Cell Reports Physical Science, Volume *2*

Supplemental information

Solid-state packing dictates the unexpected solubility of aromatic peptides

Santu Bera, Xuewei Dong, Bankala Krishnarjuna, Shannon A. Raab, David A. Hales, Wei Ji, Yiming Tang, Linda J.W. Shimon, Ayyalusamy Ramamoorthy, David E. Clemmer, Guanghong Wei, and Ehud Gazit

1. Supplemental Experimental Procedures

Preparation of peptide assemblies. For assembly, peptides in the required concentration were dissolved in double distilled water by vigorous vortexing for 2 min. The peptide solutions were then incubated at 18°C for two weeks with frequent shaking before examination.

Dynamic light scattering (DLS). Eight hundred µL of the sample solution at required concentration was introduced into a DTS1070 folded capillary cell (Malvern, Worcestershire, U.K.), and the size was measured using a Zetasizer Nano ZS analyzer (Malvern Instruments, Malvern, UK) at 25.0 °C and a backscatter detector (173°). Three measurements were performed and averaged for accuracy.

Scanning electron microscopy (SEM). A 5 µL aliquot was allowed to dry on a microscope glass cover slip at ambient conditions over night and coated with Au. SEM images were recorded using a JSM-6700F FE-SEM (JEOL, Tokyo, Japan) operating at 10 kV.

Fourier-transform infrared (FTIR) spectroscopy. A 30 µL aliquot of the peptide solution was deposited onto disposable KBr infrared sample cards (Sigma-Aldrich, Rehovot, Israel), which were then allowed to dry under vacuum. The samples were saturated twice with 30 µL of D_2O and vacuum dried. FTIR spectra were collected using a nitrogen purged Nicolet Nexus 470 FTIR spectrometer (Nicolet, Offenbach, Germany) equipped with a deuterated triglycine sulfate (DTGS) detector. Measurements were performed using a 4 cm^{-1} resolution and by averaging 64 scans. The absorbance maxima

values were determined using an OMNIC analysis program (Nicolet). The background was subtracted using a control spectrum.

Circular dichroism (CD) spectroscopy. CD spectra were collected using a Chirascan spectrometer (Applied Photophysics, Leatherhead, UK) fitted with a Peltier temperature controller set to desired temperature, using quartz cuvettes with an optical path length of 0.1 mm (Hellma Analytics, Müllheim, Germany). Absorbance of the sample was kept within the linear range of the instrument during measurement. Data acquisition was performed in steps of 1 nm at a wavelength range of 190 to 260 nm with a spectral bandwidth of 1.0 nm and an averaging time of 3 s. The spectrum of each sample was collected three times and averaged. Baseline was similarly recorded for phosphate buffer and subtracted from the samples spectra. Data processing was performed using Pro-Data Viewer software (Applied Photophysics, Leatherhead, UK).

Powder X-ray diffraction (XRD). The lyophilized peptide powder was dissolved in double distilled water and allowed to self-assemble by incubation at 18°C for four weeks. The sample was then centrifuged for 10 min at 6000 rpm and the solution was decanted to remove non-assembled peptide molecules. The assembled structures were lyophilized and poured inside a glass capillary 0.5 mm in diameter. X-ray diffraction was collected using a Bruker D8 Discover theta/theta diffractometer with liquid-nitrogen-cooled intrinsic Ge solid-state linear position detector.

Crystal preparation and data collection. Crystals used for data collection were grown using the vapor diffusion method. The dry peptide was first dissolved in water, at a concentration of 5 mg/ml. Then, 50 µL was deposited into a series of 8x40 mm vessels.

Each tube was sealed with Parafilm®, in which a single small hole was pricked using a needle. The samples were placed inside a larger vessel filled with 2 mL of acetonitrile. The systems were ultimately capped and incubated at 18 °C for several days. Needle-like crystals grew within 7-8 days. For data collection, crystals were coated in paratone oil (Hampton Research), mounted on a MiTeGen cryo-loop and flash frozen in liquid nitrogen. Diffraction data were collected at 100 K on a Rigaku XtaLabPro with a Dectris PilatusR 200K-A detector using CuKα radiation λ= 1.54184Å.

Processing and structural refinement of crystal data. The diffraction data were processed using CrysAlisPro 1.171.39.22a. Structure was solved by direct methods in SHELXT-2016/4.[1] The refinements were performed with SHELXL-2016/4 and weighted full-matrix least-squares against $|F^2|$ using all data. Atoms were refined independently and anisotropically, with the exception of hydrogen atoms, which were placed in calculated positions and refined in a riding mode. Crystal data collection and refinement parameters are shown in Supplementary Table 1 and the complete data can be found in the cif file as supplementary information. The crystallographic data have been deposited in the CCDC with no. 1942277 for AAF.

Ion mobility spectrometry-mass spectrometry (IMS-MS) experiments. To fabricate electrospray emitters, uncoated borosilicate glass capillaries (ID: 1.2 mm OD: 1.5 mm) were purchased from Sutter Instrument Co. (Novato, CA). The 10 cm long capillaries were pulled with a Sutter p-97 micropipette puller to produce electrospray emitters with 1 µm tip sizes. The tripeptide solutions (each prepared at 10 mM in Milli-Q water at pH 6.8) were inserted into the back of a pulled emitter and a 0.25 mm platinum wire was

inserted into the solution. An ESI potential between 1-2 kV was connected to the platinum wire to generate electrospray ions. A custom-made 4-meter drift tube coupled to a time-of-flight (ToF) mass spectrometer was used for IMS-MS analysis. The instrument is shown in **Figure S10**. A detailed description of IMS theory and instrumentation have been detailed previously.[2,3] Briefly, ions produced by ESI enter the IMS-MS instrument through a narrow capillary and are stored in an hourglass-shaped ion funnel (F1) until being pulsed into the drift region by an electrostatic gate (G1). The ion packet then traverses the drift region which is filled with a neutral buffer gas (~3 Torr He) and has a constant electric field drop (~12 $V \cdot cm^{-1}$). After every meter of separation, the diffuse ion packet is radially focused by ion funnels with applied RF potentials (F2/F3/F4/F5). Ions then exit the drift tube through a differentially pumped region and are pulsed into an orthogonal reflection-geometry ToF-MS where they are separated by *m*/*z*.

NMR spectroscopy. Samples were prepared by dissolving lyophilized peptides in 90% H_2O/10% ^2H_2O at 10 mM concentration. AAF and FAA were also prepared at 40 mM concentration. All the NMR experiments were performed on a Bruker 500 MHz NMR spectrometer equipped with a triple-resonance TXI (5 mm with xyz gradient) probe operated at 25 °C.[4] One-dimensional (1D) ^1H and two-dimensional (2D) ^1H-^1H TOCSY and 2D ^1H-^1H NOESY were recorded for ^1H chemical shift assignments. [^{13}C-^1H]-HSQC and [^{15}N-^1H]-HSQC spectra were recorded for ^{13}C and ^{15}N chemical shift assignments, respectively. 1D ^1H NMR spectra were also recorded at different temperatures ranging from 5 to 30 °C. All NMR data were processed in Bruker TopSpin (4.0.6) and analyzed using the Bruker TopSpin/CcpNmr analysis software. DOSY experiments were recorded using a simulated echo sequence with a 3-9-19 pulse sequence for water suppression. The

DOSY data were processed/analyzed in phased mode using T_1/T_2 analysis module in Bruker TopSpin.

All-atom molecular dynamic simulations (All-atom MD)

All-atom MD simulations of 60 AAF/FAA/AFA systems were carried out in an isothermal-isobaric (NPT) ensemble using the GROMACS-2016.4 software package[5] in combination with OPLS-AA force field.[6] 60 AAF/FAA/AFA molecules were randomly placed in an 8×8×8 nm^3 cubic box filled with 15796 TIP4P water molecules,[7] as the initial state of each simulated system. Na^+ and Cl^- ions were also added to the simulation boxes. The bond lengths of peptides and water molecules were constrained using the LINCS[8] and SETTLE[9] algorithms, respectively, allowing an integration time step of 2 fs. The peptide and non-peptide (water molecules and counterions) groups were separately coupled to an external heat bath using a velocity rescaling coupling method,[10] maintaining the temperature at 310 K. The pressure was kept at 1 bar using the Parrinello-Rahman method.[11] Electrostatic interactions were calculated using the particle mesh Ewald (PME) method with a real space cutoff of 1.4 nm.[12] The van der Waals interactions were calculated using the same cutoff of 1.4 nm. An important consideration in choosing a cutoff for the calculation of electrostatic and van der Waals interactions is a tradeoff between accuracy and computational cost, larger cutoff will improve accuracy, albeit an increased computational cost. In order to improve accuracy, we choose 1.4 nm as the cutoff for both electrostatic and van der Waals interactions in our MD simulations. In fact, a cutoff of 1.4 nm for electrostatic and van der Waals interactions has been used in extensive computational studies on peptides, polymers and chemical compounds.[13-16] Verlet cutoff-scheme was used for neighbor searching. To equilibrate the systems after

energy minimization, simulations were performed in the 100 ps NVT MD run first, followed by the 100 ps NPT MD run. Subsequently, three individual 500 ns MD simulations were carried out for the AAF/FAA/AFA systems.

Coarse-grained molecular dynamic simulations (CG-MD). CG-MD simulations on 720 AAF/FAA/AFA systems were performed using the GROMACS-2018.3 software package[5] in combination with the MARTINI coarse-grained model (version 2.1).[17,18] The mapping from the all-atom model of the AAF/FAA/AFA molecules to the CG model and the interaction types of the CG beads are shown in Figure S18. The AAF/FAA/AFA molecule can be divided into two groups: main chain and side chain (the aromatic ring), which are colored in yellow and blue, respectively (Figure S18). Each AAF/FAA/AFA molecule was represented by six CG beads: three beads for the main chain and three beads for the aromatic side chain. Water molecules were represented using P4 interaction types of beads. In the initial state of three simulated systems, 720 AAF/FAA/AFA molecules were randomly placed in a solution containing 40000 water beads, yielding a peptide concentration of ~65 mg/mL. Electrostatic interactions were calculated using the PME method with a real space cutoff of 1.2 nm,[12] and the same cutoff was used for van der Waals interactions. The solute and solvent were separately coupled to an external heat bath using a velocity rescaling coupling method,[10] and a pressure bath using the Parrinello-Rahman method.[11] After 200 ps NVT MD run and the 800 ps NPT MD run, three microsecond-long (3 μs) MD simulations were performed on 720 AAF/FAA/AFA systems.

Analysis methods for MD simulations

Data analyses were performed using in-house-developed codes and tools implemented in the GROMACS package.

