

**Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels**

Pim W. J. M. Frederix,<sup>†‡</sup> Gary G. Scott,<sup>†</sup> Yousef M. Abul-Haija,<sup>†</sup> Daniela Kalafatovic,<sup>†</sup> Charalampos G. Pappas,  
Nadeem Javid, Neil T. Hunt,<sup>‡</sup> Rein V. Ulijn,<sup>†</sup> and Tell Tuttle<sup>†\*</sup>

**Experimental methods*****Coarse grain molecular dynamics***

Tripeptide coordinate files were created using VMD scripting tools<sup>1</sup> and converted to CG representation in the MARTINI force field (version 2.2) using martinize.py.<sup>2</sup> As the secondary structure needs to be defined in the force field, the flag `-ss = EEE` was used, (`E` = extended  $\beta$ -sheet) leading to Qa, Nda and Qd beads for the backbone particles. This choice was made as  $\beta$ -sheet-like conformations are often observed in peptide nanostructures (see e.g. ref.<sup>3</sup>), but it should be noted that for the case of tripeptides this is not expected to bias the conclusions as the same choice was made for all peptides. An alternative choice would be to use random coil (`R` or `C`) parameters, but in MARTINI v. 2.2, the relevant differences between `E` and `R` bonded interactions are small: a slightly elevated backbone-backbone-backbone angle and force constant (134° and 25 kJ/mol vs 127° and 20 kJ/mol, respectively), and a different particle type (Nda for `E`, P5 for `R`).<sup>11</sup> The latter leads to an attractive interaction of the middle amino acid backbone with other particle types that is on average  $\sim$ 0.5 kJ/mol lower. Although these changes could impact the aggregation rate and the final state, the objective of the current study is to rapidly reduce a large number of potential tripeptides into a select few that are able to form hydrogelators and as such we believe our choice of the secondary structure flag based on the likely final state is appropriate.

Using the GROMACS code version 4.5.3,<sup>4</sup> a cubic box of  $13 \times 13 \times 13$  nm<sup>3</sup> containing 300 zwitterionic tripeptides was created giving a peptide concentration of 0.23 mol L<sup>-1</sup> in standard CG water, with side chains in their most prevalent charge state at pH 7. Note that the computational concentration is one order of magnitude higher than the experimental one to accelerate the assembly process and thus decrease computational cost. Periodic boundary conditions were used. LJ interactions were shifted to zero in the range 0.9–1.2 nm, and electrostatic interactions in the range 0.0–1.2 nm for all simulations (no Particle Mesh Ewald method was used). A relative dielectric constant  $\epsilon_r = 15$  was used in standard CG water simulations for screening of the electrostatic interactions, while 2.5 was used for simulations in polarizable water. The box was energy minimized for 5000 steps or until forces on atoms converged to under 200 pN. The minimized box was subsequently equilibrated for 500,000 steps of 25 fs, using the Berendsen algorithms<sup>5</sup> to keep temperature ( $\tau_T = 1$  ps) and pressure ( $\tau_P = 3$  ps) around 303 K and 1 bar, respectively. Bond lengths in aromatic side chains and the backbone-side chain bonds in I, V and Y were constrained using the LINCS algorithm.<sup>6</sup> The total simulation for this initial screening phase equates 12.5 ns, but this equates to roughly 50 ns ‘effective time’, due to the smoothness of the CG potentials.<sup>7,8</sup> All times reported in this paper take into account this speedup factor. For the tripeptides selected for further study, the water in the solvated energy-minimized box was converted to polarizable water (PW)<sup>9</sup> to better account for charge screening. This system was then energy-minimized again and run in the NPT ensemble for  $4 \cdot 10^6$  steps, or 400 ns effective time. The AP was calculated at both 50 and 400 ns, analogous to the initial screening phase (see main text). In PW, 50 ns appears an appropriate simulation length and we extrapolate that timescale to normal CG water, but the data is not entirely conclusive as there are differences in AP score at the 50 ns and 400 ns time marks for a small number of peptides in the PW simulations (see Supplementary Table 1). Finally, for the systems with experimental information available, a similar, but larger simulation was carried out for 1200 ns using 1200 peptides in a box of  $24 \times 24 \times 24$  nm<sup>3</sup> (peptide concentration 0.14 M in standard CG water) to allow the formation of larger structural features like tubes, spheres or fibers where appropriate. This simulation was extended further to 10  $\mu$ s for KYF.

For the morphological analysis, the GROMACS tools were employed to determine the largest cluster of molecules, align the cluster along its principle axes and determine the moment of inertia in three dimensions. The relative moments of inertia along x, y and z axes, together with visual inspection, determined the

predominant morphology of the cluster. The rules of thumb for determination of the morphology are  $I_x \approx I_y \approx I_z$  describes a spherical structure (aspect ratio < 2), while  $I_x < I_y \approx I_z$  indicates a fibrous / oblong morphology (visual inspection and aspect ratio < 3 distinguish between the two)). Two molecules belong to the same cluster when any two atoms on the different molecules are closer than 0.5 nm, taking into account periodic boundary conditions.

For KYF, a backmapping procedure was performed using *backward.py* (v. 0.1, accessed 09/07/2014, ref.<sup>10</sup>) The mapping was performed on the result of a 9.6  $\mu$ s simulation on the 300 peptide box using the GROMOS 53a6 force field. The coordinates of the structure were linearly scaled with a factor 1.1 to reduce the effect of close contacts due to the high density of the fibre. The backmapped structure was then relaxed by a sequence of minimizations and equilibrations in the NVT ensemble as suggested in the published protocol<sup>10</sup>, all in the GROMACS 4.6.3 package using periodic boundary conditions:

1. 10,000 step steepest descent minimization with all non-bonded interactions ignored
2. 10,000 step steepest descent minimization with all non-bonded interactions switched on
3. 50,000 step NVT simulation with a time step of 0.2 fs
4. 50,000 step NVT simulation with a time step of 0.5 fs
5. 50,000 step NVT simulation with a time step of 1.0 fs
6. 5,050,000 step NVT simulation with a time step of 2.0 fs
7. Solvation using SPC (single point charge) water, and neutralization using Cl<sup>-</sup> ions
8. 5,000 step steepest descent minimization
9. 5,000,000 step NPT simulation in water with a time step of 2.0 fs for equilibration
10. 25,000,000 step NPT simulation in water with a time step of 2.0 fs for analysis

Step 1-6 were performed in vacuum using LINCS constraints on all bonds and ref\_t set to 200 K with a coupling constant of 0.1 ps. Step 9 was performed in water, using LINCS constraints on all bonds involving hydrogen. Peptides and water + ions were coupled to a temperature bath of 298 K using the V-rescale algorithm ( $\tau_T = 1.0$  ps). Pressure was kept around 1.0 bar ( $\tau_p = 1.5$  ps) using the Berendsen algorithm. The neighbour list was updated every 10 steps. Electrostatics were handled using the reaction field method with  $\epsilon_{rf} = 62$ . The cut-offs for coulomb and vanderwaals forces were set at 1.4 nm, with a cut-off of 0.9 nm for short-range interactions.

#### *Peptide synthesis*

GGG was purchased from Sigma Aldrich. All other tripeptides were synthesized using standard Fmoc solid phase peptide synthesis with HBTU activation. Briefly, Fmoc Wang resin loaded with the first amino acid was allowed to swell in DMF for 15 min (repeated twice). Fmoc deprotection was performed in 20% piperidine in DMF (repeated twice, for 5 min and 15 min). HBTU activation was performed with 3 eq. of Fmoc-amino acid and 3 eq. of HBTU in DMF (5 ml for every gram of resin), with 6 eq. DIPEA. Coupling was performed at RT for 3 h, after thorough washes with DMF and DCM. Final cleavage was obtained using a mixture of TFA/TIPS/water (95:2.5:2.5). The process was repeated to couple the third amino acid. Resulting peptides were allowed to be precipitated in cold ether or lyophilized when appropriate. The precipitated peptides were allowed to dry under vacuum overnight. The purity of each peptide was verified by HPLC. The peptides' identity was verified by ESI-MS and 1H NMR.

#### *HPLC*

A 30  $\mu$ l sample was injected onto a Macherey-Nagel C18 column (Intersil Phenyl for KFD and KHD) with a length of 250 mm and an internal diameter of 4.6 mm and 5-mm fused silica particles at a flow rate of 1 ml/min. The eluting solvent system (in all cases with 0.1% v/v TFA added) had a linear gradient of 20% (v/v) acetonitrile in water for 4 min, gradually rising to 80% (v/v) acetonitrile in water at 35 min. This concentration was kept constant until 40 min when the gradient was decreased to 20% (v/v) acetonitrile in water at 42 min. The purity of each peak was determined by UV detection at 214 nm.

#### *FTIR*

Samples were contained in a standard IR transmission cell (Harrick Scientific) between two 2 mm CaF<sub>2</sub> windows, separated by a polytetrafluoroethylene (PTFE) spacer of 50 μm thickness. Spectra were recorded on a Bruker Vertex 70 spectrometer by averaging 25 scans at a spectral resolution of 1 cm<sup>-1</sup>. Spectra were corrected for absorption from a phosphate buffer blank sample and absorptions from trifluoroacetic acid (TFA).

#### TEM

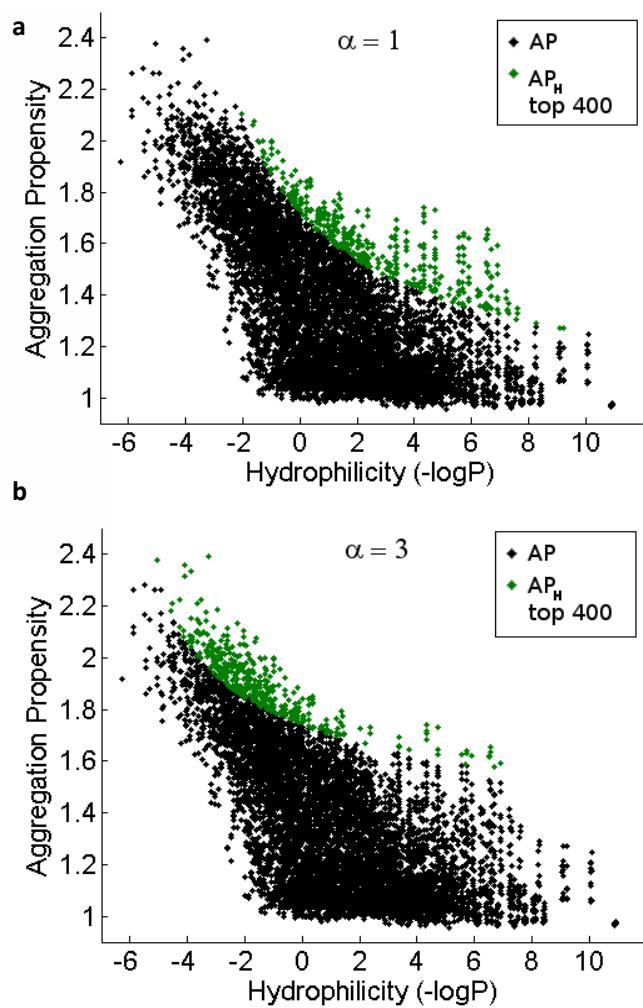
Carbon-coated copper grids (200 mesh) were glow discharged in air for 30 s. The support film was touched onto the gel surface for 3 s and blotted down using filter paper. Negative stain (20 ml, 1% aqueous methylamine vanadate (Nanovan; Nanoprobe) was applied and the mixture was blotted again using filter paper to remove excess. The dried specimens were then imaged using a LEO 912 energy filtering transmission electron microscope operating at 120kV fitted with 14 bit/2 K Proscan CCD camera

#### Diffusion Ordered NMR Spectroscopy (DOSY)

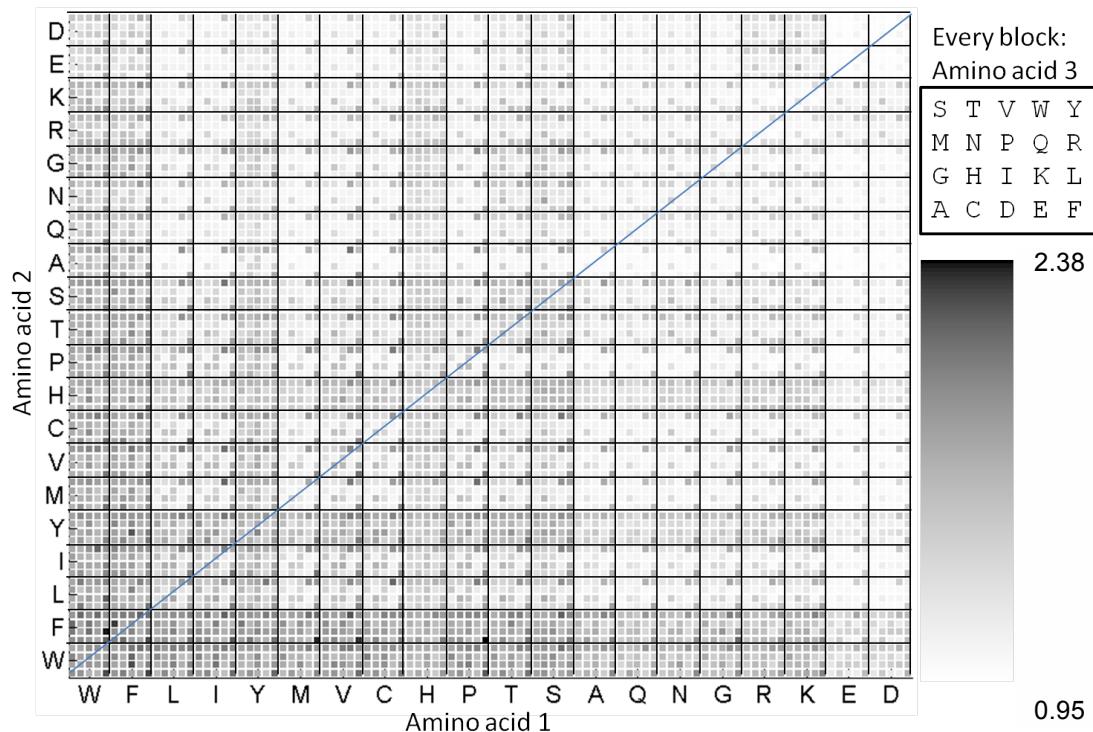
Peptide solutions were prepared in D<sub>2</sub>O at 10 mM peptide concentration as a compromise between solubility and NMR signal intensity. DOSY spectra were acquired at 600 MHz using a Bruker Avance 600 spectrometer using the Bruker microprogram *dstebp3sat* at 298 K. The eddy current delay (Te) was set to 5 ms. The diffusion time was adjusted to 100 ms. The duration of the pulse field gradient, δg, was optimized in order to obtain 5% residual signal with the maximum gradient strength with the resulting δ value of 3.6 ms. The pulse gradient was increased from 2 to 95% of the maximum gradient strength using a linear ramp 16k data points in the F2 dimension (20 ppm) and 16 data points in the F1 dimension were collected. Final data sizes were 16k×128.

#### Dynamic Light Scattering (DLS)

The dynamic light scattering (DLS) measurements were carried out on 10 mmol/L peptide solutions (i.e. below the critical gelation concentration) by using a 3 DDLS spectrophotometer (LS instruments, Fribourg, Switzerland) using vertically polarized He-Ne laser light (25 mW with wavelength of 632.8 nm) with an avalanche photodiode detector at an angle of 90° at 25 °C. Intensity autocorrelation functions were recorded and analyzed by means of the cumulant method in order to determine the intensity weighted diffusion coefficients D and the average hydrodynamic radius R<sub>h</sub> by using the Stokes-Einstein equation,  $R_h = k_B T / 6\pi\eta D$ , where k<sub>B</sub> is the Boltzmann constant, T is the absolute temperature and η is the solvent viscosity at the given temperature.

Supporting Figures

**Supplementary Figure 1.** AP as a function of hydrophilicity for all 8000 tripeptides. Green dots represent the top 400 tripeptides from the hydrophobicity-corrected  $AP_H$  score with alternative values for coefficient  $\alpha$ . **a**,  $\alpha = 1$ . **b**,  $\alpha = 3$ . For  $\alpha = 2$  see main text Fig. 2.

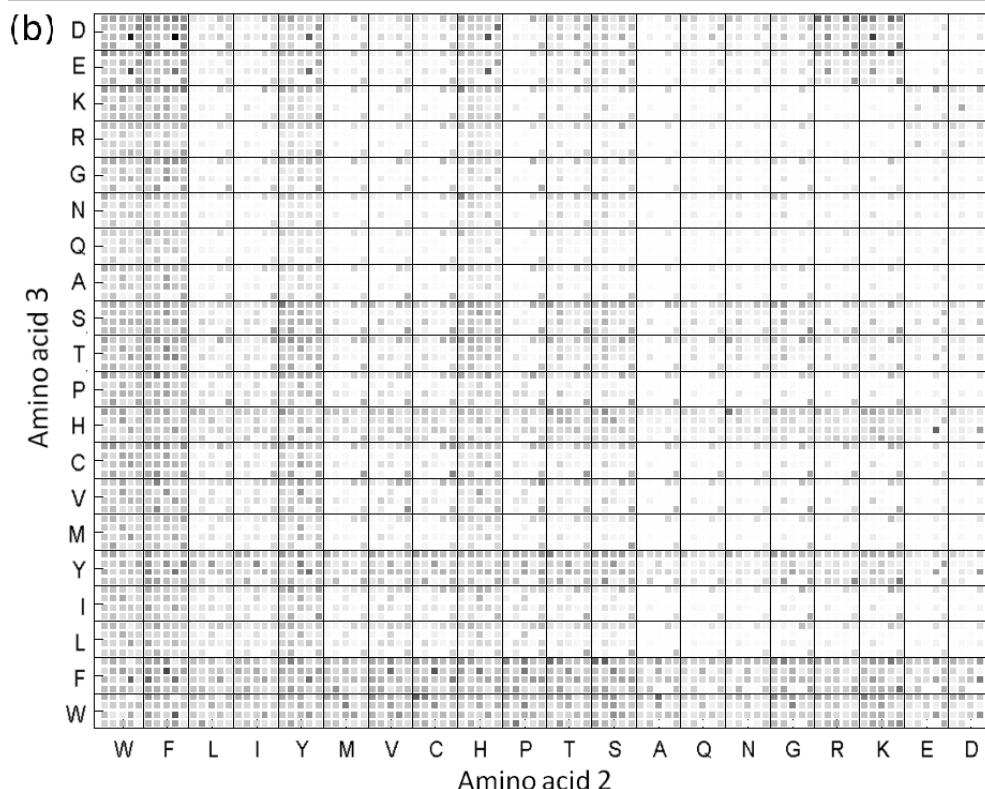
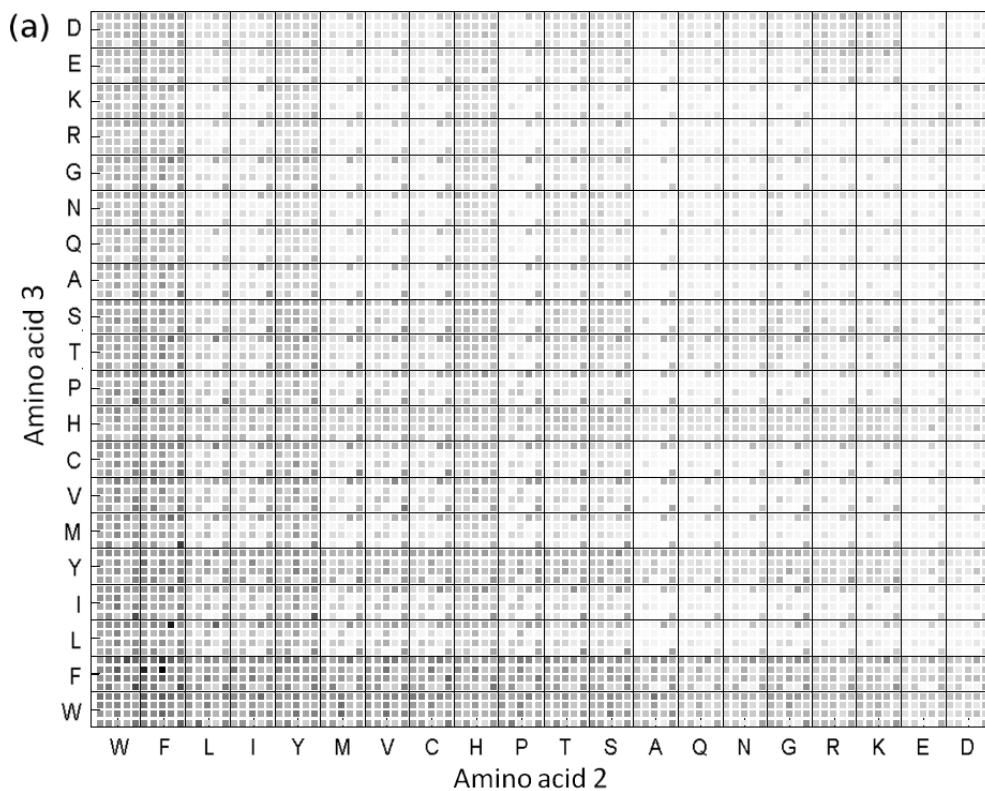


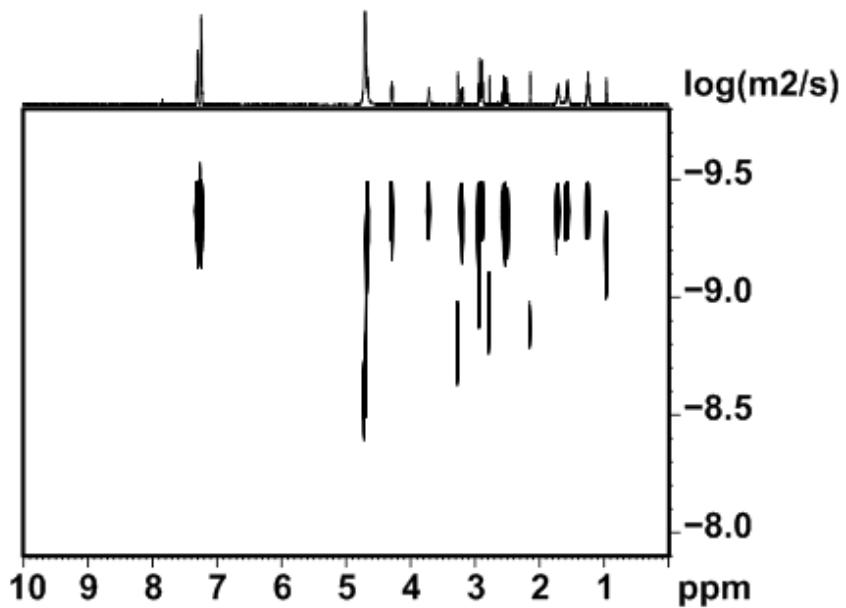
**Supplementary Figure 2.** AP score for all 8000 combinations of three amino acids after a 50 ns simulation. Within every rectangle, the third amino acid is represented by the position of the coloured square at the locations indicated in the legend on the right.

