Supplementary Information

An Intramolecular Tryptophan-Condensation Approach for Peptide Stapling†

Eunice Y.-L. Hui,*[a] Bhimsen Rout,[b] Yaw Sing Tan,[c] Chandra S. Verma,[c,d,e] Kok-Ping Chan[a] and Charles W. Johannes*[a]

[a]*Institute of Chemical and Engineering Sciences, Agency for Science, Technology and Research, Singapore-138665.*
[b]*Institute of Medical Biology, Agency for Science, Technology and Research, Singapore-138648.*
[c]*Bioinformatics Institute, Agency for Science, Technology and Research, Singapore 138671.*
[d]*Department of Biological Sciences, National University of Singapore, Singapore 117543*
[e]*School of Biological Sciences, Nanyang Technological University, Singapore 637551*

* Corresponding author.

Email: eunicehui@ices.a-star.edu.sg

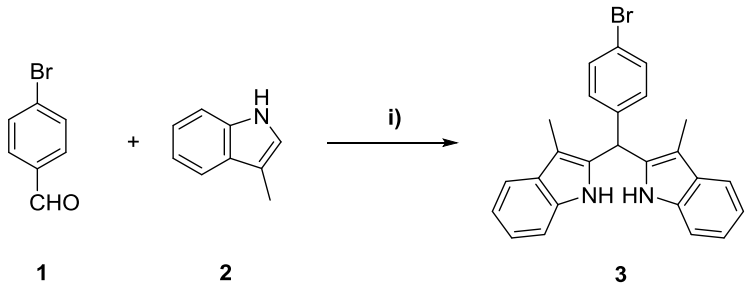
Table of Contents

General Information

Chemicals and anhydrous solvents were obtained from Sigma Aldrich and were used without further purification. Spectroscopic grade solvents were purchased from Sigma Aldrich. Anhydrous solvents were transferred using oven-dried syringes. Thin-layer chromatography (TLC) was carried out on 0.25 mm E. Merck silica gel plates (60F-254). Chromatograms were visualized with cerium ammonium molybdate solution (CAM). E. Merck silica gel (60, particle size 0.040 – 0.063 mm) was used for flash column chromatography. For peptide synthesis, all amino acids and resins were purchased from Advanced Chemtech (Louisville, KY). Fmoc-threonine, serine, glutamic acid and tyrosine were t-butyl protected. Fmoc-tryptophan was used without the Boc protecting group. All peptides were synthesied on PAL-NovaSyn TG resin which was obtained from Novabiochem (San Diego, CA) with a loading capacity of 0.22 mmol g^{-1}. NMR spectra were recorded on Bruker Avance III 400MHz spectrometer in MeOD-d_4, CD_2Cl_2 and DMSO-d_6. Data are reported in the following order: chemical shifts are given (δ); multiplicities are indicated as s (singlet), d (doublet), t (triplet), q (quartet) and m (multiplet). Low-resolution mass spectra (LRMS) were acquired with a Waters 3100 Mass spectrometer. High-resolution mass spectra (HRMS) were recorded on an Agilent ESI-TOF mass spectrometer at 3500 V emitter voltage. Exact m/z values are reported in Daltons. Liquid Chromatography Mass Spectrometry (LCMS) was performed using a Waters 3100 Mass spectrometer and Waters 2545 binary gradient system fitted with a C12 analytical column (Phenomenex® 4.6 x 150mm, 4 µm,) employing a binary gradient of solution A (0.1% Formic acid in H_2O) and solution B (0.1% Formic acid in acetonitrile).The purification of peptides was performed by semi-preparative reverse phase HPLC using C12 semi-preparative column (Phenomenex® 10 x 250mm, 4 µm) employing a binary gradient of solution A (0.1% TFA in H_2O) and solution B (0.1% TFA in acetonitrile). The purity of peptides was ascertained by analytical reverse phase HPLC using analytical column (Phenomenex® 4.6 x 150mm, 4 µm) employing a binary gradient of solution A (0.1% TFA in H_2O) and solution B (0.1% TFA in acetonitrile). The HPLC purified fractions were freeze dried using a Labconco lyophiliser at -84 ^oC and 0.01 mbar vacuum. Emission spectrum measurement was performed on a Perkin Elmer Envision 2104 Multilabel Reader using 96-well plates. Energy minimisations and MD simulations were performed by the PMEMD module of AMBER 14,[7] using the ff14SB[8] force field for the peptide residues and the generalised AMBER force field for the stapled residues. Circular dichroism (CD) measurements were performed using a Jasco J-810 spectrophotometer.

Detailed Methods and Analytical Data

Synthetic scheme for the skatole condensed product



1 **2** **3**

Scheme S1 Synthesis of model substrate **3** for the optimization of tryptophan residue condensation. i) acid.

Experimental procedures

2,2'-((4-bromophenyl)methylene)bis(3-methyl-1*H*-indole) (3). Optimisation of conditions (Table 1).

