

Aggregation of Small Peptides Studied by Molecular Dynamics Simulations

Dagmar Flöck, 1* Giulia Rossetti, 1 Isabella Daidone, 2 Andrea Amadei, 3 and Alfredo Di Nola 1

¹Department of Chemistry, University of Rome "La Sapienza", Rome 00185, Italy

Peptides and proteins tend to ag-**ABSTRACT** gregate under appropriate conditions. The amyloid fibrils that are ubiquitously found among these structures are associated with major human diseases like Alzheimer's disease, type II diabetes, and various prion diseases. Lately, it has been observed that even very short peptides like tetra and pentapeptides can form ordered amyloid structures. Here, we present aggregation studies of three such small polypeptide systems, namely, the two amyloidogenic peptides DFNKF and FF, and a control (nonamyloidogenic) one, the AGAIL. The respective aggregation process is studied by all-atom Molecular Dynamics simulations, which allow to shed light on the fine details of the association and aggregation process. Our analysis suggests that naturally aggregating systems exhibit significantly diverse overall cluster shape properties and specific intermolecular interactions. Additional analysis was also performed on the previously studied NFGAIL system. Proteins 2006;65:914-921. © 2006 Wiley-Liss, Inc.

Key words: conformational changes; peptide aggregation; human calcitonin; human islet amyloid polypeptide

INTRODUCTION

Proteins and polypeptide chains tend to aggregate under appropriate conditions. 1,2 While these aggregates initially have mostly been associated with diseases, $^{3-5}$ including Alzheimer's, Huntington's, Parkinson's, type II diabetes, or the transmissible prion disorders, it has now become evident that almost any protein and peptide is subject to aggregation given the right environmental conditions such as pH, temperature, and ionic strength. 1,2 Many of these form amyloid-like fibrils which show a distinct X-ray fiber diffraction pattern corresponding to a well ordered β -sheet secondary structure. Since morphologically similar aggregates evolve from proteins unrelated in sequence and/or structure it is widely believed that general principles underlie the oligomerization and fibrillation.

The role of small soluble oligomers in disease and as precursors of amyloid formation is widely under discussion and investigation. For example, it could be shown recently that specific Amyloid β protein assemblies in the

brain impair memory.⁶ The knowledge of the structural properties of these oligomers would present the possibility to target these assemblies and abort the disease (in this case Alzheimer's) before permanent structural changes have developed. Therefore, the understanding of the very early stages of amyloid formation which we aim to elucidate here with computer simulations becomes highly desirable.

An exciting new entry to this field was the discovery that very short aromatic peptides can form well-ordered amyloid-like fibrils in the shape of nanotubes.^{7,8} These self-assembled tubular nanostructures can then be utilized to produce discrete nanowires with a long persistence length and a wide range of applications in nanotechnology. From a computational point of view, short peptides are very interesting because all-atom molecular dynamics (MD) simulations with explicit water of these many particle systems are feasible. Furthermore, since the systems are relatively simple we can hope to get a deeper insight in aggregation kinetics and dynamics. While single particle simulations provide only insights regarding the monomer fluctuations in the given external conditions, simulations of such multiparticle systems can monitor the profound conformational changes which are driven by interpeptide interactions.

Previously we have reported a MD study⁹ of the initial self-assembly stages of the $\mathrm{NH_2-NFGAIL-COOH}$ peptide, the core recognition motif of the islet amyloid polypeptide (IAPP) associated with type II diabetes. Structural analysis of the cluster showed the formation of a flat ellipsoid-shaped cluster with the single peptides in preferred parallel alignment facilitated by intermolecular aromatic interactions. Here, we present MD studies on three small peptides: (a) DFNKF, the central motif of the human calcitonin; (b) FF, a motif of the amyloid β protein; and (c) the nonamyloidogenic AGAIL, a single point

²Computational Molecular Biophysics Group, IWR, University of Heidelberg, 69120 Heidelberg, Germany

³Department of Chemical Science, University of Rome "Tor Vergata", Rome 00133, Italy

Grant sponsor: European Community Training and Mobility Research Network Project "Protein (mis) folding"; Grant number: HPRN-CT-2002-00241.

