Electronic Supplementary Material (ESI) for Soft Matter.

Supporting information

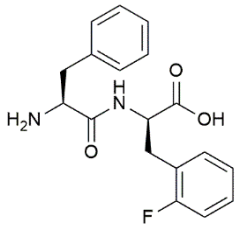
Single-atom substitution enables supramolecular diversity from dipeptide building blocks

E. Scarel, O. Bellotto, P. Rozhin, S. Kralj, M. Tortora, A. V. Vargiu, R. De Zorzi, B. Rossi,* and S. Marchesan*

Email: barbara.rossi@elettra.eu, smarchesan@units.it

Table of contents

S1. L-Phe-D-(2F)-Phe spectroscopic data



^{1}H NMR (400 MHz, DMSO-d_6) δ (ppm): **8.90** (1H, d, *J* = 8.8 Hz, NH), **7.27** (5H, m, Ar), **7.16** (1H, ddd, *J* = 1.2Hz, *J* = 8.8 Hz, *J* = 10.4 Hz, ArF), **7.11** (1H, ddd, *J* = 1.2 Hz, *J* = 7.4 Hz, *J* = 7.4 Hz, ArF), **7.06** (2H, m, ArF), **4.61** (1H, ddd, *J* = 5.2 Hz, *J* = 8.8 Hz, *J* = 14.0 Hz, CHαF), **3.97** (1H, dd, *J* = 4.4 Hz, *J* = 8.8 Hz, CHα), **3.17** (1H, dd, *J* = 5.2 Hz, *J* = 14.0 Hz, CHβF), **2.85 1H,** (dd, *J* = 4.4 Hz, *J* = 14.0 Hz, CHβF), **2.85** (1H, dd, *J* = 8.8 Hz, *J* = 14.0 Hz, CHβ), **2.59** (1H, dd, *J* = 8.8 Hz, *J* = 14.0 Hz, CHβ). **^{13}C NMR** (100 MHz, DMSO-d_6) δ (ppm): **172.2** (1C, COOH), **168.1** (1C, CONHR), **160.8** (1C, d, ^{1}J = 242.6 Hz, ArF C_1), **134.7** (1C, Ar C_1), **131.9** (1C, d, ^{3}J = 4.4 Hz, ArF C_5), **129.4** (2C, Ar, $C_{3,5}$), **129.0** (1C, d, ^{3}J = 8.1 Hz, ArF C_3), **128.5** (2C, Ar $C_{2,6}$), **127.1** (1C, Ar C_4), **124.2** (d, ^{4}J = 3.6 Hz, 1C, ArF C_4), **123.9** (1C, d, ^{2}J = 15.2 Hz, ArF C_2), **115.2** (1C, d, ^{2}J = 21.5 Hz, ArF C_6), **53.3** (1C, Cβ), **52.1** (1C, Cβ), **37.0** (1C, Cα), **31.0** (1C, Cα). **MS (ESI): $C_{18}H_{19}FN_2O_3$** exact mass: 330.14 m/z 331.2 $(M+H)^+$, 329.2 $(M-H)^-$

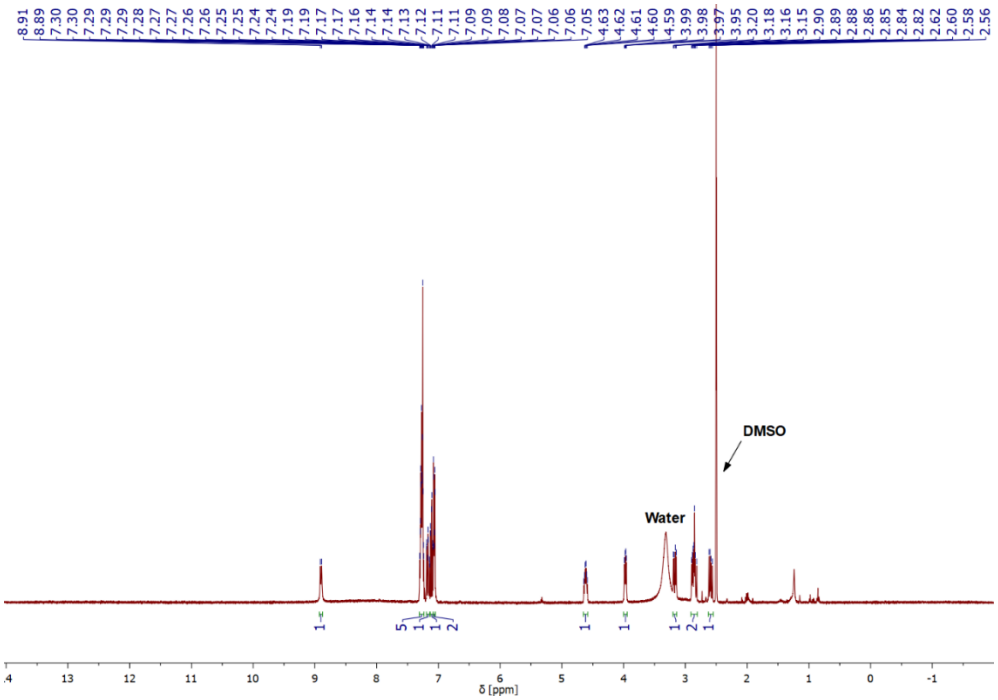


Fig. S1.^{1}H NMR spectrum of L-Phe-D-(2F)-Phe.

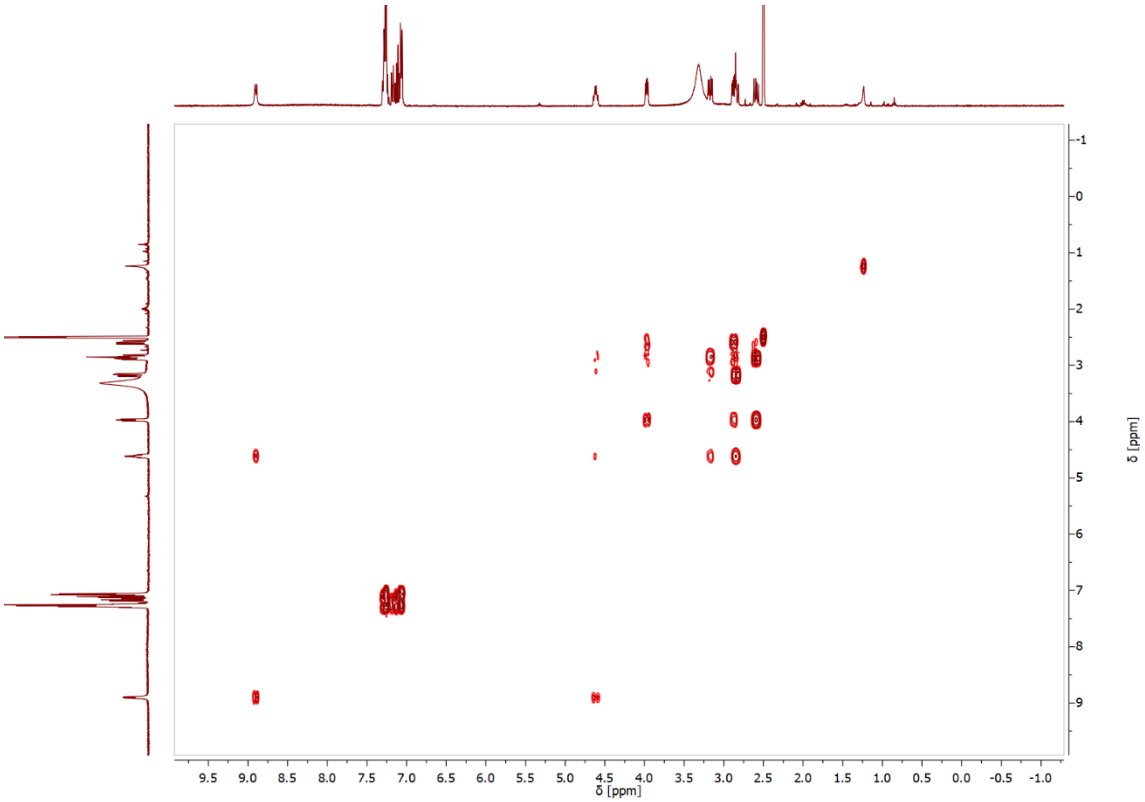


Fig. S2. gCOSY 2D NMR spectrum of L-Phe-D-(2F)-Phe.

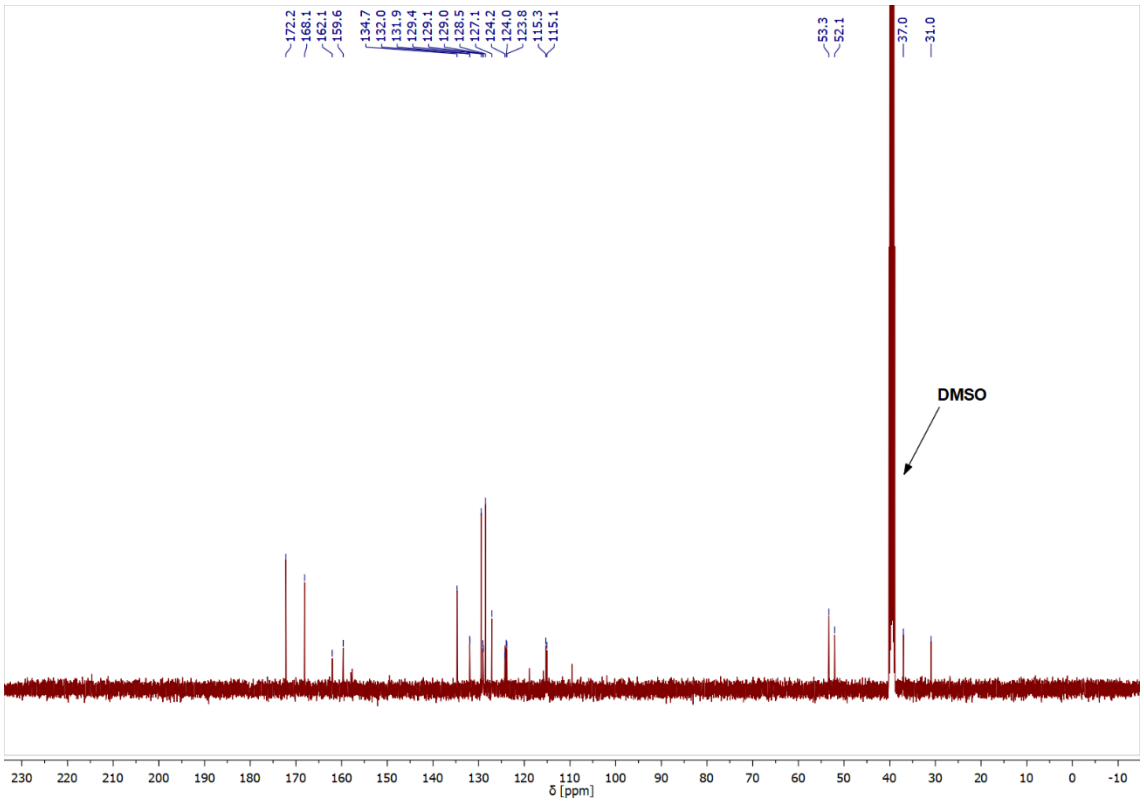


Fig. S3. ^{13}C NMR spectrum of L-Phe-D-(2F)-Phe.

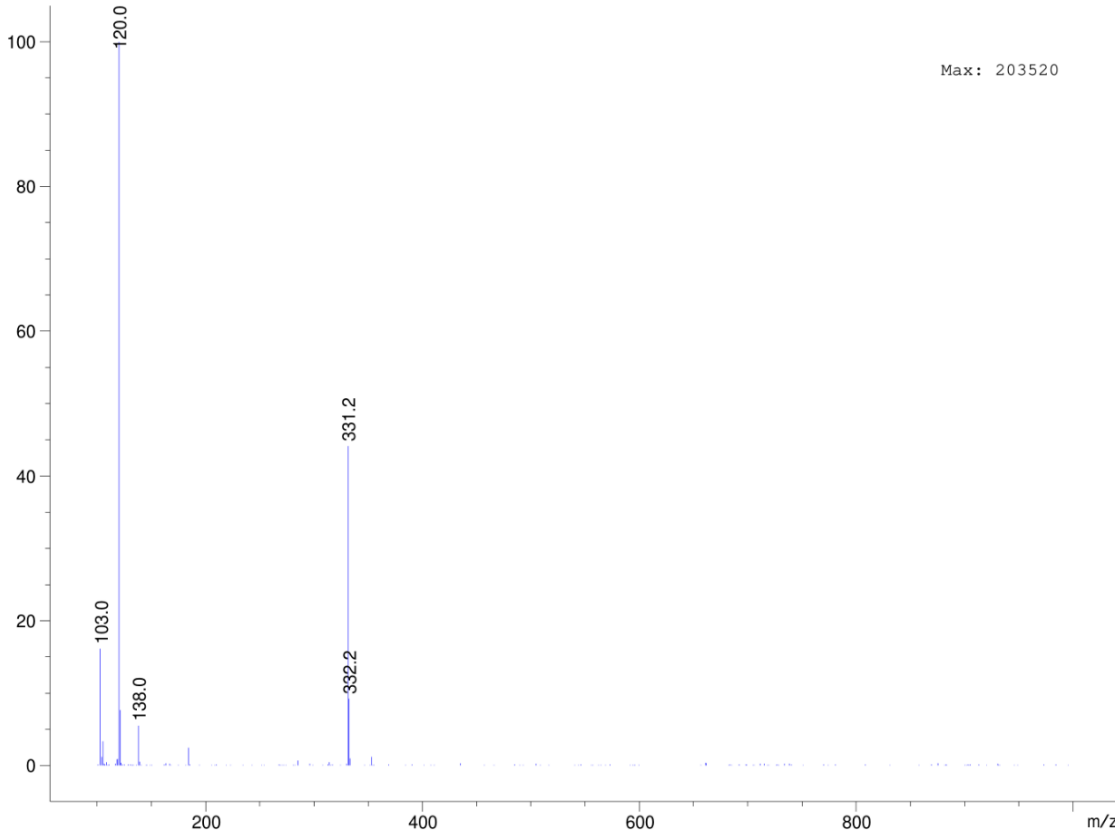


Fig. S4 ESI-MS spectrum of L-Phe-D-(2F)-Phe (positive ion mode).

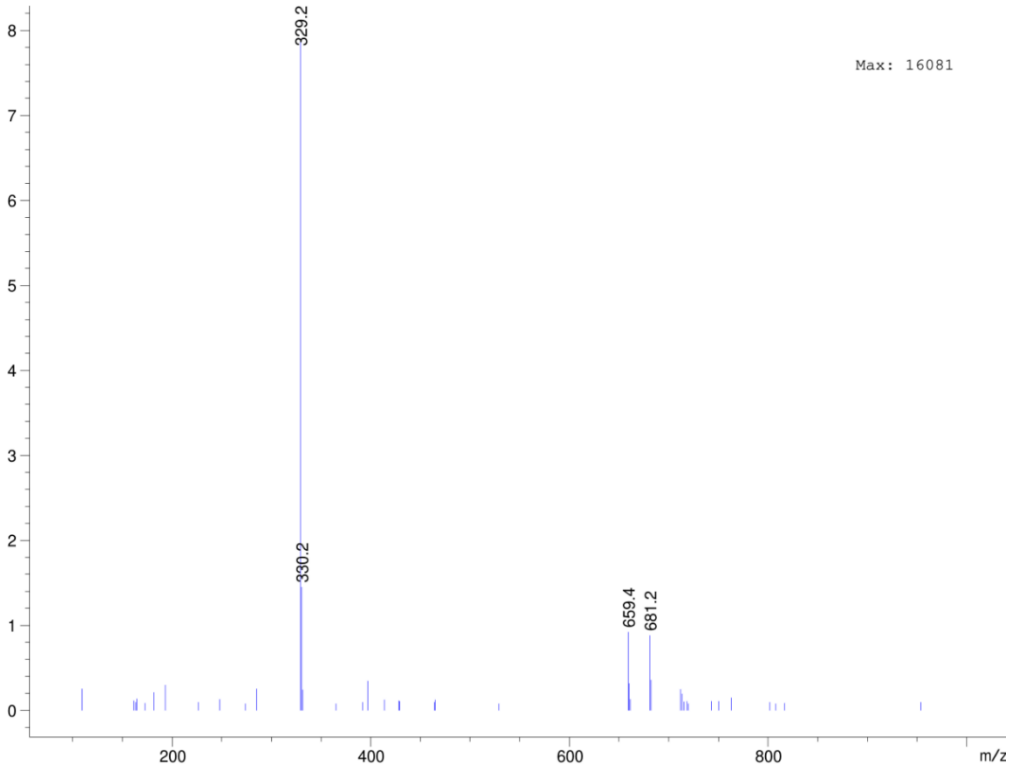
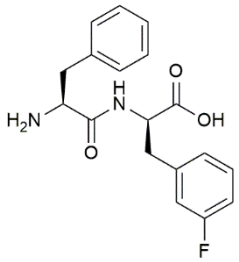


Fig. S5. ESI-MS spectrum of L-Phe-D-(2F)-Phe (negative ion mode).

