Electronic Supplementary Material (ESI) for Chemical Science.

Supporting Information

Nature-inspired dimerization as a strategy to modulate neuropeptide pharmacology exemplified with vasopressin and oxytocin

Zoltan Dekan[‡], Thomas Kremsmayr[‡], Peter Keov, Mathilde Godin, Ngari Teakle, Leopold Dürrauer, Huang Xiang, Dalia Gharib, Christian Bergmayr, Roland Hellinger, Marina Gay, Marta Vilaseca, Dennis Kurzbach, Fernando Albericio, Paul F. Alewood, Christian W. Gruber* and Markus Muttenthaler*

Table of Contents

MATERIALS

Reagents and kits

All reagents obtained commercially were used without further purification. The protected Fmoc-amino acid derivatives, Rink amide resin (0.52 mmol/g) and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Iris Biotech GmbH (Marktredwitz, Germany). 2-Chloro-trityl chloride resin (1.5 mmol/g) was obtained from Chem-Impex International (USA). Dichloromethane (DCM), N,N-dimethylformamide (DMF), N,N-diisopropylethylamine (DIEA) and trifluoroacetic acid (TFA) were peptide synthesis grade and supplied by Auspep (Melbourne, Australia). Triisopropylsilane (TIPS), sodium nitrite ($NaNO_2$), hydrazine monohydrate (N_2H_4 64-65 %), sodium 2-mercaptoethanesulfonate (MesNa) and α-cyano-4-hydroxycinnamic acid matrix were supplied by Merck/Sigma Aldrich (Darmstadt, Germany). S-trityl-3-mercaptopropionic acid S-acetamidomethyl-3-mercaptopropionic acid, Boc-Cys(Npys)-OH and dVDAVP ((Deamino-Cys^1,Val^4,D-Arg^8)-Vasopressin) as control peptide were purchased from Bachem (Bubendorf, Switzerland). Acetonitrile (ACN) for HPLC analysis and purification was obtained from VWR International (Darmstadt, Germany). Dulbecco's Modified Eagle's Medium (DMEM) and penicillin/streptomycin were sourced from Thermo Fisher Scientific (Scoresby, Australia). Foetal Bovine Serum (FBS) was sourced from GE Life Sciences (Parramatta, Australia). The IP-1 assay kit was purchased from CisBio (Codolet, France), the FLIPR Calcium-4 Assay Kit from Molecular Devices (Sunnyvale, CA, USA), FuGENE HD Transfection reagent from Promega Corporation (Madison, WI, USA), and HEK293 and COS-1 cells from American Type Culture Collection (ATCC, Manassas, VA, USA). All other reagents and solvents were obtained from Sigma Aldrich (Sydney, Australia) in the highest available purity.

SYNTHETIC METHODS

Analytical reversed-phase high-performance liquid chromatography (RP-HPLC)
Standard analysis: Using a C_{18} column (Kromasil 300-5-C_{18} 4.6 x 50 mm) on a Dionex Ultimate 3000 system with UV detection at 214 and 280 nm. Runs were performed at a flow rate of 1 mL/min using linear gradients from 5-65 %B in 30 min. Solvent A: ddH_2O + 0.1% TFA. Solvent B: ACN + 0.08% TFA.

Mass spectrometry (MS)
Standard analysis: Performed on a Waters 3100 single quadrupole mass spectrometer with electrospray ionization (ESI) in positive ion mode over a range of 400-2000 Da. Samples (5–15 μL) were injected at a flow rate of 1 mL/min (50% ACN/0.05% TFA in water).

HR-MS analysis: Performed on a Bruker maXis UHR-TOF (Qq-TOF) with ESI in positive ion mode. Samples were introduced *via* direct infusion in ACN/MeOH + 1% H_2O and a flow rate of 3 μL/min.

Preparative RP-HPLC
Peptides were purified using a C_{18} column (Kromasil 300-10-C_{18} 250 x 21.2 mm) on a Waters AutoPurification HPLC system with UV detection at 214 nm. Purifications were performed at 20 mL/min using ddH_2O + 0.1% TFA as solvent A, ACN + 0.08% TFA as solvent B and a linear gradient either from 5-45 %B in 40 min or from 5-55 %B in 50 min.

Peptide concentration quantification

Peptide concentrations were determined based on peak area detected at 214 nm by analytical RP-HPLC (Phenomenex - Prodigy 100 Å, 5 µm, 4.6 x 50 mm, 2 mL/min, gradient 0–90% B in 10 min) against two peptide standards (OT, VP) with known peptide content established by amino acid analysis. Using the Beer-Lambert Law, the peptide concentrations were calculated on the basis of absorbance of standards and samples using calculated extinction coefficients.

Peptide synthesis

Peptide precursors were synthesised on a Symphony automated peptide synthesizer (Protein Technology, Tucson, USA) using standard Fmoc protocols.[1] Amino acid side chains were protected as Asn(Trt), Arg(Pbf), Cys(Acm/Trt), Gln(Trt), and Tyr(tBu). Fmoc-Rink amide resin was used for peptides with a C-terminal amide and 2-chloro-trityl resin for the hydrazide precursors of N- to C-cyclic compounds (0.2 mmol scale). For the latter, the 2-chloro-trityl resin was manually loaded with hydrazine using a 5% solution of hydrazine monohydrate in DMF two times for 30 min and unreacted functional groups on the resin were then capped by treatment with 5% MeOH/DMF for 10 min.[2] The first amino acid (Gly, 5 eq.) was directly coupled onto the hydrazine loaded resin using HBTU/DIEA (5 eq/10 eq) activation in DMF (final loading: ~0,5 mmol/g; determined *via* dibenzofulvene–piperidine adduct quantification at 301 nm[3]). Linear peptide chains were further assembled as stated above to give an N-terminal Cys for intramolecular ligation. Fmoc deprotection was performed with 30% piperidine/DMF for 1 min followed by a 2 min repeat. DMF-washes were performed 10 times after each coupling as well as after each deprotection step. After chain assembly and final Fmoc deprotection, the peptide resins were washed with 50% MeOH in DCM and dried under N_2 gas. Peptide cleavage was performed at 25 ºC in TFA:H_2O:TIPS (90:5:5) for 1.5 h. TFA was evaporated using N_2, and cold diethyl ether was added to precipitate the peptide. The precipitate was filtered, washed a second time with cold ether, dissolved in 50% ACN / 0.1% TFA in water, and lyophilised.

The dVDAVP precursor was manually assembled using Boc chemistry with *in-situ* neutralisation containing an a Cys(Npys) at position 6, since no N-terminal dCys(Npys) building block was available.[4] Boc deprotections were accomplished using neat TFA (2 x 1 min). Couplings were performed in DMF using 4 eq. of Boc-amino acid/HBTU/DIEA (8:8:11) relative to resin loading for 10 min. Amino acid side-chains were protected as Asn(Xan), Arg(Tos), Cys(Acm/Npys), Gln(Xan), and Tyr(2-BrZ). Cleavage from the resin and removal of side chain protecting groups was achieved by treatment with 90% HF/10% *p*-cresol at 0°C for 1 h. Following evaporation of HF, products were precipitated and washed with cold Et_2O and lyophilised from 50% ACN/0.1% TFA/H_2O.

Monomeric peptide synthesis

Folding of the purified monomeric reduced peptides was carried out in 0.1 M NH_4HCO_3 buffer, pH 8.2, at 25ºC for 24 h at a peptide concentration of 100 µM. Oxidative folding was monitored by analytical RP-HPLC and ESI-MS, and the folded peptides were purified using preparative RP-HPLC. The oxidised peptides were of >95% purity, as determined by analytical RP-HPLC.

Parallel and antiparallel dimer synthesis

Formation of first interchain disulfide bond – parallel homodimers: crude C(Acm)/C(SH) peptide was dissolved (10 mg/mL) in 6M GdnHCl/0.2M NH_4HCO_3 (pH 8.2) and was stirred with exposure to air for three days. The reaction was monitored by analytical HPLC and ESI-MS, and the product was isolated using preparative RP-HPLC.

Formation of first interchain disulfide bond – antiparallel dimers:[5] Crude C(Acm)/C(SH) and C(Npys)/C(Acm) peptides were combined in an equimolar ratio in helium-purged 6 M GdnHCl/0.2 M NH_4OAc (pH 4.5)

(10 mg/mL) and stirred under argon for 16 h. The reaction was monitored by analytical HPLC and ESI-MS, and the product was purified using preparative HPLC.

Formation of second interchain disulfide bond: The bis-Cys(Acm) containing peptide dimers prepared above were dissolved (1 mg/mL) in an aqueous solution of 20 mM HCl/80% MeOH. A solution of I_2 in MeOH (10 eq.) was added and stirred for 1h. The reaction was monitored by analytical HPLC and ESI-MS, and the product was purified using preparative HPLC.

Cyclic dimer synthesis

Cyclization and formation of first disulfide bond[6]

Crude linear peptide hydrazides (2 mM) were dissolved in a phosphate buffer at pH 3 (0.2 M NaH_2PO_4, 6 M guanidine hydrochloride) and activated for 15 min at -20°C (NaCl/ice bath) using $NaNO_2$ (10 eq., 0.5 M aqueous solution). MesNa (40 eq) was dissolved in a phosphate buffer at pH 7.5 (0.2 M NaH_2PO_4, 6 M guanidine hydrochloride, pH adjusted with aqueous NaOH) and the solution was added to the activated peptide (final peptide concentration 200 µM). The final pH was adjusted to 7.3-7.5 and the mixture was stirred at 25°C until complete cyclization and consecutive one-pot folding was indicated by means of HPLC and MS analysis (typically overnight). The cyclic Acm-peptides with one disulfide bond were isolated using preparative HPLC.

Formation of second disulfide bond

The cyclic bis-Cys(Acm) containing peptide dimers with one disulfide bond prepared above were dissolved in an aqueous solution of 20 mM HCl/80% MeOH (0.5 mg/mL). A solution of I_2 in MeOH (10 eq.) was added and stirred for 30 min. The reaction was monitored by analytical HPLC and ESI-MS. Upon completion the reaction was quenched by addition of ascorbic acid and the product was purified using preparative HPLC.