^{*}Correspondence to: Dagmar Flöck, Department of Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, Rome 00185, Italy. E-mail: floeck@caspur.it

Received 5 April 2006; Revised 21 June 2006; Accepted 28 June 2006

Published online 15 September 2006 in Wiley InterScience (www. interscience.wiley.com). DOI: 10.1002/prot.21168

mutation of the central motif (NFGAIL) of the human IAPP, as a control. $\label{eq:notation}$

The three peptide systems that were chosen for this study have been thoroughly studied experimentally. The peptide DFNKF forms well ordered fibrils similar to the aggregates of the human calcitonin. In a recent solidstate NMR study, Naito et al. 10 demonstrated that this central core region of calcitonin (DFNKF) arranges in an antiparallel β-sheet. They propose that this arrangement is driven by electrostatic interaction (and therefore pH dependent) and then stabilized by parallel π - π interactions between the phenylalanine rings. The ability of this peptide to build aggregates was previously demonstrated by Gazit and coworkers. With the help of single point mutations they already stressed the importance of aromatic interactions for the aggregation of this peptide. 11 Nussinov and coworkers have computationally explored the stability of prearrangement of small oligomers of DFNKF (with termini NH2 and COOH-comparable to the experimental studies of Gazit et al. which were performed in organic solvent). They found that their prearranged parallel β -sheet becomes more stable as the number of added strands increases. 12 In a second computational study they also proposed a possible atomic model for a DFNKF protofilament with a specific peptide strand $association. \\^{13}$

It has been experimentally shown that NFGAIL aggregates and builds structures very similar to those which are built by the entire peptides. 14 They are also just as toxic for the B cells of the pancreas. Recently, Azriel and Gazit¹⁵ have performed an alanine scan of the peptide to analyze the influence of single residues on the fibril morphology and the kinetics of polymerization. They found that the phenylalanine plays a critical role for aggregation, while the polar asparagine seems to influence mainly the kinetics of polymerization. In fact, the substitution of the phenylalanine with an alanine (NAGAIL) has been shown to make the aggregation of the peptide into amyloid fibrils impossible in vitro. In a combined experimental and computational investigation, 16 Zanuy et al. demonstrated that the phenylalanine side chains stabilize the macromolecular structure through their chemical character and their restricted flexibility when interacting with aliphatic residues. This finding is in good agreement with the results of MD simulation studies^{17,18} where the early aggregation steps of NFGAIL and NAGAIL were investigated.

The Alzheimer's β amyloid diphenylalanine structural motif has been shown to assemble rapidly into ordered semicrystalline structures when diluted into aqueous solution.⁸ The tubular nanostructures formed by the dipeptides exhibit a green-gold bifringence when stained with Congo red dye, which is consistent with an organization that may be similar to that of amyloid structures. A computational investigation of the FF motif aggregation has (to our knowledge) not yet been performed.

Here, in this MD study, we investigate the possible initial steps in the aggregation and nucleation of the different peptides. As in our previous study on the aggregation

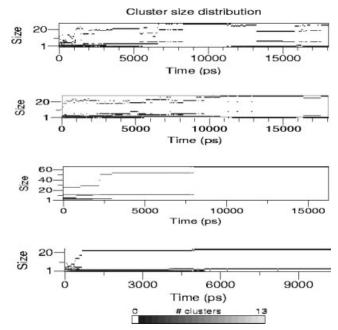


Fig. 1. Cluster analysis along time for (a) DFNKF 340 K, (b) DFNKF 300 K, (c) FF 340 K, and (d) AGAIL 340 K. The vertical axis represents the number of peptides in the cluster.

of NFGAIL, 9 different from most computational studies, our starting point is not a preformed aggregate of peptides but a solution with quasi-random position and orientation of the peptides which mimic local microscopic conditions that can favor the formation of aggregates on allatom, explicit water MD accessible time.