**Supplementary Figure 3 (next page).** Alternative representation of AP (**a**) and  $AP_H$  (**b**) scores for all 8000 combinations of three amino acids after a 50 ns simulation. Within every rectangle, the *first* amino acid is represented by the position of the coloured square at the locations indicated in the legend on the right.

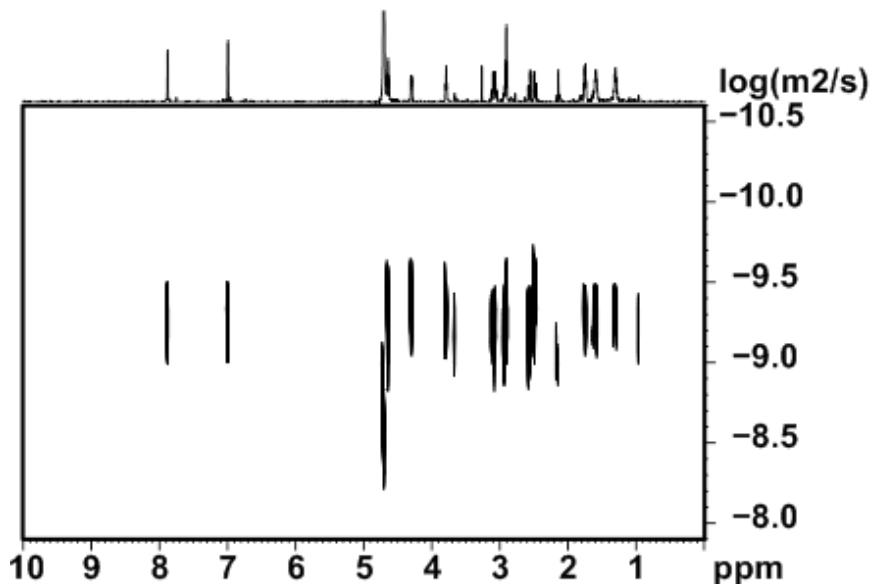
Every block:  
Amino acid 1

S	T	V	W	Y
M	N	P	Q	R
G	H	I	K	L
A	C	D	E	F

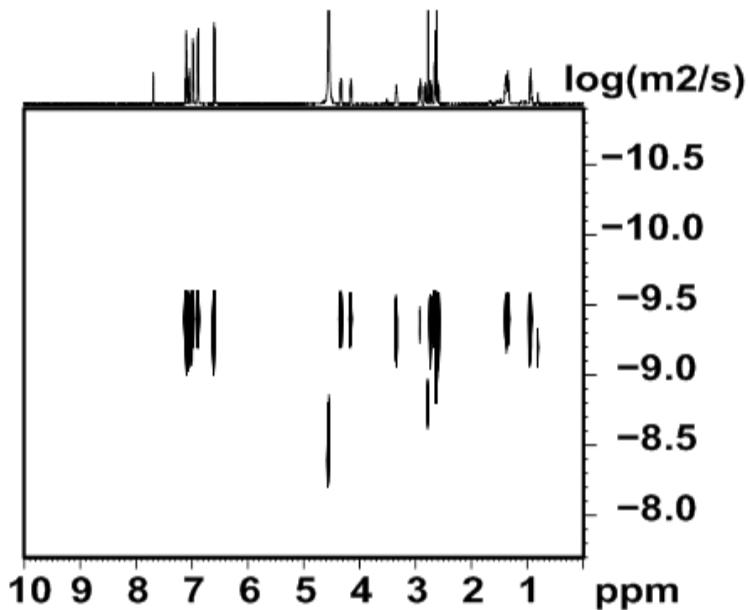




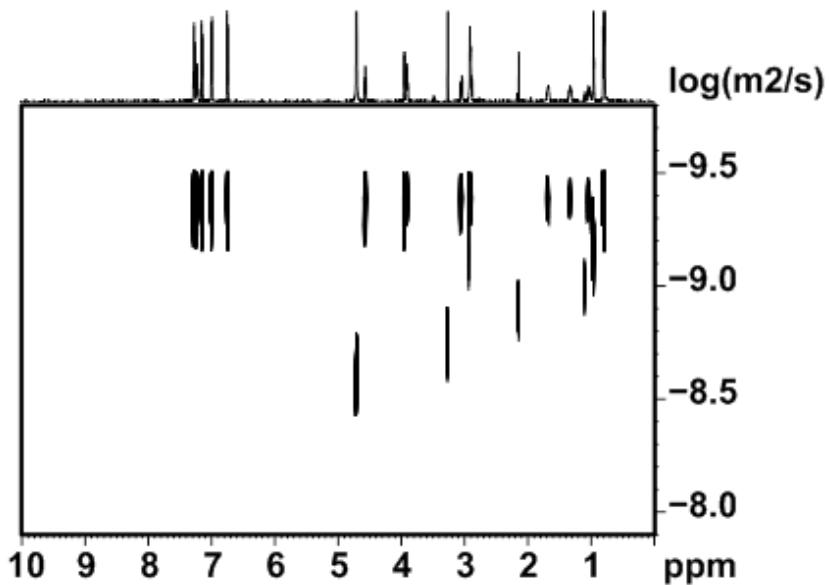
**Supplementary Figure 4:** DOSY NMR spectrum of 10 mM of KHD in D<sub>2</sub>O at pH 7 at 298K.



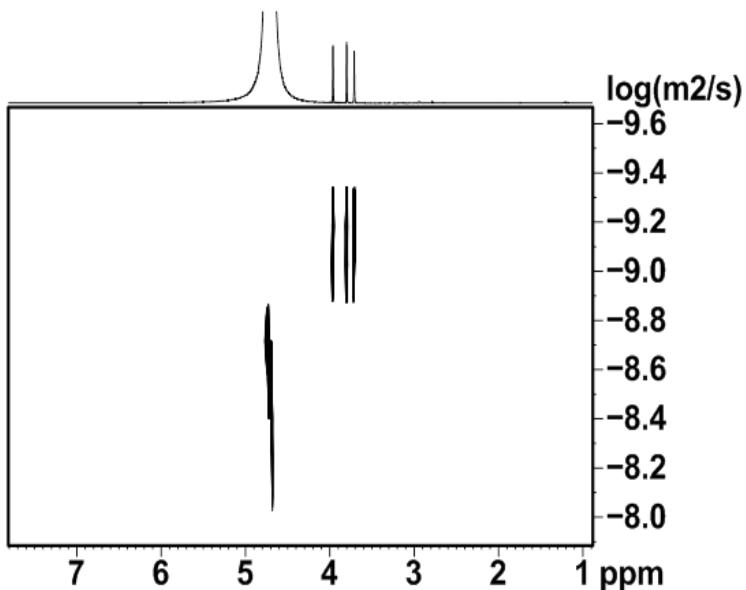
**Supplementary Figure 5:** DOSY NMR spectrum of 10 mM of KFD in D<sub>2</sub>O at pH 7 at 298K.



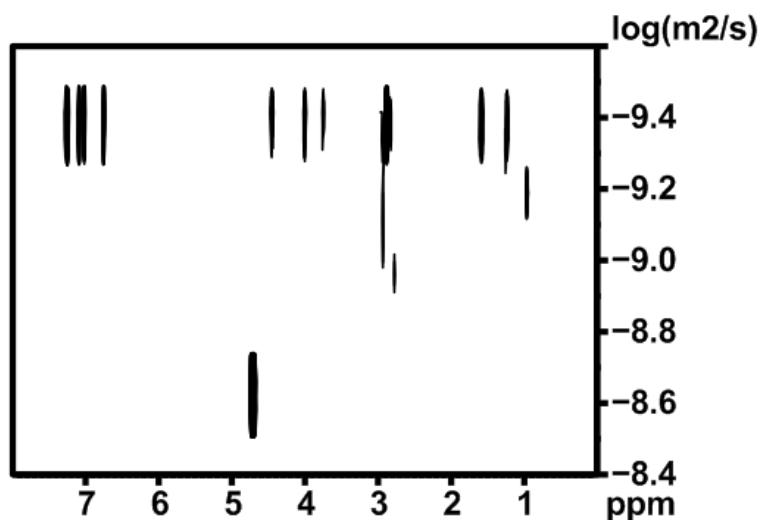
**Supplementary Figure 6:** DOSY NMR spectrum of 10 mM of KYF in  $\text{D}_2\text{O}$  at pH 7 at 298K.



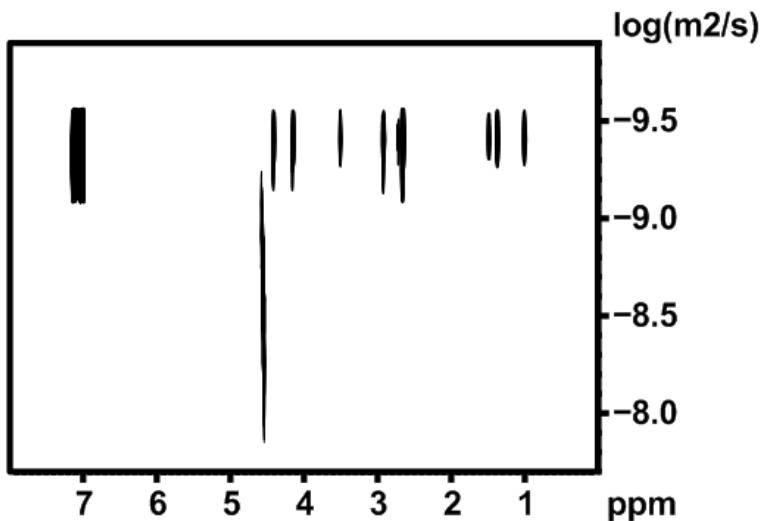
**Supplementary Figure 7:** DOSY NMR spectrum of 10 mM of YFI in  $\text{D}_2\text{O}$  at pH 7 at 298K.



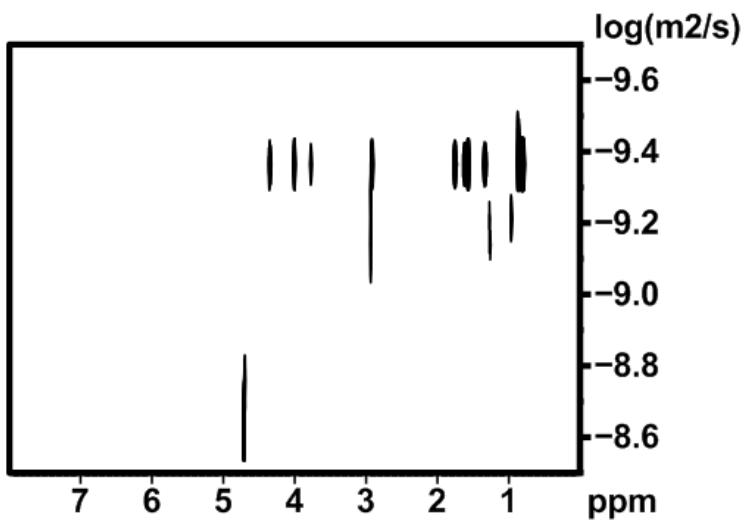
**Supplementary Figure 8:** DOSY NMR spectrum of 10 mM of GGG in  $D_2O$  at pH 7 at 298K. Note that the diffusion constant for GGG was determined from the aliphatic backbone protons, as this peptide lacks aromatic groups.



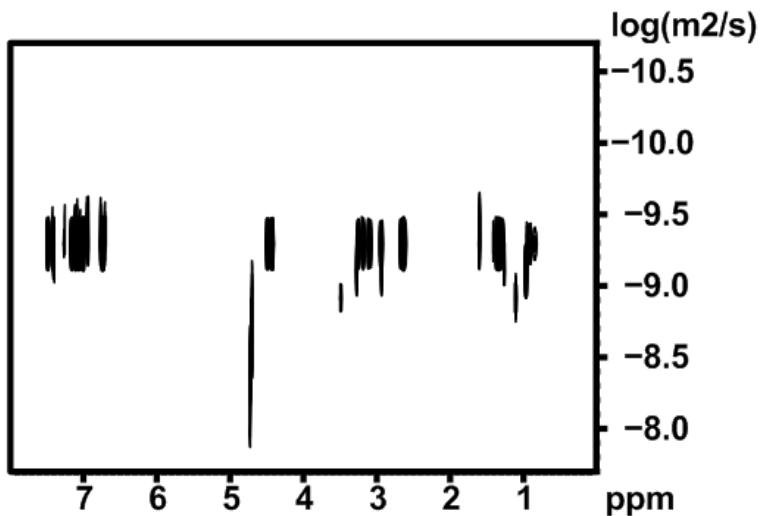
**Supplementary Figure 9:** DOSY NMR spectrum of 10 mM of FYK in  $D_2O$  at pH 7 in 298K.



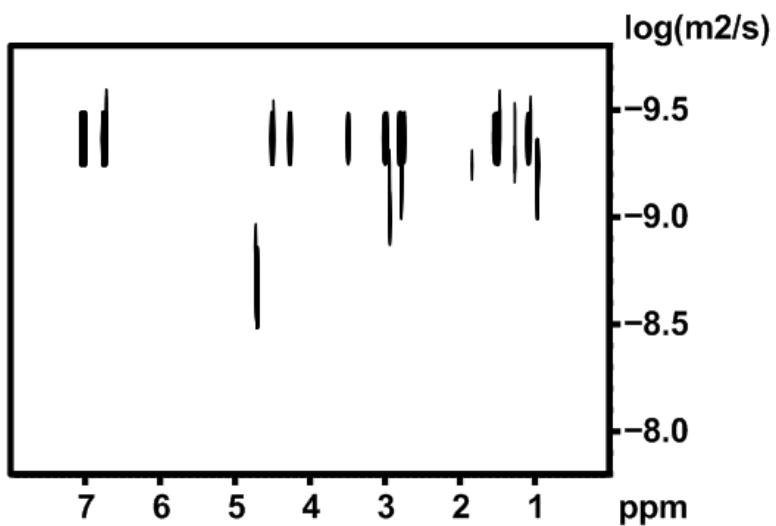
**Supplementary Figure 10:** DOSY NMR spectrum of 10 mM of KFF in  $D_2O$  at pH 7 in 298K.



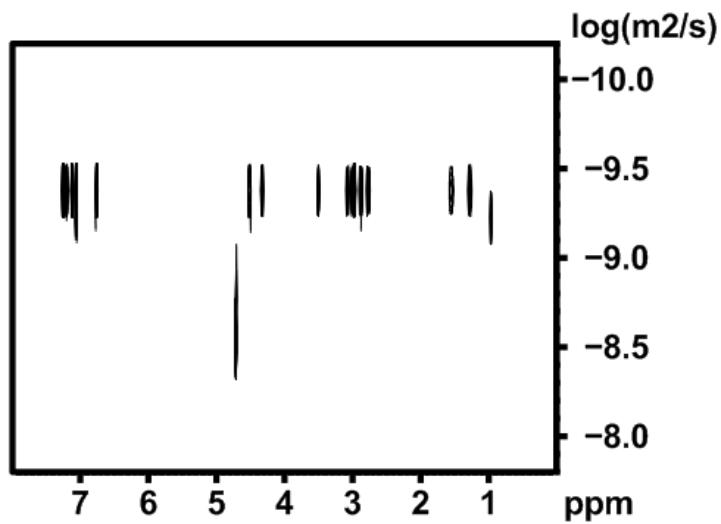
**Supplementary Figure 11:** DOSY NMR spectrum of 10 mM of KLL in  $D_2O$  at pH 7 in 298K. Note that the diffusion constant for GGG was determined from the aliphatic backbone protons, as this peptide lacks aromatic groups.



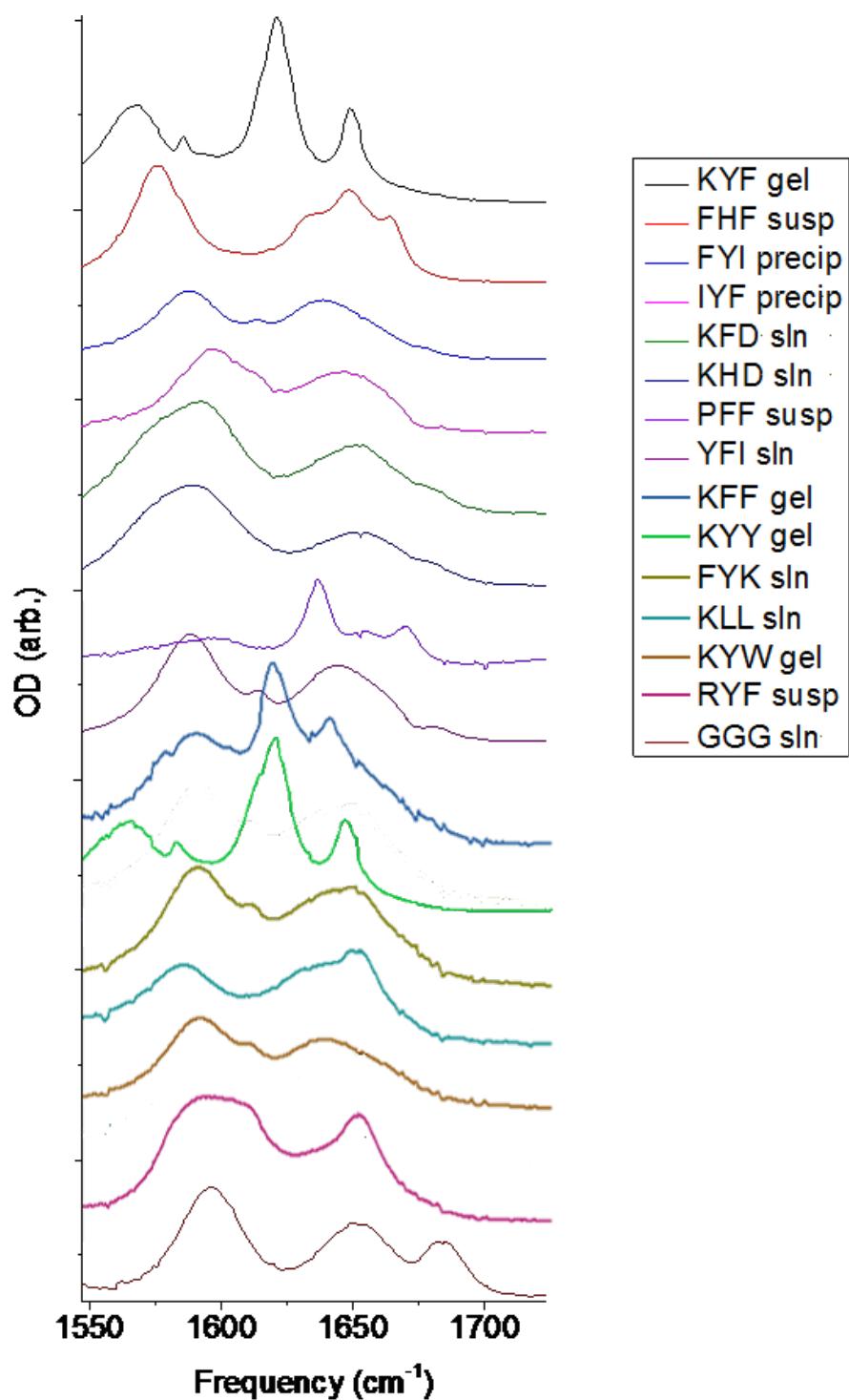
**Supplementary Figure 12:** DOSY NMR spectrum of 10 mM of KYW in D<sub>2</sub>O at pH 7 in 298K.



**Supplementary Figure 13:** DOSY NMR spectrum of 10 mM of KYY in D<sub>2</sub>O at pH 7 in 298K.



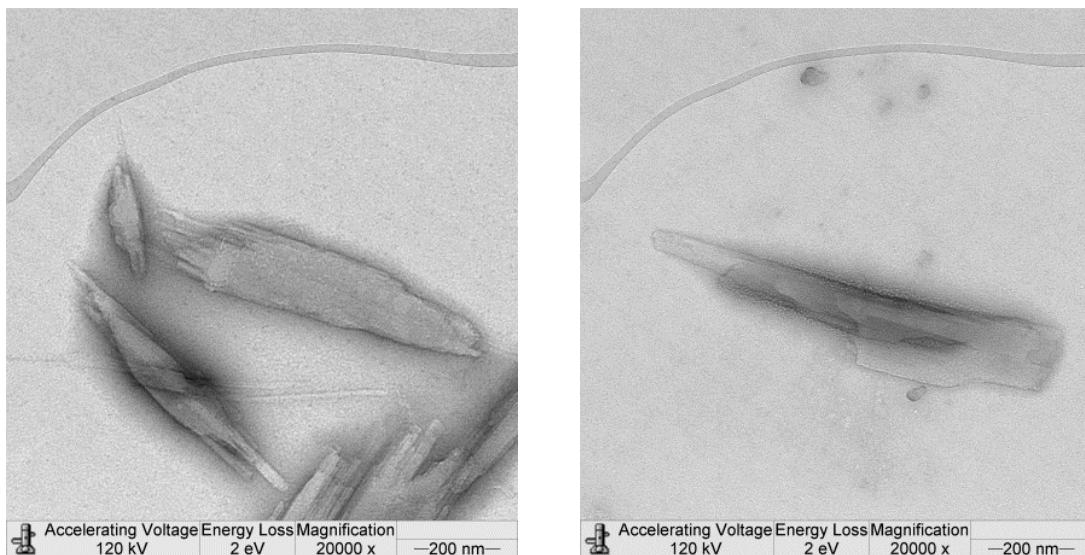
**Supplementary Figure 14:** DOSY NMR spectrum of 10 mM of RYF in D<sub>2</sub>O at pH 7 in 298K.



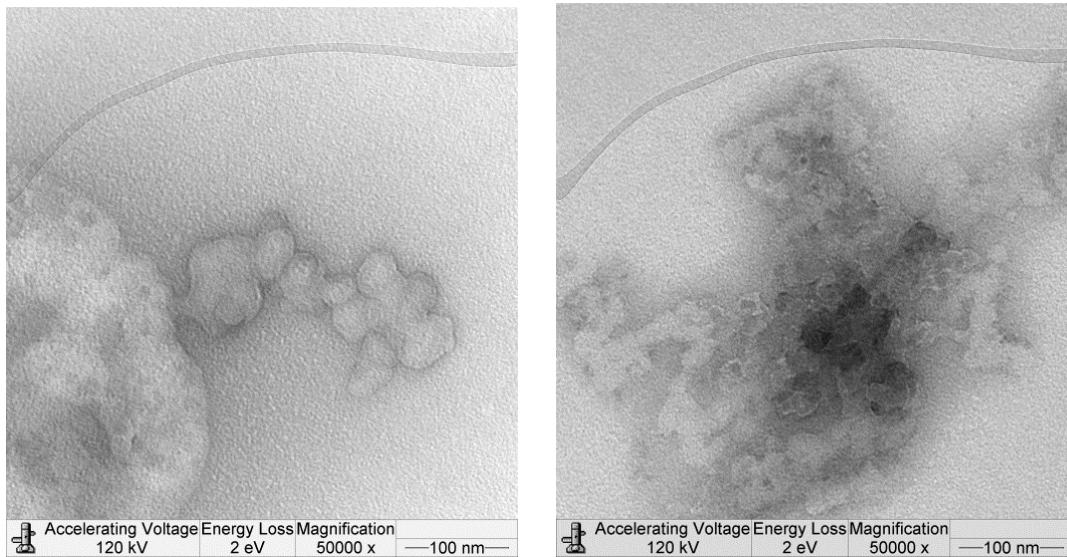
**Supplementary Figure 15.** FTIR absorption spectra of 30 mM peptide solutions in pD 8.01 M sodium phosphate buffer in  $\text{D}_2\text{O}$  and KYF hydrogel in  $\text{D}_2\text{O}$  at pD 7.2. Path length 50  $\mu\text{m}$ . Spectra have been corrected for absorption of atmospheric water, liquid  $\text{D}_2\text{O}$  background absorption and residual trifluoroacetic acid from synthesis. The KYF and KYY spectra have been divided by 5 for clarity.

