Electronic Supplementary Information

Identification of protease substrates by combinatorial profiling on tentagel beads

Jacob Kofoed^{‡a} and Jean-Louis Reymond^{*a}

^{*a*} Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, CH-3012 Berne, Switzerland. Fax: +41 31 631 80 57; Tel: +41 31 631 43 25; E-mail: jeanlouis.reymond@ioc.unibe.ch; Website: http://www.dcb.unibe.ch/groups/reymond.

[‡] Present adress: Diabetes Protein and Peptide Chemistry, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark.

Experimental section

General. Reagents were purchased in the highest quality available from Fluka, Sigma, Bachem, Novabiochem, NeoMPS or Aldrich. All solvents used in reactions were bought in p.a. quality or distilled and dried prior to use. Solvents for extractions were distilled from technical quality. Sensitive reactions were carried out under nitrogen or argon, the glassware being heated under high vacuum. Preparative RP-HPLC (flow rate 100 mL min⁻¹) was performed with a Waters Delta Prep 4000 system with a Waters Prepak Cartridge (500 g) as column and Waters 486 Tunable Absorbance Detector. Semi-preparative RP-HPLC (flow rate 4 mL min⁻¹) was performed with a Water 510 Pump operated with a Waters Automated Gradient Controller and Jasco UV-2075 Plus Detector on a Vydac 218 TP (1.0 cm × 25 cm) column. Analytical RP-HPLC (flow rate 4 mL min⁻¹) was performed on Waters 600E systems with Waters Atlantis (4.6 mm × 100 mm, dC18, 5 mm) column, UV detection with Waters 996 photodiode array detector). Data recording and processing was done with Waters Empower2 software. Eluents for all systems were: A: water and 0.1% TFA; D: acetonitrile, water and TFA (3/2/0.1%). Analytes were quantified using external standards. MS and HRMS analyses were provided by the mass spectrometry service of the Department of Chemistry and Biochemistry, University of Berne. ¹H and ¹³C NMR spectra were recorded on Bruker AC 300 (300 MHz) and DRX 500 or Avance 500 (500 MHz) instruments. Chemical shifts (d) are given in ppm referring to solvent residual peak, coupling constants (J) in Hertz (Hz). Solid phase peptide chemistry was performed in polypropylene syringes fitted with a polyethylene frit and a teflon stopcock and stopper.

Coupling of the Fmoc-protected amino acids. The resin was washed and swollen inside the reactor with DCM (2×5 mL) and DMF (1×5 mL). The appropriate resin (library: TentaGel HL (0.65 mmol g⁻¹), single sequences: resin was acylated with 2.5 equivalents of *N*-Fmoc amino acid in the presence of 2.5 equivalents of DIPCDI and 2.5 equivalents of HOBt in DMF. After 1 hr the resin was washed ($3 \times each$) with DMF, DCM, and MeOH, and controlled with the TNBS (trinitrobenzenesulfonic acid) or chloranil test followed by acetylation.

Cleavage of the Fmoc protecting group. The Fmoc protecting group was removed with 5 mL of a solution of DMF-piperidine (4:1, v/v) for 10 min. After filtration, the procedure was repeated and then washed $(3 \times \text{each})$ with DMF, DCM and MeOH.

N-Acetylation. The resin was acetylated with a solution of acetic acid anhydride-DCM (1:1, v/v) for 10 min. After filtration, the procedure was repeated and then washed (3 × each) with DMF, DCM and MeOH.

TFA cleavage. The cleavage was carried out using TFA-H₂O-TIS (triisopropylsilane) as a (95:2.5:2.5, v/v) solution for 6 hrs.

Resin mixing and splitting. The resin was suspended in DMF-DCM (2:1, v/v), and mixed *via* nitrogen bubbling for 15 min, and then distributed in four equal portions.