Our analysis shows that the mutated peptide system (AGAIL) clusterizes in a significantly different way with respect to the two naturally aggregating ones. While DFNKF and NFGAIL build extended cluster structures where the single peptides tend to arrange perpendicular to the main cluster axis, AGAIL peptides arrange into a bulky cluster with no preferred orientation of the peptides within the cluster. We propose that these geometric cluster properties are typical for the initial stages of protein aggregation. Given the limited set of systems investigated these results remain suggestive of the early stage of aggregation.

METHODS

MD Simulations

The three peptide systems were simulated in explicit aqueous solution. (a) DFNKF, a pentapeptide fragment of human calcitonin was simulated with charged termini at 300 and 340 K. The two simulations were performed for 18 ns each. The peptides were immersed in a rhombic dodecahedral box (x=7.1 nm, y=6.7 nm, z=5.8 nm) containing 8123 SPC water molecules. Twenty-six replicas were equally distributed within the simulation box. The starting conformation of the single peptides was taken from a previous single peptide simulation. This conforma-

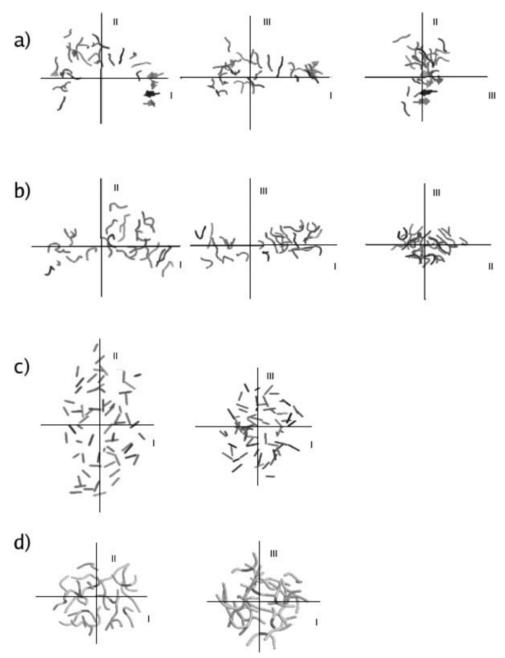


Fig. 2. Graphical representation of the clusters projected into different perpendicular planes for (a) DFNKF 340 K, (b) DFNKF 300 K, (c) FF 340 K, and (d) AGAIL 340 K.

tion corresponds to the cluster of configurations most visited during that simulation. (b) AGAIL is a mutant form of a fragment of the human IAPP. The simulation was performed for 10 ns at 340 and 300 K. The peptide was immersed in a rhombic dodecahedral box (x=9.2 nm, y=8.7 nm, z=7.5 nm) containing 19203 SPC water molecules. Twenty-six replicas were equally distributed within the simulation box with random orientation of the single peptides. (c) The dipeptide FF was immersed in a rhombic dodecahedral box (x=9.4 nm, y=8.9 nm, z=7.7 nm) containing 23955 SPC water. Sixty-four replicas were equally

distributed within the simulation box with random orientation of the single peptides. All atomic structures were modeled with the builder module of the molecular modeling software InsightII. 19 The simulation was performed at $340~\rm K$.

All simulations were performed and analyzed with the GROMACS software package and the GROMOS96 united atom force field^{20,21} G43a1. The starting conformation were modeled with the molecular graphics package InsightII.¹⁹ Neutral pH conditions were realized by setting the protonation states of the ionizable residues according

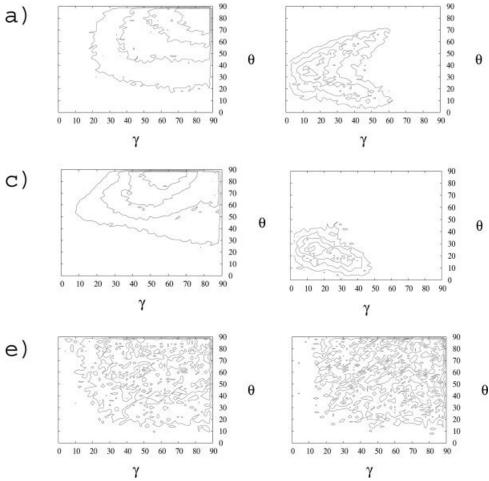


Fig. 3. Analysis of Phe–Phe orientation for DFNKF. γ , the angle between the two ring surface normals and θ , the angle between normal and the vector connecting the two geometrical centers of the aromatic rings. (a) 340 K, within the same peptide. (b) 340 K, between different peptides. (c) 300 K, within the same peptide. (d) 300 K, between different peptides. (e) Analysis of Phe–Phe orientation for FF: within the same peptide. (f)Analysis of Phe–Phe orientation for FF: between different peptide.