Method A:
To a mixture of skatole (2 eq.) and 4-bromobenzaldehyde (1 eq.) in CH_2Cl_2 (2 mL) was added acids (methanesulfonic acid, (1*S*)-(+)-Camphor-10-sulfonic acid, p-toluenesulfonic acid or trifluoroacetic acid, 5 eq.) and the resulting mixture was stirred at 40 °C for 6 h. Dichloromethane was added and the organic layer was washed with water (1x). The aqueous layer was neutralised with saturated $NaHCO_3$ and extracted with dichloromethane. Combined organic extracts were washed with brine (1x) and dried over anhydrous Na_2SO_4. The mixture was filtered and solvent was removed under reduced pressure. The crude product was purified using flash column chromatography (petroleum ether: diethyl ether 95:5) to give bis-indole **3** as a white solid (0-72%). ^{1}H NMR (400 MHz, MeOD) δ 7.51 – 7.41 (m, 4H), 7.29 – 7.22 (m, 2H), 7.12 – 6.93 (m, 8H), 6.02 (s, 1H), 2.16 (s, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.31, 135.54, 132.86, 132.31, 130.35, 129.66, 122.08, 121.50, 119.79, 118.73, 111.06, 109.17, 40.57, 29.85, 8.69. HRMS: (ESI TOF) *m/z*: $[M-H]^-$ Calcd for $C_{25}H_{20}BrN_2$ *m/z* 427.0815; Found 427.0818.

Method B:
To a solution of skatole (42 mg, 0.32 mmol, 2 eq.) in concentrated HCl (1 mL) was added 4-bromobenzaldehyde (29 mg, 0.16 mmol, 1 eq.) at room temperature in a seal tube. The reaction mixture was stirred at 100 °C for 3 h. Dichloromethane was added and the organic layer was washed with water (1x). The aqueous layer was neutralised with saturated $NaHCO_3$ and extracted with dichloromethane (1x). Combined organic extracts were washed with brine (1x) and dried over anhydrous Na_2SO_4. The mixture was filtered and solvent was removed under reduced pressure. The crude product was purified using

flash column chromatography (petroleum ether: diethyl ether 95:5) to give bis-indole **3** as a white solid (4%).

Method C:
To a mixture of skatole (42 mg, 0.32 mmol, 2 eq.) and 4-bromobenzaldehyde (29 mg, 0.16 mmol, 1 eq.) in EtOH (0.3 mL) was added sulfuric acid (10 µL) and the resulting mixture was stirred at room temperature for 18 h. Dichloromethane was added and the organic layer was washed with water (1x). The aqueous layer was neutralised with saturated $NaHCO_3$ and extracted with dichloromethane. Combined organic extracts were washed with brine (1x) and dried over anhydrous Na_2SO_4. The mixture was filtered and solvent was removed under reduced pressure to give bis-indole **3** as a white solid (85%).

Procedures and methods regarding the solid phase synthesis of peptides[1]

i. Resin Swelling
The peptides were synthesised manually, using solid phase and Fmoc chemistry at 0.25 mmol scale using PAL-NovaSyn TG resin (0.22 mmol g^{-1}). The dry resin was swelled with NMP for 15 min before use.

ii. Deprotection of *N*-Fmoc protecting groups
N-terminal Fmoc protecting groups were removed by the addition of piperidine (20% v/v) in NMP for 15 min, followed by rinsing the resin with NMP (5 × 5 mL).

iii. Kaiser Test[2]
The Kaiser Test was employed for the determination of the successful coupling or deprotection for most of the residues. A small number of resin beads were taken out and placed in a vial, followed by the addition of two drops of each of the three solutions in the following order: 1) Ninhydrin (5% w/v) in Ethanol, 2) Phenol (80% w/v) in Ethanol, 3) pyridine. The solution was then heated to ca. 120 °C for 2 min. A successful coupling gave no change in the colour of the beads, whereas bright blue beads illustrate a successful deprotection.

iv. Coupling of Amino Acids with DIC and HOAt
Fmoc-protected amino acids (4.0 equiv according to initial loading of the resin), diisopropylcarbodiimide (DIC, 3.0 equiv) and 1-Hydroxy-7-azabenzotriazole (HOAt, 4.0 equiv.) were dissolved in NMP (0.5 M). The mixture was pre-activated for 7 min and added to the resin for 1 h. The reagents were drained away and the resin was washed with NMP (2 × 5 mL).

v. *N*-terminal acetylation
Peptides were acetylated using a mixture of acetic anhydride, DIPEA and DMF (2:1:2, 5mL) for 60 min. The resin was drained, washed with NMP (2 × 5 mL), CH_2Cl_2 (2 × 5 mL) and MeOH (2 × 5 mL) and dried using compressed air.

vi. Solid phase condensation

The resin (0.0625 mmol scale) was swelled in DCE for 15 min in a microwave vial before reaction. 4-bromobenzaldehyde (69 mg, 0.38 mmol, 6 equiv) and (1*S*)-(+)-Camphor-10-sulfonic acid (145 mg, 0.62 mmol, 10 equiv) was added. The beads were then heated at 60 °C for 3 h with agitation via a stream of N_2.

