Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

SUPPORTING MATERIALS

Synthesis of Protein Mimics with Non-Linear Backbone Topology by Combined Recombinant, Enzymatic and Chemical Synthesis (CRECS) Strategy

Stephan Pritz, Oliver Kraetke, Annerose Klose, Jana Klose, Sven Rothemund, Klaus Fechner, Michael Bienert, Michael Beyermann*

General:

The template T, Biotin-K[Z-GGGQK(Mhx-Ahx-GEGK(Dde)GEGKGEG)GEG]-amide, ECD2 (*cyclo*(CGVQLTVSPEVHQSNVAWSRLG)), ECD3 (*cyclo*(CGIGKLHYDNEKSWFGKRPGVYT DYG)), and ECD4 (*cyclo*(CGVNPGEDEVSRVVFIYFNSFG)) were synthesised using the Fmoc-strategy and a standard protocol for automated solid-phase synthesis (Applied Biosystems, 433a). Peptides were purified by preparative RP-HPLC (purity >95% according to analytical RP-HPLC with 220 nm uv detection) and their identity was confirmed by ESI-mass spectrometry.
Unless otherwise noted, mass accuracy of ESI-mass spectrometry was better than 60 ppm.

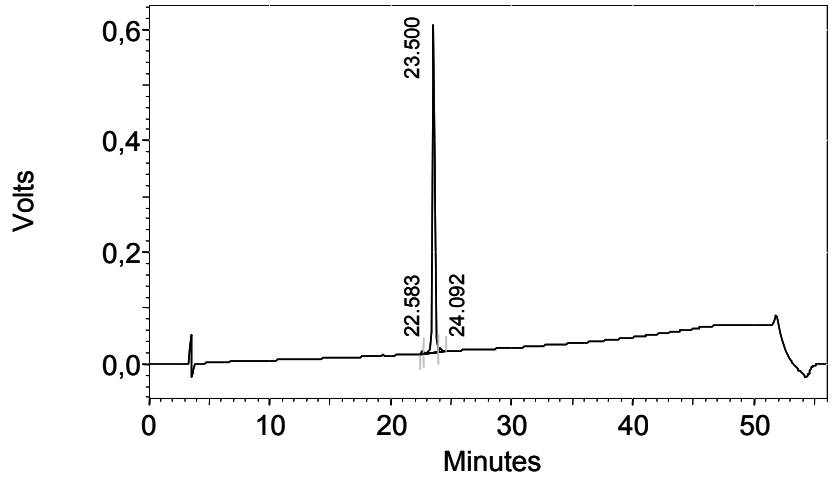
Preparation of the receptor loops by SPPS and native chemical ligation (exemplarily for ECD4):

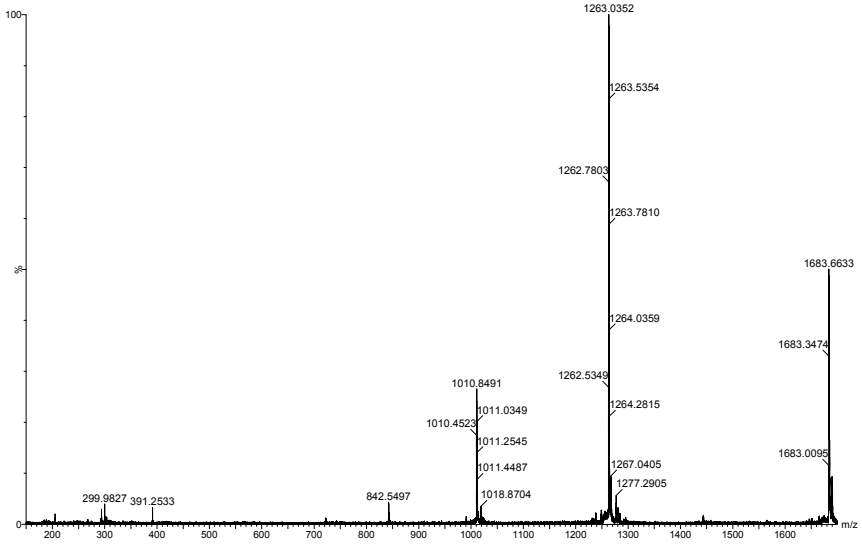
SPPS was carried out on Trt-Tentagel resin (Rapp-Polymere, Germany, 0.25 mmol/g, 1.5 g) using Fmoc-strategy (double couplings with 3 equi. of Fmoc-aa/ HBTU/ 6 equi. DIPEA). The N-terminal cysteine was coupled as BOC-Cys(Trt)-OH. The fully protected peptide was cleaved from the peptide-resin with 20 mL of HOAc/TFE/DCM (2/2/6) for 1 h at rt. After addition of n-hexane (100 mL), the solution was evaporated to dryness and the product was lyophilised twice from dioxane (yield: 1.09 g (94%)). Thiol ester was prepared according to a recently published procedure.[12] Briefly, 1.09 g (0.29 mmol) protected ECD4, 66.5 mg (0.435 mmol) HOBt, 167 mg (1 mmol) *p*-acetamidothiophenol (Aatp), and 68 µL (0.435 mmol) diisopropylcarbodiimide (DIPCDI) dissolved in 50 mL of DCM were reacted for 2 h at rt. After 2 h, 4 h, 6 h, and 8 h same quantities of DIPCDI and Aatp were again added and the reaction mixture was stirred overnight. The reaction mixture was then evaporated to dryness, the resulting product was dissolved in 20 mL of a mixture of water(5)/ phenol(5)/ tris-isopropylsilane(2)/ TFA(88) for deprotection (3 h at rt). After evaporation (to 5 mL), the thiol ester was precipitated with 250 mL of diethylether, washed with diethylether, ACN, diethylether and dried (yield: 770 mg). 50 mg of crude product gave 12.6 mg of the purified thiol ester by preparative HPLC (Vydac C-8, 300×25 mm, linear gradient 30-80% B in 70 min, A: 0.1 % TFA, B: 80% ACN-0.1% TFA). Cyclisation (NCL) was accomplished by dissolving 60 mg of the thiol ester and 60 mg TCEP in 30 mL of 0.5 M aqueous sodium bicarbonate and stirring for 1 h at rt.

