Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

for

"*Effects of mutations in de-novo designed synthetic amphiphilic β-sheet peptides on fibrils self-assembly*"

Yoav Raz[1,4], Boris Rubinov[1,4], Maayan Matmor[2], Hanna Rapaport[3,4], Gonen Ashkenasy[1,4,*] and Yifat Miller[1,4,*]

[1]Department of Chemistry, [2]Department of Materials Engineering, [3]Department of Biotechnology Engineering, [4]Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beér-Sheva 84105, Israel.

*Corresponding authors:
ymiller@bgu.ac.il ; gonenash@bgu.ac.il

Materials and Methods

1. Computational Details

All simulations were performed using the high-performance computational facilities of the Miller lab in the BGU HPC computational center. The support of the BGU HPC computational center staff is greatly appreciated.

Construction of the self-assembled conformers of peptides M and N

In order to study the self-assembly of peptides **M** and **N**, we constructed a conformational ensemble of 4 initial arrangements, while considering all possible organizations within the bilayer. We chose a total of 24 monomers for each peptide in the bilayer level, with 12 monomers in each layer.We considered four major types of conformational variabilities in the bilayer for each of the **M** and **N** peptides: (1) β-strands backbone alignments in antiparallel orientations in each layer and in antiparallel orientation within the two layers (M1 and N1); (2) β-strands backbone alignments in antiparallel orientations in each layer and in parallel orientation within the two layers (M2 and N2); (3)β-strands backbone alignments in parallel orientations in each layer and in parallel orientation within the two layers (M3 and N3); (4) β-strands backbone alignments in parallel orientations in each layer and in antiparallel orientation within the two layers (M4 and N4).

It should be noted here that in this study we did not consider heterogeneous combinations of the aforementioned four conformational variability. It may be reasonable to consider a combination of various arrangements within one fibril, and future studies in our lab may examine such combinations.

Molecular dynamics (MD) simulations procedure

MD simulations of the solvated variant models M1-M4 and N1-N4 were performed in NPT ensembles using the NAMD program[1] with the CHARMM27 force-field[2,3] for 60 ns. The models were explicitly solvated with TIP3P water molecules.[4,5] The Langevin piston method[1,6,7] with a decay period of 100 fs and a damping time of 50 fs was used to maintain a constant pressure of 1 atm. The temperature (310 K) was controlled by Langevin thermostat with a damping coefficient of 10 ps^{-1}.[1] The short-range van der Waals (VDW) interactions were calculated using the switching function, with a twin range cutoff of 10.0 and 12.0 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a cut-off of 12.0 Å for all simulations.[8,9] The equations of motion were integrated using the leapfrog integrator with a step of 2 fs. All initial variant models were energy minimized and then solvated in a TIP3P water box with a minimum distance of 15 Å from any edge of the box to any peptide atom. Any water molecule within 2.5 Å of the variant model was removed. Counter-ions (Na^+) were added at random locations to neutralize the self-assembled peptides' charge.

The solvated systems were energy minimized for 2000 conjugated gradient steps, where the distance between the β-sheets in the peptides was fixed in the 2.2 – 2.5 Å range. The counter-ions and water molecules were allowed to move. The hydrogen atoms were constrained to the equilibrium bond using the SHAKE algorithm.[10] The minimized solvated systems were heated at 200 K, where all atoms were allowed to move. Then, the systems were heated from 200 K to 250 K for 300 ps and equilibrated at 310 K for 300 ps. All simulations ran for 60 ns, and structures were saved every 10 ps for analysis. These conditions (310 K and 60 ns of timescales) were applied to test the stabilities of all the variant models.

Analysis details

To obtain the relative stability and the populations of the variant models of the self-assembled variant models, the variants' trajectories of the last 5 ns were first extracted from the explicit MD simulation, excluding water molecules. The solvation energies of all systems were calculated using the Generalized Born Method with Molecular Volume (GBMV).[11,12] In the GBMV calculations, the dielectric constant of water was set to 80.0. The hydrophobic solvent-accessible surface area (SASA) term factor was set to 0.00592 kcal/mol·Å^2. Each variant was minimized by 1000 cycles and the conformation energy was evaluated by grid-based GBMV. The minimization does not change the conformations of each variant, but only relaxes the local geometries due to thermal fluctuations which occurred during the MD simulations.

A total of 2000 conformations (500 conformations for each of the 4 examined conformers) for each self-assembled peptide design model (PEFEFEFEFEFEP and PEFEFACEFEFEP) were used to construct the free energy landscapes of the conformers and to evaluate the conformer probabilities by using Monte Carlo (MC) simulations. In the first step, one conformation of conformer i and one conformation of conformer j were randomly selected. Then, the Boltzmann factor was computed as $e^{-(Ej - Ei)/KT}$, where E_i and E_j are the conformational energies evaluated using the GBMV calculations for the respective conformations i and j, K is the Boltzmann constant and T is the absolute temperature (298 K used here). If the Boltzmann factor value is larger than a random number, the move from conformation i to conformation j is allowed. After 1 million steps, the conformations visited for each conformer were counted. Finally, the relative probability of a conformer n was evaluated as: $P_n = N_n / N_{total}$, where P_n is the population of conformer n, N_n is the total number of conformations visited for the conformer n, and N_{total} the total number of steps. The advantages of using the MC simulations to estimate conformer probability rely on the facts that the MC simulations have good numerical stability and allow transition probabilities among several conformers to be controlled. The populations of the conformers are only indicative.

Calculations of the populations of all four **M** models show that model M1 occupies higher populations than each one of the 3 other **M** models. Two self-assembled designed models of **N**,

N1 and N2, have preference than models N3 and N4 models, since N1 and N2 illustrate relatively higher populations than N3 and N4.

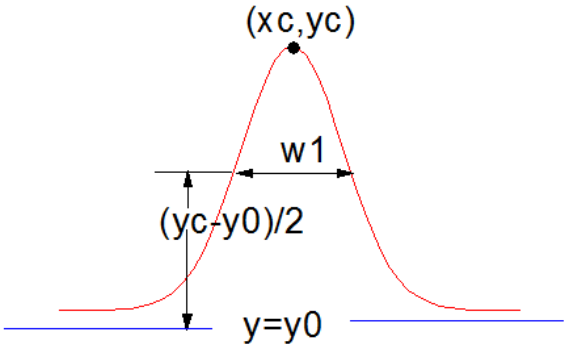
We further examined the stability of the self-assembled peptides by following changes in the number of hydrogen bonds between the β strands, with the hydrogen bond cutoff set to 2.5 Å, and by monitoring the changes in the inter-sheet distance (Cα backbone-backbone distance) in the hydrophobic core domain of all the amphiphilic self-assembled peptides. The core domain is defined as the Cα backbone-backbone distance between two residues in position 7 (which is the "center" of the self-assembled amphiphilic peptide) of each layer.

To estimate the pitch of each model studied here, we computed for each model the length of each fibril-like model and multiplied it by the angle, considering three atoms between Cα atoms of residues Pro1 Phe7/Cys7 of peptide at one end of the fibril and Pro1 at the second end. For the computed values we considered the average values of the total simulations of 60 ns.

Analysis of populations for models M1-M4 and N1-N4

To compare the result of the relative conformational energies obtained from GBMV method, we used Origin Pro 8 to analyze the distribution of the conformational energies. All 500 values of each conformer extracted from the last 5 ns of the simulations had been used to characterize the distribution by histograms (seen in black boxes in Figures S3 and S4). Two peaks had been chosen to fit a Gaussian function for each peak for all conformers. The integration of the fitted Gaussians had been achieved by Origin Pro 8.

The definition for the Gaussian function we used is:



A>0
offset:y0=1
center:xc=2
width:w=1.5
area:A=5

yc=y0+A/(w*sqrt(pi/2))
w=w1/sqrt(ln(4))

$$y = y_0 + \frac{A}{w\sqrt{\pi/2}} e^{-2\frac{(x-x_c)^2}{w^2}}$$

Where X_c is the median values, A is the amplitude, w is the width and y_0 are the initial Cartesian coordinates.

Due to some overlap of the peaks, we compared the integrated values observed from the fitted Gaussian and the integrated values obtained from the Riemann integral of the data set of all 500 conformational energies for each model. The deviation between the integrated values obtained from the fitted Gaussians and the integrated values obtained from the Riemann integral was less than 1% for all of the conformers, excluded for N4 that has 5.4% of deviation. These deviations indicate on a good estimation for the integrated peak.