S2. L-Phe-D-(3F)-Phe spectroscopic data



^{1}H-NMR (400 MHz, DMSO-*d*$_6$), δ (ppm): **8.89** (1H, d, *J* = 8.4Hz, NH), **7.29** (4H, m, ArF), **7.04** (5H, m, Ar), **4.59** (1H, ddd, *J* = 4.8Hz, *J* = 8.8Hz, *J* = 8.4Hz, CHαF), **4.00** (1H, dd, *J* = 8.4Hz, *J* = 4,8Hz, CHα), **3.12** (1H, dd, *J* = 13.8Hz, *J* = 4.8 Hz, CH$_\beta$F), **2.85** (2H, dd dd, *J* = 14.0Hz, *J* = 13.8Hz, *J* = 8.8Hz, *J* = 4.8Hz, CH$_\beta$F and CH$_\beta$F), **2.64** (1H, dd, *J* = 14.0Hz, *J* = 8.4Hz, CH$_\beta$). **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$), δ (ppm): **172.2** (1C, COOH), **168.0** (1C, CONHR), **163.3-160.9** (1C, d, ^{1}J = 241.9 Hz, ArF C$_1$), **140.0** (1C, d, ^{3}J = 7.5 Hz, ArF C$_5$), **134.6** (1C, Ar C$_4$), **130.2** (1C, d, ^{3}J = 8.3 Hz, ArF C$_3$), **129.4** (2C, Ar C$_3$,C$_5$), **128.5** (2C, Ar C$_2$C$_6$), **127.2** (1C, Ar C$_1$),**125.5** (1C, d, ^{4}J = 2.8 Hz, ArF C$_4$), **116.0** (1C, d, ^{2}J = 20.8 Hz, ArF C$_6$), **113.5** (1C, d, ^{2}J = 20.6 Hz, ArF C$_2$), **53.2** (2C, Cβ), **37.0** (1C, Cα), **36.7** (1C, Cα). **MS (ESI): C$_{18}$H$_{19}$FN$_2$O$_3$** exact mass: 330.14 m/z 331.2 $(M+H)^+$, 329.2 $(M-H)^-$

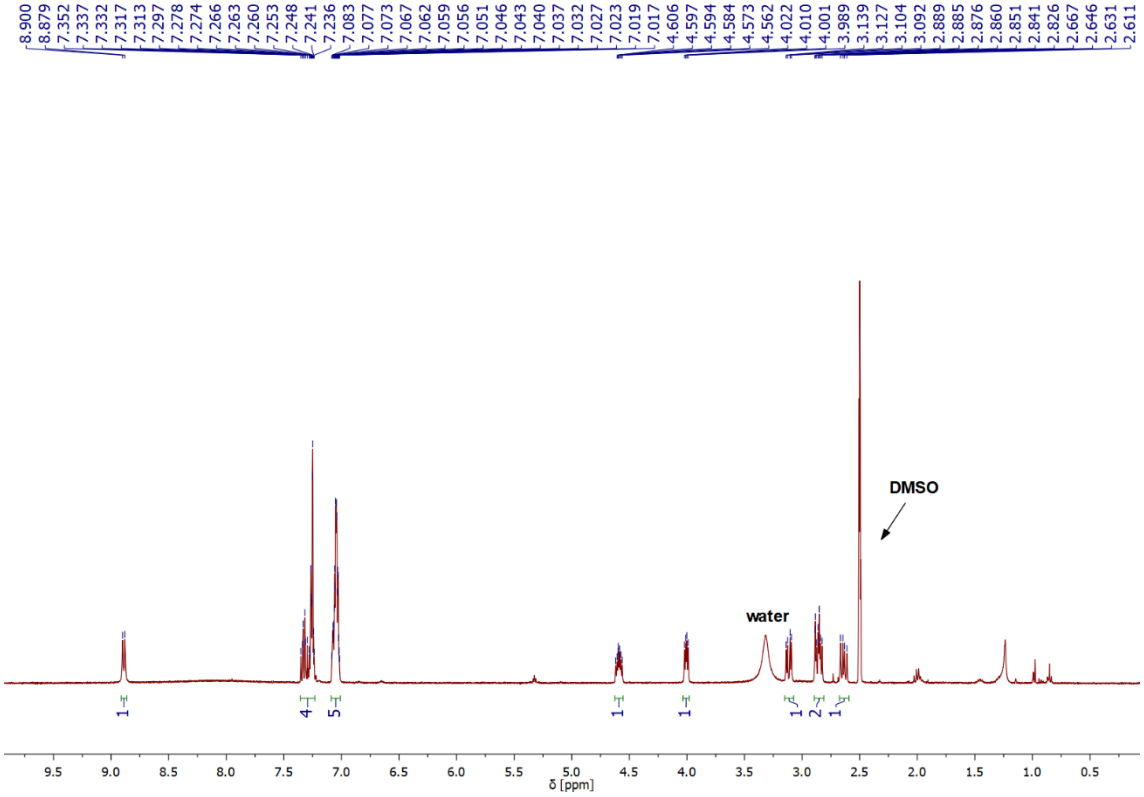


Fig. S6 ^{1}H NMR spectrum of L-Phe-D-(3F)-Phe.

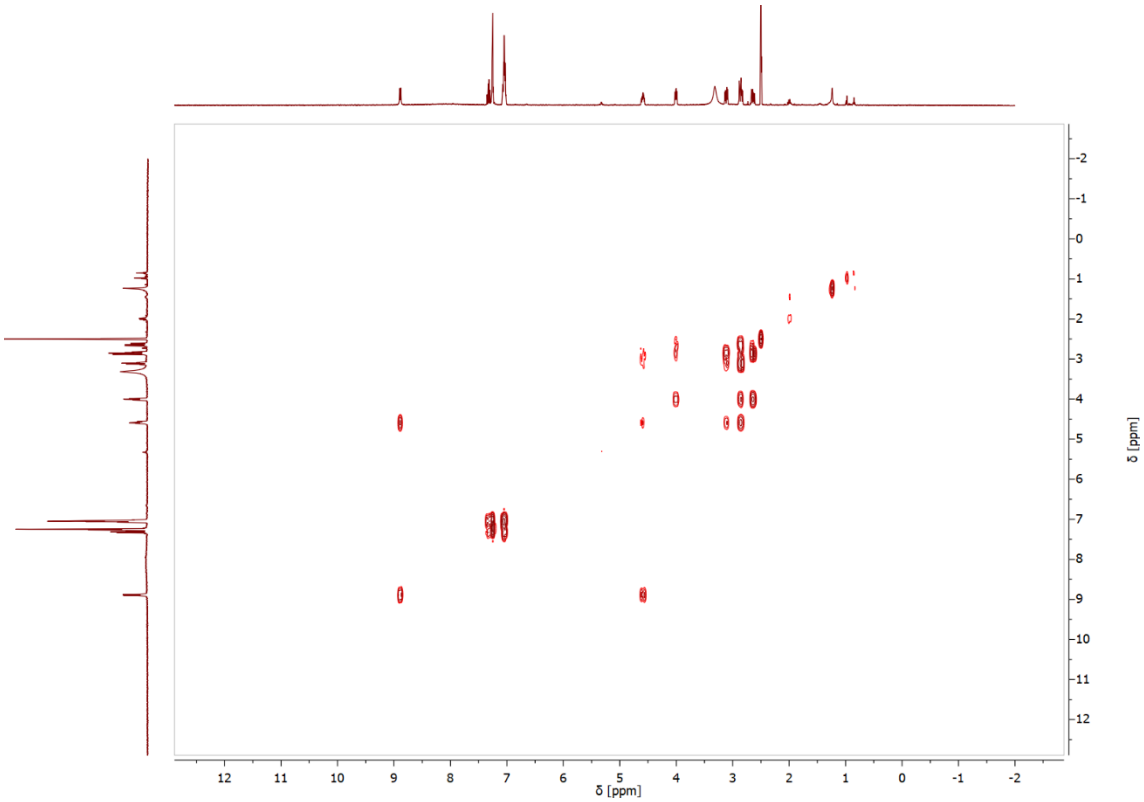


Fig. S7 gCOSY 2D-NMR spectrum of L-Phe-D-(3F)-Phe.

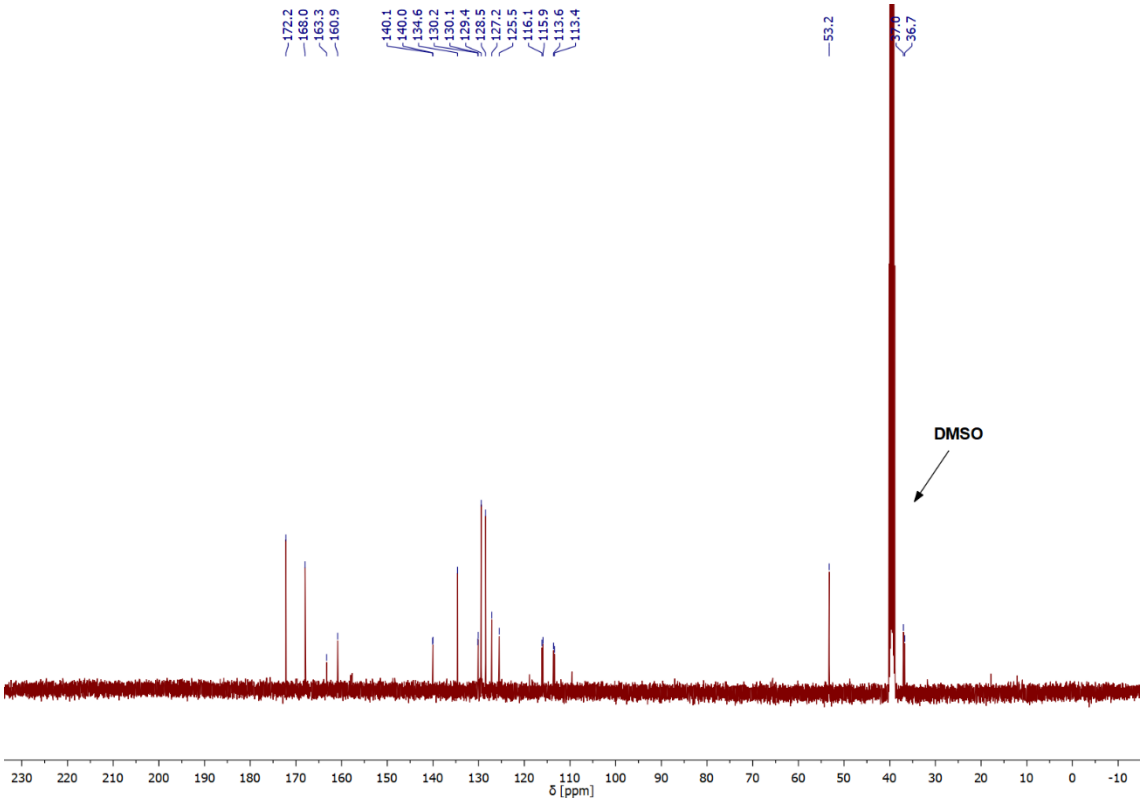


Fig. S8 ^{13}C NMR spectrum of L-Phe-D-(3F)-Phe.

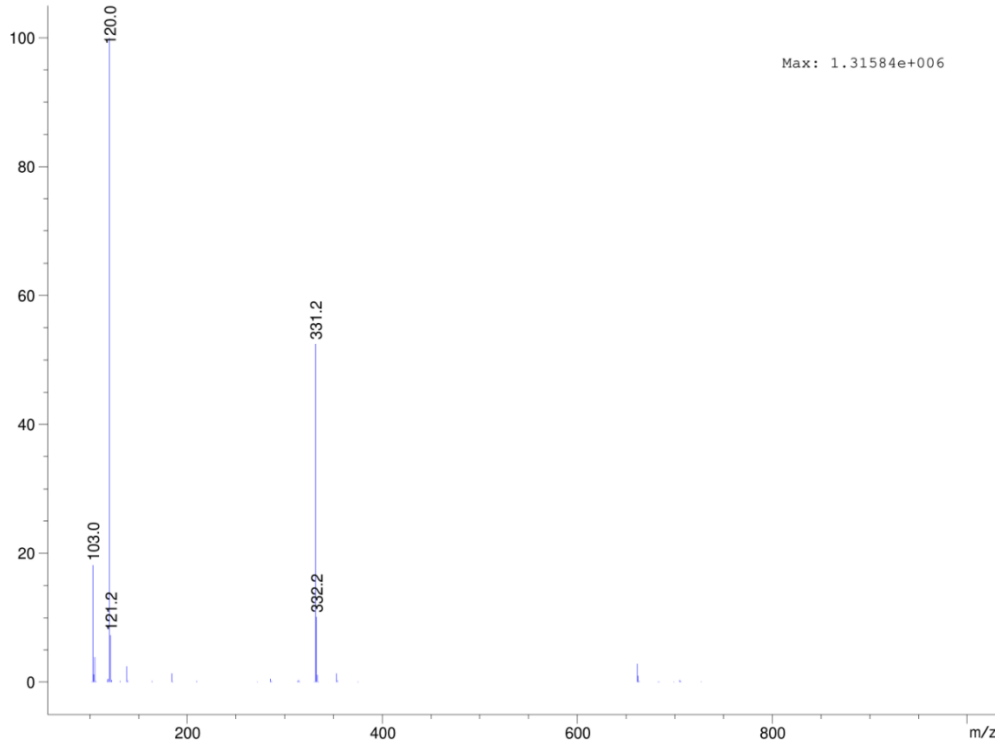


Fig. S9 ESI-MS spectrum of L-Phe-D-(3F)-Phe (positive ion mode).

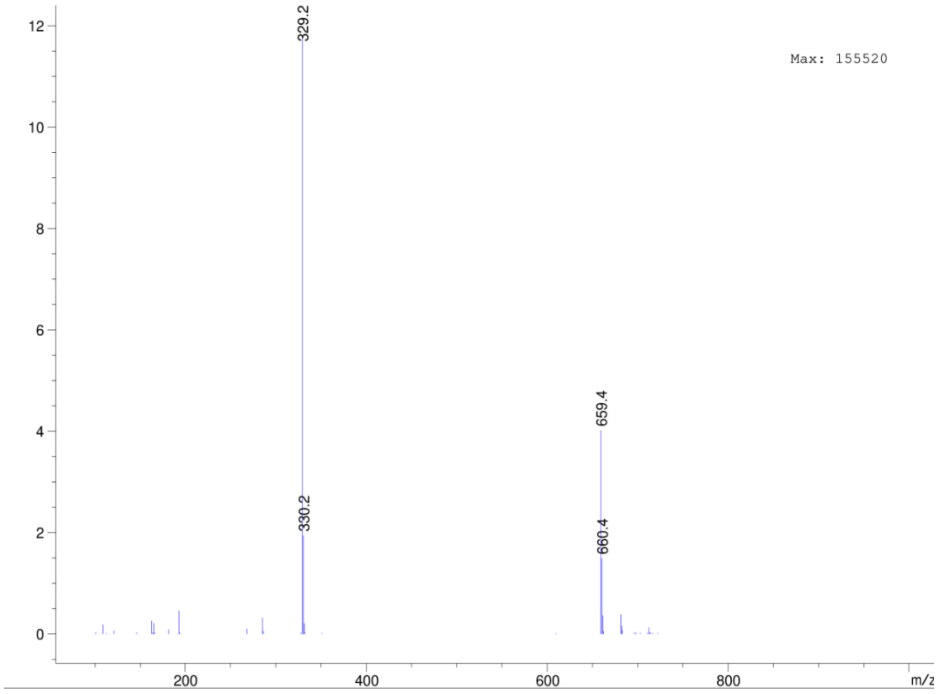
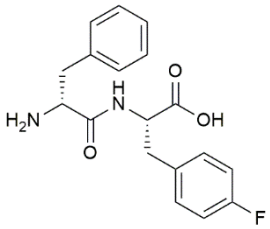


Fig. S10 ESI-MS spectrum of L-Phe-D-(3F)-Phe (negative ion mode).