After addition of 300 µL HCl_{conc}, the solution was evaporated, the precipitate was separated by centrifugation, washed with water, ACN and diethylether, and lyophilised from ACN/water (yield: 27 mg of ECD4).

Preparation of the three-loops-template construct T-ECD4-3-2-NH_2 (see Fig.2):

Thiol-maleimide ligation, step A: To 33 mg of template T (12.5 µmol) and 37 mg (15 µmol) of ECD4 dissolved in 0.4 mL DMSO, 1.6 mL water and 0.25 mL of 1 M $NaHCO_3$, pH 8.5, were added. After stirring for 1 h at rt, the product T-ECD4 was formed quantitatively and purified by RP-HPLC.

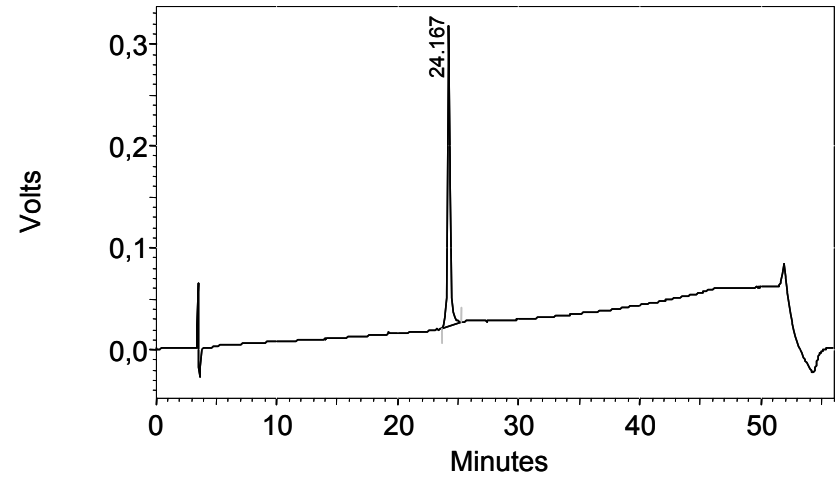


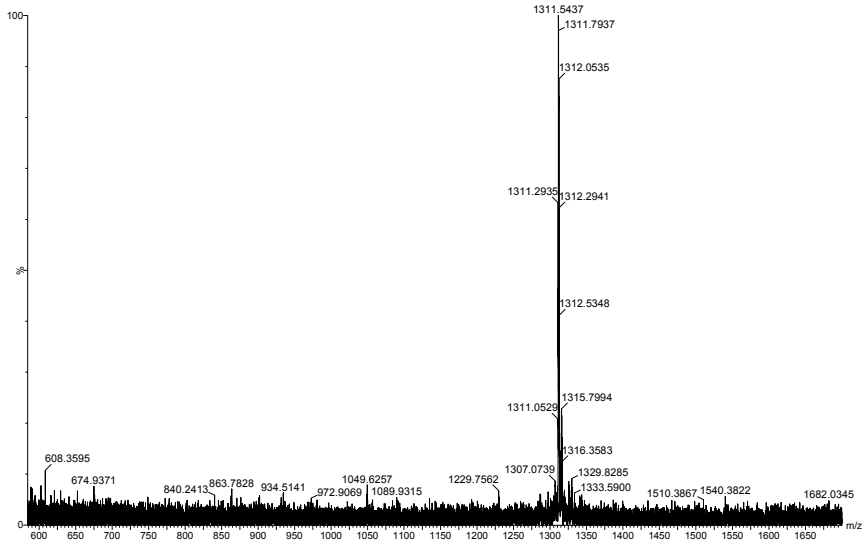


	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+3H]^{3+}$	1683.45	1683.34