Estimation of the helicity pitches values

The helicity pitch value for each of the fibril conformers was calculated using the following equation:

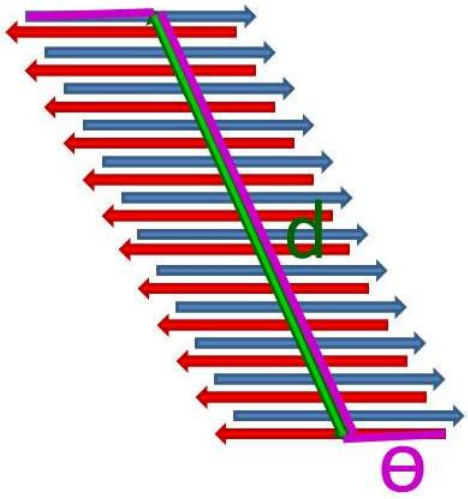
$$hp = \frac{360}{\theta} \times d$$

Using θ and d as schematically illustrated in Scheme S1.

θ – The dihedral angle that forms between four Cα atoms of designated residues of peptides at the ends of the analyzed structure. For M1 model, θ includes the Cα atoms of residues Pro1, Phe7 at one end peptide and Phe7, Pro13 at the other end peptide. For N models, it includes residues Pro1, Cys7 at one end peptide and Cys7, Pro13 at the other end peptide.

d – The distance between Cα atoms of residues 7 located at the end peptides, along the fibril axis.

We have used θ and d average values of 3000 conformations of each model, i.e. the entire ensemble of 30 ns simulations.



Scheme S1.

2. Experimental procedure

β-sheet characterization by CD

A solution with the desired concentration of peptide **M** or **N** in water (pH ~5.5) or ammonium bicarbonate buffer (0.01 M, pH 7) was sonicated and then allowed to equilibrate for 20 minutes at 25 ^{o}C. Spectra were taken on a Jasco-815 CD spectropolarimeter, using a quartz cell with 1.0 mm path length and 4 second averaging times. The CD signals resulting from buffer alone were subtracted from the spectrum of each peptide solution.

Analysis of fibril formation by thioflavin T (ThT) staining

A ThT stock solution (1.0 mM) was prepared by dissolving 1.6 mg ThT in 5 ml Millipore water, filtered through a 0.2 μm filter, and then diluted 10 times with 20 mM MOPS buffer. Solutions of desired concentrations of peptide **M** or **N** were sonicated, and added to the ThT solution (total volume of 100 μL) prior to equilibration and measurements. The fluorescence spectra were recorded on a Varian Cary Eclipse fluorescence spectrometer, 96-microwellplate reader. The excitation wavelength was 440 nm (slit width 5.0 nm) and the emission recorded between 460-560 nm. To correct the signals, the background fluorescence of the ThT solution without peptide was subtracted from the fluorescence intensities obtained for the samples.

Fibril imaging by atomic force microscopy (AFM)

Silicon substrates with native oxide layer were cleaned using a freshly prepared Piranha solution (solution of 3:7 30% H_2O_2 and concentrated H_2SO_4) for 30 min followed by three times immersion in Millipore water for 10 min. (Caution: Piranha is a very strong oxidant and reacts violently with many organic materials.) The peptide solution in buffer pH 7 was put on the surface (10 μL for 10 seconds), followed by drying under nitrogen flow. Atomic force microscopy (AFM) (Solver-Pro, NTMDT, Ru) topography images were acquired using noncontact tips (NSG03 NT-MDT, Ru (1.74 N m−1, 90 kHz, respectively)). Characterization of **M** and **N** fibrils were each made with two samples. Several fibril structures ($n \geq 3$) from each sample were analyzed for acquiring the statistically relevant data needed to calculate the fibrils width and pitch.

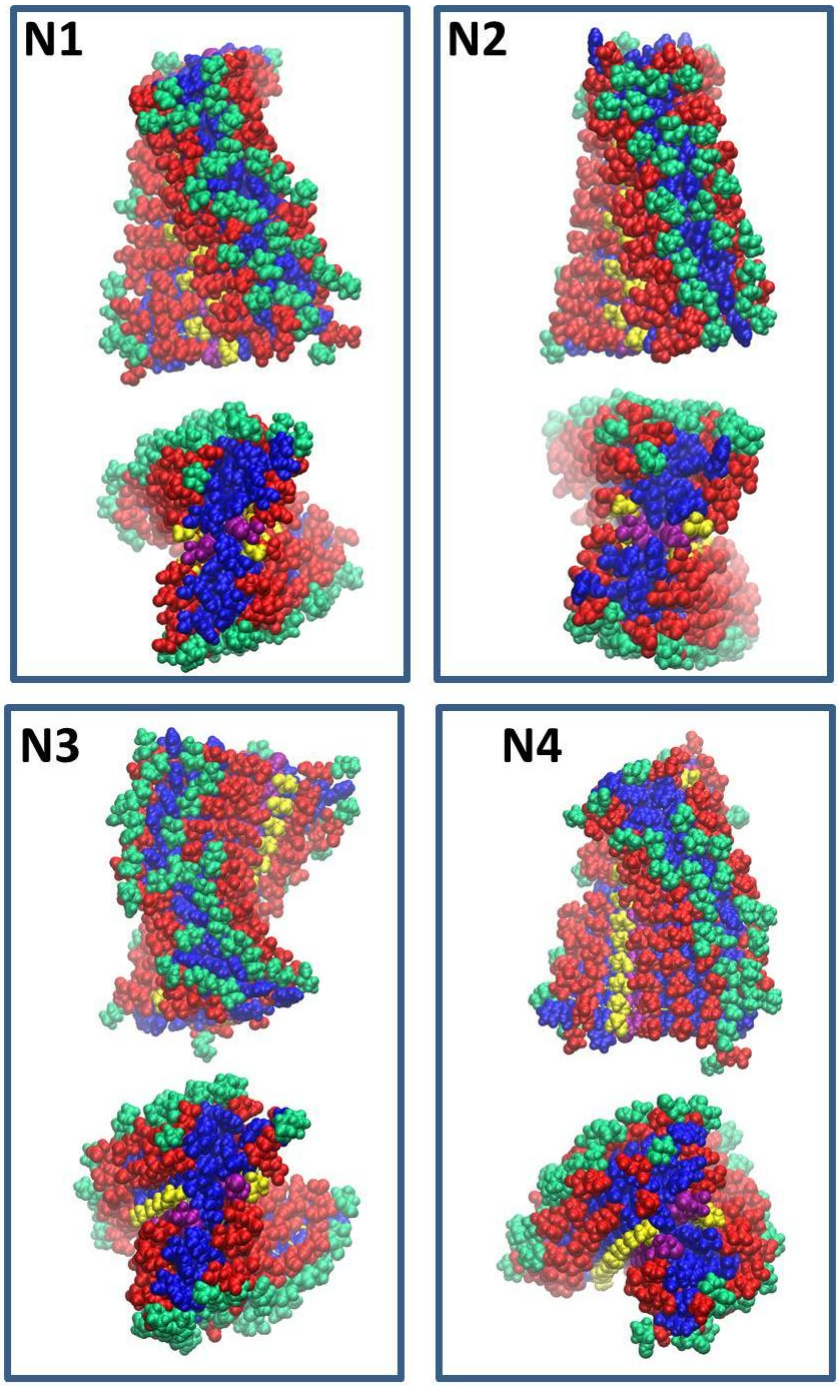


Figure S1: Illustration of the simulated constructed conformers of peptide **N** ^{1}PEFEFACEFEFE^{13}P (a view from the top and from the side of the fibril-like morphology): conformer N1 has two antiparallel monolayers arranged in antiparallel orientation with each other; conformer N2 has two antiparallel monolayers arranged in parallel orientation with each other; conformer N3 has two parallel monolayers arranged in parallel orientation with each

other; conformer N4 has two parallel monolayers arranged in antiparallel orientation with each other. Pro (green), Phe (red), Glu (blue), Ala (yellow), Cys (purple).

