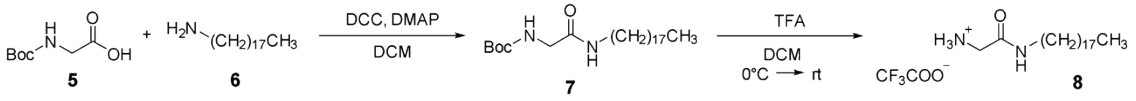
STRUCTURE OPTIMIZATION OF LIPOPEPTIDE ASSEMBLIES FOR ALDOL REACTIONS IN AN AQUEOUS MEDIUM

Bruna M. Soares, Pedro T. Sodré, Andrea M. Aguilar, Barbara B. Gerbelli,[a] Juliane N. B. D. Pelin, Karina B. Argüello, Emerson R. Silva, Marcelo A. Farias, Rodrigo V. Portugal, Carsten Schmuck, Maurício D. Coutinho-Neto, Wendel A. Alves.

1. Experimental procedures for lipopeptides synthesis

1.1. Solution phase synthesis



tert-butyl (2-(octadecylamino)-2-oxoethyl)carbamate (**7**) (Guan e Zhao, 2000)

To a solution of Boc-glycine (**5**, 0.88 g, 5.0 mmol) and octadecylamine (**6**, 1.4 g, 5.0 mmol) in anhydrous dichloromethane (40.0 mL), 4-(dimethylamino) pyridine (DMAP, 0.13 g, 1.1 mmol) was added, at 0 °C. After 30 minutes, N, N'-dicyclohexylcarbodiimide (DCC, 1.2 g, 5, 5 mmol) was added. 30 minutes later, the reaction mixture was allowed to reach the room temperature, and it remained under stirring for 5 days. After TLC analysis, DCC and DMAP were added again. The next day, the reaction mixture was filtrated and the solvent was removed under reduced pressure. The residue was purified by flash silica column chromatography technique, using a mixture of hexane:ethyl acetate (6:4) as eluent. The product **7** was obtained as a white solid with 78% yield (1.7 g, 3.9 mmol), and it was characterized by ^{1}H NMR, ^{13}C NMR and MS.

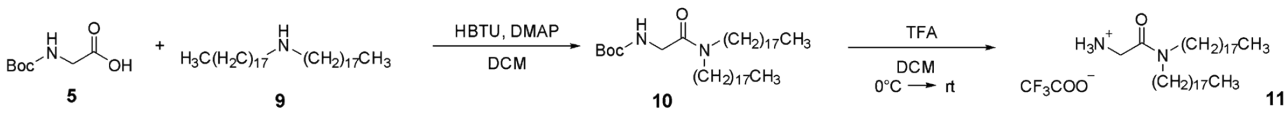
^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 6.18 (br s, 1H, N*H*-amide), 5.19 (br s, 1H, N*H*-amide Boc), 3.75 (d, *J* = 6.0 Hz, 2H, NHCH_2CO), 3.24 (dd, *J* = 6.0 Hz, 9.0 Hz, 2H, OCNHCH_2), 1.45 (s, 9H, C(CH_3)$_3$), 1.51-1.40 (m, 11H, C(CH_3)$_3$ and OCNHCH$_2$CH_2), 1.24 (s, 30H, CH_2-alkyl chain), 0.87 (t, *J* = 6.0 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, $CDCl_3$):** δ (ppm) 169.4 (*C*ONH(CH$_2$)$_{17}$), 156.2 (C(CH$_3$)$_3$O*C*O), 80.4 (*C*(CH$_3$)$_3$), 49.3 (NH*C*H$_2$CO), 44.8 (NH*C*H$_2$CH$_2$), 39.7-22.8 (16x *C*H$_2$), 28.4 (C(*C*H$_3$)$_3$), 14.2 (*C*H$_3$). **MS (ESI+, *m*/*z*):** 449.4058 [M+Na]$^{+}$; calculated for $C_{25}H_{50}N_2NaO_3^{+}$: 449.3714.

2-(octadecylamino)-2-oxoethan-1-aminium 2,2,2-trifluoroacetate (**8**) (Dai et al., 2014)

To a solution of **7** (0.64 g, 1.5 mmol) in anhydrous DCM (17.0 mL), trifluoroacetic acid (TFA, 0.9 mL, 11.6 mmol) was added slowly at 0 °C, under inert nitrogen atmosphere. Then, the ice bath and the inert gas flow were removed and the reaction mixture was stirred for 19 hours. After adding a portion of methanol (20.0 mL) and removing the solvent under reduced pressure, the deprotected product was obtained as a white solid in quantitative yield (0.73 g,

1.65 mmol). The product was characterized by ^{1}H NMR, ^{13}C NMR and MS, and it was submitted to the next step without further purification.

^{1}H NMR (300 MHz, DMSO): δ (ppm) 8.32 (t, *J* = 6.0 Hz, 1H, N*H*-amide), 8.04 (br s, 3H, NH_3^+), 3.51 (br s, 2H, CH_2CO), 3.10 (q, *J* = 6.0 Hz, 2H, NHCH_2), 1.52-1.38 (m, 2H, NHCH$_2$CH_2), 1.23 (s, 30H, CH_2-alkyl chain), 0.85 (t, *J* = 6.0 Hz, 3H, CH_3). **^{13}C NMR (75 MHz, DMSO):** δ (ppm) 165.6 (*C*O), 47.5 (*C*H$_2$CO), 33.3-22.1 (17x *C*H$_2$), 13.9 (*C*H$_3$). **MS (ESI+, *m*/*z*):** 327.3645 $[M+H]^+$; calculated for $C_{20}H_{43}N_2O^+$: 327.3370.



tert-butyl (2-(dioctadecylamino)-2-oxoethyl)carbamate (**10**) (Miller et al., 2005; Kamaly et al., 2008)

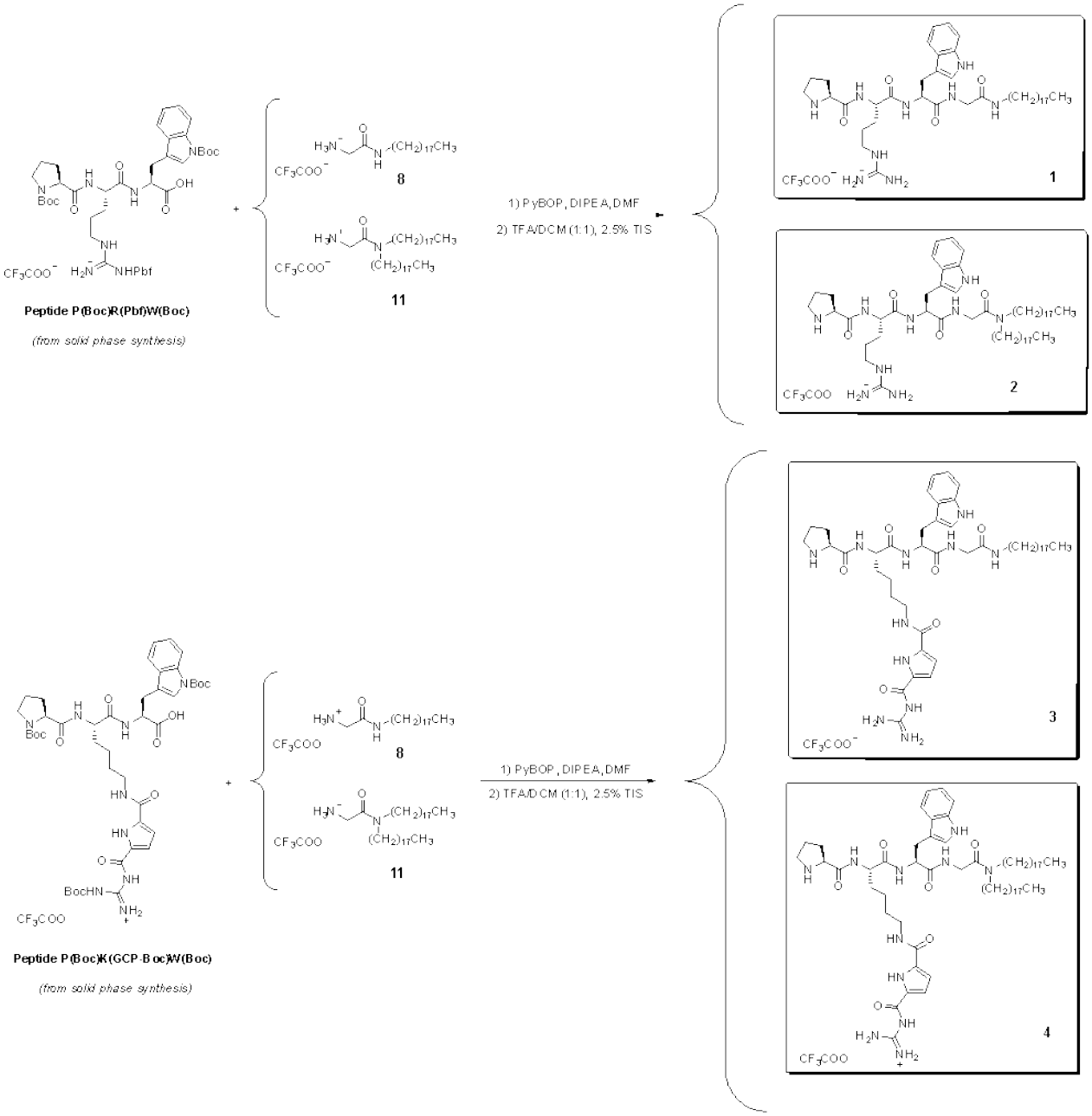
To a solution of Boc-glycine (**5**, 0.88 g, 5.0 mmol) and dioctadecylamine (**9**, 2.6 g, 5.0 mmol) in anhydrous dichloromethane (40.0 mL), HBTU (2.3 g, 6.0 mmol) and DMAP (1.8 g, 15.0 mmol) were added. The reaction mixture remained under stirring for 19 hours under nitrogen atmosphere at room temperature. Then the reaction mixture washed with water (3x) and extracted with DCM/methanol (2x). After concentration under reduced pressure, it was obtained a yellow oil. The oil was solubilized in diethyl ether and washed with citric acid 7%, water and brine. After drying with magnesium sulphate and filtration, the solvent was removed under reduced pressure. The residue was purified by flash silica column chromatography technique, using a mixture of hexane:ethyl acetate (7:3) as eluent. The product **10** was obtained as a colourless oil with 77% yield (2.6 g, 3.9 mmol), and it was characterized by ^{1}H NMR, ^{13}C NMR and MS.

