doi:10.1016/j.jmb.2004.11.002 J. Mol. Biol. (2005) 345, 1213–1227

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

A Comparative Study of Amyloid Fibril Formation by Residues 15–19 of the Human Calcitonin Hormone: A Single β-Sheet Model with a Small Hydrophobic Core

Nurit Haspel[1], David Zanuy[2], Buyong Ma[3], Haim Wolfson[1] and Ruth Nussinov[3,4]*

[1]*School of Computer Science Faculty of Exact Sciences Tel Aviv University, Tel Aviv 69978, Israel*

[2]*Laboratory of Experimental and Computational Biology NCI-Frederick, Bldg 469, Rm 151, Frederick, MD 21702 USA*

[3]*Basic Research Program SAIC-Frederick, Inc., Laboratory of Experimental and Computational Biology NCI-Frederick, Bldg 469 Rm 151, Frederick, MD 21702 USA*

[4]*Department of Human Genetics, Sackler Faculty of Medicine, Sackler Institute of Molecular Medicine, Tel Aviv University, Tel Aviv 69978 Israel*

**Corresponding author*

Experimentally, the human calcitonin hormone (hCT) can form highly stable amyloid protofibrils. Further, a peptide consisting of hCT residues 15–19, DFNKF, was shown to create highly ordered fibrils, similar to those formed by the entire hormone sequence. However, there are limited experimental data regarding the detailed 3D arrangement of either of these fibrils. We have modeled the DFNKF protofibril, using molecular dynamics simulations. We tested the stabilities of single sheet and of various multi sheet models. Remarkably, our most ordered and stable model consists of a parallel-stranded, single β-sheet with a relatively insignificant hydrophobic core. We investigate the chemical and physical interactions responsible for the high structural organization of this single β-sheet amyloid fibril. We observe that the most important chemical interactions contributing to the stability of the DFNKF organization are electrostatic, specifically between the Lys and the C terminus, between the Asp and N terminus, and a hydrogen bond network between the Asn side-chains of adjacent strands. Additionally, we observe hydrophobic and aromatic π stacking interactions. We further simulated truncated filaments, FNKF and DFNK. Our tetra-peptide mutant simulations assume models similar to the penta-peptide. Experimentally, the FNKF does not create fibrils while DFNK does, albeit short and less ordered than DFNKF. In the simulations, the FNKF system was less stable than the DFNK and DFNKF. DFNK also lost many of its original interactions becoming less organized, however, many contacts were maintained. Thus, our results emphasize the role played by specific amino acid interactions. To further study specific interactions, we have mutated the penta-peptide, simulating DANKF, DFNKA and EFNKF. Here we describe the model, its relationship to experiment and its implications to amyloid organization.

Keywords: amyloid conformation; β-sheet; calcitonin; network of interactions; hydrophobic core

Introduction

The process of protein misfolding and fibrillar aggregation is associated with a large number of pathologically unrelated human diseases such as Alzheimer, Parkinson and Huntington,[1–4] all sharing the appearance of insoluble fibrillar deposits of proteins.[5] The proteins involved in these processes are unrelated in sequence or in function. Despite this unrelatedness, the general similarity in morphology has led to many attempts to obtain generalized models that would explain and characterize amyloid fiber formation.[6,7] It is generally assumed that the amyloid fiber organization is compatible with a cross-β structure.[8] However, the details of the organization and the interactions are not always clear, since usually no direct 3D information is available. Understanding the amyloid ultra-structure organization should yield clues into the chemical mechanisms that lead to the formation of these invasive fibers.[9] Current

Abbreviations used: hCT, human calcitonin hormone; MD, molecular dynamics.

E-mail address of the corresponding author: ruthn@ncifcrf.gov

experimental techniques can provide only either short-range distances or mesoscopic levels of organization, but are unable to provide high-resolution data. Recently, solid-state NMR studies were able, in some cases, to provide information regarding the backbone and side-chain arrangement of a single amyloid strand or sheet[10,11] but no such information regarding the internal organization of the human calcitonon (hCT) system is available. From the computational standpoint, achieving the goal of finding a structural model for a molecular system is a challenging task. The increase in the computational capabilities now allows us to enlarge the size of the systems that can be studied by atomistic molecular simulations. Within this context, molecular dynamics (MD) simulations have become a major tool for understanding conformational and aggregation phenomena at the molecular level.[12,13] Of particular note, using MD simulations, the first kinetic stages of β-sheet formation are currently being intensively explored.[12,14]

Our group has been studying the amyloid problem from a different starting point. We assume that the final fiber micro organization holds clues to the amyloid aggregation. Experimentally, it is well established that fibril growth initiates from a minimal pre-formed, soluble self-assembled seed complex that must retain its organization in water. This is essential in order for the seed to assume the role of a structural template for the final growth process. Thus, rather than sampling many micro states of aggregation of a very small number of molecular assemblies, we explore possible final arrangements of molecular complexes and evaluate their relative stabilities.[15–18] Thus, we do not attempt to directly assess the process of amyloid formation, which would imply simulating time length far longer than the current computational capabilities. Instead, we model the structure of the minimal seed organization. Further, due to computational limitations, we do not study the final shape and size, but a number of molecular species that could mimic the minimal organization requirements. Once we have built different combinations, through high temperature MD simulations we establish the relative stabilities of the models. Those models that are able to retain some degree of self organization for the time simulated, would be good candidates to study the structural properties of the amyloid fibers formed by a given sequence. This strategy is based on systematic exploration of the molecular self-assembly of a system consisting of a small number of copies of the same peptide. This strategy, combined with the analysis of the explored models, has proven very useful. Based on the models, we have been able to interpret some experimental results and to suggest an elongation mechanism for the proto filament of a prion peptide segment (residues 113–120).[16] The encouraging results have led us to explore other sequences: we rationalized the observed structural preferences of Aβ peptide segments[17,19] and predicted the micro structural differences of the islet amyloid polypeptide segments 22–27 and 22–29[15,18] and the role of some critical amino acid residues in the peptide segment 22–27.[15] In all cases, the organization of the smallest ordered structure always implied lateral association, either of sheets in the shorter peptides, or of peptide segments of the longer sequences.[15]

Here, we extend our study of the structural and chemical properties to the amyloid forming peptide fragment DFNKF (residues 15–19 of human calcitonin, hCT) and some of its tetra and penta peptide single point mutations. Human calcitonin[20] is a 32 amino acid residue polypeptide hormone, produced by the thyorid C-cells. It is related to calcium homeostasis. Amyloid fibril formation of hCT is associated with medullary carcinoma of the thyroid.[21,22] It was shown that hCT has little secondary structure at room temperature. However, hCT fibrils were found to be highly ordered, consisting of both α-helix and β-sheet elements. Experimentally, the penta peptide fragment DFNKF is known to form fibrils similar to those formed by the entire 32-residue hormone.[23,24] Studies performed on truncated segments of DFNKF showed that the DFNK tetra peptide was also capable of forming fibrils, although bifurcated and much shorter than the ones obtained for the penta peptide. On the other hand, the FNKF segment did not form detectable amyloid fibrils.

Based on our studies on the other amyloid systems, we assumed that due to its short length the DFNKF protofilament also had to consist of at least two laterally associated sheets. We tested DFNKF models with practically all potential arrangements of parallel/antiparallel strands within sheets, and of sheets with respect to each other. Our extensive efforts spanned 18 MD explicit water simulations with over 72 nanoseconds. Nevertheless, neither models with two sheets of four strands in each nor those with three sheets of three strands per sheet, had sufficient stability to explain the experimental results. Detailed examination of the more stable multi-sheet systems within these illustrated that for those models that contained parallel-stranded sheets, while the sheets separated from each other, the sheets themselves kept largely intact. Hence, we concluded that the most probable organization for the protofilament structure of the DFNKF peptide is a single sheet layer. This has led us to test a long, parallel, nine-stranded single layer β-sheet. The model consisting of a single β-sheet with only a small hydrophobic core maintained the highest degree of bulk β-sheet organization and preserved the largest amount of backbone and side-chain interactions during the simulation. Moreover, preliminary results (H. H. (G) Tsai *et al.*, unpublished results) of parallel tempering simulations over a temperature range of 300–600 K have further confirmed that a single sheet with parallel strands is the most stable organization for this short peptide. This raised the interesting question of the origin of the chemical

interactions stabilizing this single layer β-sheet in water.

Here we present a meticulous characterization of the conformational and structural properties of this model and its validation using the available experimental data. We found that the parallel-stranded single β-sheet DFNKF was stabilized largely by electrostatic interactions and hydrogen bonds with π stacking and hydrophobic interactions also playing a role. Finally, we tested the effect of specific single point mutations of the DFNKF sequence. Experimentally, when the two phenylalanine residues of DFNKF are replaced by alanine no fibrils are detected.[24] Partial experimental data are also available regarding the effect of changing only one of the Phe residues to Ala. The question arises as to whether both Phe residues are equally critical for the amyloid formation or whether one of them contributes more than the other to the ability of DFNKF to form fibers. Similarly, if charge is the important factor in the stability of this single layer sheet, then it is interesting to examine whether a substitution of Asp by Glu at the N terminus (DFNKF to EFNKF) would lead to an equally stable fibril. Consequently, we have also simulated EFNKF to look into the question of charge *versus* side-chain length at this position.

