Syntheses of Cyanophycin Segments for Investigations of Cell-Penetration

Marcel Grogg§
Donald Hilvert
Albert K. Beck
Dieter Seebach*

Laboratorium für Organische Chemie (LOC), Department Chemie und Angewandte, Biowissenschaften (D-CHAB), ETH-Zürich, Hönggerberg HCI, Vladimir-Prelog-Weg 3, 8093 Zürich, Switzerland seebach@org.chem.ethz.ch

§ This author did all the experiments

Published as part of the 50 Years SYNTHESIS – Golden Anniversary Issue

Received: 15.06.2018
Accepted: 15.06.2018
Published online: 28.06.2018
DOI: 10.1055/s-0037-1610202; Art ID: ss-2018-z0415-op

License terms: 

Abstract Novel guanidinium-rich oligopeptide derivatives R-[Adp(X)]8-NH2 are described, which consist of an octa-aspartic acid backbone with argininyalted side chains that are derived from the biopolymer cyanophycin [H-(Adp)n-OH]. The Fmoc-Adp(X,Pbf)-OH building blocks for solid-state peptide synthesis (SSPS) of Adp octamers were prepared from Fmoc-Arg(Pbf)-OH and Fmoc-Asp-OAll. Coupling on PAL resin provided four octamers with and without N-terminal fluorescent groups (FAM) and C-terminal amide groups. Milligram quantities of Adp-octamers were isolated after preparative HPLC purification. The structure of the novel guanidinium-rich oligomers is unique insofar as the side chains of the Asp8-backbone include both a guanidino and a carboxylic acid group, the influence of which will be tested with the corresponding ester and amide derivatives that were synthesized in parallel. Unusual cell-penetrating properties of the Adp-octamers are expected.

Key words guanidinium-rich oligopeptides, β/α-Asp-Arg-dipeptide building block, biopolymer cyanophycin, solid-state peptide synthesis, cell-penetrating peptides

Arginine- and lysine-rich natural (Tat1, Penetratin2) and unnatural peptides (oligo-arginines and other guanidinium-rich compounds)3,4 are cell-penetrating peptides (CPPs) and can carry a large variety of cargoes into prokaryotic and eukaryotic cells.3–5 A schematic representation is shown in Figure 1 A; the number of guanidinium groups is usually between 4 and 14. For a more detailed discussion and additional literature, see the citations in references3–5 and in the introduction of our recent paper on cell penetration, herbicidal activity, and in vivo toxicity of guanidinium-rich compounds.6

Despite all the activity in this field there has been no attention paid by the CPP community, so far, to the biopolymer cyanophycin (Figure 1 B), a guanidinium-rich natural product, which was discovered in characteristic granules in blue-green algae by the Italian botanist Antonio Borzi in 1887 and chemically identified by R. D. Simon in 1971.7 In recent years, the biopolymer cyanophycin, a temporary microbial nitrogen storage material of cyanobacteria, has been studied most comprehensively by the group of A. Steinbüchel.8,9 The polymer and its dipeptidic building block can be produced using industrial equipment ‘on any desired scale’.10 Cyanophycin is a polyaspartic acid argininyalted on the carboxylic acid groups of the side-chains, and the building block is a dipeptide with aspartic acid incorporated as a β-amino acid10 (Figure 1 B). For simplicity, we use the three-letter code Adp for the cyanophycin building block.6 Since
peptides with an N-terminal β2-1-amino acid residue are not cleaved by common aminopeptidases,10–12 H-Adp-OH should be quite stable under physiological conditions.

To be able to find out whether cyanophycin segments with a length typical of CPPs (vide supra)3–5 have cell penetrating properties we decided to synthesize octamer derivatives (cf. Figure 1B, with n = 8) by conventional solid-state peptide synthesis (SSPS) using Fmoc chemistry.

For this purpose, the readily available dipeptide H-Adp-OH9 [with permission of A. Sallam (Cysal GmbH, Technologiehof Münster) and of dp] (Figure 2) looked like a convenient starting material, but this would have required its modification by selectively (!) protecting the guanidino and the carboxylic acid group in the Arg-residue and by putting an Fmoc group on the N-terminus of the Asp residue. Instead, we synthesized suitably protected dipeptide derivatives 2 with free amino groups, to which the Asp moiety was attached by reaction with Fmoc-Asp-OAllyl under peptide-coupling conditions to produce the three Fmoc-Adp(Pbf,X)-OAllyl derivatives 4a–c. De-allylation with phenylsilane/Pd(PPh3)4 led to the building blocks Fmoc-Adp(Or-Bu,Pbf)-OH (5), Fmoc-Adp(OMe,Pbf)-OH (6), and Fmoc-Adp(NMe2,Pbf)-OH (7), ready for SSPS (see Scheme 2). Overall yields of up to 40% could be attained for the four steps from Fmoc-Arg(Pbf)-OH (1) to the Adp building blocks 5–7 (for details, see Experimental part). Compound 5 with a t-Bu ester group was actually prepared as precursor to Adp-octamers 8a and b with free carboxylic acid groups, formed concomitantly with removal of the peptide from the resin by trifluoroacetic acid (TFA).