^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 5.55 (br s, 1H, N*H*-amide); 3.93 (s, 2H, NHCH_2CO); 3.30 (dd, *J* = 6.0 Hz, 9.0 Hz, 2H, OCNCH_2); 3.13 (dd, *J* = 6.0 Hz, 9.0 Hz, 2H, OCNCH_2); 1.44 (s, 9H, C(CH_3)$_3$); 1.58-1.44 (m, 13H, C(CH_3)$_3$ and 2x OCNHCH$_2$CH_2); 1.25 (s, 60H, 2x CH_2-alkyl chains); 0.90-0.85 (m, 6H, 2x CH_3). **^{13}C NMR (75 MHz, $CDCl_3$):** δ (ppm) 167.8 (*C*ON), 156.0 (C(CH$_3$)$_3$O*C*O), 79.6 (*C*(CH$_3$)$_3$), 47.1 (NH*C*H$_2$CO), 46.3 (N*C*H$_2$CH$_2$-alkyl chain 1), 42,3 (N*C*H$_2$CH$_2$-alkyl chain 2), 32.1-22.8 (32x *C*H$_2$), 28.5 (C(*C*H$_3$)$_3$), 14.2 (2x *C*H$_3$). **MS (ESI+, *m*/*z*):** 701.7554 $[M+Na]^+$; calculated for $C_{43}H_{86}N_2NaO_3^+$: 701.6531.

2-(dioctadecylamino)-2-oxoethan-1-aminium 2,2,2-trifluoroacetate (**11**) (Dai et al., 2014)

Same procedure as already described for compound **8**. The product **11** was obtained as a white solid in quantitative yield (1.2 g, 1.8 mmol). The product was characterized by ^{1}H NMR and MS, and it was submitted to the next step without further purification.

^{1}H NMR (300 MHz, $CDCl_3$): δ (ppm) 3.89 (s, 2H, CH_2CO), 3.29 (dd, J = 6.0 Hz, 9.0 Hz, 2H, NCH_2-alkyl chain 1), 3.12 (dd, J = 6.0 Hz, 9.0 Hz, 2H, NCH_2-alkyl chain 2), 1.58-1.40 (m, 4H, 2x $NHCH_2CH_2$), 1.25 (s, 60H, 2x CH_2-alkyl chains), 0.90-0.85 (m, 6H, 2x CH_3). **MS (ESI+, *m*/*z*):** 579.7119 $[M+H]^+$; calculated for $C_{38}H_{79}N_2O^+$: 579.6187.



Lipopeptide PRWG-C_{18} (**1**): Representative procedure for the coupling between the peptide and alkyl chain

To a solution containing the tripeptide obtained from solid phase synthesis (0.55 mmol) and the alkyl chain compound (0.66 mmol) in dimethylformamide (DMF, 10 mL) was added (Benzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate (PyBOP, 0.61 mmol) and N, N-diisopropylethylamine (DIPEA, 1.1 mmol). After 2 days under stirring, new portions of PyBOP and DIPEA were added, in the same amounts. After around 18 hours, the solvent was removed under reduced pressure and cold water was added to the residue. Then, filtration and washing with cold water was performed under reduced pressure. Subsequently, the residue was solubilized in DCM (5 mL), and to this solution were added triisopropylsilane (TIS, 250 μL) and TFA (5 mL). After 26 hours of stirring, a portion of methanol was added and the solvent was removed under reduced pressure. This step was repeated. The crude product was obtained as a yellow oil, which was submitted to purification in reverse phase MPLC (H_2O → MeOH, 0.05% TFA). In that way, the PRWG-C_{18} lipopeptide (**1**) was obtained in 37% yield (0.16 g, 0.19 mmol) and was characterized by high resolution mass spectroscopy.

MS (ESI+, *m*/*z*): 766.6480 $[M+H]^+$; calculated for $C_{42}H_{72}N_9O_4^+$: 766.5702. **HRMS (ESI+, *m*/*z*):** 766.5734 $[M+H]^+$; calculated for $C_{42}H_{72}N_9O_4^+$: 766.5702 (Error: 4.2 ppm).

Lipopeptide PRWG-$(C_{18})_2$ (**2**)

In the end of the process, a portion of methanol was added to the crude product and the solvent was removed under reduced pressure. For this peptide, this step was repeated with diethyl ether. The crude product was obtained as a mixture of white solid and oil, which was submitted to purification in reverse phase MPLC. In that way, the PRWG-$(C_{18})_2$ lipopeptide (**2**) was obtained in 76% yield (0.44 g, 0.39 mmol) and was characterized by high resolution mass spectroscopy.

MS (ESI+, *m*/*z*): 1018.9888 $[M+H]^+$; calculated for $C_{60}H_{108}N_9O_4^+$: 1018.8519. **HRMS (ESI+, *m*/*z*):** 1018.8561 $[M+H]^+$; calculated for $C_{60}H_{108}N_9O_4^+$: 1018.8519 (Error: 4.1 ppm).

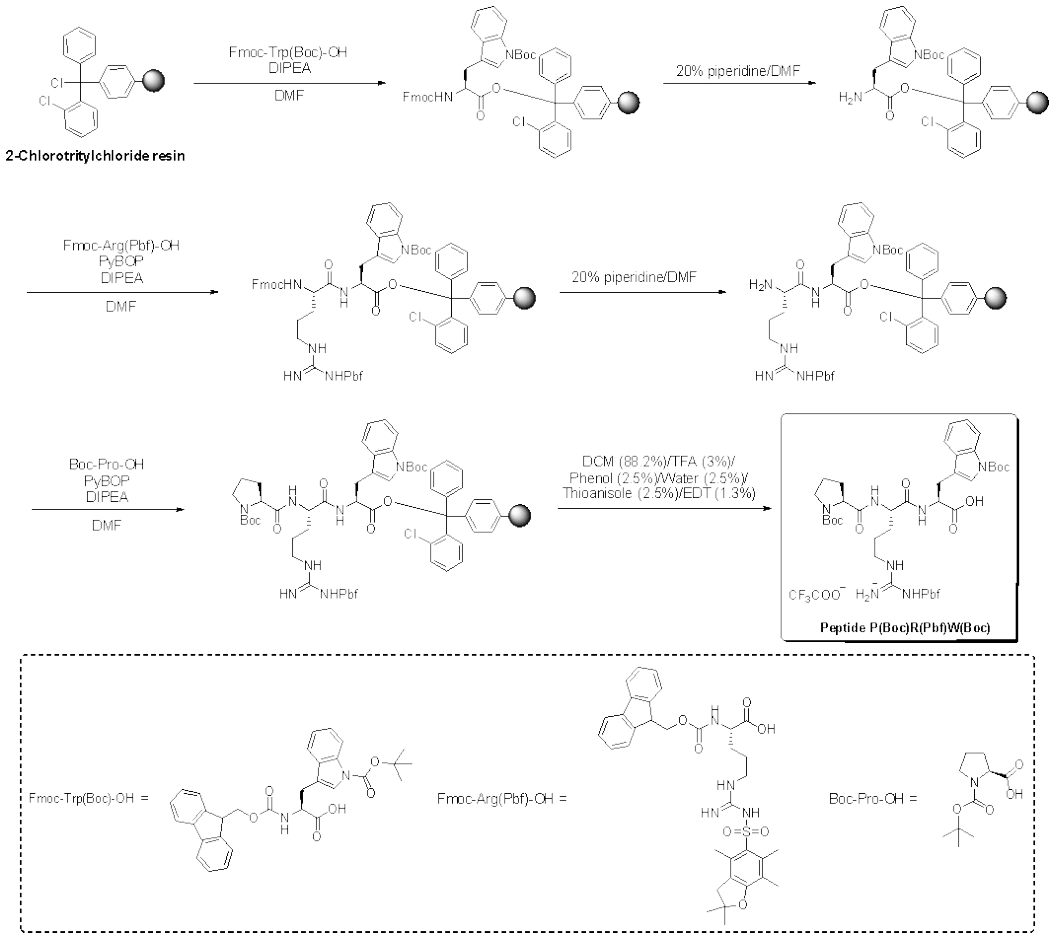
Lipopeptide PK(GCP)WG-C_{18} (**3**)