to their pK_a 's. Simulations were performed in the NVT ensemble, with temperatures kept close to the desired value by a weak coupling to an external heat bath. ²² The coupling constant was chosen to match the time step. The LINCS algorithm was used to constrain bond lengths. ²³ Long-range electrostatic interactions were calculated using the particle-mesh Ewald (PME) method. ²⁴ For the PME calculation the spacing of the Fourier transformed grid was set to 0.12 nm and the relative strength of the electrostatic interaction at the cutoff to 1×10^{-5} . The cutoff radius for the Lenard-Jones interactions was set to 0.9 nm. After a subsequent energy minimization of solvent and peptides and a 100 ps relaxation of the solvent, the system temperature was gradually increased from 50 K to the desired values over 100 ps.

RESULTS

DFNKF: The two simulations of DFNKF performed at 300 and 340 K and with charged termini were performed

for 18 ns each starting from "random" distribution in the simulation box with various conformations taken from a MD simulation of the single peptide. In Figure 1, the cluster analysis of the trajectories is reported. Peptides are considered to be in a cluster when the minimal atomic distance between two elements of the cluster does not exceed 0.35 nm. The systems tend to aggregate after a very short simulation time. A large cluster forms which tends to break into two or three parts but reforms towards the end of the simulation [see Fig. 1(a,b)]. The cluster at higher temperature takes on a curved extended structure [see Fig. 2(a)], while for the simulation at 300 K [Fig. 2(b)] the cluster forms an extended cylinder shape. The analysis of the Phe-Phe orientation (performed over the time while the largest cluster is stable) is shown in Figure 3. Interestingly, the orientation between two phenylalanines (Phe) within the same peptide tends to be rather perpendicular, while phenylalanines of different peptides tend to adopt a rather parallel orientation (as has been found in a recent solid-state NMR study¹⁰). The peptides also tend to

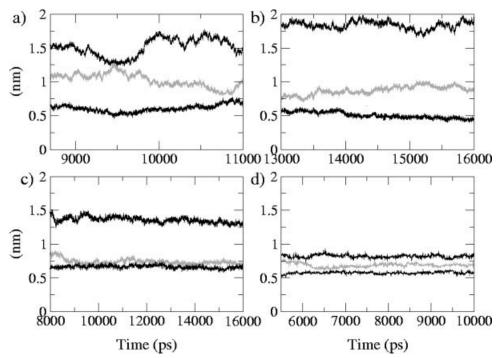


Fig. 4. Time evolution of the eigenvalues of the covariance matrix for (a) DFNKF 340 K, (b) DFNKF 300 K, (c) FF 340 K, (d) AGAIL 340 K.

TABLE I. Values λ_1, λ_2 , and λ_3 for the Different Peptide Systems

	λ_1	σ	λ_2	σ	λ_3	σ (nm)
DFNKF 340 K	1.8380	0.0005	0.8749	0.0005	0.5109	0.0003
FF 340 K	1.3650	0.0005	0.7675	0.0004	0.6637	0.0004
NFGAIL 340 K	1.3460	0.0014	0.8995	0.0004	0.5201	0.0002
AGAIL 340 K	0.8200	0.0007	0.6873	0.0009	0.5750	0.0005

arrange in a fashion prone to build antiparallel β -sheets. This finding corresponds well to the experimental finding that charged interpeptide interactions (in addition to the hydrophobic interactions) are responsible for antiparallel stacking in hCT (human calcitonin) fibrils. 10

FF: The simulation of FF at 340 K was performed for 16 ns starting from a "random" distribution of the single peptides in the simulation box. In Figure 1(c), the cluster analysis of the trajectory is reported. After 3 ns of simulation time the single molecules built a system of two clusters. After 8 ns a single stable cluster is formed which adopts an extended cylindrical shape [see Fig. 2(c)]. The analysis of the Phe–Phe orientation [see Fig. 3(e,f)] shows that the orientation between two phenylalanines tends to be rather perpendicular regardless of the interacting residues being from the same or different peptides.