S3. D-Phe-L-(4F)-Phe spectroscopic data



^{1}H-NMR (400 MHz, DMSO-d_6), δ (ppm): **8.83** (1H, d, *J* = 8.4Hz, NH), **7.24** (5H, m, Ar), **7.10** (2H, m, ArF), **7.03** (2H, m, ArF), **4.53** (1H, ddd, *J* = 8.8Hz, *J* = 8.4Hz, *J* = 4.8Hz, CHαF), **4.00** (1H, dd, *J* = 8.0Hz, *J* = 4.8Hz, CHα), **3.06** (1H, dd, *J* = 14.0Hz, *J* = 4.8 Hz, $CH_{\beta}F$), **2.88** (1H, dd, *J* = 14.0Hz, *J* = 4.8Hz, CH_{β}), **2.82** (1H, dd, *J* = 14.0Hz, *J* = 4.8Hz, $CH_{\beta}F$), **2.66** (1H, dd, *J* = 14.0Hz, *J* = 8.0Hz, CH_{β}). **^{13}C-NMR** (100 MHz, DMSO-d_6), δ (ppm): **172.3** (1C, COOH), **167.9** (1C, CONHR), **162.4 – 160.0** (1C, d, ^{1}J = 240.9 Hz, ArF C_1), **134.6** (1C, Ar C_4), **133.3** (1C, d, ^{4}J = 3.0 Hz, ArF C_4), **131.2** (2C, d, ^{3}J = 8.0 Hz, ArF $C_{3,5}$), **129.5** (2C, Ar $C_{3,5}$), **128.5** (2C, Ar $C_{2,6}$), **127.2** (1C, Ar C_1), **115.0** (2C, d, ^{2}J = 21.0 Hz, ArF $C_{2,6}$), **53.6** (1C, Cβ), **53.2** (1C, Cβ), **37.0** (1C, Cα), **36.3** (1C, Cα). **MS (ESI): $C_{18}H_{19}FN_2O_3$** exact mass: 330.14 m/z 331.2 $(M+H)^+$, 329.2 $(M-H)^-$

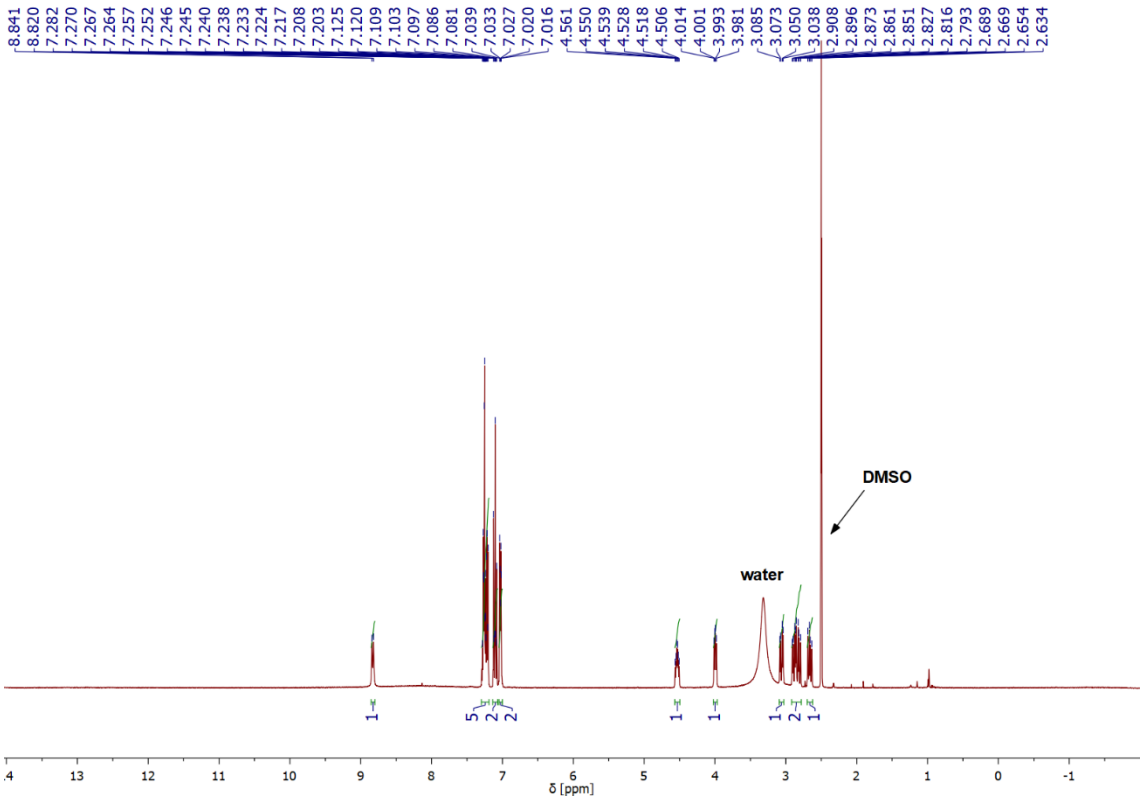


Fig. S11. ^{1}H NMR spectrum of D-Phe-L-(4F)-Phe.

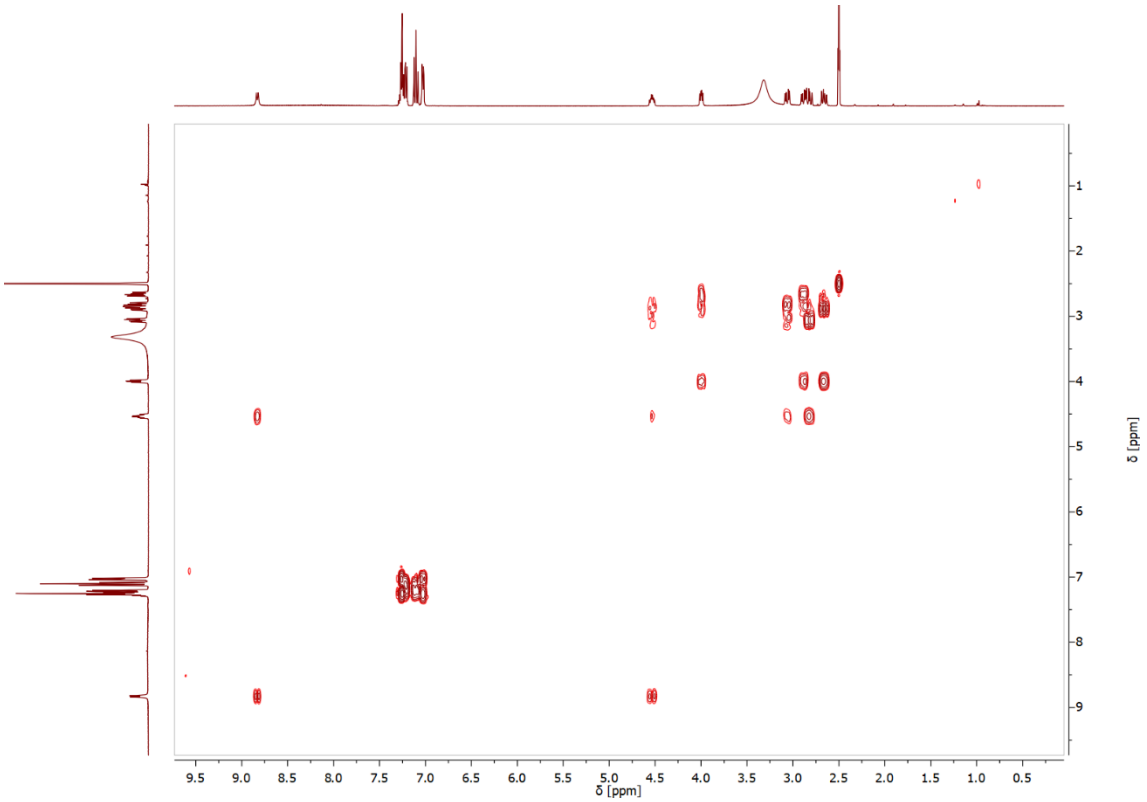


Fig. S12. gCOSY 2D-NMR spectrum of D-Phe-L-(4F)-Phe.

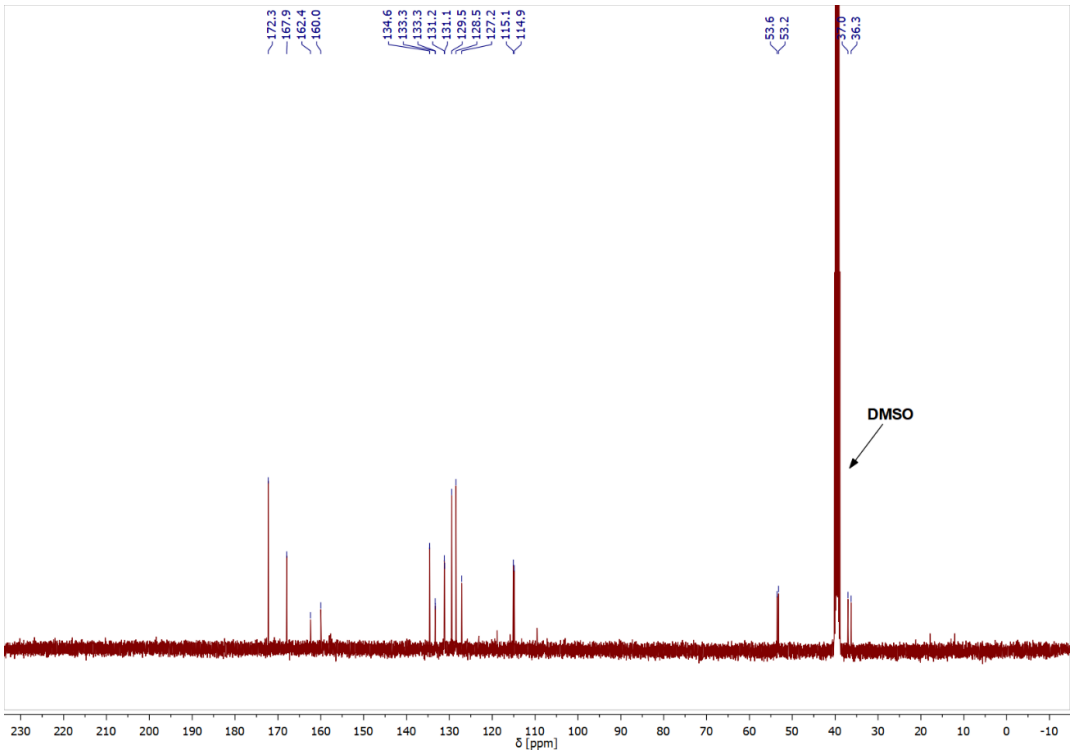


Fig S13. ^{13}C NMR spectrum of D-Phe-L-(4F)-Phe.

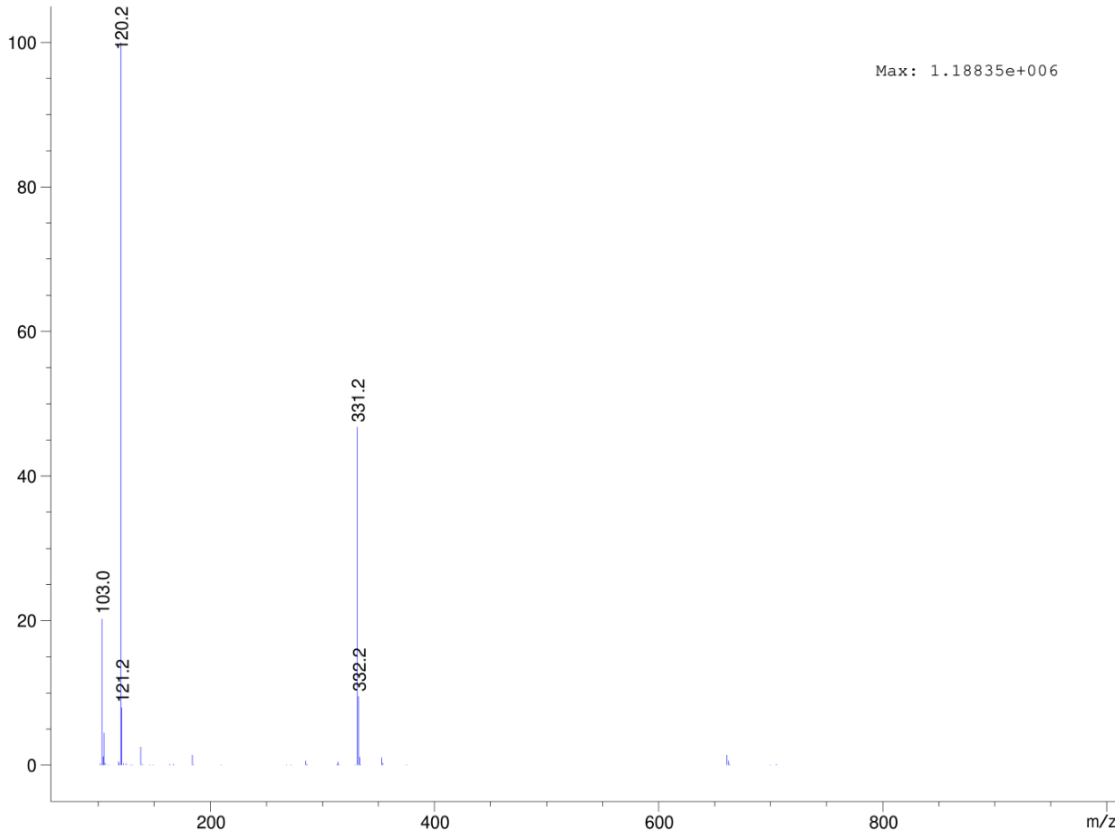


Fig. S14. ESI-MS spectrum of D-Phe-L-(4F)-Phe (positive ion mode).