The outer surface protein A (OspA) contains a three-stranded single layer hydrophilic β-sheet flanked by two globular domains that interact with the edge strands. Koide *et al*.[25] have enlarged the sheet by duplicating the β-hairpin to yield two constructs, the first consisting of five and the second of seven β-strands. They have shown that in both designed constructs, the single layer β-sheet which is largely devoid of a hydrophobic core is relatively stable. This particular structure led them to speculate that single sheets may be involved in amyloid formation. The high stability of our DFNKF model suggests that depending on the sequence, for a single sheet with a small hydrophobic core to be stable, the interaction of edge strands with globular domains is not an absolute requirement. Further, in the designed experimental OspA sheet, the strands were covalently joined by a hairpin. Here, the absence of chain connectivity between the strands allows them larger freedom to adopt a conformation that optimizes their interactions. This may explain why in the designed construct a sheet with two inserted hairpin structures was less stable than a sheet with one added hairpin. In our case here, amyloid elongation through strand addition stabilizes the protofibril. Sheets with more strands are more stable, probably due to this optimization.

Results and Discussion

The DFNKF penta-peptide multi-sheet systems: bulk organization

Our initially built models consisted of two four-stranded sheet octamers and three three-stranded sheet nonamers. Octa-peptides and nona-peptides are good candidates for simulations, since their small size makes them suitable for atomic studies. The two and three-sheet structure enables some strands to be hidden from water and some to be exposed, thus allowing us to study the effect of different surroundings of both the peptide and the solvent on the association process. The basic set of models differed in the organization of the strands within the sheets (parallel or antiparallel) and of the sheets with respect to one another (again, parallel or antiparallel; see Supplementary data for more details of the two and three-sheet structures). Each model was simulated for four nanosecond at 350 K. To quantitatively assess the structural bulk stability of the models and to give a general view of the stability, we used geometric criteria, the retention of the inter-strand and inter-sheet distances during the simulation. We previously developed these criteria[15,18,26] since using existing criteria such as RMSD cannot give us an absolute assessment of the structural stability, only a measurement relative to a known structure, which is unavailabe in our case. These criteria are detailed in the Supplementary data.

After testing many possible two and three-sheet models with different arrangements of parallel/antiparallel strands and sheets, we realized that the strands should probably be parallel within the sheets. Several indications pointed to this conclusion: the inter-strand distance parameters remained much more constant for the parallel models than for the antiparallel models. In addition, in the antiparallel models, the ability to retain the main-chain hydrogen bonds appeared to be seriously compromised compared to the parallel models. The parallel stranded models showed an interesting phenomenon: at a very early stage of the simulation the two sheets got separated from one another due to steric hindrance and each sheet continued to be rather stable on its own, behaving like a self-containing system. The geometry of each sheet got self optimized at a very early stage of the simulation, with the strands of each sheet twisting with respect to one another (a classical characteristic of a cross-β-structure), remaining twisted throughout the entire simulation. This is visible when comparing the conservation of the inter-sheet distances with the intra-sheet distances in both types of models. The former presents a higher degree of variation than the latter, providing an indirect indication of a loose association, i.e. a loss of organization. As a consequence, each sheet separately exhibits the characteristic interactions of the single sheet model.

Due to the fact that none of the two-sheet models remained stable enough throughout the simulation, we decided to test models composed of three sheets with three strands in each. One advantage of this arrangement is that there is one sheet (the middle one), which is completely buried between the other two, shielded from the solvent. Since most of the

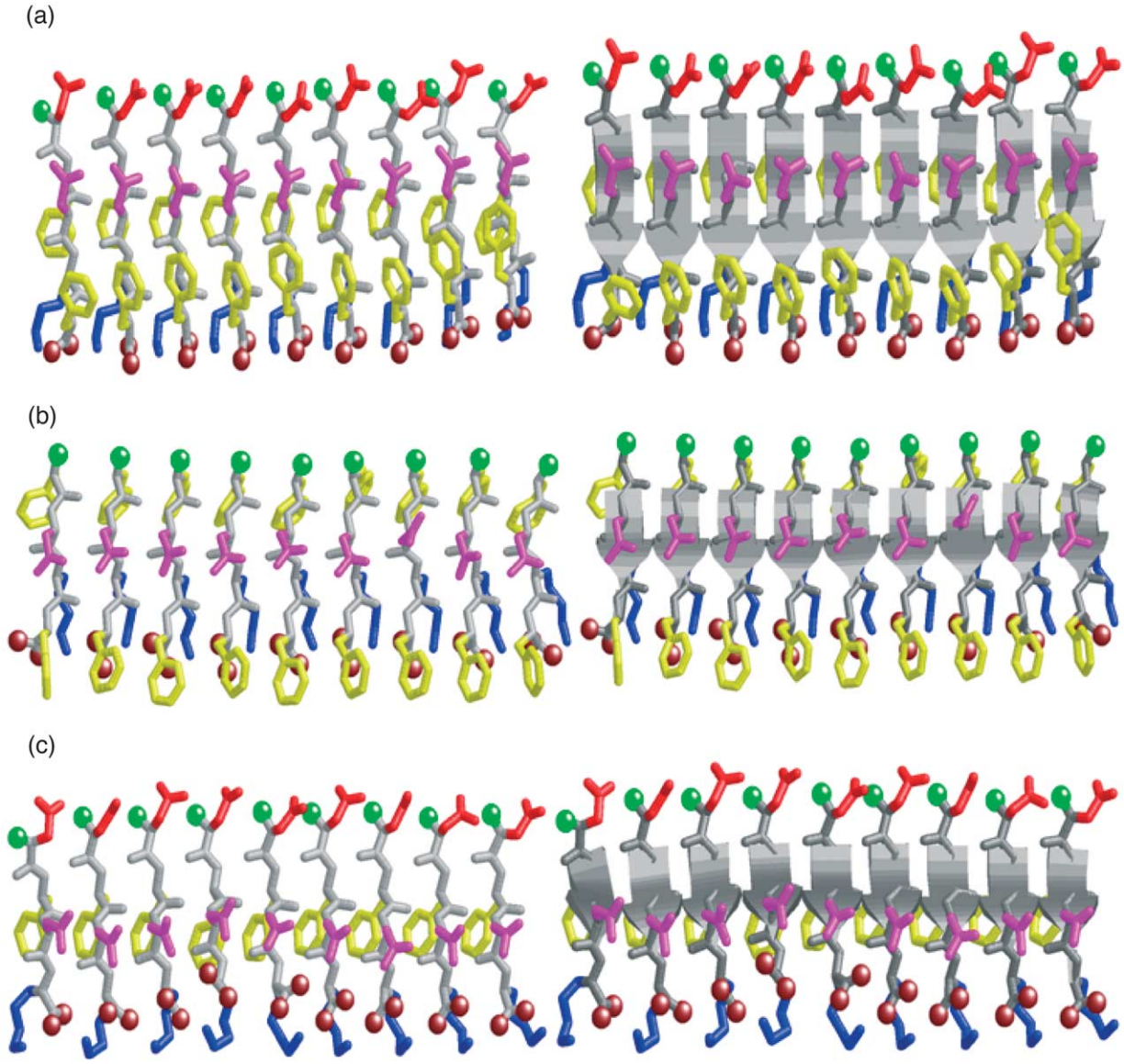

Figure 1 (*legend next page*)

residues have hydrophilic side-chains, it is mainly relevant to the Phe aromatic rings. Having at least one buried sheet allows us to study the possible hydrophobic interactions taking place in the system. Further, in three-sheet models the complex nature of amyloid formation may illustrate some differences in the relative stabilities when the number of associated sheets in the molecular dynamics models is changed. For example, the study of the related amyloidogenic NFGAIL system[15] presents a case where the association of more sheets contributed dramatically to the stability of the amyloid model system: while two-sheet models presented a low capacity to maintain an acceptable degree of order, three-sheeted models showed a more than acceptable ordered organization. All our three-sheet models had parallel strands within the sheets and most had antiparallel sheets (except for one, where two of the sheets are parallel). Here, as in the two-sheet case, the inter-sheet geometrical and chemical criteria were less constant than the intra-sheet ones, showing again a clear tendency to a looser organization between sheets. In addition, the bulk structure of the three-sheet models was less stable than the two-sheet models and tended to dissociate more quickly. Based on the results of the two and three-sheet systems, we inferred that adding more sheets does not contribute to the stability of the system. We realized that each sheet formed its own system of interactions, which led us to conclude that the system may be a one-sheet system. We decided, based on this observation, to simulate a long one-sheet system of parallel strands.