AGAIL: The two simulations of AGAIL performed at 300 and 340 K were performed for 10 ns each starting from "random" distributions. In the following, we only report the analysis of the simulation at 340 K, since both simulations show very similar overall features. In Figure 1(d) the cluster analysis of the trajectory is reported. After less than 1 ns of simulation time the single molecules have build a stable system of two clusters, one of 24 and

TABLE II. Values d_1 and d_2 for the Different Peptide Systems

	d_1	σ	d_2	σ
DFNKF 340 K	1.5011	0.0024	0.2393	0.0002
FF 340 K	1.7182	0.0011	0.2921	0.0003
NFGAIL 340 K	1.4964	0.0022	0.2316	0.0001
AGAIL 340 K	1.1954	0.0024	0.4557	0.0005

one of 2 particles. The distribution of the Φ , Ψ backbone dihedral angles over the whole simulation time shows that also for this system the β -sheet region is most populated (data not shown) and therefore the peptides are mostly in a rather extended conformation. The cluster adapts a sphere-shaped structure. In opposition to the previously published NFGAIL system, here we find a less ordered cluster with specific characteristics (see last paragraph of this section).

To access general features of the different clusters the eigenvalues of the principal geometrical axes were calculated for all simulations. These eigenvalues which can account for the geometrical shape of the cluster are calculated along the MD trajectories after the formation of a unique stable cluster and are reported in Figure 4. For

TABLE III. Principal Geometrical Axis of the Single Peptide

	x-axis	σ	y-axis	σ	z-axis	σ (nm)
DFNKF 340 K	0.4523	0.0002	0.1078	0.0002	0.0341	0.0002
FF 340 K	0.1950	0.0005	0.0363	0.0001	0.0140	0.0001
NFGAIL 340 K	0.4461	0.0002	0.1598	0.0001	0.0854	0.0001
AGAIL 340 K	0.3530	0.0001	0.1090	0.0001	0.0819	0.0001

an overall characterization of the cluster shape we define:

$$d_1 = \lambda_1/\lambda_2 \tag{1}$$

and

$$d_2 = \lambda_3/(\lambda_1 + \lambda_2) \tag{2}$$

where λ_1 , λ_2 , and λ_3 are the time average of first, second, and third eigenvalues respectively.

 d_1 shows the relation of the lengths of the cluster towards its widths, while d_2 sets the thickness of the cluster in relation to the sum of the lengths and widths. For a perfect sphere $d_1 = 1$ and $d_2 = 0.5$. The values for the analyzed peptide systems are given in Table I.

As can be seen in Figure 4(a–c) and from the values d_1 and d_2 in Table II the clusters for DFNKF, FF, and NFGAIL are characterized by a large first eigenvector with respect to the other two. Also d_1 and d_2 indicate that these clusters are long extended flat structures. On the other side, AGAIL [see Fig. 4(d)] clusterizes in a rather spherical shape.

The extension of the single peptides is investigated with a variation of the covariance analysis (considering the backbone atoms of each peptide) where the geometric center of the single molecule is taken as a reference point. In each case the average value for the lifetime of each cluster is shown in Table III. It is clearly visible that the amyloidogenic species DFNKF, FF, and NFGAIL exhibit elongated peptides, while the control peptide AGAIL shows a more bent structure.