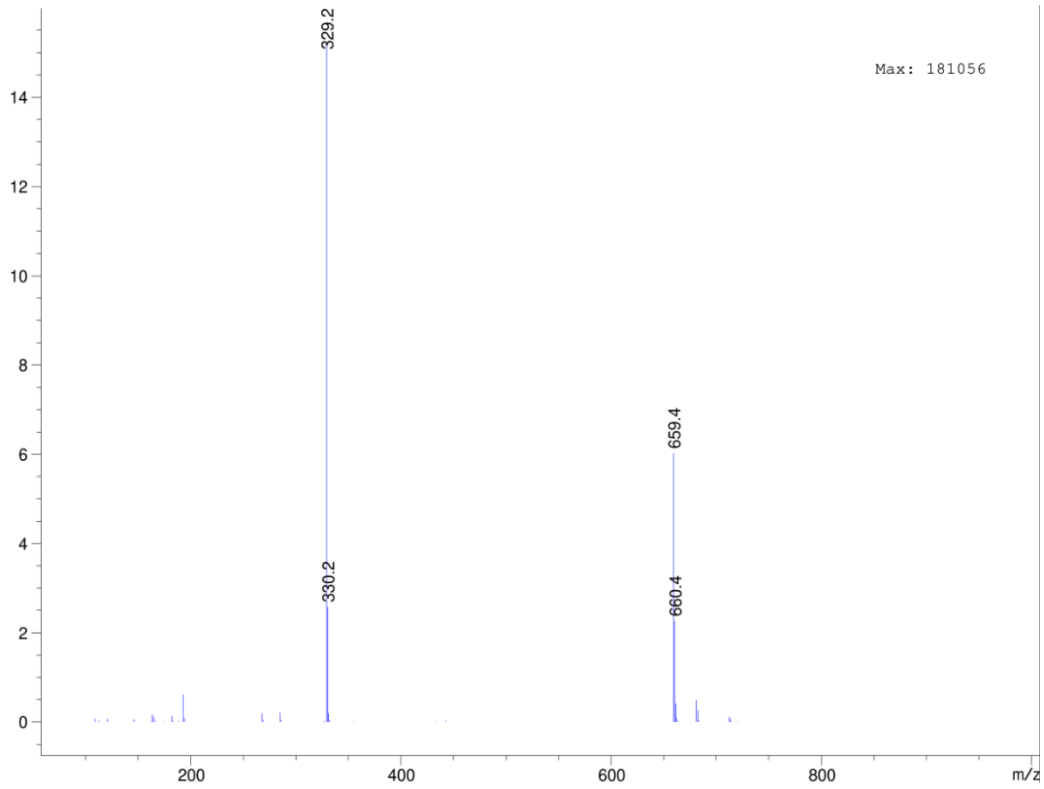
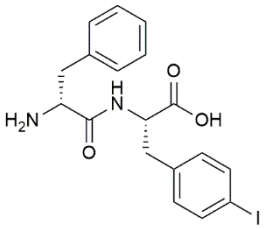


Fig. S15 ESI-MS spectrum of D-Phe-L-(4F)-Phe (negative ion mode).

S4. D-Phe-L-(4I)-Phe spectroscopic data



^{1}H NMR (400 MHz, DMSO-d_6), δ (ppm): **8.83** (1H, d, *J* = 8.4Hz, NH), **7.62** (2H, d, J_o = 8.4Hz, ArI), **7.27** (3H, dd, J_o = 5.2Hz, J_m = 1.6Hz, Ar), **7.00** (2H, m, Ar), **6.98** (2H, d, J_o = 8.4Hz, ArI), **4.52** (1H, ddd, *J* = 8.8Hz, *J* = 8.4Hz, *J* = 4.8Hz, CHαI), **4.01** (1H, dd, *J* = 8.0Hz, *J* = 4.8Hz, CHα), **3.02** (1H, dd, *J* = 13.8Hz, *J* = 4.8Hz, CHβI), **2.90** (1H, dd, *J* = 14.0Hz, *J* = 4.8Hz, CHβ), **2.78** (1H, dd, *J* = 13.8Hz, *J* =8.8Hz, CHβI), **2.67** (1H, dd, *J* = 14.0Hz, *J* = 8.0Hz, CHβ). **^{13}C NMR** (100 MHz, DMSO-d_6), δ (ppm): **172.2** (1C, COOH), **168.0** (1C, CONHR), **137.0** (3C, Ar), **134.6** (1C, Ar), **131.7** (2C, Ar), **129.5** (2C, Ar), **128.5** (2C, Ar), **127.1** (1C, Ar), **92.7** (1C, CI), **53.3** (1C, Cβ), **53.2** (1C, Cβ), **37.0** (1C, Cα), **36.6** (1C, Cα). **MS (ESI): $C_{18}H_{19}IN_2O_3$** exact mass: 438.04 m/z 439.0 $(M+H)^+$, 437.2 $(M-H)^-$

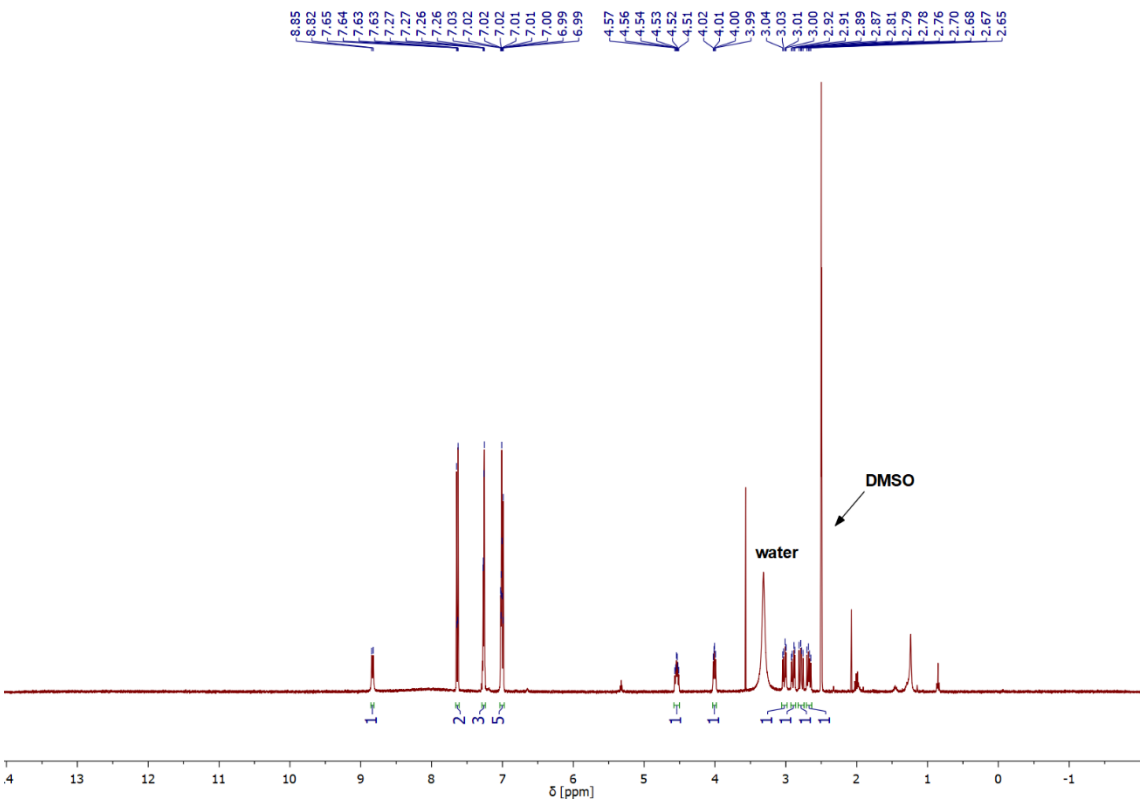


Fig. S16. ^{1}H NMR spectrum of D-Phe-L-(4I)-Phe.

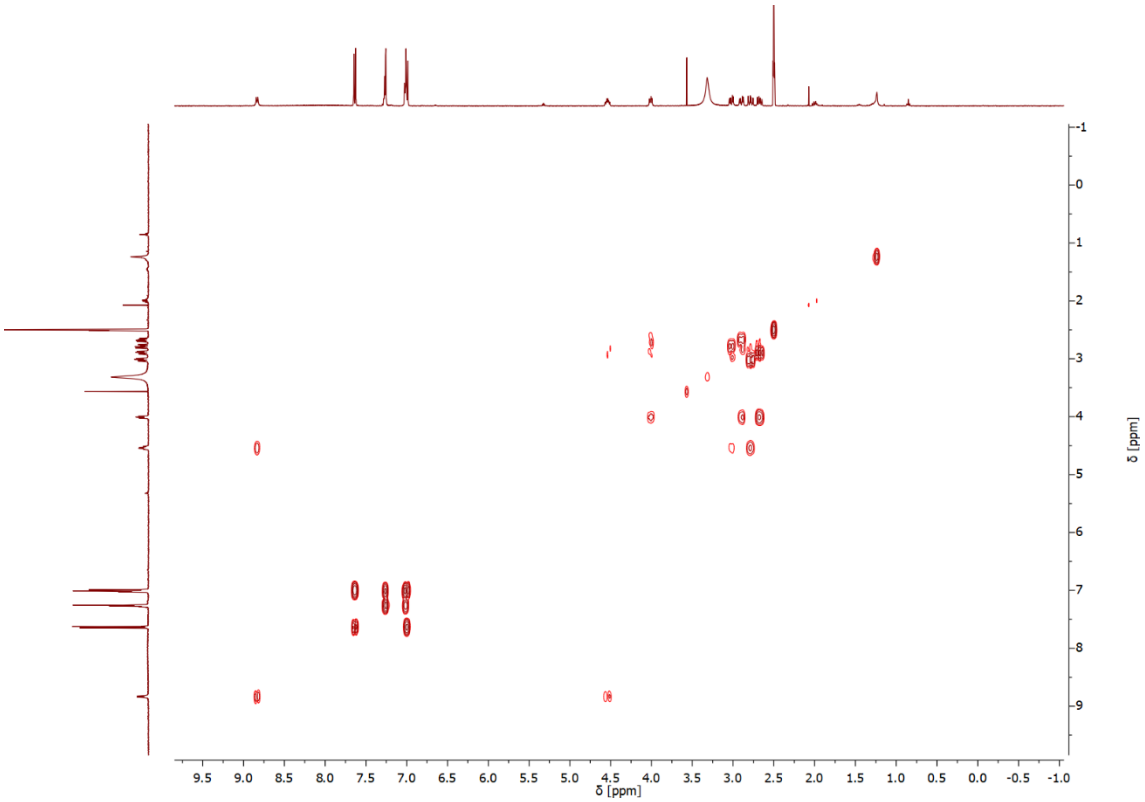


Fig. S17. gCOSY 2D-NMR spectrum of D-Phe-L-(4I)-Phe.

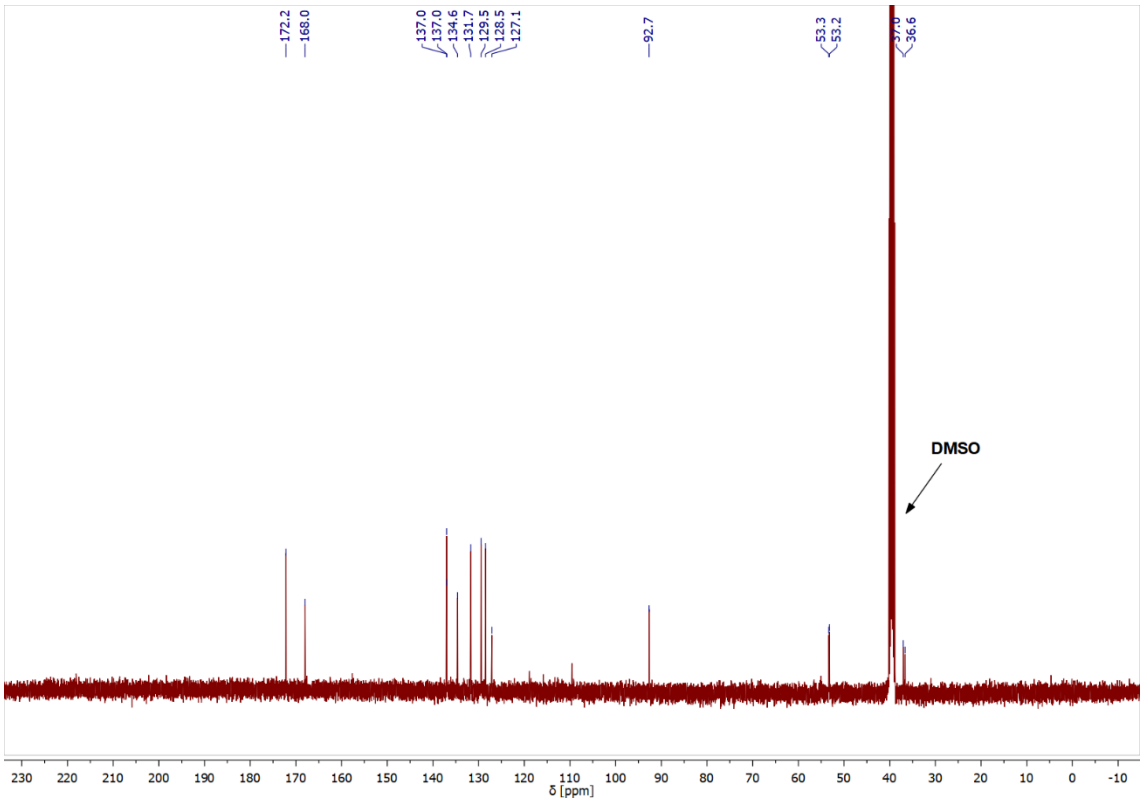


Fig. S18. ^{13}C NMR spectrum of D-Phe-L-(4I)-Phe.

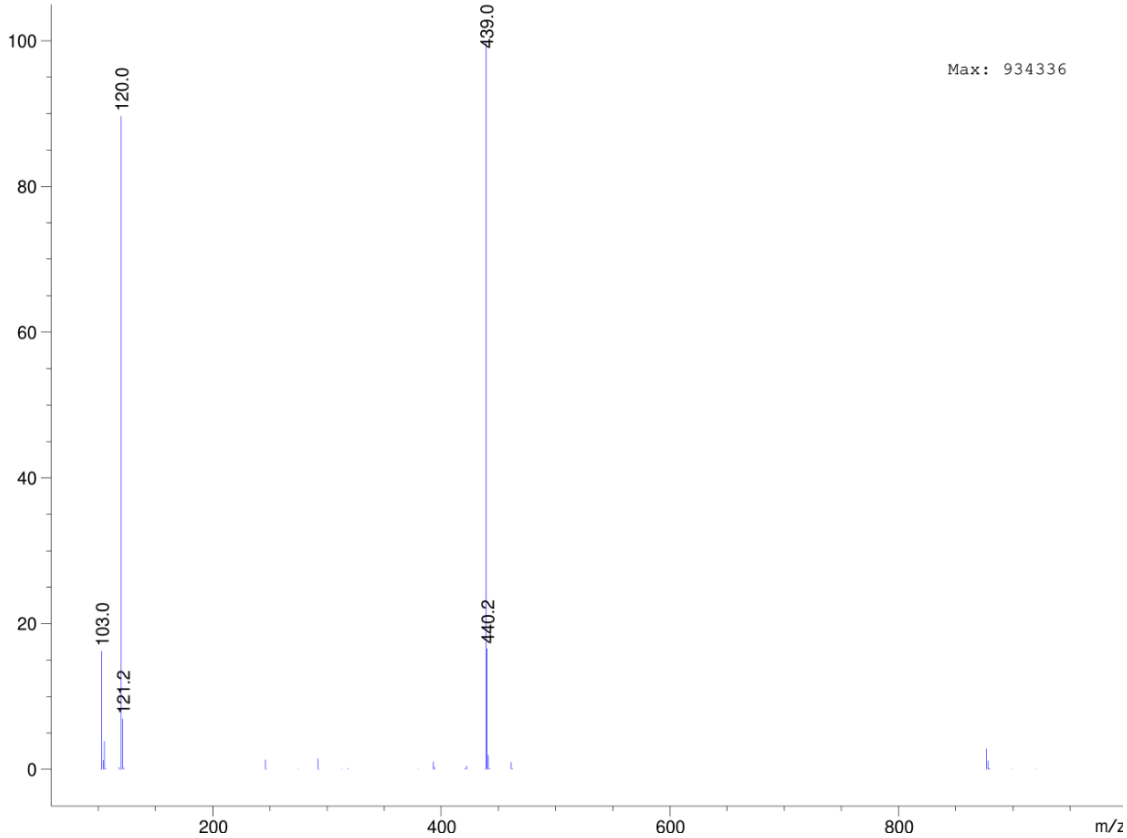


Fig. S19. ESI-MS spectrum of D-Phe-L-(4I)-Phe (positive ion mode).