The DFNKF penta-peptide one-sheet system: bulk organization and interactions

The one sheet model was designated Mono1. Figure 1(a) presents an atomistic representation of

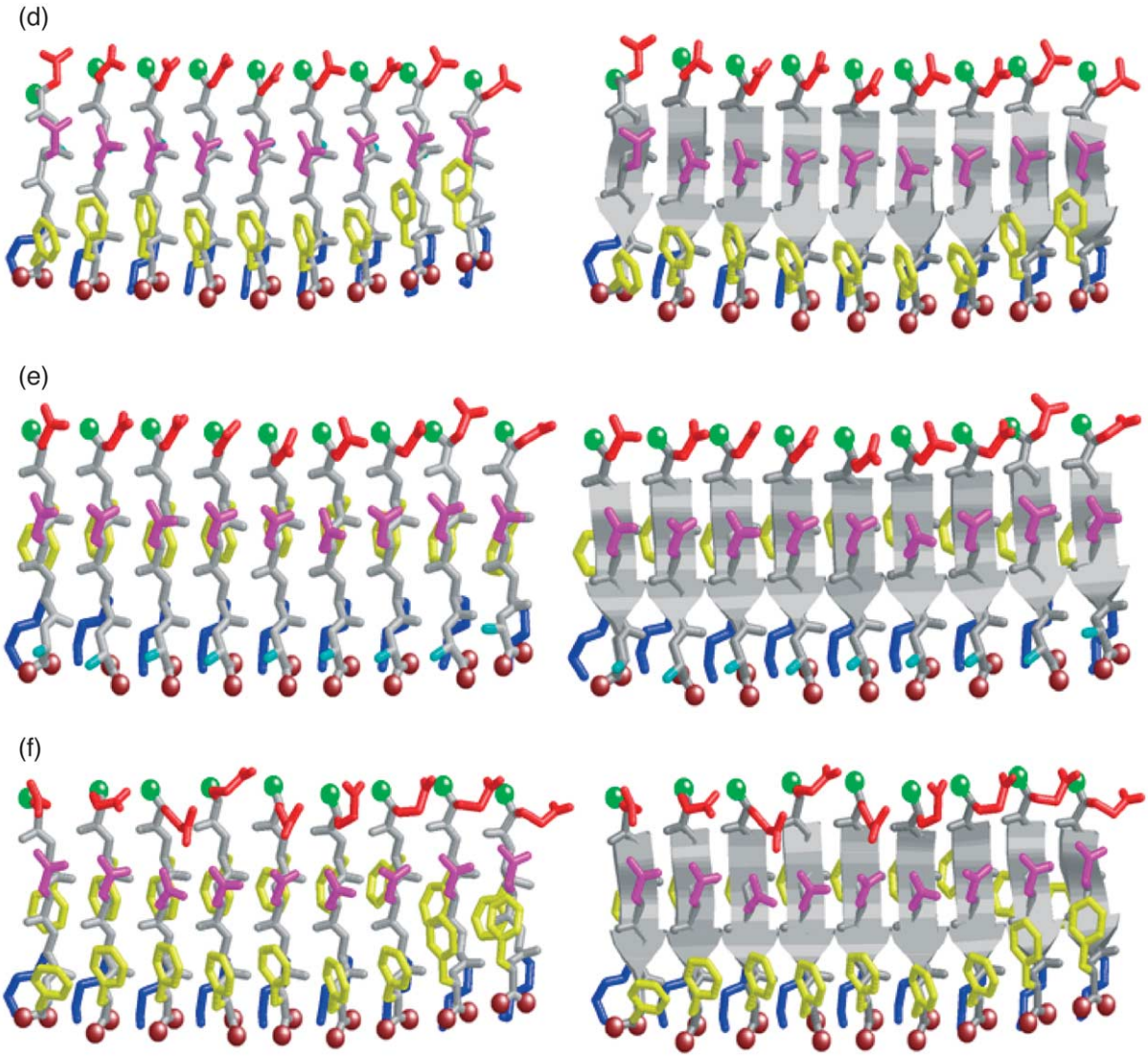

Figure 1. Atomistic representation of the one-sheet models. In all representations hydrogen atoms have been deleted for clarity. Main-chain atoms are depicted in gray while side-chains are colored depending on their chemical nature. Asp is represented in red, Phe in yellow, Asn in magenta, Lys in blue, Ala in cyan, N termini in green space-fill and C termini in brown space-fill. Each model is depicted in two different representations. On the left side atoms and bonds are represented by sticks whereas on the right side the main-chain conformation is depicted by arrows representation, which also shows the different inter strand orientations for each model. From the top to the bottom: DFNKF, Mono1 (a), FNKF (b), DFNK (c), DANKF (d), DFNKA (e) and EFNKF (f).

the minimized model. The strands stack upon each other, with no twist between them. A nine-stranded system was selected based on the assumption that a sheet of this size provides sufficient information about the side-chain interactions underlying the system, while saving computational time that may be very long for a larger system with more strands. Such a sheet should also be large enough to maintain sufficient regularity of the β-sheet organization. It should be mentioned that we did not build a fiber but a protofilament. Thus, its diameter is much smaller than that observed by electron microscopy (EM) (mesoscopic size). What is observed in EM is the association of few protofilaments to form a filament.

The one-sheet parallel-stranded model was very stable throughout the entire simulation time, compared to the other models. The only significant change in the organization of this single sheet model was the spontaneous twist of the strands with respect to one another after less than 100 picoseconds of simulation. The strands remained twisted thereafter, with the β-sheet organization being very stable and maintaining a high degree of structural order. Table 1 provides the average twist angles between two adjacent strands for the single sheet models at different times of the simulation. The average twist angle was around 20° and changed only a little during the simulation.

We analyzed the structural parameters of the system and the stabilizing interactions. In terms of bulk structure, the structural analysis provides clear

Table 1. Twist angle (deg.) of the stable one-sheet models during the simulation

	Mono1	DANKF[a]	DFNKA[a]	EFNKF(a)[a]	EFNKF(b)
100 picoseconds	19.5875	21.0429	23.4571	19.2429	20.5
1 nanosecond	18.7875	28.2	23.0429	21.7714	20.3375
2 nanoseconds	18.075	35.8143	39.2286	22.3143	20.7125
3 nanoseconds	23.6875	45.6	43.9571	20.3857	20.5625
4 nanoseconds	19.2	45.8857	41.2286	26.7857	19.75

[a] The data refer to strands no. 1–8 only.

indications to the high organization of the protofilament. The main structural parameter used to initially evaluate the organization of the one-sheet model was the percentage of main-chain hydrogen bonds (with respect to the minimized structure) retained by the system. As seen in Figure 2(a), model Mono1 retained at least 90% of its hydrogen bonds throughout the entire simulation.

Analyses of the side-chain interactions further rationalize this high stability of the DFNKF single-layer sheet arrangement. We observe a network of hydrogen bonds between adjacent Asn side-chains. This interaction functions as a pivot that holds the system together. Figure 3 shows the evolution in the number of Asn side-chain hydrogen bonds during the simulation of the single-sheet model. Two different distance and angle criteria (detailed in the Figure legend) are employed to locate weaker hydrogen bonds and dipole interactions. These parameters oscillated more than the backbone structural characteristics, since the Asn side-chain is more flexible and free to move than the backbone atoms. Furthermore, side-chains are more exposed to the solvent and tend to form hydrogen bonds

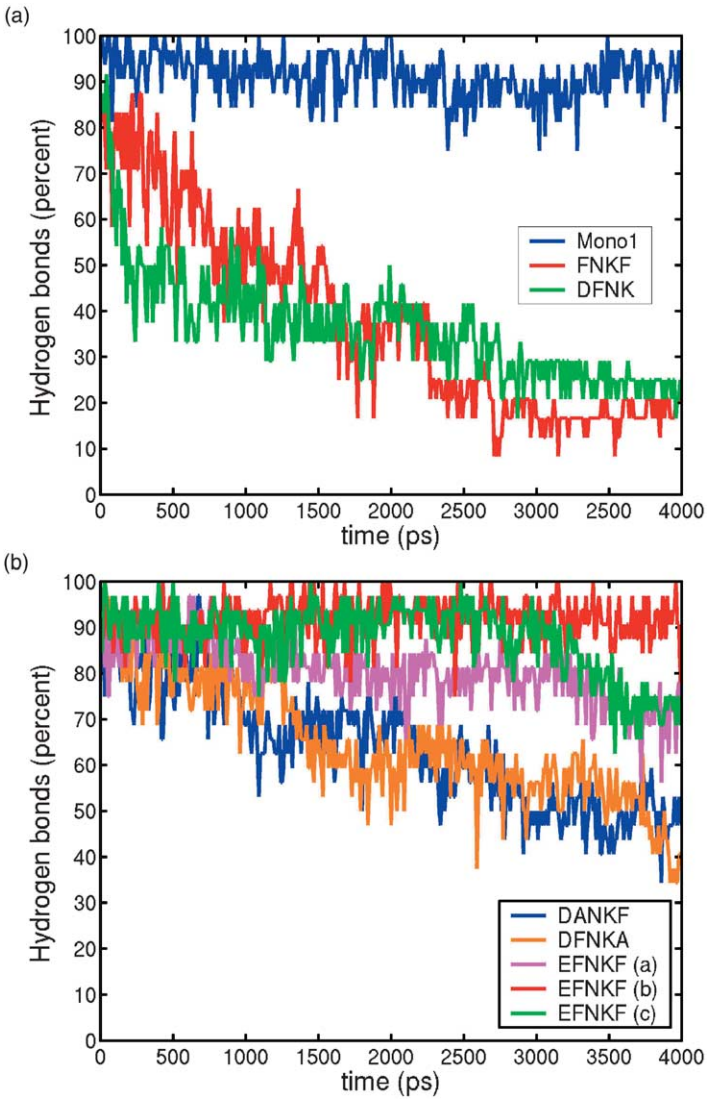

Figure 2. The fraction of main-chain hydrogen bonds during the simulation. The percentage is with respect to the minimized stage. Each model is represented by a different color (see inset legend).