For all-atom MD simulation data, a hydrogen bond (H-bond) was considered to be formed when (i) the distance between N and O was smaller than 0.35 nm and (ii) the angle of N − H ⋯ O (or O − H ⋯ N) was larger than 150°. The SASA fraction of Ala residues and Phe side chains was defined as the percentage of the SASA of Ala or Phe side chain relative to the SASA of all AAF/FAA/AFA molecules at each time point. Residue-pair contact probabilities were used to estimate the inter-peptide interactions. The angle between two aromatic rings refers to the angle between the normal vectors of the two rings. If the angle was larger than 90°, the supplementary angle was used as the angle between the two aromatic rings. Two aromatic rings were considered to form π-π stackings when their centroid distance is within 0.7 nm.[19] The 2D free energy landscape was constructed using the formula –RTln[P(angle, centroid distance)], where P (angle, centroid distance) is the probability of a stacking pattern to have a certain value of angle and centroid distance. Probability of the parallel aromatic stacking pattern was calculated using a distance cutoff of 0.5 nm.

The data in the last 300 ns of all-atom MD trajectories were used to calculate the probability density function, the free energy landscape and the probability of the parallel stacking pattern. Trajectory visualization and graphical structure analysis were performed using the VMD[20] and PyMOL[21] software suite.

2. Critical soluble concentration of tripeptides

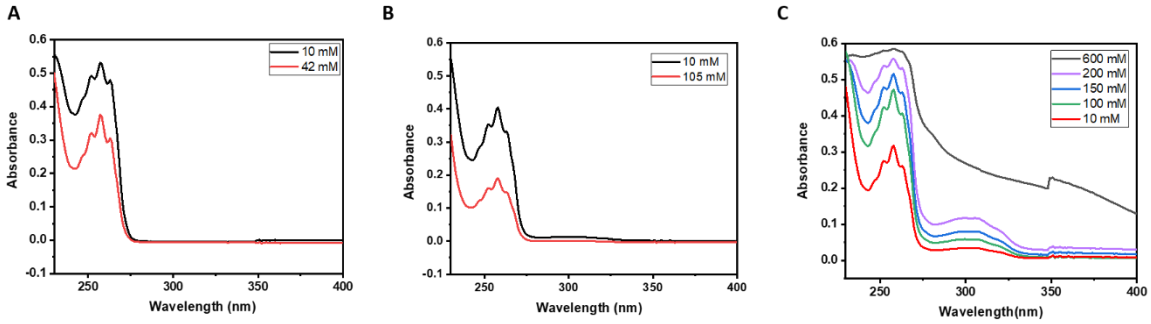


Figure S1: UV/Vis characterization of tripeptides. UV/Vis spectra of (**A**) FAA, (**B**) AAF and (**C**) AFA at different concentrations showing increasing absorbance intensity with rise of concentration without any shift of peak position.

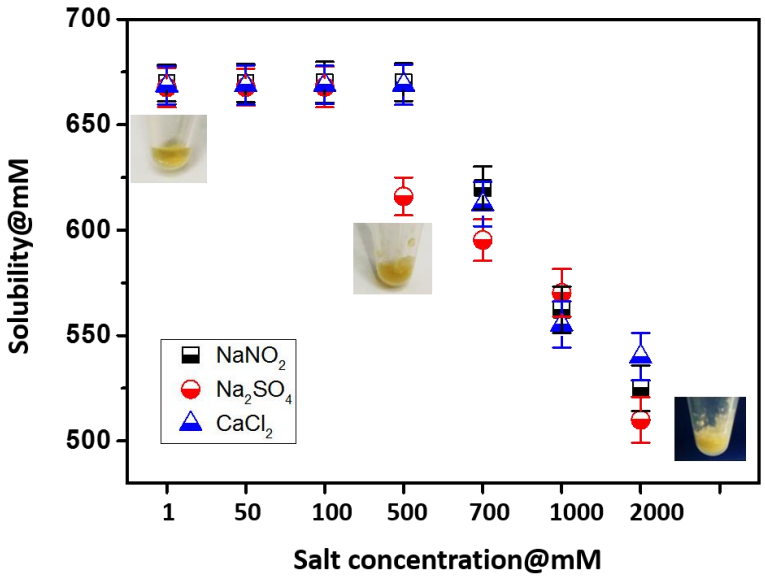


Figure S2: Critical soluble concentration of AFA in the presence salts. Solubility was measured in presence of different concentrations of various salts such as $NaNO_2$, Na_2SO_4 and $CaCl_2$. The effect of salts was quite similar even though changing the cation as well as anion. Solubility decreased only at very high concentration of salt (nearly 500 mM or above) due to commonly known "salting out" effect.

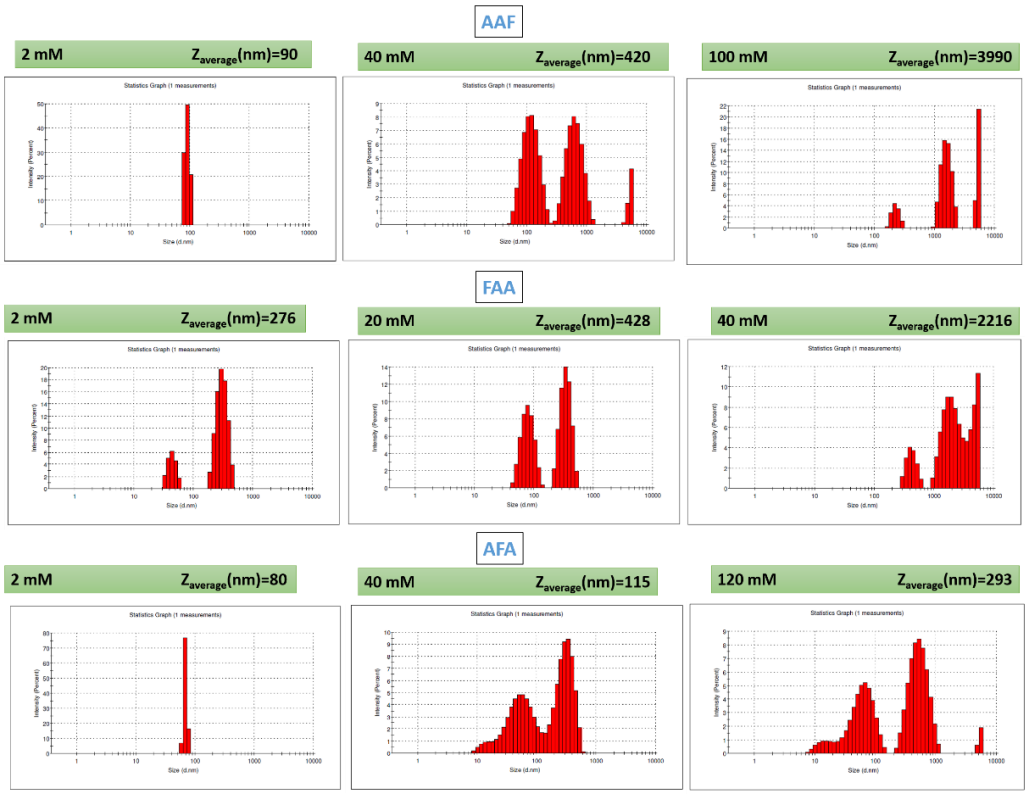


Figure S3: DLS characterization of tripeptides. The results showing the average hydrodynamic diameter of assembled structures at different concentrations.

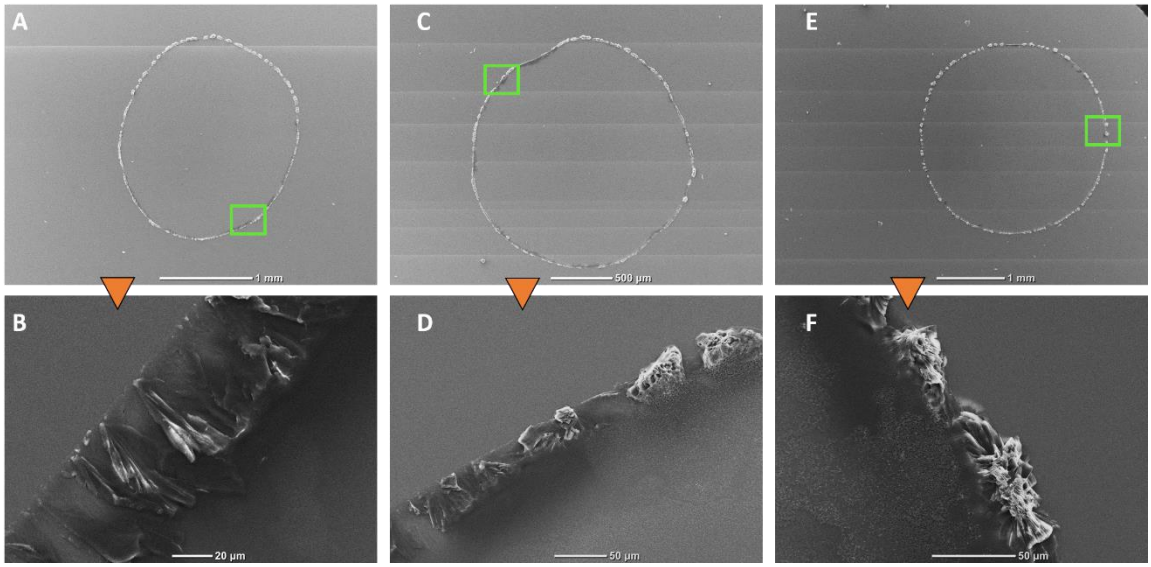


Figure S4: SEM images of AFA. The micrographs showing the resultant architectures formed by 2 mM AFA at pH 6.8 after two weeks of self-assembly. The upper panel represents the entire area of the solution drop and the corresponding zoom-in area marked by a green square is shown in the lower panel.

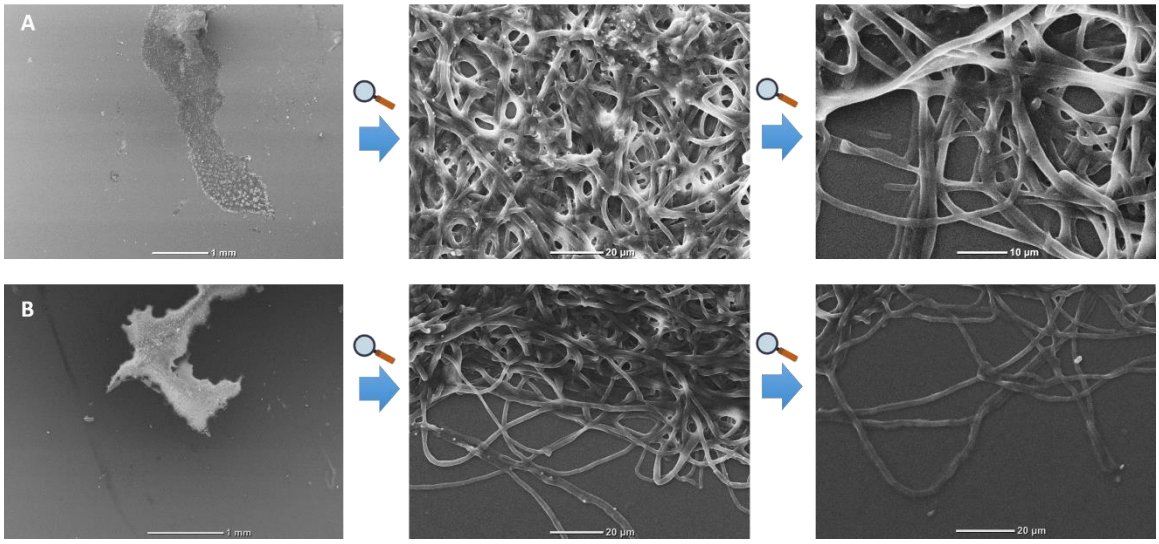


Figure S5: SEM images of AAF and FAA. The micrographs of (**A**) AAF and (**B**) FAA were recorded at 2 mM concentration and pH=6.8.