We also investigated the single peptide orientation with respect to the main cluster axes (see Fig. 6). Therefore, two angles α and β (see Fig. 5) are defined to characterize the orientation of the first principal geometrical axis of the single peptides in the reference frame given by the three principal geometrical axes of the cluster. Interestingly, for DFNKF we find that the peptides tend to orientate themselves perpendicular to the main cluster axis [see Fig. 6(a,b)]. A secondary structure analysis revealed that for the 340 K simulation an antiparallel β -sheet (as found in the NMR study 10) is formed which is stable for 1.5 ns. In a similar fashion the FF peptides tend to arrange perpendicular to the main cluster axis [see Fig. 6(c)], while the AGAIL peptides don't show such a preferential orientation [Fig. 6(d)].

DISCUSSION AND CONCLUSION

It has been found that the systems that naturally aggregate in the conditions given here, form particular clus-

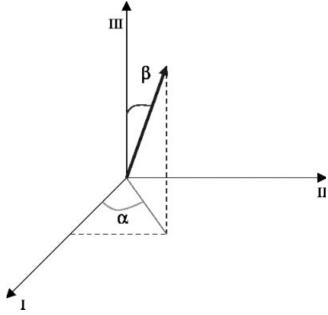


Fig. 5. Definition of the angles α and β to characterize the orientation of the first principal geometrical axis of the single peptides in the reference frame given by the three principal geometrical axes of the cluster.

ter types during the simulation of the first aggregation steps. In particular, DFNKF, FF, and NFGAIL build long flat aggregated structures, while the clusters of AGAIL adopt a more bulky geometrical shape. For DFNKF, FF, and NFGAIL we find the tendency of the peptides to adopt an extended conformation while aggregating (see Table III). In a previous study9 we also demonstrated that this tendency of the peptides to be elongated is independent of the starting conformation of the single peptides (α -helix, β -turn, extended). The interactions between aromatic residues seem to play a critical role in aggregation, which corresponds well with findings of other experimental and computational studies.^{8,9,18} For example, in the case of the IAPP motif it has been shown that the interaction of aromatic side-chains plays a major role in the determination of the aggregation kind. While the substitution of the Phe with a less hydrophobic Trp still results in self-organization of the peptides into amyloidlike structures the Phe → Ala mutation significantly reduces their amyloidogenic potential. Recently, two high resolution structures of amyloid assemblies involving aromatic interactions have been published.^{25,26} In fibrilforming tau fragments tyrosine-tyrosine interactions stabilize the intersheet layers and this particular contact is proposed as a possible aromatic target for drugs. 25 In our

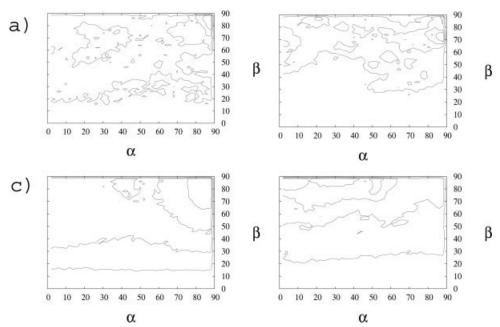


Fig. 6. Density map of the α , β distribution along the trajectories: (a) DFNKF 340 K, (b) DFNKF 300 K, (c) FF 340 K, (d) AGAIL 300 K.

simulated assemblies we find predominately parallel $\pi-\pi$ stacking between the Phe residues of different peptides, where these tend to form antiparallel β -sheets. Serpell and coworkers determined the detailed structure of a highly stable amyloid fiber composed of a sequence-designed polypeptide including four phenylalanines. 26 The Phe interactions between the adjacent antiparallel strands are dominated by parallel-displaced rings and hence these results compares well with our findings. Our analysis suggests that extended flat aggregation structures, where the peptides arrange perpendicular to the main cluster axis and optimize their aromatic side-chain interactions, might precede the typical protofilament and fibril arrangements which are formed over longer time-scales.

ACKNOWLEDGMENTS

The authors thank the computer center CASPUR (Rome, Italy), E. Gazit for useful discussions, and R. Sacco for his contribution to the work.