Analysis: IR absorption in the amide I region ( $1600\text{-}1700\text{ cm}^{-1}$ ) is sensitive to secondary structure of proteins<sup>11</sup> and short peptides.<sup>3</sup> For short peptides in solution, a weak, broad absorption centred around  $1650\text{ cm}^{-1}$  is observed, together with a broad absorption around  $1595\text{ cm}^{-1}$ , which are assigned to amide (CO-NH) and carboxylate (COO<sup>-</sup>) vibrational modes, respectively. These modes can be observed for peptides GGG, KFD, KHD, IYF, FYI, FYK, KLL, RYF and KYW in Supplementary Figure 15. When aggregation takes places via intermolecular hydrogen bonding of the amide groups, the  $1650\text{ cm}^{-1}$  typically narrows and shifts to lower frequency, while the  $1595\text{ cm}^{-1}$  broadens or decreases in intensity due to protonation or salt bridge formation. This was observed for PFF and, to a lesser extent, FHF, KLL and KYW. For KYF, KFF and KYY, a dramatic change in the amide region was noticed upon gelation with intense absorption at  $1621$  and  $1649\text{ cm}^{-1}$  for the amide groups. In KYF and KYY, the salt-bridged carboxylate groups were found at  $1568\text{ cm}^{-1}$ . These observations indicate strong interactions between N-terminus or lysine side chains and the C-terminus and strong intramolecular hydrogen bonding between amide modes, both suggesting a well-ordered peptide nanostructure.

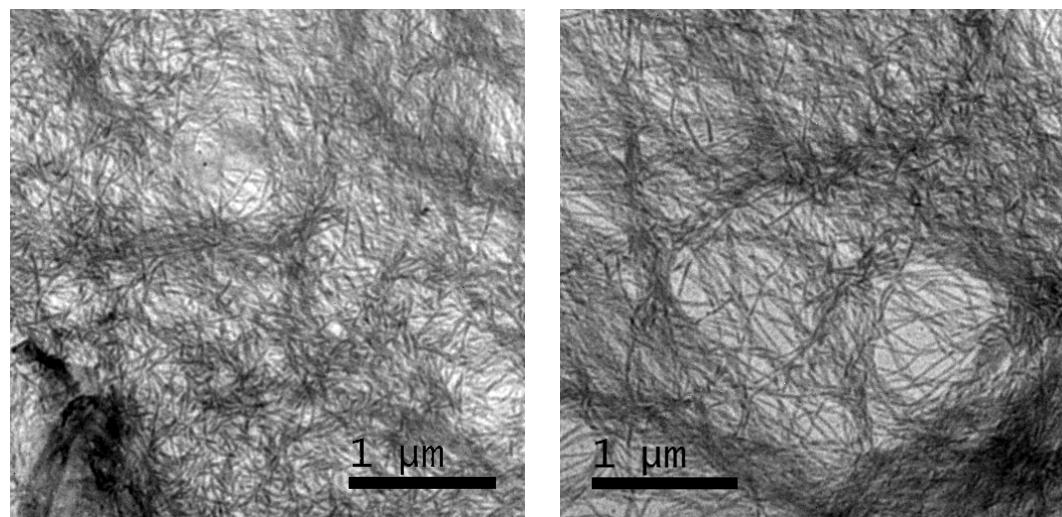
Furthermore, broad contributions at  $1580\text{ cm}^{-1}$  in KFD and KHD originate from aspartic acid side chain absorption, while weak absorptions at  $1611\text{ cm}^{-1}$  in IYF, FYI and KYF originate from a tyrosine side chain vibrational mode and a broad absorption around  $1608\text{ cm}^{-1}$  in RYF is assigned to the arginine side chain.<sup>11</sup> In the spectra of KFD, KHD, KYF and PFF a small residual absorption from trifluoroacetic acid can be found around  $1676\text{ cm}^{-1}$ .



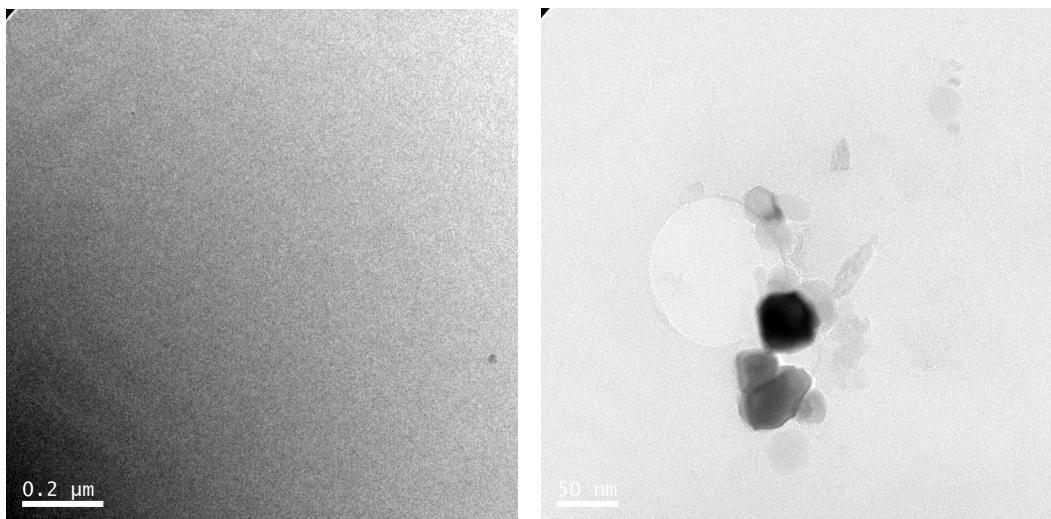
**Supplementary Figure 16.** Additional TEM images of PFF



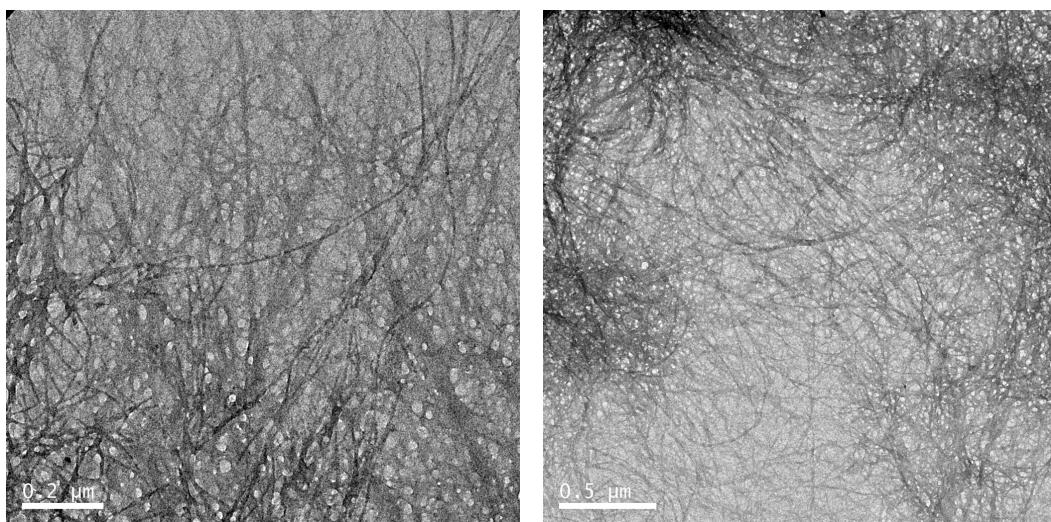
**Supplementary Figure 17.** Additional TEM images of KFD



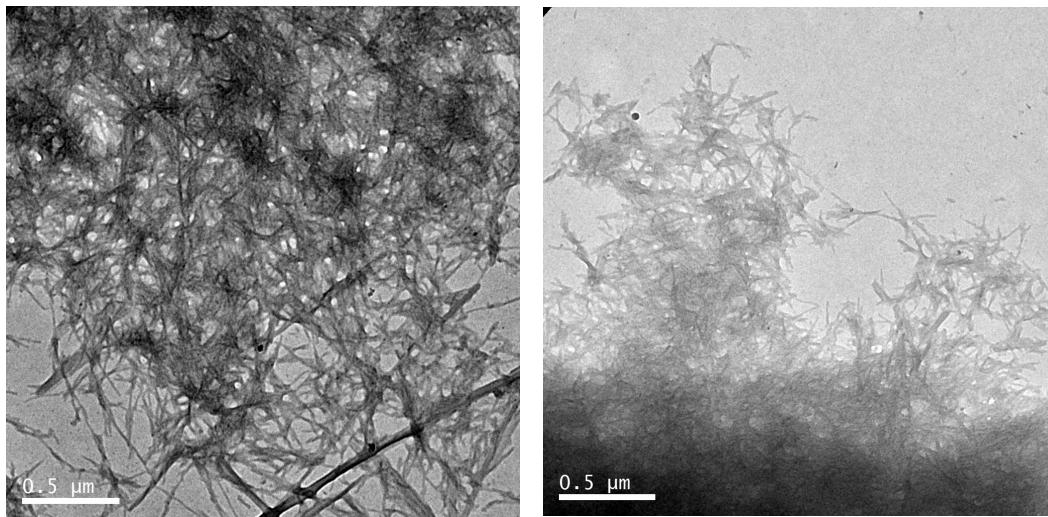
**Supplementary Figure 18.** Additional TEM images of KYF gel



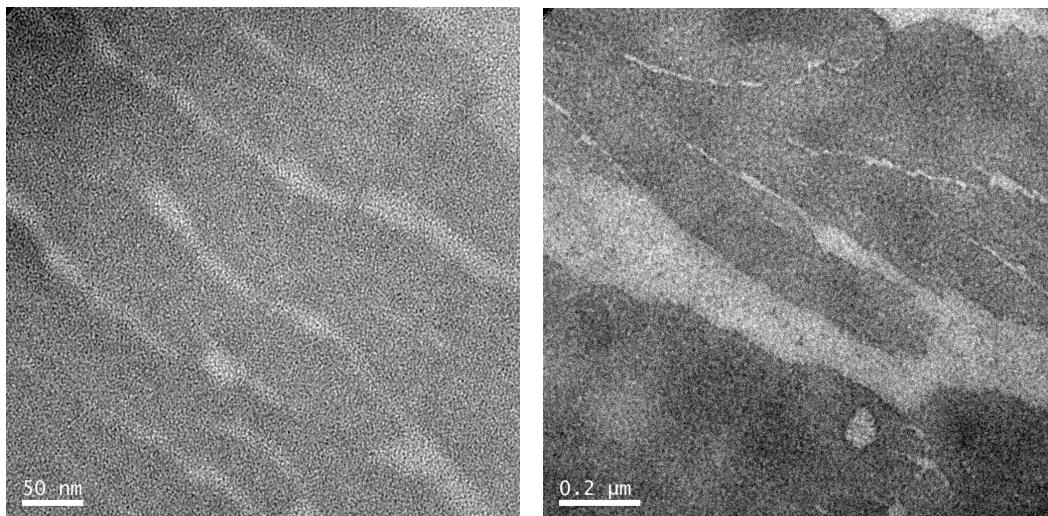
**Supplementary Figure 19.** TEM images of GGG solution



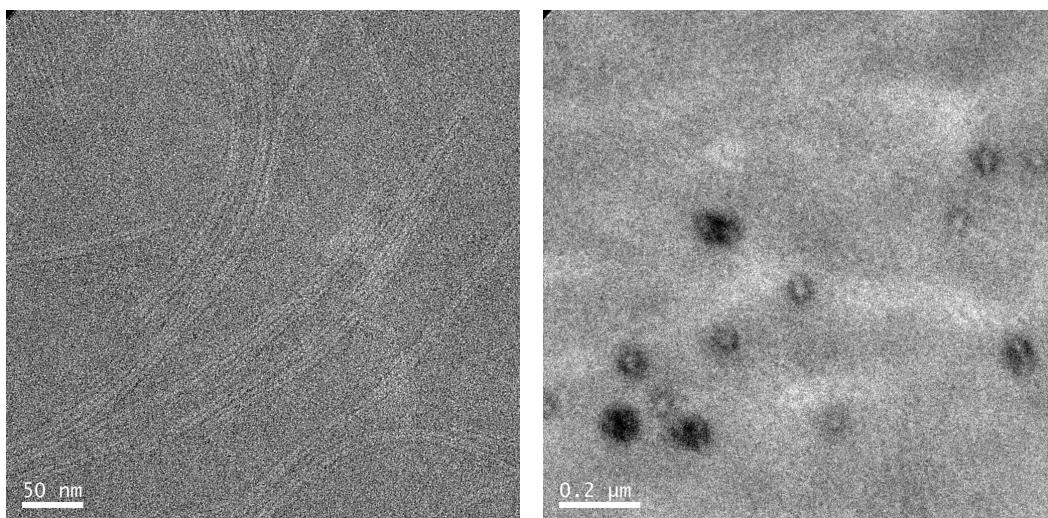
**Supplementary Figure 20.** TEM images of KFF gel

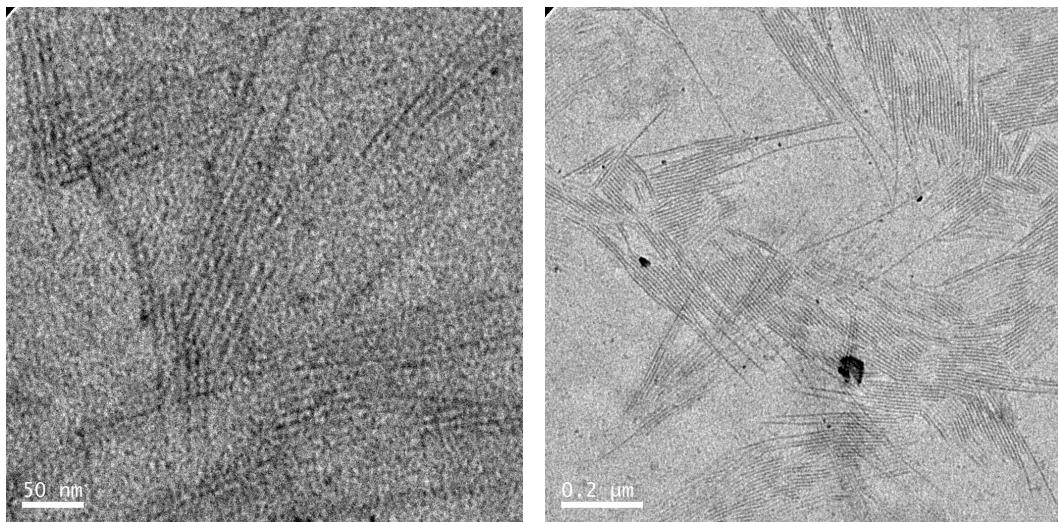
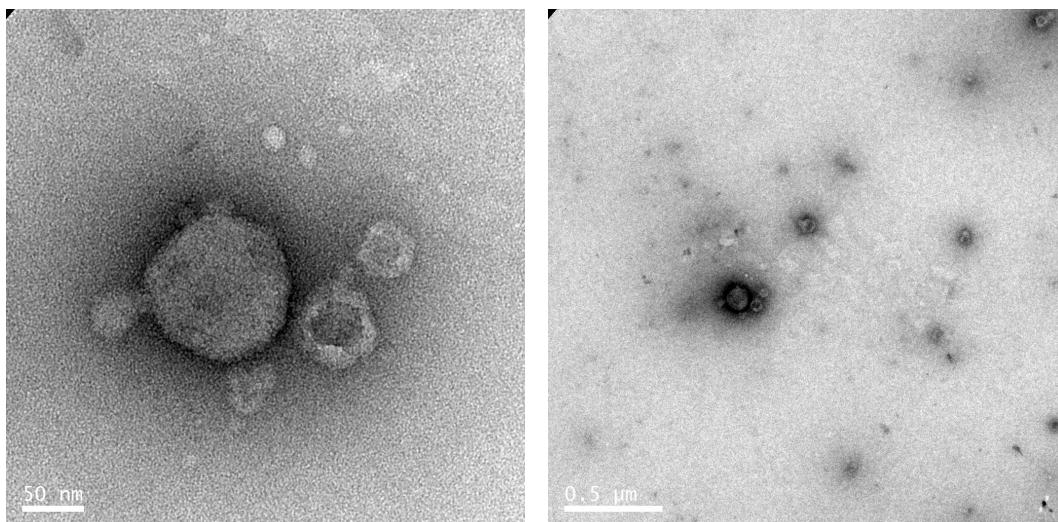


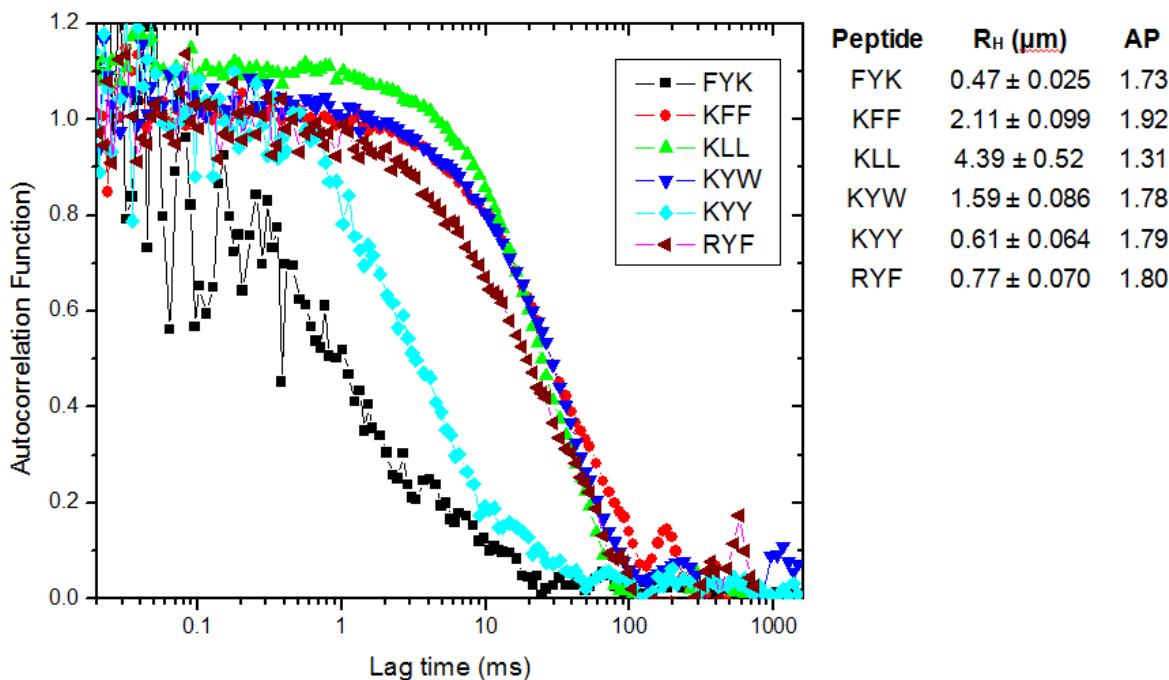
**Supplementary Figure 21.** TEM images of KYY gel



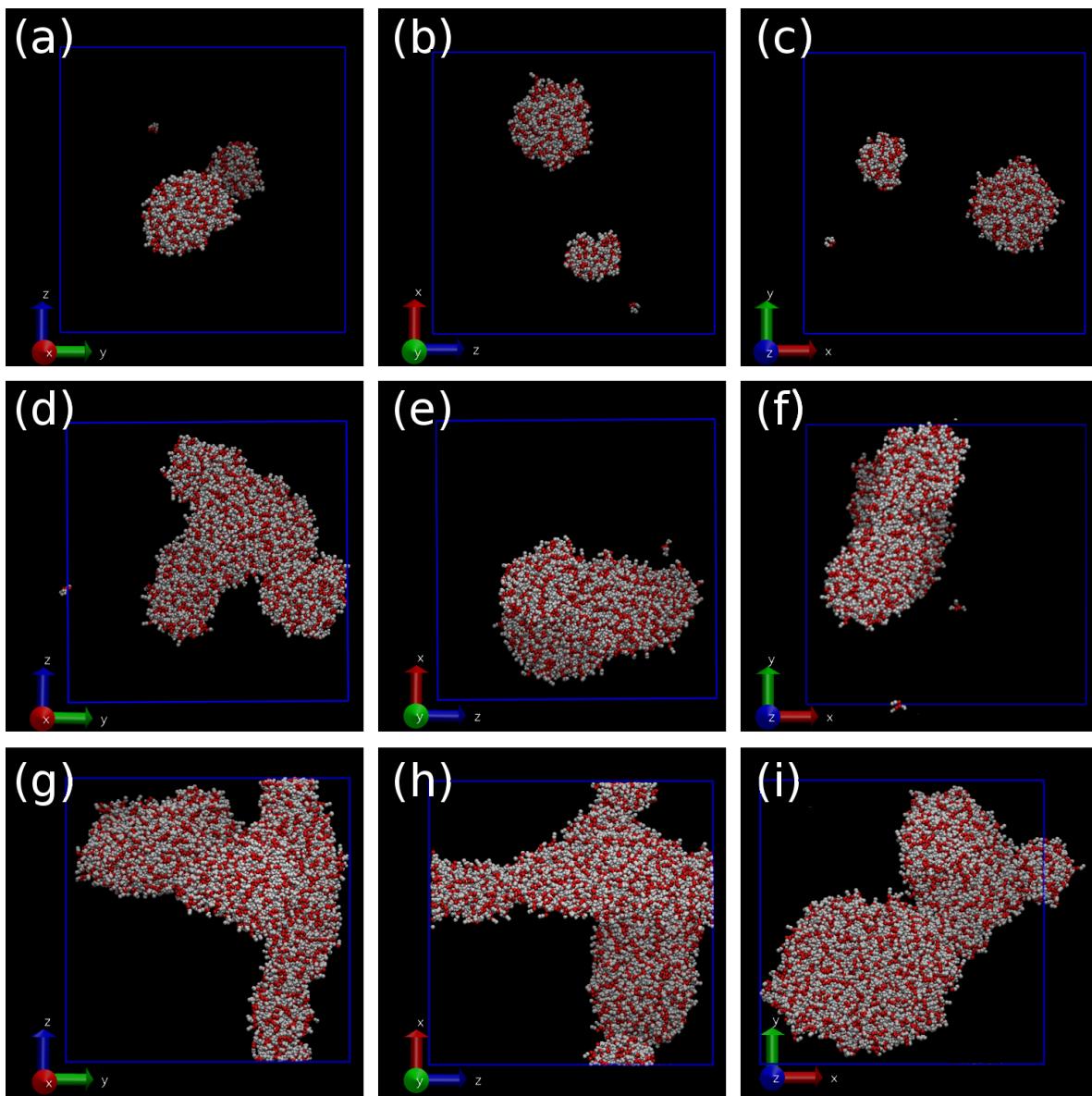
**Supplementary Figure 22.** TEM images of KYW gel