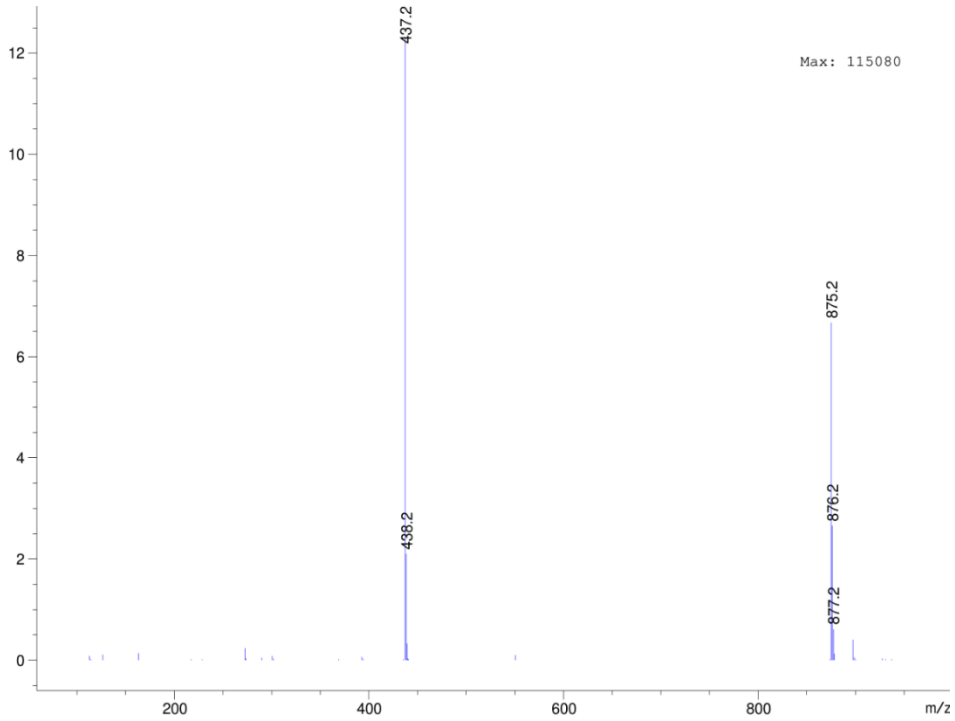


Fig. S20. ESI-MS spectrum of D-Phe-L-(4I)-Phe (negative ion mode).

S5. LC-MS traces

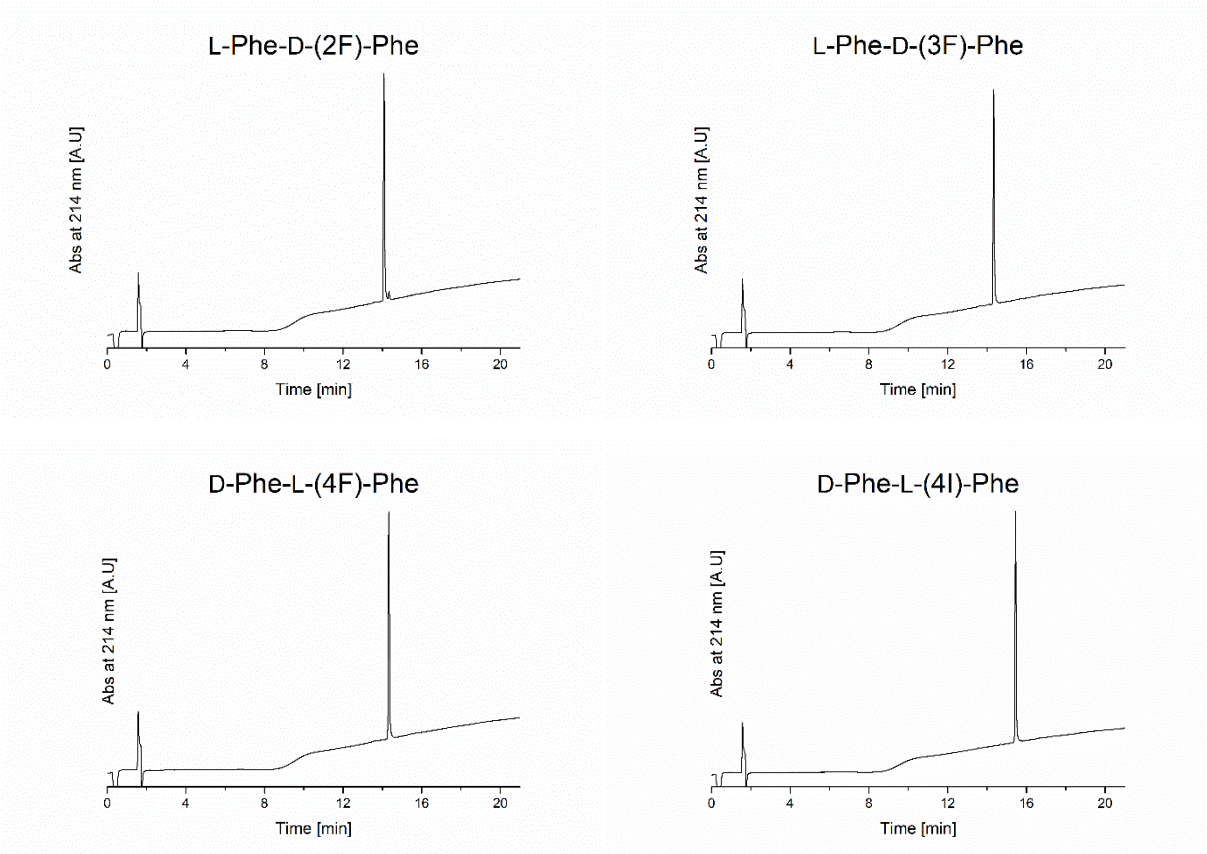


Fig. S21. LC traces of the four peptides. Retention times are reported in the **Table S1.**

Halogen Substitution	**Retention time (min) for C-terminal substitution (this work)**	**Retention time (min) for N-terminal substitution (*ACS Nano* 2020, *14*, 16951)**
2-F	14.2	14.1
3-F	14.4	14.4
4-F	14.4	14.4
4-I	15.5	15.5

Table S1. Retention time of heterochiral Phe-Phe dipeptides from chromatography of **Fig. S21**, and their comparison with their gelling analogs halogenated at the N-terminal position (describe din *ACS Nano* 2020, *14*, 16951). The HPLC method, which was run on an analytical C-18 column (Luna, Phenomenex), with a flux of 0.3 ml/min, consisted of a mixture MeCN/water with a gradient from 5% to 95% MeCN (with 0.05% TFA) over 20 minutes.

S6. Gel thermoreversibility tests

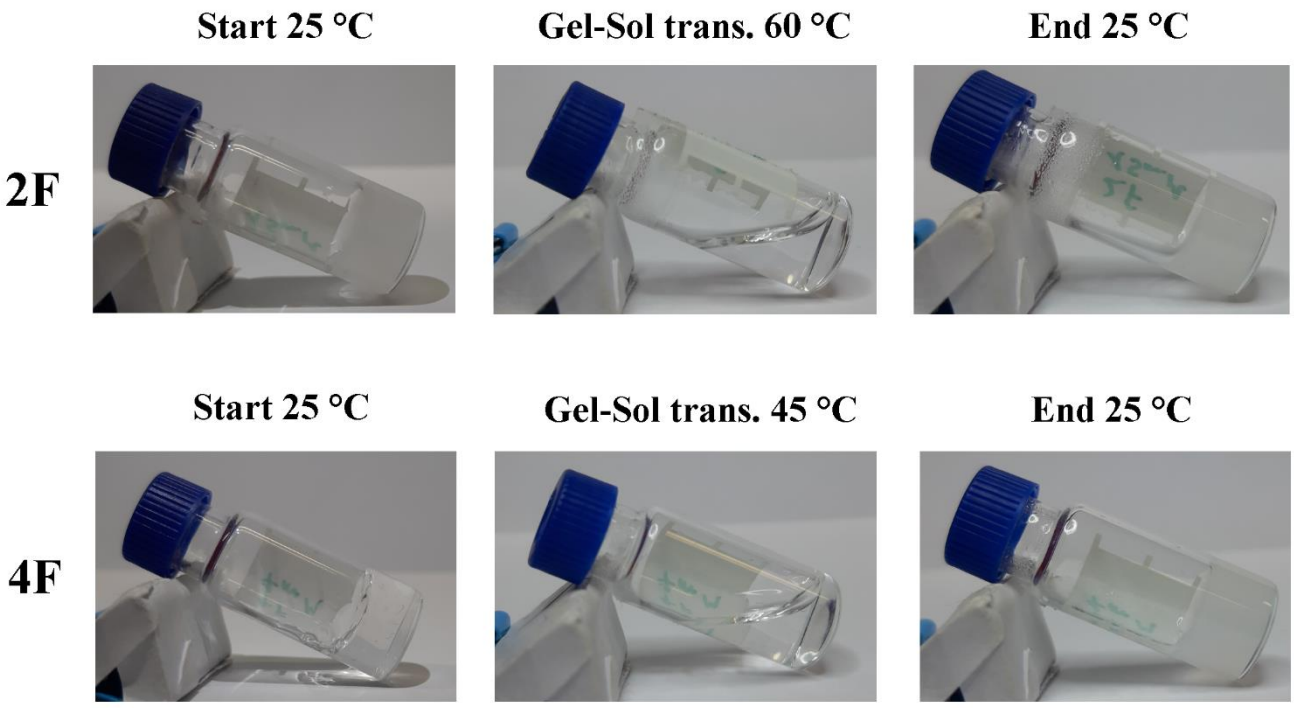


Fig. S22. Gel thermoreversibility tests for 2-F and 4-F at their mgc.

S7. TEM micrographs

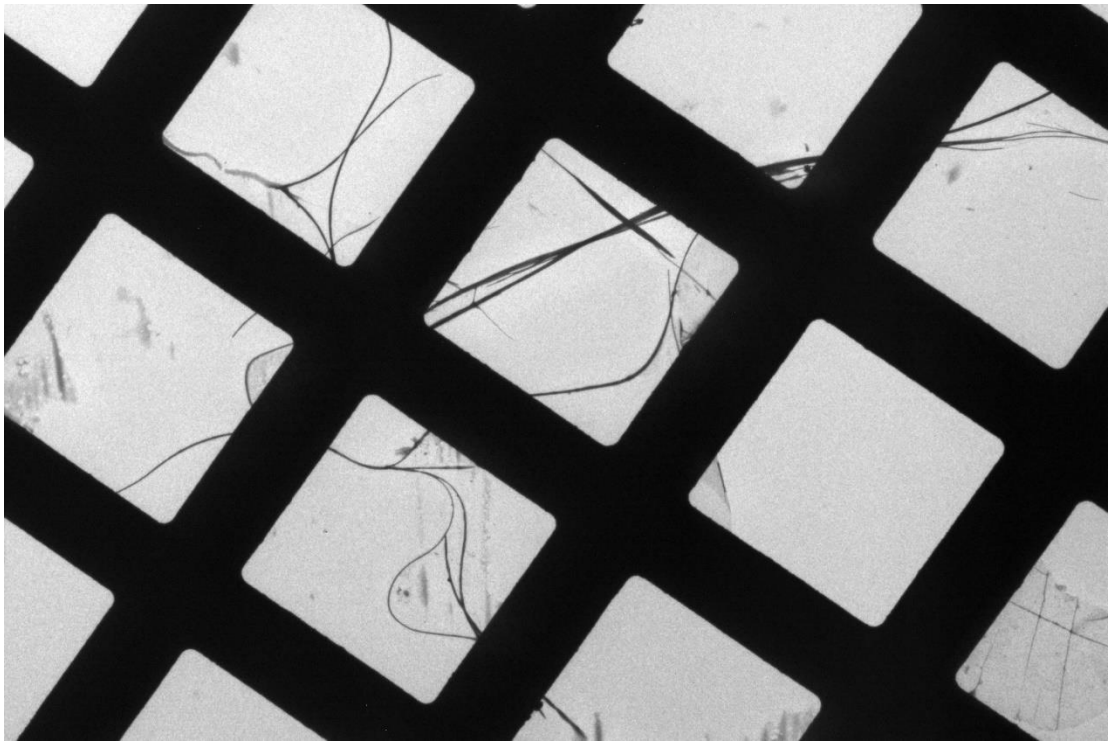


Fig. S23. Low magnification (Low-MAG) TEM micrograph showing visible curved microfibres for 2-F gel in the grid.

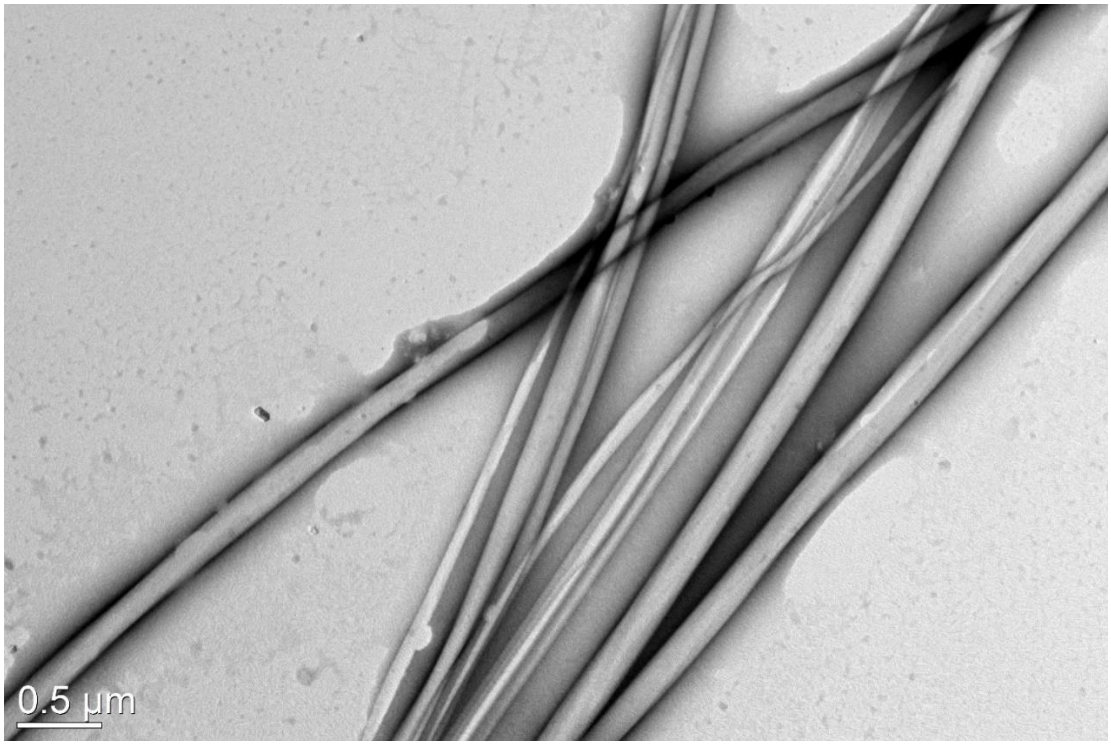


Fig. S24. TEM image of microfibers for 2-F gel.

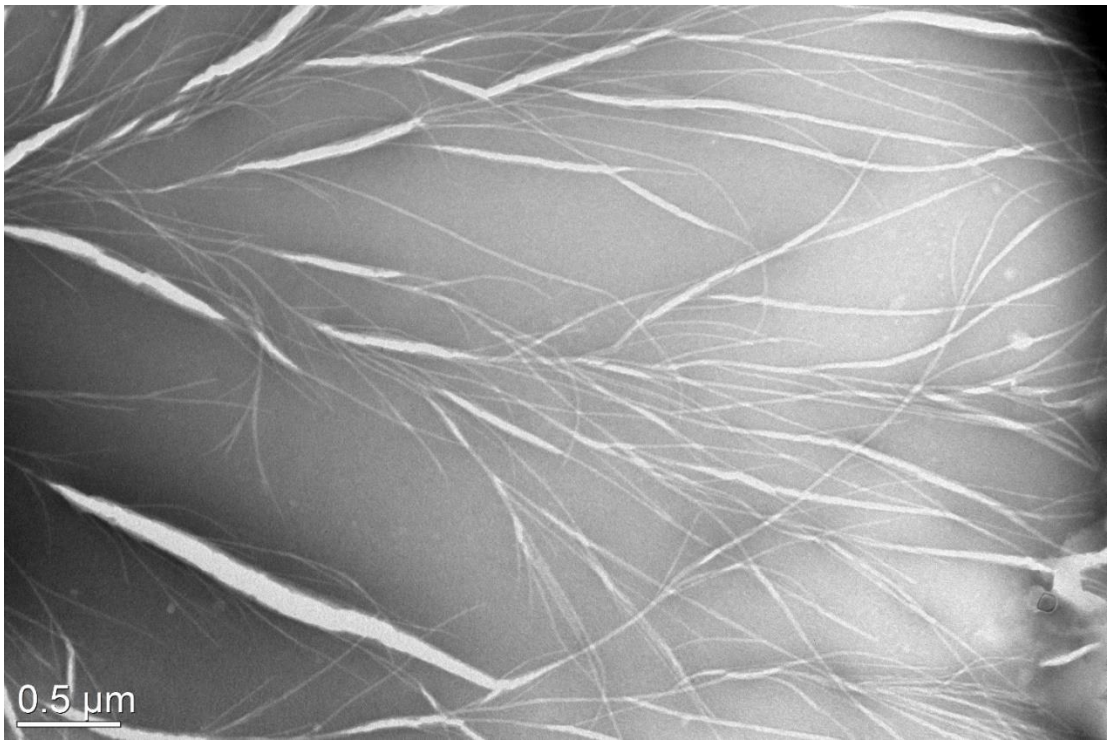


Fig. S25. TEM image of 2-F bundling and branching fibrils into microfibers.