vii. Cleavage and deprotection of PAL-NovaSyn TG resins

After solid phase condensation was complete, the resin was washed with CH_2Cl_2 (2 × 5 mL) and dried using compressed air. Peptides were then simultaneously cleaved and side-chain deprotected with TFA:triisopropylsilane:water (95:2.5:2.5, 4 mL) for 2 h followed by filtration and precipitation with diethyl ether:pentane (1:1 v/v, 20 mL). The precipitate was collected by centrifugation (4000 rpm × 7 min), dried and re-dissolved in a mixture of MeCN and water and lyophilised.

viii. Peptide Purification

Peptides were purified by semi-preparative scale HPLC using a Jupiter Proteo semi-preparative column on an increasing gradient of solvent A and B at a flow rate of 5 mL min^{-1}. Crude peptides were suspended in DMSO at an approximate concentration of 10 mg mL^{-1}. Purification runs injected a maximum of 900 µL of crude peptide solution and were allowed to run for 30 min, with solvent B increasing from 40-70% or 20-50% and the eluent scanned at 220 nm. After purification, the fractions were analysed using an analytical HPLC fitted with Jupiter Proteo analytical column at a flow rate of 1 mL min^{-1} and lyophilised.

Procedure for solution phase condensation

To a solution of peptide **4d** (10 mg, 0.01 mmol, 1 eq.) and 4-bromobenzaldehyde (14 mg, 0.08 mmol, 6 eq.) in CH_2Cl_2 (1 mL) was added (1S)-(+)-Camphor-10-sulfonic acid (60 mg, 0.26 mmol, 20 eq.) and the resulting mixture was stirred at 40 °C for 6 h. The crude product was dried and then suspended in DMSO at an approximate concentration of 10 mg mL^{-1}.Purification was done using semi-preparative scale HPLC.

MS chromatogram for the formation of product 3

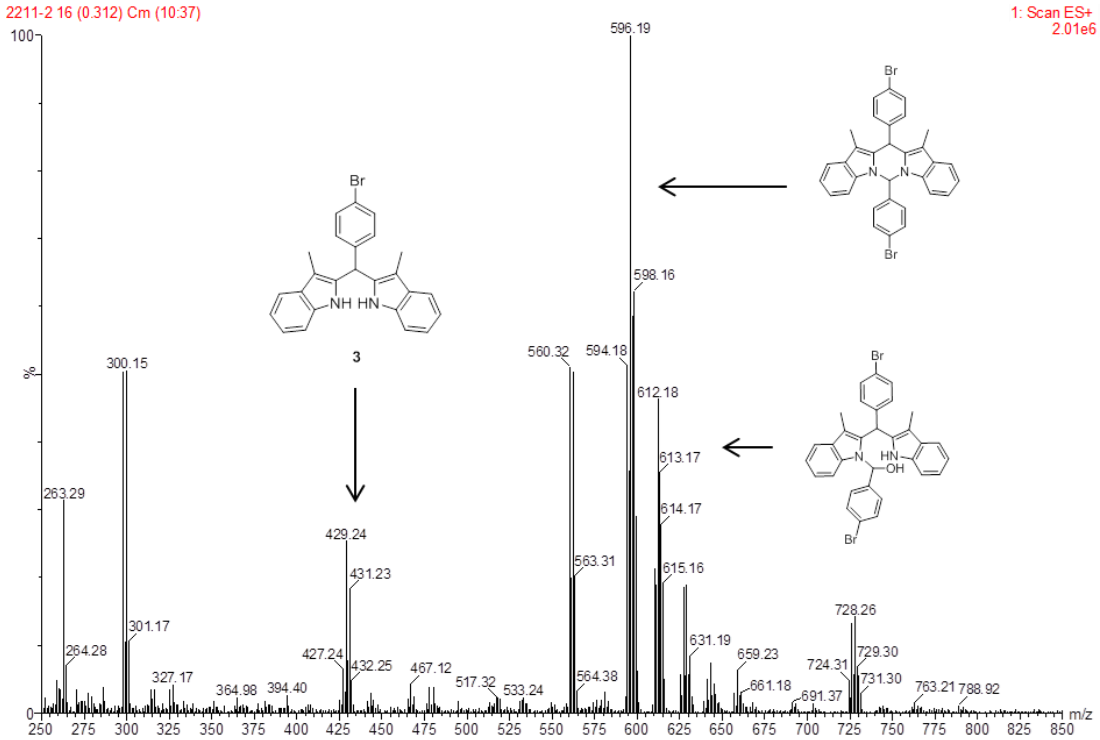


Fig. S1 Section of the low resolution MS chromatogram indicating the formation of **3** together with postulated side products.