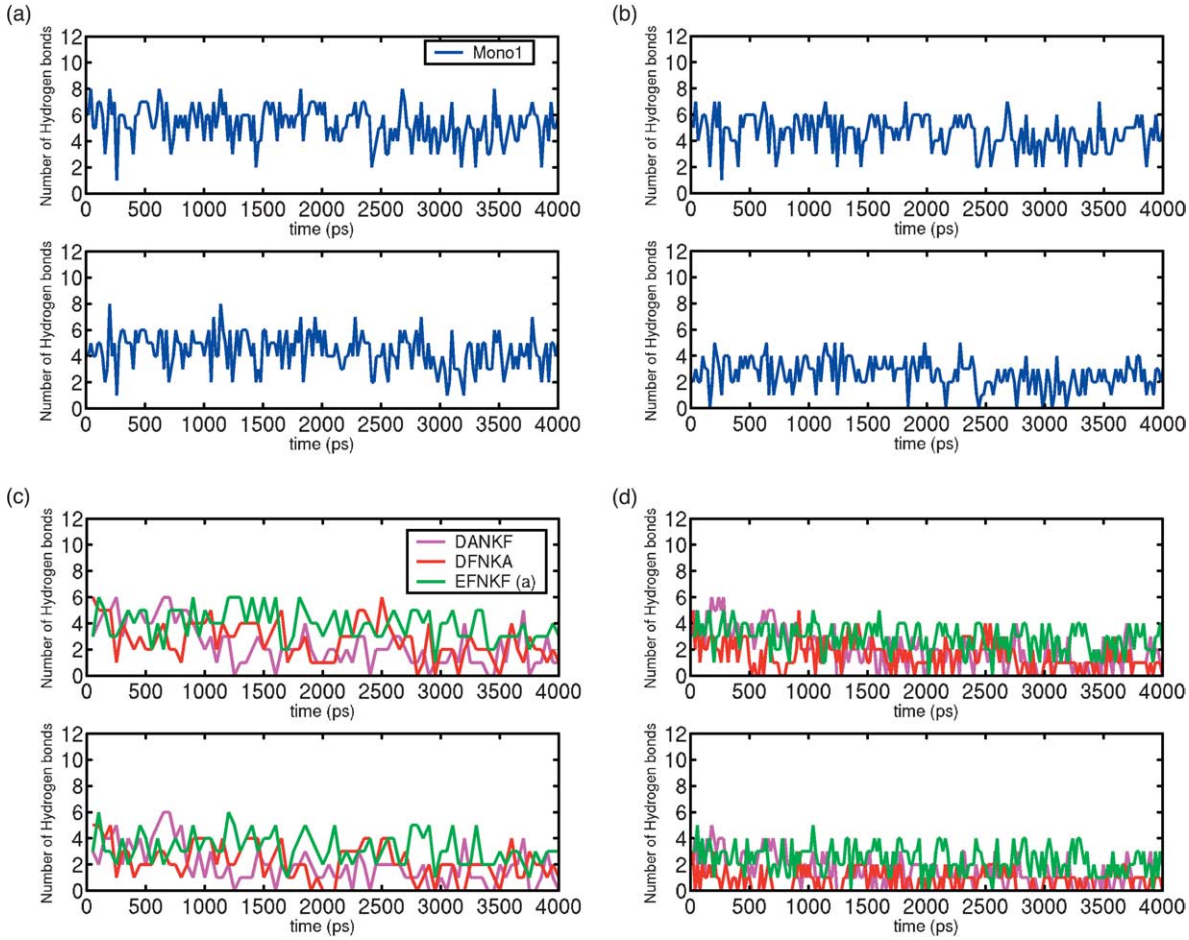

Figure 3. The evolution of the total ((a) and (c)) and preserved ((b) and (d)) number of Asn-Asn side-chain hydrogen bonds during the simulation of the one-sheet models. Each model is plotted in a different color (detailed in the inset legends). At each subfigure two analyses were performed. Top: maximum donor–acceptor distance 2.5 Å, minimum donor–hydrogen–acceptor angle 120°. Bottom: maximum donor–acceptor distance 2.8 Å, minimum donor–hydrogen–acceptor angle 90°.

with water. Hence, hydrogen bonds between Asn side-chains of adjacent strands tend to break and form more freely than backbone hydrogen bonds. However, it is still clear from Figure 3 that at most times, there are more than four side-chains (that is, more than half of the possible side-chain pairs) involved in hydrogen bonds. This is a further indication of the association of the strands in a stable and well organized β-sheet.

The DFNKF structure is stabilized mainly by a network of side-chain–main-chain electrostatic interactions. The number of total and preserved electrostatic interactions during the simulation are plotted in Figure 4(a) and (b) on the top. The bottom parts of Figure 4(a) and (b) show the total and preserved number of electrostatic interactions, with the interactions of each of the two side-chain termini plotted in different colors. In the case of model Mono1, almost all possible 34 electrostatic interactions between the charged side-chains and the respective termini existed during the simulation time. However, it is visible that the Asp–N terminus interaction is better preserved and oscillates less than the Lys–C terminus interaction. The clue to this observation lies in the fact that the Lys side-chain is longer and has more degrees of freedom, especially since it resides in a hydrophobic environment created by the aromatic rings. The Lys side-chain has a long hydrophobic tail that prefers the hydrophobic environment and a hydrophilic end that prefers the charged environment induced by the C terminus. In addition, the C terminus is located in the fifth position while the Lys occupies the fourth position of the peptide. Thus, the distance between the Lys side-chain and the C terminus is greater than between the Asp and the N terminus, which are located at the same position. Therefore, salt bridges formed between the Lys side-chain and the C terminus are less stable, more inclined to dissociate and re-associate.

Clearly, Phe-mediated interactions may provide another key to the high structural organization of the DFNKF single-sheet layer model. We analyzed the total and preserved (with respect to the minimized structure) number of aromatic interactions between the strands (Figure 5). Model

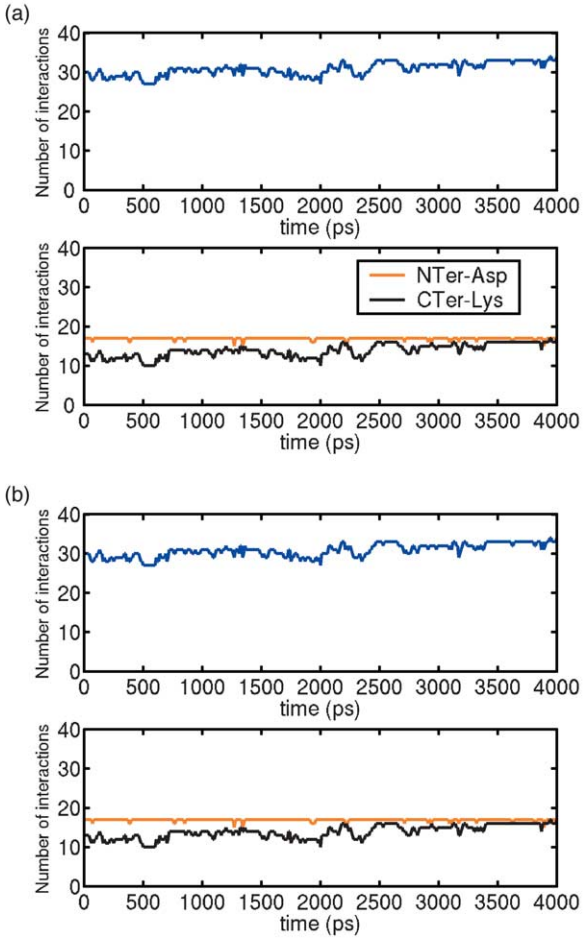

Figure 4. The evolution of the total (a) and preserved (b) number of electrostatic interactions during the simulation of the one-sheet model. Each section is divided into two subplots. Top: the total number of interactions. Bottom: the number of preserved interactions. Each type of interaction is plotted in a separate color (see the inset legend).

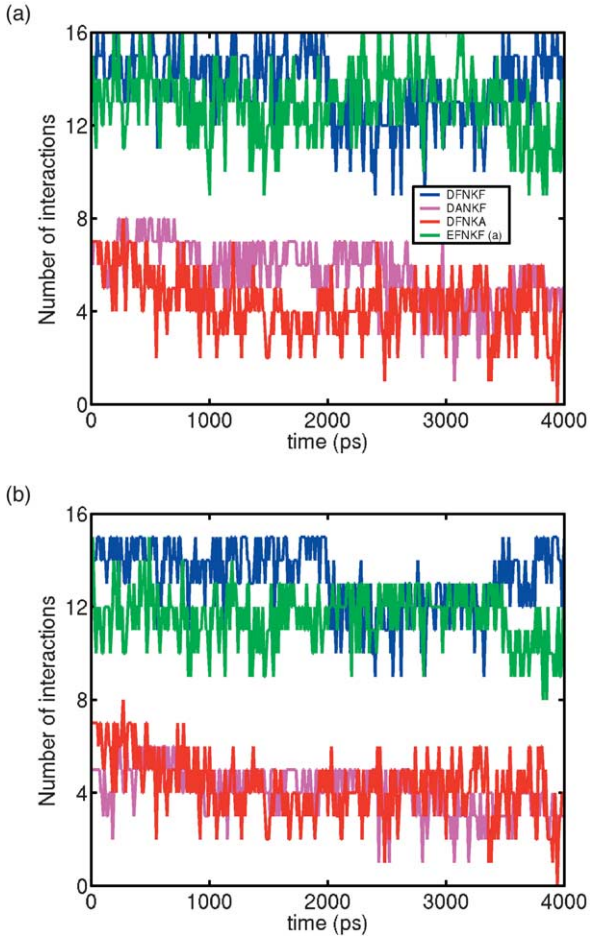

Figure 5. The evolution of the aromatic interactions during the simulation of the one-sheet models. Top: number of total aromatic interactions for the one-sheet models. Bottom: number of preserved aromatic interactions for the one-sheet models. Each model is represented by a different color (see the inset legend).