4-Carboxybenzaldehyde 2-{ethyl-[4-(4{nitrophenylazo)phenyl]-amino}ethyl ester (3). A solution of 4-carboxybenzaldehyde (1) (100 mg, 0.67 mmol), Disperse Red 1 (2) (210 mg, 0.67 mmol) and DMAP (81 mg, 0.67 mmol) in DCM was cooled in an ice-bath before EDC (141 mg, 0.74 mmol) was added. The reaction was allowed to heat up to rt overnight. The reaction mixture was extracted with brine and evaporated to dryness. The red residue was flash chromatographed (n-hexane-ethyl acetate (2:1, v/v)) to yield 3 as a red waxy solid (30 mg, 10%). Rf = 0.21 (n-hexane-ethyl acetate (2:1, v/v)). IR (neat) $\tilde{V} = 2973$, 1718, 1703, 1596, 1585, 1512, 1382, 1333, 1310, 1266, 1198, 1131, 1100, 1071, 1015, 856, 824, 756 cm-1. 1H NMR (300 MHz, CDCl3) = 10.01 (s, 1H), 8.27 (d, 2H, J = 7.16 Hz), 8.04 (d, 2H, J = 7.16

6.60 Hz), 7.76 (m, 6H), 6.81 (d, 2H, J = 9.42 Hz), 4.49 (t, 2H, J = 6.62 Hz), 3.72 (t, 2H, J = 6.62 Hz), 3.51 (q, 2H, J = 6.97 Hz), 1.12 (t, 3H, J = 6.97 Hz) ppm. 13C NMR (300 MHz, CDCl3) = 191.4, 169.9, 155.9, 153.5, 150.7, 142.9, 141.9, 135.0, 130.3, 129.6, 124.7, 122.6, 111.6, 62.4, 58.5, 42.3, 12.4 ppm. ESI MS(+): calcd for C24H24N4O5+ 447.17, found 447.19.

On-bead proteolytic assays. 50 mg library resin was swollen overnight in 100 mM bis-tris buffer at the indicated pH (see below). The swelling solution was removed by filtration. 1 mL of a solution of the protease of interest (1 mg/ml)1 in 100 mM bis-tris buffer was added and resin was shaken for 18 hrs and washed extensively with bis-tris buffer, DMSO, DMF, MeOH, DCM, MeOH, DMF and finally with bis-tris buffer again ($3 \times \text{each}$). Then 1 mL of 10 mM solution of **3** in THF/H₂O/AcOH (90:5:5, v/v/v) was added and the resin was shaken for 1 hrs. Afterwards NaBH₃CN (6 mg, 0.09 mmmol) was added and the resin was shaken for 1 hrs. The resin was washed extensively with bis-tris buffer, DMSO, DMF, MeOH, DMF and finally with bis-tris buffer, DMSO, DMF, MeOH, DCM, MeOH, DMF and finally with bis-tris buffer again ($3 \times \text{each}$) and a suspension of the resin in DMF was transferred to a silica gel plate and the beads were observed under a microscope. Single red colored beads were transferred via a syringe needle to amino acid analysis vials. The following commercial enzymes were used: Trypsin from pig pancreas 1645 U mg⁻¹ (Fluka 82495); Subtilisin from *Bacillus licheniformis* 10.5 U mg⁻¹ (Fluka 85968); α -Chymotrypsin from bovine pancreas 74.6 U mg⁻¹ (Fluka 27270); Pepsin from porcine stomach mocusa 4500 U mg⁻¹ (SIGMA p-6887).

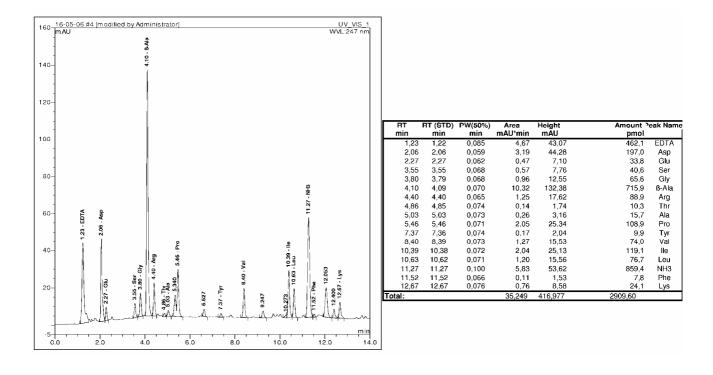
Peptide Synthesis. The peptides were re-synthesized on Rink amide NovaGelTM (highloading PEG-PS resin 0.63 mmol g⁻¹) using the same coupling conditions as described for the linear library synthesis. After TFA cleavage the peptides were precipitated with methyl *tert*butyl ether and dissolved in water-acetonitrile mixture.