Analytical LCMS Chromatogram of Peptide

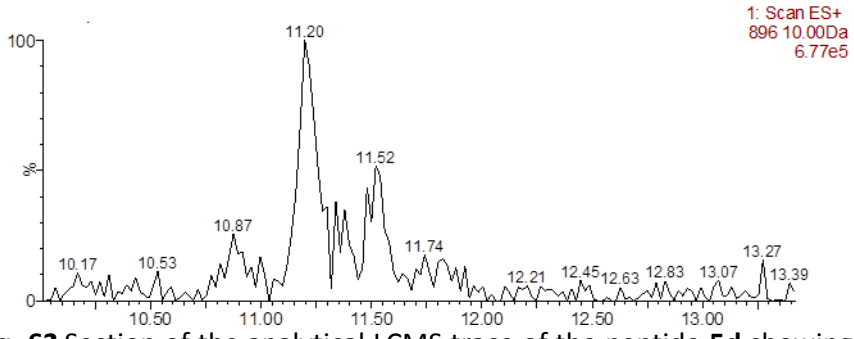


Fig. S2 Section of the analytical LCMS trace of the peptide **5d** showing multiple peaks in the UV trace with mass corresponding to the target peptide.

HPLC purification of peptides

Purification of peptides were performed using semi-preparative HPLC with a detector at 220 nm with 1% TFA in H_2O as solvent A and 0.1% TFA in acetonitrile as solvent B. Peptides **4a-4g** and **5a-5g** were purified using a solvent gradient of 40-70%B (**Table S1**) and the rest of the peptides were purified using a solvent gradient of 20-50% B (**Table S2**). Purity of the fractions was ascertained by analytical HPLC using same gradient of solvents. Fractions of similar purity were combined and lyophilised.

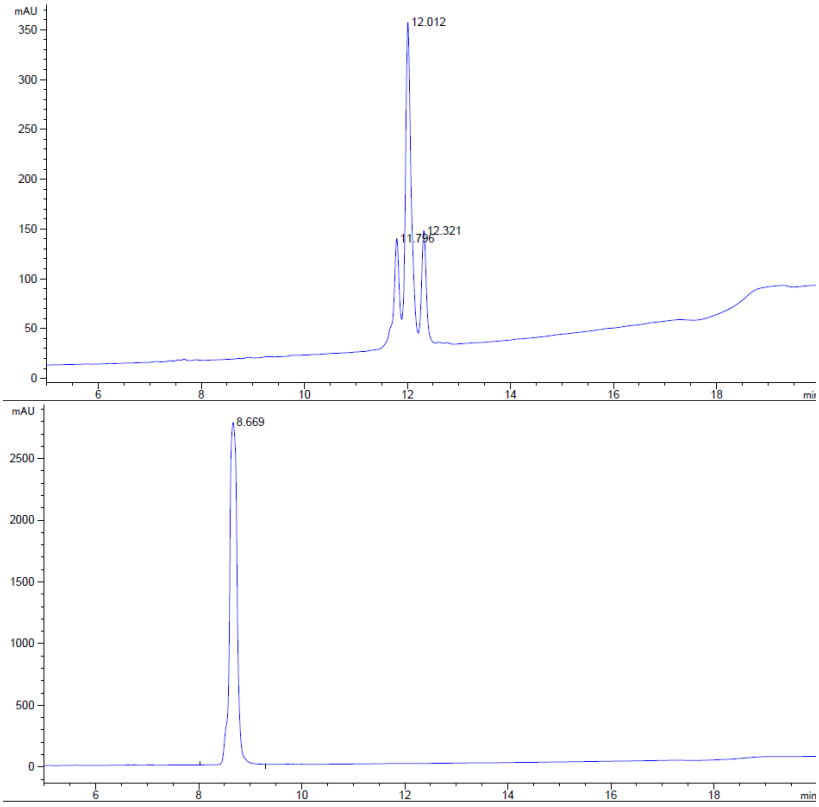


Fig. S3 HPLC chromatograms of purified stapled peptide **5d** (top) and its crude linear precursor **4d** (bottom) monitored at 220 nm

Time (min)	Solvent B (%)	Flow rate (mL/min)
0	40	5
2	40	5
32	70	5
33	90	5
35	90	5
36	40	5
40	40	5

Table S1 Solvent gradient of 40-70%B

Time (min)	Solvent B (%)	Flow rate (mL/min)
0	20	5
2	20	5
32	50	5
33	90	5
35	90	5
36	20	5
40	20	5

Table S2 Solvent gradient of 20-50%B

Peptide	Retention Time	HPLC peak splitting
4a	8.786	s
4b	8.953	s
4c	8.275	s
4d	8.722	s
4e	7.423	s
4f	7.828	s
4g	10.667	s
4h	10.525	s
4i	10.628	s
5a	11.661	s
5b	11.395	s
5c	11.744	d
5d	12.054	d
5e	10.234	s
5f	10.688	s
5g	13.909	d
5h	13.740	d
5i	13.738	d

Table S3 HPLC retention time of peptides and peak splitting pattern. s-singlet, d-doublet.