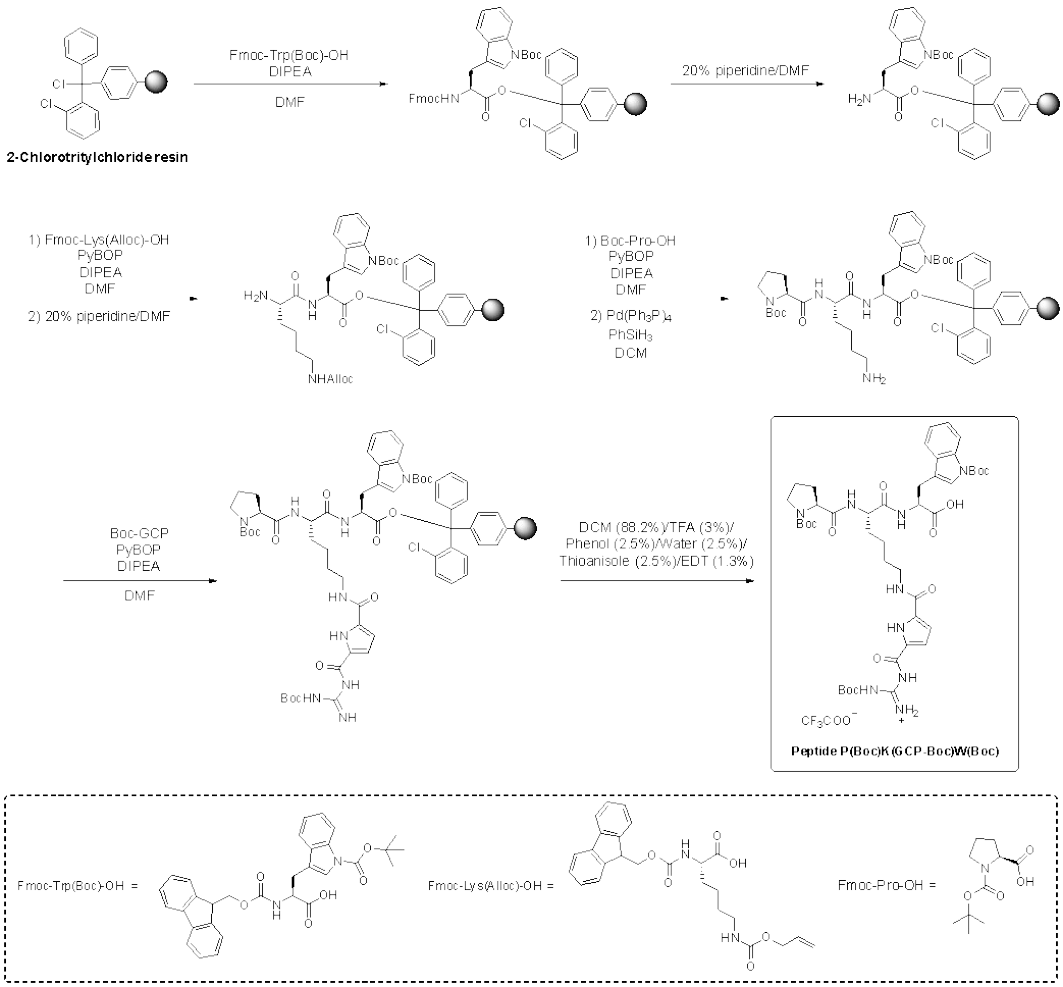
In the end of the process, a portion of methanol was added to the crude product and the solvent was removed under reduced pressure. For this peptide, this step was repeated three times. The crude product was obtained as a salmon solid, which was submitted to purification in reverse phase MPLC. In that way, the PK(GCP)WG-C_{18} lipopeptide (**3**) was obtained in 52% yield (0.13 g, 0.125 mmol) and was characterized by high resolution mass spectroscopy. **MS (ESI+, *m*/*z*):** 916.7328 $[M+H]^+$; calculated for $C_{49}H_{78}N_{11}O_6^+$: 916.6131. **HRMS (ESI+, *m*/*z*):** 916.6115 $[M+H]^+$; calculated for $C_{49}H_{78}N_{11}O_6^+$: 916.6131 (Error: 1.7 ppm).

Lipopeptide PK(GCP)WG-$(C_{18})_2$ (**4**)

In the end of the process, a portion of methanol was added to the crude product and the solvent was removed under reduced pressure. For this peptide, this step was repeated twice. The crude product was obtained as an orange solid, which was submitted to purification in reverse phase MPLC. In that way, the PK(GCP)WG-$(C_{18})_2$ lipopeptide (**4**) was obtained in 81% yield (0.28 g, 0.22 mmol) and was characterized by high resolution mass spectroscopy. **MS (ESI+, *m*/*z*):** 1169.0139 $[M+H]^+$; calculated for $C_{67}H_{114}N_{11}O_6^+$: 1168.8948. **HRMS (ESI+, *m*/*z*):** 1168.8995 $[M+H]^+$; calculated for $C_{67}H_{114}N_{11}O_6^+$: 1168.8948 (Error: 4.0 ppm).

1.2. Solid phase synthesis





1.2.1. General procedures for the SPPS

Standard solid phase peptide synthesis was performed in Schlenck glass vessels equipped with a frit and a stopper. The reaction mixtures were shaken on the orbital shaker under argon atmosphere. The solution was removed by filtration after the reaction and the resin beads were washed with the solvent.

Kaiser Test

To detect the free amino groups on the resin, two separate solutions of Ninhydrin (1 g) and phenol (40 g), both in ethanol (10 mL), were prepared for the *Kaiser* test. Then a few resin beads were taken from the reaction and added to a mixture of the two solutions (0.3 mL from each one), which was heated for 1 min at 100-110 °C. Resin beads with free amino

functionalities (-NH_2) were colored dark blue or red while the resin beads without amino functions stayed white.

Fmoc Deprotection

Fmoc protecting group was removed by treatment with 20% piperidine in DMF (2x 30 min). Then the resin was thoroughly washed with DMF (6x 5 min). The completion of the Fmoc deprotection with the formation of free amino groups was monitored by a positive *Kaiser* test and repeated if necessary.

Alloc Deprotection

The removal of Alloc protecting group was achieved by treatment with $Pd(PPh_3)_4$ (0.1 eq.) in the presence of $PhSiH_3$ (24 eq.) and DCM for 2 hours. Then the resin was washed with DCM (4x 5 min) and DMF (2x 5 min). The completion of the Alloc deprotection was monitored by a positive *Kaiser* test and repeated if necessary.

Coupling Procedure

Initially, the 2-chlorotritylchloride resin was swollen in DMF for 2 h. Then the first amino acid (2 eq.) was attached to the resin (1 eq.) using DIPEA (3 eq.) as the base in DMF (from swelling) under argon atmosphere at room temperature by shaking the reaction mixture overnight. Then the resin was washed with DMF (3x 5 min). The resin was capped using MeOH (HPLC grade) for 15 minutes, and then washed with DMF (3x 5 min). Finally, the Fmoc protecting group was removed by agitation in 20% piperidine/DMF as described above. The next couplings were promoted using the amino acid (2 eq.), PyBOP (2.1 eq.) as the coupling reagent and DIPEA (3 eq.) as the base in DMF. After the washing step using DMF (3x), the attachment of the amino acid was monitored by a negative *Kaiser* test. The coupling was repeated if necessary.