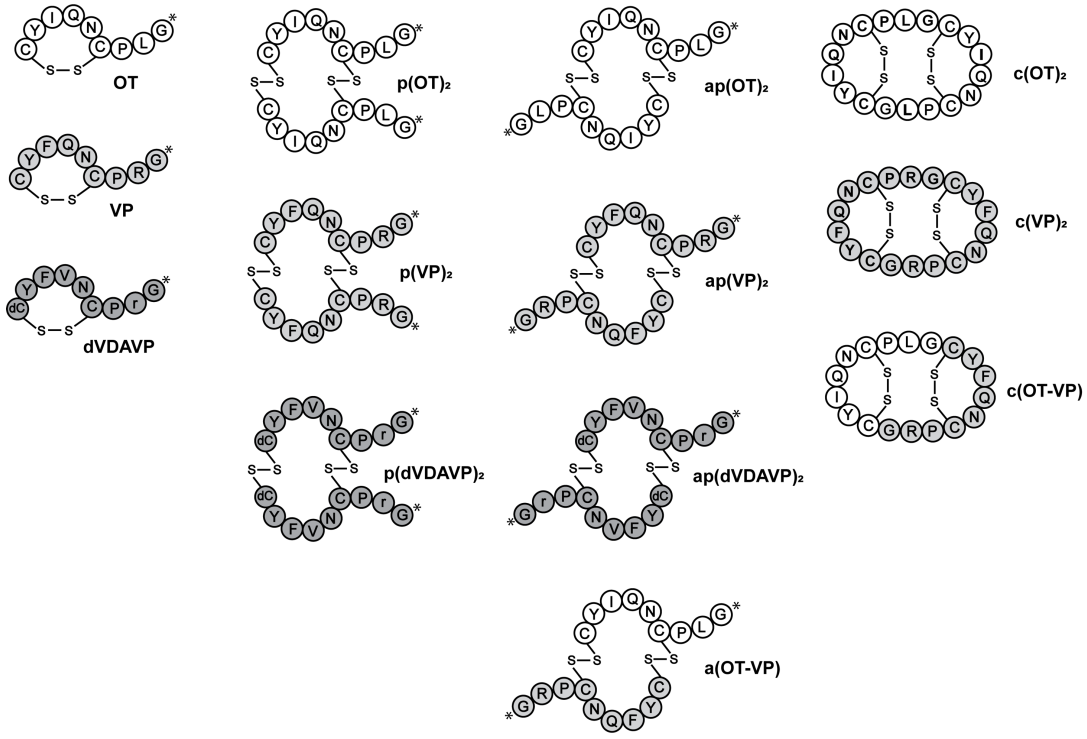


Figure S1. Schematic structures of vasopressin, oxytocin, dVDAVP and their dimeric constructs. Individual amino acids are represented as circles (* indicates C-terminal amide, dC, deaminocysteine; p, parallel; ap, antiparallel, c, cyclic).

PHARMACOLOGY

Cloning, cell culture, transfection, and membrane preparation for functional assays

Human (h) OTR, $V_{1a}R$, $V_{1b}R$, and V_2R cDNA sequences were inserted into the pEGFP-N1 plasmids (Clontech, Mountain View, California, USA) using SacI/HindIII and BamHI restriction sites with or without an additional stop-codon before the GFP coding region (to yield a C-terminal GFP fusion protein). Stably transfected HEK293 cell lines[7] or transient expression of receptors in HEK293 or COS-1 cells were established using the jetPrime transfection protocol (Polyplus transfection), Lipofectamine 2000 (Invitrogen) or FuGENE HD (Roche) transfection reagents, as per the respective manufacturers' protocols. Cells expressing the receptors were either used for functional assays or harvested, and membranes were prepared as described previously.[7]

Radioligand displacement assays

To measure the affinity of the dVDAVP analogues to the four human OT/VP receptors, we used radioligand displacement assays, as previously described.[8] Briefly, isolated membranes from receptor-expressing cells (25 µg/reaction) were incubated with radioactive agonists [^{3}H]OT or [^{3}H]VP (0.25 nM–2 nM, according to the K_D of the ligand to its respective receptor) (PerkinElmer, Waltham, Massachusetts, USA), respectively, and several concentrations of competing peptide. The reaction was stopped by filtration over glass fibre filters using a cell harvester. Non-specific binding was determined in the presence of 1 µM VP or OT, respectively. Specific binding was calculated as the difference of total and non-specific binding. Concentration-dependent affinity data are presented as affinity constants (K_i) calculated following Cheng-Prusoff (Table S1). Binding data are presented as bar-charts and were normalised to the total bound radioligand in the absence of competing ligand. Each data point was measured in quadruplicate (Figure S2).

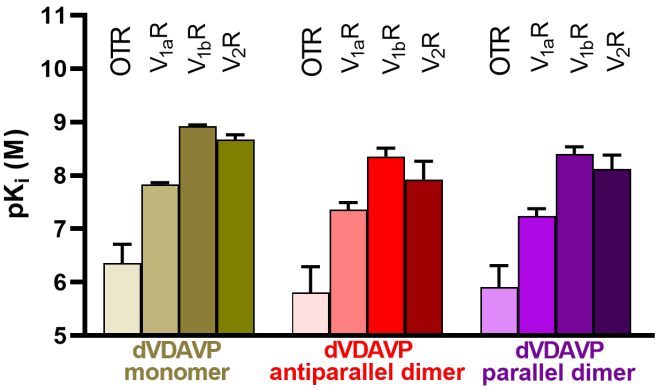


Figure S2. Binding affinities of the dVDAVP derivatives at the four VP/OT receptors.

Table S1. Binding affinities (K_i) of ligands using radioligand displacement assays at $V_{1a}R$, $V_{1b}R$, V_2R and OTR.

	hV$_{1a}$R	hV$_{1b}$R	hV$_2$R	hOTR
Ligand	**K_i (nM)** ± SEM	**K_i (nM)** ± SEM	**K_i (nM)** ± SEM	**K_i (nM)** ± SEM
dVDAVP	16.5 ± 1.2#	1.4 ± 0.2	1.8 ± 0.5#	224 ± 96%
ap(dVDAVP)$_2$	53.4 ± 35.5#	4.4 ± 1.2	8.6 ± 8.2#	1093 ± 351
p(dVDAVP)$_2$	76.0 ± 41.7#	3.3 ± 0.4	4.2 ± 1.1#	744 ± 285

N > 3 experiments; # N=2 K_i (nM) ± SD

Measurement of cAMP accumulation

cAMP accumulation at V_2R was measured as described previously[9] using kits from both PerkinElmer (Waltham, Massachusetts, USA) for the OT/VP and dVDAVP dimers and CisBio (France) for the cyclic dimers according to manufacturers' guidelines. Briefly, HEK293 cells, 38–48 h post-transfection, were stimulated for 20 and 30 min, respectively, by a range of peptide concentrations. The data were normalised to % of maximum VP response.

Measurement of inositol phosphate one (IP-1) accumulation

HEK293 cells were cultured in DMEM supplemented with 10% FBS and penicillin/streptomycin (100 U mL^{-1}) and plated in 6-well plates prior to transfection and transfected as described above. 4–6 h following transfection, cells were dissociated by trituration and plated into white opaque 384-well plates (CellStar®, Greiner, Kremsmünster, Austria) at ~10,000 cells per well. Cells were cultured for 38–48 h in a humidified incubator at 37°C and an atmosphere of 5% CO_2, and then assayed for second messenger production. Verification of receptor expression was assessed *via* microscopic examination of membrane localised GFP fluorescence. Quantitative measurements of receptor-mediated IP-1 were performed by competitive immunoassay utilising the IP-One assay kit (CisBio, France).[10] HEK293 cells were prepared as described above and then assayed as per the manufacturer's protocol. Briefly, at the time of analysis, the culture media was removed from cells, and then replaced with IP-1 stimulation buffer (HEPES 10 mM, $CaCl_2$ 1 mM, $MgCl_2$ 0.5 mM, KCl 4.2 mM, NaCl 146 mM, glucose 5.5 mM, LiCl 50 mM, pH 7.4). Cells were allowed to equilibrate in the stimulation buffer at 37°C for 15 min, followed by addition of peptide ligands (to final concentrations of 0.1 pM – 10 μM) for subsequent stimulation for 1 h at 37°C. The stimulation was terminated by lysis and the simultaneous addition of homogenous time-resolved fluorescence resonance energy transfer reagents. The lysates were incubated for a minimum of 1 h at 25°C. Fluorescence emission measurements at 620 nm and 665 nm were performed using a FlexStation3 plate reader (Molecular Devices, Sunnyvale, CA) or Spark Multimode plate reader (Tecan, Männedorf, Switzerland) at an excitation wavelength of 340 nm. Results were analysed as a ratio of fluorescence intensities of 665 nm to 620 nm.

FLIPR assay measuring intracellular Ca^{2+} responses

COS-1 cells were cultured and transiently transfected with OTR/V_{1a}R/V_{1b}R DNA ratio 1:6 following the manufacturer's protocol (FuGENE HD, Roche). 24 h post-transfection, cells were seeded at a density of 18,000 cells/well in black-walled imaging plates (Corning, Lowell, MA, USA) and maintained for another 24 h at 37°C in a 5% humidified CO_2 incubator. An assay measuring Ca^{2+} responses was performed 48 h post-transfection. Briefly, at the time of analysis the culture media was removed from the cells. The cells were loaded with the Calcium-4 No-wash dye by diluting the lyophilised dye in physiological salt solution (PSS: 140 mM NaCl, 11.5 mM glucose, 5.9 mM KCl, 1.4 mM $MgCl_2$, 1.2 mM NaH_2PO_4, 5 mM $NaHCO_3$, 1.8 mM $CaCl_2$, 10 mM HEPES, pH 7.4), and incubated for 30 min at 37°C in a 5% humidified CO_2 incubator. Intracellular Ca^{2+} responses were measured in response to increasing concentrations of agonist peptides (10 pM – 100 mM), in a Fluorometric Imaging Plate Reader (FLIPR; Molecular Devices, Sunnyvale, CA) using a cooled CCD camera with excitation at 470–495 nM and emission at 515–575 nM. Camera gain and intensity were adjusted for each plate to yield a minimum of 1000 arbitrary fluorescence units (AFU) baseline fluorescence. Prior to addition of the agonist, 10 baseline fluorescence readings were taken, followed by fluorescent readings every second for 300 s. Concentration-response curves were established by plotting Percentage of Control after initial calculation of Delta F/F0 values, where F0 is the base-line level of fluorescence and Delta F is the change in fluorescence from the baseline level, against agonist concentration using Prism.