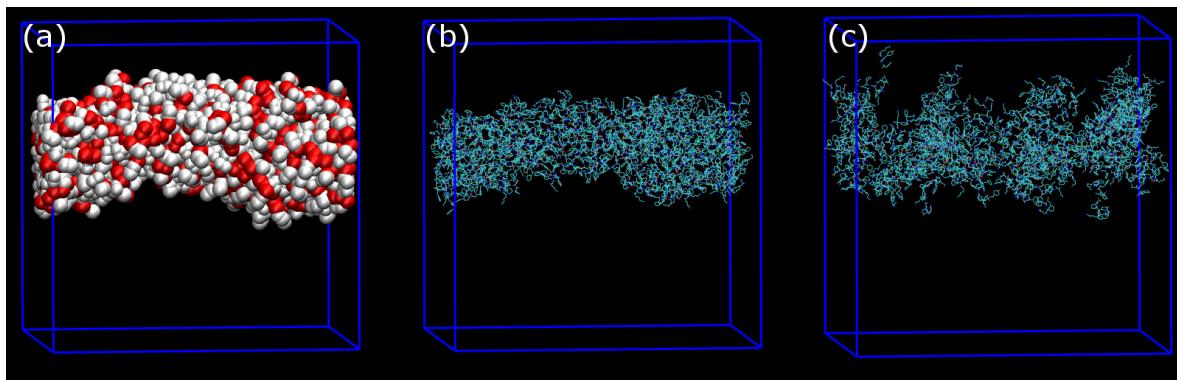
**Supplementary Figure 23.** TEM images of KLL solution**Supplementary Figure 24.** TEM images of RYF suspension**Supplementary Figure 25.** TEM images of FYK solution



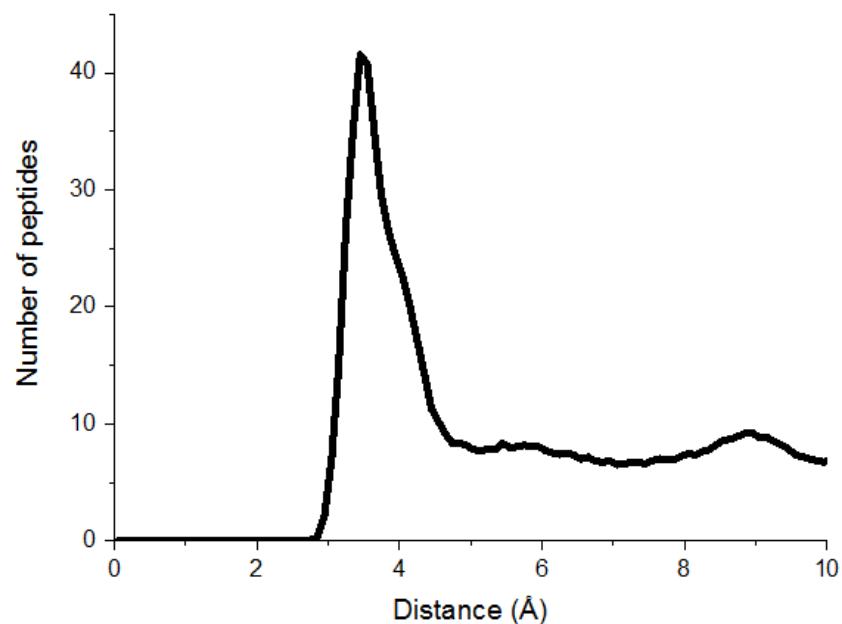
**Supplementary Figure 26.** Additional DLS autocorrelation functions for FYK, KFF, KLL, KYW, KYY and RYF (10 mM at pH 7). The extracted hydrodynamic radii are displayed top right.



**Supplementary Figure 27.** Result of 4800 ns simulations of KYF tripeptides in 24 x 24 x 24 nm boxes. Periodic boundaries are indicated in blue. (a-c) 300 peptides (0.03M) viewed along x, y and z-axes. (d-f) 1200 peptides (0.15M) viewed along x, y and z-axes. (g-i) 2400 peptides (0.3M) viewed along x, y and z-axes.



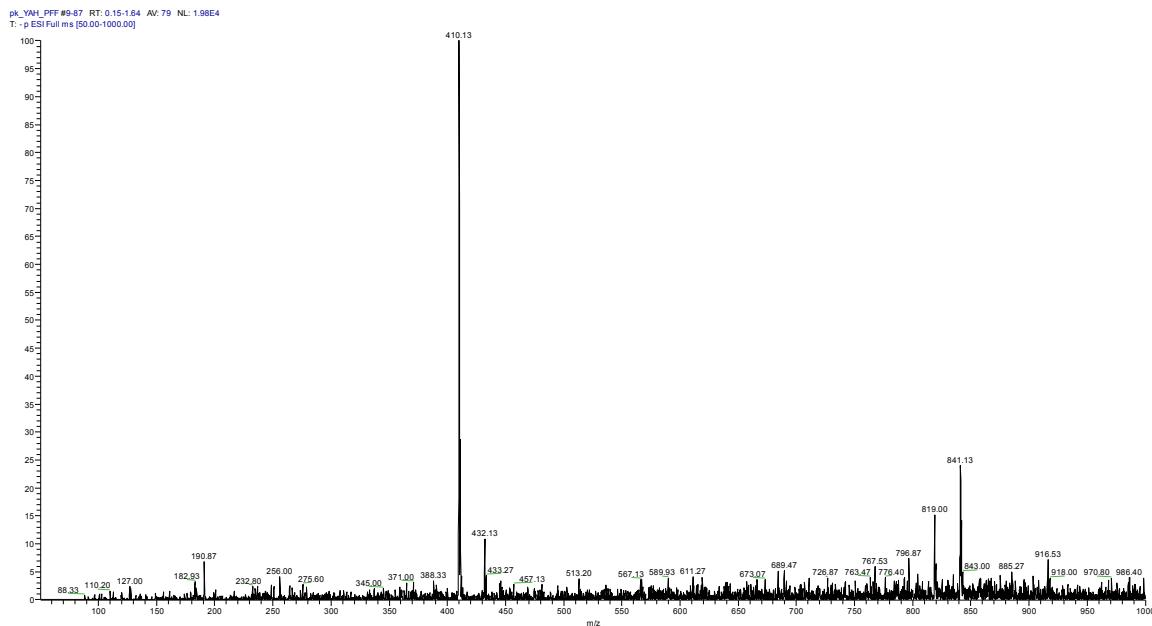
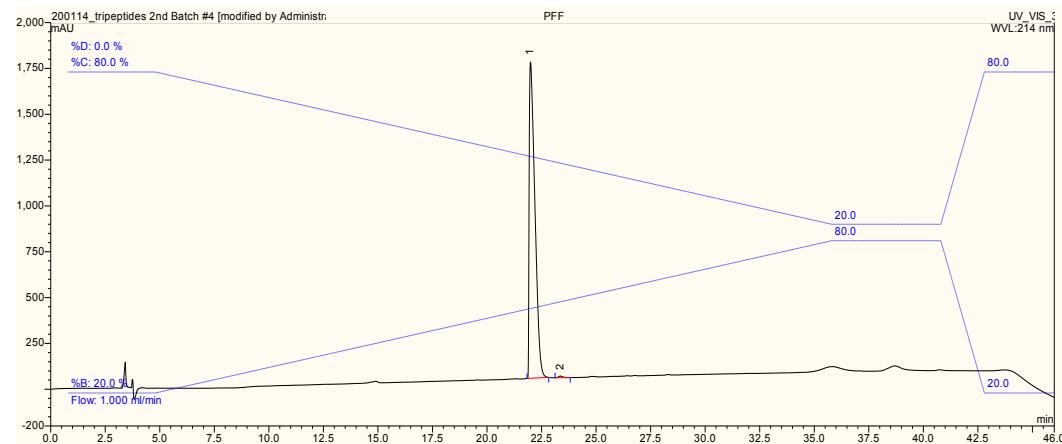
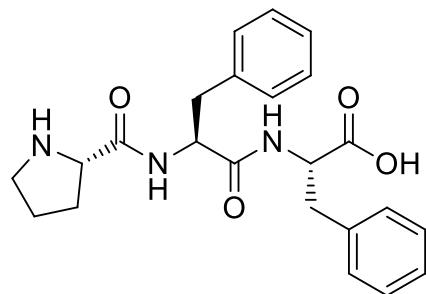
**Supplementary Figure 28.** Backmapping of the 4800 ns result of 300 KYF peptides at 0.23M. (a) CG simulation result. (b) Backmapped and minimized structure. (c) Result of 30 ns simulation on backmapped structure showing multiple protrusions from the original fibre. In each panel, the blue cube indicates the periodic boundaries in the simulation.

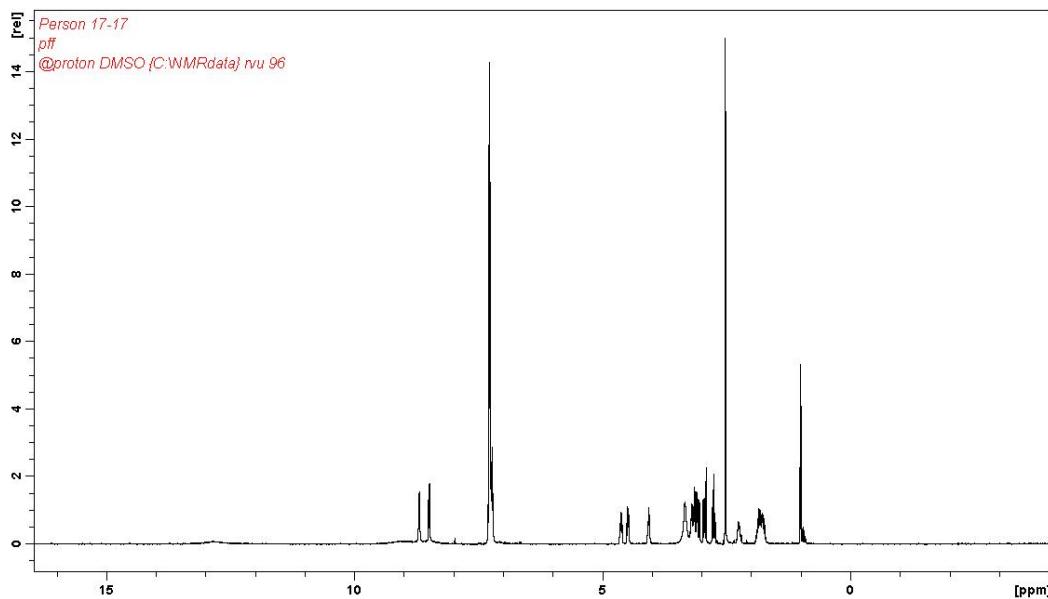


**Supplementary Figure 29.** Radial distribution function of lysine nitrogen atoms (both side chain and terminal amine) to phenylalanine carboxylate carbon, averaged over the final 25 ns of the MD simulation. The maximum at 3.5 Å indicates the abundance of salt bridges.<sup>12</sup>

## Structure, HPLC, Mass, NMR of peptides

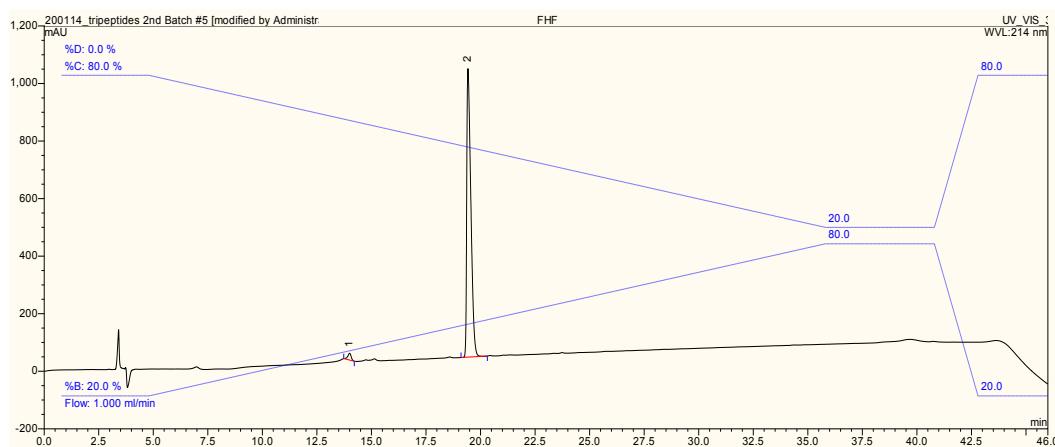
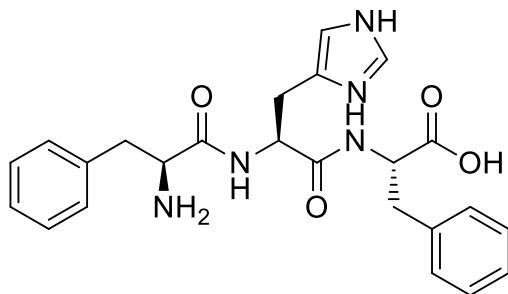
## Pro-Phe-Phe: (Purity: 99.6%)

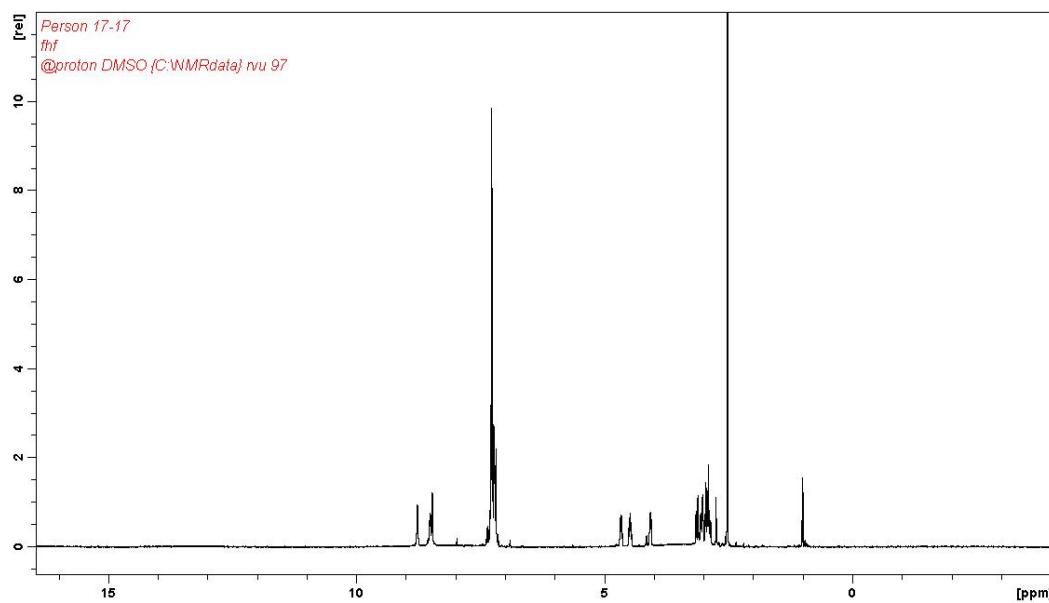
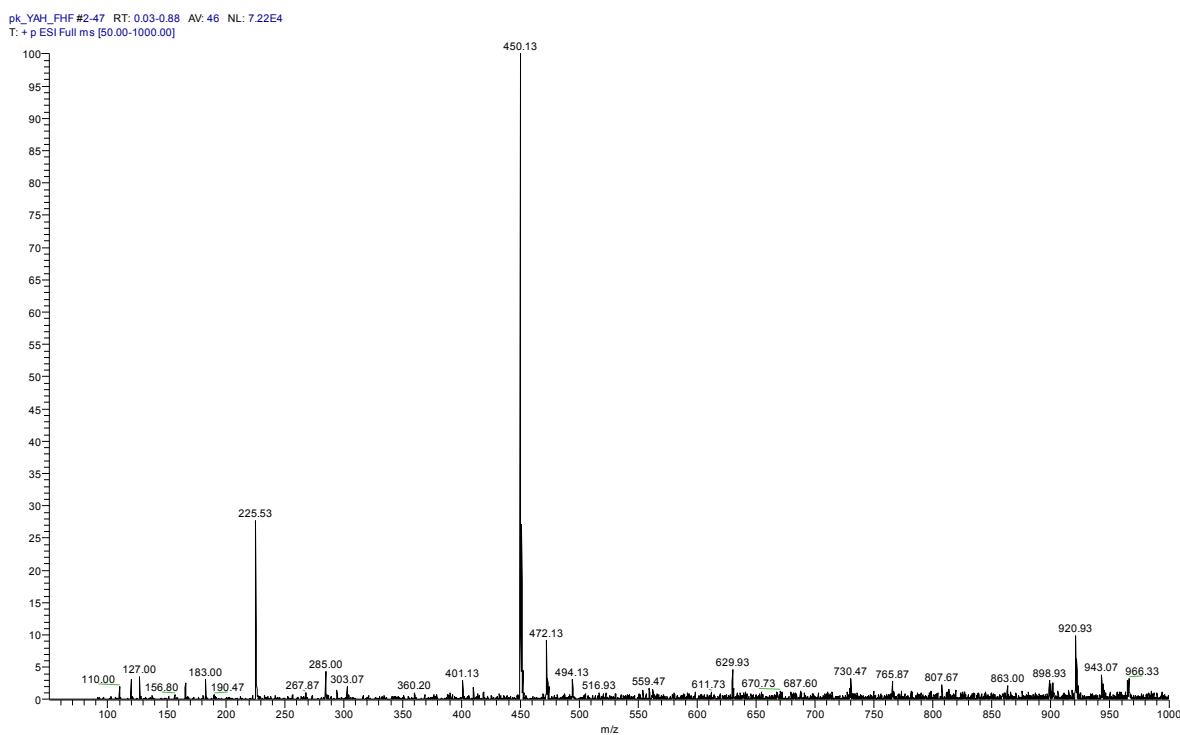




**Supplementary Figure 30.** Structure, HPLC trace, Mass Spec and NMR of PFF

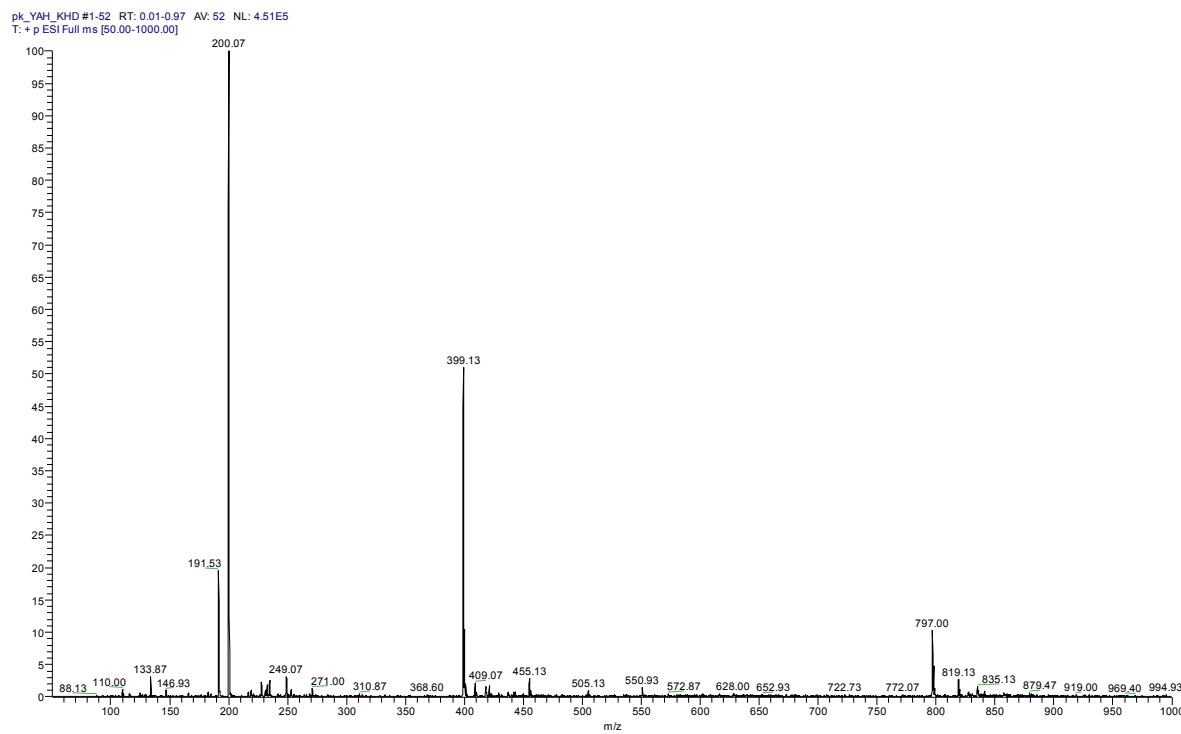
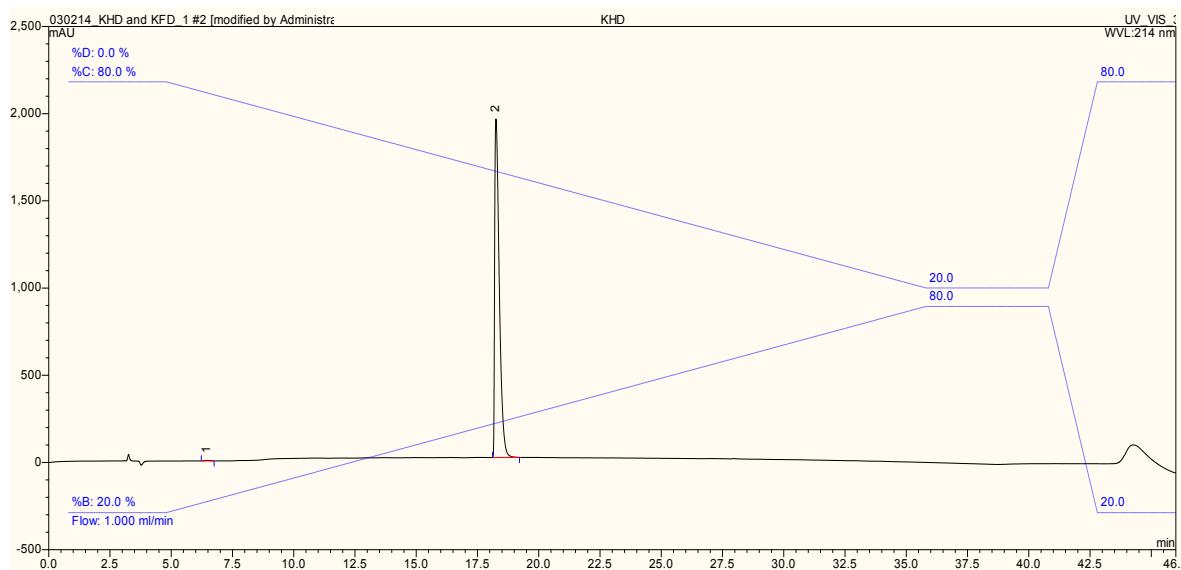
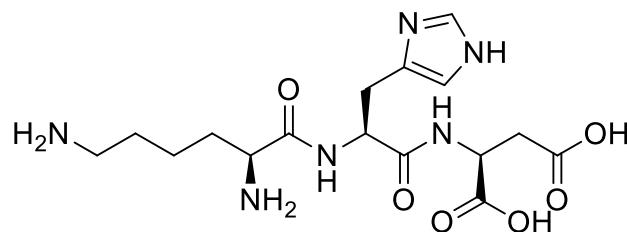
**Phe-His-Phe (Purity: 98.1%)**