Solution phase proteolytic cleavage. Peptides 10-17 were conditioned as 2 mM stock solutions in water. The proteolysis was started by adding 5 μ L of a freshly prepared 1 mg mL⁻¹ stock solution2 of the protease in 100 mM bis-tris buffer to a mixture of 50 μ L peptide stock solution and 45 μ L 100 mM bis-tris buffer. The assay concentration under these conditions was 1 mM for the peptide substrate and 50 μ g mL⁻¹ for the protease. The following pH were used: Trypsin: pH 8, Subtilisin: pH 6.5, α -Chymotrypsin: pH 8, Pepsin: pH 4. The reactions were analyzed by RP-HPLC. Flow rate: 3.0 mL min⁻¹, 100% A, 0%D to 100% D, 0% 0 in 15

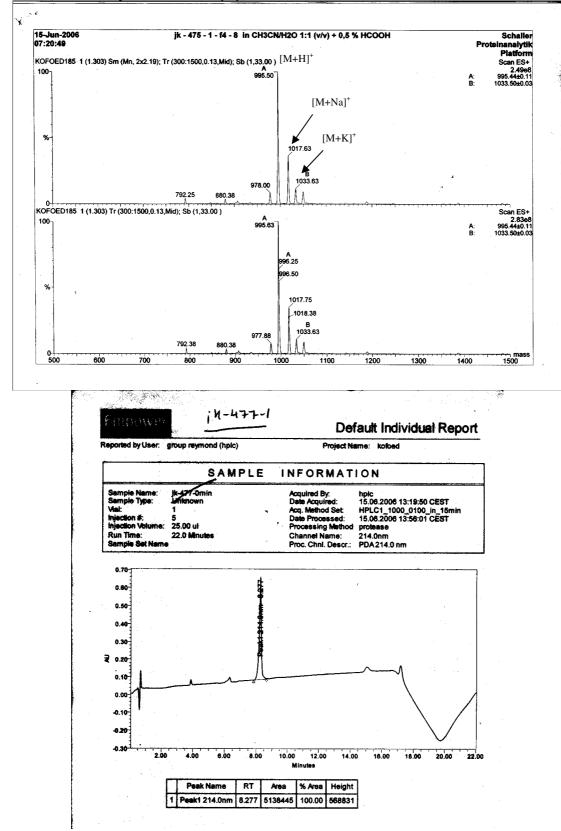
min. The crude proteolytic mixture was also subjected to ESI MS(+) analysis. The peptide fragments were identified using a computer generated database of possible mass fragments. For the Meprin-alpha the provided stock solution was diluted appropriately.

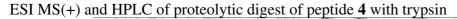
Bead analysis. The beads were transferred via a sterile syringe needle to amino acid analysis vial and hydrolyzed with aqueous HCl (6 M) at 110°C for 22 hrs. Quantitative AAA was performed by HPLC after derivatization with phenyl isothiocyanate. The detection limit for such analysis is usually 50 pmoles. We used tentagel beads of 90 micrometer diameter with 0.63 mmol/g loading. Sequences were assigned from AAA using the TAGSFREE program available from the author's website (http://www.dcb.unibe.ch/groups/reymond/). The user provides an input file <name>.csv describing the library as follows: Each line lists a different amino acid (in 3 letter code) followed by the usage vector as a series of 1's and 0's for each position up to the planned number of variable positions all separated by semicolons. The program handles "unique pair" designs or any other design. An excel file listing unique pairs is provided as a help for library design. Running the TAGSFREE program on the input file <name>.csv generates an output file <name>LIB.txt in which all possible AAA and the corresponding sequences are listed in the order of AAA and an output file <name>STAT.txt which contains the analysis of the generated library. The AAA of any bead is written in form of a series of integers (0, 1, 2) indicating the relative amounts of each building block in the order used in the input file. Decoding is carried out by typing an AAA in the search window of a text editor capable of handling the large library file. The search function automatically goes to the line(s) containing the AAA, which shows the corresponding sequence(s).