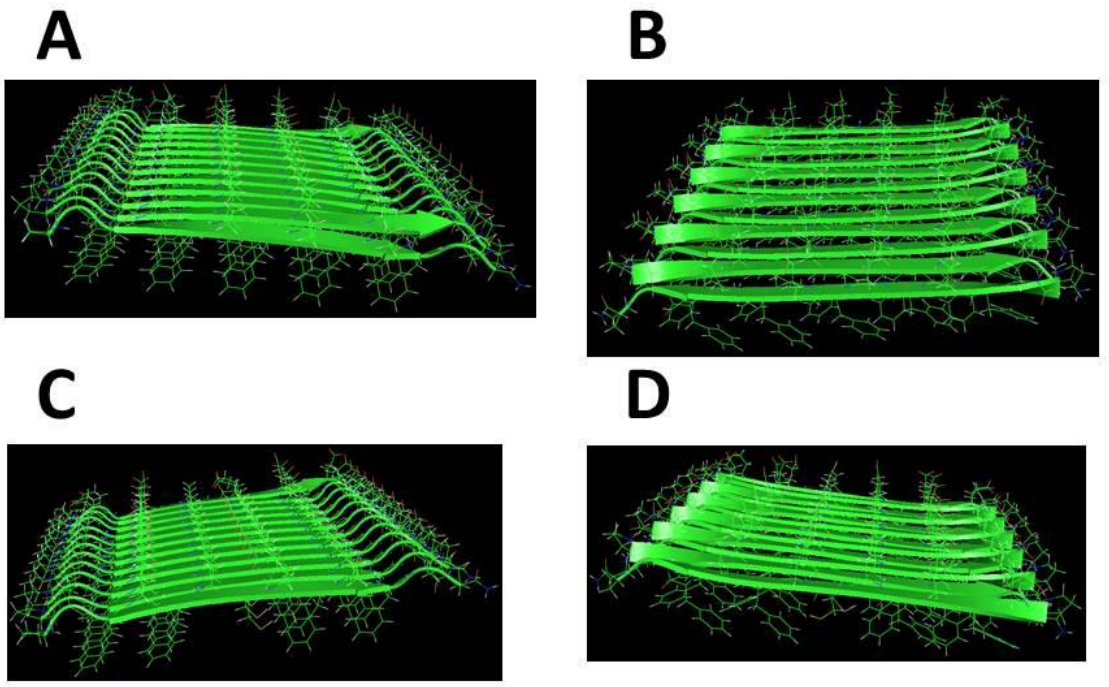


Figure S2: Illustration of the monomer arrangements of peptides ^{1}PEFEFEFEFEFE^{13}P (**M**) in parallel (A) and antiparallel (B), and ^{1}PEFEFACEFEFE^{13}P (**N**) in parallel (C) and antiparallel (D) β-sheets.

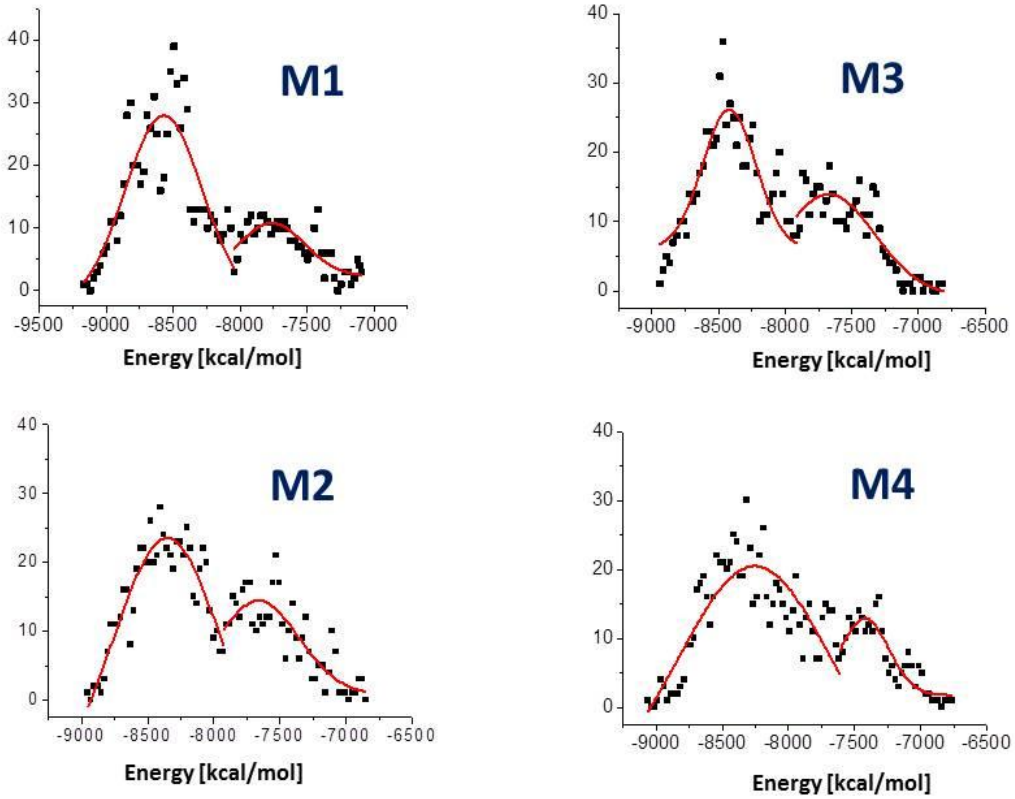


Figure S3: Scatter charts of the 500 conformations obtained from the GBMV energy values extracted from the last 5 ns of each conformer: M1-M4 (black boxes). The scatter charts represent the "histograms" of the number of conformations in energies' ranges. The fitted curves (red line) were computed directly from Origin Pro 8.

