolak, W. (2019). Beyond amino acid sequence: disulfide bonds and the origins of the extreme amyloidogenic properties of insulin's H-fragment. The FEBS journal, 286(16), 3194-3205.

[61] Dec, R., & Dzwolak, W. (2020). Extremely Amyloidogenic Single-Chain Analogues of Insulin’s H-Fragment: Structural Adaptability of an Amyloid Stretch. Langmuir, 36(41), 12150-12159.

[62] Dec, R., Puławski, W., & Dzwolak, W. (2021). Selective and stoichiometric incorporation of ATP by self-assembling amyloid fibrils. Journal of Materials Chemistry B, 9(41), 8626-8630.

[63] Dec, R., & Dzwolak, W. (2021). A tale of two tails: Self-assembling properties of A-and B-chain parts of insulin's highly amyloidogenic H-fragment. International Journal of Biological Macromolecules, 186, 510-518.

[64] Koliński, M., Dec, R., & Dzwolak, W. (2021). Multiscale Modeling of Amyloid Fibrils Formed by Aggregating Peptides Derived from the Amyloidogenic Fragment of the A-Chain of Insulin. International Journal of Molecular Sciences, 22(22).