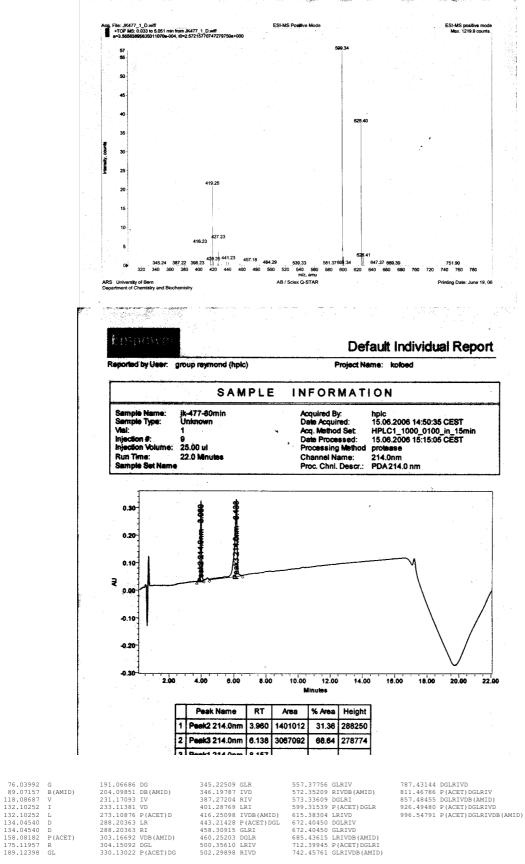
Ac-Pro-Asp-Gly-Leu-Arg-Ile-Val-Asp-Bla-NH₂ (4). From NovaGelTM(160 mg, 0.63 mmol g⁻¹), 4 was obtained as colorless foamy solid after preparative HPLC purification (35 mg, 31%, as TFA-salt); anal. RP-HPLC (80% A, 20% D to 20% A, 80% D in 15 min): $t_{\rm R} = 6.68$ min; ESI MS(+): calcd for C₄₃H₇₃N₁₃O₁₄: 995.54, found: 995.50.



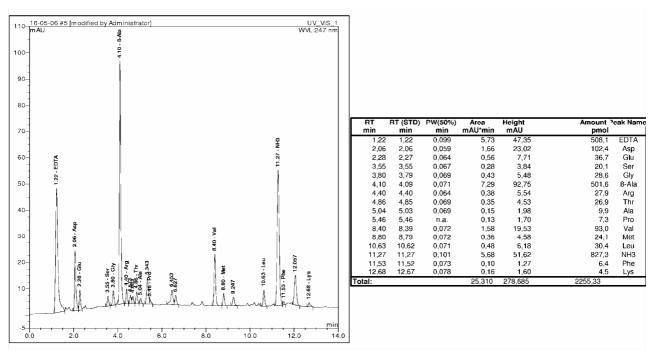
ESI-MS(+) and analytical HPLC of peptide 4