Statistics and data analysis

All data resulted from at least three independent experiments (N ≥ 3) and are presented as the mean ± standard error of the mean (SEM), except where otherwise stated. Curve fitting was performed by non-linear regression in GraphPad Prism (GraphPad Software Inc., La Jolla, California, USA) using a three-parameter modified Hill equation with a slope set to unity: $Y=Bottom+(Top-Bottom)/(1+10^{LogEC50-X})$, where X corresponds to the log of concentration; values for Top, Bottom and LogEC50 were initially set to 1 and not constrained. Statistical analysis, where applicable, was performed using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test or the Newman-Keuls multiple comparison test, or by two-way ANOVA with Bonferroni correction.

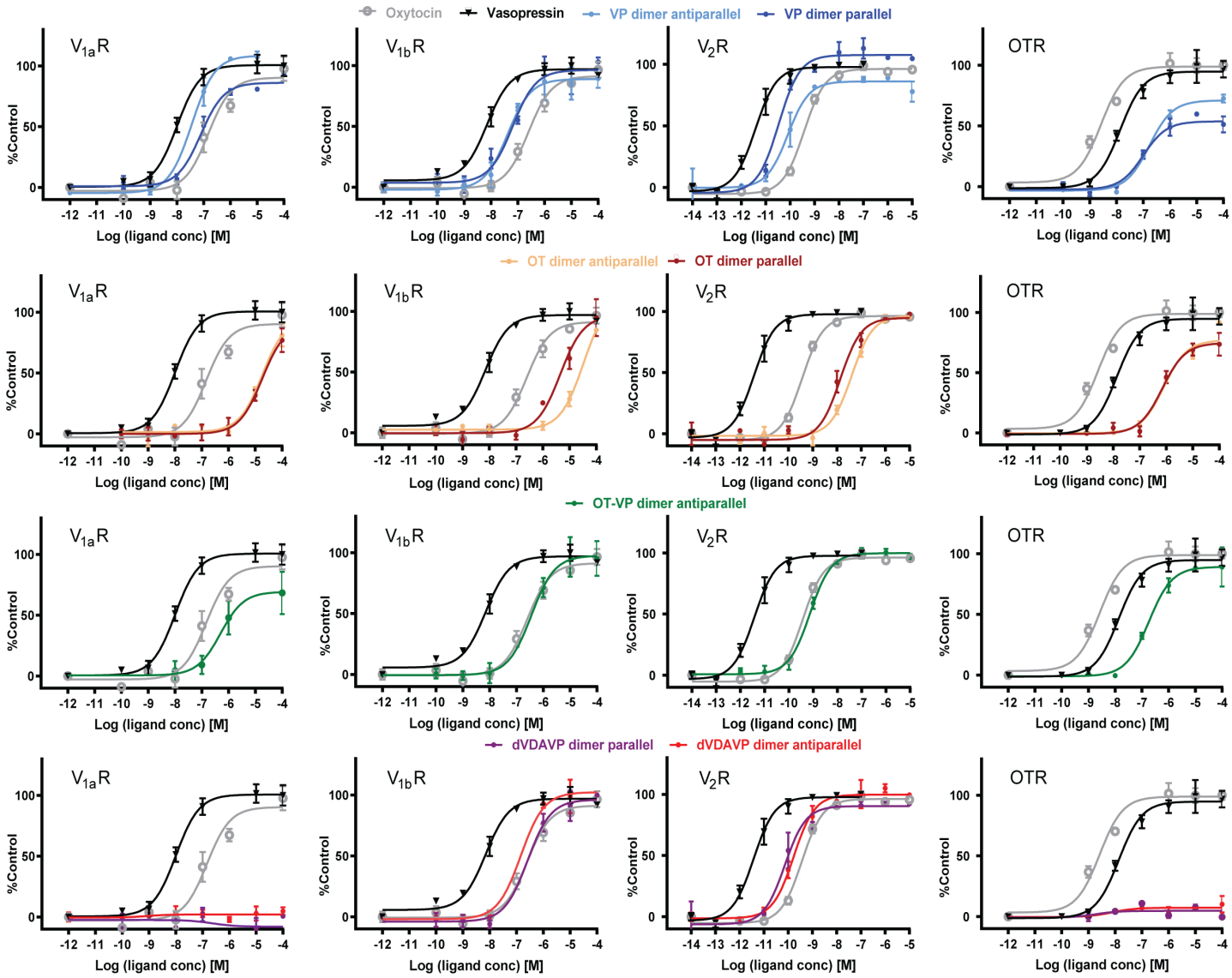


Figure S3. Concentration-response curves of parallel and antiparallel dimers at the individual receptor types using the IP-one (OTR, $V_{1a}R$, $V_{1b}R$) and cAMP (V_2R) second messenger assays. Receptors were heterologously expressed in HEK293 cells; experiments were performed in technical triplicates. N ≥ 3.

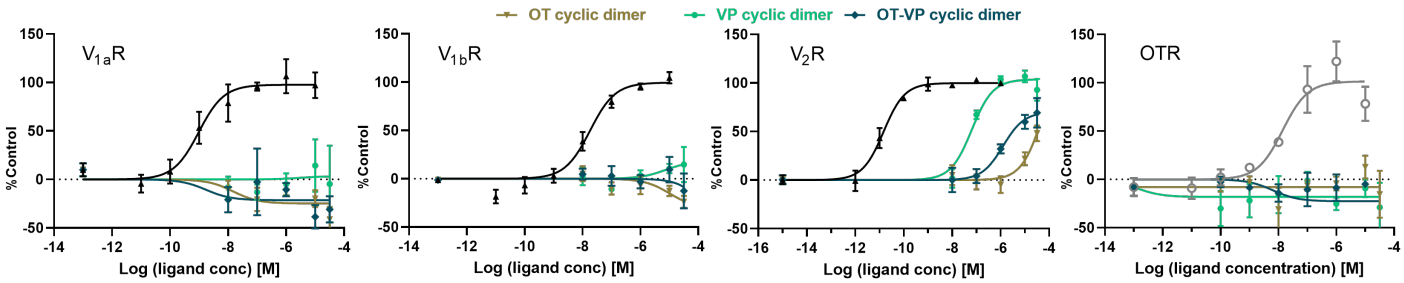


Figure S4. Concentration-response curves of the N- to C-terminal cyclised dimers at the individual receptor types using the IP-one ($V_{1a}R$, $V_{1b}R$, OTR) and cAMP (V_2R) second messenger assays. Receptors were heterologously expressed in HEK293 cells; experiments were performed in technical triplicates. N ≥ 3.

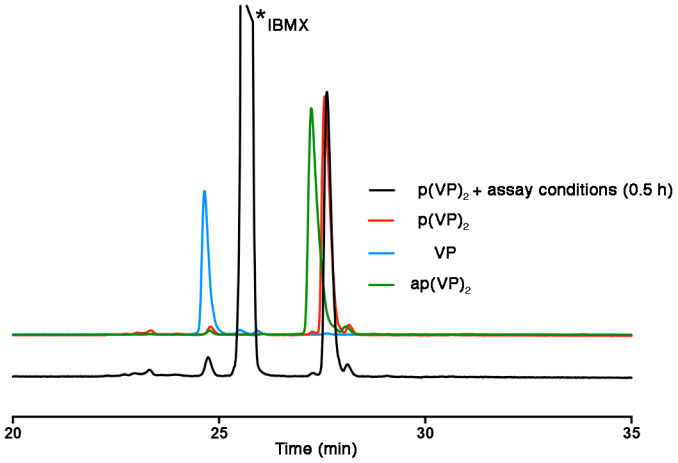


Figure S5. Analytical C_{18}-RP-HPLC study of parallel VP homodimer before and after the V_2R cAMP assay. HPLC comparison with monomeric VP and antiparallel VP homodimer shows that the $p(VP)_2$ remains intact after 30 min. IBMX, 3-isobutyl-1-methylxanthine, cAMP assay component.

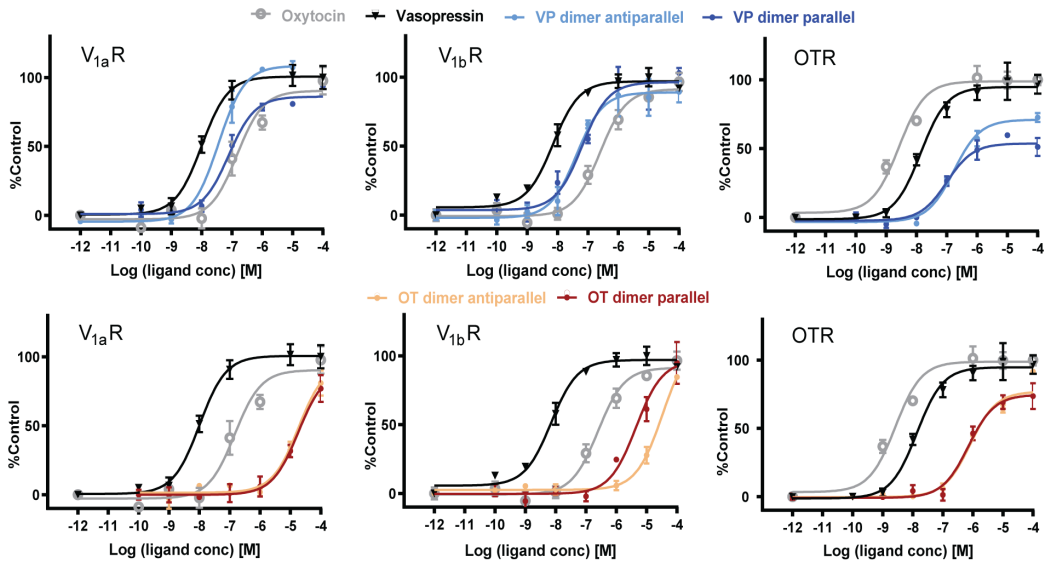


Figure S6. Concentration-response curves of parallel and antiparallel OT and VP homodimers at the individual receptor types ($V_{1a}R$, $V_{1b}R$, OTR) using a Ca^{2+} FLIPR bioassay. Receptors were heterologously expressed in COS-1 cells; experiments were performed in technical triplicates. N ≥ 3.