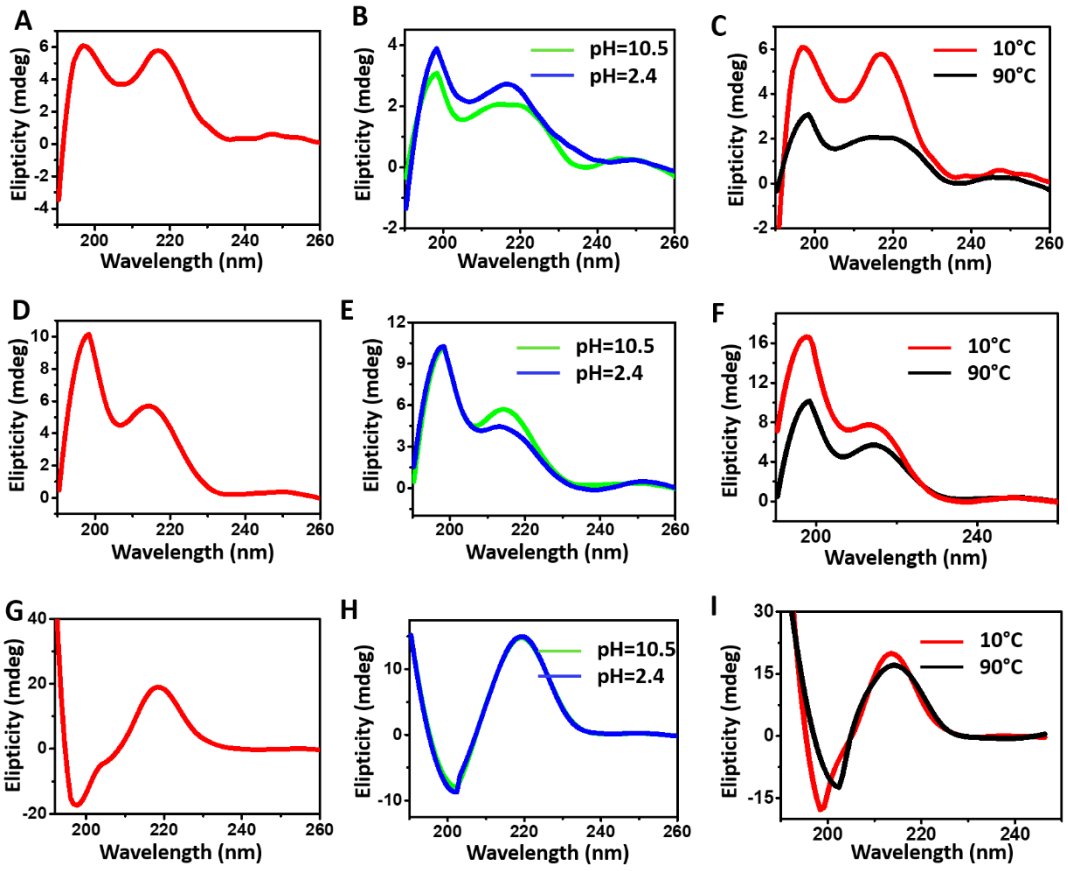


Figure S6: CD spectroscopy of the tripeptides. The spectra of (**A**-**C**) AFA, (**D**-**F**) AAF and (**G**-**I**) FAA were measured under different conditions. (**A**,**D**,**G**) The solution secondary structure at 2 mM concentration and pH 6.8. (**B**,**E**,**H**) At high (pH=10.5) and low (pH=2.4) pH values, the CD spectra did not exhibit any marked change, indicating preservation of the secondary structure. (**C**,**F**,**I**) Upon raising the temperature from 10°C to 90°C, all three peptides showed partial unfolding as the intensity of the major peaks decreased slightly.

3. X-ray diffraction of tripeptides

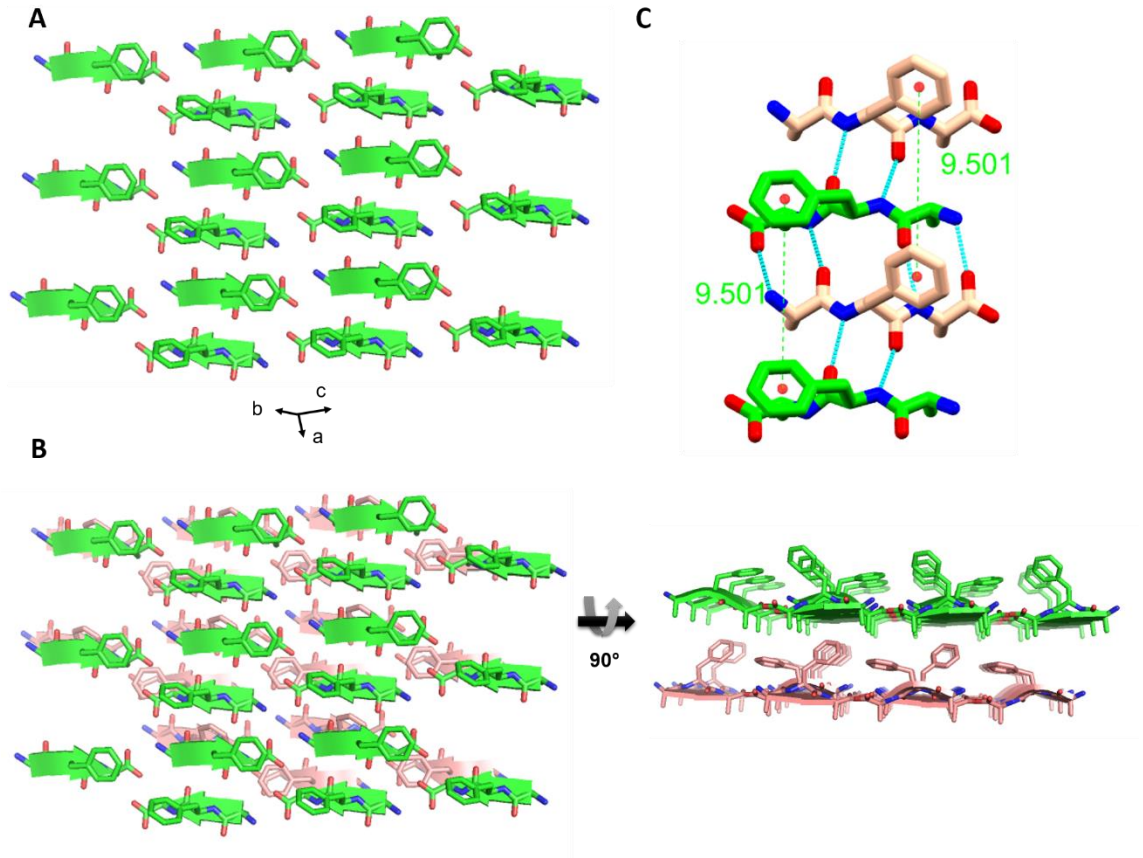


Figure S7: Single crystal structure of AFA. (**A**) Single sheet. (**B**) Hydrophobic interaction between two adjacent sheets. (C) The closest possible interacting distance between two aromatic rings of F is 9.5 Å.

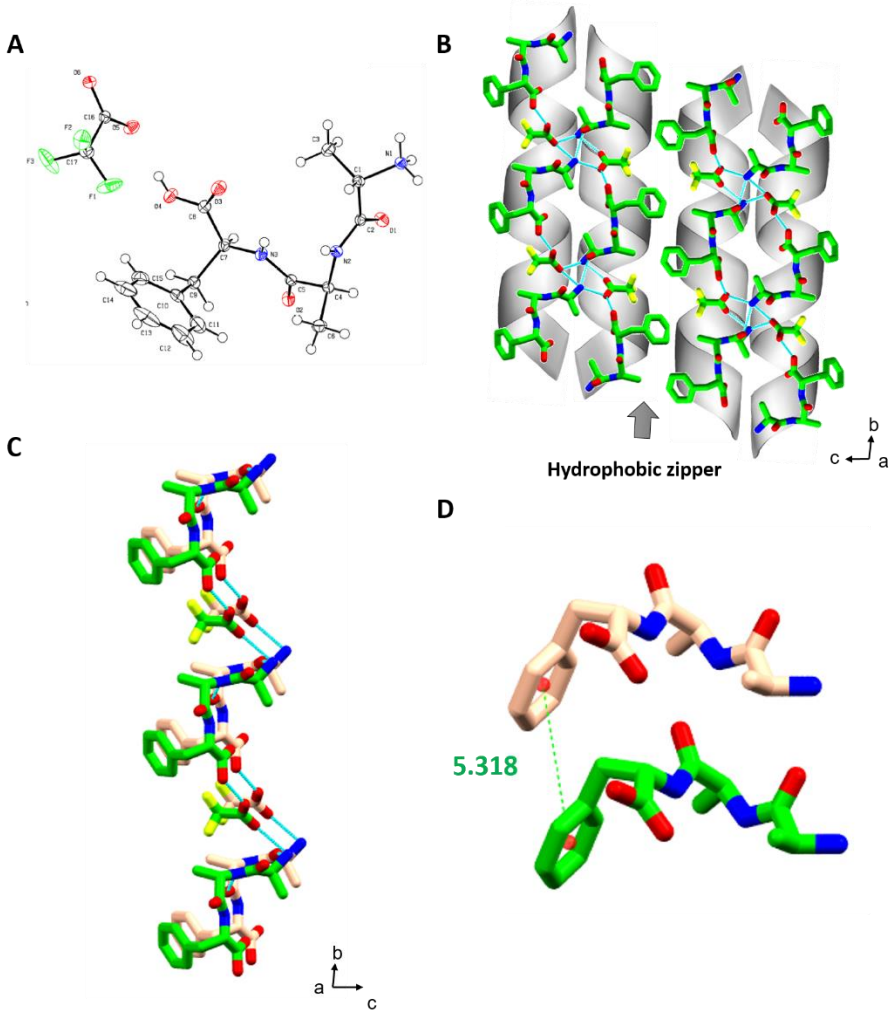


Figure S8: Single crystal structure of AAF. (**A**) The ORTEP diagram of the asymmetric unit in 50% probability displacement ellipsoids. (**B**) Stabilization of nearby helical dimers through hydrophobic zipper-like interactions of aromatic rings. (**C**) Stacking of adjacent helixes in the crystallographic *a*-direction. (**D**) The closest distance of π-π stacking between aromatic rings.