Cleavage from the Resin

After the completion of all the coupling steps (and Fmoc deprotection if necessary), the resin was thoroughly washed with DCM (3 ×5 min), MeOH (3 ×5 min) and again DCM (3 ×5 min). Then the resin was dried under vacuum for 1 h. The cleavage of the product from the

2-chlorotritylchloride resin was achieved by adding a mixture of 88.2% DCM/3% TFA/2.5% Phenol/2.5% H_2O/2.5% Thioanisole/1.3% EDT. Then the suspension was shaken for 15 minutes at room temperature and washed with DCM for 10-15 seconds. After that, an acid removal step was promoted with rotoevaporation (31°C) and additions of dichloromethane and methanol sequentially. It was obtained a residue to which a mixture of diethyl ether and hexane or only diethyl ether was added to precipitate the peptide. The sample was left in the fridge overnight. The supernadant was removed and the sample was submitted to rotoevaporation. The crude product was analyzed by mass spectroscopy. Finally, the material was purified by RP_{18}-MPLC using appropriate conditions (H_2O/MeOH + 1% TFA), and the product was analyzed by mass spectroscopy again.

Peptide P(Boc)R(Pbf)W(Boc)

The crude peptide was precipitated using a mixture of hexane and diethyl ether. It was purified by reverse phase MPLC (H_2O – MeOH), and characterized by MS.

MS (ESI+, *m*/*z*): 910.5110 $[M+H]^+$; calculated for $C_{45}H_{64}N_7O_{11}S^+$: 910.4379.

Peptide P(Boc)K(GCP-Boc)W(Boc)

The crude peptide was precipitated using diethyl ether. It was purified by reverse phase MPLC (H_2O – MeOH), and characterized by MS.

MS (ESI+, *m*/*z*): 908.5391 $[M+H]^+$; calculated for $C_{44}H_{62}N_9O_{12}^+$: 908.4512.

1.3. Molecular dynamics simulation protocol

The simulation consisted of the following steps: system minimization, NVT equilibration, NPT equilibration, and NPT production. Particular attention should be paid to the initial minimization and equilibration steps, as these systems typically need to relax their initial configuration to remove close contacts in the structure. Hence, our initial NVT equilibration procedure restricts the lipidic group of 2 kcal/Å2 to ensure the system's stability during the heating phase to 300K for 200 ps. Next, a short restrained 300 ps NPT step at P=1ATM is performed, followed by 5 ns of unrestrained molecular dynamics at T=300K and P = 1 atm. A production of 100ns is performed for each of the simulated systems using the GPU implementation of AMBER 18. We used a Berendsen barostat with a 1.0 ps relaxation time

and pressure equals 1 atm, a Langevin thermostat with a 2 ps^{-1} collision frequency, and a timestep of 0.02 fs. Shake was used to constrain the Hydrogen to their equilibrium positions.

1.4. Quantum mechanical calculations

pKa calculations: Free energies for the deprotonation processes (□G) were computed for protonated L-proline methyl ester (PME) and Trifluoroacetic acid (TFA) in three distinct environments, water, cyclohexanone and N-heptane. The results are summarized in table S2, reporting both □G and pKa. Geometries were optimized, followed by a frequency calculation to obtain the free energy. Computed results used bare and two water molecules (2W values) in H bond configurations with the solute (shown in figure S11). Solvation energies were calculated using the SMD method. □G and pKa were computed using a ma-def2-TZVP basis sets and B3LYP functional as implemented in Orca code version 4.2.1 in all calculations. □G and pKa results for the bare solute used a proton solvation energy in water of -265.9 $kcal \cdot mol^{-1}$.

Proton transfer reactions: Table S4 shows results for the proton transfer reaction of L-proline methyl ester and trifluoroacetic anion with and without microsolvation. In the microsolvation approach, explicit water molecules are added as proton donors and acceptors to the solute molecules. Solvation energies were computed using the SMD method as show in Figure S11.

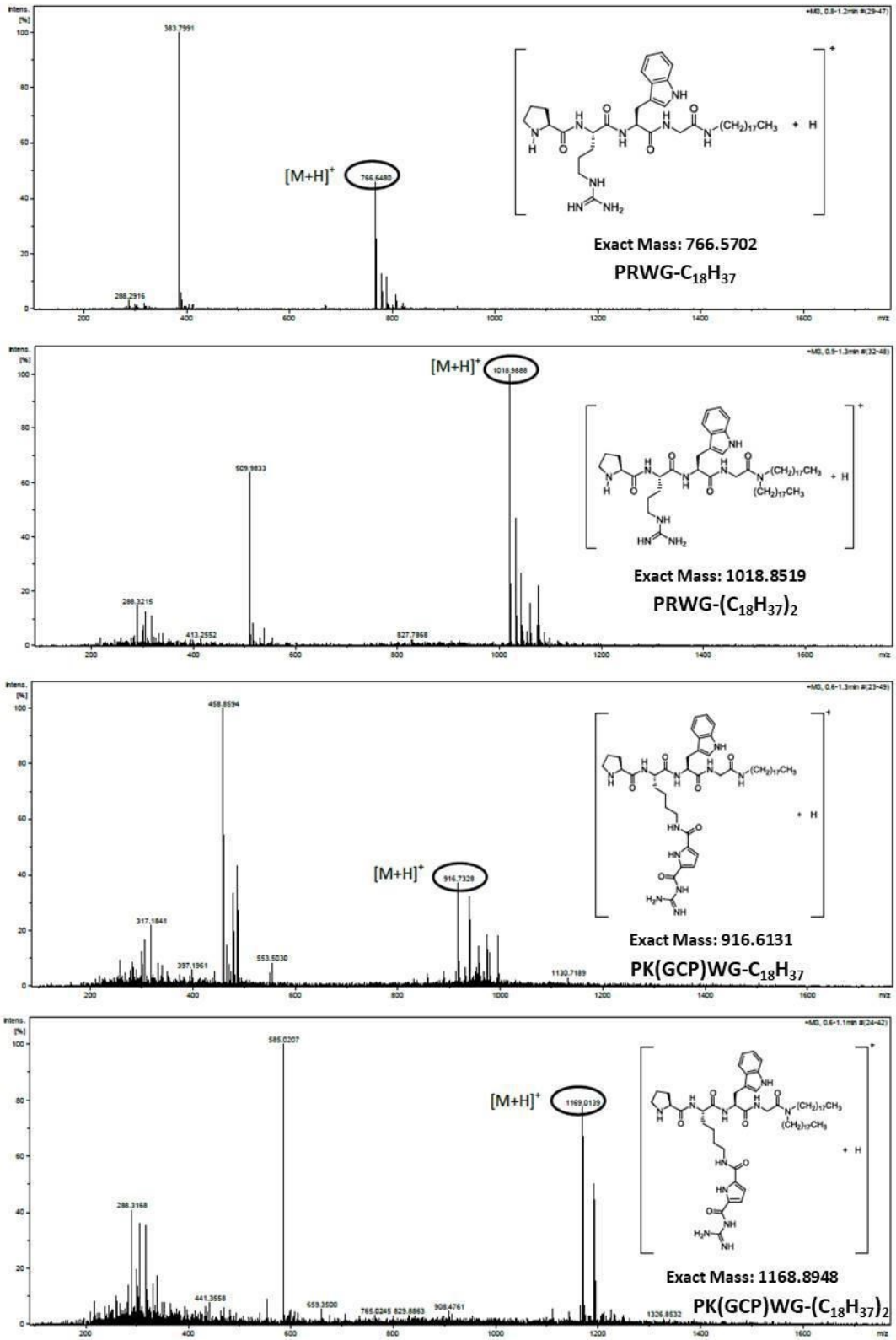


Figure S1. Mass spectrometry analysis of PRWG($C_{18}H_{37}$), PRWG($C_{18}H_{37}$)$_2$, PK(GCP)WG-$C_{18}H_{37}$ and PK(GCP)WG-($C_{18}H_{37}$)$_2$.