High Resolution Mass Spectrometry Data for Peptides

Peptide	Sequence	Obs	Exp
4a	Ac-^{1}WWALL^{5}-NH_2	751.3929	751.3902
4b	Ac-^{1}WAWLL^{5}-NH_2	751.3915	751.3902
4c	Ac-^{1}WALWL^{5}-NH_2	729.4063	729.4083
4d	Ac-^{1}WALLW^{5}-NH_2	751.3889	751.3902
4e	Ac-^{1}EYWALLW^{7}-NH_2	1043.5	1043.496
4f	Ac-^{1}YWALLWS^{7}-NH_2	1001.485	1001.486
4g	Ac-^{1}ALWALLW^{7} -NH_2	935.5113	935.5114
4h	Ac-^{1}WNGRW^{5}-NH_2	759.3696	759.3685
4i	Ac-^{1}WRGDW^{5}-NH_2	760.3498	760.3525
5a	Ac-^{1}**WW**ALL^{5}-NH_2	917.3316	917.332
5b	Ac-^{1}**W**A**W**LL^{5}-NH_2	917.3323	917.332
5c	Ac-^{1}**W**AL**W**L^{5}-NH_2	895.3478	895.3501
5d	Ac-^{1}**W**ALL**W**^{5}-NH_2	917.3392	917.332
5e	Ac-^{1}EY**W**ALL**W**^{7}-NH_2	1211.441	1211.438
5f	Ac-^{1}Y**W**ALL**W**S^{7}-NH_2	1147.445	1147.445
5g	Ac-^{1}AL**W**ALL**W**^{7}-NH_2	1081.47	1079.455
5h	Ac-^{1}**W**NGR**W**^{5}-NH_2	927.3087	927.3093
5i	Ac-^{1}**W**RGD**W**^{5}-NH_2	926.2858	926.2943
W denotes the presence of staple involving tryptophan residue.			

Masses are in m/z units and the **Obs**erved and **Exp**ected values quoted

Circular Dichroism

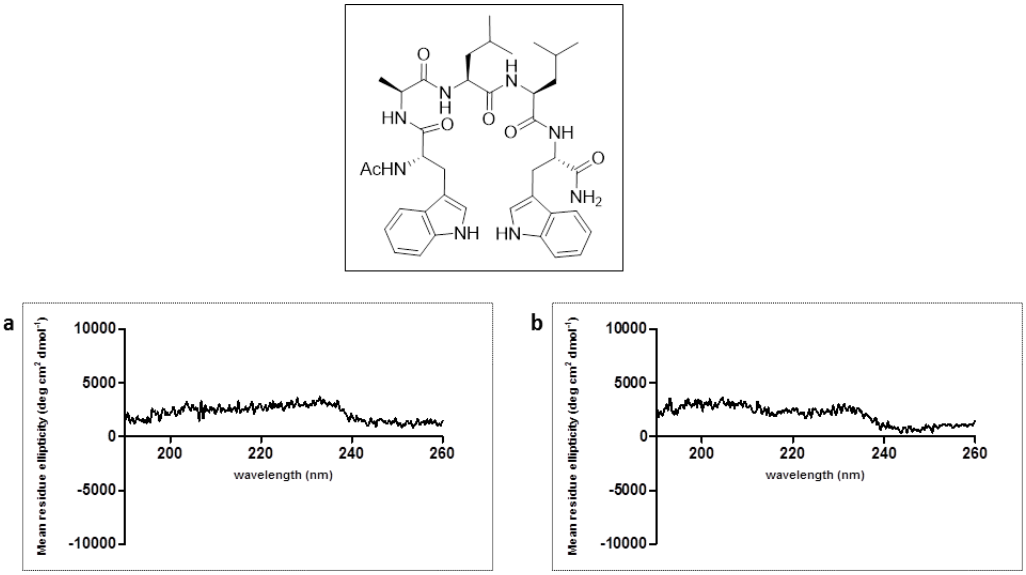


Fig. S4 Circular dichroism spectra of **4d**. **a**, At 140 µM in a buffer of 25 mM Na_2HPO_4 (pH 7). **b**, At 140 µM in a buffer of 25 mM Na2HPO4 (pH 7) (90%) and 10% of TFE.