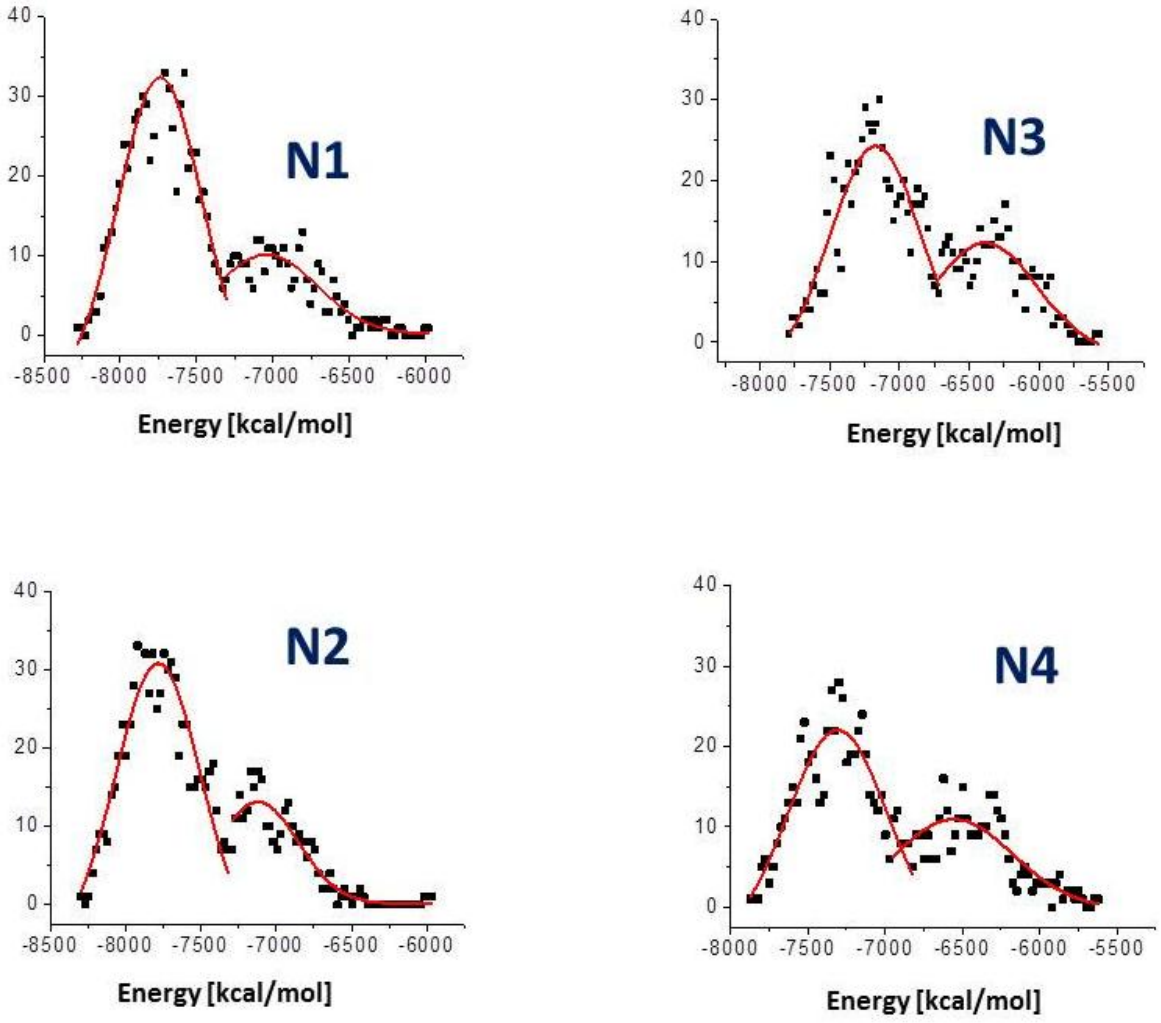


Figure S4: Scatter charts of the 500 conformations obtained from the GBMV energy values extracted from the last 5 ns of each conformer: N1-N4 (black boxes). The scatter charts represent the "histograms" of the number of conformations in energies' ranges. The fitted curves (red line) were computed directly from Origin Pro 8.

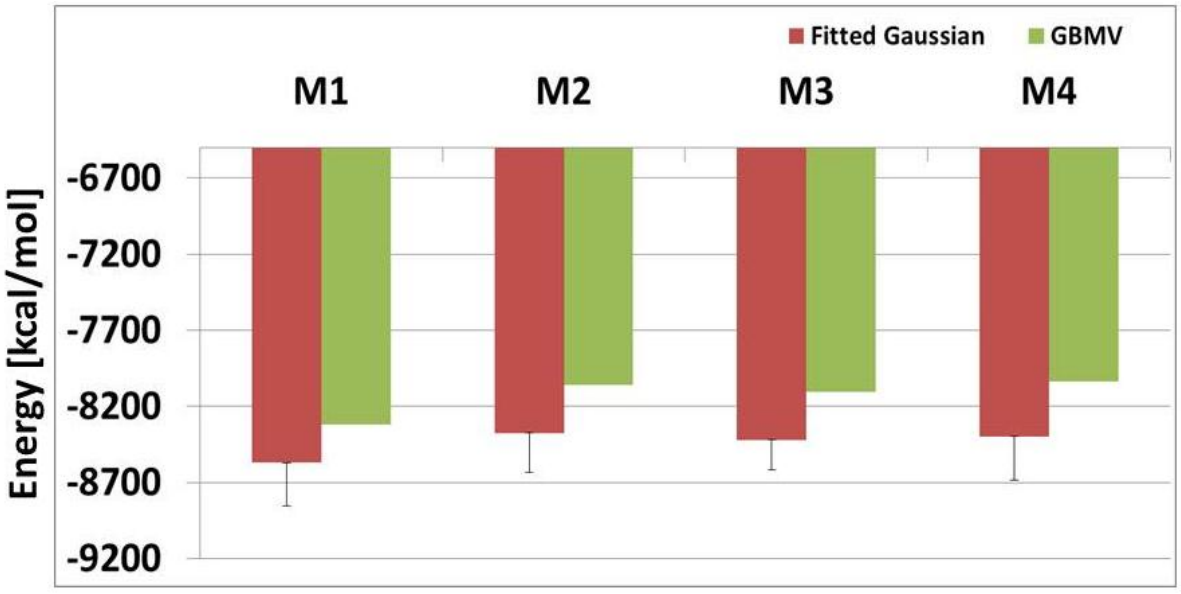


Figure S5: A comparison between the averaged energies of the 500 conformations using the GBMV method and the energies obtained from the primary peak of models M1-M4 using origin's peak fitting function with the standard deviations (Figure S3).

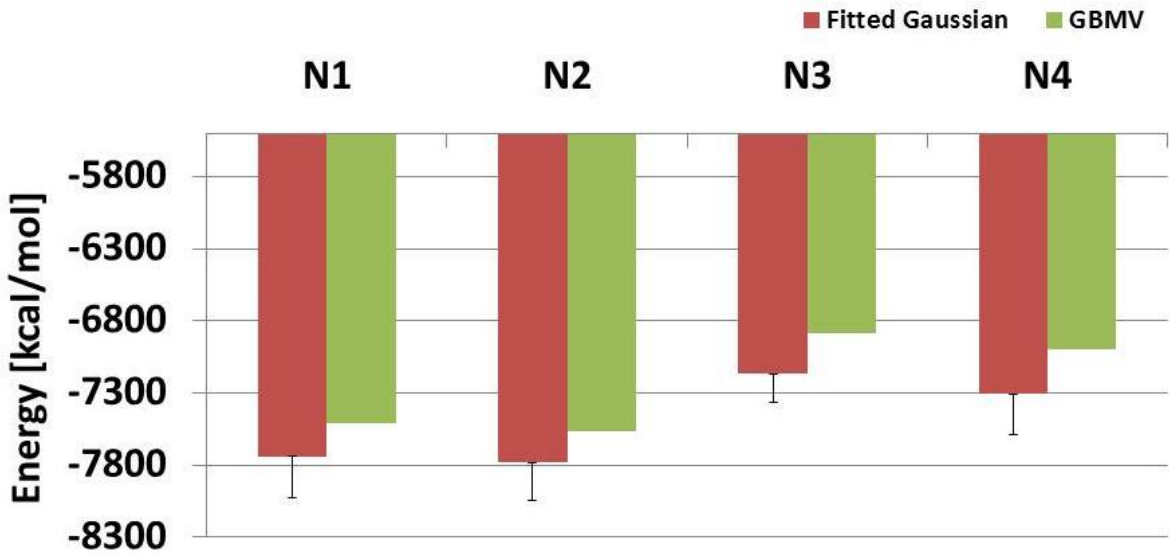


Figure S6: A comparison between the averaged energies of the 500 conformations using the GBMV method and the energies obtained from the primary peak of models N1-N4 using origin's peak fitting function with the standard deviations (Figure S4).

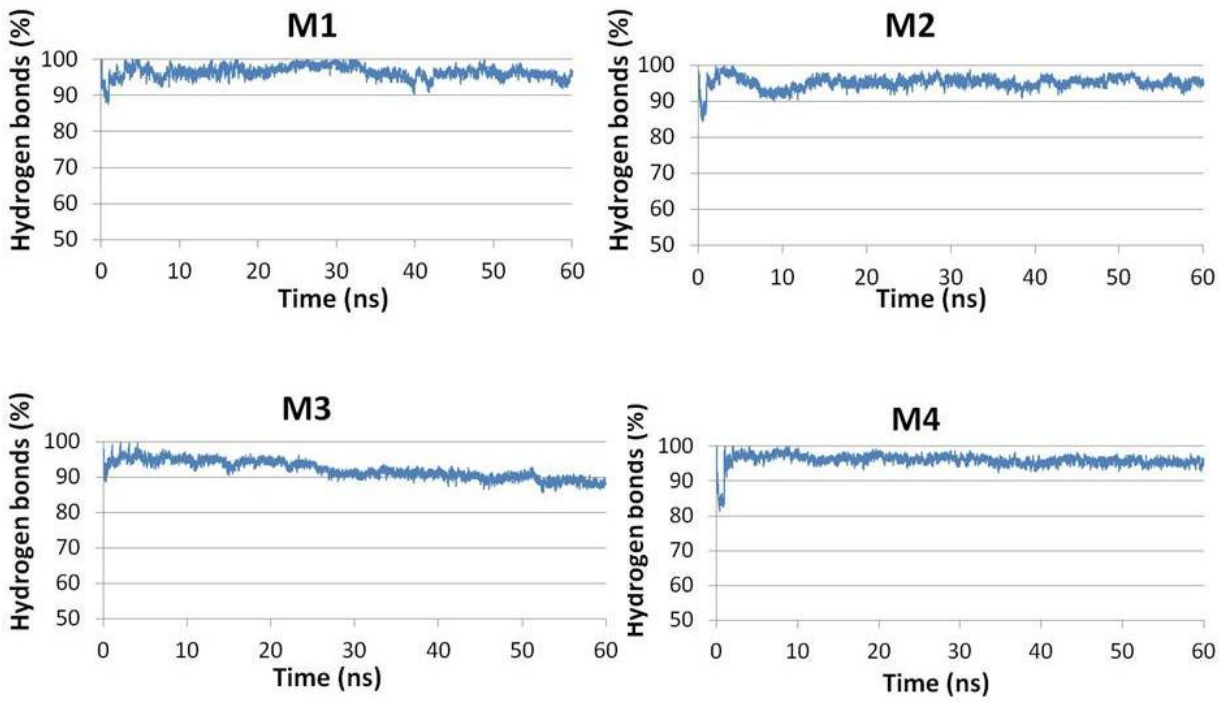


Figure S7: Fraction of the number of hydrogen bonds (in percentage) between all β-strands as compared with the initial structures of M1-M4.