Mono1 retained a fairly constant number of aromatic interactions throughout the simulation, most of which were preserved. However, the number of aromatic interactions oscillated more than the other structural and chemical characteristics. As seen in Figure 5, there has been a certain drop in the number of aromatic interactions toward the middle of the simulation. Despite that, the overall structural organization of the system remained very stable and highly ordered, suggesting that although Phe-mediated interactions are important, they are not the main stabilizing force of the bulk β-sheet structure. A general overview of the interactions can be obtained by inspection of the molecular models. Figure 6 depicts the front and side views of the conformational evolution of Mono1 during the simulation process. Clearly, the model maintained a highly ordered twisted β-sheet organization. Looking at the snapshots, we see that the twist between the strands enables the formation of hydrophobic interactions between the side-chains. These hydrophobic interactions enable the Phe rings to associate and stack. Without this twist, the Phe aromatic rings are more exposed to the water. Additionally, a critical driving force to the twisting of the strands is the optimization of the salt bridges between the charged side-chains and the termini.

Mutant study: simulation and prediction of specific changes in amino acids

In an attempt to validate our results, we simulated the two truncated tetra-peptide filaments, FNKF and DFNK. Experimentally, the FNKF does not create fibrils while DFNK does, albeit short and less ordered than DFNKF.[24] Both our models consisted of a single β-sheet with nine parallel strands, similar to the original DFNKF system. Our assumption was that the fibrils created by DFNK are similar to those of DFNKF, that is, a

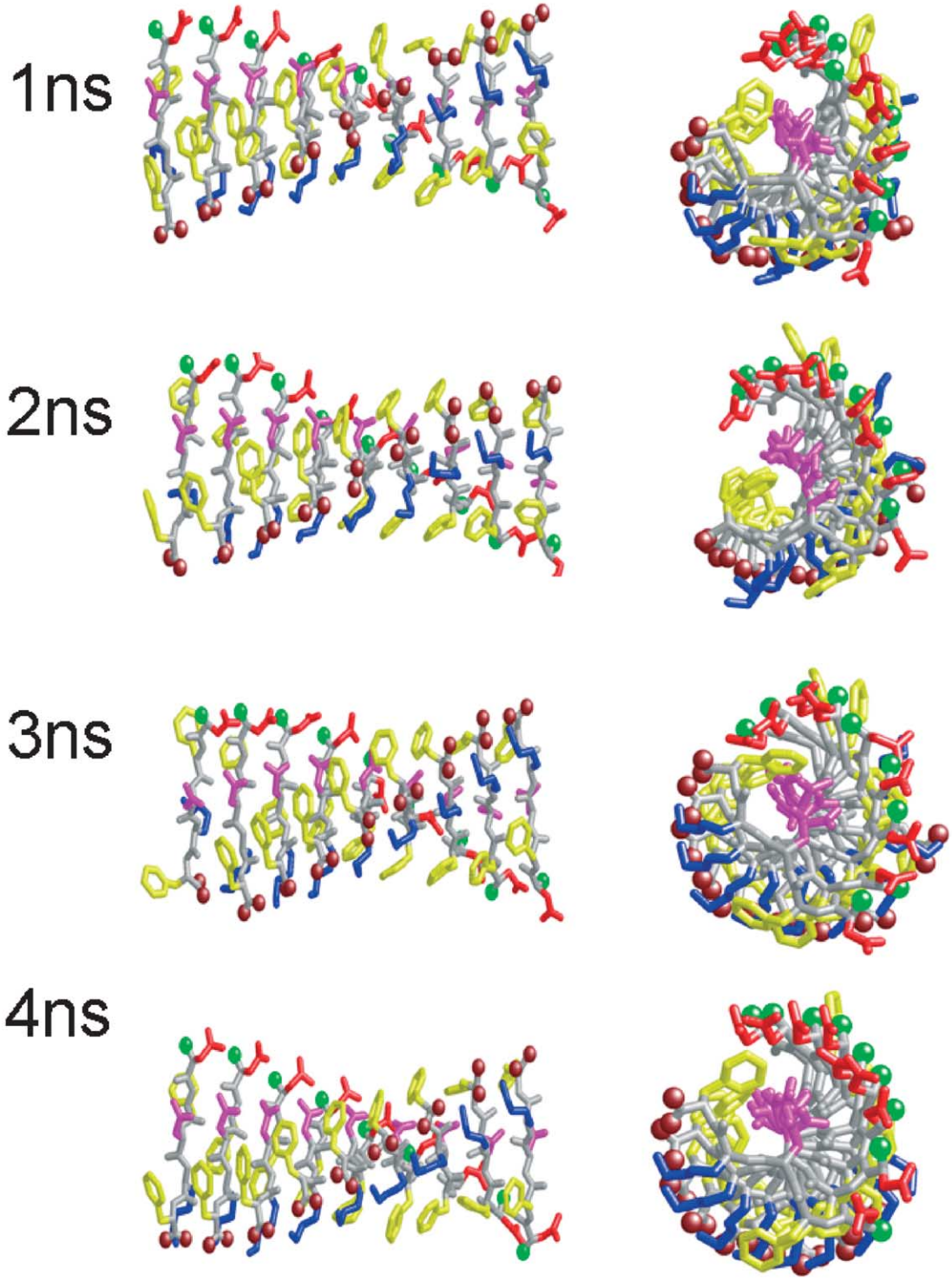

Figure 6. The conformational evolution of model Mono1 after one, two, three and four nanosecond (top to bottom). Right, front view. Left, side view. For the color coding, see Figure 1.

single-sheet layer with parallel strands. On the other hand, we note that this assumption for the DFNK model may possibly bias our results for this truncated mutant.

The minimized FNKF and DFNK systems can be seen in Figure 1(b) and (c), respectively. Consistent with experiment, the FNKF system was less stable than the DFNK and the DFNKF penta-peptide. FNKF lost the β-sheet organization at a very early stage of the simulation with the strands becoming an amorphous aggregate. However, DFNK also lost many of its original interactions and became less organized. Despite that, many of the original contacts were maintained and as a result the bulk β-structure was much better preserved than that of the FNKF, especially in the middle strands. Our