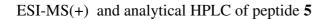


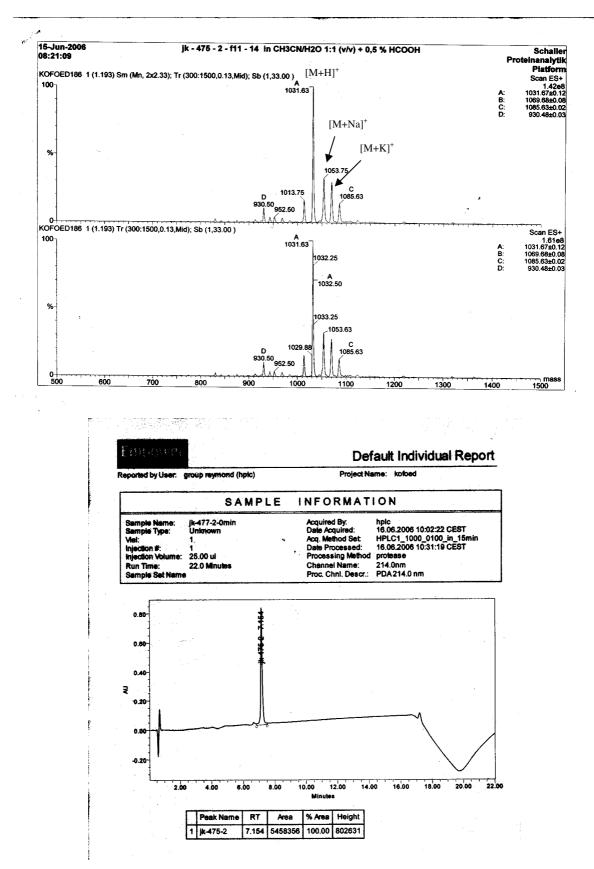


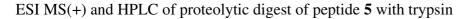


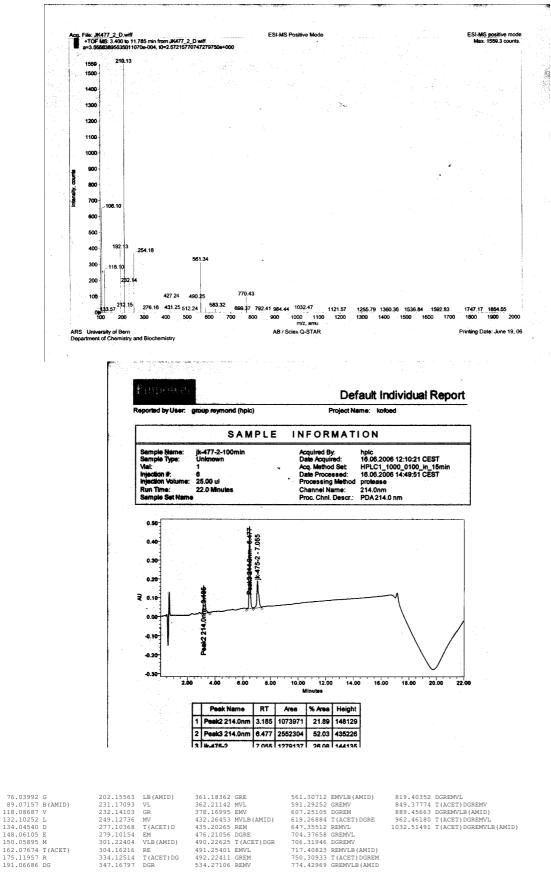
Ac-Thr-Asp-Gly-Arg-Glu-Met-Val-Leu-Bla-NH₂ (5). From NovaGelTM(160 mg, 0.63 mmol g⁻¹), 5 was obtained as colorless foamy solid after preparative HPLC purification (27 mg, 31%, as TFA-salt); anal. RP-HPLC (80% A, 20% D to 20% A, 80% D in 15 min): $t_{\rm R} = 5.04$ min; ESI MS(+): calcd for C₄₂H₇₃N₁₃O₁₅S: 1031.51, found: 1031.63.