$[M+4H]^{4+}$	1262.84	1262.78
$[M+5H]^{5+}$	1010.47	1010.45
$[M+6H]^{6+}$	842.22	842.24
M	5047.34	5047.2 ± 0.3

Figure 6: Characterisation of T-ECD4 by analytical HPLC (top) and ESI-MS (bottom).

Introduction of maleimidohexanoic acid, step B: To 61 mg (12 µmol) of T-ECD4 and 32 mg (0.1 mmol) Mhx-OSu dissolved in 3 mL DMSO, 0.5 mL water and 0.1 mL of 1 M $NaHCO_3$, pH 8.5, were added. After stirring for 1 h at rt, the product T-ECD4-Mhx was formed quantitatively and purified by RP-HPLC.

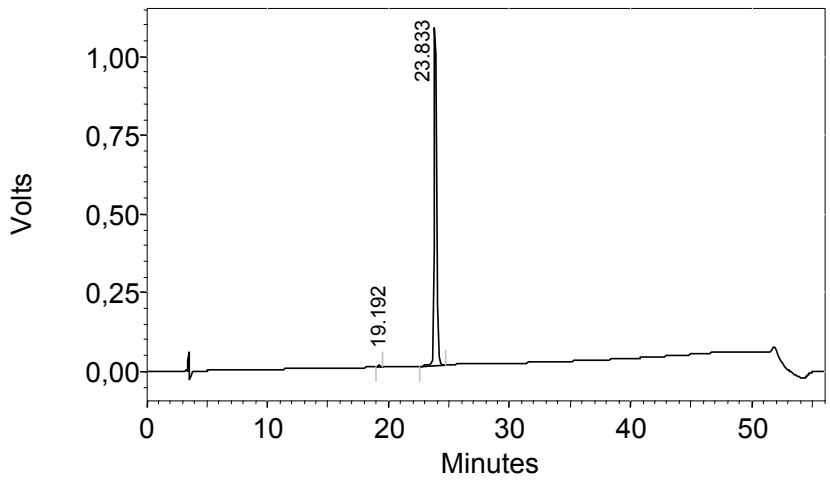


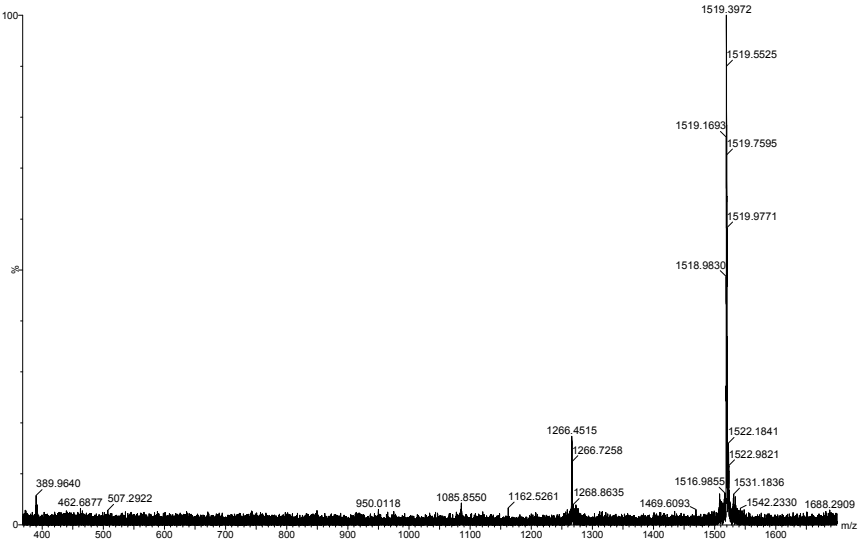


	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+4H]^{4+}$	1311.10	1311.07
M	5240.42	5240.3 ± 0.3

Figure 7: Characterisation of T-ECD4-Mhx by analytical HPLC (top) and ESI-MS (bottom).