As resin for the SSPS we used N-alkylated PAL,13 the Fmoc-groups of the growing chains were removed with piperidine in DMF, and the couplings were achieved with HATU/Hünig base (DIPEA) in DMF. The same conditions were employed for attachment of the N-terminal fluorescent FAM label (Scheme 3). Release of the peptide chains from the resin and removal of the Pbf protecting groups was performed with TFA–H2O–TIS, and the products were purified by preparative HPLC. Milligram amounts of the octa-Adp-carboxamides 8–10 (Schemes 2, 3) were synthesized in this way.

We also prepared the octa-Adp-amides with methyl ester 9 and amide groups 10 in the side chains in order to be able to compare the biological activities of Adp-octamers with and without a possible internal neutralization of positively charged guanidinium by negatively charged carboxylate groups (see 8a and b and formulae in Figure 1B).

The determination of i.v. toxicities and the cell-penetrating properties of octa-Adp derivatives described herein are reported in ref.6 and will be described in a separate paper,14 respectively.

Protected amino acids and the PAL resin were purchased from Bachem, HATU from Aaptec, 5(6)-carboxyfluorescein (5/6-FAM) from abcr and all other chemicals were purchased from Sigma Aldrich. All reagents were used as received, solvents were technical grade, and the reactions were run in open flasks fitted with PFTE...
coated magnetic stir bars at r.t., unless otherwise noted. Peptide couplings were carried out in ISOLUTE® Double fritted filtration column, 15 mL 20 μm PE (reaction vessel, Biotage) at r.t. The building blocks for peptide synthesis were activated in 4 mL screw vial 45 × 14.77 mm (activation vessel, BGB) closed with PFTE lined cap 13-425 (Thermo Scientific) at r.t. Analytical TLC was performed with Merck 60 F254 pre-coated aluminum silica plates and visualized by UV detection (254 nm). Flash column chromatography (FC) was performed using SiliCycle (SilaFlash® P60, 230–400 mesh particle size) silica gel. All fractions collected by FC were analyzed by TLC to identify the different compounds. Melting points were recorded on a Büchi melting point B-540 device.

IR spectra were recorded on a PerkinElmer Spectrum Two FT-IR spectrometer at r.t. using ATR as the sampling technique. NMR spectra were recorded on a Bruker Advance-III 400 MHz spectrometer in the NMR Service at the Laboratory of Organic Chemistry (LOC), ETH Zurich. 1H NMR spectra were recorded relative to the residual solvent peak (CDCl3 δH = 7.26, DMSO- d6 δH = 2.50) and reported as follows: chemical shift (ppm), multiplicity (standard abbreviations; ovlp: overlap), coupling constant (Hz), and integration. 13C NMR spectra were recorded relative to residual solvent peaks (CDCl3 δC = 77.0, DMSO- d6 δC = 39.5). All 1H and 13C signals were assigned via HSQC and HMBC experiments. 19F NMR (D2O) spectra were recorded with 1H decoupling.

The LCMS runs were performed with a Waters Acquity UPLC system equipped with an H-class quaternary solvent manager, an H-class sample manager FTN with sample organizer, a PDA detector, a SQ detector 2, and a 1.7 μm 2.1 × 50 mm BEM C18 UPLC column. Eluent system: H2O and MeCN containing 0.1% HCO2H using a flow of 1 mL/min.

Scheme 2  Deallylation of 4a–c to give the SSPS-building blocks Fmoc-Adp(X,Pbf)-OH 5–7 for assembly to the target octa-Adp derivatives without (8a, 9a, and 10a) and with (8b, 9b, and 10b) a FAM fluorescence label.

Scheme 3  Assemblies of 8 Adp-building blocks by SSPS on PAL resin (Bachem, 200–400 mesh, 0.28 mmol/g) to give the octamers 8–10 with C-terminal amide groups on milligram scale. TIS: Trisopropylsilane.
(Gradient: 0–0.2 min, 5% MeCN; 0.2–1.5 min, 5–80% MeCN; 1.5–2 min, 80–100% MeCN; 2–2.2 min, 100% MeCN; 2.2–2.3 min, 100–5% MeCN; 2.3–3 min, 5% MeCN).

All semi-preparative HPLC runs were performed with a Waters preparative 150 LC system equipped with a 2545 quaternary gradient module, a 2489 UV/visible detector, a Fraction Collector III and 5 µm 20 × 150 mm Reprosil-pure 120 C18 AQ column (Dr. Maisch GmBH, Basel, Switzerland). Eluent system: H2O and MeCN containing 0.1% TFA, using a flow of 10 mL/min (Gradient: 0–0.5 min, 30% MeCN; 0.5–70, 30–90% MeCN; 70–75 min, 90% MeCN. Gradient 2: 0–5 min, 10% MeCN; 5–45 min, 10–90% MeCN; 45–55 min, 90% MeCN).

All analytical HPLC runs were performed with a Dionex Ultimate 3000 system equipped with a 3000 pump module, a 3000 Autosampler, a 3000 RS Variable Wavelength Detector, and a 3.5 µm XBridge C18 column. Eluent system: H2O and MeCN added with 0.1% TFA, using a flow of 1 mL/min (Gradient: 0–1 min, 5% MeCN; 1–20 min, 5% MeCN; 20–23 min, 95% MeCN).

All self-measured MALDI-TOF spectra were recorded on a Bruker microflex benchtop MALDI-TOF system. High-resolution mass spectrometry (HRMS) was performed on a Bruker maXis UHR-TOF by electrospray ionization (ESI) or a Bruker solariX by matrix-assisted laser desorption/ionization (MALDI) by the Molecular and Biomolecular Analysis Service (MoBiAS) of the LOC at ETH Zurich. MoBiAS also performed elemental analysis.