REFERENCES

- Thirumalai D, Klimov DK, Dima RI. Emerging ideas on the molecular basis of protein and peptide aggregation. Curr Opin Struct Biol 2003;13:146–159.
- Stefani M, Dobson CM. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. J Mol Med 2003;81:678–699.
- Harper JD, Wong SS, Lieber CM, Lansbury PT. Observation of metastable Ab amyloid protofibrils by atomic force microscopy. Chem Biol 1997;4:119–125.
- Selkoe DJ. Alzheimer's disease: genotypes, phenotypes and treatments. Science 1997;275:630,631.

- Prusinger S. Prions. Proc Natl Acad Sci USA 1998;95:12580– 12585.
- Lesné S, Koh MT, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallagher M, Ashe KH. A specific amyloid-b protein assembly in the brain impairs memory. Nature 2006;440:352–357.
- Görbitz CH. Nanotube formation by hydrophobic dipeptides. Chem Eur J 2001;7:5153–5159.
- 8. Reches M, Gazit E. Casting metal nanowires within discrete self-assembled peptide nanotubes. Science 2003;300:625–627.
- Colombo G, Daidone I, Gazit E, Amadei A, Di Nola A. Molecular dynamics simulation of the aggregation of the core recognition motif of the islet amyloid polypeptide in explicit water. Proteins 2005;59:519–527.
- Naito A, Kamihira M, Inoue R, Saito H. Structural diversity of amyloid fibril formed in human calcitonin as revealed by sitedirected ¹³C solid-state NMR spectroscopy. Magn Reson Chem 2004;42:247–257.
- Reches M, Porat Y, Gazit E. Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. J Biol Chem 2002;277:35475–35480.
- Tsai H, Zanuy D, Haspel N, Gunasekaran K, Ma B, Tsai C, Nussinov R. The stability and dynamics of the human calcitonin amyloid peptide DFNKF. Biophys J 2004;87:146–158.
- Zanuy D, Haspel N, Tsai H, Ma B, Gunasekaran K, Wolfson HJ, Nussinov R. Side chain interactions determine the amyloid organization: a dingle layer β-sheet molecular structure of the calcitonin peptide segment 15–19. Phys Biol 2004;1:89–99.
- Kapuriotu A, Schmauder A, Tenidis K. Structure-based design and study of nonamyloidogenic, double N-methylated IAPP amyloid core sequences as inhibitors of IAPP amyloid formation and cytotoxicity. J Mol Biol 2002;312:339–350.
- 15. Azriel R, Gazit E. Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. An experimental support for the key role of the phenylalanine residue in amyloid formation. J Biol Chem 2001;276:34156–34161.
- Zanuy D, Porat Y, Gazit E, Nussinov R. Peptide sequence and amyloid formation: molecular simulations and experimental study of a human islet amyloid polypeptide fragment and its analogs. Structure 2004;12:439–455.
- Wu C, Lei H, Duan Y. Formation of partially ordered oligomers of amyloidogenic hexapeptide (NFGAIL) in aqueous solution observed in molecular dynamics simulations. Biophys J 2004;87:3000–3009.

- Wu C, Lei H, Duan Y. The role of Phe in the formation of well-ordered oligomers of amyloidogenic hexapeptide (NFGAIL) observed in molecular dynamics simulations with explicit solvent. Biophys J 2005;88:2897–2906.
- Insight II User Guide, version 2.1.0. San Diego: Biosym Technologies; 1992.
- van der Spoel D, van Drunen R, Berendsen HJC. GROningen MAchine for chemical simulation. Department of Biophysical Chemistry, BIOSON Research Institute, Nijenborgh, Groningen, 1994
- van Gunsteren WF, Billeter S, Eising A, Hünenberger P, Kruger P, Mark AE, Scott W, Tironi I. Biomolecular simulations: the GRO-MOS96 manual and user guide. Zürich, Groningen: Biomos b.v.; 1996
- 22. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
- Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
- Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. J Chem Phys 1993;98: 10089–10092.
- 25. Inouye H, Sharma D, Goux WJ, Kirschner DA. Structure of core domain of fibril-forming PHF/Tau fragments. Biophys J 2006;90:1774–1789.
- Makin OS, Atkins E, Sikorski P, Johansson J, Serpell LS. Molecular basis for amyloid fibril formation and stability. Proc Natl Acad Sci USA 2005;102:315–320.