Thiol-maleimide ligation, step C: To 40.4 mg (7.8 µmol) of T-ECD4-Mhx and 28.2 mg (12 µmol) of ECD2 dissolved in 2 mL DMSO, 0.25 mL water and 50 µL of 1 M $NaHCO_3$, pH 8.5, were added. After stirring for 1.5 h at rt, the product was formed quantitatively and purified by RP-HPLC.

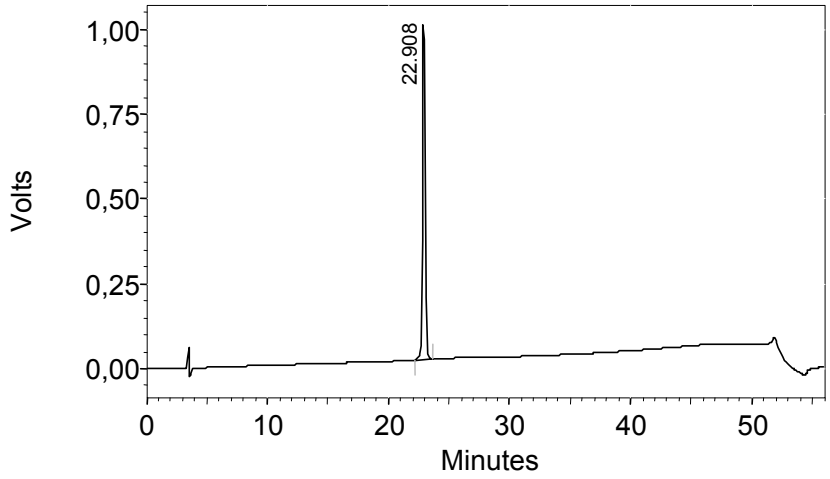


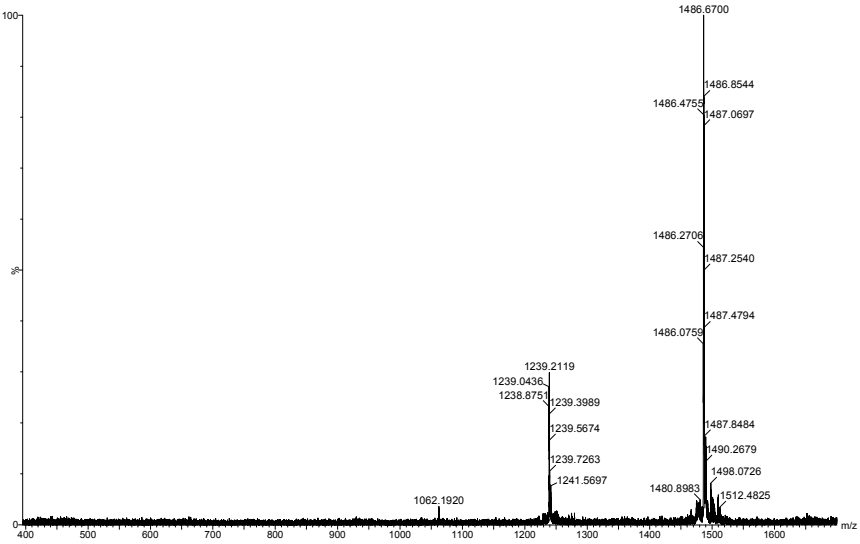


	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+5H]^{5+}$	1518.72	1518.80
$[M+6H]^{6+}$	1265.76	1265.85
M	7588.59	7589.0 ± 0.5

Figure 8: Characterisation of T-ECD4-2 by analytical HPLC (top) and ESI-MS (bottom).

Dde removal, step D: To 63.5 mg (8.36 µmol) of T-ECD4-2 dissolved in 50 mL of ACN/water (1/1), 0.25 % aqueous ammonia was added until pH 6.3 was reached, and the solution was lyophilised. The product was dissolved in 3.92 mL DMF and 80 µL hydrazine was added. After stirring for 10 min at rt, the deblocking was complete, 200 µL TFA were added and the product T-ECD4-2-NH_2 was purified by RP-HPLC.