Table S1: Data collection and refinement statistics

Experimental details:

Crystal data	AAF
Chemical formula	C17 H22 F3 N3 O6
Mr	421.37
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
a (Å)	5.31786(5)
b (Å)	12.0314(1)
c (Å)	31.7122(3)
α (°)	90
β (°)	90
γ (°)	90
V ($Å_3$)	2028.99(3)
Z , *Z*'	4
μ (mm^{-1})	1.061
Temperature (K)	100 (2)
Data collection	
Diffractometer	Rigaku XtaLAB AFC12 (RINC): Kappa dual offset/far
Wavelength (Å)	1.54184
Crystal size (mm)	0.262x0.032x0.025
*T*min , *T*max	0.960,0.974
$N_{measured}$(unique)	17797(4384)
$N_{observed}$ [$I > 2\sigma(I)$]	4237
R_{int}	0.0406
θ_{max} (°)	80.155
Refinement	
R[F2 > 2σ(F2)]	0.0349
wR	0.0361
wR[*F*2 > 2σ(*F*2)]	0.0879
wR(*F*2)	0.0892
Goodness-of-fit	1.037
No. of reflections	4384
No. of parameters	266
No. of restraints	0
H-atom treatment	H-atom parameters constrained

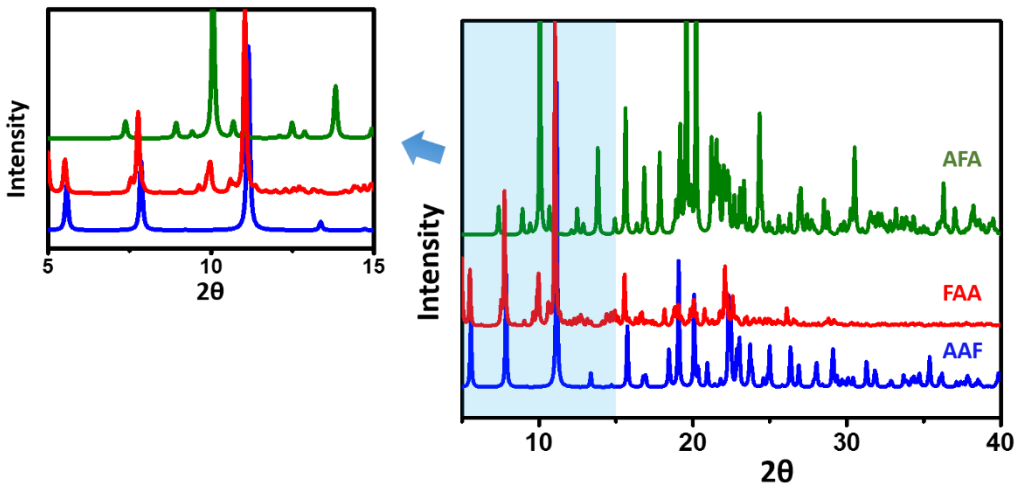


Figure S9: Powder X-ray diffraction of the tripeptides. The region in the graph highlighted by cyan has been enlarged in the left side. The AAF and FAA show quite similar diffraction pattern indicating their similar molecular arrangement in the atomic level and higher order packing.

4. Aggregation characterization using ESI–IMS–MS

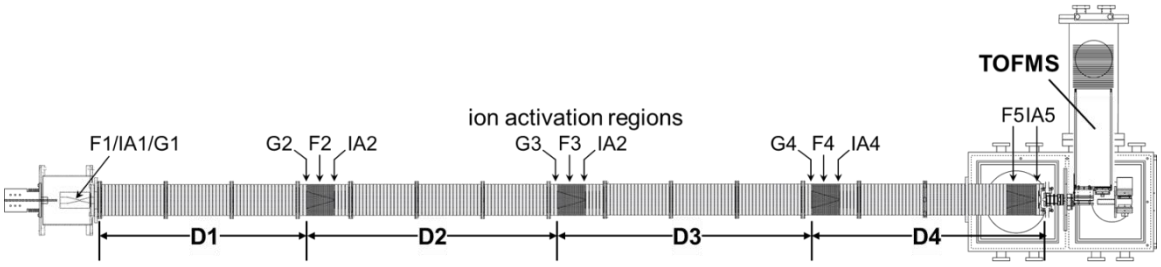


Figure S10: Schematic illustration of the 4m IMS-MS instrument used for tripeptide experiments.

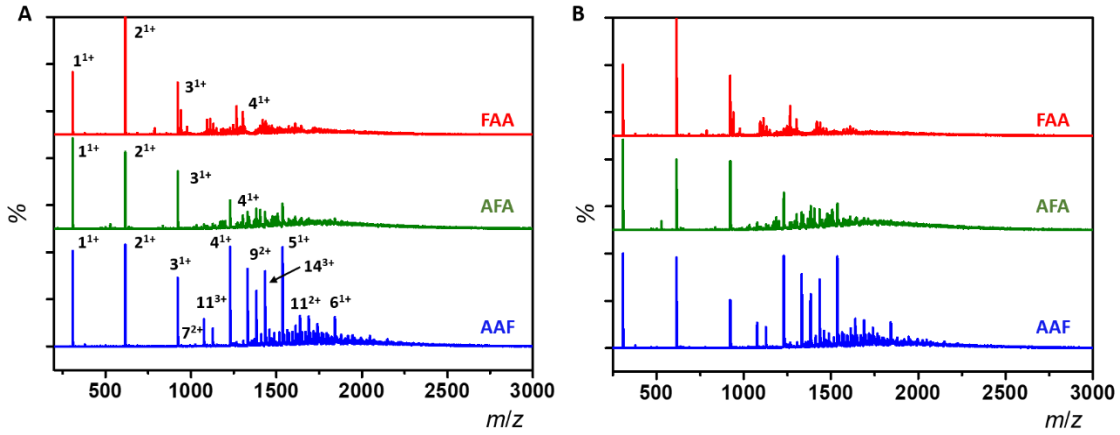


Figure S11: ESI-MS mass spectrum of the tripeptides. The spectrum were recorded at 10 mM peptide concentration and at two different time point after preparation of samples, after 1 day (**A**) and after two weeks (**B**). The numbers above the peaks denote the oligomer order, with the positive-charge state of ions in superscript. The two spectra are showing quite similar peak position and intensity, indicating presence of similar oligomers.

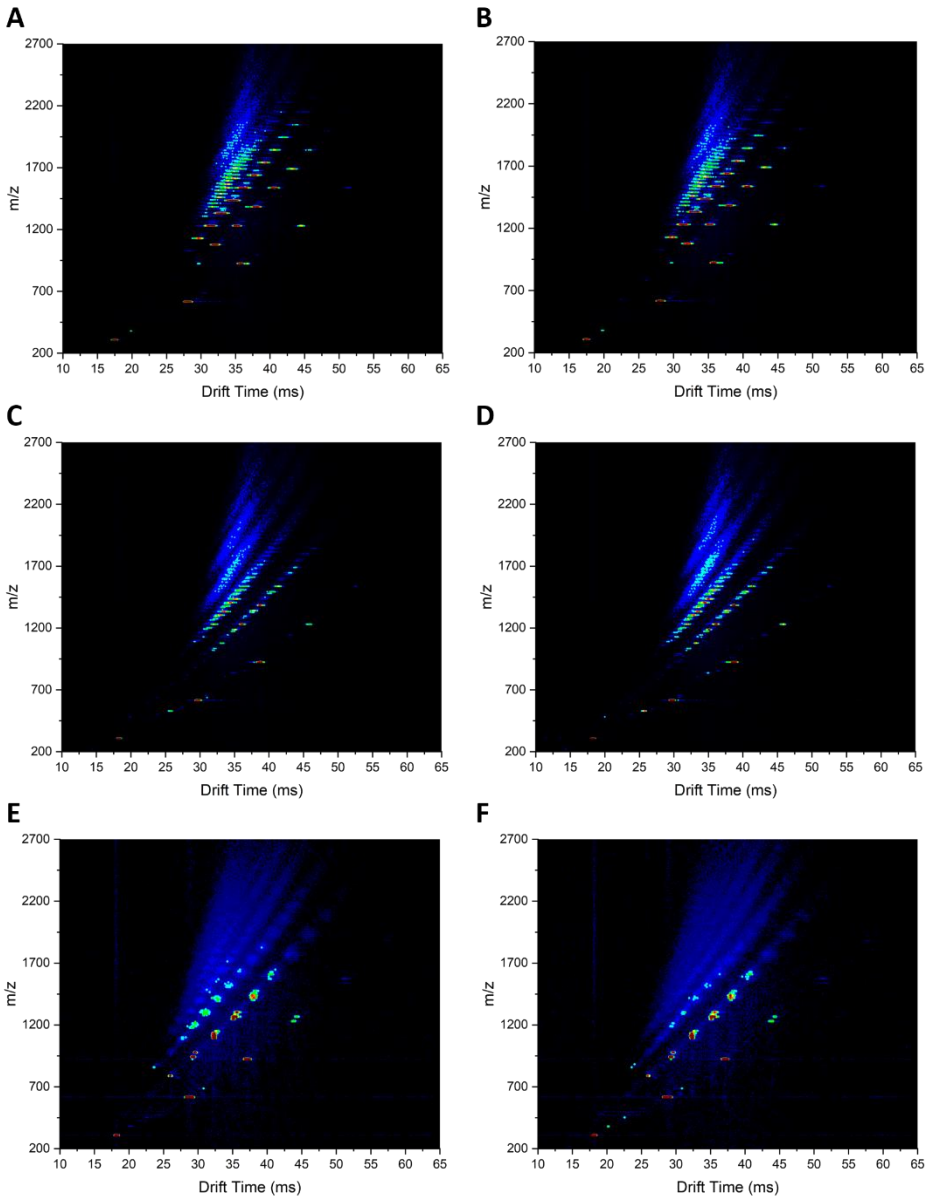


Figure S12: ESI–IMS–MS 2-dimensional plot of the tripeptides. The images of (**A**,**B**) AAF, (**C**,**D**) AFA and (**E**,**F**) FAA monomers through to oligomers. The experiments were carried out at 10 mM peptide concentration and at two different time points after preparation of samples, after 1 days (**A**,**C**,**E**) and after two weeks (**B**,**D**,**E**). However, the 2-dimensional plot showed quite a similar pattern during this time scale.