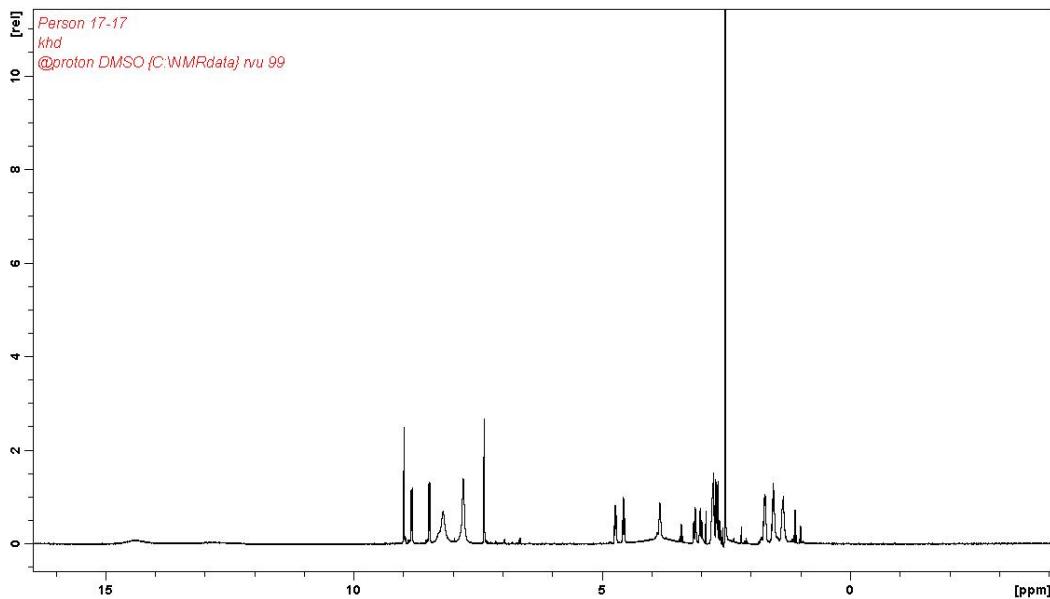




**Supplementary Figure 31.** Structure, HPLC trace, Mass Spec and NMR of FHF

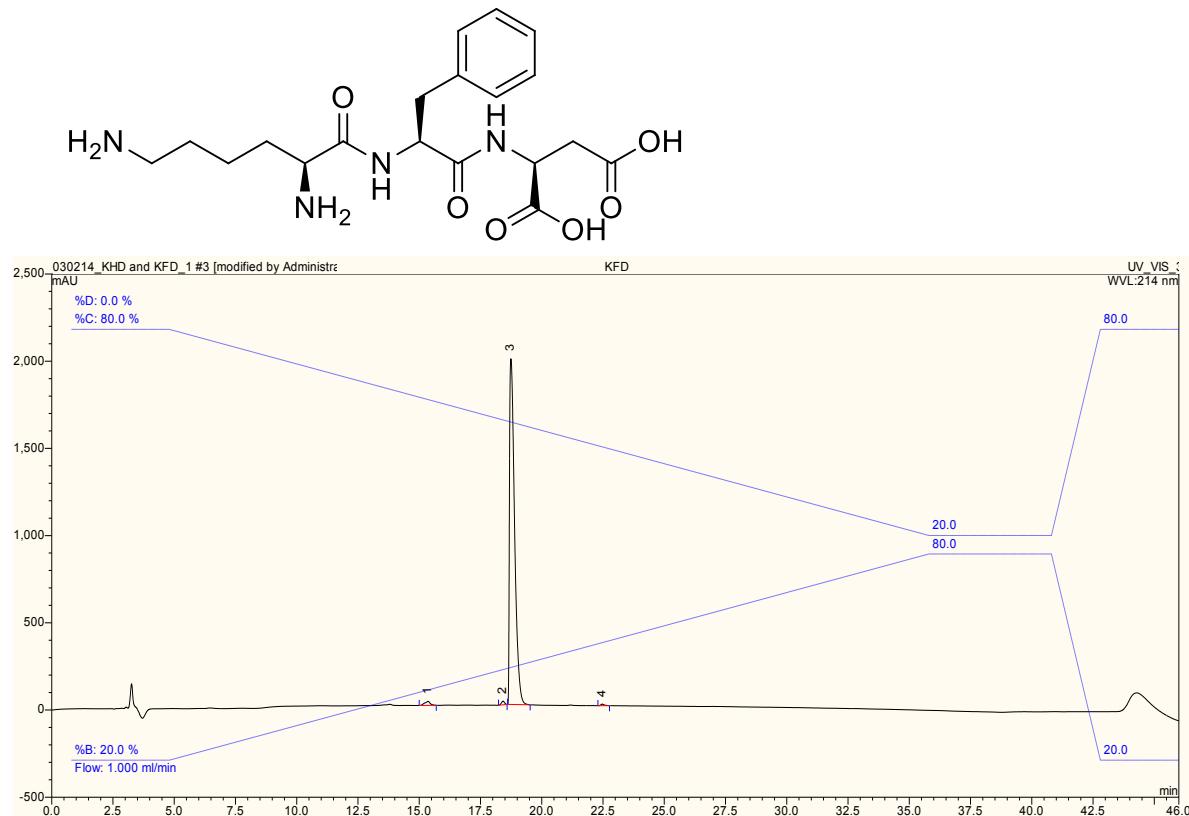
**Lys-His-Asp (Purity: 99.8%)**

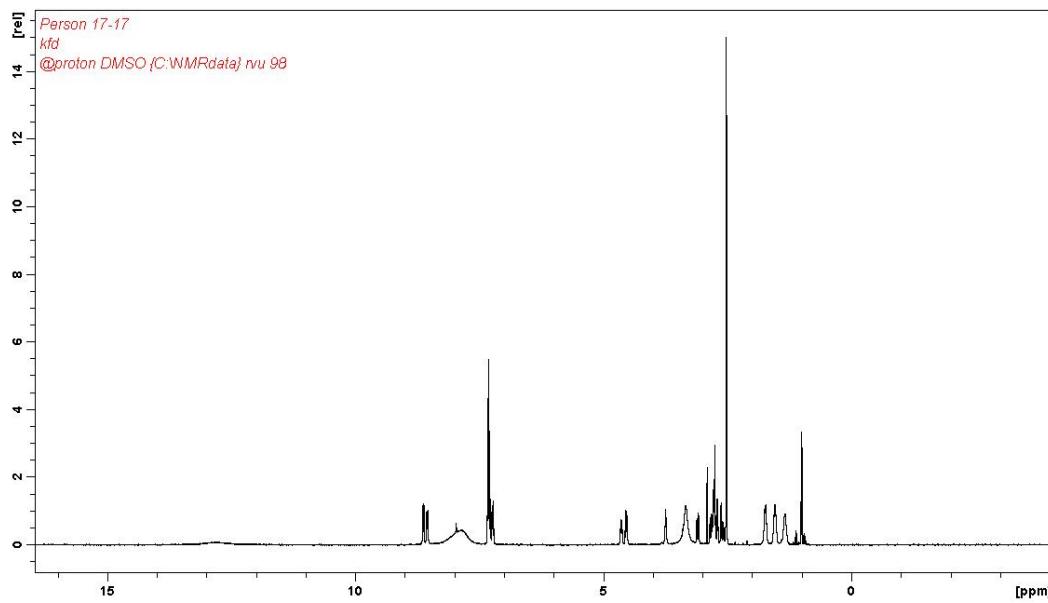
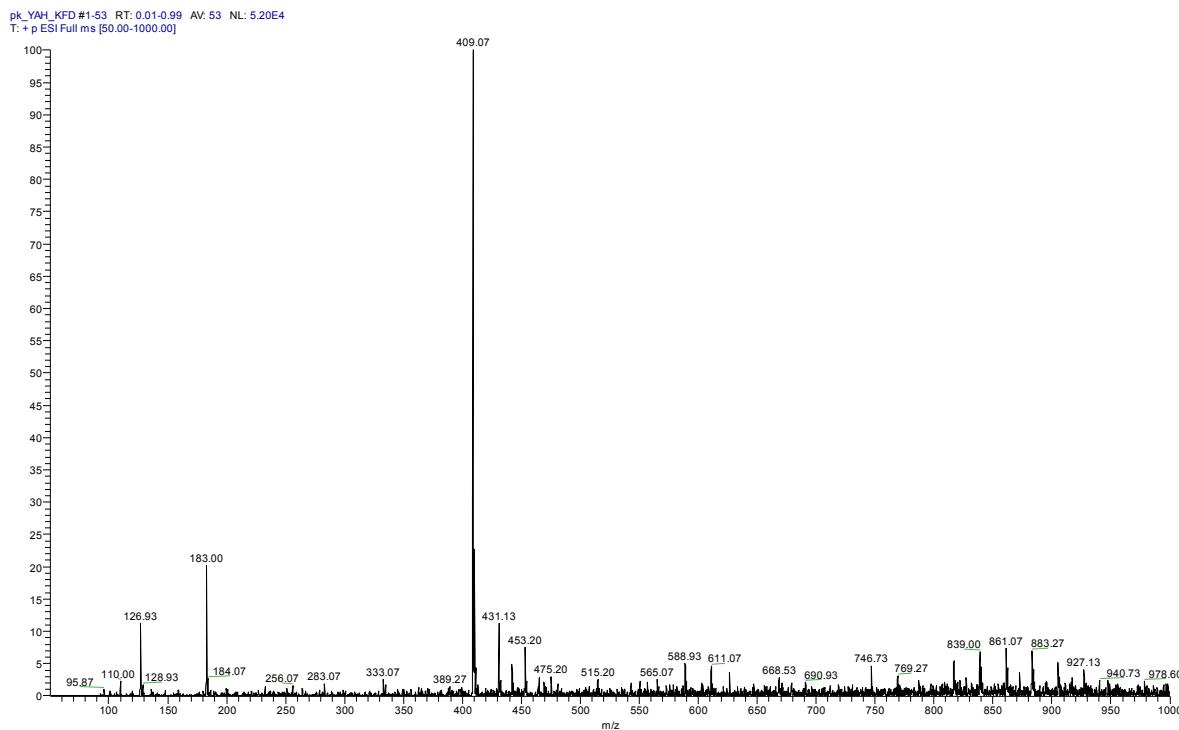




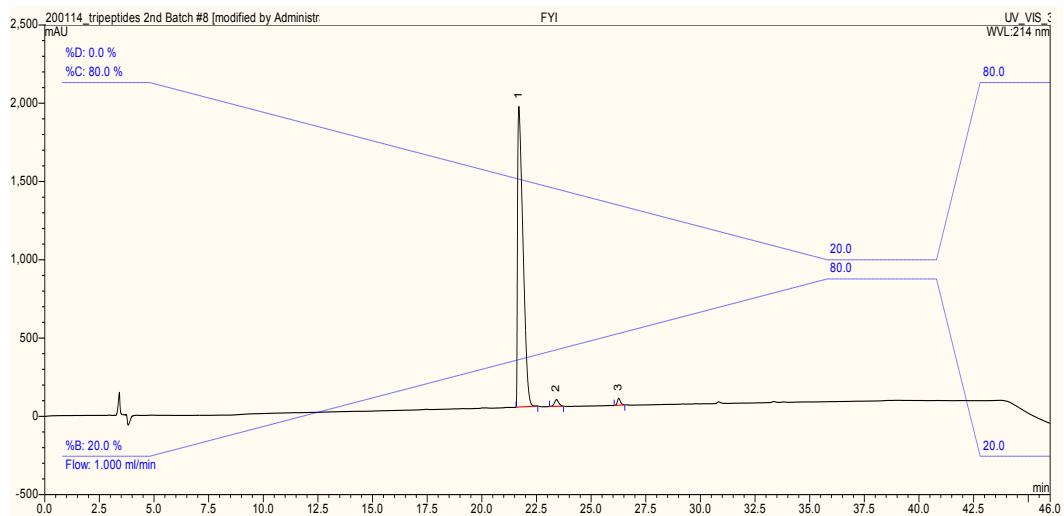
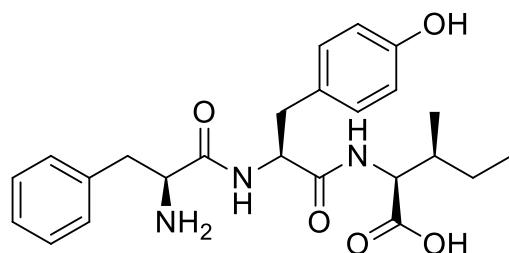
**Supplementary Figure 32.** Structure, HPLC trace, Mass Spec and NMR of KHD

**Lys-Phe-Asp (Purity: 97.8%)**

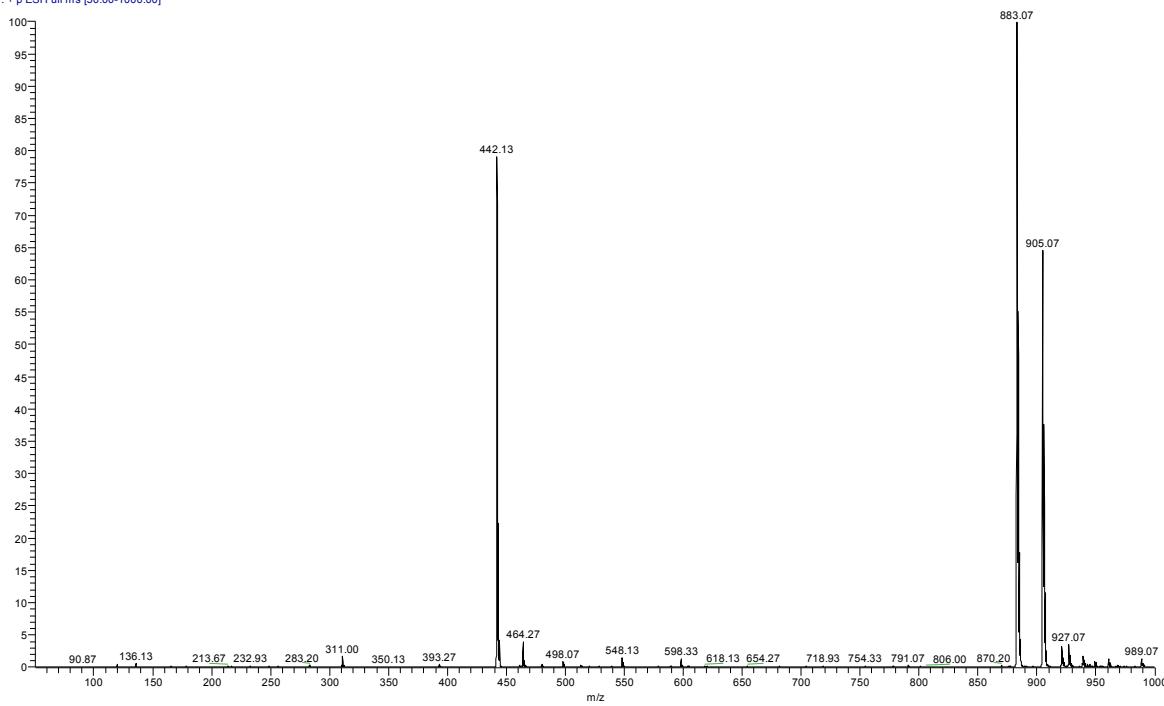


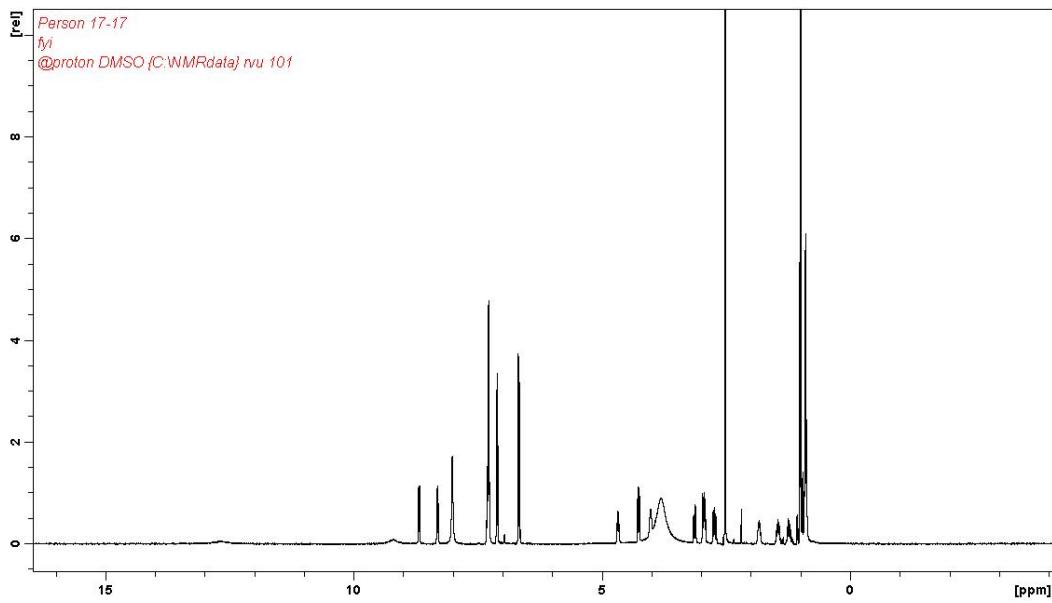


**Supplementary Figure 33.** Structure, HPLC trace, Mass Spec and NMR of KFD

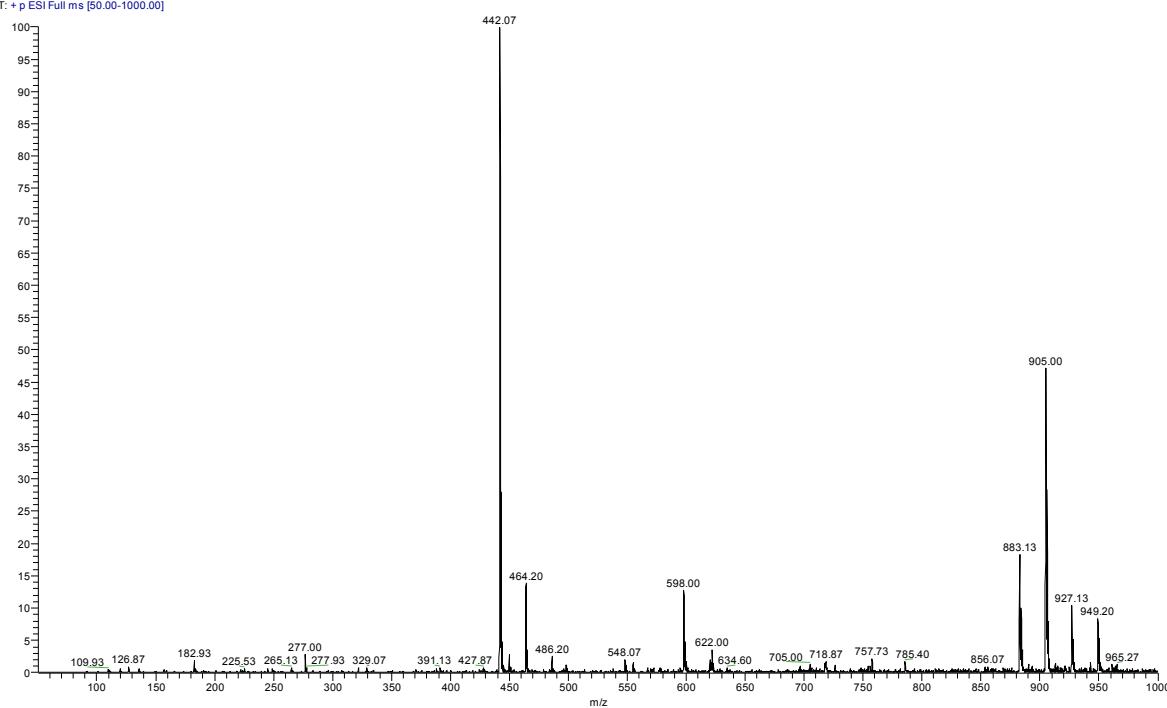
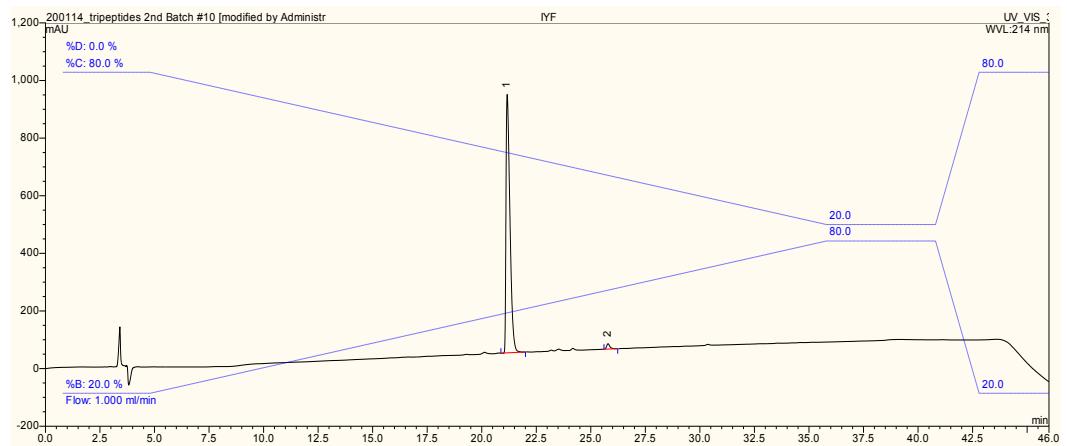
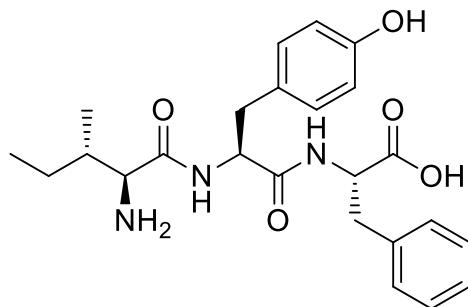
**Phe-Tyr-Ile (Purity: 97.0%)**