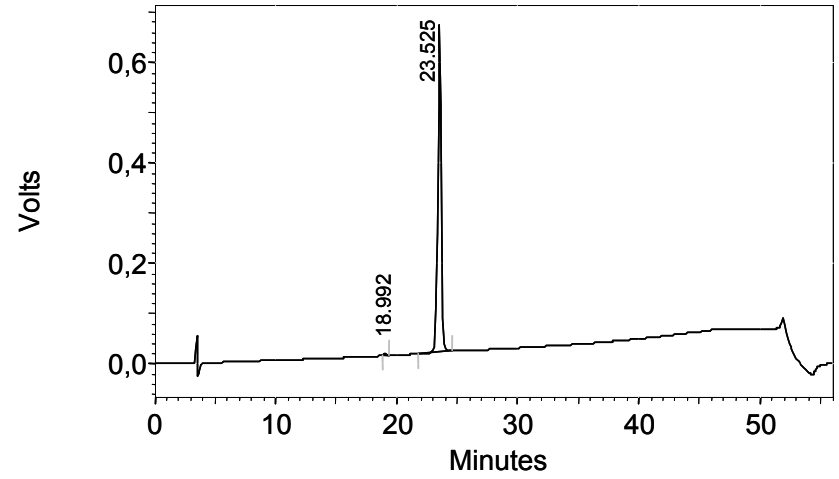


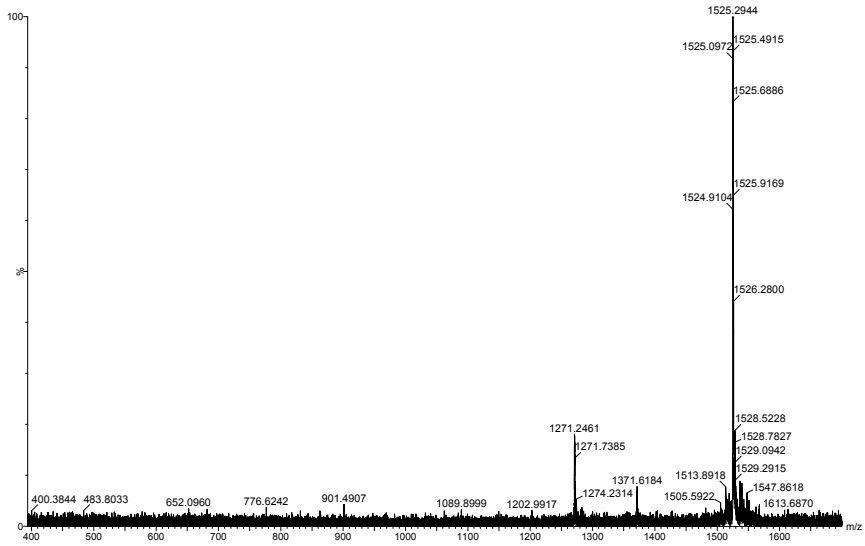


	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+5H]^{5+}$	1485.90	1485.87
$[M+6H]^{6+}$	1238.42	1238.40
$[M+7H]^{7+}$	1061.64	1061.62
M	7424.50	7424.3 ± 0.4

Figure 9: Characterisation of T-ECD4-2-NH_2 by analytical HPLC (top) and ESI-MS (bottom).

Introduction of maleimidohexanoic acid, step E: To 51 mg (6.9 µmol) of T-ECD4-2-NH_2 and 3.32 mg (0.01 mmol) Mhx-OSu dissolved in 3 mL DMSO, 500 µL water and 100 µL of 1 M $NaHCO_3$, pH 8.5, were added. After stirring for 90 min at rt, the product was formed quantitatively and purified by RP-HPLC.