Table S2: Self-assembled clusters of AAF obtained from ESI–IMS–MS experiment.

Family	symbol	m/z	N	Z	Assembly
+1	1	308	1	1	$[1(AAF)+H]^{+}$
	2	615	2	1	$[2(AAF)+H]^{+}$
	3	922	3	1	$[3(AAF)+H]^{+}$
	4	1229	4	1	$[4(AAF)+H]^{+}$
	5	1536	5	1	$[5(AAF)+H]^{+}$
	6	1843	6	1	$[6(AAF)+H]^{+}$
+2	6	922	6	2	$[6(AAF)+2H]^{2+}$
	7	1076	7	2	$[7(AAF)+2H]^{2+}$
	8	1229	8	2	$[8(AAF)+2H]^{2+}$
	9	1383	9	2	$[9(AAF)+2H]^{2+}$
	10	1536	10	2	$[10(AAF)+2H]^{2+}$
	11	1690	11	2	$[11(AAF)+2H]^{2+}$
	12	1843	12	2	$[12(AAF)+2H]^{2+}$
	13	1997	13	2	$[13(AAF)+2H]^{2+}$
+3	10	1024	10	3	$[10(AAF)+3H]^{3+}$
	11	1127	11	3	$[11(AAF)+3H]^{3+}$
	12	1229	12	3	$[12(AAF)+3H]^{3+}$
	13	1331	13	3	$[13(AAF)+3H]^{3+}$
	14	1434	14	3	$[14(AAF)+3H]^{3+}$
	15	1536	15	3	$[15(AAF)+3H]^{3+}$
	16	1638	16	3	$[16(AAF)+3H]^{3+}$
	17	1741	17	3	$[17(AAF)+3H]^{3+}$
	18	1843	18	3	$[18(AAF)+3H]^{3+}$
	19	1945	19	3	$[19(AAF)+3H]^{3+}$
	20	2048	20	3	$[20(AAF)+3H]^{3+}$
	21	2150	21	3	$[21(AAF)+3H]^{3+}$
	17	1306	17	4	$[17(AAF)+4H]^{4+}$

+4	18	1383	18	4	$[18(AAF)+4H]^{4+}$
	19	1459	19	4	$[19(AAF)+4H]^{4+}$
	20	1536	20	4	$[20(AAF)+4H]^{4+}$
	21	1613	21	4	$[21(AAF)+4H]^{4+}$
	22	1690	22	4	$[22(AAF)+4H]^{4+}$
	23	1766	23	4	$[23(AAF)+4H]^{4+}$
	24	1843	24	4	$[24(AAF)+4H]^{4+}$

Table S3: Self-assembled higher order clusters of AFA obtained from ESI–IMS–MS experiment.

Family	symbol	m/z	N	Z	Assembly
+1	1	308	1	1	$[1(AFA)+H]^{+}$
	2	615	2	1	$[2(AFA)+H]^{+}$
	3	922	3	1	$[3(AFA)+H]^{+}$
	4	1229	4	1	$[4(AFA)+H]^{+}$
	5	1536	5	1	$[5(AFA)+H]^{+}$
+2	7	1076	7	2	$[7(AFA)+2H]^{2+}$
	8	1229	8	2	$[8(AFA)+2H]^{2+}$
	9	1383	9	2	$[9(AFA)+2H]^{2+}$
	10	1536	10	2	$[10(AFA)+2H]^{2+}$
	11	1690	11	2	$[11(AFA)+2H]^{2+}$
	12	1843	12	2	$[12(AFA)+2H]^{2+}$
	13	1997	13	2	$[13(AFA)+2H]^{2+}$
+3	12	1229	12	3	$[12(AFA)+3H]^{3+}$
	13	1331	13	3	$[13(AFA)+3H]^{3+}$
	14	1434	14	3	$[14(AFA)+3H]^{3+}$
	15	1536	15	3	$[15(AFA)+3H]^{3+}$

Table S4: Self-assembled higher order clusters of FAA obtained from ESI–IMS–MS experiment.

Family	symbol	m/z	N	Z	Assembly
+1	1	308	1	1	$[1(FAA)+H]^{+}$
	2	615	2	1	$[2(FAA)+H]^{+}$
	3	922	3	1	$[3(FAA)+H]^{+}$
	4	1229	4	1	$[4(FAA)+H]^{+}$
	5	1536	5	1	$[5(FAA)+H]^{+}$
+2	6	922	6	2	$[6(FAA)+2H]^{2+}$
	6	940	6	2	$[6(FAA)+K]^{2+}$

5. Aggregation analysis by NMR spectroscopy

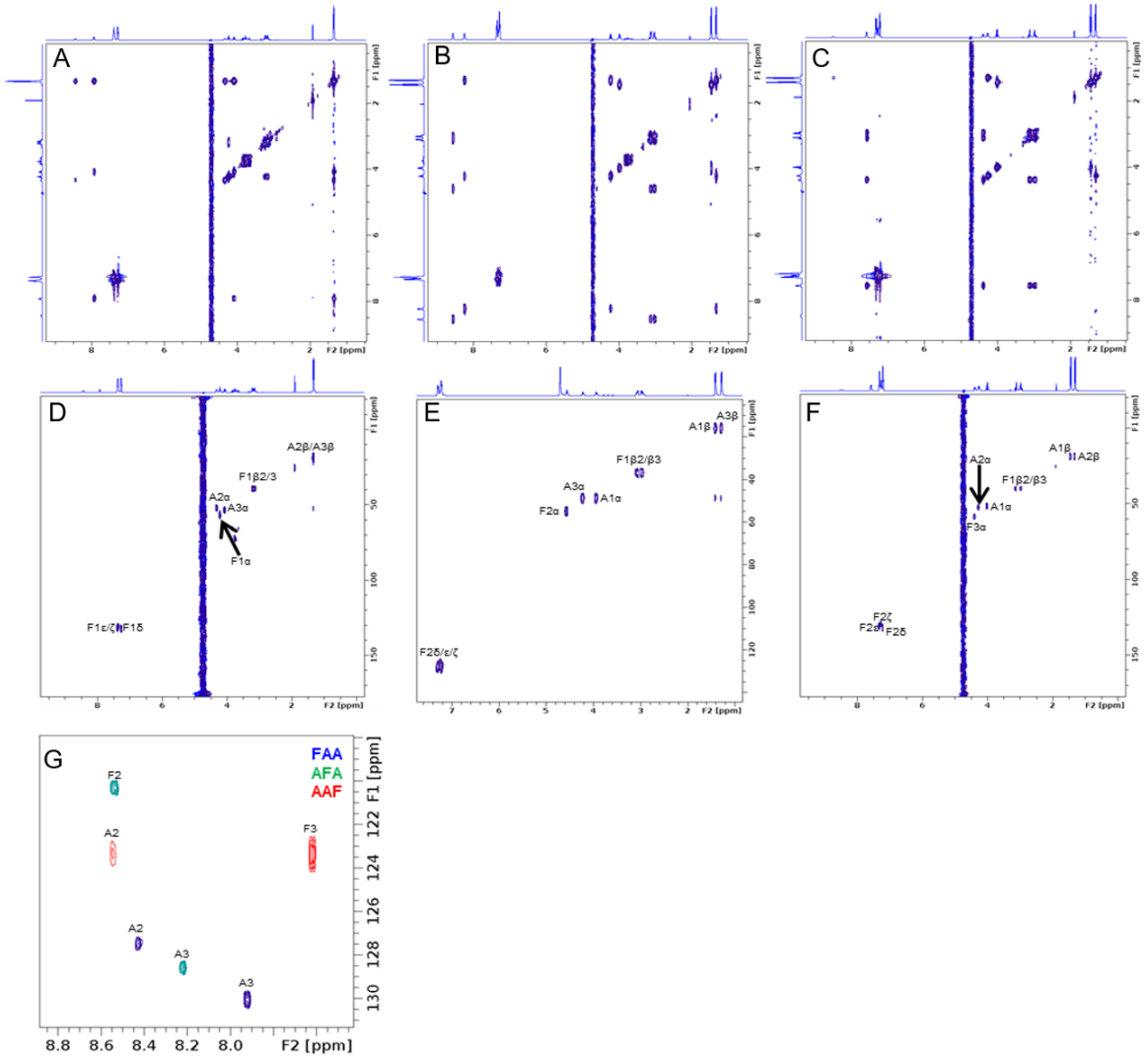


Figure S13. Two-dimensional NMR spectra of tripeptides. The [^{1}H-^{1}H]-TOCSY (**A-C**) and [^{13}C-^{1}H]-HSQC (**D-F**) NMR spectra of FAA (**A,D**), AFA (**B,E**) and AAF (**C,F**) recorded in 100 % $^{2}H_2O$ or 90 % H_2O/ 10 % $^{2}H_2O$. (**G**) An overlay of [^{15}N-^{1}H]-HSQC spectra obtained from all three tripeptides.

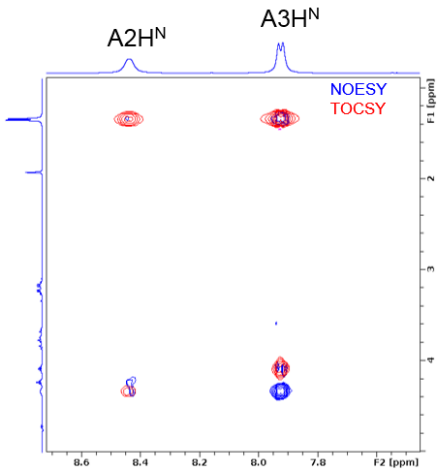


Figure S14. Two-dimensional NMR spectra of FAA. The [^{1}H-^{1}H]-NOESY and [^{1}H-^{1}H]-TOCSY NMR spectra of FAA recorded in 90 % H_2O/ 10 % ^2H_2O.