Table S2. Potencies of ligands tested via Ca^{2+} FLIPR bioassay at human OTR, $V_{1a}R$ and $V_{1b}R$.

	hV_{1a}R	**hV_{1b}R**	**hOTR**
Ligand	**EC_{50}(nM)**±SEM	**EC_{50}(nM)**±SEM	**EC_{50}(nM)**±SEM
VP	8.2±1.5	4.5±1.5	18.2±9.8
VP dimer parallel	53.1±15.6	49.9±18.5	99.5±12.1
VP dimer antiparallel	60.1±15.0	64.0±28.9	157.1±28.3
OT	149.9±101.7	324.7±37.0	5.7±4.0
OT dimer parallel	>10000	4759.0±353.5	620.5±167.6
OT dimer antiparallel	>10000	>10000	828.7±8.7

N ≥ 3 independent experiments

STRUCTURE DETERMINATION OF SPONTANEOUSLY FORMED dVDAVP DIMER

RP-HPLC and MALDI-MS analysis for dVDAVP dimer determination

RP-HPLC analysis was performed using a Dionex Ultimate 3000 HPLC (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Separations were performed using a Phenomenex Kinetex C_{18} (100 Å, 2.6 µm, 150 x 3 mm) stationary phase and a mobile phase consisting of solvent C (0.1% TFA in water) and solvent D (90% ACN/0.1% TFA in water). The flow rate was adjusted to 0.3 mL/min, and separation was performed using a linear gradient of 1%/min from 5–65% solvent D. Samples were dissolved in solvent C and injected with a volume of 40 µL. Absorbance was recorded at 214/254/280 nm. MALDI-MS spectra were generated on a MALDI-TOF/TOF 4800 analyser (ABSciex, Darmstadt, Germany) using the α-cyano hydroxy cinnamic acid matrix (saturated in 50% ACN/0.1% TFA in water) for sample preparation. Matrix and samples were mixed in a ratio of 6:1 and 0.5 µL were spotted onto the MALDI target plate. MS and MS/MS spectra were generated using the reflector mode by acquiring up to 5000 spectra. The device was internally calibrated before each experiment for both data acquisition modes.

Structure determination using ion mobility (IM-) and electrospray ionisation (ESI-)MS

Samples were reconstituted with MeOH, 1% formic acid (1:1, v/v and infused by automated nano-electrospray using a Triversa Nanomate (Advion BioSciences, Essex, UK) as the interface. The settings for spray voltage and gas pressure in the Nanomate were 1.65 kV and 0.5 psi respectively. Nano ESI-IM-MS experiments were performed on a Synapt G1 HDMS mass spectrometer (Waters Corp., Manchester, UK). The instrument was set in positive polarity and using the W TOF reflector mode. Sample cone voltage, extraction cone voltage, and source temperature were set at 20 V, 5 V and 20°C, respectively. The source pumping speed in the backing region of the mass spectrometer was reduced to a 4 mbar to avoid ion unfolding. Trap and transfer collision energies were set at 6 and 4 V, respectively. Trap gas and IM gas flows were 5 and 24 mL/min, respectively. The wave amplitude was set at 8 V. The ion mobility resolving power of the instrument used for this analysis was 10 (where the resolving power is defined as $\Omega/\Delta\Omega$; Ω=collisional cross section area). The instrument was calibrated over the m/z-range between 500 and 8000 Da using a solution of caesium iodide. MassLynx 4.1 and Driftscope 2.1 software (Waters Corp., Manchester, UK) were used for both the MS and IM processing. The purified dimer was diluted in 50 mM NH_4HCO_3, pH 8.2 and digested overnight with chymotrypsin (2%, w/w) at 37ºC. Digested sample was analysed by nano-LC-MS/MS on a NanoAcquity Ultra Performance LCTM chromatographic system (Waters Corp., Milford, USA) coupled to a LTQ-FT Ultra mass spectrometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and using a nanospray Advion interface source (Advion Triversa Nanomate, Advion Biosciences, Essex, UK). A volume of 1 µL of the dimer digest (1 µM in 1% aqueous formic acid solution) was preloaded onto a C_{18} Symmetry trap column (180 µm × 2 cm, Waters Corp., Milford, USA) at a flow rate of 15 µL/min for 3 min and eluted at 250 nL/min using a C_{18} analytical column (BEH130™ 75 µm × 25 cm, 1.7 µm, Waters Corp., Milford, USA) with a 60-min chromatographic run from 1 to 35% solvent F in 60 min (solvent E = 0.1% formic acid (FA) in water, solvent F = 0.1% FA in ACN). The spray voltage in the NanoMate source was set at 1.7 kV. The capillary voltage and tube lens on the LTQ-FT Ultra mass spectrometer were tuned to 40 V and 120 V, respectively. The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the FT with the resolution (defined at 400 m/z) set at 100,000. Up to six of the most intense ions per scan were fragmented and detected in the linear ion trap. The ion count target value was 1,000,000 for the survey scan and 50,000 for the MS/MS scan. Target ions already selected for MS/MS were dynamically excluded for 15 s. A minimal signal required to trigger the MS to MS/MS switch was set at 1,000 spectral counts. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for

fragmentation. Data were processed using the StavroX v2.4 search engine.[11] Searches were run against a FASTA database, which contained the monomer sequence. Search parameters were set considering interchain disulfide linkages (cross-links) between both Cys and deamino Cys, to interrogate the presence of digested fragments originated from the parallel or antiparallel dimer. Search parameters included chymotrypsin specificity, allowing for two missed cleavage sites. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.8 Da. Cross-linked peptides (with disulfide bridge linkages) with high scores identified from StravroX searches were validated manually.

Structure determination – Results

During oxidative folding of dVDAVP monomer we observed the spontaneous formation of a second product with similar mass (60% by analytical HPLC). To reveal the identity of this earlier-eluting folding product, we performed high-resolution MS as well as analytical RP-HPLC studies with the two folding products as well as commercially available monomeric dVDAVP (Figure S7). MALDI-MS confirmed the two characteristic ions species of the singly charged $[M+H]^+$ ion (2079.74 m/z) and the $[M+2H]^{2+}$ double charged mass signal (1040.14 m/z) of dimeric dVDAVP. By contrast synthetic dVDAVP and commercial dVDAVP displayed only the expected monomeric monoisotopic mass of the singly charged $[M+H]^+$=1040.44 m/z (Figure S7a-b). Synthetic dVDAVP and commercial dVDAVP had the same retention times, while dimeric dVDAVP eluted earlier (Figure S7c). All three peptides had the same mass and retention time upon reduction with 10 mM DTT and alkylation with 100 mM iodoacetamide (Figure S7d). MALDI-TOF/TOF experiments of the reduced and alkylated species also confirmed an identical amino acid sequence (Figure S7e). Taken together, we confirmed that the later-eluting product peak corresponded to monomeric dVDAVP, whereas the earlier-eluting more hydrophilic side product corresponded to its disulfide homodimer.

Given the possible disulfide bond connectivity between the two monomeric chains, such a dimer could adopt either a parallel or an antiparallel directionality (Figure S8a). Additionally, it could also be present as a co-eluting mixture of both. We therefore assessed the structural complexity of the dimer by ion-mobility mass spectrometry (IM-MS). This technique separates gaseous ions not only on the basis of their mass to charge ratio (m/z) but also of their size and shape.[12, 13] Therefore, it allows both the separation of conformers and of ions in different aggregation states. The isolated m/z 1040.14 ion (z=2) corresponding to the dimer resolved as a single peak in the 2D plot of ion mobility drift time *vs.* m/z, as shown by the IM-MS analysis (Figure S8b). This result confirmed the presence of only a single dimeric species, which correlated well with the single peak in the analytical RP-HPLC trace.

MS/MS experiments were performed on the intact dimer on a LTQ-FT Ultra instrument (linear ion trap with a Fourier transform ion cyclotron resonance MS detector) to reveal the dimer's directionality. Fragments obtained using collision-induced fragmentation and electron capture dissociation fragmentation techniques could not unambiguously identify the nature of the dimer and, consequently, an enzymatic digestion step was added prior to the MS/MS analysis. The nano-LC-MS/MS analysis of the chymotryptic cleavage fragments of the dimer revealed the presence of two specific antiparallel fragments (Figure S8c). No specific parallel fragments with distinct m/z values were detected. Overall, the performed MS experiments confirmed the unequivocal identity of the folding side product being the antiparallel dVDAVP homodimer.