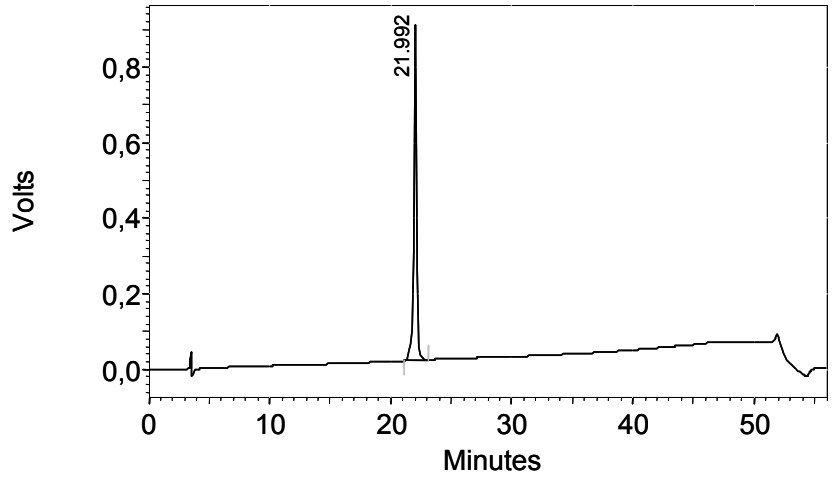


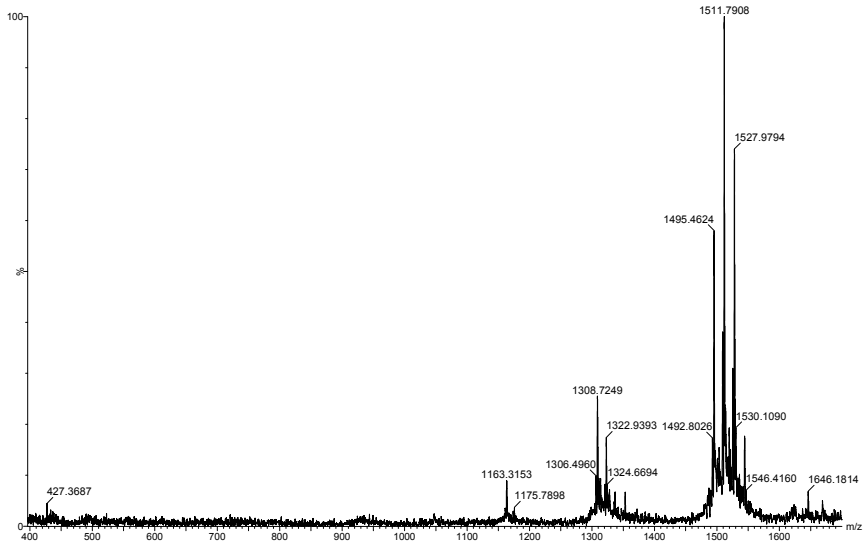
	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+5H]^{5+}$	1524.52	1524.66
$[M+6H]^{6+}$	1270.60	1270.51
M	7617.58	7617.6 ± 0.5

Figure 10: Characterisation of T-ECD4-2-Mhx by analytical HPLC (top) and ESI-MS (bottom).

Thiol-maleimide ligation, step F: To 41 mg (5.4 µmol) T-ECD4-2-Mhx and 26 mg (8.7 µmol) ECD3 dissolved in 3 mL DMSO, 0.5 mL water and 100 µL of 1 M $NaHCO_3$, pH 8.5, were added. After stirring for 1 h at rt, the ligation product was formed quantitatively and purified by RP-HPLC.

Removal of the Z-group, step G: 42.1 mg (3.98 µmol) T-ECD4-3-2 were dissolved at 0°C in the deblocking mixture (1.8 mL thioanisol, 0.9 mL ethanedithiol, 18 mL TFA, 1.8 mL TFMSA), and the reaction mixture was stirred for 10 min at 0°C. After evaporation, the product T-ECD4-3-2-NH_2 was precipitated with diethylether and subsequently purified by RP-HPLC.





	m / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+7H]^{7+}$	1495.50	1495.46
$[M+7H+TFA]^{7+}$	1511.79	1511.79
$[M+7H+2TFA]^{7+}$	1528.07	1527.98
$[M+8H]^{8+}$	1308.68	1308.72
$[M+8H+TFA]^{8+}$	1322.94	1322.94
$[M+9H]^{9+}$	1163.39	1163.32
M	10461.47	10461.2 ± 0.6

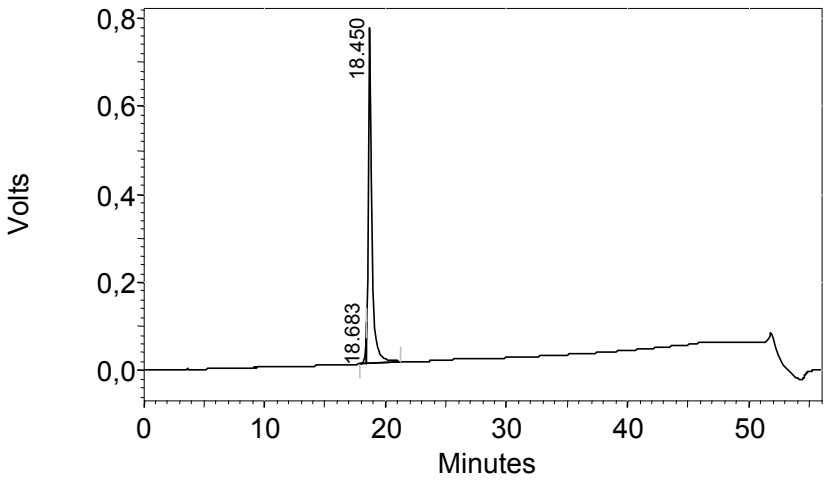
Figure 11: Characterisation of T-ECD4-3-2-NH_2 by analytical HPLC (top) and ESI-MS (bottom).