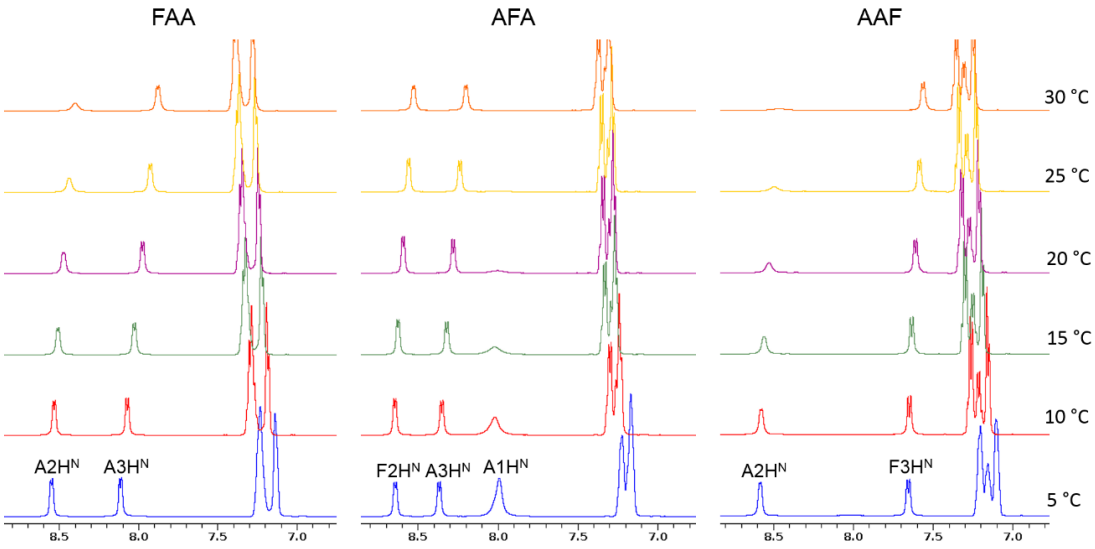


Figure S15. Temperature dependent ^{1}H NMR spectra of the tripeptides. The spectra were recorded in 90 % H_2O/ 10 % ^2H_2O at the indicated temperatures.

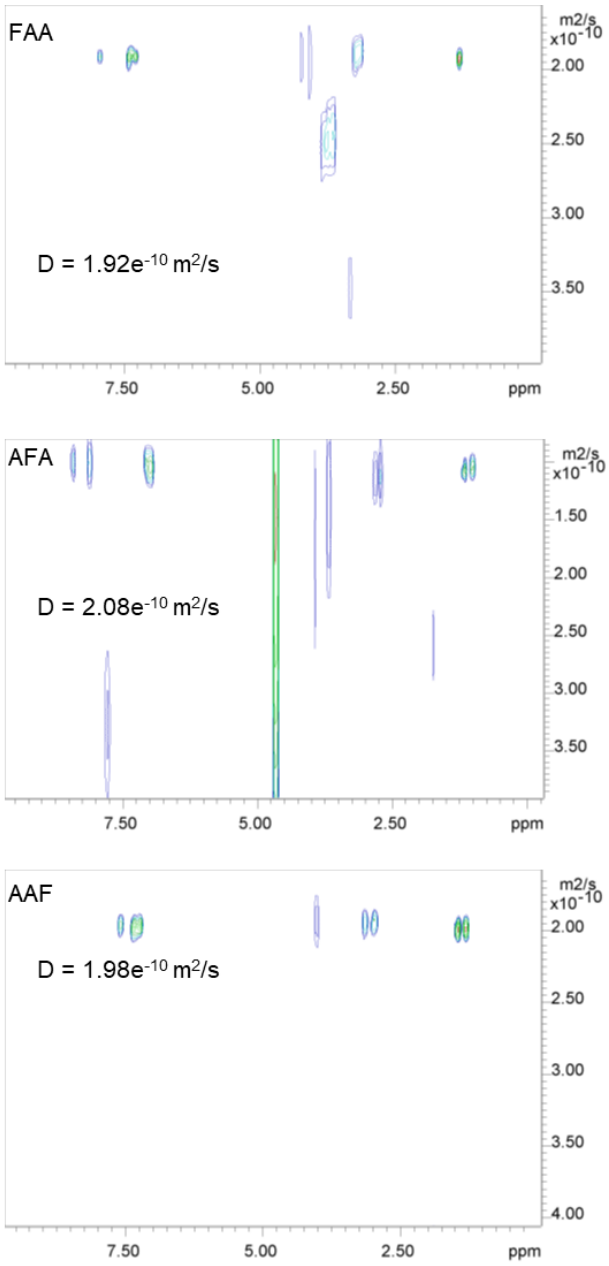


Figure S16. DOSY NMR spectra of tripeptides. The spectra were recorded in 90 % H_2O/ 10 % ^2H_2O. The experimentally determined diffusion coefficients are indicated in the corresponding spectrum for each peptide.

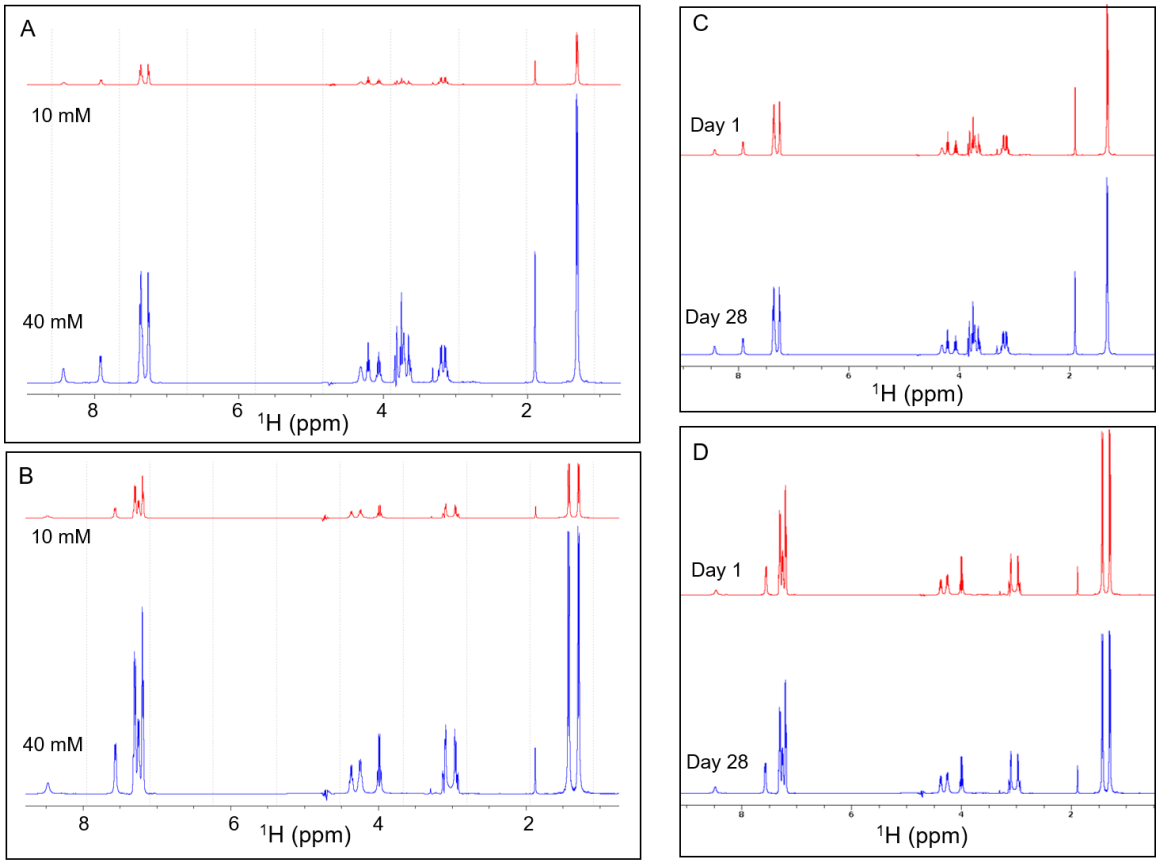


Figure S17. Concentration and time dependent ^{1}H NMR experimental results. (**A,B**) ^{1}H NMR spectra of FAA (**A**) and AAF (**B**) recorded in 90 % H_2O/ 10 % ^2H_2O at 10 mM (red) and 40 mM (blue) concentrations. (**C,D**) Time-dependent ^{1}H NMR spectra of FAA (**C**) and AAF (**D**) recorded in 90 % H_2O/ 10 % D2O at 40 mM concentration.

Table S5. $^{3}JH^{N}\alpha$ values (in Hz) and temperature coefficients for the tripeptides investigates in this study

FAA	**F1**	**A2**	**A3**
$^{3}JH^{N}\alpha$ at 5 °C	-	6.4	6.9
$^{3}JH^{N}\alpha$ at 25 °C	-	-	7.2
Temp. Coefft. (ppb/K)	-	2.8	9.2
AFA	**A1**	**F2**	**A3**
$^{3}JH^{N}\alpha$ at 5 °C	-	6.9	7.4
$^{3}JH^{N}\alpha$ at 25 °C	-	6.7	7.3
Temp. Coefft. (ppb/K)	-	4.8	6.8
AAF	**A1**	**A2**	**F3**
$^{3}JH^{N}\alpha$ at 5 °C	-	5.8	8.6
$^{3}JH^{N}\alpha$ at 25 °C	-	-	8
Temp. Coefft. (ppb/K)	-	4.8	3.6

Table S6. Translational diffusion coefficients and hydrodynamic radii for the tripeptides investigated in this study.

Peptide	**Diffusion coefficient ($e^{-10}m^2/s$)**	**Hydrodynamic radius(nm)**
FAA	1.92	1.28
AFA	2.08	1.18
AAF	1.98	1.24