**Preparation of Building Block 5**

**tert-Butyl N-’-[(9H-Fluoren-9-yl)methoxy]carbonyl]-N-’-[(2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl]argininate (2a)**

In a 100 mL round-bottomed flask, Fmoc-Arg(Pbf)-OH (6.0 g, 9.25 mmol) and DMAP (119.4 mg, 977 µmol, 0.11 equiv) were dissolved in CH2Cl2 (20 mL). Then, t-BuOH (4 mL, 41.82 mmol, 4.52 equiv) was added and the flask was cooled for 5 min on ice. To the cooled mixture, DCC (2.11 g, 10.23 mmol, 1.11 equiv) was added and the resulting solution was stirred for further 5 min on ice. Afterwards, the flask was left with stirring for 3 h 40 min at rt. The reaction mixture was then poured into 1 M aq HCl (20 mL) and extracted with CH2Cl2 (2 × 40 mL). The organic layers were combined and dried (MgSO4) before being evaporated in vacuo. The crude product was purified by PC (EtOAc–CH2Cl2, 1:1). All collected fractions were analyzed by TLC, combined, and evaporated to give compound 2a in 50% yield (3.25 g, 4.32 mmol) as a colorless powder; Rf = 0.54 (EtOAc–CH2Cl2, 1:1).

**HRMS (ESI):** m/z [M + H]+ calcd for C38H49N4O7S: 705.3316; found: 705.3316.

**tert-Butyl N-’-[(9H-Fluoren-9-yl)methoxy]carbonyl]-N-’-[(2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl]argininate (3a)**

In a 250 mL round-bottomed flask, compound 2a (3.05 g, 4.33 mmol) was dissolved in piperidine–DMF (50 mL 1:4) and stirred for 80 min. The organic solvent was then evaporated in vacuo and in the crude reaction mixture purified 2 times by PC (CH2Cl2–MeOH, 9:1). All collected fractions were analyzed by TLC, combined, and evaporated to give compound 3a (2.08 g, 4.31 mmol, quant) as a yellowish oil; Rf = 0.27 (CH2Cl2–MeOH, 9:1).

**HRMS (ESI):** m/z [M + H]+ calcd for C38H49N4O7S: 705.3316; found: 705.3316.
CH₂), 2.54 (s, 3 H, Pbf CH₃), 2.50 (s, 3 H, Pbf CH₃), 2.09 (s, 3 H, Pbf CH₃), 1.91–1.51 (m ovlp, 4 H, 2 × Arg CH₂), 1.47 (s, 6 H, 2 × Pbf CH₃), 1.45 (s, 9 H, t-CH₃).

¹¹C NMR (101 MHz, CDCl₃): δ = 171.47 (C=O amide), 171.01 (C=O allyl ester), 170.77 (C=O t-Bu ester), 161.14 (Pbf αC-O), 156.31 (Gua C=N), 153.81 (Fmoc C=O), 143.93 (Fmoc αC=C), 143.85 (Fmoc γC=C), 141.41 (Fmoc αC=C), 140.95 (Pbf αC=C), 134.98 (Pbf αC=C), 131.38 (allyl C=C), 127.88 (Fmoc αC=C), 127.23 (Fmoc αC=C), 125.85 (Pbf αC=C), 125.29 (Fmoc αC=C), 120.12 (Fmoc αC=C), 119.11 (allyl C=C), 118.74 (Pbf αC=C), 87.58 (Pbf CH₃), 83.53 (t-CH₃ C), 67.51 (Fmoc CH₂), 66.56 (allyl CH₂), 51.66 (α CH), 50.80 (α CH), 47.22 (Fmoc CH), 43.06 (Pbf CH₃), 40.90 (Asp CH₂), 37.89 (Arg CH₂), 30.93 (Arg CH₂), 28.68 (Pbf CH₂), 28.05 (t-CH₃ C), 24.24 (Arg CH₃), 19.38 (Pbf CH₃), 17.75 (Pbf CH₂), 12.59 (Pbf CH₃).

LCMS (ESI): m/z = 862.3 (100%) [M + H⁺]; tᵣ = 1.73 min.


Methyl N⁴-[[(9H-Fluoren-9-yl)methoxy][carbonyl]-N⁴-[(1-[(t-butoxy)-1-oxo-5-[3-[(2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl]guanidino]pentan-2-yl]asparagine (5)]

In a 100 mL round-bottomed flask under N₂, compound 4a (2.65 g, 3.08 mmol) was dissolved in CH₂Cl₂ (25 mL). Then, PhSiH₃ (750 µL, 6.16 mmol, 2 equiv) and Pd(PPh₃)₄ (313.1 mg, 0.248 mmol, 0.1 equiv) were added. The resulting solution was stirred for 2.5 h, poured in 1 M aq HCl (20 mL) and extracted with CH₂Cl₂ (3 × 40 mL). The organic layers were combined, dried (MgSO₄) and evaporated in vacuo. The crude reaction mixture was resuspended in H₂O–MeCN + 0.1% TFA and aq NaHCO₃ was slowly added to the remaining aqueous phase until saturation (vigorous gas evolution). Subsequently, the aqueous phase was dissolved in piperidine–DMF (1:4, 10 mL) and stirred for 1 h. The crude reaction mixture was purified by FC (MeOH–CH₂Cl₂, 1:9). All the collected fractions were analyzed by TLC, combined, and evaporated to give compound 5 in 95% yield (885 mg, 2.13 mmol) as a colorless powder; Rf = 0.26 (EtOAc–CH₂Cl₂, 1:1).