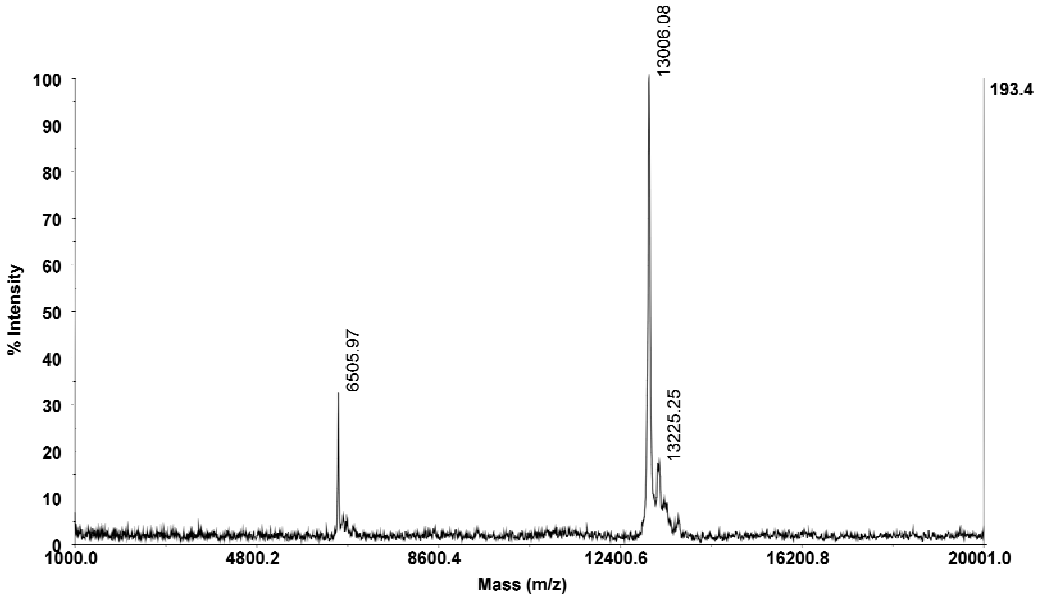
Recombinant synthesis of ECD1-LPKTGGRR, step X:

As described previously by Klose et al.[4], *E. coli* BL21 (DE3), transformed with the plasmid pET-CRF_1-N-terminus with sortase A recognition motif-LPKTGGRR, was grown in LB medium with antibiotics at 37°C until an optical density of OD_{600nm}= 0.5 – 0.7 was reached. The expression was started by addition of IPTG. After incubation for 3-4 h at 37°C, an expression yield of approximately 30-35 mg ECD1-LPKTGGRR per liter of medium was obtained. The cells were then harvested by centrifugation and subjected to lysis. Inclusion bodies, containing ECD1-LPKTGGRR, were separated by centrifugation and stored at –20°C.

Folding & purification of ECD1-LPKTGGRR, step Y:

The inclusion bodies were resuspended in 5 M GuHCl, 20 mM Tris, pH 7.5 and sonicated. After centrifugation, ECD1-LPKTGGRR was reduced with tris(2-carboxyethyl) phosphine for 2 h at rt and the solution was then dialysed against 5 M GuHCl/ 1 mM EDTA, pH 3, overnight at 8°C. Insoluble material was removed by centrifugation. The supernatant was diluted to a final protein concentration of 0.5 mg/ml and the pH readjusted to 7.5. Folding of ECD1-LPKTGGRR was achieved by dialysis against a buffer containing 1 M arginine, 100 mM TrizmaBase, 1 mM EDTA, 1 mM gluthation-reduced (GSH), and 1 mM GSSG, pH 7.5 at 10°C. The folded protein was purified by preparative RP-HPLC (Vydac C4, 250×10 mm, 5 µm, 300 Å, with a linear gradient 20-60% B within 70 min, A: 0.1 % TFA, B: 80% aqueous ACN-0.1% TFA, 10 ml/min) and lyophilised.