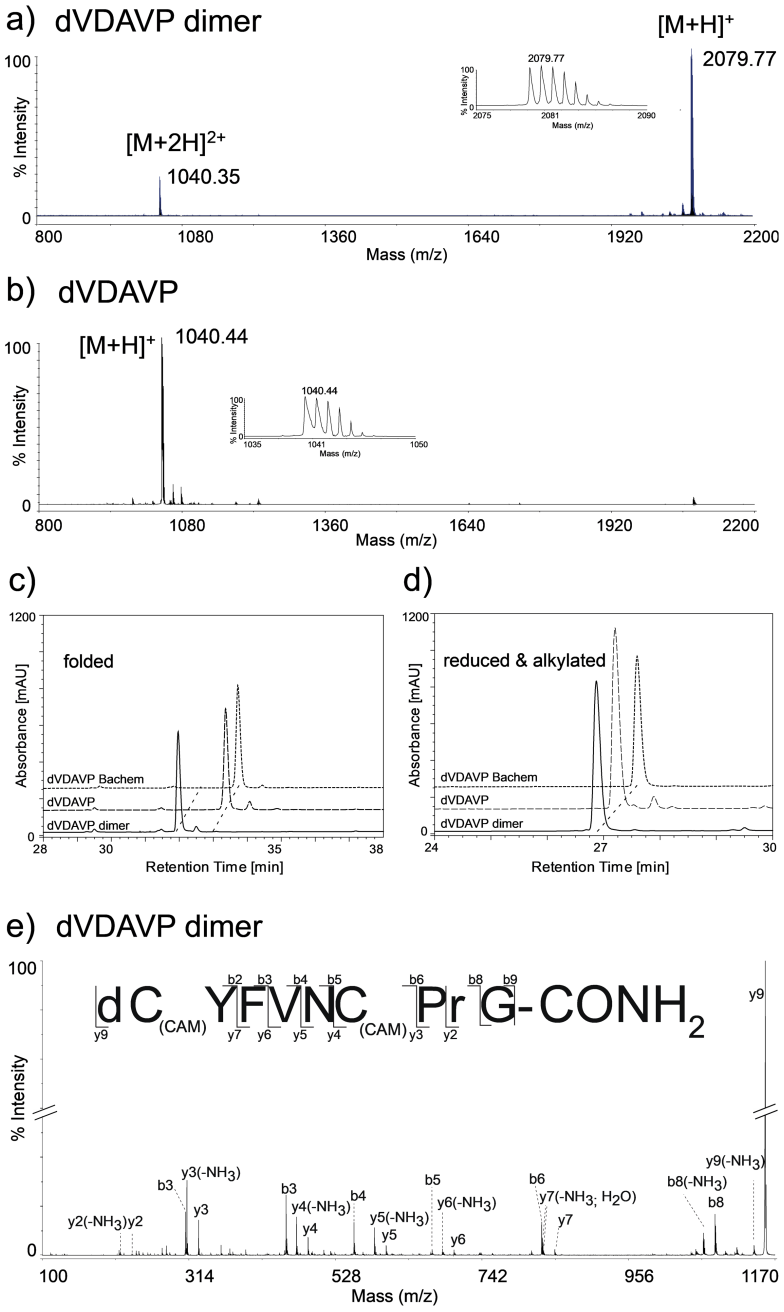


Figure S7. Characterisation of the dVDAVP analogues by analytical RP-HPLC and MS. a) Analysis of the isolated folding side product by MALDI-MS revealed two characteristic ions species: the mass signal of the singly charged $[M+H]^+$ ion (2079.74 m/z) and the $[M+2H]^{2+}$ double charged mass signal (1040.14 m/z). **b)** By contrast, synthetic and commercial dVDAVP displayed only the expected monomeric monoisotopic mass of $[M+H]^+$ = 1040.44 m/z (corresponding to dCYFVNCPrG*; d, deamino, *, C-terminal amide). **c)** Overlay of RP-HPLC chromatograms of the two folding products and commercial dVDAVP. The dimer product eluted earlier versus the co-eluting synthetic and commercial dVDAVP. **d)** All compounds had the same retention time upon reduction with 10 mM DTT and alkylation with 100 mM iodoacetamide. **e)** The two reduced and alkylated folding products were used for peptide fragmentation experiments applying a MALDI-TOF/TOF, and the amino acid sequence was elucidated *via* manual fragment ion assignment. Both peptides were found to be identical by amino acid sequence. The representative MS/MS fragmentation spectrum of the dimer is shown; ***'CAM' refers to carbamidomethyl.***

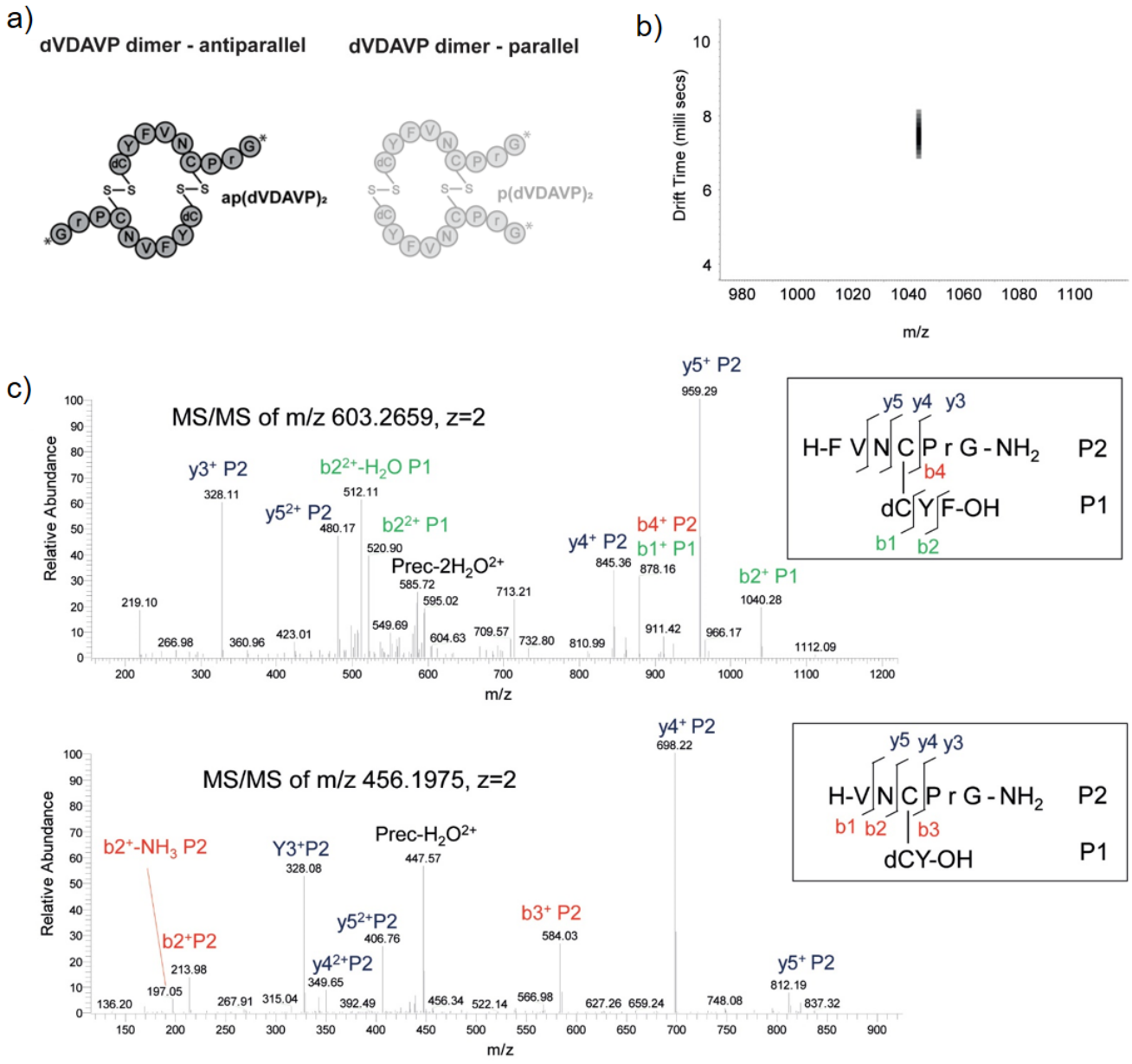


Figure S8. Structural characterisation of dVDAVP homodimer by MS. a) Schematic illustration of the two possible directionalities (parallel and antiparallel) of the spontaneously formed dVDAVP homodimer. **b)** Dimer ion mobility-MS analysis: 2D plot of ion mobility drift time versus m/z. The isolated m/z 1040 ion (z=2) corresponded to the dimer and was resolved as a single peak confirming that no other dimeric species were present. **c)** Chymotryptic digestion of the homodimer and LC-MS/MS analysis indicated the presence of two specific and unequivocal antiparallel homodimer fragments; MS/MS spectrum of m/z 603.2659 (z=2) corresponded to the antiparallel chymotryptic fragment (dCYF-OH)-(H-FVNCPrG-NH_2) (top panel) and MS/MS spectrum of m/z 456.1975 (z=2) corresponded to the antiparallel chymotryptic fragment (dCY-OH)-(H-VNCPrG-NH_2) (lower panel); no specific parallel homodimer fragments with distinct m/z values were detected. 'prec.' refers to precursor ion.

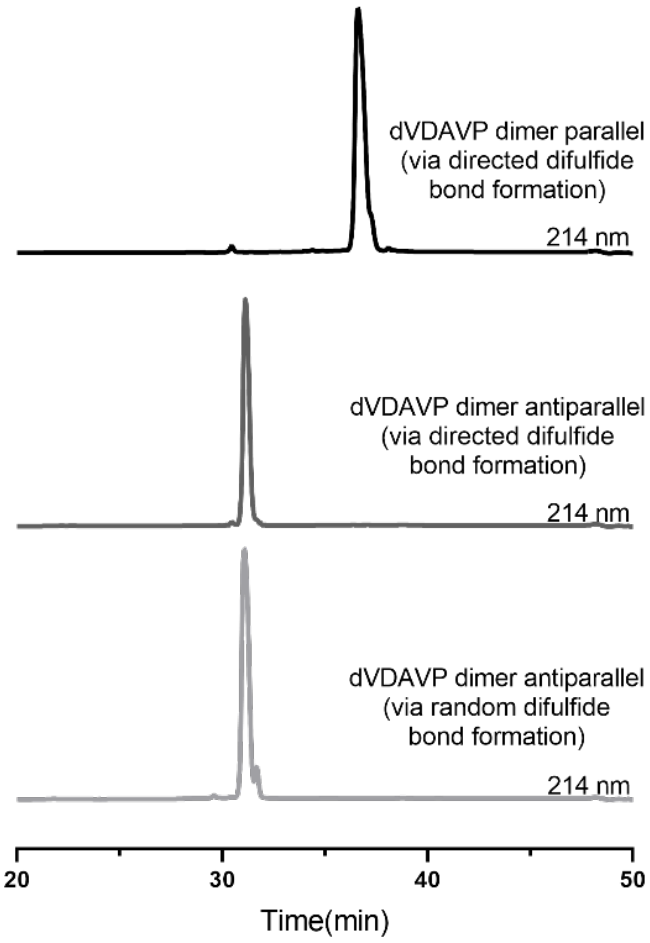


Figure S9. Analytical RP-HPLC traces of parallel and antiparallel dVDAVP homodimers, purified using gradient of 10 to 55% HPLC solvent B (90% ACN/0.043% TFA/H_2O ACN/0.043% TFA/H_2O) over 30 min.