¹¹C NMR (400 MHz, DMSO-d₆): δ = 7.73 (d, J = 7.6 Hz, 2 H, Fmoc CH₃), 7.55 (d, J = 7.1 Hz, 2 H, Fmoc CH₃), 7.37 (t, J = 7.5 Hz, 2 H, Fmoc CH₃), 7.31–7.22 (m, 2 H, Fmoc CH₂), 6.29–5.92 (m, 3 H, Gua NH), 5.67 (d, J = 8.2 Hz, 1 H, amide NH), 4.36 (d, J = 6.5 Hz, 2 H, Fmoc CH₂), 4.34–4.24 (m, 1 H, Fmoc CH), 4.16 (t, J = 6.9 Hz, 1 H, Arg α CH), 3.70 (s, 3 H, OCH₃), 3.20 (ddt, J = 18.8, 13.5, 7.1 Hz, 2 H, Arg CH₂), 2.90 (s, 2 H, Pbf CH₂), 2.57 (s, 3 H, Pbf CH₃), 2.50 (s, 3 H, Pbf CH₃), 2.07 (s, 3 H, Pbf CH₃), 1.92–1.63 (m, 2 H, Arg CH₂), 1.63–1.51 (m, 2 H, Arg CH₂), 1.42 (s, 6 H, 2 × Pbf CH₃).

Preparation of Building Block 6

Methyl N⁴-[[[(9H-Fluoren-9-yl)-methoxy][carbonyl]-N⁴-[[2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl]argininate (2b)]

In a 50 mL round-bottomed flask, Fmoc-Arg(Pbf)-OH (1.65 g, 2.54 mmol) was dissolved in neat SOCl₂ (2 mL) and stirred for 30 min. Then the flask was placed on an ice bath and MeOH (10 mL) was added dropwise under vigorous stirring. The mixture was stirred overnight at r.t. Then, it was poured into 1 M aq HCl (20 mL) and extracted with CH₂Cl₂ (3 × 40 mL). The organic layers were combined and dried (MgSO₄) before being evaporated in vacuo. The crude reaction mixture was purified by FC (EtOAc–CH₂Cl₂, 1:1). All collected fractions were analyzed by TLC, combined, and evaporated to give compound 2b in 86% yield (1.45 g, 2.54 mmol) as a colorless powder; Rf = 0.26 (EtOAc–CH₂Cl₂, 1:1).

¹¹H NMR (400 MHz, CDCl₃): δ = 7.73 (d, J = 7.6 Hz, 2 H, Fmoc CH₃), 7.55 (d, J = 7.1 Hz, 2 H, Fmoc CH₃), 7.37 (t, J = 7.5 Hz, 2 H, Fmoc CH₃), 7.31–7.22 (m, 2 H, Fmoc CH₂), 6.29–5.92 (m, 3 H, Gua NH), 5.67 (d, J = 8.2 Hz, 1 H, amide NH), 4.36 (d, J = 6.5 Hz, 2 H, Fmoc CH₂), 4.34–4.24 (m, 1 H, Fmoc CH), 4.16 (t, J = 6.9 Hz, 1 H, Arg α CH), 3.70 (s, 3 H, OCH₃), 3.20 (ddt, J = 18.8, 13.5, 7.1 Hz, 2 H, Arg CH₂), 2.90 (s, 2 H, Pbf CH₂), 2.57 (s, 3 H, Pbf CH₃), 2.50 (s, 3 H, Pbf CH₃), 2.07 (s, 3 H, Pbf CH₃), 1.92–1.63 (m, 2 H, Arg CH₂), 1.63–1.51 (m, 2 H, Arg CH₂), 1.42 (s, 6 H, 2 × Pbf CH₃).

In a 100 mL round-bottomed flask, compound 4b (1.91 g, 1.46 mmol) was dissolved in CH2Cl2 (10 mL) and N2 was bubbled through the solution for 10 min. Then, PhSH (4 mL, 32.45 mmol, 22.52 equiv) and Pd(PPh3)4 were added (66.6 mg, 57.6 mmol, 0.04 equiv). The solution was stirred for 2 h. Then, the mixture was poured into 1 M aq HCl (20 mL) and extracted with CH2Cl2 (3 × 40 mL). The organic layers were combined and dried (MgSO4) before being evaporated in vacuo. The crude reaction mixture was separated by FC (CH2Cl2–MeOH, 95:5 + 0.5% TFA). All the fractions containing the product (TLC test) were combined and evaporated. Compound 6 was isolated by semi-preparative HPLC with 3 injections using gradient 2 in 51% yield (575 mg, 0.74 mmol) as a colorless powder; mp 118 °C; Rf = 0.29 (MeOH–CH2Cl2, 5:9:5 + 0.5% TFA).

IR (neat): 3332 (w), 2972 (w), 2936 (w), 1723 (m), 1663 (s), 1574 (m), 1544 (m) 1450 (m), 1372 (w), 1388 (w), 1331 (m), 1289 (m), 1293 (m), 1200 (s), 1140 (s), 1091 (s), 992 (m), 849 (m), 781 (m), 760 (m), 741 (s), 722 (m), 641 (s), 613 (s) cm⁻1.