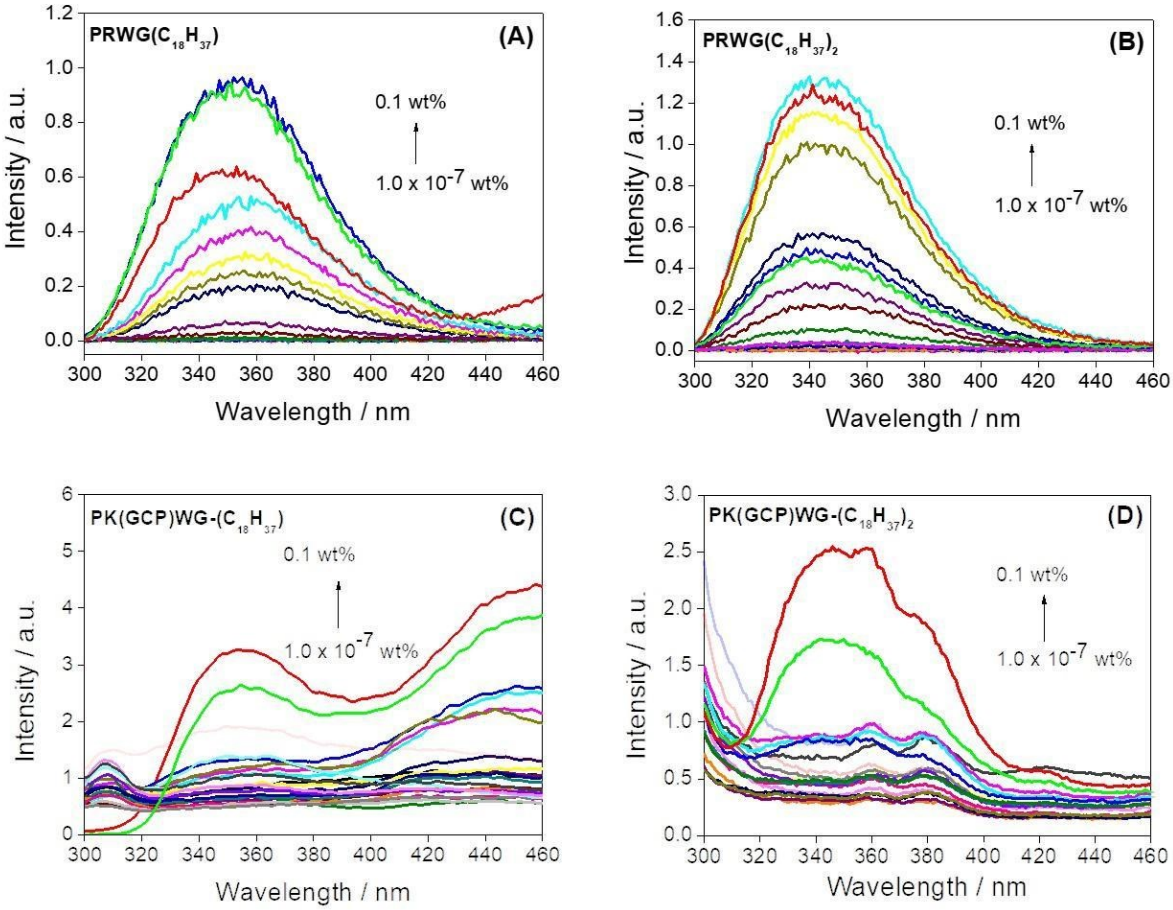


Figure S2. Fluorescence emission spectra at native conditions of (A) PRWG-$C_{18}H_{37}$, (B) PRWG-$(C_{18}H_{37})_2$, (C) PK(GCP)WG-$C_{18}H_{37}$, and (D) PK(GCP)WG-$(C_{18}H_{37})_2$ considering the intrinsic tryptophan emission.

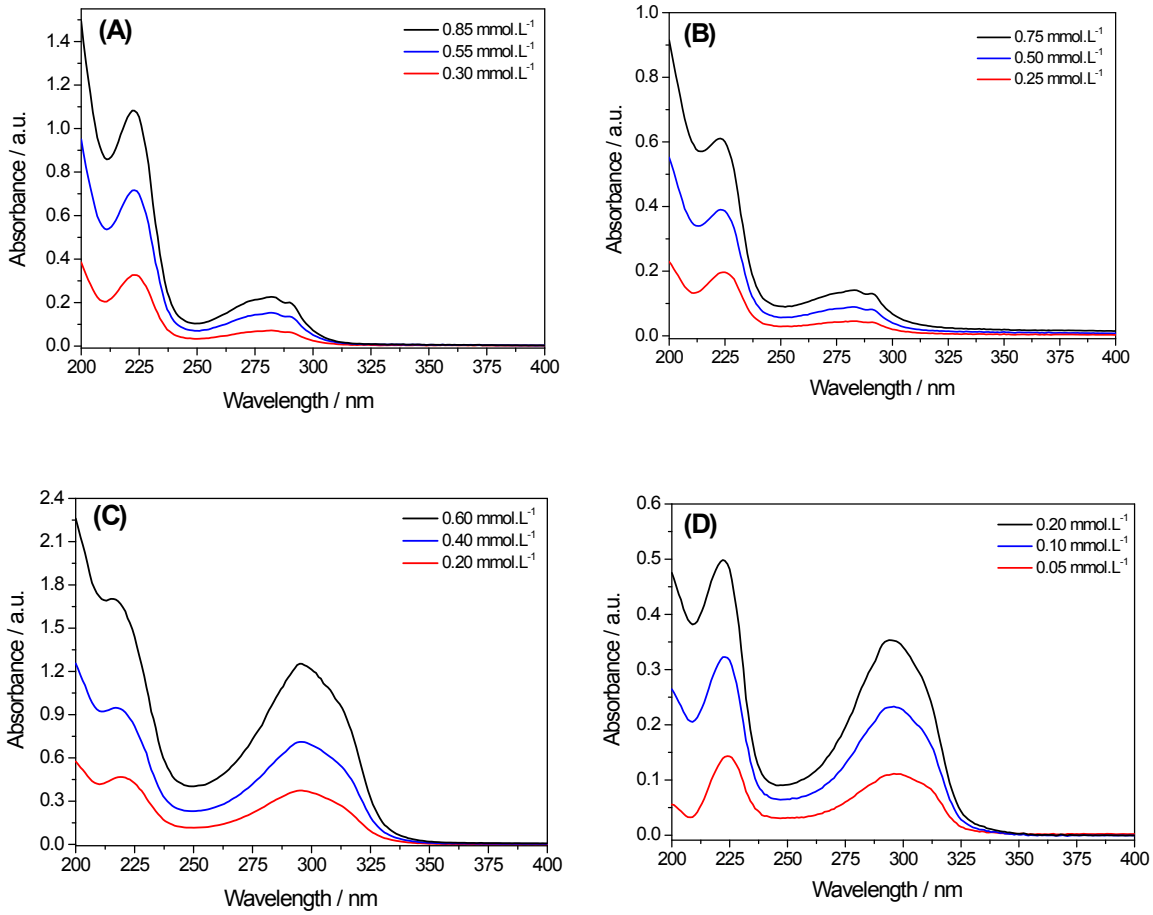


Figure S3. UV-Vis absorption spectra of (A) PRWG-$C_{18}H_{37}$, (B) PRWG-$(C_{18}H_{37})_2$, (C) PK(GCP)WG-$C_{18}H_{37}$, and (D) PK(GCP)WG-$(C_{18}H_{37})_2$ in aqueous solutions.

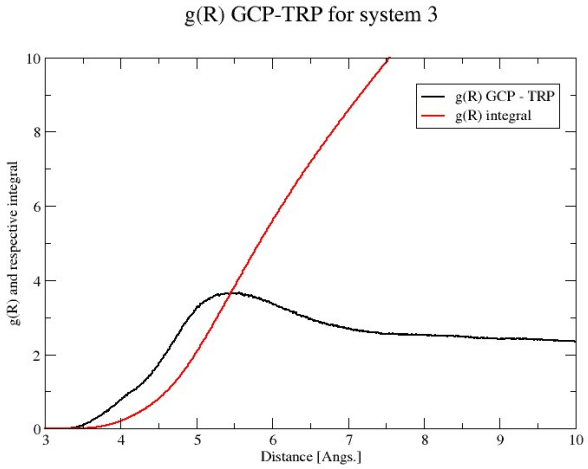


Figure S4. Radial distribution function and its integral for the tryptophan ring and GCP pair.

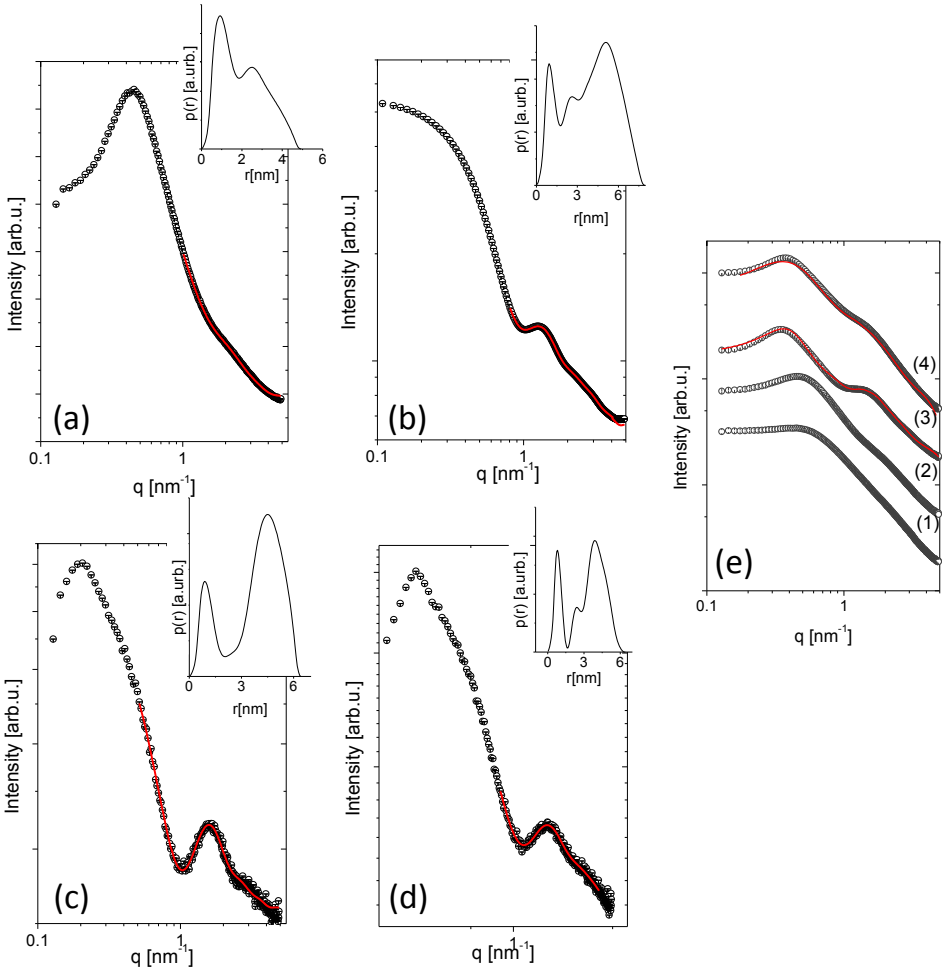


Figure S5. SAXS data with IFT fitting with the *p*(r) as a function of ratio for the compounds (A) (**1**), (B) (**2**), (C) (**3**) and (D) (**4**). (E) SAXS scattering with 1wt%.