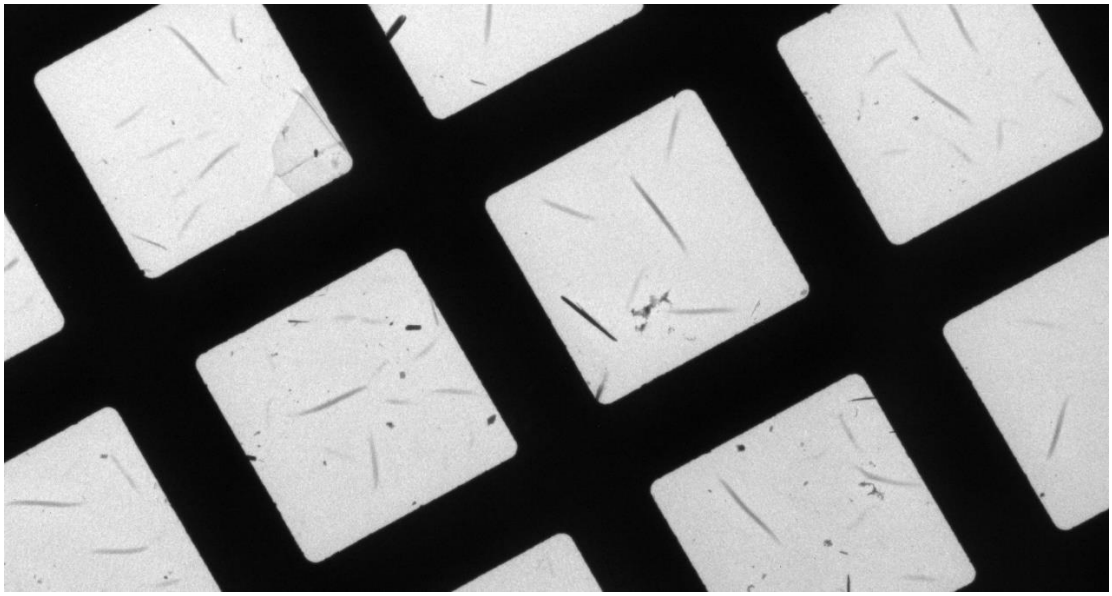


Fig. S26. Low magnification (Low-MAG) TEM image of microcrystals of 3-F dispersion.

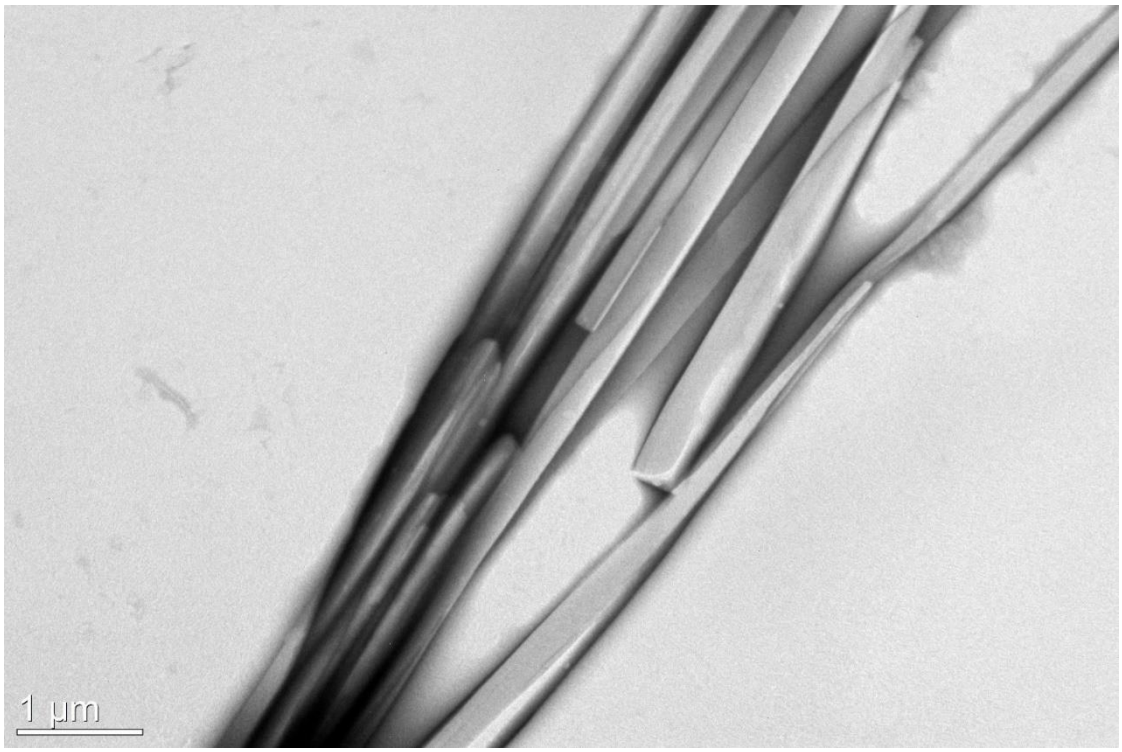


Fig. S27. TEM image of 3-F dispersion.

Fig, S28. Low magnification (Low-MAG) TEM image of 4-I microcrystals' dispersion on the grid.

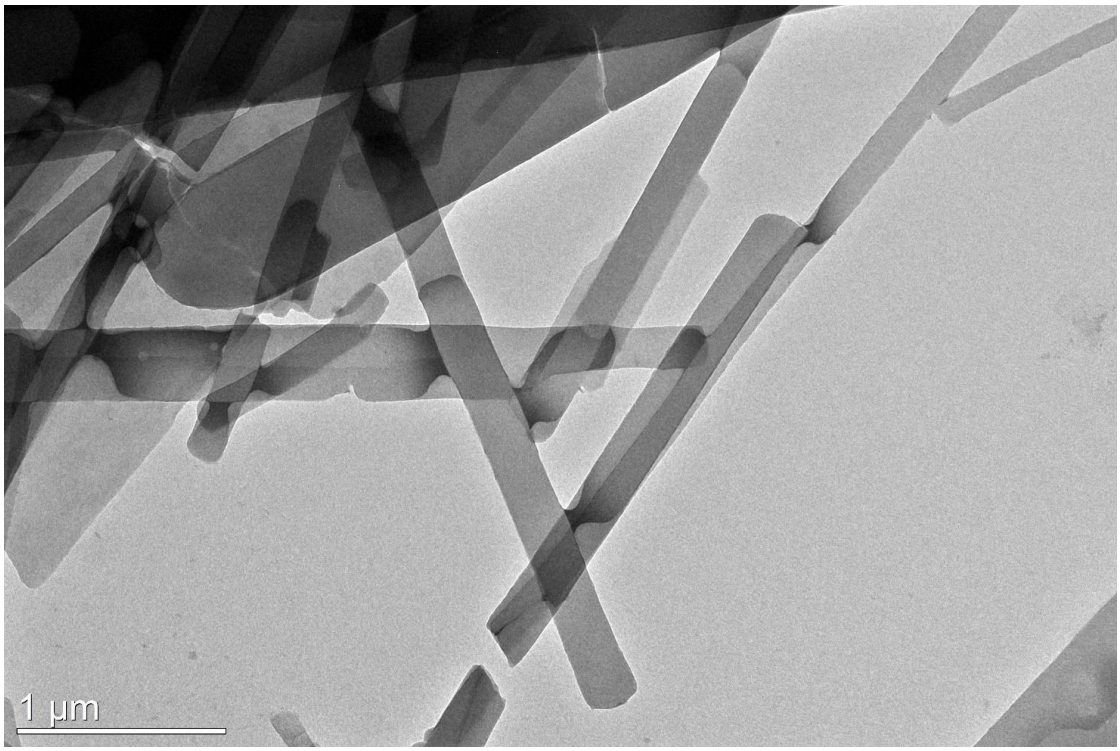


Fig. S29. TEM image of 4-I dispersion.

S8. Rheology data – frequency sweeps

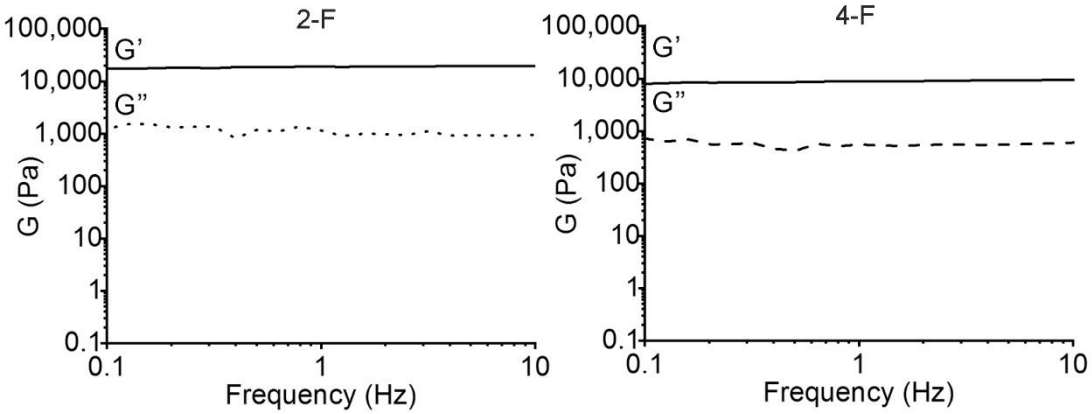


Fig. S30. Frequency sweeps for 2-F and 4-F at their mgc.

S9. Crystallographic data

D-Phe-L-(4-I)-Phe (denoted as 4-I in the MS)

Description.

The asymmetric unit contains a molecule of the peptide in its zwitterion form, two molecules of water in general positions and a molecule of water in a special position (2-fold symmetry axis). A total of 4 molecules of peptide, related by symmetry operators of the *C2* space group, are present in the unit cell (Figure S31).

The crystal packing shows a separation between layers with prevalence of hydrophobic groups, *i.e.* the phenyl moieties, and regions with hydrophilic interactions involving the backbone of the peptides and the solvent molecules (Figure S32). In particular, in the hydrophilic layers, peptides form piles in the crystallographic direction *b*, held together by hydrophilic interactions, most of which are mediated by solvent molecules (Figure S33). In the hydrophobic layers, no specific π-π interactions are recognizable. The iodine atoms are equally spaced along the *a* crystallographic direction, with an average distance of 4.6 Å (Figure S34).

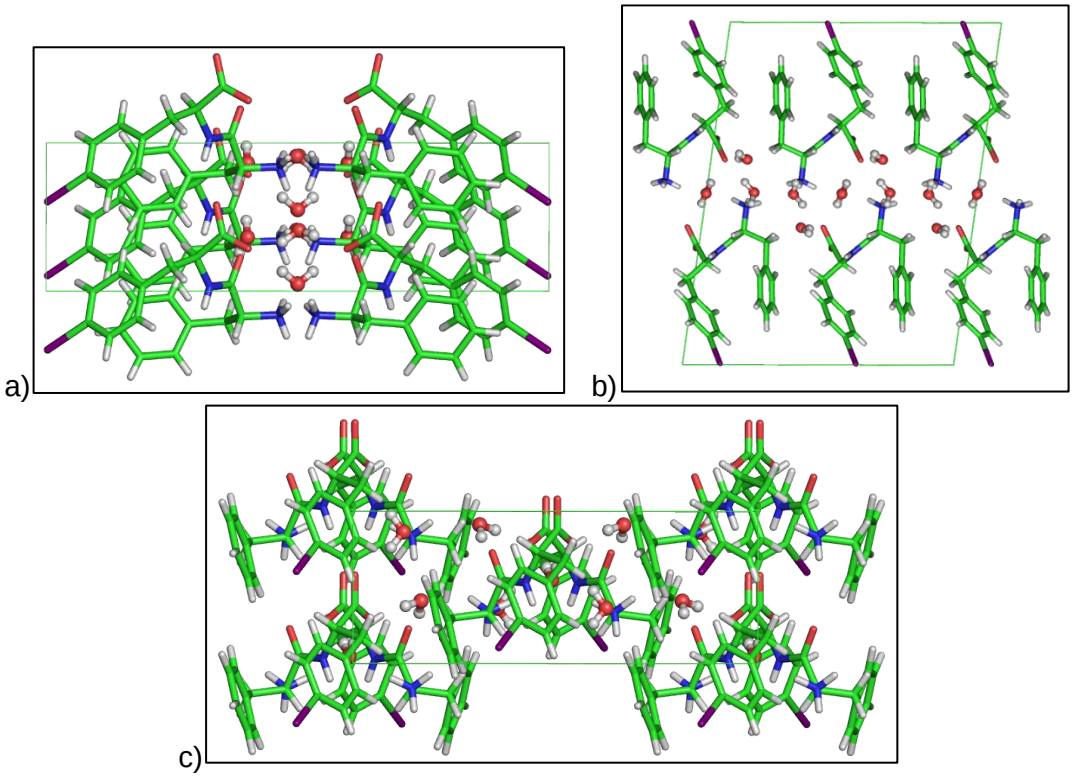


Fig. S31: Unit cell of crystals of 4-I. Crystal packing in the crystal of the peptide 4-I, grown from a phosphate buffer solution. Views along the *a* crystallographic axis (a), the *b* crystallographic axis (b) and the *c* crystallographic axis (c). Peptide molecules are shown as sticks, solvent molecules are shown as spheres. Carbon atoms are shown in green, oxygen atoms in red, nitrogen atoms in blue, iodine atoms in magenta, hydrogen atoms in white.

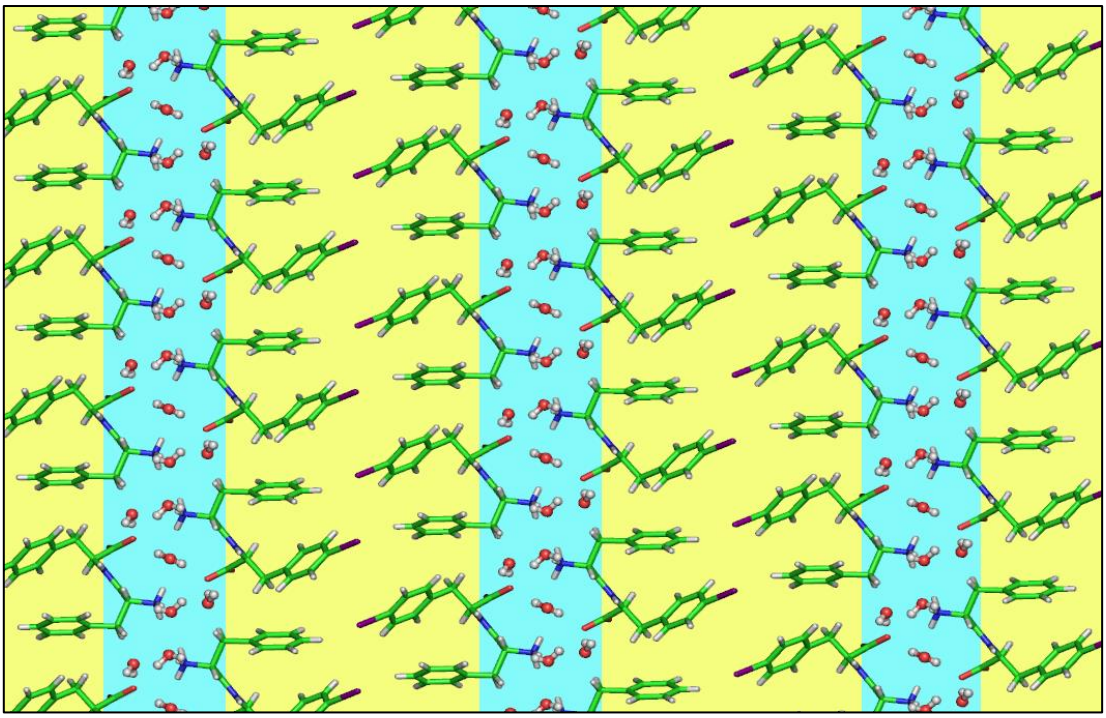


Fig. S32: Packing of peptide and solvent in the structure of 4-I. Hydrophilic and hydrophobic layers in the crystal of 4-I. Hydrophobic layers are shown with a yellow background, hydrophilic layers with a blue background. Crystal packing is shown along the *b* crystallographic direction.