6. Molecular dynamic simulations

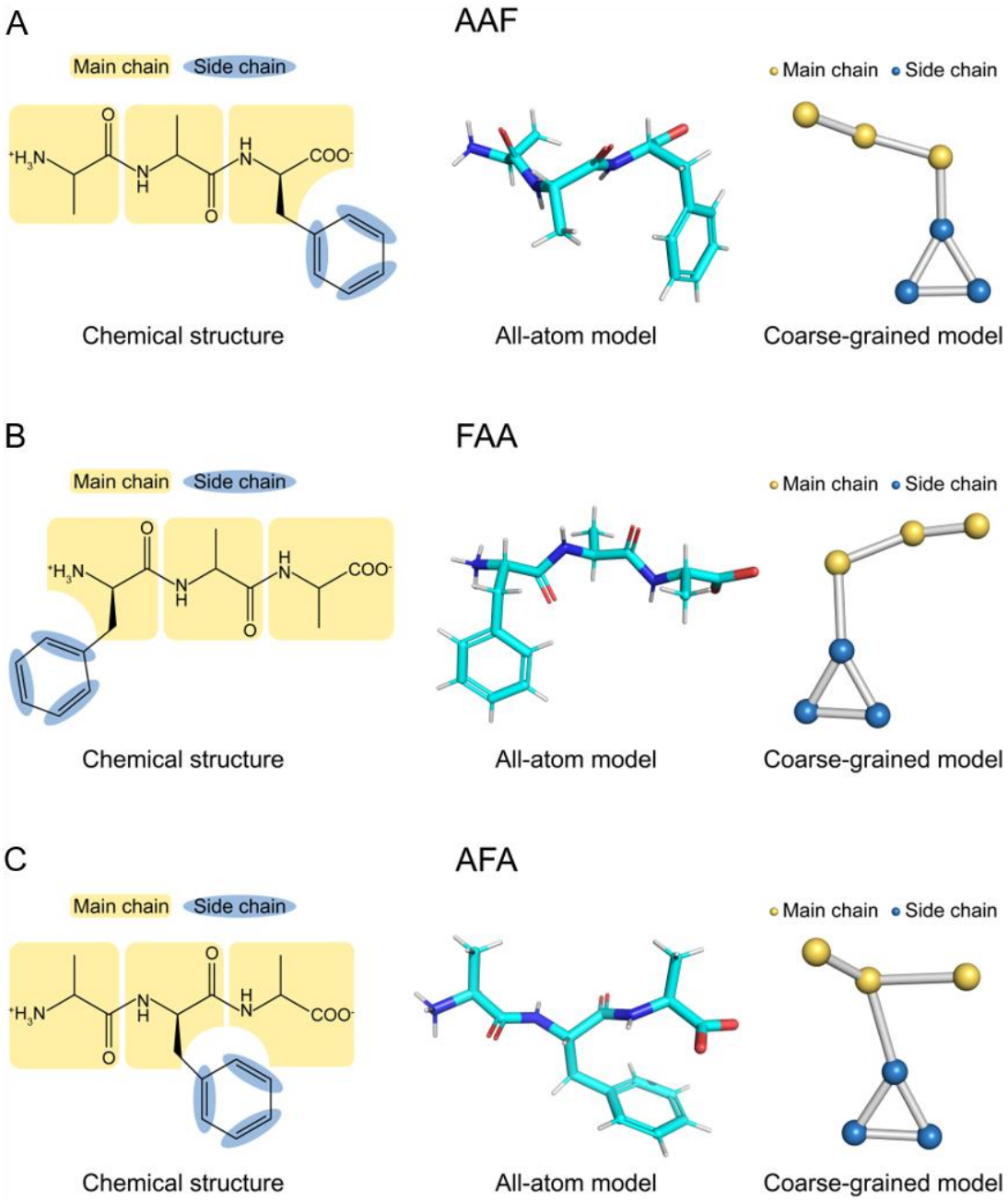


Figure S18: Chemical structure, all-atom model and coarse-grained model of tripeptides. (**A**) AAF, (**B**) FAA and (**C**) AFA molecules.

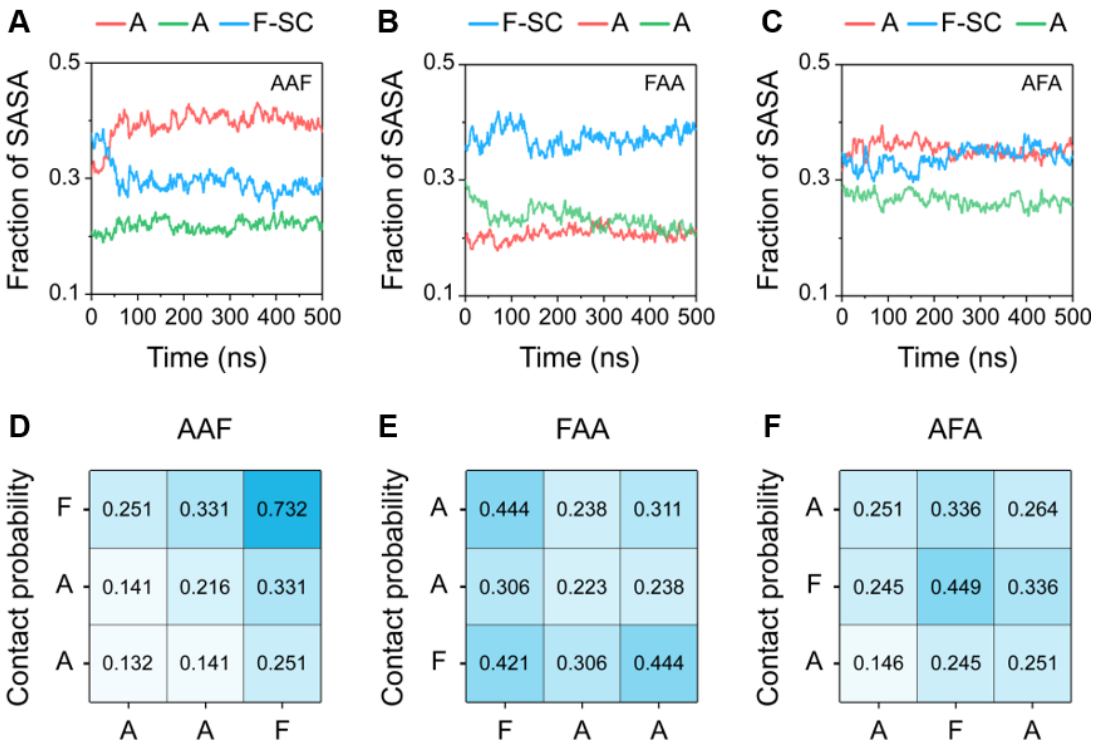


Figure S19: 6. Molecular dynamic simulations of tripeptides. (A-C) Time evolution of the SASA fraction of Ala residues and Phe side chain in AAF, FAA and AFA. (**D-F**) Contact probability for each residue pair of AAF, FAA and AFA.

7. X-ray structure analysis of AFF

A

Torsion angle	ϕ	ψ
Molecule A	-144.20	150.81
Molecule B	-127.95	135.10

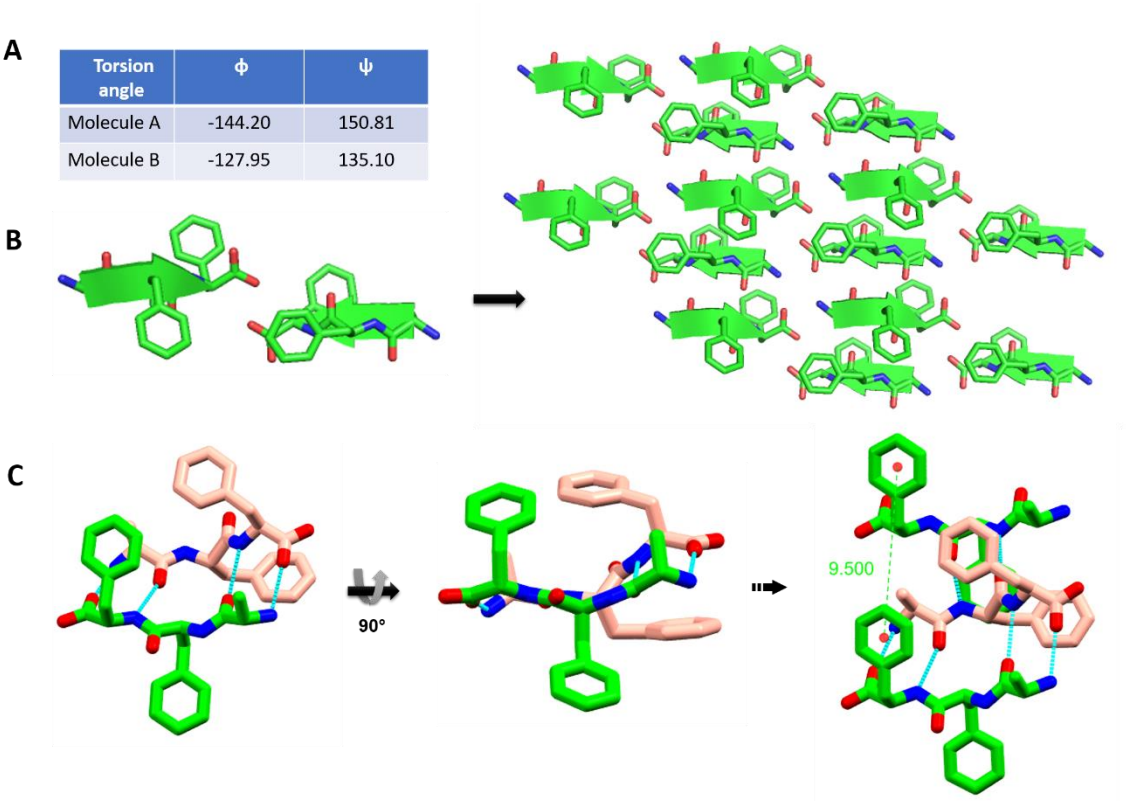


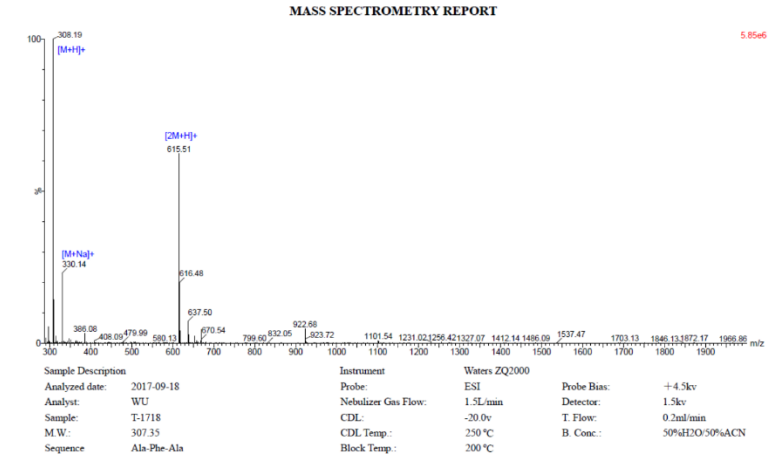
Figure S20: Crystal structure of AFF (CCDC ref no. 1862583[22]). (**A**) Torsion angle of Phe_2 for the two different molecules present in the asymmetric unit. (**B**) The antiparallel β-sheet structure. (**C**) Absence of any π-π interactions in the solid state packing as observed from different directions.

Table S7: Solid-state packing dependent solubility of peptides obtained through screening of peptide library at different length-scale.