MALDI-MS	*m* / *z* calcd. (monoisotopic)	*m* / *z* found (monoisotopic)
$[M+H]^+$	13006.6	13006.1
$[M+2H]^{2+}$	6503.8	6506.0
M	13005.6	13007 ± 4

Figure 12: Characterisation of ECD1-LPKTGGRR by analytical HPLC (top) and MALDI-MS (bottom).

Treatment of the product with α-chymotrypsin and LC-MS analysis of the peptide fragments showed the identity and the correct disulfide pattern of the ECD1-LPKTGGRR.

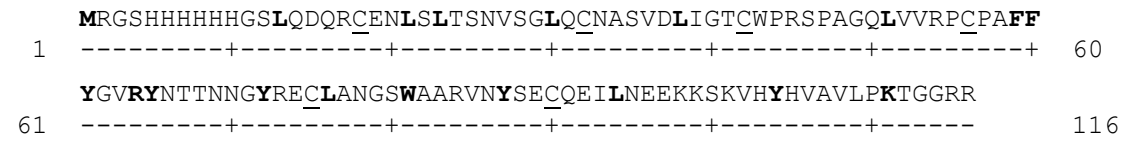


Figure 13: Sequence of ECD1-LPKTGGRR with N-terminal His_6-Tag and C-terminal sortase-recognition motif. Cysteine residues are underlined, cleavage sites found by chymotryptic digest are marked in bold (cleavages occurred C-terminally of marked amino acids).

Table 1: Linear fragments found for the chymotryptic digest of in vitro folded ECD1-LPKTGGRR.

fragment	sequence	monoisotopic mass calcd.	monoisotopic mass found	deviation in ppm
2-13	RGSHHHHHHGSL	1397.656	1397.645	-8
22-23	SL	218.127	218.128	6
22-30	SLTSNVSGL	876.455	876.455	-1
24-30	TSNVSGL	676.339	676.336	-4
61-64	YGVR	493.265	493.265	0
61-65	YGVRY	656.328	656.325	-5
62-65	GVRY	493.265	493.264	-2
66-72	NTTNNGY	782.320	782.318	-2
77-81	ANGSW	533.223	533.224	2
82-87	AARVNY	692.361	692.356	-7
95-104	NEEKKSKVHY	1260.646	1260.644	-2
105-111	HVAVLPK	762.475	762.469	-8
105-116	HVAVLPKTGGRR	1289.768	1289.759	-7
112-116	TGGRR	545.303	545.301	-4

All expected cysteine-free fragments being larger than one amino acid were found. Additionally, some fragments with a missed cleavage site could be identified (e.g. fragment 22-30). The protein was also cleaved at two residues nontypical for chymotrypsin (Arg^{64} and Lys^{111}), possibly due to contamination of the enzyme with trypsin. No linear fragment containing cysteine was found, i.e. all cysteines were involved in disulfide bonding.

For determination of the disulfide pattern, the masses of all possible combinations of disulfides (homodimers included) were calculated and the respective masses searched for in the LC-MS run. Only three disulfide-bridged fragments were found corresponding to the disulfide pattern (C1-C3, C2-C5, C4-C6) reported for the soluble CRF_1-NT.[4]

Table 2. Disulfide-bridged fragments found for the chymotryptic digest of in vitro folded ECD1-LPKTGGRR.

Fragment	Cysteine pattern	Monoisotopic mass calcd.	Monoisotopic mass found	Deviation in ppm
14-21 and 39-51	Cys^{18}-Cys^{42} (C1-C3)	2387.111	2387.075	-15
31-38 and 73-76	Cys^{32}-Cys^{75} (C2-C5)	1365.601	1365.596	-4
52-59 and 88-94	Cys^{56}-Cys^{90} (C4-C6)	1705.816	1705.802	-9

Sortase-mediated ligation of ECD1-LPKTGGRR and T-ECD4-3-2-NH_2:

A mixture of ECD1-LPKTGGRR (238 µM), T-ECD4-3-2-NH_2 (238 µM) and sortase A (5.7 µM) dissolved in Tris-buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM $CaCl_2$, 0.95 M urea, pH 7.5) was shaken at 23°C. After two days a product yield of ~ 60% (as determined by HPLC) was attained which did not further increase. The crude product was purified by RP-HPLC (Prontosil C-18, 250×8 mm, C18, 300 Å, 5 µm, eluent A: 0.1% TFA, eluent B: 80% ACN-0.1% TFA, linear gradient 30-80% B in 70 min, 2 ml/min). Lyophilisation of the product fractions yielded 1.678 mg of the desired product (36.4% yield) which was characterised by HPLC and MS (Fig. 4).