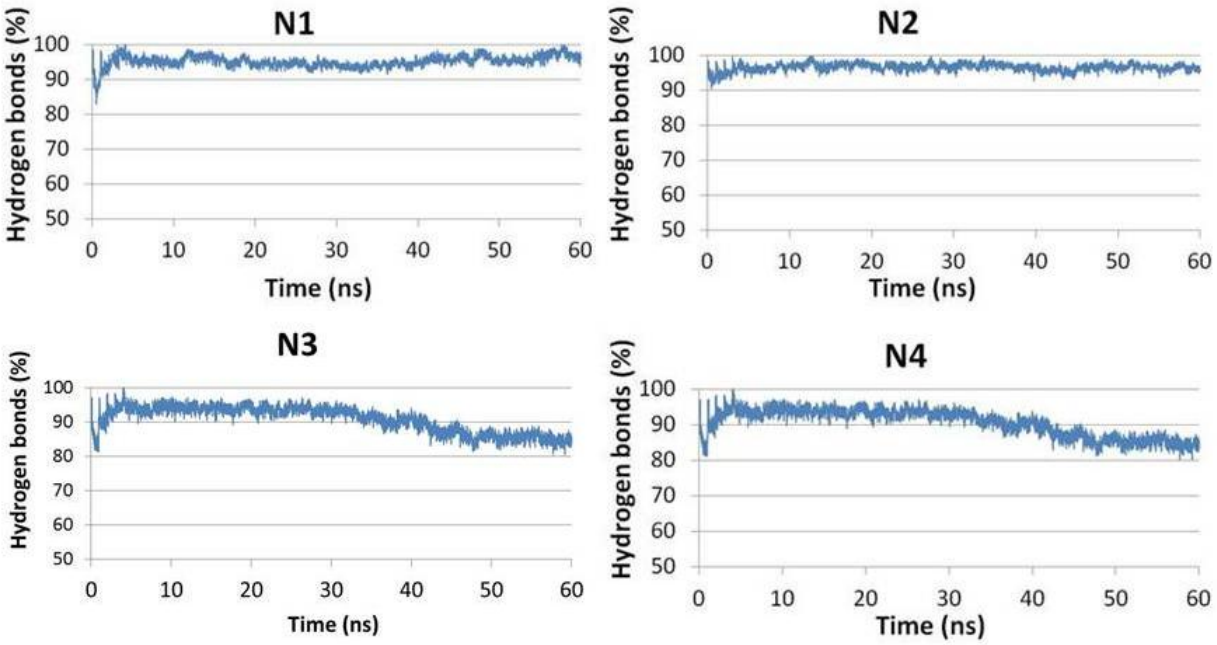


Figure S8: Fraction of the number of hydrogen bonds (in percentage) between all β-strands as compared with the initial structures of N1-N4.

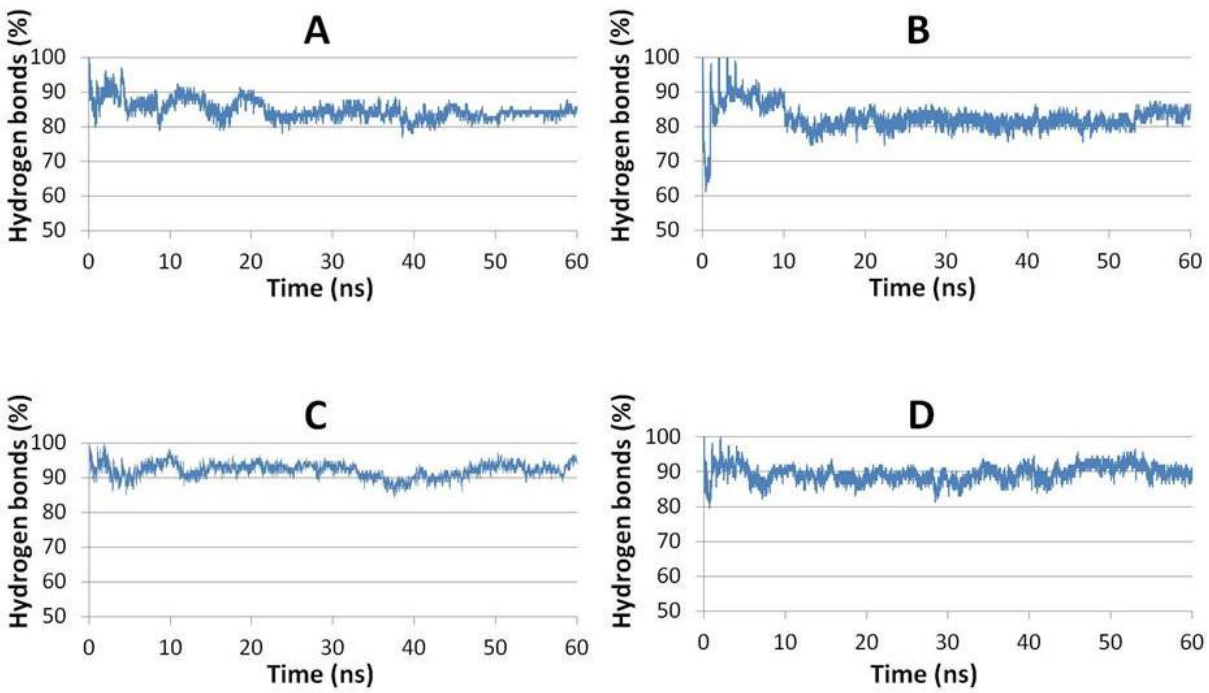


Figure S9: Fraction of the number of hydrogen bonds (in percentage) between all β-strands as compared with the initial structures of the monolayer of the peptides ^{1}PEFEFEFEFEFE^{13}P in antiparallel (A) and parallel (B), and ^{1}PEFEFACEFEFE^{13}P in antiparallel (C) and parallel (D).

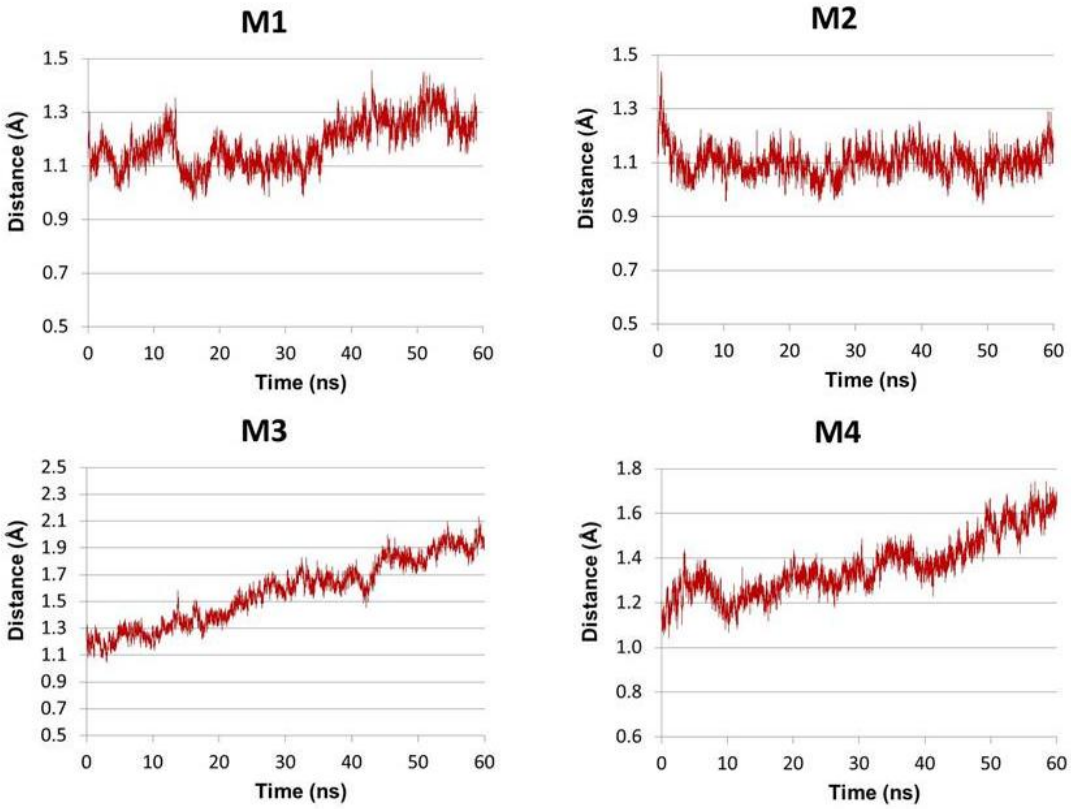


Figure S10: The RMSDs of M1-M4.