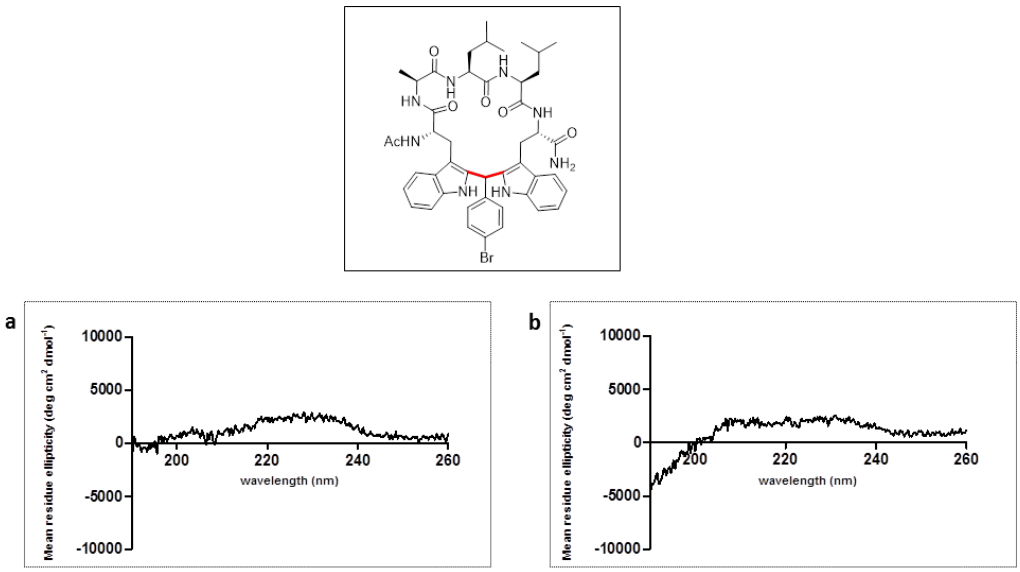


Fig. S5 Circular dichroism spectra of **5d**. a) At 140 µM in a buffer of 25 mM Na_2HPO_4 (pH 7). b) At 140 µM in a buffer of 25 mM Na_2HPO_4 (pH 7) (90%) and 10% of TFE.

8. Enzymatic Degradation

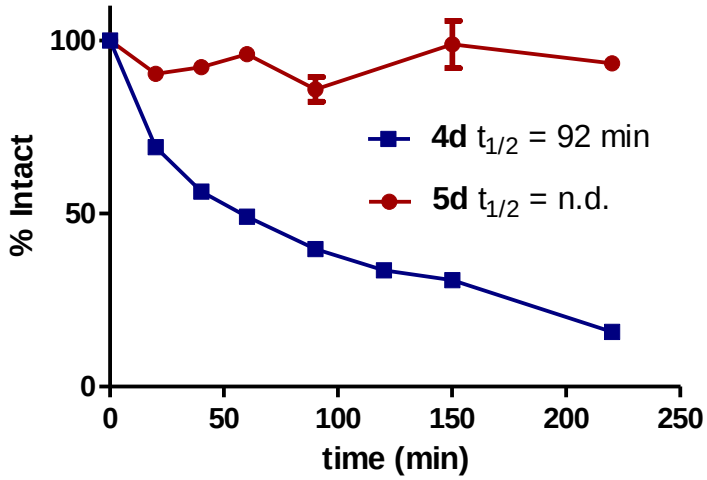


Fig. S6 Proteolytic degradation assay of stapled peptides **5d** and its linear precursors **4d**. With respect to the linear compounds **4d**, α-chymotrypsin cleavage products (from the hydrolysis on the *C*-terminal side of tryptophan)[3] were observed. Significant different behaviour was observed for the corresponding stapled peptides **5d** which remained unaltered and only traces of the corresponding C-terminal hydrolysis of tryptophan were detected.

9. Molecular dynamics simulations

The initial structure of **4d** was taken from the Protein Data Bank (PDB) structure 3LNZ[4] (MDM2 in complex with the N8A mutant of the PMI peptide) for molecular dynamics (MD) simulations. Peptide **4d** was generated by truncating both ends of the PMI-N8A peptide such that the core WALLS sequence is retained, mutating the C-terminal serine to tryptophan, and capping the N- and C-termini of the truncated sequence with acetyl and N-methyl groups respectively. Using PyMOL,[5] the bisindole linker was modelled onto **4d** to generate the stapled peptide **5d**. Each peptide was then solvated with TIP3P[6] water molecules in a periodic truncated octahedron box, such that the solvent box walls are at least 15 Å away from the peptide.

For each of the peptides, three independent MD simulations using different initial atomic velocities were carried out. Energy minimisations and MD simulations were performed by the PMEMD module of AMBER 14,[7] using the ff14SB[8] force field for the peptide residues and the generalised AMBER force field for the stapled residues. Atomic charges for the stapled residues were derived using the R.E.D. Server,[9] by fitting restrained electrostatic potential (RESP) charges[10] to a molecular electrostatic potential (MEP) computed by the Gaussian 09 program[11] at the HF/6-31G* level of theory. The SHAKE[12] algorithm was applied to constrain all bonds to hydrogen atoms, thus allowing for a time step of 2 fs. Non-bonded interactions were truncated at 9 Å while electrostat-

ic interactions were accounted for by the particle mesh Ewald method.[13] Each system underwent 500 cycles of steepest-descent energy minimisation, followed by 500 cycles of conjugate-gradient energy minimisation. The system was then gradually heated to 300 K over 50 ps at constant volume before equilibration at a constant pressure of 1 atm for another 50 ps. Harmonic positional restraints with a force constant of 2.0 kcal mol^{-1} $Å^{-2}$ were placed on the non-hydrogen atoms of the peptide during these energy minimisation and equilibration steps. Unrestrained further equilibration (2 ns) and production (200 ns) runs were subsequently carried out at 300 K and 1 atm, using a Langevin thermostat[14] to maintain the temperature and a Berendsen barostat[15] to maintain the pressure.