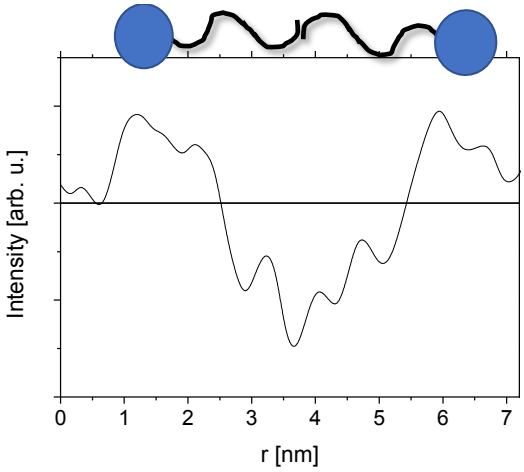


Figure S6. Contrast profile of the Cryo-TEM image for compound (**3**) with 1 wt%.

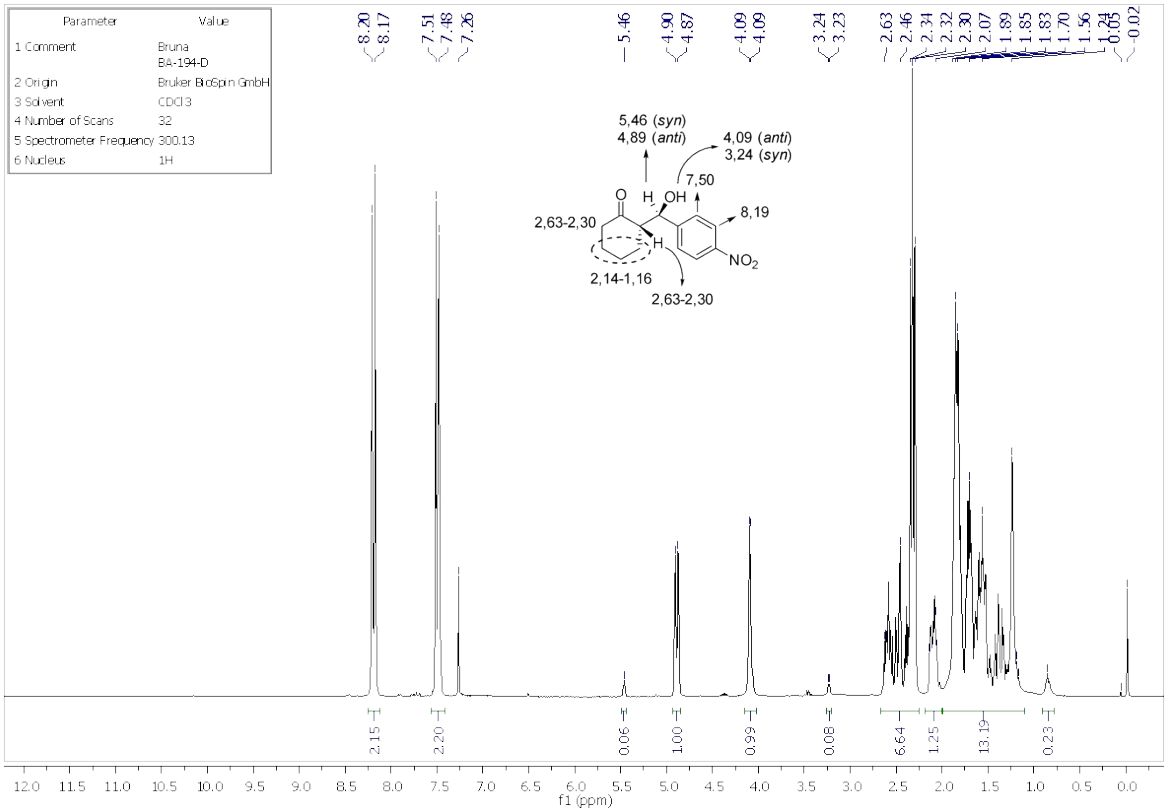
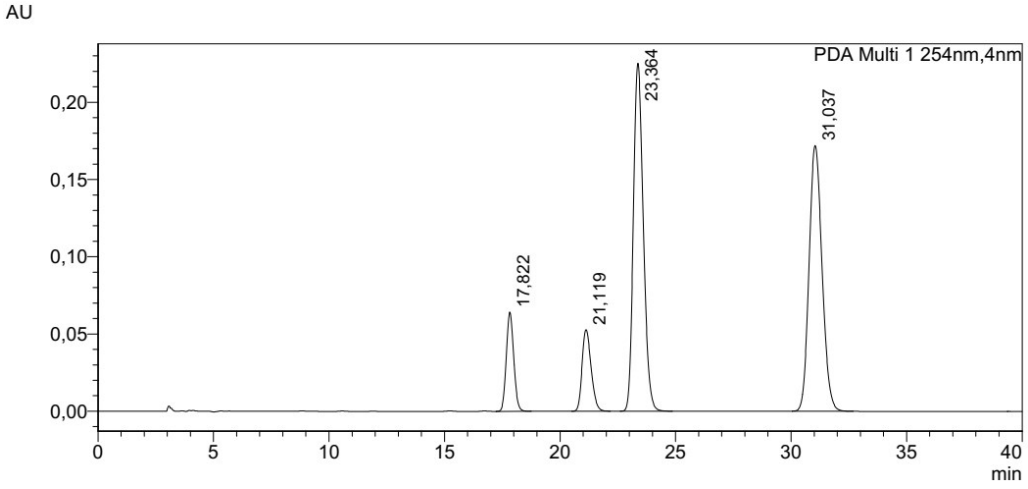


Figure S7. Representative ^{1}H NMR spectra of crude aldol products are described in Table 2.

a) Racemic mixture

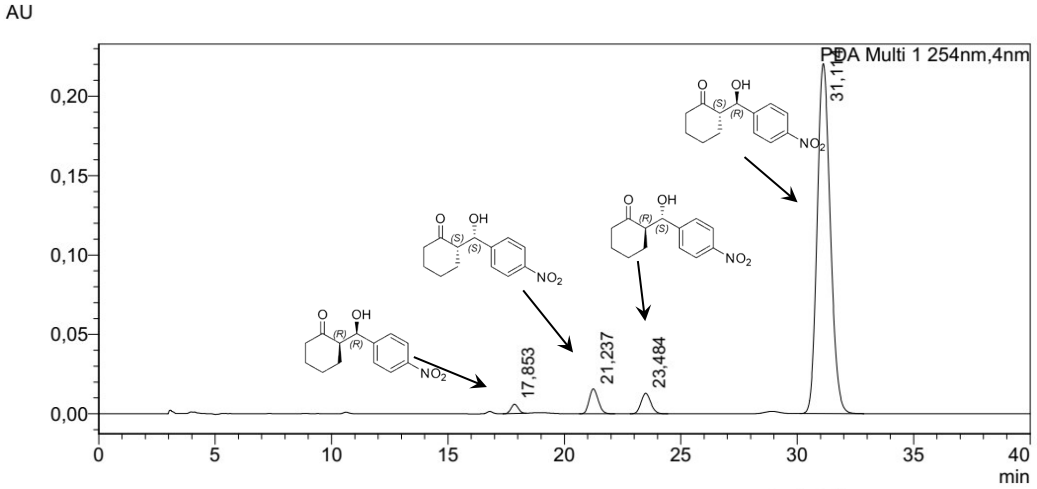


Peak Table

PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Area%	Height%
1	17,822	1440417	64311	8,825	12,505
2	21,119	1437553	52808	8,808	10,269
3	23,364	6726865	225190	41,215	43,788
4	31,037	6716522	171959	41,152	33,438
Total		16321358	514269	100,000	100,000

*b) Aldol reaction catalyzed by lipopeptide **4** (representative).*



Peak Table

PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Area%	Height%
1	17,853	126515	5871	1,320	2,303
2	21,237	427083	15779	4,456	6,189
3	23,484	378106	12911	3,945	5,064
4	31,114	8652741	220394	90,279	86,444
Total		9584445	254955	100,000	100,000

Figure S8. Chiral-phase HPLC chromatogram for aldol products. Conditions: Chiralpak AD-H, hexane/2-propanol (90/10), 30ºC, 1.0 mL·min-1, λ= 254 nm.