DIMER STRUCTURAL CHARACTERIZATION, MODELING AND RECEPTOR DOCKING

Circular dichroism (CD) spectroscopy

CD measurements were performed on a Jasco J-810 spectropolarimeter, using a quartz cell of path length 0.1 cm. Spectra were recorded in the range of 260–185 nm at a rate of 20 nm/min, 1 nm data pitch, 2 s response, and 1 nm bandwidth, with 5 scans averaged for each sample. Peptide concentrations were 50 and 25 μM for monomers and dimers respectively in 10 mM sodium phosphate (pH 7.4).

Modelling of parallel and antiparallel dimer structures

To model the parallel and antiparallel dVDAVP (**dC**YF**V**NCP**r**G*) homodimers, we first predicted the structure of VP (CYFQNCPRG*) using the PEP-FOLD web interface;[14] VP was selected as a starting point since PEP-FOLD only works with natural amino acids. The obtained VP structure was mirrored along the axis formed by the two SH-residues and dimerized, either in a parallel (p) or an antiparallel (ap) fashion (after rotating one monomer by 180°), by forming two intermolecular disulfide bonds. The obtained $p(VP)_2$ and $ap(VP)_2$ structures were then remodelled into $p(dVDAVP)_2$ and $ap(dVDAVP)_2$ using built-in routines of the YASARA software package. The obtained structures were subsequently fed into molecular dynamics (MD) simulations.

MD-simulations were performed using the YASARA software-package. The AMBER 03 force field was employed with periodic boundary conditions.[15] Non-bonded interactions were cut off at 10.5 Å. Long-range Coulombic interactions were treated by a smoothed particle-mesh Ewald method.[16] Non-native amino acids were built using YASARA and semi-quantum-mechanically parameterised (YAPAC-AM1). The models of the polymers were built according to the polymers used for the experimental studies. The modelled structures were subject to energy minimisation in vacuum, subsequently randomly placed in the simulation box and solvated by water at pH 7, charge neutralised by addition of 1 % NaCl, and again minimised (steepest descent minimisation followed by simulated annealing). The chosen time increment was 2 fs. MD trajectories of 10.0 μs length were accumulated at 37°C for the two systems. Intermolecular forces were recalculated at every second simulation sub-step. Temperature rescaling was employed with a set-temperature of 37°C. The box dimensions (cubic of 37 Å side length) were controlled so as to yield a solvent pressure of 1 bar. Snapshots of the simulations were taken every 10,000 fs. The described method is likely to yield structures that fluctuate around local energetic minima and the presented data should be interpreted that way. The structures presented in the main text correspond to the final structures of the respective MD runs.

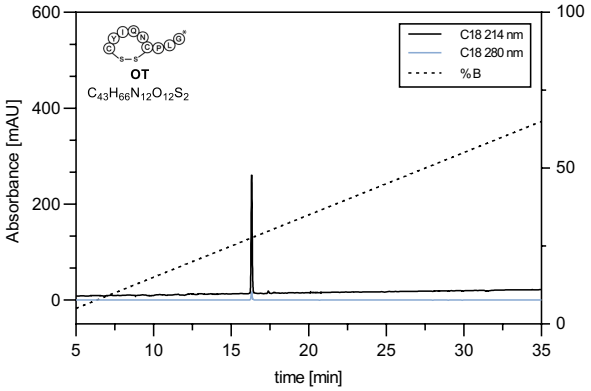
Receptor docking

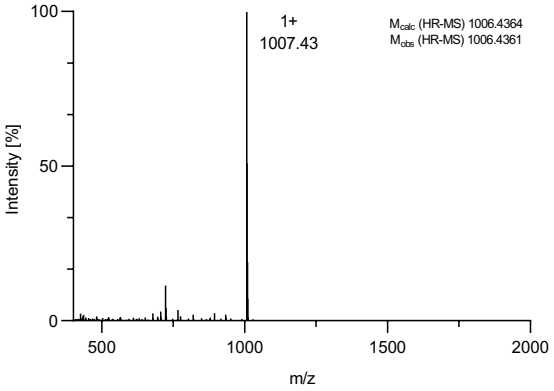
Insights into dimeric ligand-receptor interactions were generated using VINA[17] as implemented in YASARA, the dimer structures reported in the main text and a recently reported receptor homology model of $hV_{1a}R$.[18] 25 structures were generated for each dimer and an energetically favourable structure was selected for further energy minimisation. The selected receptor-ligand complexes were therefore subject to further molecular dynamics simulations, each for 1 μs in explicit solvent following the protocol detailed in the preceding paragraph.

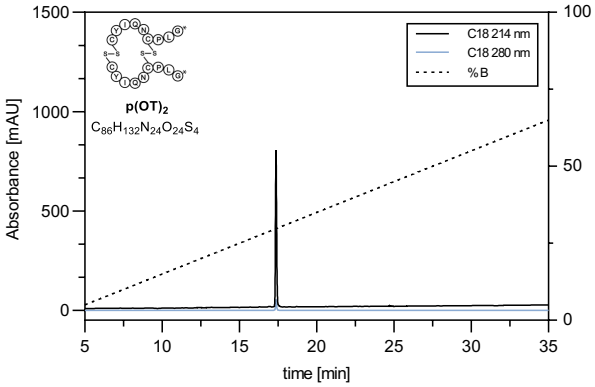
These previously identified $hV_{1a}R$ residues[18] were employed for the docking experiments: 7, 11, 50, 54, 57, 58, 61, 64, 65, 68, 81, 84, 85, 88, 91, 92, 133, 137, 140, 150, 152, 153, 155, 161, 164, 165, 168, 172, 208, 212, 215, 218, 219, 222, 223, 234, 238, 239, 242, 246, 249. Their mean spatial coordinates were calculated, and the result used to centre the binding pocket within a 10x10x10 nm large box as specified for the VINA runs.

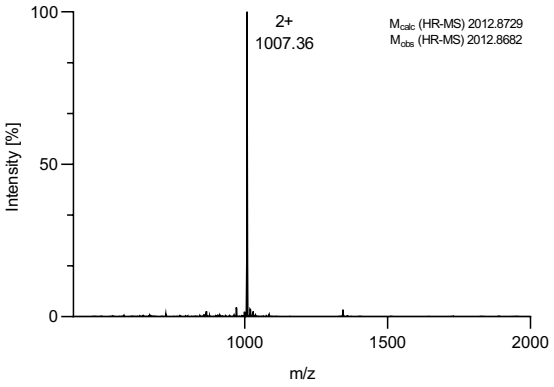
COMPOUND ANALYSIS

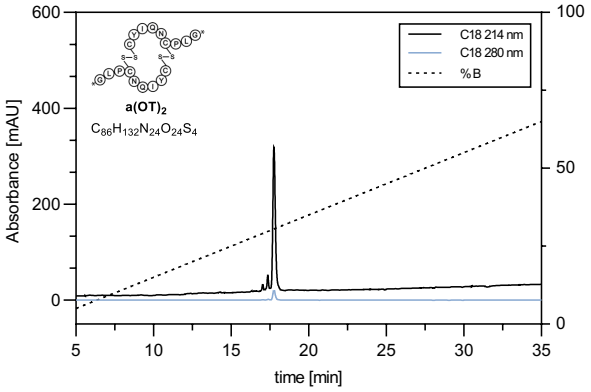
Analytical HPLC and MS data of synthesised monomeric and dimeric peptide structures