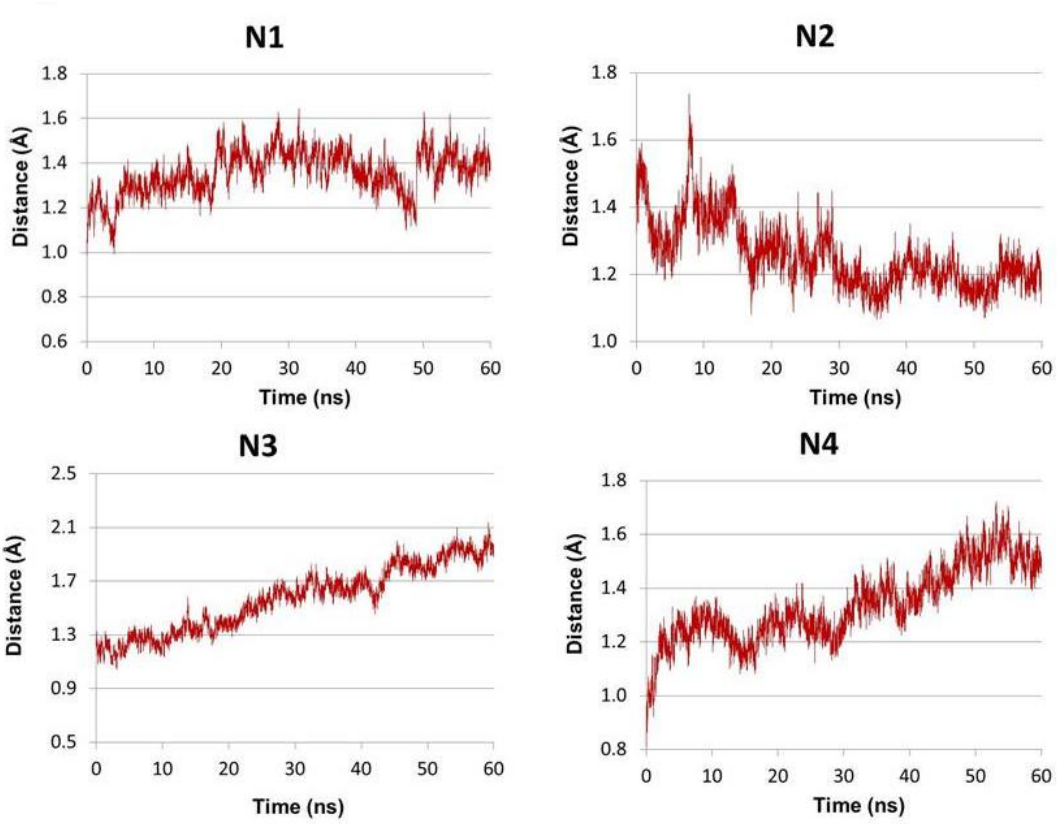


Figure S11: The RMSDs of N1-N4.

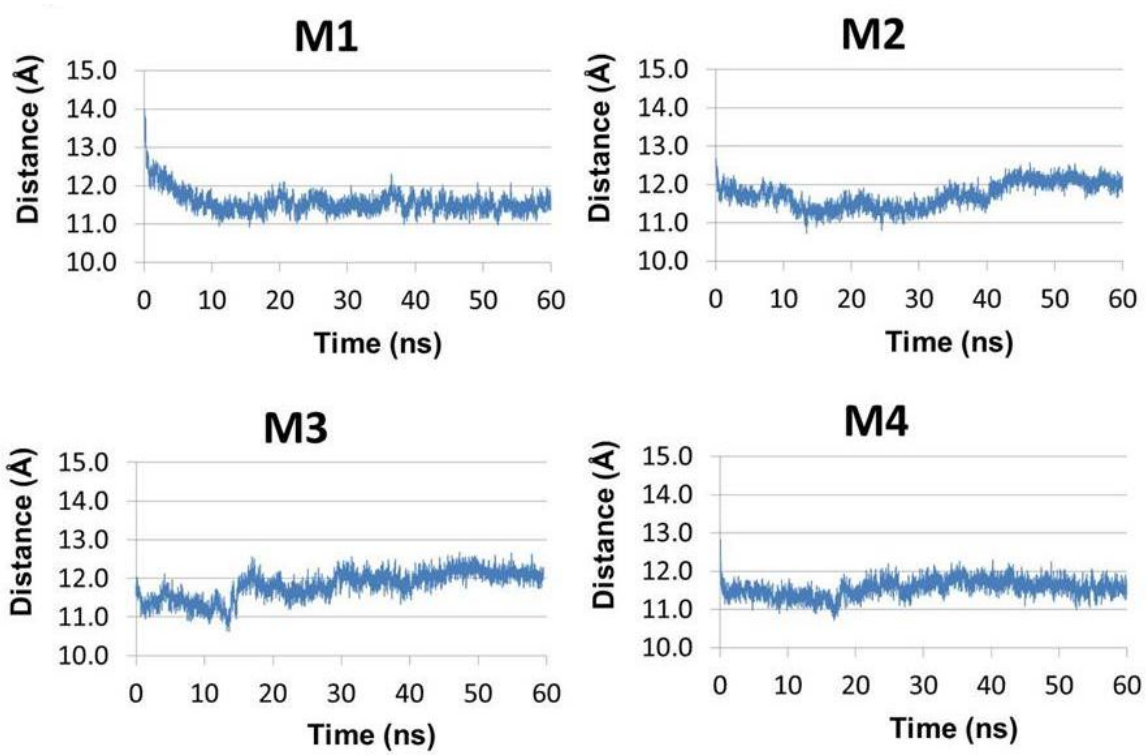


Figure S12: Averaged Cα backbone-backbone distance for models M1-M4 throughout the molecular dynamics (MD) simulations.