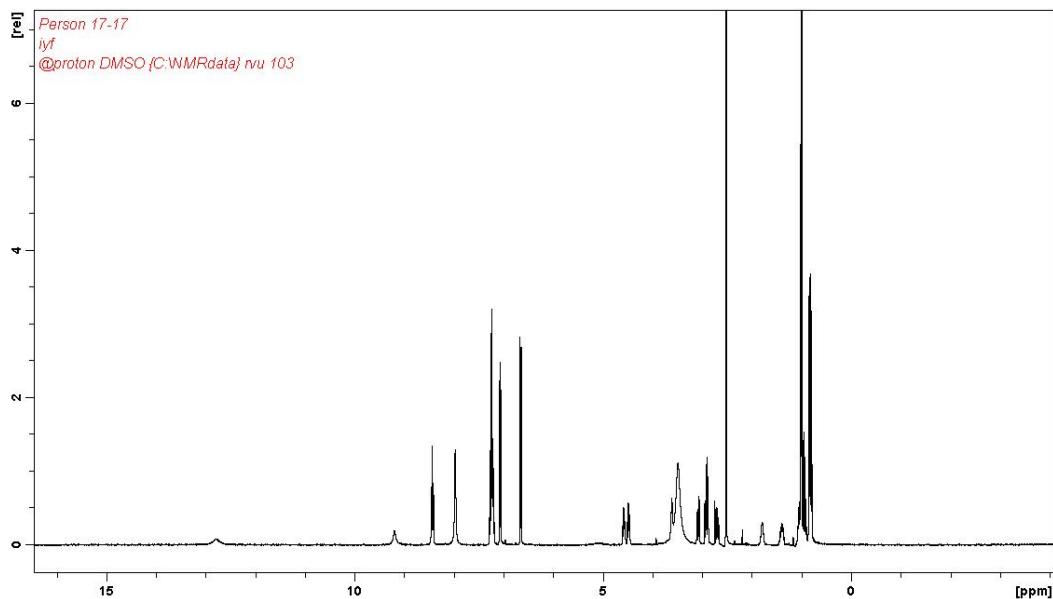
pk\_YAH\_FYI#16-48 RT: 0.28-0.94 AV: 28 NL: 8.93E5  
T: + ESI Full ms [50.00-1000.00]





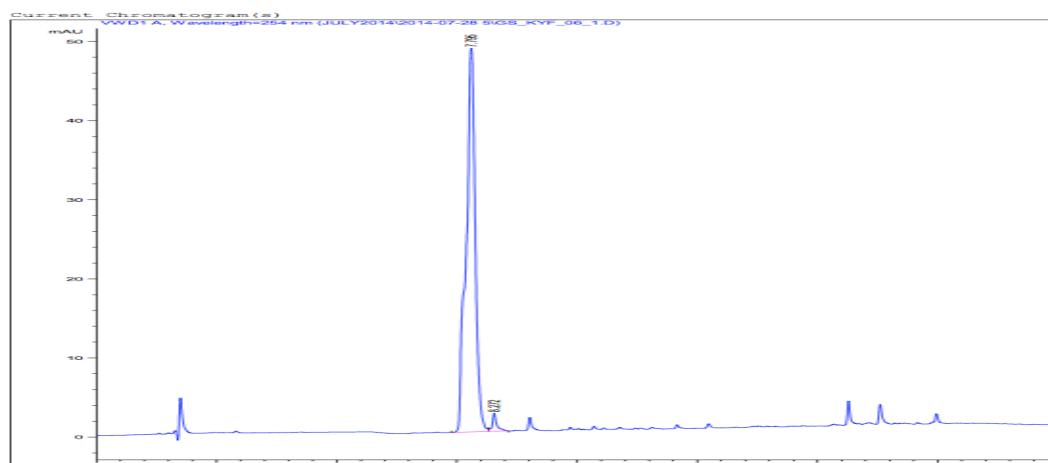
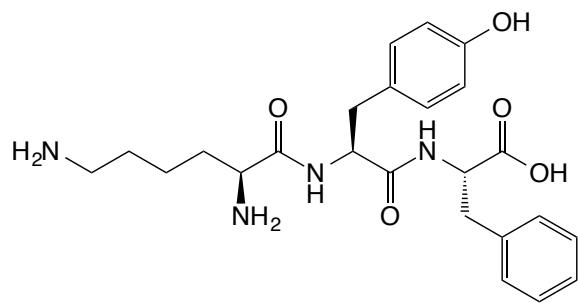
**Supplementary Figure 34.** Structure, HPLC trace, Mass Spec and NMR of FYI

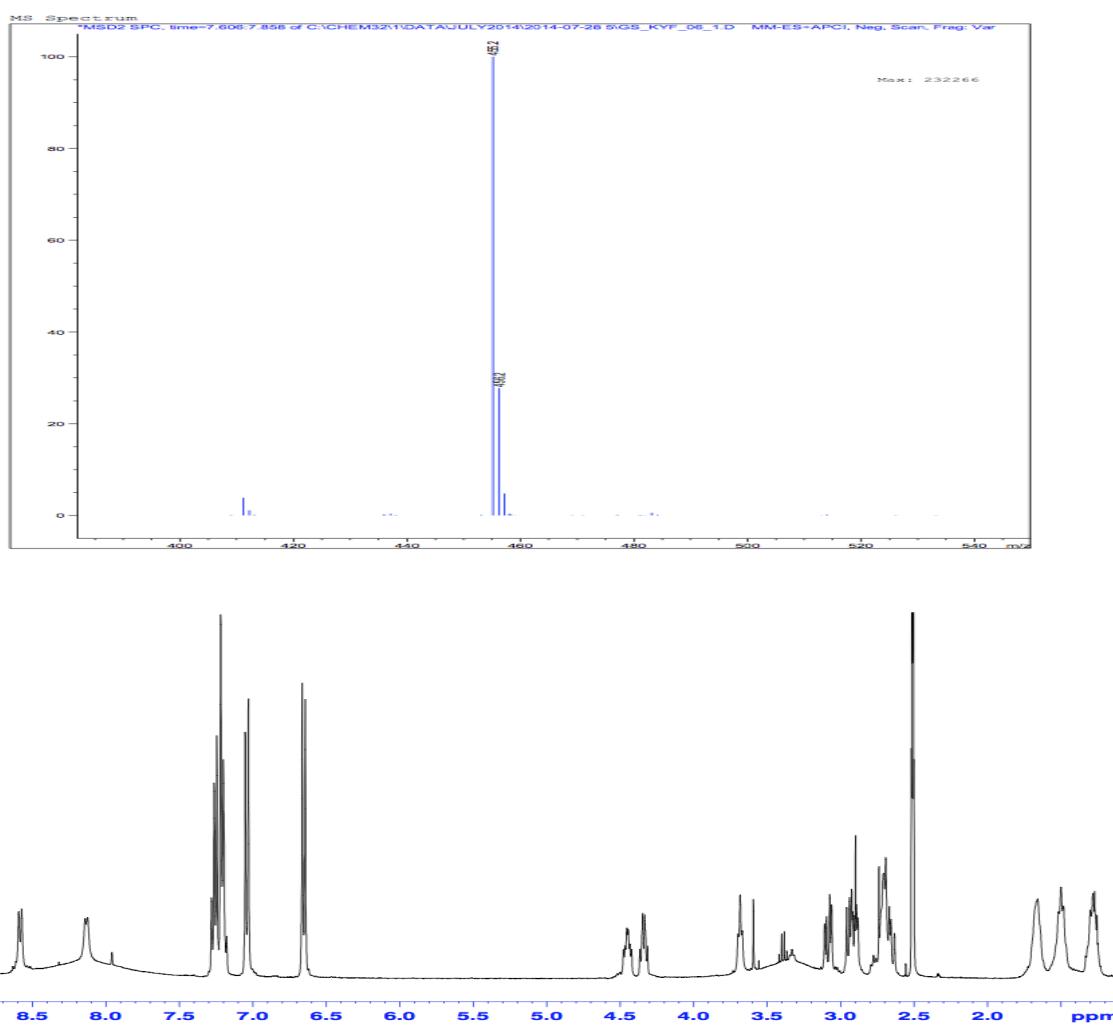
**Ile-Tyr-Phe (Purity: 98.3%)**



**Supplementary Figure 35.** Structure, HPLC trace, Mass Spec and NMR of IYF

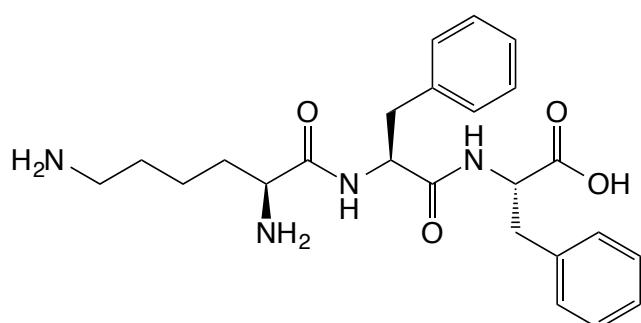
**Lys-Tyr-Phe (Purity: 94.2%)**

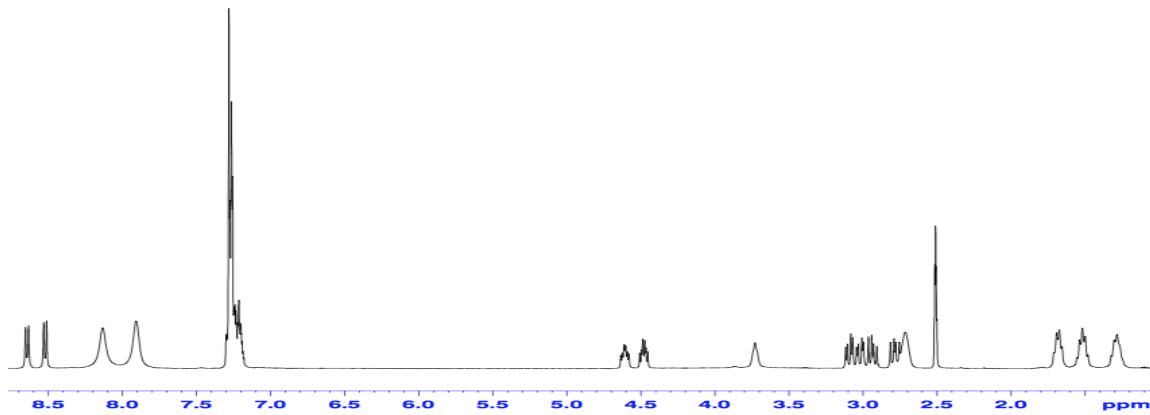
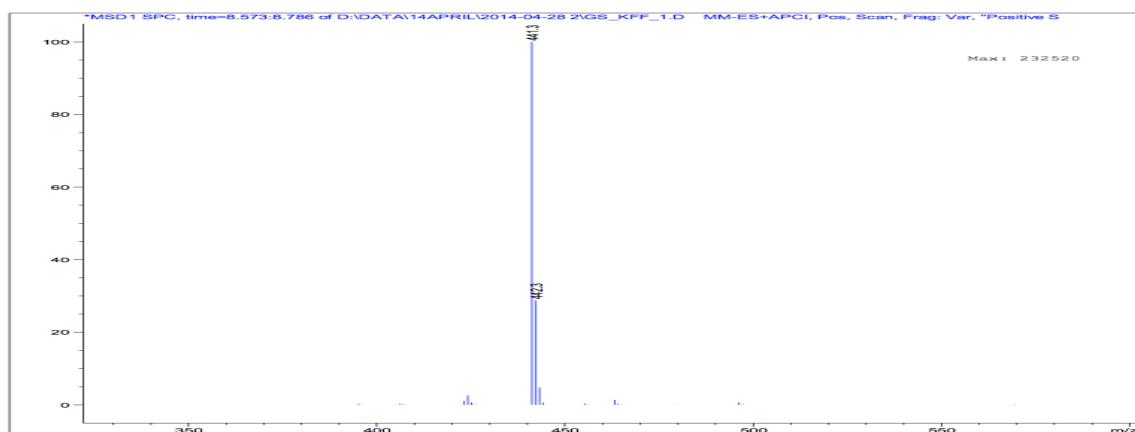
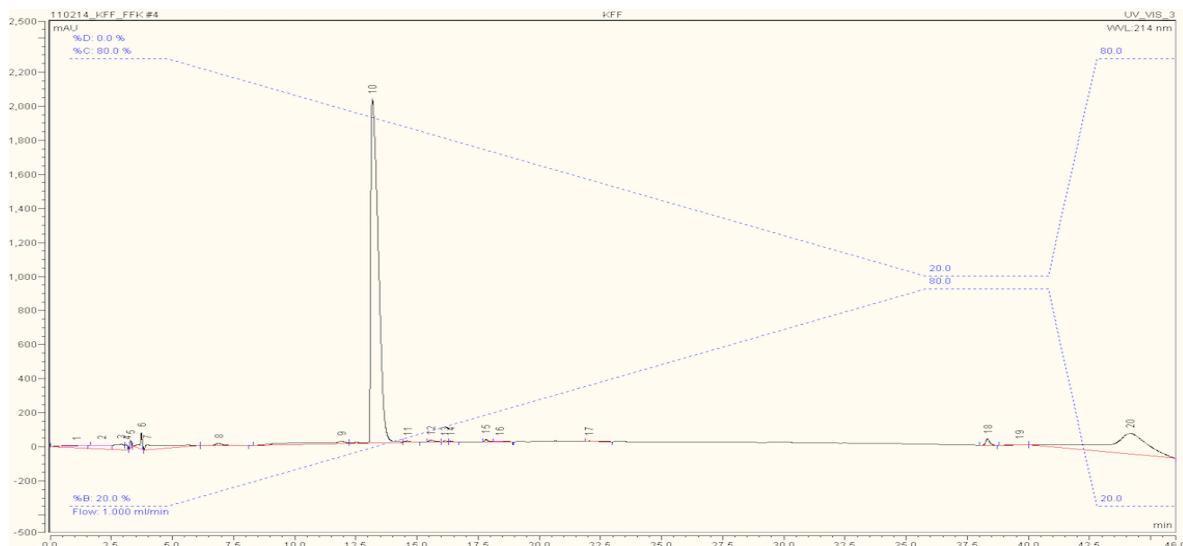


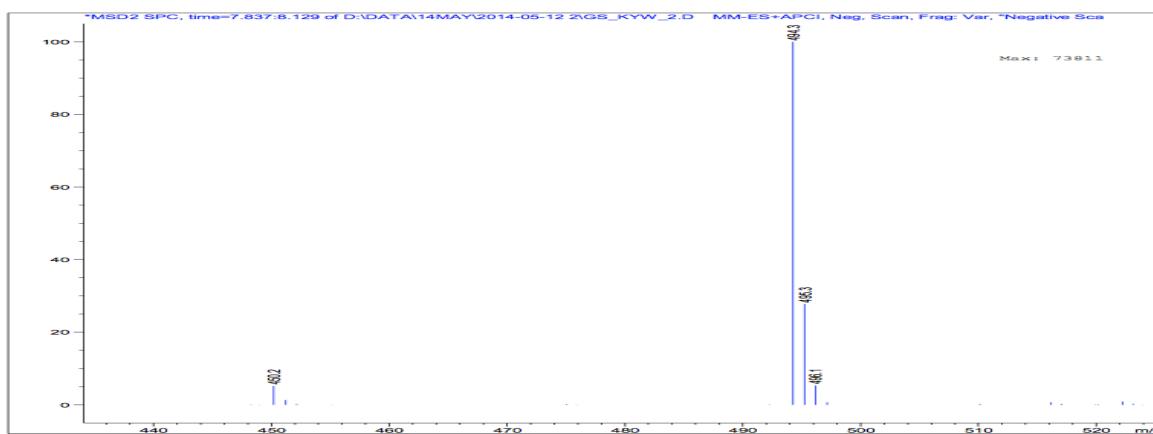
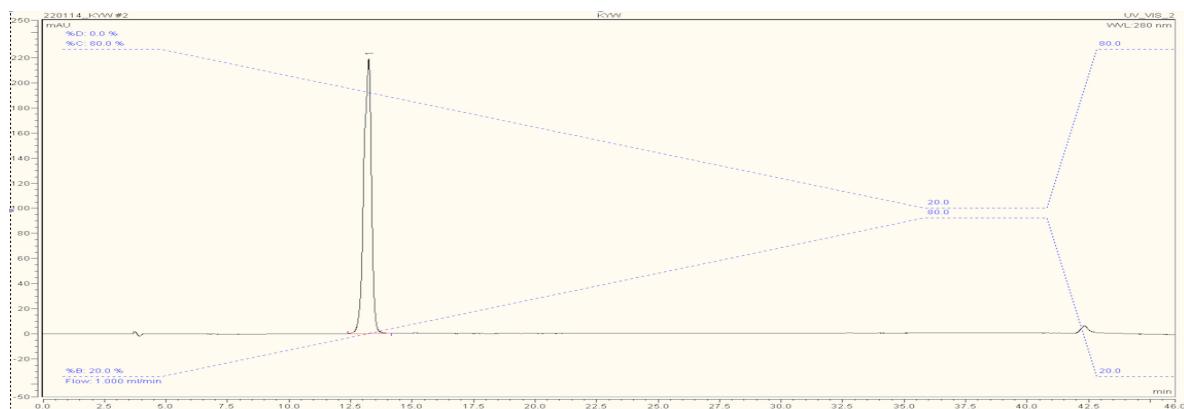
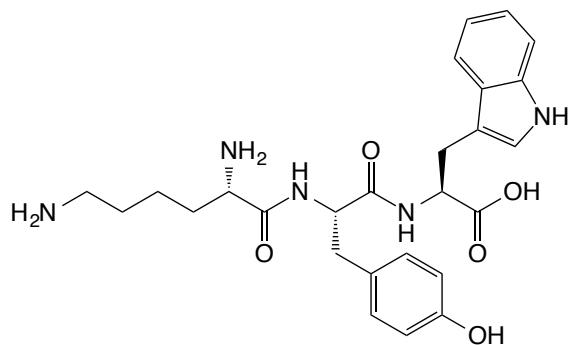


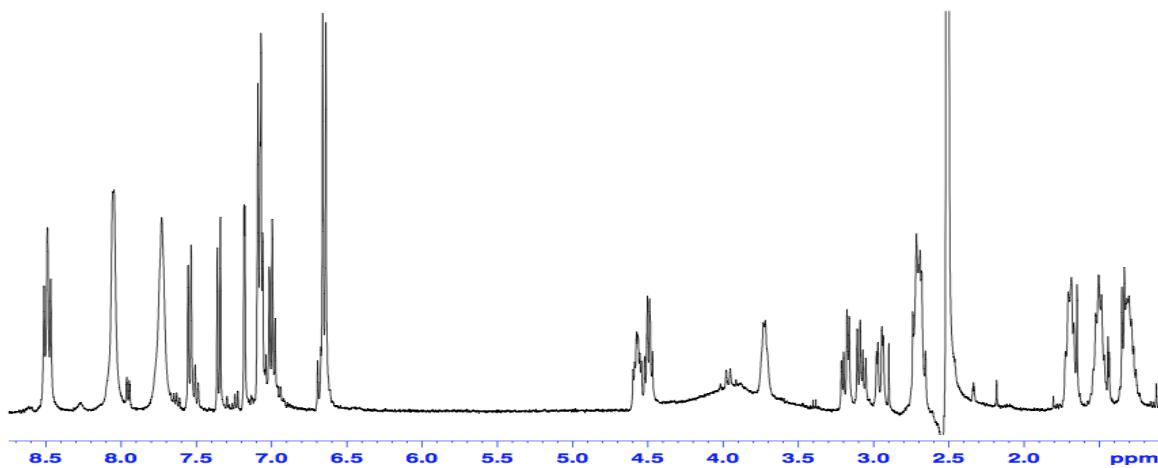
**Supplementary Figure 36.** Structure, HPLC trace, Mass Spec and NMR of KYF

**Lys-Phe-Phe (Purity: 97.7%)**

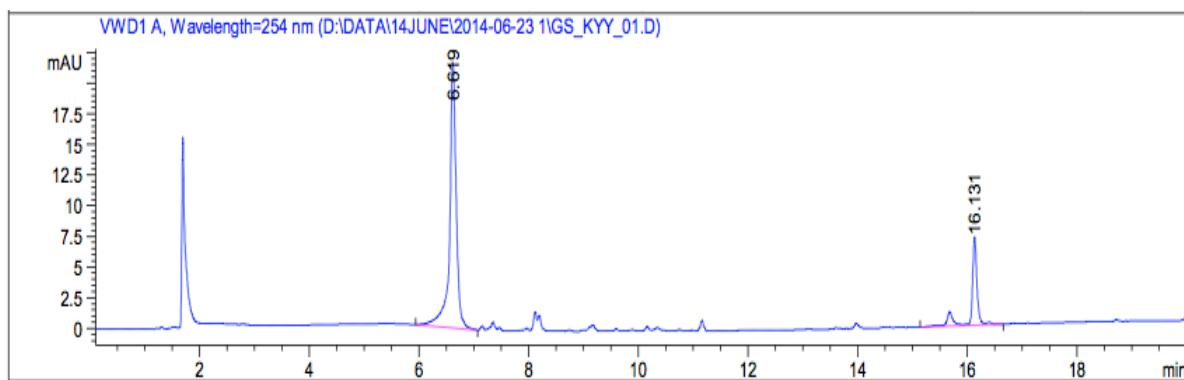
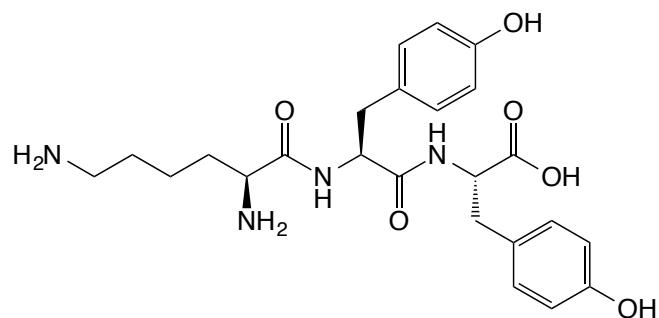