1H NMR (400 MHz, DMSO-d6): δ = 8.31 (d, J = 7.4 Hz, 1 H, amide NH), 7.88 (d, J = 7.4 Hz, 2 H, Fmoc CH), 6.79 (d, J = 7.5 Hz, 2 H, Fmoc CH2), 7.46 (d, J = 8.3 Hz, 1 H, amide NH), 7.41 (t, J = 7.4 Hz, 2 H, Fmoc CH2), 7.36–7.28 (m, 2 H, Fmoc CH2), 6.69 (s, 1 H, Gua NH), 6.41 (s, 1 H, Gua NH), 4.34 (td, J = 8.1, 5.3 Hz, 1 H, Fmoc CH3), 4.30–4.14 (m, ovlp, 4 H Fmoc CH2 + 2 H CH2), 3.59 (s, 3 H, OCH3), 3.01 (d, J = 6.1 Hz, 2 H, Arg CH2), 2.94 (s, 2 H, Pbf CH2), 2.69–2.52 (m, 2 H, Asp CH2), 2.47 (s, 3 H, Pbf CH3), 2.41 (s, 3 H, Pbf CH3), 1.73–1.47 (m, 2 H, Arg CH2), 1.39 (s, 6 H, 2 × Pbf CH3), 1.47–1.28 (m, 2 H Arg CH2).

13C NMR (101 MHz, DMSO-d6): δ = 172.97 (C=O amide), 172.35 (C=O Me ester), 169.31 (CO2H), 157.45 (Pbf CH2), 156.02 (Pbf CO), 155.76 (Gua C≡N), 143.79 (Pbf CH2), 143.75 (Pbf CH2), 140.70 (Pbf CH2), 140.69 (Fmoc CH2), 137.25 (Pbf CH2), 134.10 (Pbf CH2), 131.43 (Pbf CH2), 127.63 (Fmoc CH2), 127.07 (Fmoc CH2), 125.23 (Fmoc CH2), 124.33 (Pbf CH2), 121.00 (Pbf CH2), 116.26 (Fmoc CH2), 86.30 (Pbf CH2), 65.70 (Fmoc CH3), 51.80 (α CH), 51.79 (OCH3), 50.49 (Fmoc CH), 46.59 (α CH), 42.44 (Pbf CH3), 39.73 (Arg CH2 ovlp with DMSO signal), 36.59 (Asp CH2), 28.28 (Pbf CH2), 28.20 (Arg CH2), 25.40 (Arg CH2), 18.94 (Pbf CH3), 17.58 (Pbf CH3), 12.26 (Pbf CH3).

LC (ESI): m/z = 778.9 (100%) [M + H]+; tR = 1.65 min.


Preparation of Building Block 7

(9H-Fluoren-9-yl)methyl-(1-(dimethylamino)-1-oxo-5-[3-[2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl]sulfonyl]guanidino)pentan-2-yl](carbamate (2c)

In a 50 mL round-bottomed flask, Fmoc-Arg(Pbf)-OH (1.65 g, 2.54 mmol) was dissolved in neat SOCl2 (2 mL). The solution was stirred for 30 min. Then, SOCl2 was evaporated using a gentle N2 flow. Afterwards, 2 M NHMe2 in THF (11.5 mL, 22.44 mmol, 10 equiv) was added dropwise by cooling the mixture on ice. The mixture was stirred overnight at rt, poured into 1 M aq HCl (20 mL) and extracted with CH2Cl2 (3 × 40 mL). The organic layers were combined and dried (MgSO4) before being evaporated in vacuo. The crude reaction mixture was purified by FC (CH2Cl2–MeOH, 95:5). All fractions collected were analyzed by TLC.
by TLC, combined, and evaporated to afford compound 2c in 68% yield (1.05 g, 2.29 mmol) as a yellowish solid; $R_f = 0.41$ (CH$_2$Cl$_2$–MeOH, 95:5).

1H NMR (400 MHz, CDCl$_3$): $\delta = 7.74$ (d, $J = 7.5$ Hz, 2 H, Fmoc CH$_\alpha$), 7.56 (dd, $J = 7.3$, 3.5 Hz, 2 H, Fmoc CH$_\alpha$), 7.42–7.34 (m, 2 H, Fmoc CH$_\alpha$), 7.31–7.24 (m, 2 H, Fmoc CH$_\alpha$), 6.21 (s, 1 H, Gua NH), 6.12 (s, 2 H, Gua NH), 6.04 (d, $J = 8.2$ Hz, 1 H, amide NH), 4.72–4.56 (m, 1 H, Fmoc CH), 4.44–4.26 (m, 2 H, Fmoc CH$_2$), 4.16 (t, $J = 6.9$ Hz, 1 H, Arg $\alpha$ CH), 3.40–3.08 (m, 2 H, Arg $\alpha$ CH), 3.01 (s, 3 H, NCH$_3$), 2.93 (s, 3 H, NCH$_3$), 2.92 (s, 2 H, Pbf CH$_2$), 2.57 (s, 3 H, Pbf CH$_3$), 2.23 (s, 2 H, NH$_2$), 2.08 (s, 3 H, Pbf CH$_3$), 1.74–1.60 (m, 2 H, Arg $\alpha$ CH), 1.45 (s, 6 H, 2 × Pbf CH$_3$), 1.25 (s, 2 H, Arg $\alpha$ CH).