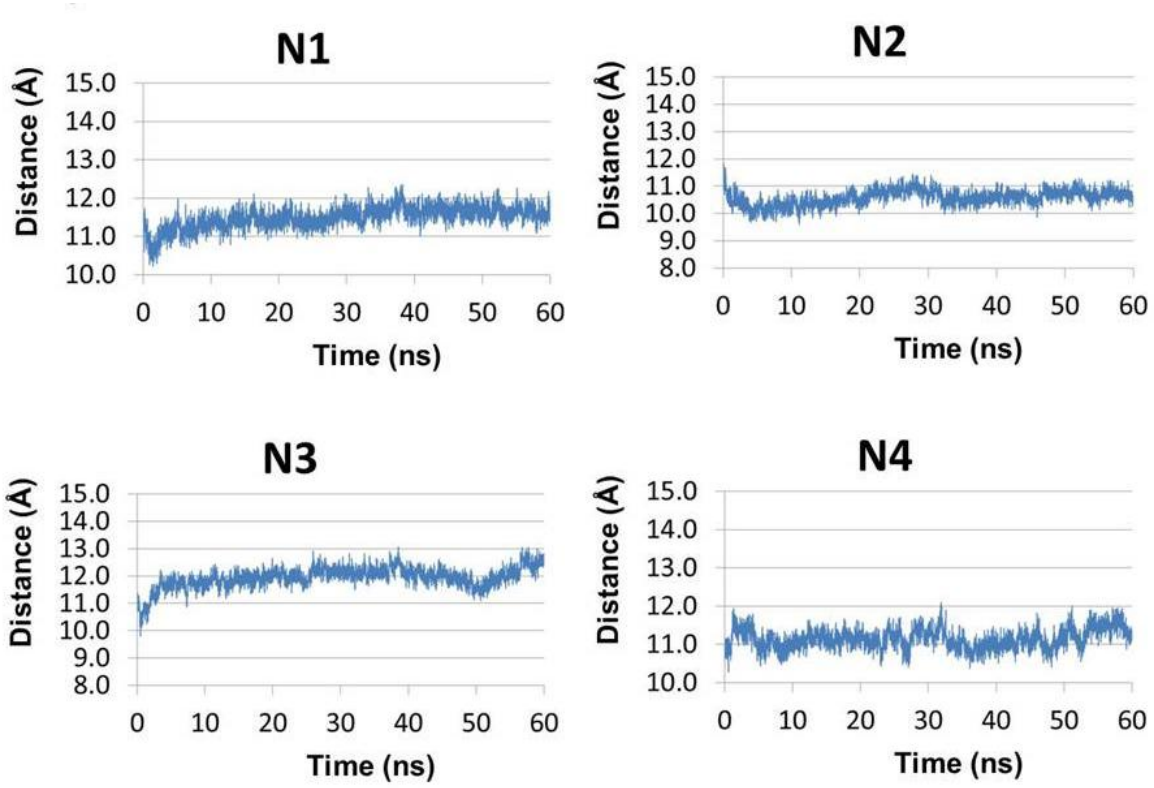


Figure S13: Averaged Cα backbone-backbone distance for models N1-N4 throughout the molecular dynamics (MD) simulation

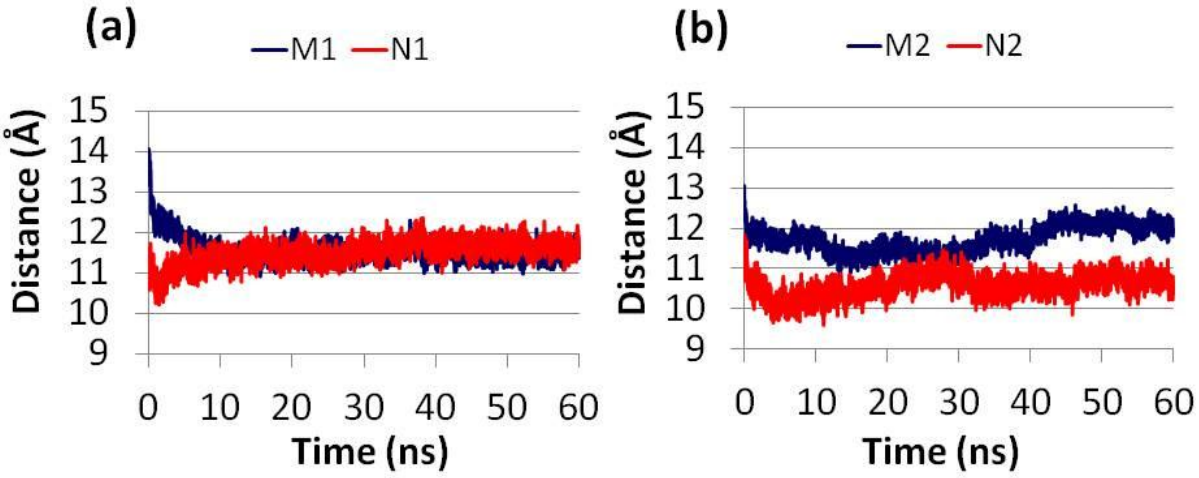


Figure S14: Average Cα backbone-backbone distances of conformers M1 and N1 (a) and M2 and N2 (b), throughout the MD simulations.

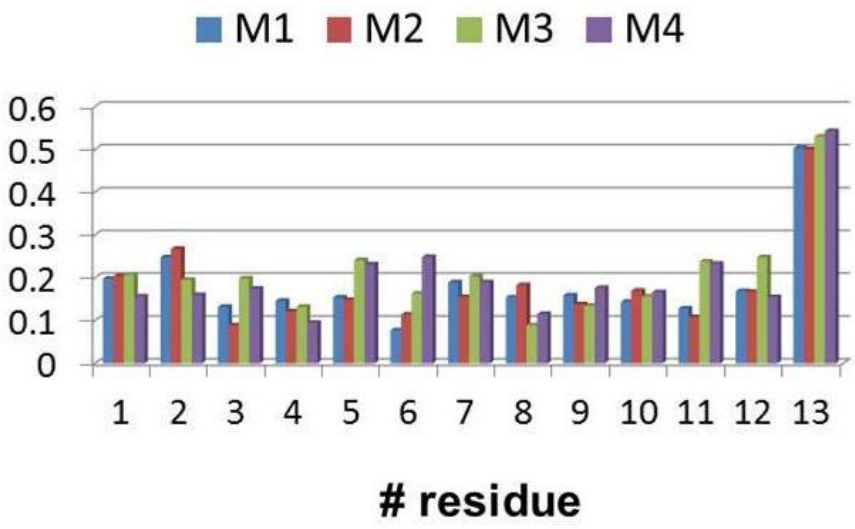


Figure S15: Average number of water molecules around each side chain Cβ carbon (within 4 Å) for models M1-M4. The largest difference between the four models is observed for residue Glu6: in M4 this residue is the more solvated than the others and in N1 this residue is least solvated. Probably, the organizations of parallel β-sheet in the monolayer lead to the Glu6 to be exposed to the water solution.

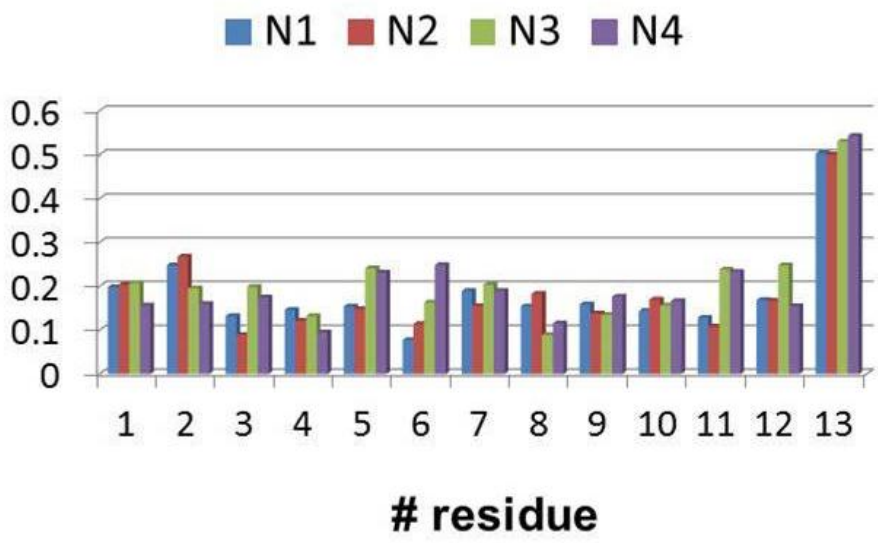


Figure S16: Average number of water molecules around each side chain Cβ carbon (within 4 Å) for models N1-N4. The largest difference between the four models is observed for residue Ala6: in N4 this residue is the more solvated than the others and in M1 this residue is least solvated. Probably, the organizations of parallel β-sheet in the monolayer lead to the Ala6 to be exposed to the water solution.