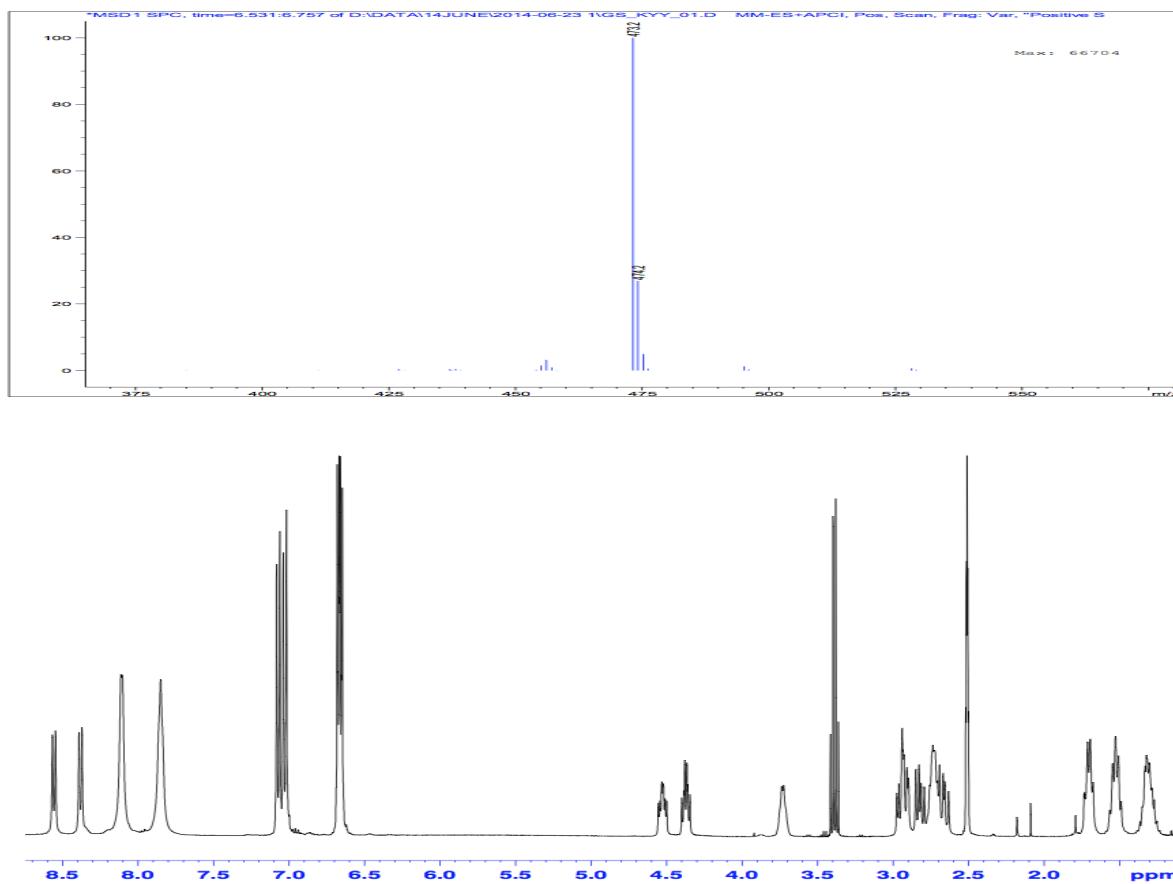


**Lys-Tyr-Trp (Purity: 99.2%)**

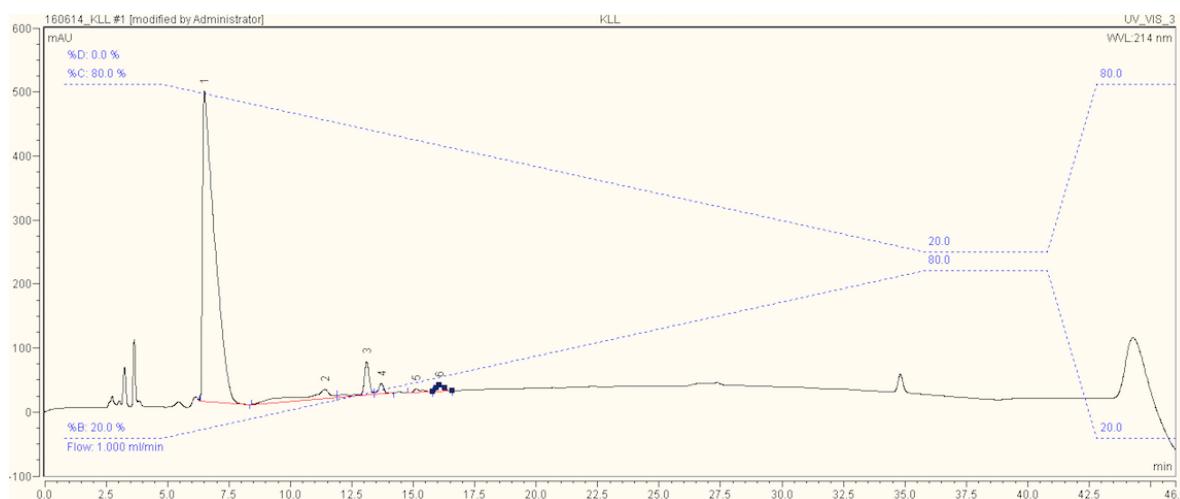
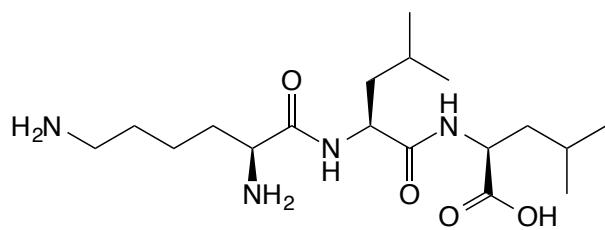


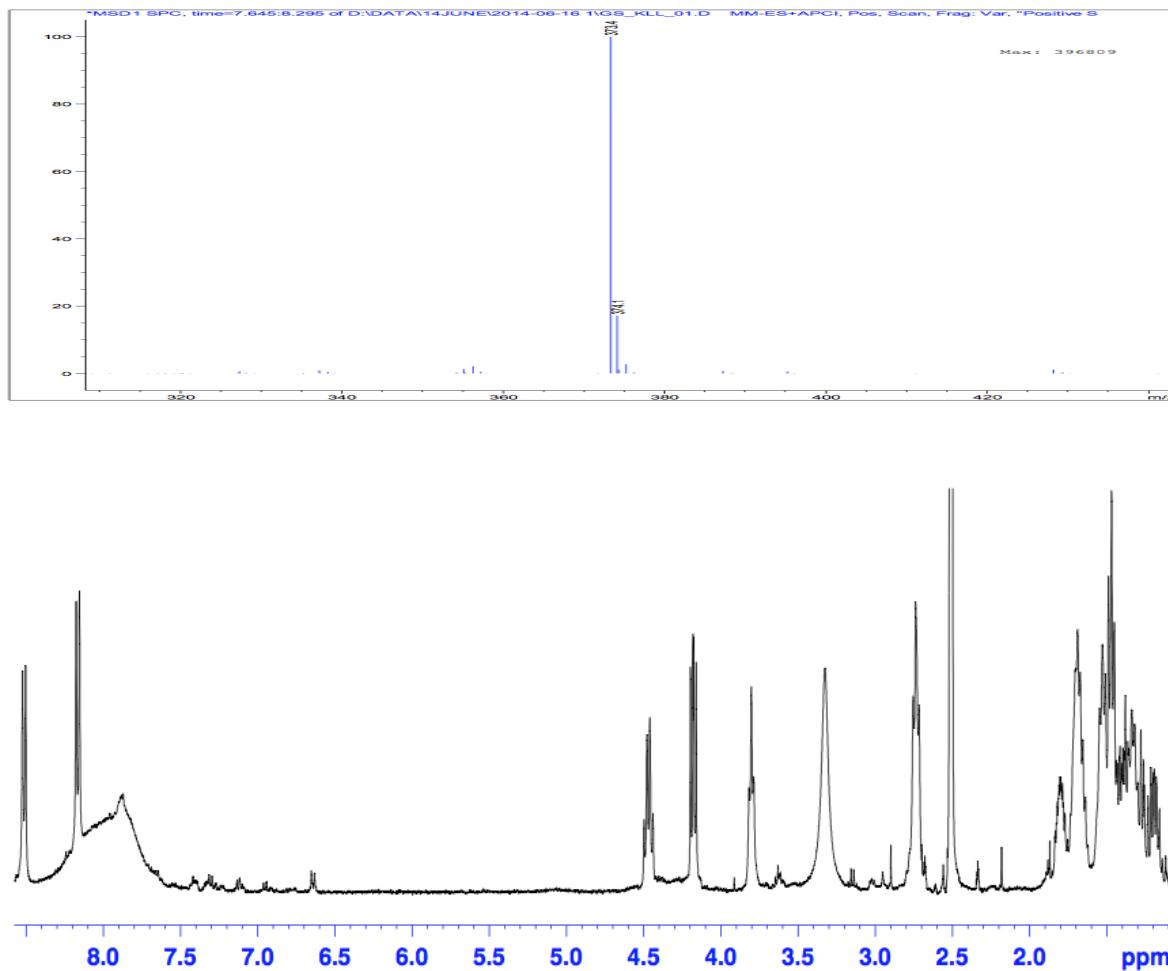
Lys-Tyr-Tyr (Purity: 84.9%)



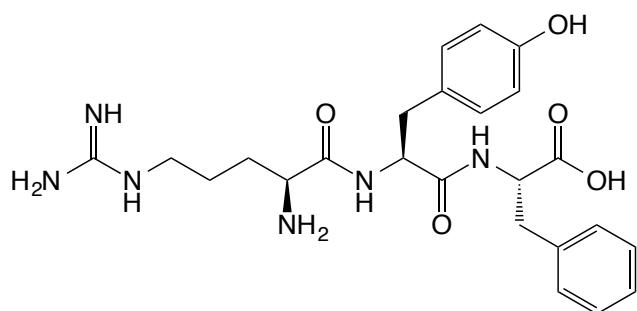


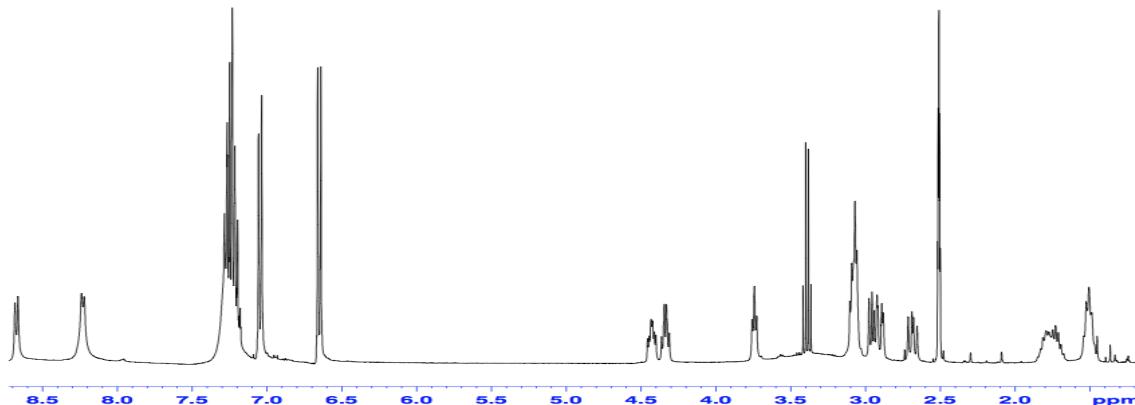
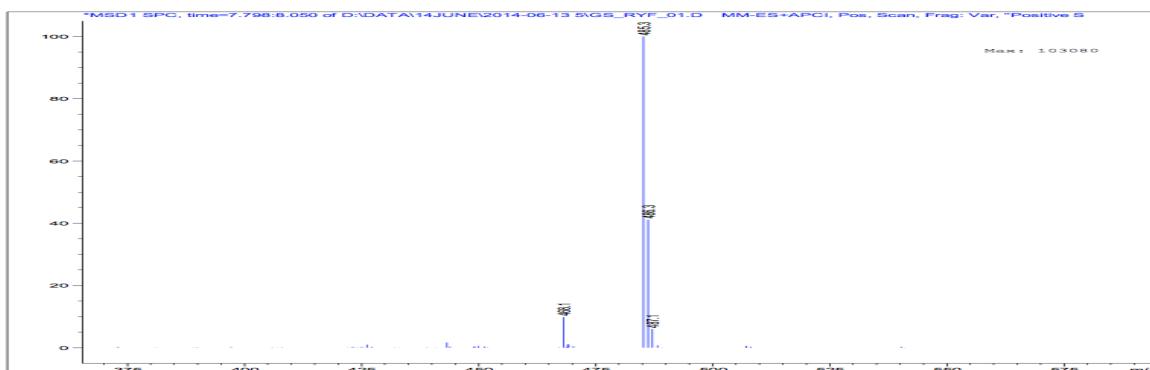
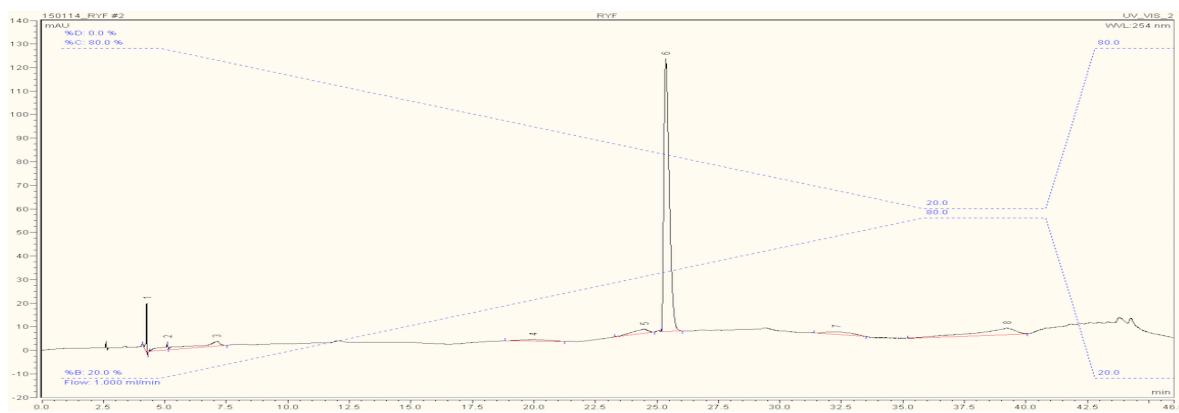
**Lys-Leu-Leu (Purity: 96.1%)**



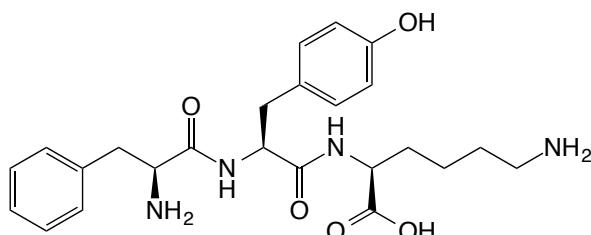


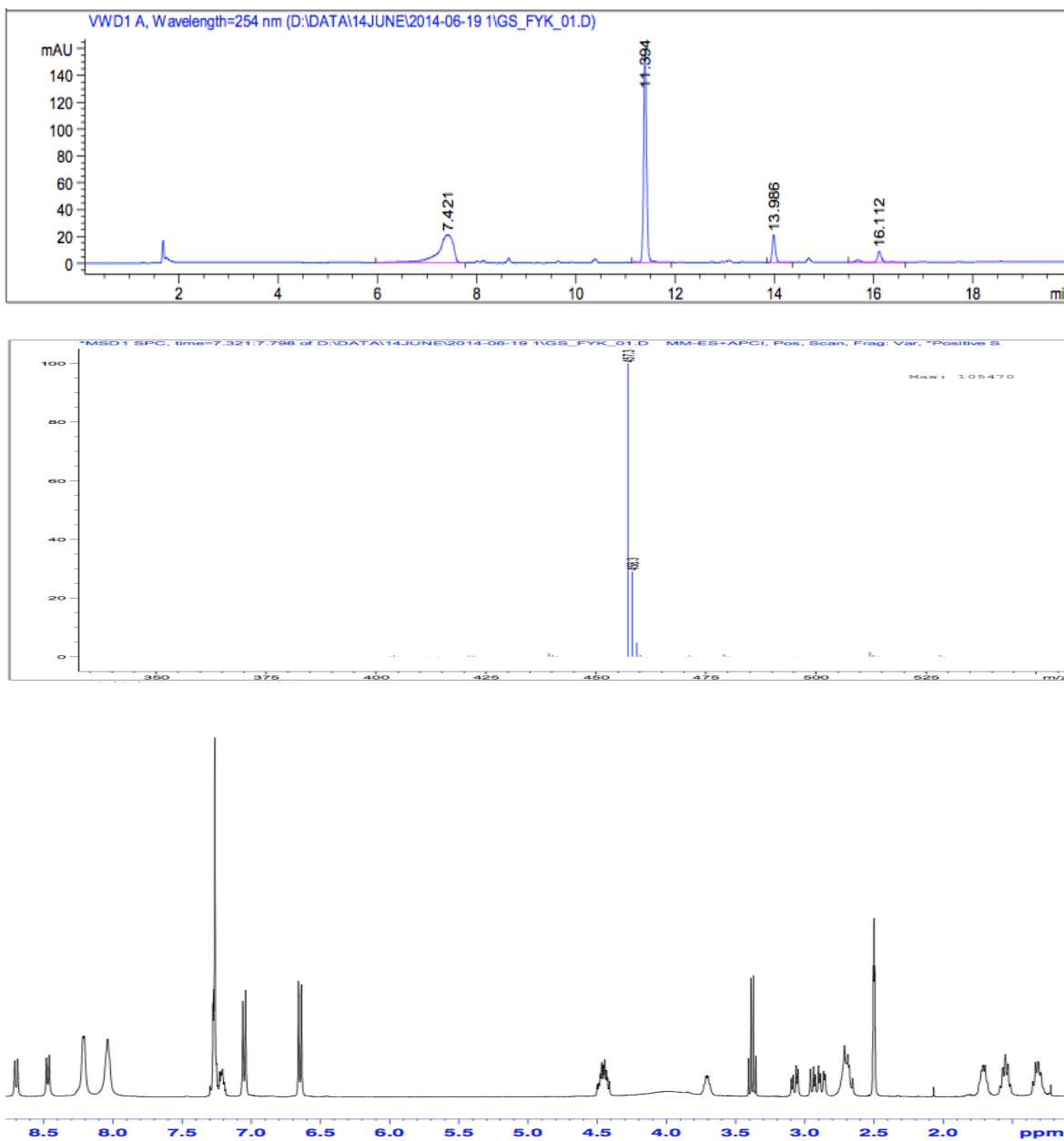
**Arg-Tyr-Phe (Purity: 93.1%)**





### Phe-Tyr-Lys (Purity: 56.3%)





### Supporting Tables

**Supplementary Table 1.** AP scores after 50, 400 and 1200. Note that the AP at 1200 ns was calculated at a different concentration and should therefore not be compared to the AP at 50 or 400 ns. Numbers higher than the initial cut-off of AP = 2 have been coloured green.

Trip.	AP 50 ns <sup>a</sup>	AP 50 ns <sup>b</sup>	AP 400 ns <sup>b</sup>	AP 1200 ns <sup>c</sup>

CFF	2.05	2.21	3.04	3.70
DFN	1.17	1.10	1.11	1.28
ECG	1.01	1.02	0.99	1.05
FFF	2.26	2.44	2.42	3.44
FFV	2.03	2.21	2.84	3.91
KFG	1.56	1.39	1.57	1.94
LFF	2.07	2.41	2.70	3.34
VFF	2.33	2.24	2.92	3.82
VYV	1.88	1.78	2.44	2.67
PFF	2.39	2.20	3.51	3.70
FYI	2.22	2.12	2.24	3.19
FHF	2.09	1.95	3.22	3.17
YFI	1.97	1.93	2.15	3.11
IYF	1.89	2.26	2.18	3.24
KYF	1.85	1.70	1.83	3.19
KFD	1.73	1.40	1.70	2.33
KHD	1.63	1.44	1.92	2.51
KFF	1.92	1.72	2.01	2.66
KYY	1.79	1.58	1.88	3.03
KYW	1.78	1.76	1.99	2.87
KLL	1.31	1.06	1.08	1.38
FYK	1.73	1.63	1.67	2.59
RYF	1.80	1.60	1.98	2.66
GGG	1.07	1.07	1.04	1.11

<sup>a</sup>: 300 peptides in standard CG water, 0.23 M, <sup>b</sup>: 300 peptides in polarizable CG water, 0.23 M, <sup>c</sup>: 1200 peptides in standard CG water, 0.14 M

**Supplementary Table 2.** Moments of inertia along x, y and z axes, the number of molecules in the largest cluster, the number of clusters, aspect ratio and observed morphology from the final frames of the 1200 ns MD simulations on 1200 peptides in a box.

Tripe	$I_x (10^6 \text{ amu}^* \text{ nm}^2)$	$I_y (10^6 \text{ amu}^* \text{ nm}^2)$	$I_z (10^6 \text{ amu}^* \text{ nm}^2)$	# mol. in largest cluster	number of clusters	Aspect ratio ( $I_z / I_x$ )	Main morphology
CFF	3.82	5.69	6.82	791	6	1.79	Spherical
DFN	0.02	0.03	0.03	31	453	2.02	N/A
ECG	0.00	0.00	0.00	4	1053	2.77	N/A
FFF	1.70	1.80	2.40	396	7	1.41	Sphere/plate
FFV	2.61	6.38	6.52	752	2	2.49	Oblong
FHF	5.14	13.83	14.54	802	8	2.83	Fibrous <sup>a</sup>
FYI	2.68	6.07	6.30	743	5	2.35	Oblong
FYK	1.78	6.14	6.38	483	6	3.59	Fibrous
GGG	0.00	0.00	0.00	21	904	3.69	N/A
IYF	0.91	1.09	1.19	328	5	1.30	Spherical
KFD	2.70	17.80	18.20	561	5	6.75	Fibrous <sup>a</sup>
KFF	0.98	5.26	5.48	406	7	5.57	Fibrous
KFG	0.38	1.15	1.27	182	44	3.37	Small/fibrous
KHD	14.66	60.67	69.42	1164	2	4.74	Fibrous <sup>a</sup>
KLL	0.18	0.84	0.90	132	221	4.93	N/A
KYF	4.36	6.60	9.19	623	5	2.11	Fibrous <sup>a</sup>
KYW	3.15	8.58	10.67	491	5	3.38	Fibrous <sup>a</sup>
KYY	1.25	2.19	2.58	362	6	2.06	Spherical
LFF	4.24	11.67	12.20	1002	3	2.88	Oblong
PFF	2.66	4.64	5.17	693	2	1.95	Spherical/oblong
RYF	1.34	6.67	6.86	406	8	5.13	Fibrous
VFF	2.68	3.85	4.22	664	4	1.57	Spherical
VYV	0.71	0.78	0.86	311	10	1.21	Small/spherical
YFI	1.24	2.18	2.31	433	8	1.86	Spherical

<sup>a</sup>: Due to branching the aspect ratio becomes unrepresentatively low

**Supplementary Table 3.** Number of hydrogen bonds averaged over the final 25 ns of the MD simulation of the backmapped simulation of 300 KYF peptides at 0.23 M. A donor-acceptor heteroatom distance cut-off of 3.3 Å and donor-hydrogen-acceptor angle cut-off of 30 degrees were used. Termini are defined as using the N-terminus nitrogen or C-terminus oxygen atoms as donor or acceptor, while backbone is defined as the two amide groups.