13C NMR (101 MHz, CDCl$_3$): $\delta = 156.79$ (Gua C=O), 156.20 (Fmoc C=O), 143.87 (Fmoc $\alpha$C=C), 143.74 (Fmoc $\alpha$C=C), 141.43 (Fmoc $\alpha$C=C), 139.80 (Fmoc $\alpha$C=C), 133.14 (Pbf $\alpha$C=C), 128.62 (Pbf $\alpha$C=C), 127.12 (Pbf $\alpha$C=C), 127.07 (Pbf $\alpha$C=C), 125.12 (Fmoc $\alpha$C=C), 124.72 (Pbf $\alpha$C=C), 121.15 (Fmoc $\alpha$C=C), 115.79 (Pbf $\alpha$C=C), 68.49 (Pbf $\alpha$C=C), 66.43 (allyl CH$_2$), 47.26 (Arg $\alpha$ CH), 43.36 (Pbf CH$_2$), 41.13 (Arg $\alpha$ CH), 37.21 (NCH$_3$), 35.96 (NCH$_3$), 30.79 (Arg $\alpha$ CH), 28.72 (Pbf CH$_3$), 24.71 (Arg $\alpha$ CH), 19.41 (Pbf CH$_3$), 18.05 (Pbf CH$_3$), 12.61 (Pbf CH$_3$); (TFA traces).

HRMS (ESI): $[M + H]^+$ calcd for C$_{21}$H$_{36}$N$_5$O$_4$S: 454.2483; found: 454.2483.

**2-Amino-N,N-dimethyl-5-[3-[[2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl]sulfonyl]guanidino]pentamidine (3c)**

In a 100 mL round-bottomed flask, compound 2c (989 mg, 1.46 mmol) was dissolved in piperidine–DMF (1:4, 10 mL). The solution was stirred for 1 h. Then, 1 M aq HCl (30 mL) was added, and the mixture was extracted with CH$_2$Cl$_2$ (3 × 30 mL). The combined organic layers were combined, dried (MgSO$_4$) before being evaporated in vacuo. Further purification via FC (CH$_2$Cl$_2$–MeOH, 96:4) and TLC analysis of the collected fractions gave, after evaporation, compound 4c in 77% yield (938 mg, 1.13 mmol) as a colorless solid; $R_f = 0.30$ (MeOH–CH$_2$Cl$_2$, 9:5).

1H NMR (400 MHz, CDCl$_3$): $\delta = 7.76$ (d, $J = 7.6$ Hz, 2 H, Fmoc CH$_\alpha$), 7.60 (t, $J = 6.2$ Hz, 2 H, Fmoc CH$_\alpha$), 7.40 (t, $J = 7.5$ Hz, 2 H, Fmoc CH$_\alpha$), 7.30 (t, $J = 7.5$ Hz, 2 H, Fmoc CH$_\alpha$), 7.22 (d, $J = 7.0$ Hz, 1 H, amide NH), 6.58 (s, 3 H, Gua NH), 5.94 (d, $J = 7.3$ Hz, 1 H, amide NH), 5.85 (td, $J = 10.8$, 5.2 Hz, 1 H, allyl C=CH$_2$), 5.29 (d, $J = 17.1$ Hz, 1 H, allyl C=CH$_2$), 5.23 (d, $J = 10.5$, 1 H, allyl C=CH$_2$), 4.80 (t, $J = 7.9$ Hz, 1 H, $\alpha$ CH), 4.73–4.65 (m, 1 H, $\alpha$ CH), 4.65–4.53 (m, 2 H, $\alpha$ CH), 4.48–4.16 (m ovlp, 3 H, Fmoc CH + Fmoc CH$_2$), 3.36 (d, $J = 154.1$ Hz, 2 H, Arg CH$_2$), 3.02 (s, 3 H, NCH$_3$), 2.96 (s, 3 H, NCH$_3$), 2.95 (s, 2 H, Pbf CH$_3$), 2.14–2.80 (m, 2 H, Asp CH$_2$), 2.55 (s, 3 H, Pbf CH$_3$), 2.51 (s, 3 H, Pbf CH$_3$), 2.09 (s, 3 H, Pbf CH$_3$), 1.64 (dd ovlp, $J = 24.2$, 18.0, 11.3 Hz, 4 H, 2 × Arg $\alpha$ CH), 1.46 (s, 6 H, 2 × Pbf CH$_3$).

13C NMR (101 MHz, CDCl$_3$): $\delta = 171.07$ (C=O amide), 170.94 (C=O allyl ester), 160.84 (Pbf $\alpha$C=C), 156.09 (Fmoc C=O), 153.83 (Gua C=O), 143.79 (Fmoc $\alpha$C=C), 141.29 (Fmoc $\alpha$C=C), 140.64 (Pbf $\alpha$C=C), 134.89 (Pbf $\alpha$C=C), 131.24 (allyl C=C), 127.76 (Fmoc $\alpha$C=C), 127.29 (Pbf $\alpha$C=C), 127.10 (Fmoc C=O), 125.64 (Pbf $\alpha$C=C), 125.18 (Fmoc $\alpha$C=C), 120.00 (Fmoc $\alpha$C=C), 119.02 (allyl C=C), 118.53 (Pbf $\alpha$C=C), 87.38 (Pbf $\alpha$C=C), 67.36 (Fmoc CH$_2$), 66.43 (allyl CH$_2$), 50.66 ($\alpha$ CH), 48.12 ($\alpha$ CH), 47.09 (Fmoc CH$_3$), 42.97 (Pbf CH$_3$), 40.66 (Arg CH$_3$), 37.80 (Asp CH$_3$), 37.00 (NCH$_3$), 36.02 (NCH$_3$), 30.45 (Arg CH$_3$), 28.56 (Pbf CH$_3$), 24.08 (Arg $\alpha$ CH), 19.27 (Pbf CH$_3$), 17.65 (Pbf CH$_3$), 12.47 (Pbf CH$_2$); (TFA traces).