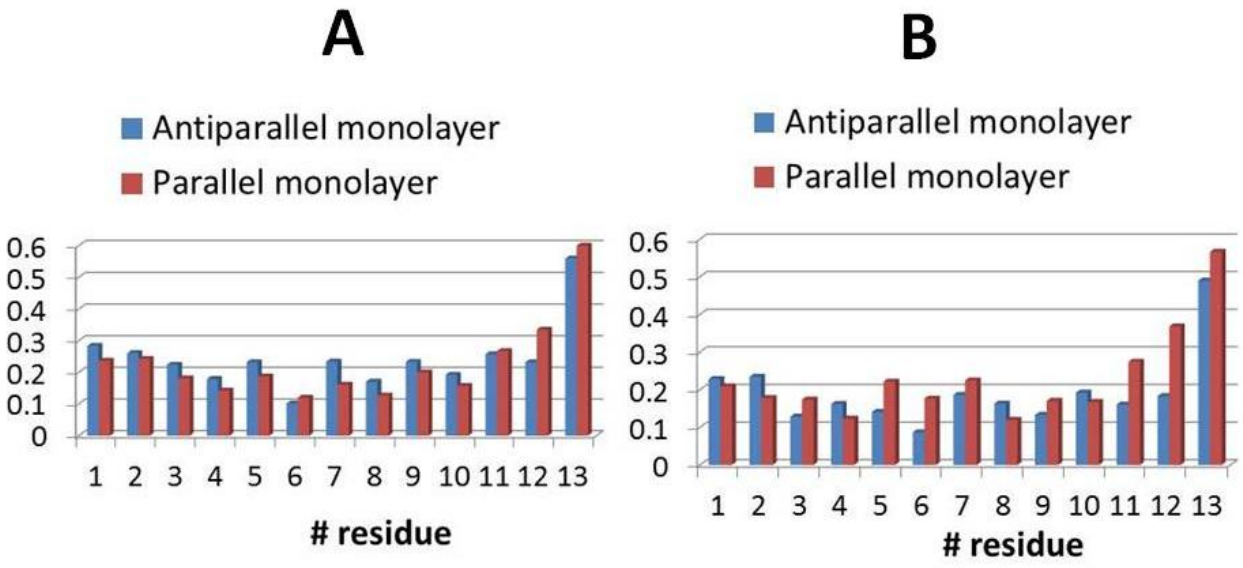


Figure S17: Average number of water molecules around each side chain Cβ carbon (within 4 Å) for monolayer models of the peptide M: ^{1}PEFEFEFEFEFE^{13}P (A) and the peptide N: ^{1}PEFEFACEFEFE^{13}P (B).

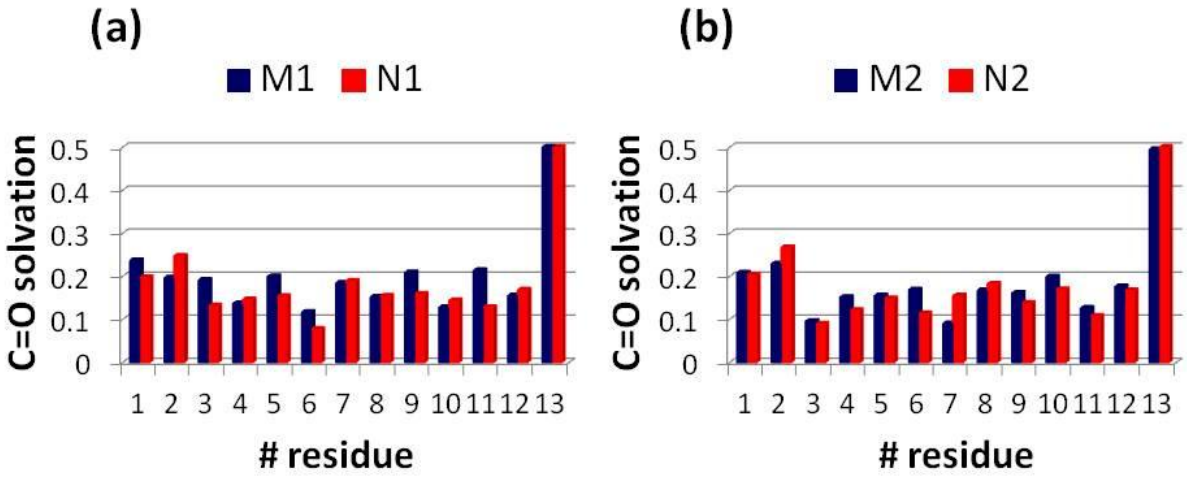


Figure S18: Average number of water molecules around each side chain Cβ carbon (within 4Å) for M1 and N1 (a), and M2 and N2 (b).

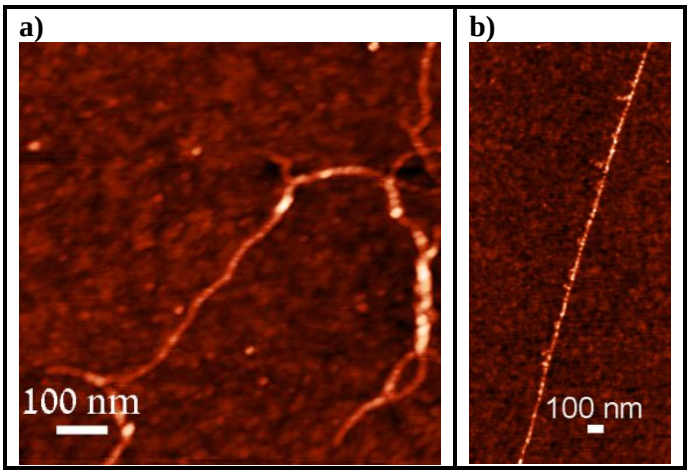


Figure S19: Additional AFM topography images of individual fibril structures formed by peptides **M** (a) and **N** (b). Images were taken after at least 30 minutes of equilibration. These images, and additional fibril images, were used to calculate the reported structural parameters, fibril width and helix pitch.

Table S1: The conformational energies of the simulated self-assembly models M1-M4 and N1-N4.

Conformer	Energy [kcal/mol][a]	Populations (%)
M1	**-8321.5** (468.3)	**31.0**[b]
M2	-8059.3 (467.3)	22.7[b]
M3	-8106.5 (462.3)	24.1[b]
M4	-8036.8 (494.7)	22.2[b]
N1	**-7513.9** (431.0)	**29.0**[b]
N2	**-7567.1** (404.5)	**36.3**[b]
N3	-6884.3 (482.5)	16.9[b]
N4	-7000.6 (494.0)	17.8[b]

[a] Conformational energies were computed using the GBMV calculations (Refs.[11,12]). Standard deviation values are given in parenthesis.
[b] We have previously noted that relatively small differences in the populations (~5% or more) can imply preference of one population among the entire ensemble.[13-16]

Table S2: The conformational energies of the simulated variant monolayer models of the peptide M: ^{1}PEFEFEFEFEFE^{13}P and the peptide N: ^{1}PEFEFACEFEFE^{13}P.

Peptide	Model arrangement	Energy (kcal/mol)
^{1}PEFEFEFEFEFE^{13}P	Antiparallel	-4084.29 (253.24)
^{1}PEFEFEFEFEFE^{13}P	Parallel	-3422.55 (297.62)
^{1}PEFEFACEFEFE^{13}P	Antiparallel	-3511.29 (245.82)
^{1}PEFEFACEFEFE^{13}P	Parallel	-3449.05 (251.48)

*Conformational energies were computed using the GBMV calculations (Refs.[11,12]) Standard deviation values are presented in parenthesis.

Table S3: The solvation energies of the simulated variant models of the peptide ^{1}PEFEFEFEFEFE^{13}P and ^{1}PEFEFACEFEFE^{13}P

Model	Energy (kcal/mol)
M1	-112232.37 (692.10)
M2	-116535.21 (711.87)
N1	-74655.87 (488.23)
N2	-73975.39 (464.24)

*Solvation energies were computed using the GBMV calculations (Refs.[11,12]) Standard deviation values are presented in parenthesis.

References:

(1) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283.

(2) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(3) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(4) Mahoney, M. W.; Jorgensen, W. L. *Journal of Chemical Physics* **2000**, *112*, 8910.

(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926.

(6) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *Journal of Chemical Physics* **1994**, *101*, 4177.

(7) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *Journal of Chemical Physics* **1995**, *103*, 4613.

(8) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993**, *98*, 10089.

(9) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *Journal of Chemical Physics* **1995**, *103*, 8577.

(10) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Computational Physics* **1977**, *23*, 327.

(11) Lee, M. S.; Feig, M.; Salsbury, F. R.; Brooks, C. L. *Journal of Computational Chemistry* **2003**, *24*, 1348.

(12) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *Journal of Chemical Physics* **2002**, *116*, 10606.

(13) Miller, Y.; Ma, B.; Nussinov, R. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 9490.

(14) Miller, Y.; Ma, B.; Tsai, C. J.; Nussinov, R. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 14128.

(15) Miller, Y.; Ma, B.; Nussinov, R. *Journal of the American Chemical Society* **2011**, *133*, 2742.

(16) Miller, Y.; Ma, B.; Nussinov, R. *Biophysical journal* **2009**, *97*, 1168.