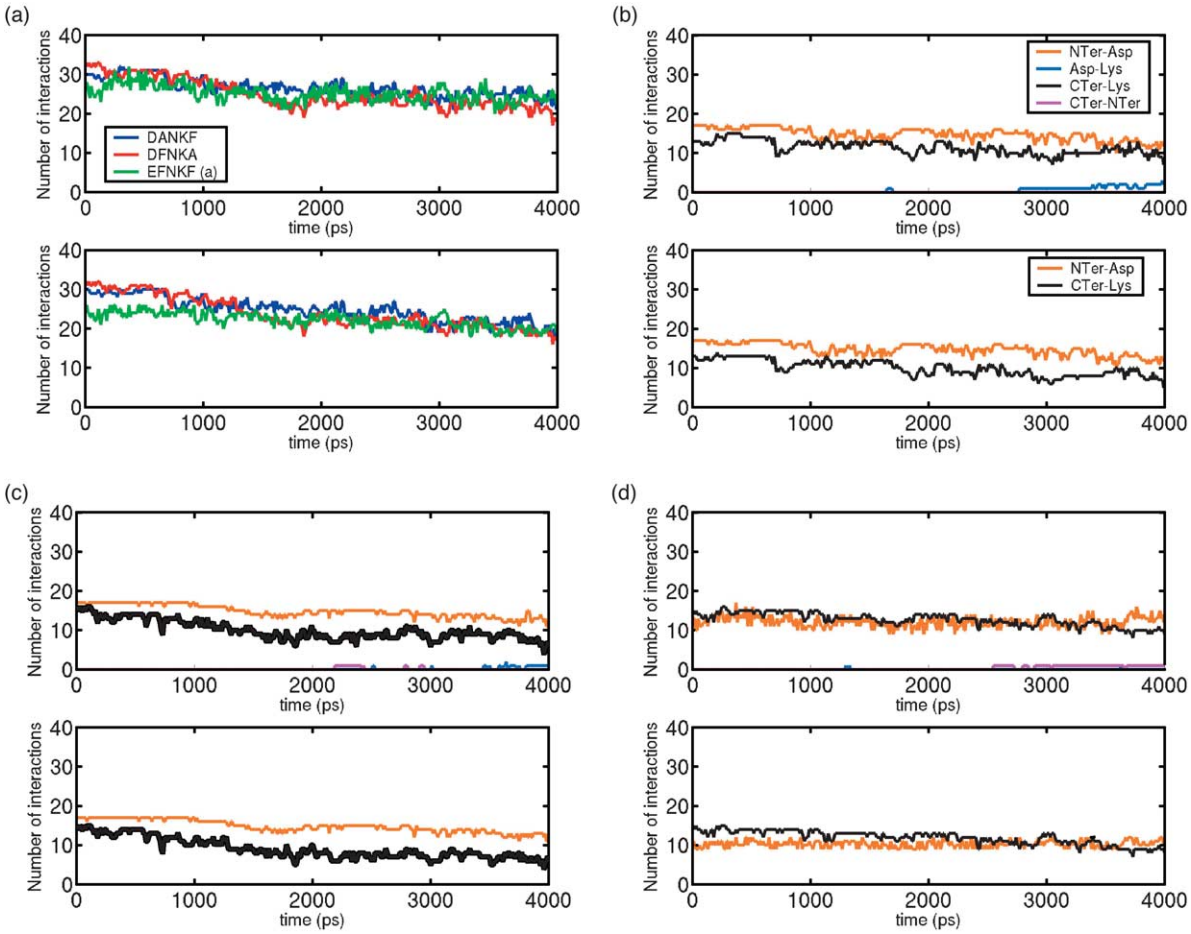

Figure 7. The evolution of the number of electrostatic interactions during the simulation of the one sheet mutant models. Each section is divided into two subplots. Top: The total number of interactions. Bottom: The number of preserved interactions. (a) The total number of interactions. Each model is plotted in a different color (see inset legend). (b)–(d) The interactions are separated according to the interacting charged groups. Each type of interaction is plotted in a separate color (see inset legends): (b) DANKF; (c) DFNKA; (d) EFNKF (rotamer a).

results with the tetra-peptides are able to partially explain the experimental observations. They rationalize the capability of DFNK to possibly create short and non-canonical fibrils as opposed to the FNKFs inability to form fibrils. Furthermore, they emphasize the role played by specific amino acid interactions. The removal of each of the terminal amino acids disrupts the ability of the sequence to form ordered fibrils and to maintain the bulk organization. However, the removal of Asp from the sequence has a larger destabilizing effect on the organization of the β-sheet structure than the removal of the Phe5 from the sequence, probably due to the greater importance of the electrostatic interactions of the Asp residue with the N terminus as compared to the Phe-mediated interactions. The DFNK system did not create highly ordered fibers either in the experiment or in the simulations. However, it is also conceivable that its topology may be different from that of the DFNKF used in the simulations.

The tetra-peptide study further highlighted the role played by each amino acid in the sequence. Replacing one amino acid by another may affect the bulk organization of the structure by changing the chemical interactions taking place in the system. In this context, it is interesting to study the effects of mutating specific amino acids on the ability of the peptide to form fibrils. Thus, we have further simulated five penta-peptidic mutated filaments, DANKF, DFNKA and three versions of EFNKF (designated EFNKF a,b,c) differing only in the initial Glu side-chain conformation. Each mutant is simulated with nine-stranded single-layer β-sheet. The rationale in simulating three different EFNKF models was to test the effect of the starting point on the stability of the model. Figure 1 shows the atomistic representation of the minimized mutants. Figure 1(d) shows the DANKF system, Figure 1(e) the DFNKA system and Figure 1(f) shows the EFNKF (c).

As expected, both DANKF and DFNKA were more structurally stable than the DFNK and FNKF tetra-peptides simulated above. Both were able to maintain a larger β-sheet core and a more organized structure. In addition, they preserved a larger percentage of their original side-chain and main-chain interactions. However, both were

considerably less stable than the original wild-type DFNKF. The three EFNKF mutants were more stable than the Phe-Ala mutants. In all three EFNKF systems, the β-sheet structure maintained a considerable amount of structural organization. However, they lost more of their bulk organization and initial interactions than DFNKF. The main-chain hydrogen bond percentages are depicted in Figure 3(b). Clearly, the differences between the mutants are not as great as the differences between the DFNKF model and the tetra-peptides. In all the mutants, the main-chain hydrogen bond loss was not as dramatic either. However, it is visible that the EFNKF mutants lost less main-chain hydrogen bonds than the Phe-Ala mutants.

The side-chain parameters draw a similar picture with regard to Asn-Asn side-chain hydrogen bonds (see Figure 3(c) and (d)), electrostatic interactions (Figure 4) and aromatic interactions (Figure 5). However, in the case of EFNKF, the Glu–N-terminal electrostatic interaction tends to be less preserved than the Asp–N-terminal electrostatic interaction in the DFNKF and in the Phe-Ala mutants (Figure 7(b)–(d)). Figure 8 shows the conformational evolution of the two Phe-Ala mutants and a representative of the Asp-Glu systems during the simulations. Both the Phe-Ala mutants were more stable and organized than the tetra-peptides but less than the original DFNKF (see Figure 6).

The simulations of the Phe-Ala mutants lead us to conclude that both phenylalanine residues play significant roles in maintaining the protofibril structure. However, we found no significant difference in the ability to preserve the bulk organization between the two Phe-Ala models. An additional study in our group[27] has recently observed that DFNKA is more likely to create ordered fibrils than DANKF. This prediction was subsequently confirmed experimentally (E. Gazit, personal communication). Further, looking at Figure 8(c), we may obtain a clue as to the specific contribution of Asp to the increased stability of DFNKF. The EFNKF systems are less stable than DFNKF since the Glu side-chain is longer and therefore cannot interact simultaneously with the N termini of its own strand and of the adjacent strand, whereas the shorter Asp can. However, this conformational effect is weaker than the hydrophobic effect that contributed to the relative disorganization of the Phe-Ala mutants. Thus, all Asp-Glu systems are much more organized.

Figure 8 provides a general overview of the simulation of the mutants. As in the DFNKF case, all the mutants formed an initial twist among the strands at a relatively early stage of the simulation (less than 100 picoseconds). Table 1 provides the twist angles between two adjacent strands for the mutant models at different times of the simulation. The average twist angle was around 20° for the stable models. In the Phe-Ala mutants, the angle tended to be larger. In these models the absence of the Phe rings gave the main-chain more freedom to twist. In the DFNK, in the Phe-Ala mutants and in one of the EFNKF mutants we excluded some of the strands from the calculation, since they drifted away from the bulk structure and therefore their twist was meaningless. The larger twist angle can provide another explanation to the decreased stability of the Phe-Ala mutants. A large twist angle between two adjacent strands may lead them to tear away from one another at a relatively early stage of the simulation by disrupting the main-chain and side-chain interactions necessary for maintaining the bulk structure. This structural disruption leads to the exposition of hydrophobic segments towards the polar solvent and the consequent destabilization of the bulk organization.

Conclusions

Koide *et al.*[25] have experimentally designed a contiguous β-sheet flanked by two globular domains. They have shown that a single sheet largely devoid of a hydrophobic core can be stable. Here, we present a highly stable single-layer parallel-stranded β-sheet model with a very small hydrophobic core for the DFNKF protofibril. The DFNKF single-sheet organization is maintained by a large number of main-chain and side-chain chemical interactions. These interactions explain why this single layer β-sheet structure can exist despite the largely absent hydrophobic core. In our model every amino acid has a particular role and each is involved in a chain of interactions, and sometimes in more than one network of interactions. Replacement or removal of each amino acid disrupts the structure and sometimes renders the creation of fibrils impossible.

The presence of only a small hydrophobic core is compensated by networks of polar interactions. The most important are electrostatic interactions between the Asp and the N terminus and between the Lys and the C terminus. These interactions provide two pivots that hold the bulk structure together at the termini and prevent the charge repulsion that could be caused between the termini of the parallel strands. The importance of the Asp–N terminus electrostatic interactions is emphasized by the fact that the DFNK tetra-peptide is capable of creating fibrils, but when the Asp is removed, as in the case of the FNKF tetra-peptide, no fibrils are detected. In the DFNKF case the Asp–N terminus electrostatic interactions are optimized by the short side-chain of Asp, adopting a conformation that allows it to interact both with the N terminus of its own strand and with that of the adjacent strand. This kind of bifurcated interaction is more difficult to obtain when the Asp is replaced by a Glu, which has a longer side-chain. This can be seen in the case of the Asp-Glu mutants, where the Glu side-chain could not adapt itself to form the bifurcated electrostatic interaction and as a result these mutants were somewhat less stable than the original DFNKF. Most of the trajectories we simulated lost a strand during the simulation, and