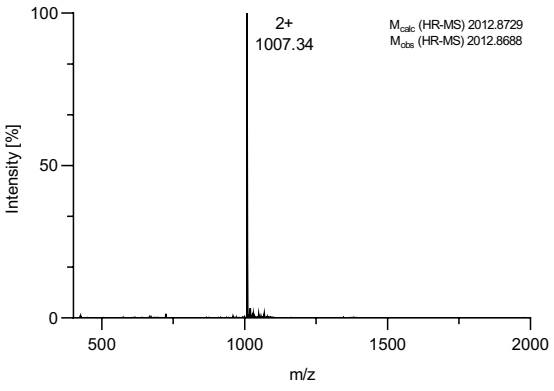


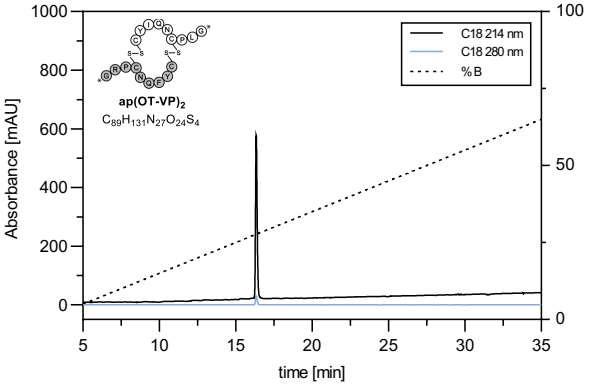


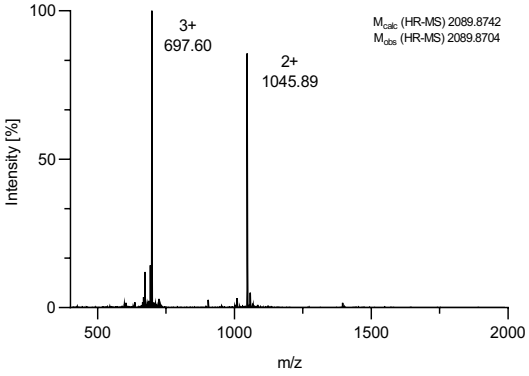


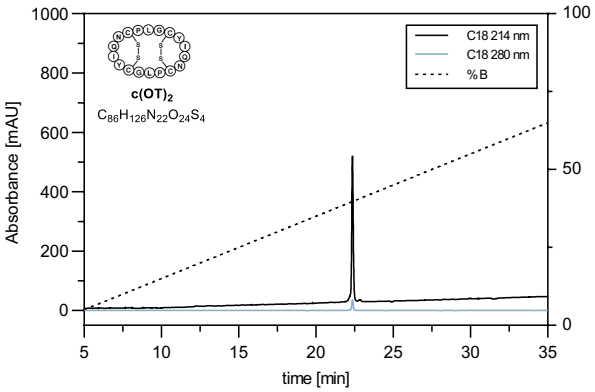


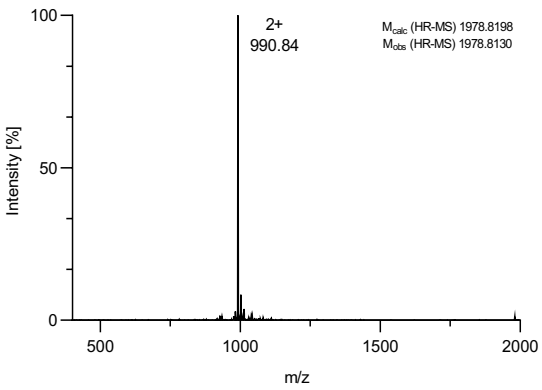


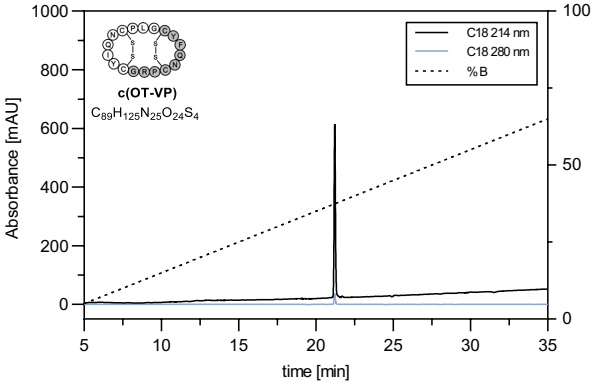


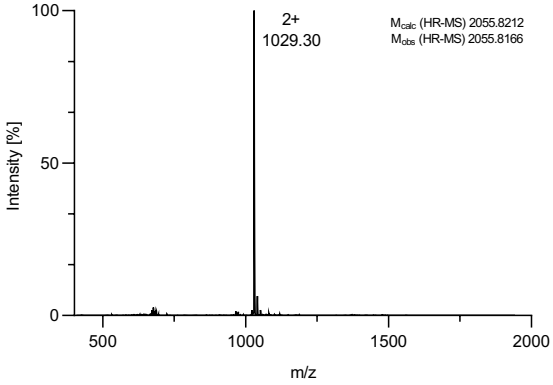

C18 214 nm
C18 280 nm
%B
ap(OT-VP)2
C89H131N27O24S4
Absorbance [mAU]
time [min]

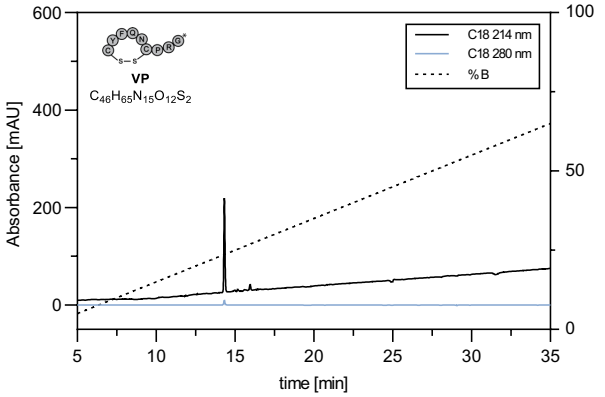

3+
697.60
2+
1045.89
Mcalc (HR-MS) 2089.8742
Mobs (HR-MS) 2089.8704
Intensity [%]
m/z

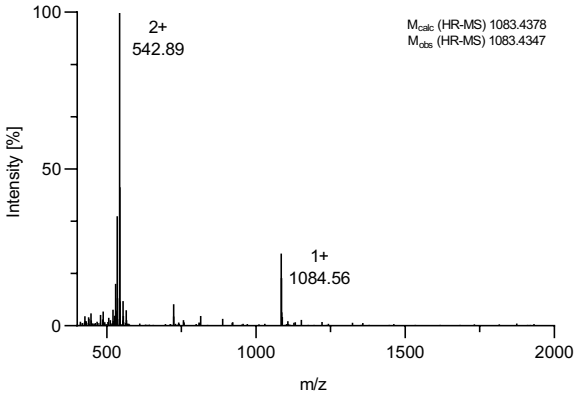

C18 214 nm
C18 280 nm
%B
c(OT)2
C86H126N22O24S4
Absorbance [mAU]
time [min]

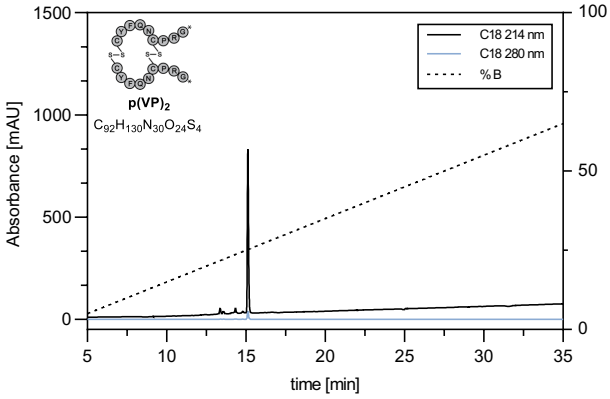

2+
990.84
Mcalc (HR-MS) 1978.8198
Mobs (HR-MS) 1978.8130
Intensity [%]
m/z

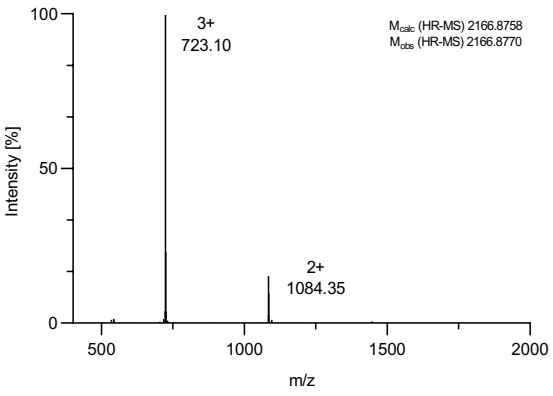

C18 214 nm
C18 280 nm
%B
c(OT-VP)
C89H125N25O24S4
Absorbance [mAU]
time [min]

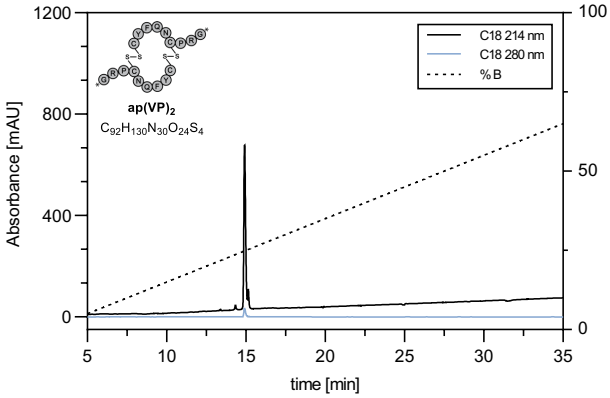

2+
1029.30
Mcalc (HR-MS) 2055.8212
Mobs (HR-MS) 2055.8166
Intensity [%]
m/z

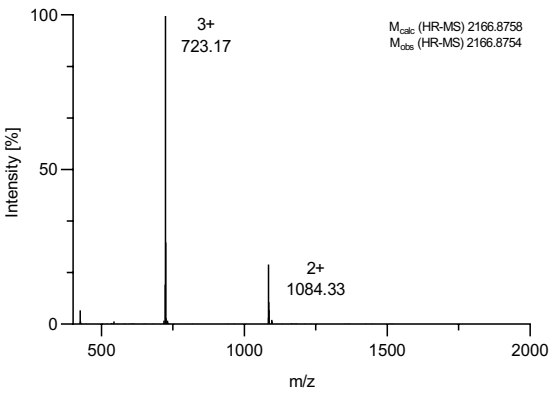

VP
$C_{46}H_{65}N_{15}O_{12}S_2$
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]

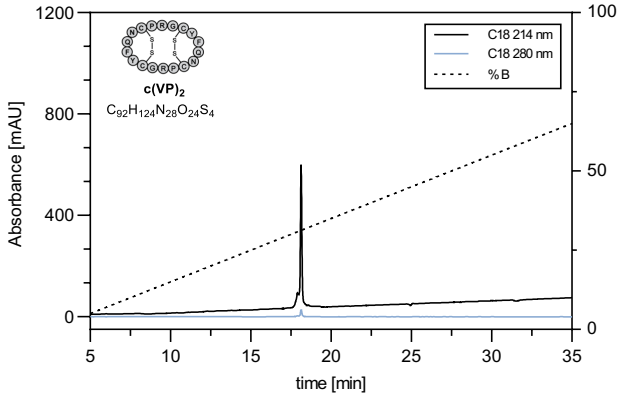

2+
542.89
1+
1084.56
M_{calc} (HR-MS) 1083.4378
M_{obs} (HR-MS) 1083.4347
Intensity [%]
m/z