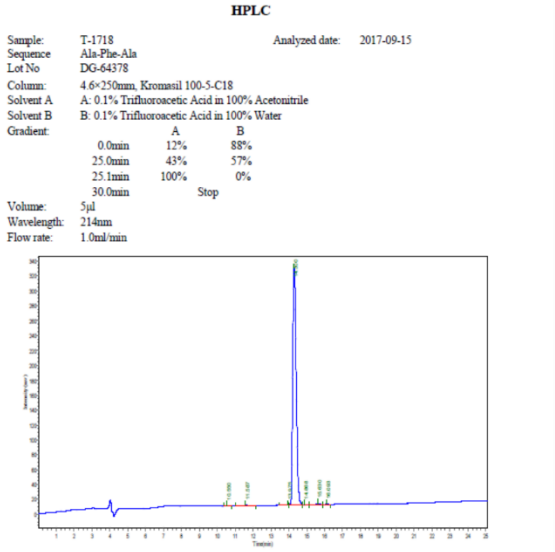
Chain length	Compound	Aromatic interaction exits?	No of aromatic group involve in stacking	Solubility (mM)	Melting point(°C)
Amino Acid	NH_3^+-F-COO^-	Yes	1	125	283
Dipeptide	NH_3^+-FF-COO^-.H_2O	Yes	2	8.1	360
	NH_3^+-FL-COO^-.H_2O	Yes	1	28.3	258
	NH_3^+-LF-COO^-.H_2O	Yes	1	19.4	218
Tripeptide	NH_3^+-FFF-COO^-	Yes	3	0.5	219
	NH_2^+-PFF-COO^-	Yes	2	2.6	297
	NH_3^+-GFF-COO^-.H_2O	Yes	2	14.5	225
	NH_3^+-AFF-COO^- .TFA.H_2O	**No**	-	**98.3 ◄**	243
	NH_3^+-AAF-COO^- .TFA	Yes	1	105	292
	NH_3^+-AFA-COO^- .TFA.AcOH.H_2O	**No**	-	**672 ◄**	257

8. Characterization of tripeptides

A



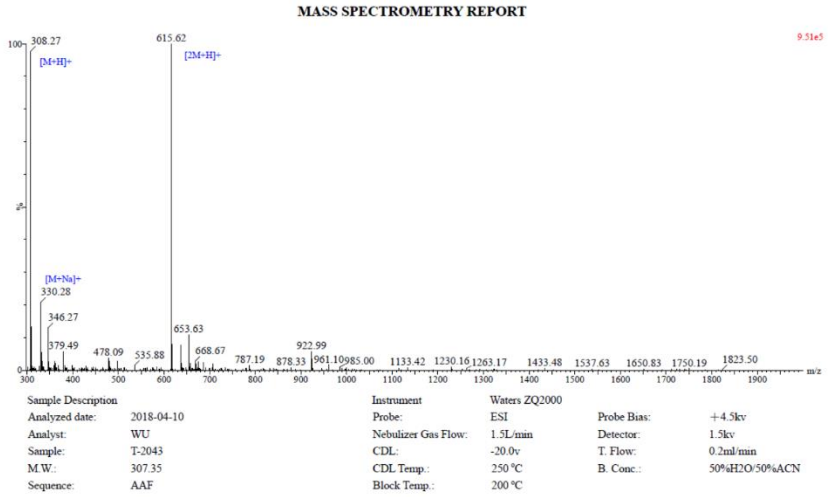
B



Peak	Time	Height	Area	Conc.
1	10.550	676.331	6652.780	0.1523
2	11.587	562.015	17937.611	0.4106
3	13.975	396.370	5749.300	0.1316
4	14.300	318425.719	4306767.500	98.5845
5	14.868	666.012	8328.713	0.1906
6	15.630	1667.626	18634.463	0.4266
7	16.093	402.934	4536.835	0.1039
Total				100.000

Figure S21: Characterization of AFA. (**A**) Mass Spectra, (**B**) HPLC trace.

A



B

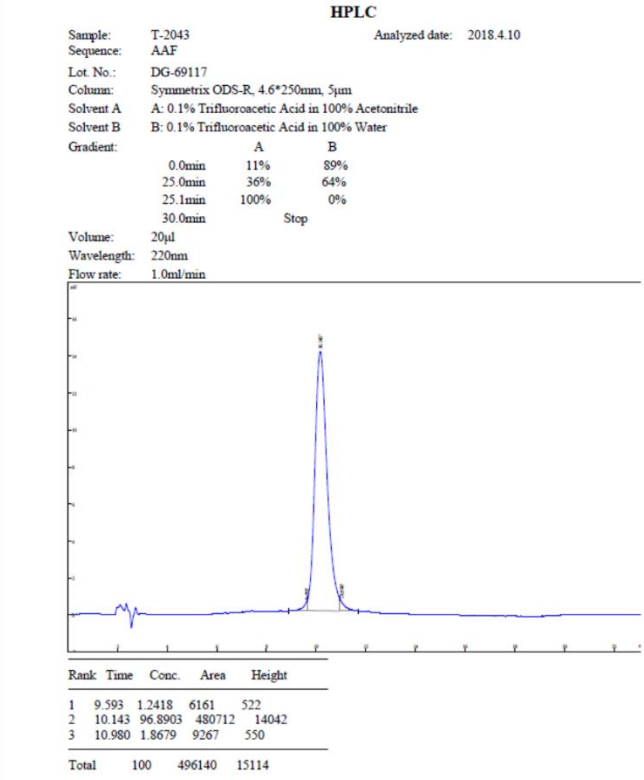
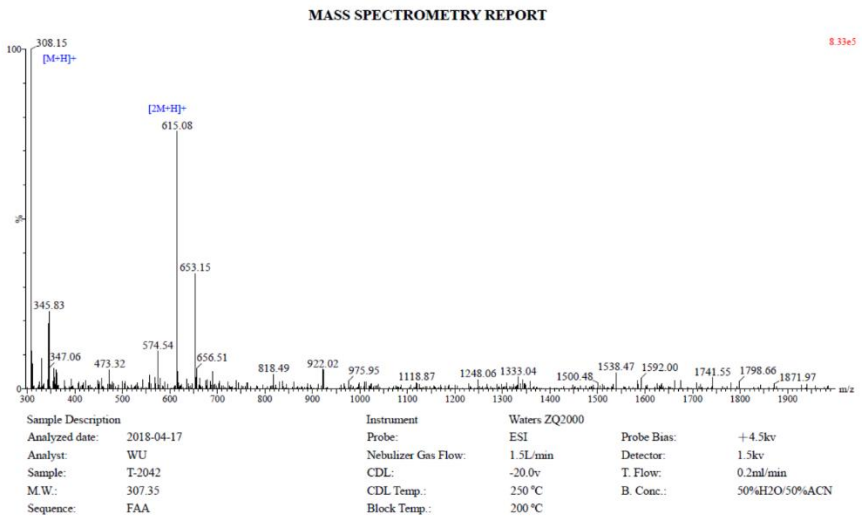


Figure S22: Characterization of AAF. (**A**) Mass Spectra, (**B**) HPLC trace.

A



B

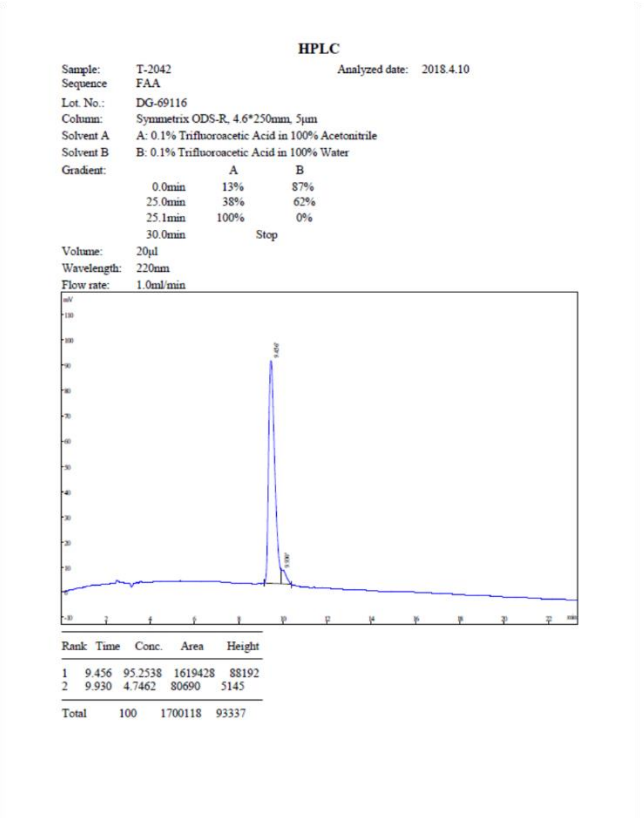


Figure S23: Characterization of FAA. (**A**) Mass Spectra, (**B**) HPLC trace.

Supplemental References

1. Sheldrick, G. (2013). SHELXL-2013. Univ. Göttingen, Göttingen, Ger.

2. Koeniger, S. L. *et al.* (2006). An IMS-IMS analogue of MS-MS. Anal. Chem. *78*, 4161–4174.

3. Merenbloom, S. I., Koeniger, S. L., Valentine, S. J., Plasencia, M. D., Clemmer, D. E. (2006). IMS-IMS and IMS-IMS-IMS/MS for separating peptide and protein fragment ions. Anal. Chem. *78*, 2802–2809.

4. Yao, S., Howlett, G. J., Norton, R. S. (2000). Peptide self-association in aqueous trifluoroethanol monitored by pulsed field gradient NMR diffusion measurements J. Biomol. NMR *16*, 109–119.

5. Abraham, M. J. *et al.* (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX *1–2*, 19–25.

6. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J. Phys. Chem. B *105*, 6474–6487.

7. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

8. Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. J. Comput. Chem. *18*, 1463–

1472.

9. Miyamoto, S., Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. J. Comput. Chem. *13*, 952–962.

10. Bussi, G., Donadio, D., Parrinello, M. (2007). Canonical sampling through velocity rescaling. J. Chem. Phys. *126*, 014101.

11. Parrinello, M., Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. *52*, 7182–7190.

12. Darden, T., York, D., Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

13. Dutta, M. S., Basu, S. (2021). Identifying the Key Residues Instrumental in Imparting Stability to Amyloid Beta Protofibrils – a Comparative Study Using MD Simulations of 17–42 Residues. J. Biomol. Struct. Dyn., 39, 431-456.

14. Lee, H., Ostadhassan, M., Sun, Z., Pu, H., Liu, B., Varma, R. S., Jang, H. W., Shokouhimher, M. (2020). RSC Adv. *10*, 37938.

15. Liu, J., Xu, Q., Jiang, J. (2019). A Molecular Simulation Protocol for Swelling and Organic Solvent Nanofiltration of Polymer Membranes. J. Membr. Sci., *573*, 639-646.

16. Kamat, S., Lin, R., Chiew, Y. C. (2019). Structure and Properties of Dicarboxylic Acids at Hexane/Water Interface: A Molecular Dynamics Study. Colloids Surf. A, *580*, 123725.

17. Monticelli, L. *et al.* (2008). The MARTINI coarse-grained force field: Extension

to proteins. J. Chem. Theory Comput. *4*, 819–834.

18. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., De Vries, A. H. (2007). The MARTINI force field: Coarse grained model for biomolecular simulations. J. Phys. Chem. B *111*, 7812–7824.

19. Burley, S. K., Petsko, G. A. (1985). Aromatic-aromatic interaction: A mechanism of protein structure stabilization. Science *229*, 23–28.

20. Humphrey, W., Dalke, A., Schulten, K. (1996). Sartorius products. J. Mol. Graph. *14*, 33–38.

21. Schrödinger, L. (2010). The PyMOL Molecular Graphics System, Version 1.3r1.

22. Bera, S., Mondal, S., Xue, B., Shimon, L. J. W., Cao, Y., Gazit, E. (2019). Rigid Helical-like Assemblies from a Self-Aggregating Tripeptide. Nat. Mater. *18*, 503–509.