Total number of H-bonds	310		
Backbone (incl. Termini) – Backbone (incl termini)	151	Termini – Termini	47
		Termini – Backbone	92
		Backbone – Backbone	12
Backbone (incl. Termini) – Side Chain	143	Termini – Side Chain	117
		Backbone – Side Chain	26
Side Chain – Side Chain	16	Tyr-Lys	16

**Supplementary Table 4.** Tripeptides with AP > 2 at 50 ns

<b>pep</b>	<b>AP</b>	<b>logP</b>	<b>AP<sub>H</sub></b>	FHF	2.09	-3.31	0.11	FFV	2.03	-3.88	0.08
PFF	2.39	-3.28	0.17	ISW	2.08	-2.75	0.13	PWI	2.03	-3.07	0.10
WFL	2.38	-5.05	0.07	WMF	2.08	-4.47	0.06	WLC	2.03	-3.36	0.09
MFF	2.36	-4.09	0.12	IYW	2.08	-3.92	0.08	VWV	2.03	-3.01	0.11
VFF	2.33	-3.88	0.13	VVF	2.08	-2.63	0.13	LWY	2.03	-4.05	0.07
FFM	2.31	-4.09	0.11	PVF	2.08	-2.03	0.15	PLW	2.02	-3.20	0.10
FWF	2.28	-5.51	0.04	CWF	2.08	-3.82	0.09	LYW	2.02	-4.05	0.07
FFF	2.26	-5.13	0.06	YFF	2.08	-4.13	0.08	IIW	2.02	-4.33	0.06
WWF	2.26	-5.89	0.02	FLF	2.08	-4.67	0.06	HFF	2.02	-3.31	0.09
FWI	2.26	-4.92	0.07	PCF	2.08	-1.59	0.17	CWM	2.02	-2.78	0.11
FYI	2.22	-3.54	0.12	WCF	2.07	-3.82	0.09	LWV	2.02	-3.80	0.08
VFW	2.22	-4.26	0.09	LPF	2.07	-2.82	0.12	TFW	2.02	-3.55	0.09
PWF	2.21	-3.66	0.12	TFF	2.07	-3.17	0.11	FFA	2.02	-2.92	0.11
IFF	2.21	-4.54	0.08	LFF	2.07	-4.67	0.06	FWY	2.02	-4.51	0.06
LCF	2.18	-2.98	0.14	FWV	2.06	-4.26	0.07	TFV	2.02	-1.92	0.14
WLL	2.18	-4.59	0.07	SCW	2.06	-1.65	0.16	FWC	2.02	-3.82	0.08
SFW	2.18	-3.34	0.12	IIF	2.06	-3.95	0.08	MWW	2.01	-4.85	0.05
IFW	2.17	-4.92	0.06	FCW	2.06	-3.82	0.08	TCW	2.01	-1.86	0.14
WFF	2.17	-5.51	0.03	WFP	2.06	-3.66	0.09	CMF	2.01	-2.40	0.12
FFW	2.17	-5.51	0.03	FYF	2.06	-4.13	0.07	IFP	2.01	-2.69	0.11
IMW	2.15	-3.88	0.10	WVF	2.06	-4.26	0.07	LIF	2.01	-4.08	0.07
FWP	2.15	-3.66	0.11	YFM	2.06	-3.09	0.11	MYF	2.01	-3.09	0.10
WIW	2.15	-5.30	0.04	YLW	2.06	-4.05	0.08	PWL	2.01	-3.20	0.10
CLW	2.15	-3.36	0.12	WFG	2.05	-2.65	0.12	FLP	2.01	-2.82	0.11
YWF	2.14	-4.51	0.07	CFF	2.05	-3.44	0.10	CFV	2.01	-2.19	0.13
GFF	2.14	-2.27	0.16	VPF	2.05	-2.03	0.14	FYL	2.01	-3.67	0.08
WFT	2.13	-3.55	0.11	VMW	2.05	-3.22	0.10	WVW	2.00	-4.64	0.05
WFM	2.13	-4.47	0.07	TYF	2.05	-2.17	0.14	FFS	2.00	-2.96	0.10
IWF	2.13	-4.92	0.05	VWW	2.05	-4.64	0.06	PFY	2.00	-2.28	0.12
WFW	2.12	-5.89	0.01	CPW	2.05	-1.97	0.15	AFW	2.00	-3.30	0.09
MFY	2.12	-3.09	0.12	FVF	2.04	-3.88	0.08	LWI	2.00	-4.46	0.06
VWF	2.12	-4.26	0.08	YFW	2.04	-4.51	0.06	LYF	2.00	-3.67	0.08
PMW	2.11	-2.62	0.14	FFT	2.04	-3.17	0.10	LWF	2.00	-5.05	0.04
WFI	2.11	-4.92	0.05	VYW	2.04	-3.26	0.10	IFY	2.00	-3.54	0.08
VAW	2.10	-2.05	0.16	YFC	2.04	-2.44	0.13	KWF	2.00	-1.00	0.16
IVW	2.10	-3.67	0.10	IWW	2.04	-5.30	0.03				
FFY	2.10	-4.13	0.08	WCW	2.04	-4.20	0.07				
SWW	2.10	-3.72	0.09	VYF	2.04	-2.88	0.11				
FFC	2.10	-3.44	0.10	ILF	2.04	-4.08	0.07				
MIF	2.10	-3.50	0.10	GFW	2.04	-2.65	0.12				
FWW	2.09	-5.89	0.01	SFF	2.04	-2.96	0.11				
MFW	2.09	-4.47	0.07	WPW	2.03	-4.04	0.07				
IWI	2.09	-4.33	0.07	IFT	2.03	-2.58	0.12				
TWF	2.09	-3.55	0.10	WLT	2.03	-3.09	0.10				

**Supplementary Table 5:**  
peptides from the top 400  
AP<sub>H</sub>

	VPF	2.05	-2.03	0.144	CSY	1.82	-0.27	0.128			
	PYY	1.96	-1.28	0.143	YFC	2.04	-2.44	0.128			
	YFK	1.83	0.38	0.143	SWP	1.93	-1.49	0.128			
<b>pep</b>	<b>AP</b>	<b>logP</b>	<b>AP_H</b>								
KFD	1.73	4.73	0.187	TKD	1.58	6.69	0.143	FVV	2.07	-2.63	0.127
KWD	1.74	4.35	0.186	FKY	1.82	0.38	0.142	SGW	1.84	-0.48	0.127
HKD	1.65	6.55	0.176	KDF	1.63	4.73	0.141	ISW	2.08	-2.75	0.127
PFF	2.39	-3.28	0.174	LCF	2.18	-2.98	0.140	CFC	1.95	-1.75	0.126
KWE	1.72	4.34	0.174	TCW	2.01	-1.86	0.140	IFG	1.94	-1.68	0.126
WKE	1.71	4.34	0.169	GFY	1.95	-1.27	0.140	KVW	1.78	0.25	0.126
KHD	1.63	6.55	0.167	TGF	1.87	-0.31	0.140	YFD	1.73	1.22	0.126
PCF	2.08	-1.59	0.167	SWE	1.73	2.00	0.139	FDF	1.78	0.22	0.126
KWF	2.00	-1.00	0.163	STY	1.84	0.00	0.139	SYH	1.81	-0.14	0.126
KFW	2.00	-1.00	0.163	TYF	2.05	-2.17	0.139	SHF	1.88	-1.14	0.125
KHE	1.62	6.54	0.161	PMW	2.11	-2.62	0.139	VFG	1.87	-1.02	0.125
TSF	1.99	-1.00	0.161	PSF	1.93	-1.11	0.139	WRE	1.63	3.35	0.125
SCW	2.06	-1.65	0.161	TFV	2.02	-1.92	0.138	KDW	1.60	4.35	0.125
WKD	1.69	4.35	0.160	SKW	1.77	1.17	0.138	PYV	1.87	-1.03	0.125
KYD	1.64	5.73	0.160	SFD	1.71	2.39	0.138	TYT	1.81	-0.21	0.124
KEH	1.62	6.54	0.159	SPF	1.93	-1.11	0.138	FYI	2.22	-3.54	0.124
GFF	2.14	-2.27	0.158	PGW	1.90	-0.80	0.137	PFY	2.00	-2.28	0.124
VAW	2.10	-2.05	0.158	FFD	1.82	0.22	0.137	YPY	1.89	-1.28	0.124
SSF	1.96	-0.79	0.155	KFT	1.75	1.34	0.136	VFD	1.71	1.47	0.124
KYE	1.63	5.72	0.155	PPW	1.99	-1.81	0.136	PFV	1.97	-2.03	0.124
STF	1.97	-1.00	0.154	RWD	1.66	3.36	0.136	SFW	2.18	-3.34	0.124
SRD	1.62	5.91	0.154	TRD	1.59	5.70	0.136	WFG	2.05	-2.65	0.124
KYY	1.79	1.38	0.152	PFT	1.94	-1.32	0.135	TFC	1.91	-1.48	0.123
PVF	2.08	-2.03	0.152	FCC	1.98	-1.75	0.135	PFS	1.87	-1.11	0.123
SKD	1.59	6.90	0.152	RFY	1.87	-0.61	0.135	STH	1.74	0.82	0.123
TFP	2.00	-1.32	0.152	PIY	1.97	-1.69	0.134	FGS	1.79	-0.10	0.123
RHD	1.63	5.56	0.151	RFD	1.64	3.74	0.134	VYY	1.95	-1.88	0.123
KYF	1.85	0.38	0.151	SHK	1.66	3.37	0.134	FRD	1.61	3.74	0.123
FKF	1.92	-0.62	0.150	SHN	1.74	1.42	0.133	VSW	1.97	-2.09	0.122
KFF	1.92	-0.62	0.150	YKD	1.58	5.73	0.133	VFE	1.70	1.46	0.122
WFD	1.88	-0.16	0.149	SGF	1.83	-0.10	0.133	KDY	1.55	5.73	0.122
WRD	1.69	3.36	0.149	HRD	1.58	5.56	0.132	CMF	2.01	-2.40	0.122
FKD	1.65	4.73	0.148	VVF	2.08	-2.63	0.131	PSY	1.79	-0.11	0.122
PHF	2.00	-1.46	0.148	TSH	1.76	0.82	0.131	YGF	1.88	-1.27	0.122
KFY	1.84	0.38	0.147	FRY	1.86	-0.61	0.130	MFY	2.12	-3.09	0.122
KFE	1.64	4.72	0.147	TTF	1.91	-1.21	0.130	MFF	2.36	-4.09	0.122
PPF	1.99	-1.43	0.147	FKW	1.89	-1.00	0.130	TFD	1.67	2.18	0.121
SFE	1.73	2.38	0.147	YPV	1.89	-1.03	0.129	SHD	1.59	4.21	0.121
SNH	1.78	1.42	0.147	PTF	1.91	-1.32	0.128	SWK	1.71	1.17	0.121
WKF	1.94	-1.00	0.146	VFF	2.33	-3.88	0.128	RDF	1.61	3.74	0.121
CPW	2.05	-1.97	0.145	CFV	2.01	-2.19	0.128	IFT	2.03	-2.58	0.121
SYS	1.84	0.21	0.145	SFY	1.98	-1.96	0.128	LPF	2.07	-2.82	0.121
				PCW	1.98	-1.97	0.128	VCF	1.98	-2.19	0.121

LFD	1.74	0.68	0.120	VFC	1.96	-2.19	0.116	PSH	1.71	0.71	0.113
KPF	1.71	1.23	0.120	TSY	1.76	0.00	0.116	SCF	1.85	-1.27	0.113
TYD	1.63	3.18	0.120	WDF	1.77	-0.16	0.116	KFV	1.71	0.63	0.112
RYD	1.58	4.74	0.120	FMC	1.98	-2.40	0.116	REY	1.56	4.73	0.112
TKF	1.70	1.34	0.120	LPY	1.91	-1.82	0.115	PHS	1.71	0.71	0.112
SVW	1.96	-2.09	0.120	FKE	1.56	4.72	0.115	FFR	1.88	-1.61	0.112
TWC	1.94	-1.86	0.120	YFG	1.86	-1.27	0.115	VYF	2.04	-2.88	0.112
CTW	1.94	-1.86	0.120	THS	1.71	0.82	0.115	FFE	1.74	0.21	0.112
KWM	1.77	0.04	0.120	WEF	1.77	-0.17	0.115	FSP	1.83	-1.11	0.112
GFW	2.04	-2.65	0.120	YCC	1.81	-0.75	0.115	KWT	1.69	0.96	0.112
GFT	1.80	-0.31	0.120	WSY	1.97	-2.34	0.115	CWM	2.02	-2.78	0.112
KYW	1.78	0.00	0.120	LSF	1.99	-2.50	0.115	WPP	1.90	-1.81	0.112
VAF	1.91	-1.67	0.119	KDH	1.52	6.55	0.115	YKF	1.73	0.38	0.112
HRE	1.55	5.55	0.119	IYC	1.91	-1.85	0.115	TSW	1.85	-1.38	0.112
IYY	2.02	-2.54	0.119	FSM	1.92	-1.92	0.115	RHE	1.53	5.55	0.111
PSW	1.89	-1.49	0.119	CPF	1.89	-1.59	0.115	GFL	1.89	-1.81	0.111
VWK	1.76	0.25	0.119	SWT	1.87	-1.38	0.115	KEY	1.53	5.72	0.111
SGY	1.72	0.90	0.119	KMW	1.76	0.04	0.115	FGF	1.95	-2.27	0.111
HKE	1.53	6.54	0.118	ASW	1.84	-1.13	0.115	FVS	1.88	-1.71	0.111
AFY	1.94	-1.92	0.118	FYT	1.95	-2.17	0.115	CTF	1.86	-1.48	0.111
SYY	1.84	-0.96	0.118	KIW	1.79	-0.41	0.115	WWK	1.85	-1.38	0.111
RWE	1.61	3.35	0.118	THF	1.86	-1.35	0.114	KIF	1.75	-0.03	0.111
CYH	1.82	-0.62	0.118	PKW	1.71	0.85	0.114	WFK	1.82	-1.00	0.111
SCY	1.79	-0.27	0.118	MHF	1.96	-2.27	0.114	KWL	1.78	-0.54	0.111
CFG	1.81	-0.58	0.118	VWH	1.98	-2.44	0.114	FWE	1.76	-0.17	0.111
KEW	1.58	4.34	0.118	VGW	1.86	-1.40	0.114	RFF	1.87	-1.61	0.111
KWC	1.73	0.69	0.118	FYK	1.73	0.38	0.114	SFG	1.75	-0.10	0.111
VFP	1.95	-2.03	0.118	CWP	1.92	-1.97	0.114	SFC	1.84	-1.27	0.111
SHS	1.71	1.03	0.118	KTF	1.68	1.34	0.114	PHV	1.76	-0.21	0.111
SWH	1.89	-1.52	0.118	EFW	1.77	-0.17	0.114	STS	1.68	1.17	0.111
PLY	1.92	-1.82	0.117	FFM	2.31	-4.09	0.114	KHW	1.70	0.82	0.111
SKE	1.52	6.89	0.117	RYF	1.80	-0.61	0.114	IGY	1.79	-0.68	0.111
TFT	1.86	-1.21	0.117	GFI	1.89	-1.68	0.114	VFH	1.92	-2.06	0.111
HKF	1.70	1.20	0.117	FRE	1.59	3.73	0.114	KSF	1.66	1.55	0.110
FTC	1.88	-1.48	0.117	PWV	1.97	-2.41	0.114	CYS	1.76	-0.27	0.110
TYI	1.89	-1.58	0.117	CGY	1.73	0.42	0.113	SKH	1.59	3.37	0.110
CLW	2.15	-3.36	0.117	PFA	1.83	-1.07	0.113	FYP	1.94	-2.28	0.110
FIS	1.98	-2.37	0.116	SWC	1.89	-1.65	0.113	YWD	1.70	0.84	0.110
PWH	1.92	-1.84	0.116	PFC	1.88	-1.59	0.113	TGS	1.65	1.86	0.110
VFK	1.73	0.63	0.116	KFH	1.69	1.20	0.113	VKW	1.73	0.25	0.110
KLF	1.77	-0.16	0.116	VFS	1.89	-1.71	0.113	TWD	1.65	1.80	0.110
CFD	1.66	1.91	0.116	CFY	1.98	-2.44	0.113	FYD	1.68	1.22	0.110
PWF	2.21	-3.66	0.116	VYV	1.88	-1.63	0.113	KFS	1.66	1.55	0.110
PVW	1.99	-2.41	0.116	PTY	1.77	-0.32	0.113	RDW	1.59	3.36	0.110
CYF	1.99	-2.44	0.116	IFP	2.01	-2.69	0.113	SSY	1.73	0.21	0.110
PCY	1.80	-0.59	0.116	SHT	1.71	0.82	0.113	PYT	1.76	-0.32	0.110

WYT	1.98	-2.55	0.110	FFA	2.02	-2.92	0.107	MFD	1.65	1.26	0.104
PAW	1.85	-1.45	0.109	WFE	1.74	-0.17	0.107	IKW	1.75	-0.41	0.104
FYC	1.96	-2.44	0.109	WFT	2.13	-3.55	0.107	WSS	1.80	-1.17	0.104
WGT	1.79	-0.69	0.109	YFE	1.66	1.21	0.106	VHS	1.71	0.11	0.104
VTW	1.94	-2.30	0.109	GVF	1.80	-1.02	0.106	TTH	1.68	0.61	0.104
FGC	1.78	-0.58	0.109	CCY	1.78	-0.75	0.106	SYP	1.73	-0.11	0.104
SFF	2.04	-2.96	0.109	TFL	1.98	-2.71	0.106	WKS	1.66	1.17	0.104
KSW	1.67	1.17	0.109	SGH	1.64	1.72	0.106	FGI	1.85	-1.68	0.104
YFM	2.06	-3.09	0.109	RYE	1.54	4.73	0.106	VMW	2.05	-3.22	0.103
WST	1.84	-1.38	0.109	IFA	1.93	-2.33	0.106	TKW	1.67	0.96	0.103
IPY	1.87	-1.69	0.109	FGV	1.80	-1.02	0.106	SWA	1.80	-1.13	0.103
TWY	1.97	-2.55	0.109	FKS	1.65	1.55	0.106	SYK	1.60	2.55	0.103
YYI	1.97	-2.54	0.109	VWV	2.03	-3.01	0.106	FFT	2.04	-3.17	0.103
PWT	1.87	-1.70	0.109	CYW	2.00	-2.82	0.106	SNY	1.68	0.60	0.103
WHP	1.89	-1.84	0.109	YFS	1.89	-1.96	0.106	SNF	1.74	-0.40	0.103
VKF	1.70	0.63	0.109	FSV	1.86	-1.71	0.106	PAF	1.79	-1.07	0.103
PFG	1.76	-0.42	0.109	TDW	1.63	1.80	0.105	GWY	1.84	-1.65	0.103
SVY	1.78	-0.71	0.108	SIY	1.83	-1.37	0.105	DKH	1.49	6.55	0.103
TFF	2.07	-3.17	0.108	KWW	1.83	-1.38	0.105	YCT	1.75	-0.48	0.103
FLP	2.01	-2.82	0.108	FFG	1.92	-2.27	0.105	CYT	1.75	-0.48	0.103
CYC	1.79	-0.75	0.108	KIY	1.67	0.97	0.105	FFS	2.00	-2.96	0.103
CGF	1.77	-0.58	0.108	FWP	2.15	-3.66	0.105	SYF	1.87	-1.96	0.103
RWY	1.81	-0.99	0.108	CSF	1.82	-1.27	0.105	SVF	1.85	-1.71	0.103
PWC	1.90	-1.97	0.108	EKW	1.55	4.34	0.105	WKI	1.74	-0.41	0.103
PPY	1.76	-0.43	0.108	CWD	1.64	1.53	0.105	PIF	1.96	-2.69	0.103
PYS	1.74	-0.11	0.108	FSL	1.95	-2.50	0.105	AFC	1.80	-1.23	0.103
SWN	1.79	-0.78	0.108	FTP	1.82	-1.32	0.105	TYS	1.72	0.00	0.103
WEK	1.55	4.34	0.108	FFC	2.10	-3.44	0.105	RHS	1.60	2.38	0.103
FHF	2.09	-3.31	0.108	TWG	1.77	-0.69	0.105	SYW	1.92	-2.34	0.103
KEF	1.54	4.72	0.108	FTL	1.97	-2.71	0.105	KWH	1.67	0.82	0.103
CKW	1.69	0.69	0.108	TKY	1.61	2.34	0.104	PYM	1.80	-1.24	0.103
RFE	1.57	3.73	0.108	PWI	2.03	-3.07	0.104	FPG	1.74	-0.42	0.103
KHF	1.67	1.20	0.108	SRW	1.71	0.18	0.104	STT	1.66	0.96	0.102
SYT	1.73	0.00	0.107	KFI	1.72	-0.03	0.104	TFK	1.64	1.34	0.102
TSS	1.67	1.17	0.107	WLT	2.03	-3.09	0.104	VYT	1.78	-0.92	0.102
FWD	1.74	-0.16	0.107	SMY	1.78	-0.92	0.104	FTI	1.95	-2.58	0.102
TKH	1.59	3.16	0.107	HDK	1.49	6.55	0.104	NSW	1.77	-0.78	0.102
FRF	1.86	-1.61	0.107	THT	1.69	0.61	0.104	TIY	1.83	-1.58	0.102
SRE	1.51	5.90	0.107	KSD	1.48	6.90	0.104	SFK	1.63	1.55	0.102
SHP	1.69	0.71	0.107	WKT	1.67	0.96	0.104	MIF	2.10	-3.50	0.102
SGS	1.63	2.07	0.107	FYG	1.81	-1.27	0.104	CWC	1.89	-2.13	0.102
SFH	1.81	-1.14	0.107	GCW	1.79	-0.96	0.104				
WKY	1.73	0.00	0.107	SRF	1.69	0.56	0.104				

**References for the SI:**

1. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
2. *martinize.py*. at <<http://md.chem.rug.nl/cgmartini/index.php/downloads/tools/204-martinize>>
3. Fleming, S. *et al.* Assessing the Utility of Infrared Spectroscopy as a Structural Diagnostic Tool for  $\beta$ -Sheets in Self-Assembling Aromatic Peptide Amphiphiles. *Langmuir* **29**, 9510–9515 (2013).
4. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
5. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
6. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
7. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
8. Marrink, S. J., de Vries, A. H. & Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
9. Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D. & Marrink, S. J. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput Biol* **6**, e1000810 (2010).
10. Wassenaar, T. A., Pluhackova, K., Böckmann, R. A., Marrink, S. J. & Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **10**, 676–690 (2014).
11. Barth, A. & Zscherp, C. What Vibrations Tell About Proteins. *Q. Rev. Biophys.* **35**, 369–430 (2002).
12. Xu, D., Tsai, C. J. & Nussinov, R. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng.* **10**, 999–1012 (1997).