LCMS (ESI): $m/z = 833.5$ (100% [M + H$^+$]); $t_R = 1.64$ min.

HRMS (ESI): $m/z$ [M + H$^+$] calcd for C$_{36}$H$_{46}$N$_5$O$_6$S: 676.3163; found: 676.3152.
IR (neat): 3330 (w), 2971 (w), 2934 (w), 1719 (s), 1651 (s), 1627 (s), 1576 (m), 1549 (m), 1450 (m), 1408 (s), 1372 (w), 1332 (w), 1292 (m), 1254 (m), 1202 (m), 1165 (m), 1135 (s), 1090 (s), 1059 (m), 996 (m), 850 (m), 782 (m), 760 (s), 741 (s), 700 (m), 641 (s), 620 (m) cm⁻¹.

1H NMR (400 MHz, DMSO-d₆): δ = 8.13 (d, J = 8.2 Hz, 1 H, amide NH), 7.88 (d, J = 7.6 Hz, 2 H, Fmoc CH₃), 7.70 (d, J = 7.4 Hz, 2 H, Fmoc CH₃), 7.50 (d, J = 8.4 Hz, 1 H, amide NH), 7.41 (td, J = 7.5, 1.1 Hz, 2 H, Fmoc CH₃), 7.32 (dt, J = 7.5, 1.5 Hz, 2 H, Fmoc CH₃), 6.64 (s, 1 H, Gua NH), 6.40 (s, 1 H, Gua NH), 4.65 (q, J = 7.9 Hz, 1 H, Fmoc CH₂), 4.34 (td, J = 8.1, 5.5 Hz, 1 H, α CH), 4.30–4.45 (m, 3 H, Fmoc CH₂ + α CH), 3.06–2.97 (m, 2 H, Arg CH₂), 2.95 (s, 3 H, NCH₃), 2.93 (s, 2 H, Pbf CH₂), 2.80 (s, 3 H, NCH₃), 2.56 (dd, J = 17.0, 6.6 Hz, 2 H, Asp CH₂), 2.46 (s, 3 H, Pbf CH₂), 2.41 (s, 3 H, Pbf CH₂), 1.99 (s, 3 H, Pbf CH₂), 1.66–1.50 (m, 2 H, Arg CH₂), 1.39 (s, 6 H, 2 × Pbf CH₃), 1.48–1.25 (m, 2 H, Arg CH₂).

13C NMR (101 MHz, DMSO-d₆): δ = 172.99 (C=O amide), 170.88 (C=O NMe₂ amide), 168.73 (CO₂H), 157.42 (Pbf CH₂C=O), 156.03 (Fmoc C=O), 148.78 (Pbf CH₂C=O), 143.75 (Fmoc CH₂C=O), 140.69 (Fmoc CH₂C=O), 137.25 (Pbf CH₂C=O), 134.14 (Pbf CH₂C=O), 131.42 (Pbf CH₂C=O), 127.62 (Fmoc CH₂C=O), 127.06 (Fmoc CH₂C=O), 125.24 (Fmoc CH₂C=O), 124.31 (Pbf CH₂C=O), 120.09 (Fmoc CH₂C=O), 116.24 (Pbf CH₂C=O), 86.28 (Pbf CH₃), 67.52 (Fmoc CH₂), 50.67 (α CH), 48.04 (Fmoc CH₂), 46.58 (α CH), 42.44 (Pbf CH₃), 39.73 (Arg CH₂, ovlp with DMSO signal), 36.75 (Asp CH₂), 36.50 (NCH₃), 35.16 (NCH₃), 28.95 (Arg CH₂), 28.28 (Pbf CH₃), 25.09 (Arg CH₂), 18.94 (Pbf CH₂), 17.58 (Pbf CH₂), 12.26 (Pbf CH₃).

LCMS (ESI): m/z = 791.9 (100%) [M + H]+; tᵣ = 1.61 min.

HRMS (ESI): m/z [M + H]+ calcd for C₉₈H₁₄₈F₂₇N₄₀O₅₀ (hygroscopic!): 2187.0588; found: 2187.0588.

Preparation of the Octamers 8–10

Solid-State Peptide Synthesis; General Procedure (GP1)

All cyanophycin octamer derivatives were synthesized using the following protocol on a 0.03 mmol scale (building block 5 for 8a, 8b, 9a, and 10a, and for 9a, b, and 7 for 10a, b).

In a 15 mL reaction vessel equipped with a valve and attached to a suction system, ca. 107 mg of dry PAL resin (0.28 mmol/g) was shaken in DMF (5 mL) for 30 min. Then, the resin was treated with piperidine–DMF (1:4, 5 mL each). The reaction vessel was transferred to the reaction vessel and the activation vessel. Then, the reaction vessel was washed 5 × 1 min with DMF (5 mL each). The Fmoc protecting group was removed by using the same procedure as described in GP1. The resin was then washed 5 × 1 min with DMF (5 mL each) and then 5 × 1 min with CH₃Cl (5 mL each). Finally, the resin was dried by suction for approximately 25 min.