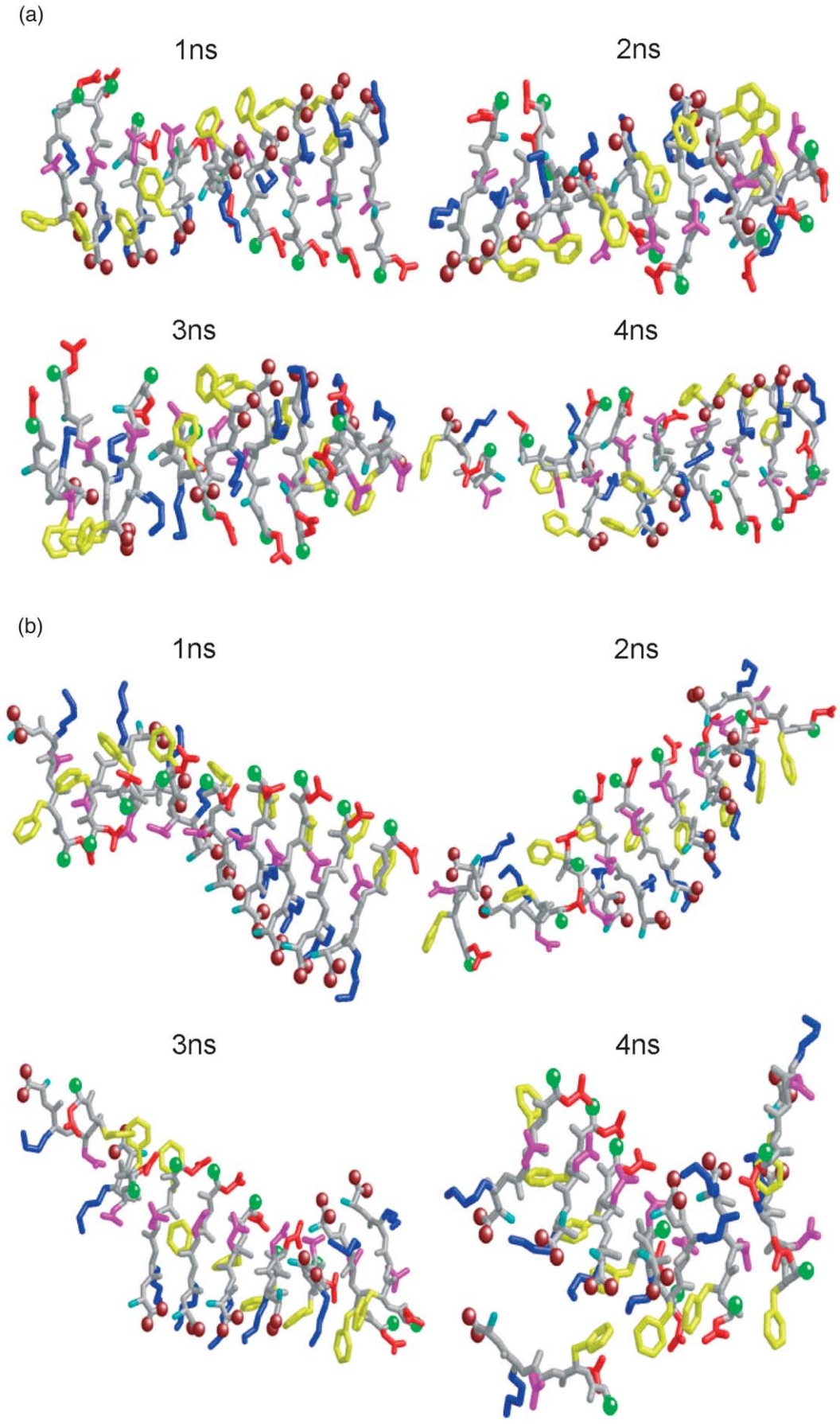

Figure 8 (*legend next page*)

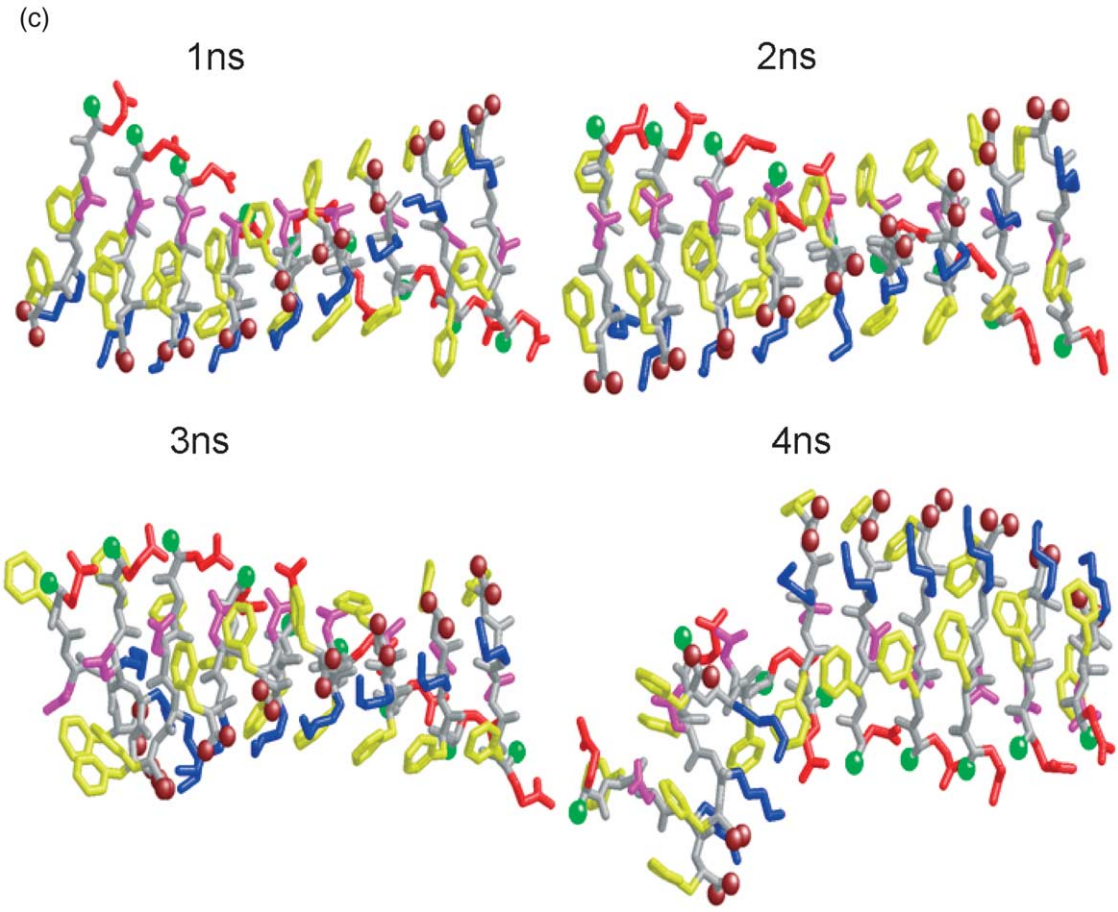

Figure 8. (a) The conformational evolution of model DANKF after one, two, three and four nanoseconds (top to bottom, left to right). (b) The conformational evolution of model DFNKA after one, two, three and four nanoseconds (top to bottom, left to right). (c) The conformational evolution of model EFNKF (c) after one, two, three and four nanoseconds (top to bottom, left to right). See Figure 1 for the colour coding.

the overall structure tended to be less organized. In addition, the Glu–N terminus interaction was less preserved than the Asp–N terminus interaction in the case of the DFNKF. The arrangement of the edge groups in electrostatic interactions provides thermal shielding (i.e. fixes the position of the strands and avoids the effect of thermal agitation), allowing the backbone amide groups to form optimal hydrogen bonds.

A second leading interaction is a network of side-chain hydrogen bonds between Asn side-chains of adjacent strands. These interactions maintain the β-sheet organization by keeping the strands close to one another in their middle, while the electrostatic interactions mentioned above, taking place near the termini, assist in maintaining the bulk β-structure by associating the two termini of the strands. These particular hydrogen bond and electrostatic networks provide a good geometrical organization that maintains the unique twisted β-sheet organization.

Our system is further characterized by Phe-mediated interactions: the stacking of the Phe aromatic rings of adjacent strands and hydrophobic interactions between the Lys4 aliphatic tail and Phe2, and between the Asn methyl group and Phe5. These last hydrophobic interactions help in shielding the hydrophobic groups from the solvent and enable the Phe residues to stack properly. Previous studies[15,19] have shown the role played by Phe when it is present in amyloidogenic sequences. In the islet amyloid polypeptide segment NFGAIL[15] the phenylalanine-mediated interactions were found to be highly important for the final fibrillar organization. Moreover, experimental data[24] showed that when the two phenylalanine residues are replaced by alanine in our segment, no fibrils are detected, indicating the cumulative importance of the Phe residues to the organization and stability of the structure. Yet, these aromatic interactions are not sufficient by themselves, as seen in the case of the FNKF, which forms no fibrils despite the presence of the two sets of aromatic rings. Despite that, the aromatic interactions are important to maintain the bulk structure, since in all the mutants where one of the Phe residues is removed (in the case of DFNK) or replaced by alanine (as in the case of DANKF and DFNKA) fibrils may be formed, albeit they are less stable.

To conclude, each amino acid in the sequence has a role, making the fibrillar structure that consists of

a single β-sheet that is largely devoid of a hydrophobic core highly stable. As shown here, the replacement of each amino acid disrupts the organization, providing clues regarding the amyloid formation process. These may facilitate the search for drugs and prevention of amyloid-related diseases.