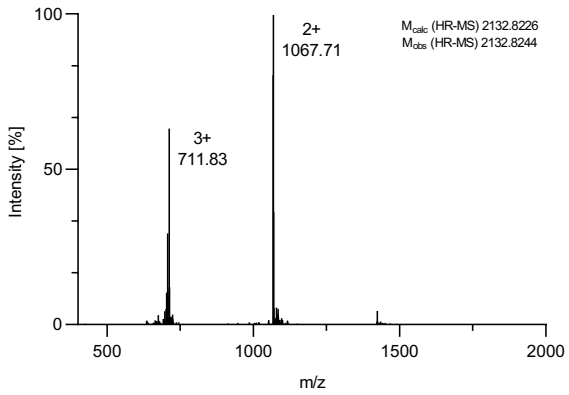

p(VP)$_2$
$C_{92}H_{130}N_{30}O_{24}S_4$
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]

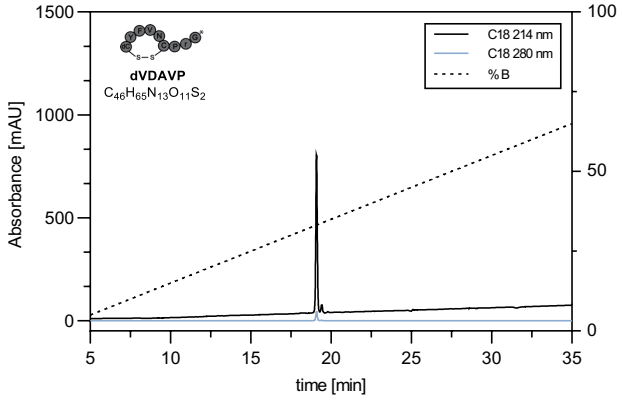

3+
723.10
2+
1084.35
M_{calc} (HR-MS) 2166.8758
M_{obs} (HR-MS) 2166.8770
Intensity [%]
m/z

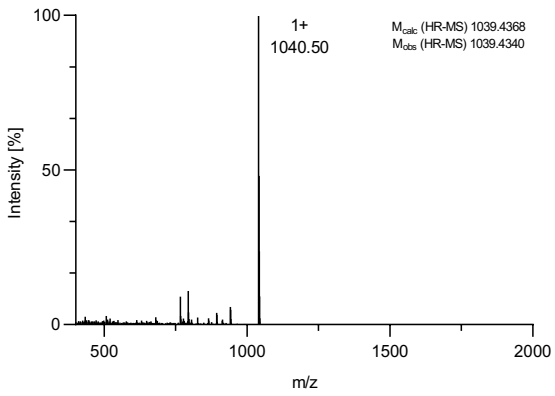

ap(VP)$_2$
$C_{92}H_{130}N_{30}O_{24}S_4$
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]

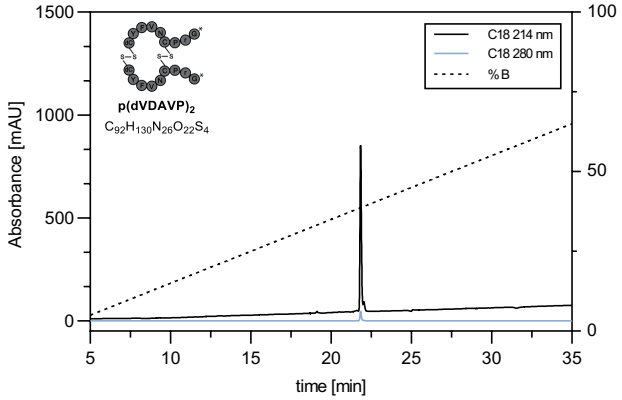

3+
723.17
2+
1084.33
M_{calc} (HR-MS) 2166.8758
M_{obs} (HR-MS) 2166.8754
Intensity [%]
m/z

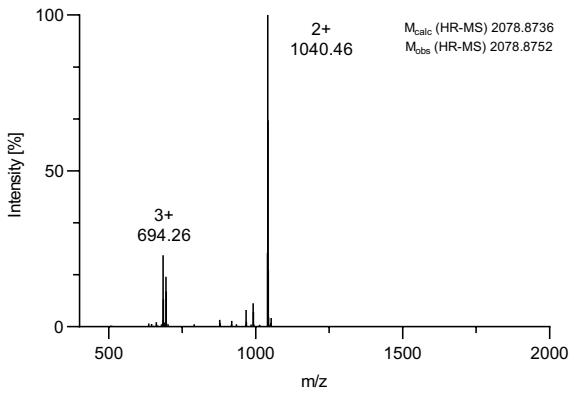

c(VP)2
C92H124N28O24S4
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]

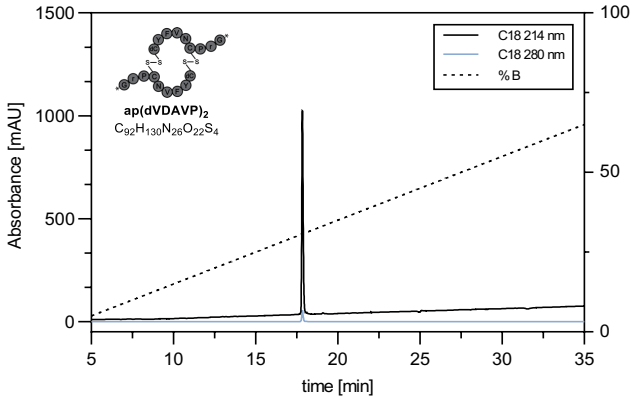

2+
1067.71
3+
711.83
Mcalc (HR-MS) 2132.8226
Mobs (HR-MS) 2132.8244
Intensity [%]
m/z

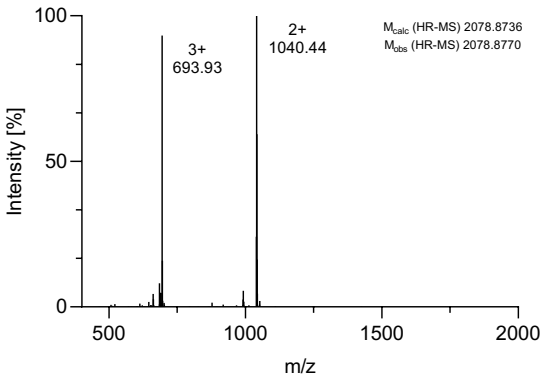

dVDAVP
C46H65N13O11S2
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]


1+
1040.50
Mcalc (HR-MS) 1039.4368
Mobs (HR-MS) 1039.4340
Intensity [%]
m/z


p(dVDAVP)2
C92H130N26O22S4
C18 214 nm
C18 280 nm
%B
Absorbance [mAU]
time [min]


2+
1040.46
3+
694.26
Mcalc (HR-MS) 2078.8736
Mobs (HR-MS) 2078.8752
Intensity [%]
m/z





REFERENCES

1. P. Alewood, D. Alewood, L. Miranda, S. Love, W. Meutermans and D. Wilson, *Methods Enzymol.*, 1997, **289**, 14-29.
2. J. S. Zheng, S. Tang, Y. K. Qi, Z. P. Wang and L. Liu, *Nat. Protoc.*, 2013, **8**, 2483-2495.
3. S. Eissler, M. Kley, D. Bachle, G. Loidl, T. Meier and D. Samson, *J. Pept. Sci.*, 2017, **23**, 757-762.
4. M. Schnolzer, P. Alewood, A. Jones, D. Alewood and S. B. Kent, *Int. J. Pept. Protein Res.*, 1992, **40**, 180-193.
5. M. Ruizgayo, F. Albericio, M. Pons, M. Royo, E. Pedroso and E. Giralt, *Tetrahedron Lett.*, 1988, **29**, 3845-3848.
6. J. S. Zheng, S. Tang, Y. Guo, H. N. Chang and L. Liu, *Chembiochem*, 2012, **13**, 542-546.
7. C. Hicks, W. Jorgensen, C. Brown, J. Fardell, J. Koehbach, C. W. Gruber, M. Kassiou, G. E. Hunt and I. S. McGregor, *J. Neuroendocrinology*, 2012, **24**, 1012-1029.
8. J. Koehbach, M. O'Brien, M. Muttenthaler, M. Miazzo, M. Akcan, A. G. Elliott, N. L. Daly, P. J. Harvey, S. Arrowsmith, S. Gunasekera, T. J. Smith, S. Wray, U. Goransson, P. E. Dawson, D. J. Craik, M. Freissmuth and C. W. Gruber, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 21183-21188.
9. L. Duerrauer, E. Muratspahic, J. Gattringer, P. Keov, H. C. Mendel, K. D. G. Pfleger, M. Muttenthaler and C. W. Gruber, *Sci. Rep.*, 2019, **9**, 19295.
10. P. Keov, Z. Liutkeviciute, R. Hellinger, R. J. Clark and C. W. Gruber, *Sci. Rep.*, 2018, **8**, 10020.
11. M. Gotze, J. Pettelkau, S. Schaks, K. Bosse, C. H. Ihling, F. Krauth, R. Fritzsche, U. Kuhn and A. Sinz, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 76-87.
12. J. L. Benesch and B. T. Ruotolo, *Curr. Opin. Struct. Biol.*, 2011, **21**, 641-649.
13. D. E. Clemmer and M. F. Jarrold, *J. Mass Spectrom.*, 1997, **32**, 577-592.
14. Y. Shen, J. Maupetit, P. Derreumaux and P. Tufféry, *J. Chem. Theory Comput.*, 2014, **10**, 4745-4758.
15. R. Salomon-Ferrer, D. A. Case and R. C. Walker, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2013, **3**, 198-210.
16. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
17. O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455-461.
18. M. G. Di Giglio, M. Muttenthaler, K. Harpsoe, Z. Liutkeviciute, P. Keov, T. Eder, T. Rattei, S. Arrowsmith, S. Wray, A. Marek, T. Elbert, P. F. Alewood, D. E. Gloriam and C. W. Gruber, *Sci. Rep.*, 2017, **7**, 41002.