Backbone atoms of the peptide residues were clustered using the MMTSB toolset.[16] The ART-2 algorithm[17,18] was used for root-mean-square-deviation-based clustering. Cut-off radii of 2.2 Å and 2.0 Å were empirically chosen to produce well-separated clusters for **4d** and **5d** respectively.

^{1}H and ^{13}C NMR Spectra

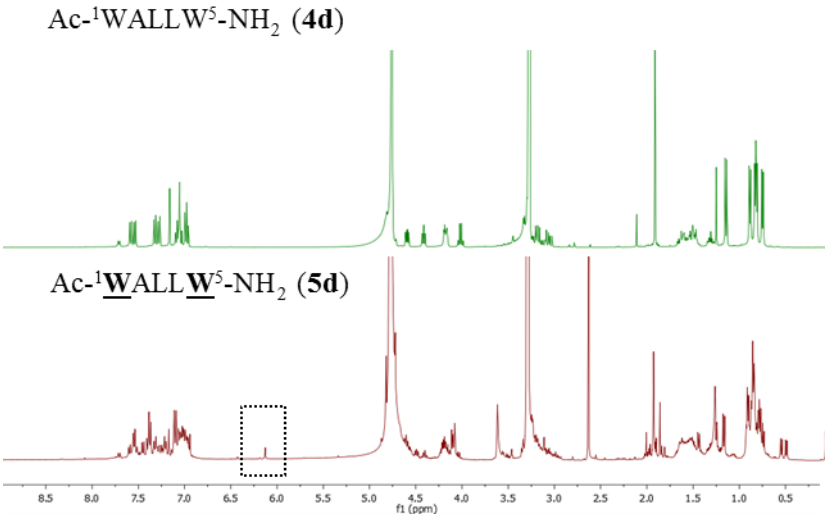


Fig. S7 ^{1}H NMR spectra of peptide **5d** (bottom) and its linear precursor **4d** (top) with diagnostic alkane C-H proton highlighted. **W** denotes the presence of staple involving tryptophan residue.

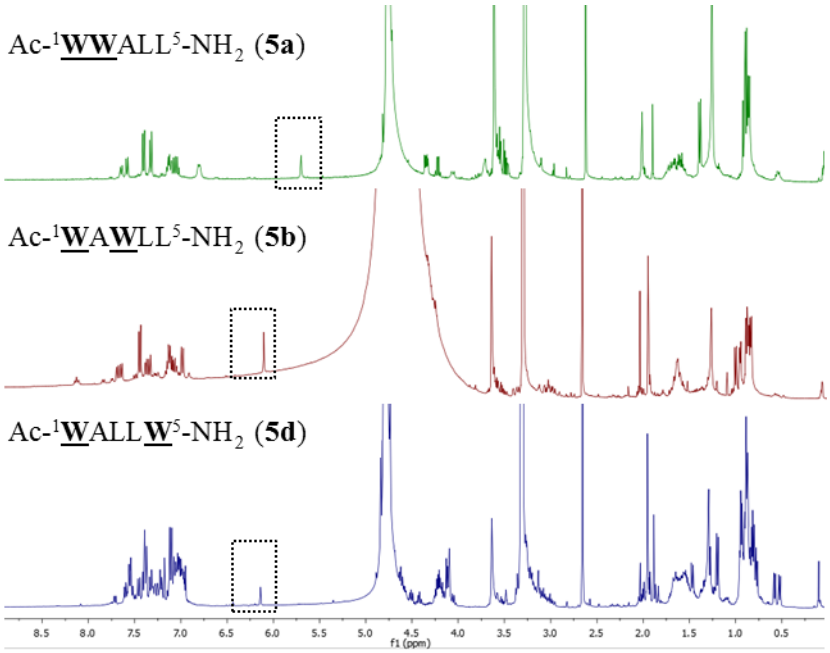


Fig. S8 ^{1}H NMR spectra of peptide **5a** (top), **5b** (middle) and **5d** (bottom) with diagnostic alkane C-H proton highlighted. **W** denotes the presence of staple involving tryptophan residue.