Table S1. Aldol reaction analysis over time.

Peptides	*Time (h)*	*Conversion (%)*	*ds*	*ee (%)*
(**1**)	0.5	33	92	85
(**2**)	0.5	16	90	83
(**3**)	0.5	31	92	88
(**4**)	0.5	24	91	85
(**1**)	1	36	92	84
(**2**)	1	20	92	86
(**3**)	1	38	93	89
(**4**)	1	32	93	86
(**1**)	2	49	91	83
(**2**)	2	30	91	82
(**3**)	2	42	93	88
(**4**)	2	55	93	86
(**1**)	4	70	91	80
(**2**)	4	74	90	74
(**3**)	4	63	92	87
(**4**)	4	68	93	88
(**1**)	6	>99	90	88
(**2**)	6	79	92	87
(**3**)	6	78	93	88
(**4**)	6	>99	92	89
(**1**)	48	>99	93	90
(**2**)	48	>99	93	90
(**3**)	48	>99	93	90
(**4**)	48	>99	94	92

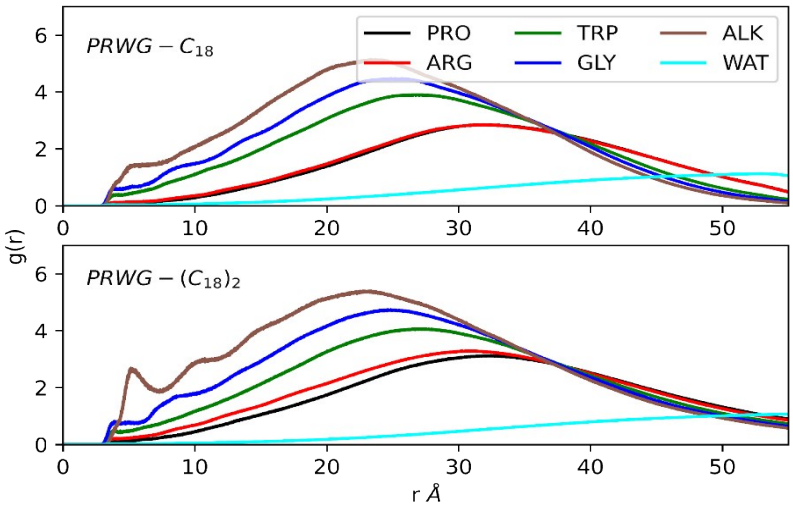


Figure S9. RDFs for the PRWG-C_{18} (**1**) (top) and PRWG-$(C_{18})_2$ (**2**) (bottom) systems show a layered structural organization of residues with respect to the micellar center, defined as the average position of carbon 1 (C1) for all residues. C1 is the first carbon of the alkyl chain. The amino acid positions are defined by the center of mass of the backbone nitrogen and C alpha atoms. Black curves denote the distributions for the C1-Proline pair; red curves the distribution for the C1-arginine pair; green curves the C1-tryptophan pair; blue curves the C1-glycine pair; and cyan curves the C1-water pair measured at the water oxygen (Wat(O)). Brow curves denote the distribution for the C1-C18 pair, where C18 is the last carbon atom of the alkyl chain.

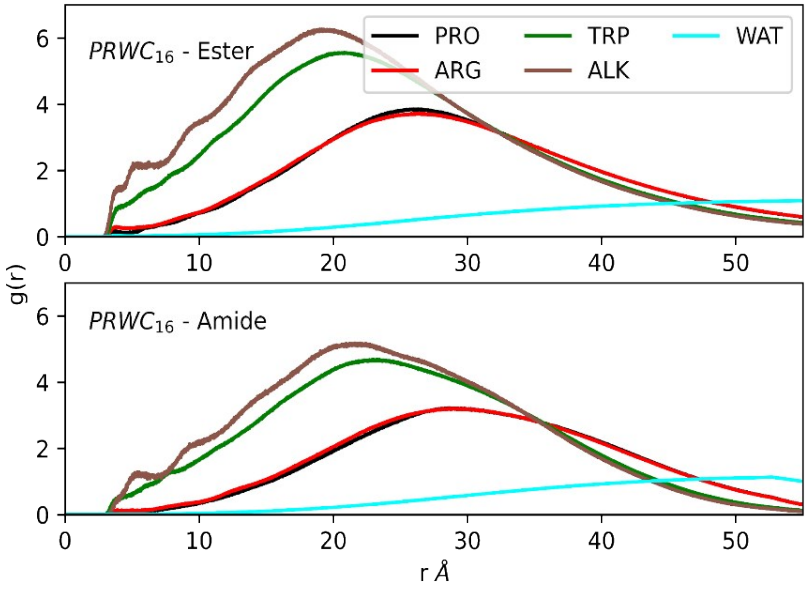


Figure S10. RDFs for systems PRW-O-C_{16} (top) and PRW-C_{16} systems show a layered structural organization of residues with respect to the micellar center, defined as the average position of carbon 1 (C1) for all residues. C1 is the first carbon of the alkyl chain. The amino acid positions are defined by the centre of mass of the backbone nitrogen and C alpha atoms. Black curves show the distributions for the C1-Proline pair; red curves the distribution for the C1-arginine pair; green curves the C1-tryptophan pair and, cyan curves the C1-water pair measured at the water oxygen (Wat(O)). Brow curves denote the distribution for the C1-C16 pair, where C16 is the last carbon atom of the alkyl chain.

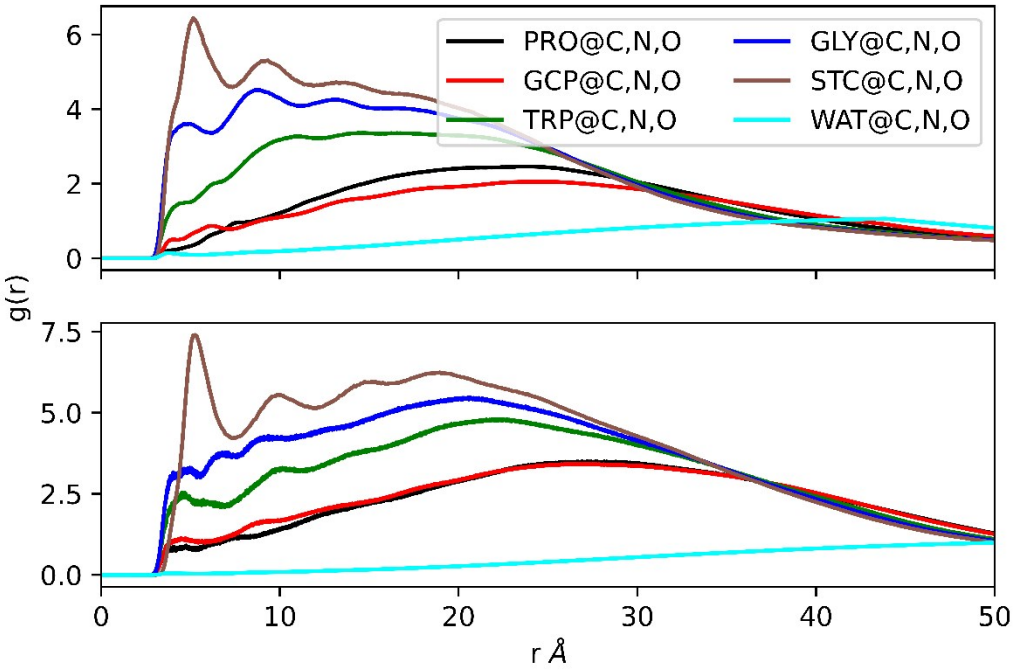


Figure S11. RDFs for systems PK(GCP)WG-$C_{18}H_{37}$ (**3**) (top) and PK(GCP)WG-$(C_{18}H_{37})_2$ (**4**) show a layered structural organization of residues with respect to the micellar center, defined as the average position of carbon 1 (C1) for all residues. C1 is the first carbon of the alkyl chain. The amino acid positions are defined by the centre of mass of the backbone nitrogen and C alpha atoms. Black curves denote the distribution for the C1-Proline pair; red curves the C1-Lys-GCP pair; green curves the C1-tryptophan pair and, cyan curves the C1-water pair measured at the water oxygen (Wat(O)). Brow curves denotes the distribution for the C1-C18 pair, where C18 is the last carbon atom of the alkyl chain.