The dry resin was transferred to a tared 100 mL round-bottomed flask and weighed. Per g of dried resin, 50 mL of cleavage cocktail was added [R-[Adp(OH)]₆-NH₂ (8): TFA–TIS–H₂O 95:2.5:2.5; R-[Adp(OMe)]₆-NH₂ (9), and R-[Adp(NMe₂)]₆-NH₂ (10): TFA–TIS–MeOH 95:2.5:2.5] and the reaction mixture was stirred at r.t. under N₂ for 4 h.

Afterwards, the resin was filtered off using a fritted glass filter and washed with neat TFA. The TFA was evaporated with N₂ flow until some material started to precipitate. Then, ice cold Et₂O was added and the resulting suspension was filtered on Celite and rinsed with ice cold Et₂O to remove cleaved protecting groups. To solubilize the peptidic material, the Celite was resuspended 3 × in H₂O–MeCN + 0.13% TFA (40 mL each, 1:1) and filtered. The filtrate was frozen, lyophilized, and stored at 4 °C before purification via semi-preparative HPLC (gradient 1). All HPLC fractions were analyzed by analytical HPLC, combined according to purity (all samples considered >95% pure), and lyophilized to yield the TFA salt form of the peptides. Peptides 8a, 9a, and 10a were obtained as colorless powders (5–15 mg each; hygroscopic). The TFA content of the salt was determined by elemental analysis of 8a, according to which 9 TFA molecules were associated with the peptide. We assumed that this was also the case for 9a and 10a. For a typical desalting procedure, vide infra.

8a

HPLC: tᵣ = 6.6 min.

MS (MALDI): m/z (%) = 2189.444 (100%) [M + H]+.


Anal. Calcd for C₉₈H₁₄₈F₂₇N₄₀O₅₀ (hygroscopic!): C, 39.79%; H, 4.66; F, 16.03; N, 17.51; O, 25.00. Found: C, 36.04%; H, 4.95%; F, 14.79%; N, 17.18.

9a

HPLC: tᵣ = 7.8 min.

MS (MALDI): m/z (%) = 2298.692 (100%) [M + H]+.


Anal. Calcd for C₈₀H₁₅₆N₄₁O₃₂ (hygroscopic!): C, 39.79%; H, 4.66; F, 16.03; N, 17.51; O, 25.00. Found: C, 36.04%; H, 4.95%; F, 14.79%; N, 17.18.

10a

HPLC: tᵣ = 7.2 min.

MS (MALDI): m/z (%) = 2403.347 (100%) [M + H]+.


FAM Coupling Procedure (Peptides 8b, 9b, 10b)

All FAM couplings were performed on a 0.015 mmol scale.

The Fmoc-peptide-resin (0.015 mmol) stored in the reaction vessel was resuspended 10 min in DMF (5 mL) and treated 3 × (2 min, 8 min, 8 min) with piperidine–DMF (1:4, 5 mL each). The reaction vessel was then washed 5 × 1 min with DMF (5 mL each). Then, in an activation vessel 5/6-FAM (33.9 mg, 0.09 mmol, 6 equiv) was predissolved in DMF (5 mL). To this solution, 0.4 M HATU in DMF (215 μL, 0.086 mmol, 5.7 equiv) and DIPEA (31.4 μL, 0.18 mmol, 12 equiv) were added...
ed and the mixture was shaken for 10 min. The content of the activation vessel was then transferred to the reaction vessel and shaken overnight (16 h). The resin was then washed 5 × 1 min with DMF (5 mL each) and treated with piperidine–DMF (1:4, 5 mL) for 30 min. After this step, resin-bound 8b, 9b, and 10b were submitted to cleavage following GP2. The TFA salts of the FAM derivatives 8b, 9b, and 10b were obtained as yellow powders (2–3 mg each; hygroscopic).

8b
HPLC: tR = 8.4 min.
MS (MALDI): m/z = 2548.919 (100%) [M + H]+.

9b
HPLC: tR = 9.8 min.
MS (MALDI): m/z = 2657.685 (100%) [M + H]+.

10b
HPLC: tR1 = 9.1 min, tR2 = 9.2 min.
MS (MALDI): m/z = 2761.782 (100%) [M + H]+.

Note: For FAM-[Adp(OH)]8-NH2 (8b) and FAM-[Adp(OMe)]8-NH2 (9b), one constitutional isomer of the FAM labeled peptide could be isolated by semi-preparative HPLC (5-FAM or 6-FAM). In the case of FAM-[Adp(NMe2)]8-NH2 (10b), only a mixture of the 2 isomers (5/6-FAM) could be isolated in pure form (see Figure 3).

Desalting the Octapeptides; Typical Procedure

Desalting of 8a
In a 50 mL falcon tube, a solution of 8a in H2O (5 mL, double deionized) was incubated at r.t. for 10 min with Amberlyst A26 (HO form, 1.5 g). The mixture was then transferred to a fritted glass filter and washed 5 times with H2O (each 20 mL, double deionized). The filtrate was collected, frozen, and lyophilized in order to give desalted 8a. For control of fluorine content a 19F NMR spectrum was recorded; there was only a tiny little signal from TFA (δ = –75.60).

References

Figure 3 Analytical chromatograms at 220 nm of the isolated peptides 8–10. The two peaks for 10b are due to separation of the two constitutional isomers with 5/6-FAM groups; in the other cases these isomers are not separated.