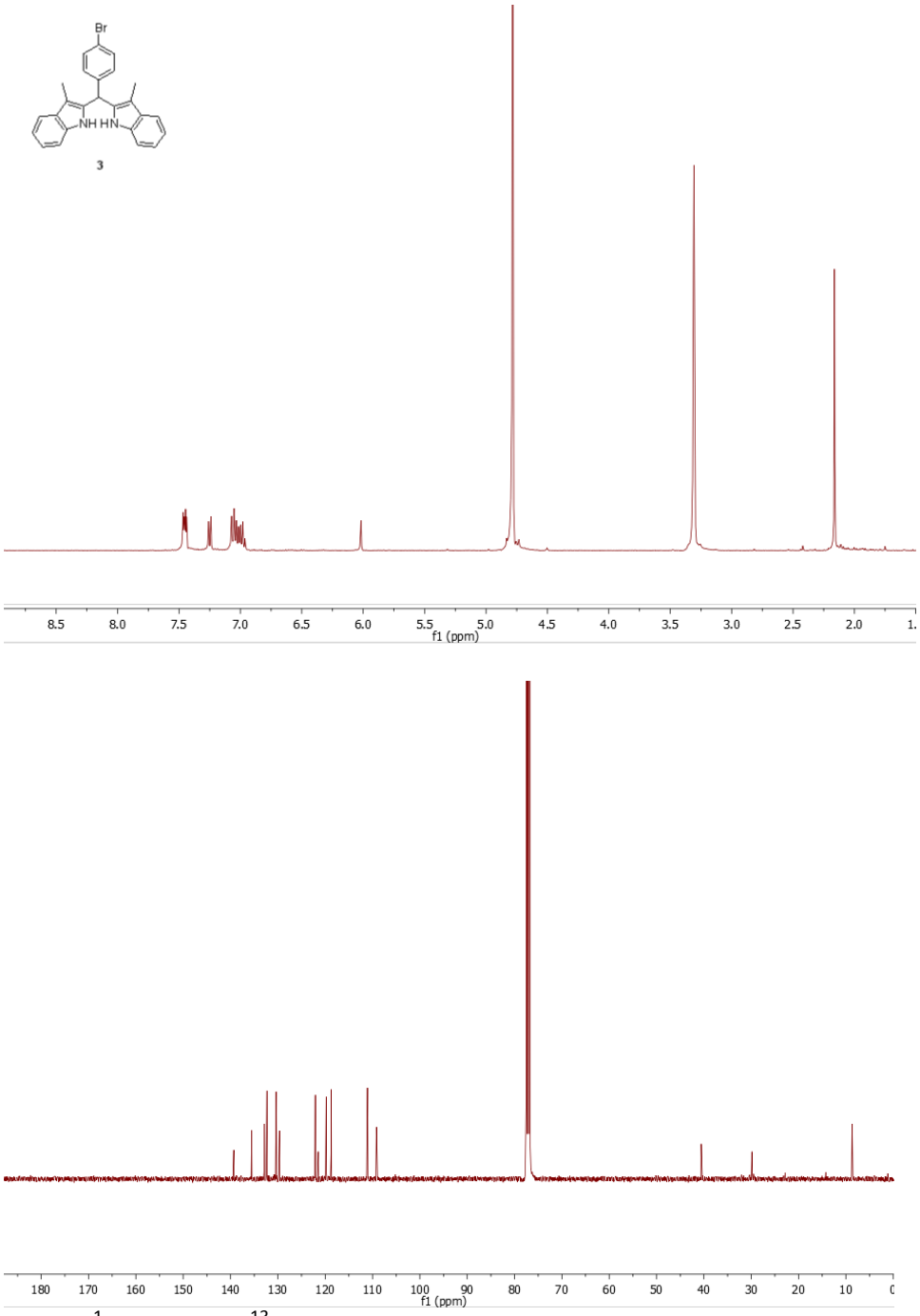


Fig. S9 ^{1}H NMR and ^{13}C NMR spectra of cyclized product **3**. The peak at 6ppm in ^{1}H NMR represents proton on tertiary carbon of skatole.

References:

1. W. C. Chan and P. D. White, *Fmoc Solid Phase Peptide Synthesis*, Oxford University Press, 2000.
2. E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Anal. Biochem.*, 1970, **34**, 595-598.
3. L. Zhou, S.M. Budge, A.E. Ghaly, M.S. Brooks and D. Dave, *Am. J. Biochem. & Biotech.*, 2011, 7, 104-123
4. C. Li, M. Pazgier, C. Li, W. Yuan, M. Liu, G. Wei, W.-Y. Lu and W. Lu, *J. Mol. Biol.*, 2010, **398**, 200-213.
5. W. L. DeLano, DeLano Scientific, San Carlos, CA, USA, 2002.
6. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
7. D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, University of California, San Francisco, 2014.
8. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696-3713.
9. E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak and F.-Y. Dupradeau, *Nucleic Acids Res.*, 2011, **39**, W511-W517.
10. W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *J. Am. Chem. Soc.*, 1993, **115**, 9620-9631.
11. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
12. J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
13. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
14. J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J. Chem. Phys.*, 2001, **114**, 2090-2098.

15. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
16. M. Feig, J. Karanicolas and C. L. Brooks, *J. Mol. Graphics Model.*, 2004, **22**, 377-395.
17. G. A. Carpenter and S. Grossberg, *Appl. Opt.*, 1987, **26**, 4919-4930.
18. M. E. Karpen, D. J. Tobias and C. L. Brooks, *Biochemistry*, 1993, **32**, 412-420.