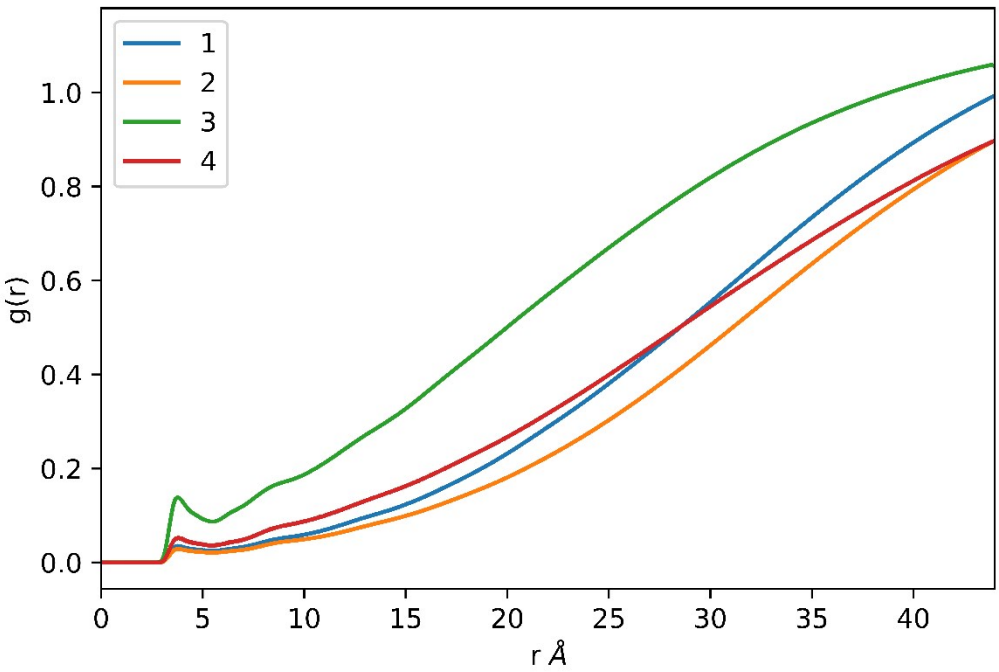


Figure S12. RDFs for systems (**1**) trough (**4**) for the cyan curves the C1-water pair measured at the water oxygen (Wat(O)) where C1 is the first carbon of the alkyl chain. The micellar center for the RDFs is defined as the average position of carbon 1 (C1) for all residues. This figure shows in detail the central part for the water distribution on the micellar systems as given by RDFs. The bulk value for g(r) is 1, therefore, a lower than 1 value for g(r) mean less water content when compared to bulk.

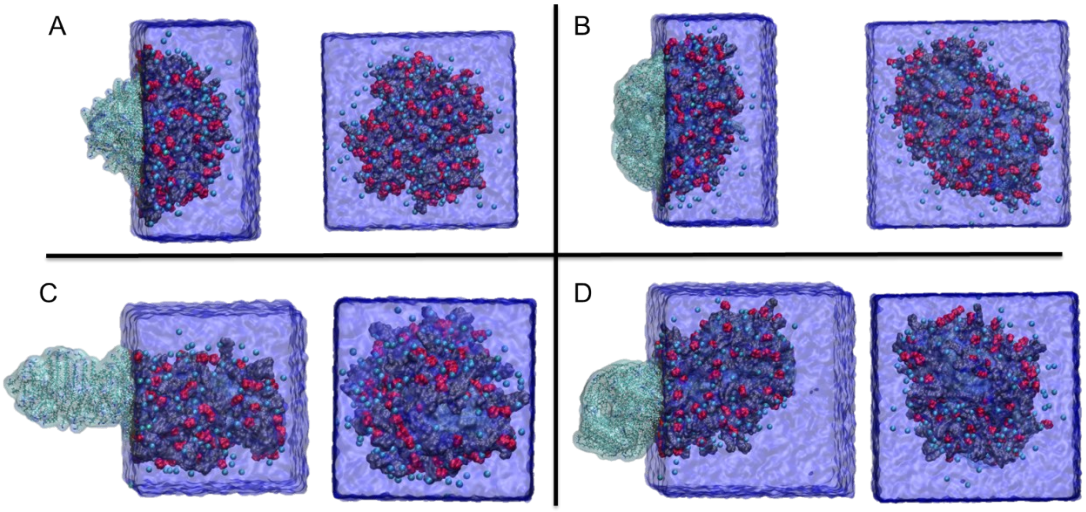


Figure S13. Final configurations after 100 ns of molecular dynamics simulations for systems (**1**) through (**4**) are presented in panels A trough D, respectively. Two views of the simulation box are displayed. Hydrophobic cores are shown using a licorice and surface representation in green. Peptides are shown using a surface representation in grey for amino-acids RWG or K(CGP)WG and red for proline. Counter ions are presented using a VDW representation in green. The water simulation box is displayed using a surface representation in blue. Simulation data is given in the text.

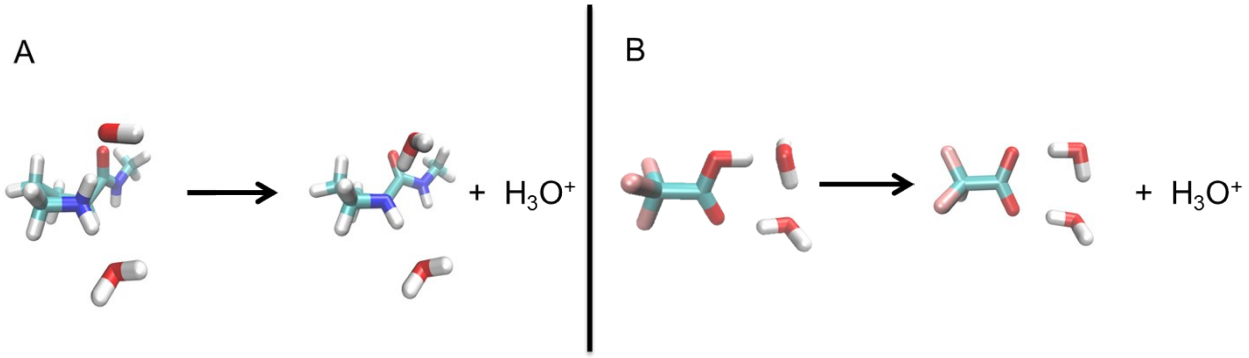


Figure S14. Equilibrium geometries used in the pKa calculations with the micro solvation approach for L-proline methyl ester and TFA in panels A and B. Two water molecules (2W values in Table S3) were explicitly added in H bond configurations with the solute molecules, while solvation energies were computed using the SMD method. Details of the procedure are given in the text.

Table S2. The radius of Gyration (in Angstrom) for the lipopeptide residues, showing both the mean value and their standard deviations. The data collection was performed for the last 90 ns of dynamics for all systems, except for PRWG(C_{18})$_2$, where we have used the last 60 ns. ROGs were computed using the same selection used for computing the RDFs.

Amino acid	Proline	Arginine	Tryptophan	Glycine	Alkyl
PRWGC_{18}					
Mean	37.819	37.678	33.104	31.314	29.750
σ	0.151	0.131	0.163	0.154	0.136
PRWG(C_{18})$_2$					
Mean	42.533	42.046	38.724	36.779	35.094
σ	0.242	0.308	0.206	0.165	0.155
PRW-O-C_{16}					
Mean	36.527	36.723	32.115	-	30.857
σ	0.255	0.376	0.378	-	0.401
PRW-NH-C_{16}					
Mean	35.590	35.431	30.509	-	29.364
σ	0.118	0.135	0.215	-	0.168

Table S3. Computed □G and pKa using a ma-def2-TZVP basis sets and B3LYP functional as implemented in Orca code version 4.2.1 in all calculations ma-def2-TZVP level of theory. PME stands for L-proline methyl ester.

System	Medium	□G (kcal/mol)	pKa
TFA	Water	-3.76	-3.06
	Cyclohexanone	3,93	2,58
	n-Heptane	25,81	18,64
TFA+2H_2O	Water	-0.35	-0,56
	Cyclohexanone	5,43	3,69
	n-Heptane	23,01	16.59
PME	Water	11.54	8.17
	Cyclohexanone	14.61	10.43
	n-Heptane	-15.31	-11.54
PME+2H_2O	Water	14.06	10.02
	Cyclohexanone	9.68	6.80
	n-Heptane	-6.42	-5.02

Table S4. Calculated □G in kcal/mol for the reaction between protonated proline and trifluoroacetic anion, resulting in neutral proline and trifluoroacetic acid.

Reaction	Water	Cyclohexanone	n-Heptane
L-ProH^+ ($2H_2O$) + TFA^-($2H_2O$) → L-Pro($2H_2O$) + TFAH ($2H_2O$)	14.41	4.25	-29.44