Methods

Computational details

The calculations were performed using the CHARMM package.[28] All the atoms of the system were considered explicitly, and the energy was calculated using the CHARMM22 force field.[29] The water molecules were represented explicitly, using the TIP3 model.[30] The simulations were performed using the NVT ensemble in orthorhombic simulation box. We chose constant volume simulations because all the trajectories were obtained at high temperature (only 23 degrees below water boiling point). By these means we could assure not losing a proper density distribution due to thermal effects. Periodic boundary conditions were applied using the nearest image convention. The box size was adjusted to fit the complex size, so that infinite dilution conditions would be maintained. For the octameric models a cubic box of 40 Å was used, for the three-sheets systems the edge size was set at 45 Å and for the one-sheet systems the box dimensions were 84 Å×45 Å×45 Å. The starting molecular structures were built using the INSIGHTII molecular package (2000, Accelrys, San Diego, CA). For any given arrangement the inter-strands distance within sheets was fixed around 5 Å. The intra-sheet twist was initially avoided, since it is highly dependent on the chemical composition and arrangement (i.e. peptide sequence). Nevertheless, our previous works have indicated that if the system is correctly relaxed, the correct twist is achieved after enough steps of energy minimization. The charge of all potential titratable groups was fixed to those values corresponding to neutral pH, i.e. the experimental conditions when DFNKF spontaneously forms fibers. Thus all aspartic acid side-chains were represented in their anionic form and all lysine side-chains in their acidic positively charged form. Both peptide edges were ionized as well, as at their physiological pH.

Before running each molecular dynamics simulation, the potential energy of each system was minimized using 2000 steepest descent steps and 600 ABNR steps. The heating protocol included 20 picoseconds of increasing temperature from 0 to the final temperature of 350 K plus 20 picoseconds of equilibration period. We perform the simulations at 350 K in order to enhance the stability differences between the models by means of thermal stress. Furthermore, using high temperature allowed us to infer some kinetic tendencies. Residue-based cutoff was applied at 12 Å, i.e. if any two molecules have any atoms within 12 Å, the interaction between them is evaluated. The SHAKE algorithm was applied to fix the bond length.[31] A numerical integration time step of two fermisecond was used for all the simulations. The non-bonded pair list was updated every 20 steps and the trajectories were saved every 500 steps (one picosecond) for subsequent analysis. Each simulation was run for a period of four nanoseconds.

Interaction analysis details

We calculated the percentage of main-chain hydrogen bonds with respect to the initial (minimized) structure. A hydrogen bond is defined by the distance between the donor and acceptor (d N–H···O) equal to or less than 2.5 Å and the angle (∟N–H···O) is more than 120°. We also calculated the hydrogen bonds between the Asp side-chains, using the same distance and angle criteria.

Another set of analyses relates to the aromatic interactions. We calculated the total number of aromatic interactions and the number of preserved intra-sheet aromatic interactions (with respect to the initial minimized structure). Two aromatic rings were considered interacting if the center of mass distance between them was less than 6 Å. We further analyzed the electrostatic interactions between two charged groups. Two charged groups were considered to interact if the distance between their interaction centers, i.e. the carbon atom of the C terminus or the Asp and the charged N atom, was 4.5 Å or less.[32,33]

Acknowledgements

We are grateful to Dr Ehud Gazit for bringing to our attention this extremely interesting peptide system and for communicating results prior to publication. We thank Drs C.-J. Tsai, K. Gunasekaran and G. Tsai for discussions. In particular, we thank Dr Jacob V. Maizel for encouragement. The computation times are provided by the National Cancer Institute's Frederick Advanced Biomedical Supercomputing Center and by the NIH Biowulf. The research of R.N. in Israel has been supported in part by the "Center of Excellence in Geometric Computing and its Applications" funded by the Israel Science Foundation (administered by the Israel Academy of Sciences), and by the Adams Brain Center. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-12400. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the US Government.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2004.11.002

References

1. Wanker, E. E. (2000). Protein aggregation in Huntington's and Parkinson's disease: implications for therapy. *Mol. Med. Today*, **6**, 387–391.

2. Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332.
3. Dobson, C. M. (2003). Protein folding and disease: a view from the first Horizon symposium. *Nature Rev. Drug Discov.* **2**, 154–160.
4. Harper, J. D. & Lansbury, P. T., Jr (1997). Models of amyloid seeding in Alzheimer's disease and Scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* **66**, 385–407.
5. Kisilevsky, R. (2000). Amyloidogenesis: unquestioned answers and unanswered questions. *J. Struct. Biol.* **130**, 99–108.
6. Gazit, E. (2002). Possible role of p-stacking in self-assembly of amyloid fibrils. *FASEB J.* **16**, 77–83.
7. Fandrich, M. & Dobson, C. M. (2002). The behavior of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* **21**, 5682–5690.
8. Sunde, M. & Blake, C. F. F. (1998). From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quart. Rev. Biophys.* **31**, 1–39.
9. Serrano, L. (2000). The relationship between sequence and structure in elementary folding units. *Advan. Protein Chem.* **53**, 49–85.
10. Tycko, R. (2003). Insights into amyloid folding problem from solid-state NMR. *Biochemistry*, **42**, 3151–3159.
11. Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M. & Griffin, R. G. (2004). High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl Acad. Sci. USA*, **101**, 711–716.
12. Klimov, D. & Thirumalai, D. (2003). Dissecting the assembly of $A\beta_{16-22}$ amyloid peptides into anti-parallel β sheets. *Structure*, **11**, 295–307.
13. Gnanakaran, S., Nymeyer, H., Portman, J., Sanbonmatsu, K. Y. & Garcia, A. E. (2003). Peptide folding simulations. *Curr. Opin. Struct. Biol.* **13**, 168–174.
14. Gsopner, 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 Sup 35. *Proc. Natl Acad. Sci. USA*, **100**, 5154–5159.
15. Zanuy, D., Ma, B. & Nussinov, R. (2003). Short peptide amyloid organization: stabilities and conformations of the islet amyloid peptide NFGAIL. *Biophys. J.* **84**, 1–11.
16. Ma, B. & Nussinov, R. (2002). Molecular dynamics simulations of alanine rich β-sheet oligomers: insight into amyloid formation. *Protein Sci.* **11**, 2335–2350.
17. Ma, B. & Nussinov, R. (2002). Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers ($A\beta_{16-22}$, $A\beta_{16-35}$, $A\beta_{10-35}$): sequence effects. *Proc. Natl Acad. Sci. USA*, **99**, 14126–14131.
18. Zanuy, D. & Nussinov, R. (2003). The sequence dependence of fiber organization: a comparative molecular dynamics study of the Islet amyloid polypeptide segments 22–27 and 22–29. *J. Mol. Biol.* **329**, 565–584.
19. Sikorski, P., Atkins, E. D. T. & Serpell, L. C. (2003). Structure and texture of fibrous crystals formed by alzheimer's Aβ(11–25) peptide fragment. *Structure*, **11**, 915–926.
20. Zaidi, M., Inzerillo, A. M., Moonga, B. S., Bevis, P. J. & Huang, C. L. (2002). Forty years of calcitonin—where are we now? A tribute to the work of Iain Macintyre, FRS. *Bone*, **30**, 655–663.
21. Butler, M. & Khan, S. (1986). Immunoreactive calcitonin in amyloid fibrils of medullary carcinoma of the thyroid gland. An immunogold staining technique. *Arch. Pathol. Lab. Med.* **110**, 647–649.
22. Berger, G., Berger, N., Guillaud, M. H., Trouillas, J. & Vauzelle, J. L. (1988). Calcitonin-like immunoreactivity of amyloid fibrils in medullary thyroid carcinomas. An immunoelectron microscope study. *Arch. A Pathol. Anat. Histopathol.* **412**, 543–551.
23. Arvinte, T., Cudd, A. & Drake, A. F. (1993). The structure and mechanism of formation of human calcitonin fibrils. *J. Biol. Chem.* **268**, 6415–6422.
24. Reches, M., Porat, Y. & Gazit, E. (2002). Amyloid fibril formation by pentapeptide and tetrapeptide segmentd of human calcitonin. *J. Biol. Chem*, **277**, 35475–35480.
25. Koide, S., Huang, X., Link, K., Koide, A., Bu, Z. & Engelman, D. M. (2000). Design of single-layer β-sheets without a hydrophobic core. *Nature*, **403**, 456–460.
26. Zanuy, D., Porat, Y., Gazit, E. & Nussinov, R. (2004). Peptide sequence and amyloid formation: molecular simulations and experimental study of a human Islet amyloid polypeptide fragment and its analogs. *Structure*, **12**, 439–455.
27. Tsai, H. H., Zanuy, D., Haspel, N., Gunasekaran, K., Ma, B., Tsai, C. J. & Nussinov, R. (2004). The stability and dynamic of the human Calcitonin amyloid peptide DFNKF. *Biophys. J.* **27**, 146–158.
28. Brooks, R. B., Bruccoleri, R. E., Olafson, B. D., Sate, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy minimization and dynamic calculations. *J. Comput. Chem.* **4**, 187–217.
29. MacKerell, J. A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Jr, Evanseck, J., Field, M. J. *et al*. (1998). All-hydrogen empirical potential for molecular modeling and dynamic studies of proteins using the CHARMM22 force field. *J. Phys. Chem. ser. B*, **102**, 3586–3616.
30. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1982). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
31. Ryckaert, J. P., Ciccoti, G. & Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341.
32. Zanuy, D., Aleman, C. & Muñoz-Guerra, S. (2002). A molecular dynamics study in chloroform solution of the stoichiometric complex formed by poly(alpha,L-glutamate) and octylmethylammonium ions. *Biopolymers*, **63**, 151–162.
33. Aleman, C. & Zanuy, D. (2001). Binding in complex ionic systems: anticooperative effects in systems stabilized by electrostatic interactions. *Chem. Phys. Letters*, **343**, 390–396.

Edited by P. T. Lansbury Jr

(Received 10 December 2003; received in revised form 19 May 2004; accepted 1 November 2004)