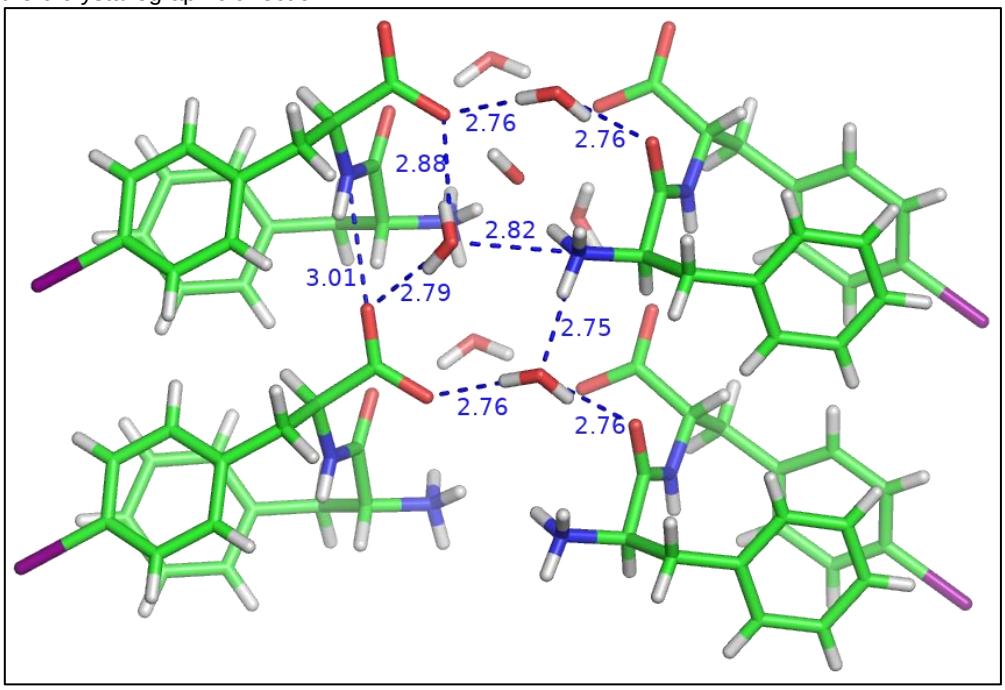


Fig. S33: Hydrogen bond and salt bridge interactions in the hydrophilic layers. Within the hydrophilic layer, interactions between peptides are mediated by hydrogen bond and salt bridge interactions between solvent molecules and peptide moieties, i.e. carboxylate groups, amino groups and amide groups. Donor-acceptor distances are reported in Å.

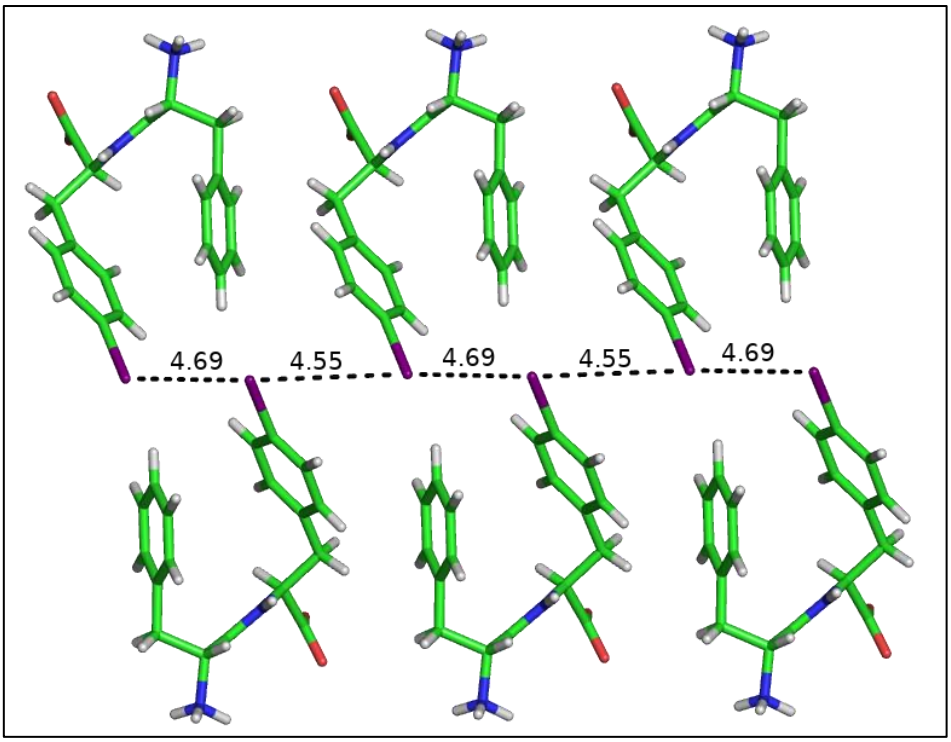


Fig. S34: Halogen-halogen distances in the structure of 4-I. Iodine-iodine distances within the hydrophobic layers are shown with black dashes and values are reported in Å.

Crystallographic details.

A needle-shaped single crystal of the peptide was collected with a loop, cryoprotected by dipping the crystal in glycerol and stored frozen in liquid nitrogen. The crystal was mounted on the diffractometer at the synchrotron Elettra, Trieste (Italy), beamline XRD1, using the robot present at the facility. Temperature was kept at 100 K by a stream of nitrogen on the crystal. Diffraction data were collected by the rotating crystal method using synchrotron radiation, wavelength 0.70 Å, rotation interval 1°/image, crystal-to-detector distance of 85 mm. A total of 360 images were collected to increase redundancy of data. Reflections were indexed and integrated using the XDS package [1], space group *C2* was determined using POINTLESS [2] and the resulting data set was scaled using AIMLESS [3]. Phase information were obtained by direct methods using the software SHELXT [4]. Refinements cycles were conducted with SHELXL-14 [4], operating through the WinGX GUI [5], by full-matrix least-squares methods on F^2. Unit cell parameters and scaling statistics are reported in Table S2. The asymmetric unit contains a molecule of the peptide, 2.5 molecule of water, one of which is located in a special position (2-fold symmetry axis). Hydrogen atoms of the peptide were added at geometrically calculated positions and refined isotropically, with thermal parameters dependent on those of the attached atom. Hydrogen atoms of the water molecules were added in the electron density considering the hydrophilic interactions with peptide and solvent molecules, and refined with geometric restraints on O-H distances (0.84±0.02Å) and H-H distance (1.34±0.04Å). During refinement, no restraints were applied on distances, angles or thermal parameters of the peptide. All the atoms, except the hydrogen atoms, within the asymmetric unit have been refined with anisotropic thermal parameters. Refinement statistics are reported in Table S2.

L-Phe-D-(3-F)-Phe (denoted as 3-F in the MS)

Description.

The asymmetric unit contains a molecule of the peptide in its zwitterion form, two molecules of water in general positions and a molecule of water in a special position (2-fold symmetry axis). A total of 4 molecules of peptide, related by symmetry operators of the *C2* space group, are present in the unit cell (Figure S35).

The crystal packing shows a separation between layers with prevalence of hydrophobic groups, i.e. the phenyl moieties, and regions with hydrophilic interactions involving the backbone of the peptides and the solvent molecules (Figure S36). In particular, in the hydrophilic layers, peptides form piles in the crystallographic direction *b*, held together by hydrophilic interactions, most of which are mediated by solvent molecules. In the hydrophobic layers, no specific π-π interactions are recognizable. The fluorine atoms are aligned along the *a* crystallographic direction, with distances of 3.2 Å and 5.9 Å (Fig. S37).

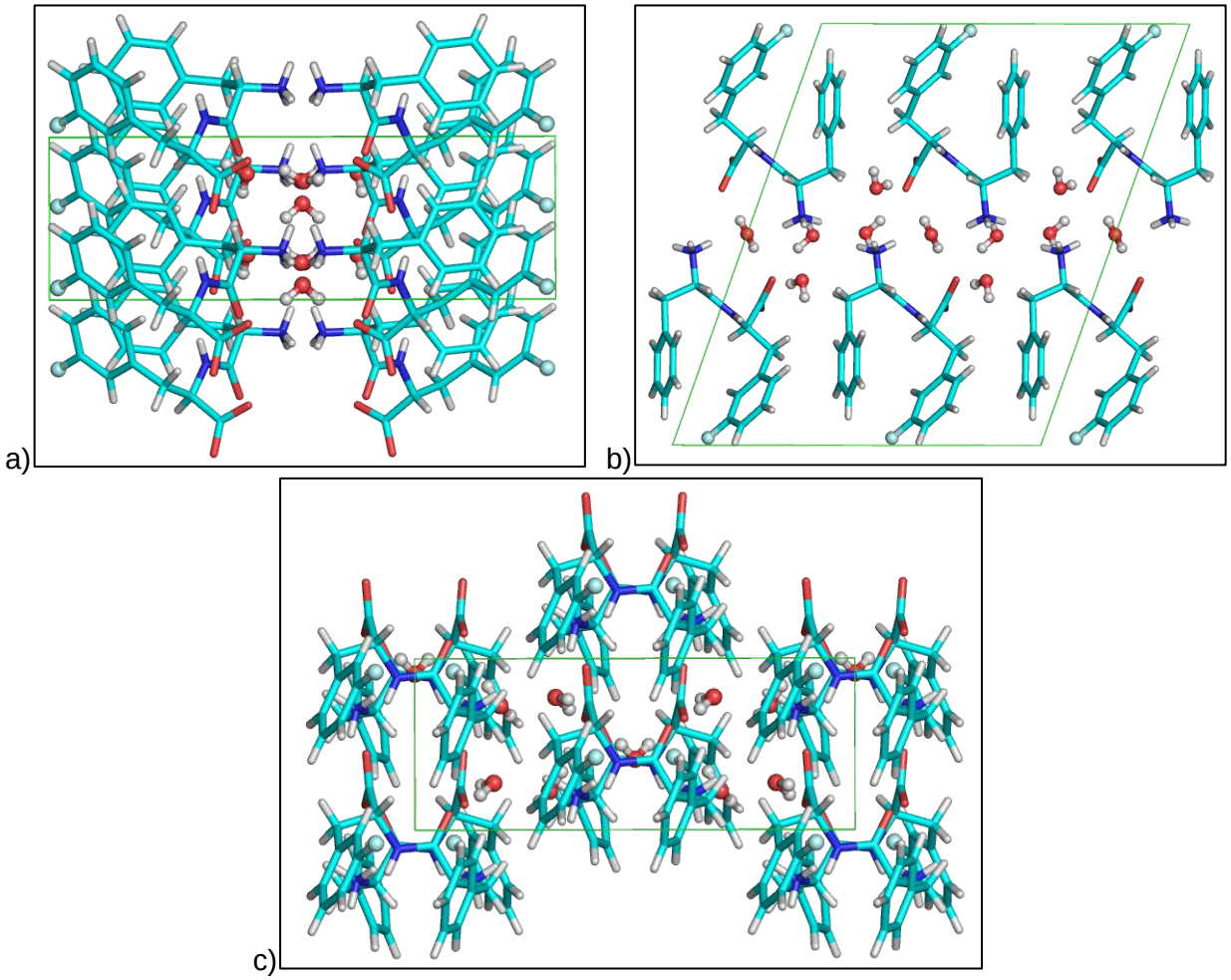


Fig. S35: Unit cell of crystals of 3-F. Crystal packing in the crystal of the peptide 3-F. Views along the *a* crystallographic axis (a), the *b* crystallographic axis (b) and the *c* crystallographic axis (c). Peptide molecules are shown as sticks, except fluorine atoms, shown as spheres; solvent molecules are shown as spheres. Carbon atoms are shown in cyan, oxygen atoms in red, nitrogen atoms in blue, fluorine atoms in light cyan, hydrogen atoms in white.

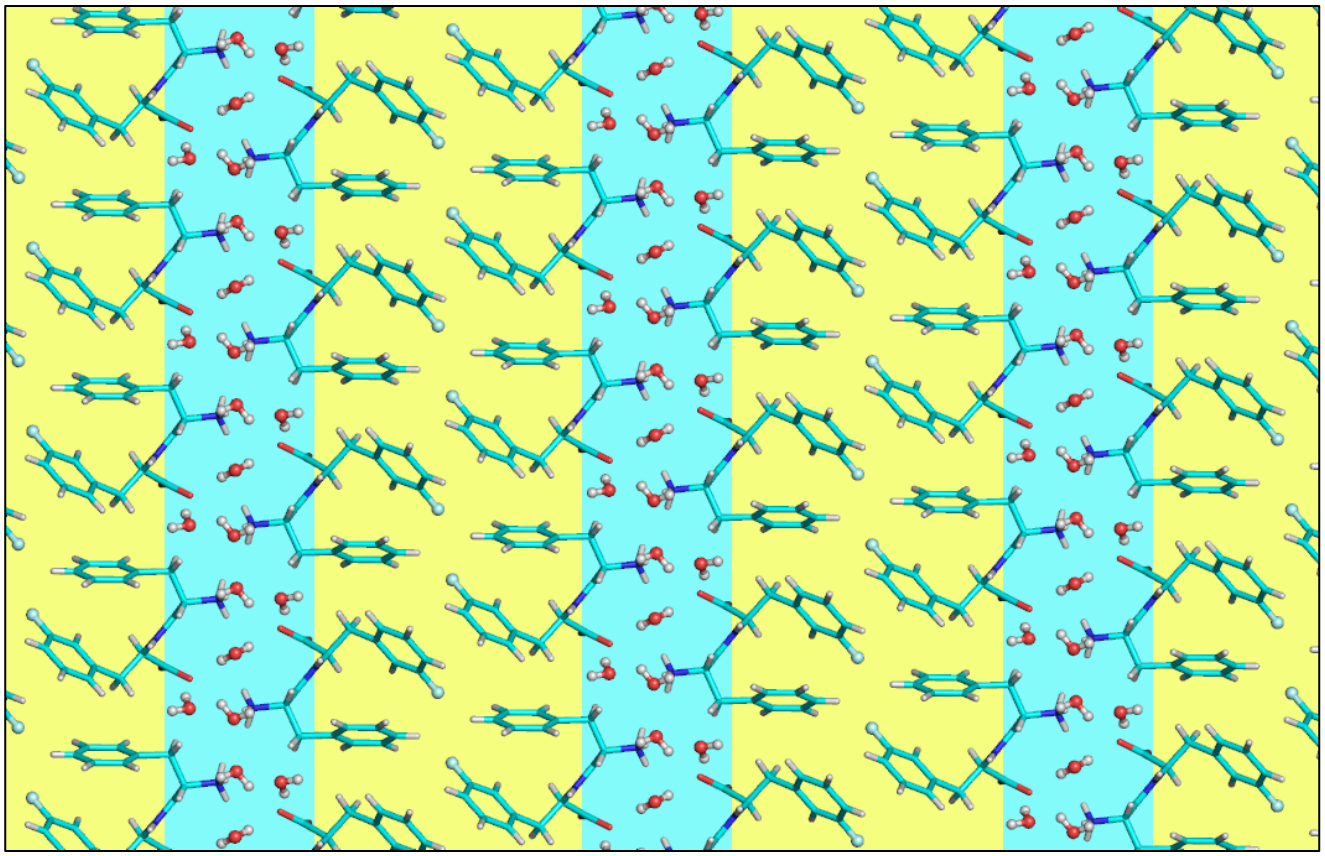


Fig. S36: Packing of peptide and solvent in the structure of 3-F. Hydrophilic and hydrophobic layers in the crystal of 3-F. Hydrophobic layers are shown with a yellow background, hydrophilic layers with a blue background. Crystal packing is shown along the *b* crystallographic direction.

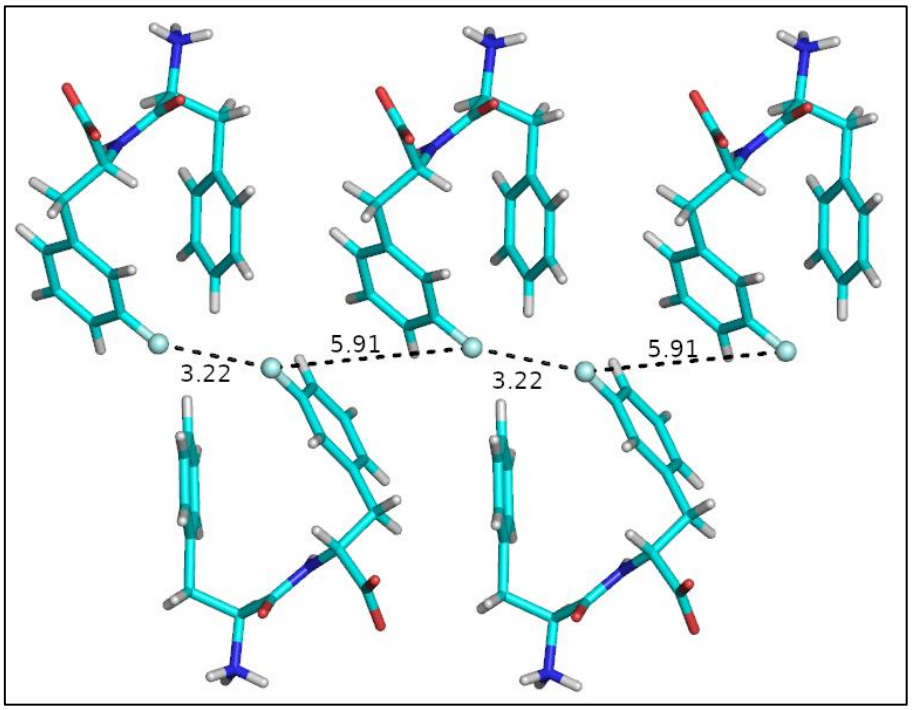


Fig. S37: Halogen-halogen distances in the structure of 3-F. Fluorine-fluorine distances within the hydrophobic layers are shown with black dashes and values are reported in Å.

Crystallographic details.

A cube-shaped single crystal of the peptide was collected with a loop, cryoprotected by dipping the crystal in glycerol and stored frozen in liquid nitrogen. The crystal was mounted on the diffractometer at the synchrotron Elettra, Trieste (Italy), beamline XRD1, using the robot present at the facility. Temperature was kept at 100 K by a stream of nitrogen on the crystal. Diffraction data were collected by the rotating crystal method using synchrotron radiation, wavelength 0.70 Å, rotation interval 1°/image, crystal-to-detector distance of 85 mm. A total of 360 images were collected to increase redundancy of data. Reflections were indexed and integrated using the XDS package [1], space group *C2* was determined using POINTLESS [2] and the resulting data set was scaled using AIMLESS [3]. Phase information were obtained by direct methods using the software SHELXT [4]. Refinements cycles were conducted with SHELXL-14 [4], operating through the WinGX GUI [5], by full-matrix least-squares methods on F^2. Unit cell parameters and scaling statistics are reported in Table S2. The asymmetric unit contains a single molecule of the peptide and 2.5 molecules of water (crystallization solvent). Hydrogen atoms of the peptide molecules were added at geometrically calculated positions and refined isotropically. Hydrogen atoms of the water molecules were manually added in the residual electron density and their position was confirmed by the analysis of the geometry of interactions. Restrains on distance (O-H distance of 0.84±0.02Å) and angles (H-H distance of 1.34±0.04Å) of the water molecules were added during refinement, using the cards DFIX and DANG of the program SHELXL-14 [4]. All the atoms within the asymmetric unit, except the hydrogen atoms, have been refined with anisotropic thermal parameters. Refinement statistics are reported in Table S2.

Table S2: Crystallographic data for 4-I and 3-F.

	4-I	**3-F**
Formula	$C_{18}H_{19}N_2O_3I_1 \cdot 2.5H_2O$	$C_{18}H_{19}N_2O_3F_1 \cdot 2.5H_2O$
Temperature (K)	100	100
Wavelength (Å)	0.7	0.7
Crystal system	Monoclinic	Monoclinic
Space group	*C 2*	*C 2*
a (Å)	16.250(3)	16.389(3)
b (Å)	6.046(1)	6.019(1)
c (Å)	20.637(4)	19.885(4)
α (°)	90	90
β (°)	97.84(3)	109.44(3)
γ (°)	90	90
V ($Å^3$)	2008.6(7)	1849.8(7)
Z, ρcalc (g/cm^3)	4, 1.598	4, 1.348
μ (mm^{-1})	1.542	0.101
F (000)	972	796
Data collection θ range	0.981 - 28.226	2.457 - 28.216
Refl. Collected / unique	16763 / 4997	15394 / 4436
Rint	0.0267	0.0341
Completeness (%)	97.5	92.8
Data/Restraints/Parameters	4997 / 9 / 261	4436 / 9 / 262
GooF	1.069	1.217
R1, wR2 [I>2σ(I)]	0.0297 / 0.0705	0.0557 / 0.1246
R1, wR2 all data	0.0307 / 0.0711	0.0557 / 0.1246

Comparison of peptide conformation and packing of 4-I and 3-F in the solid state

The conformations of 4-I and 3-F were compared by superimposing the structures. However, considering the opposite chirality of the two crystallized compounds, the structure of 4-I was first computationally inverted to represent the structure of its enantiomer, with sequence L-Phe-D-(4-I)-Phe, and then aligned with the structure of 3-F, with sequence L-Phe-D-(3-F)-Phe (Figure S38). The peptides show the same conformation as regards both the backbone and the side chains, except for the presence of the halogen atoms in different positions of the phenyl moiety of the second residue.

The analysis of the unit cell parameters, very similar except for the β angle, and space group (*C 2* for both structures) of the two crystal structures suggests that molecules are packed in a similar fashion to form the crystal. A superimposition of the supramolecular packing of the enantiomer of 4-I and 3-F shows that both peptides interact in the same was way with their polar groups within the hydrophilic layer (Figure S39). However, the difference in the β angle can be attributed to a different interaction of the hydrophobic groups that leads to a shift of neighboring hydrophilic layers in the two structures.

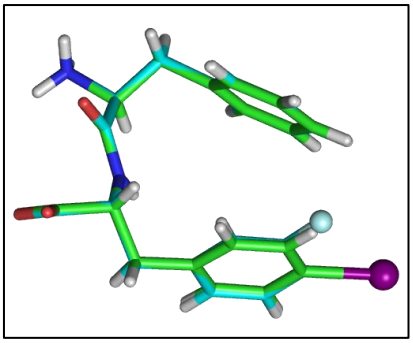


Fig. S38: Superimposition of 3-F and the enantiomer of 4-I. Enantiomer of 4-I, in green, and 3-F, in cyan, were superimposed by aligning the backbone atoms. The conformations of the two peptides show a significant similarity.

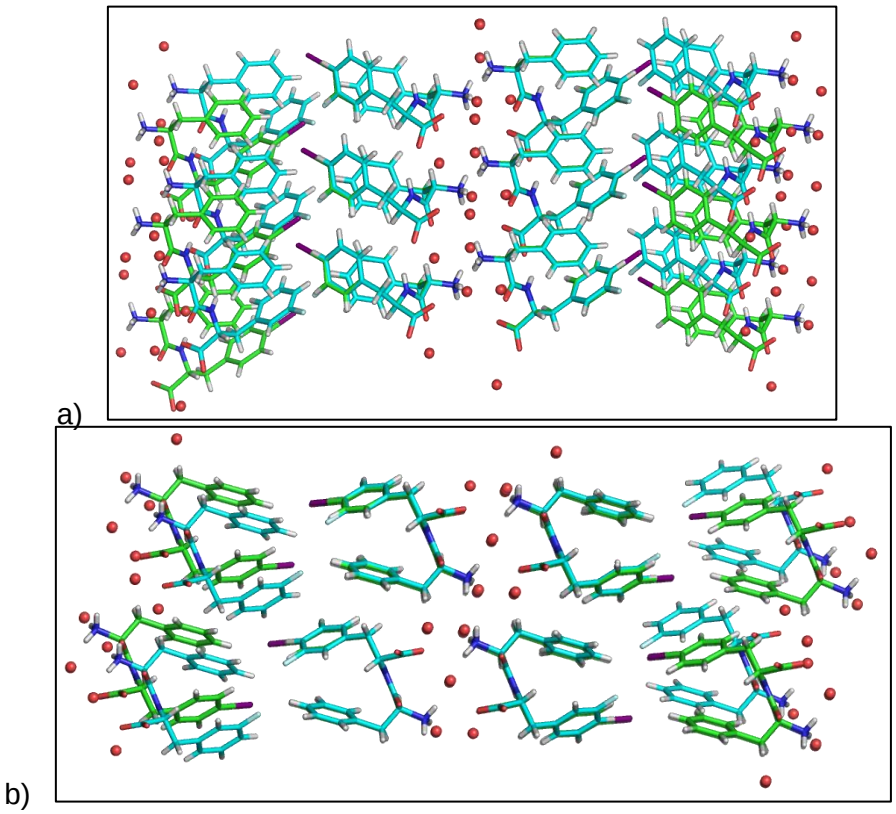


Fig. S39: Superimposition of supramolecular structures of peptide 3-F and the enantiomer of peptide 4-I. Comparison of the packing of the enantiomer of 4-I, in green, and 3-F, in cyan, shows the same pattern as regards the hydrophilic interactions, while a significant difference is observed within the hydrophobic layers. In (a), view along the *ab* diagonal direction of the unit cell; in (b), view along the *b* crystallographic direction. Peptides are represented as sticks, water molecules as spheres. Hydrogen atoms of the water molecules were omitted for clarity.

References:

[1] Kabsch, W. Acta Crystallogr., Sect. D. 2010, 66, 125−132.

[2] Evans, P. R. Acta Crystallogr., Sect. D. 2006, 62, 72−82.

[3] Evans, P. R., Murshudov, G. N. Acta Crystallogr., Sect. D. 2013, 69, 1204−14.

[4] Sheldrick, G. M. Acta Crystallogr., Sect. C. 2015, 71, 3–8.

[5] Farrugia, L.J. J. Appl. Cryst. 2012, 45, 849−854.

S10. Molecular modelling data

Molecular models of single dipeptides in implicit solvent were generated *in silico* using quantum-level calculations. Putative conformations of 2-F and 4-F were generated with Marvin version 21.20.0, ChemAxon (https://www.chemaxon.com) starting from the chemical sketches. Initial conformations of 3-F and 4-I were instead extracted from the corresponding experimental structures. Next, structural optimization was carried out with Gaussian16 [1] using the def2-SV(P) [2] basis-set and an implicit polarizable continuum model (PCM) to mimic water solvent. A harmonic analysis was performed to confirm that optimized geometries corresponded to local minima. Finally, electric dipole moments were calculated.

References:

[1] Gaussian 16, Revision A.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; *et al*. Gaussian, Inc., Wallingford CT, **2016**

[2] F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys., 2005, 7, 3297-305

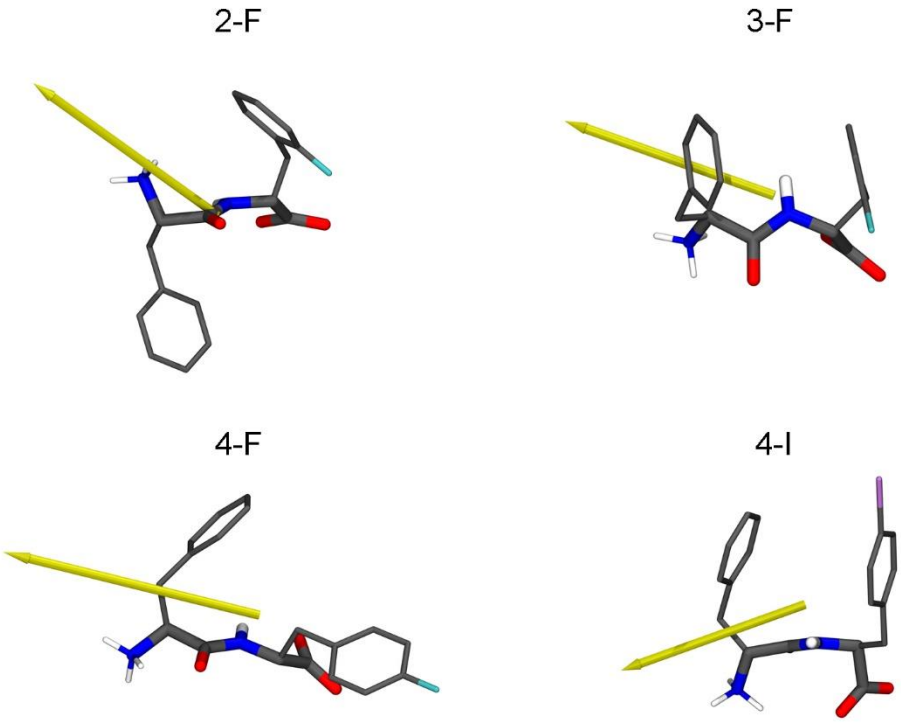


Fig. S40. Molecular dipoles orientation (yellow arrow pointing from δ- towards δ+) of the different dipeptides.

Table S3. Molecular dipole magnitude for the four dipeptides.

Molecule	Dipole [Debye]	Note
2-F	28.1	Gel
3-F	24.8	Crystal
4-F	29.2	Gel
4-I	18.7	Crystal

S11. UV Resonance Raman (UVRR) spectroscopy data

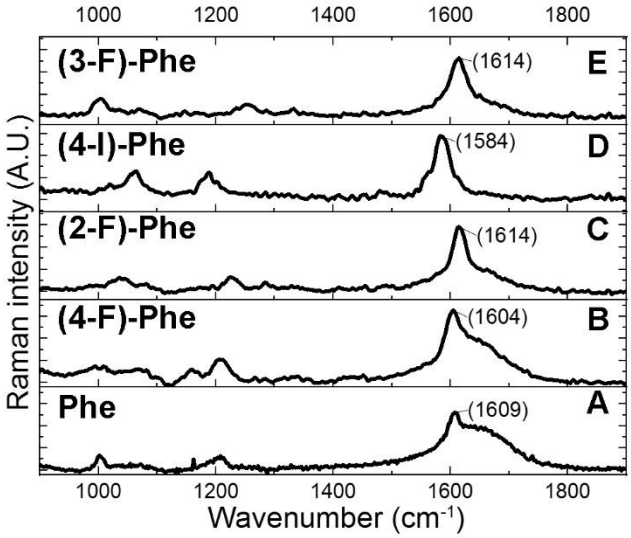


Fig. S41. UVRR spectra (excitation wavelength at 226 nm) of the halogenated amino acid precursors (devoid of the amide bond) allows unambiguous assignment of the vibrational mode of the aromatic signal in the region of interest. Samples were dissolved (1 mg/ml) in the alkaline sodium phosphate buffer (pH 11.8, 0.1 M) and measured at 298 K.
(A) L-Phe **(B)** L-(4-F)-Phe **(C)** D-(2-F)-Phe **(D)** L-(4-I)-Phe **(E)** D-(3-F)-Phe