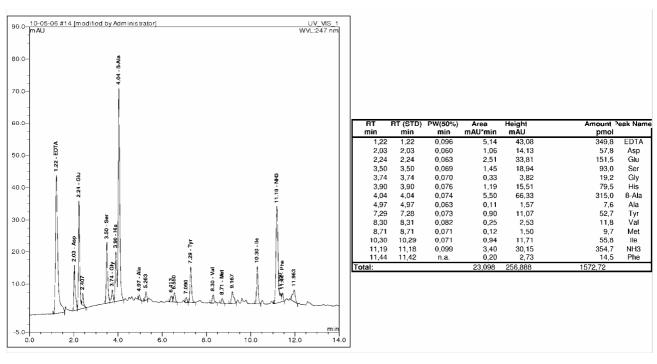


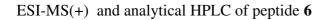


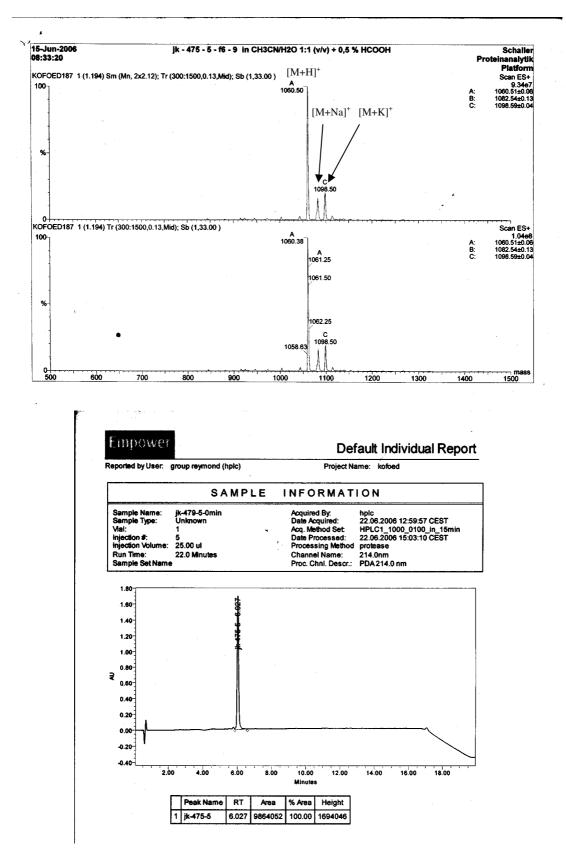




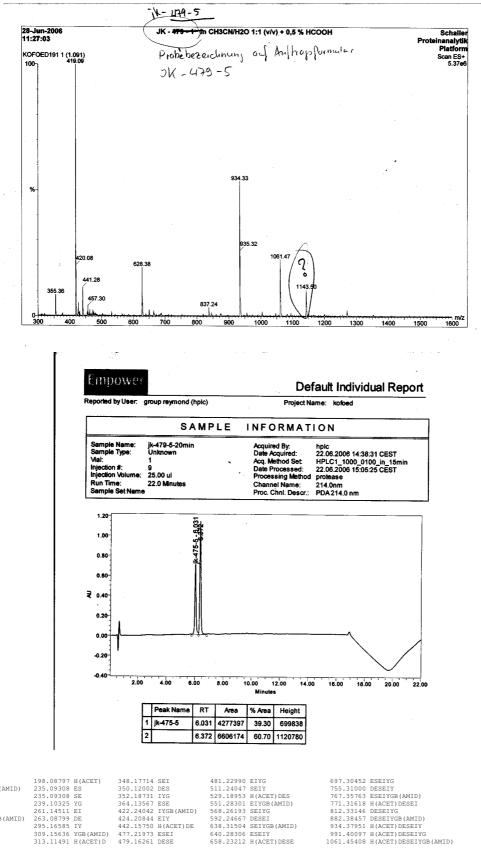
Ac-His-Asp-Glu-Ser-Glu-Ile-Tyr-Gly-Bla-NH₂ (6). From NovaGelTM (160 mg, 0.63 mmol g⁻¹), 6 was obtained as colorless foamy solid after preparative HPLC purification (51 mg, 43%, as TFA-salt); anal. RP-HPLC (80% A, 20% D to 20% A, 80% D in 15 min): $t_{\rm R} = 2.90$ min; ESI MS(+): calcd for C₄₅H₆₄N₁₂O₁₈: 1060.45, found: 1060.50.







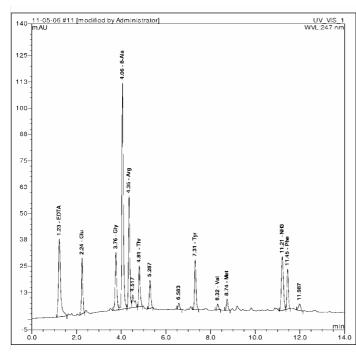
ESI MS(+) and HPLC of proteolytic digest of peptide 6 with α -Chymotrypsin



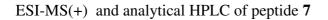
ĩą.

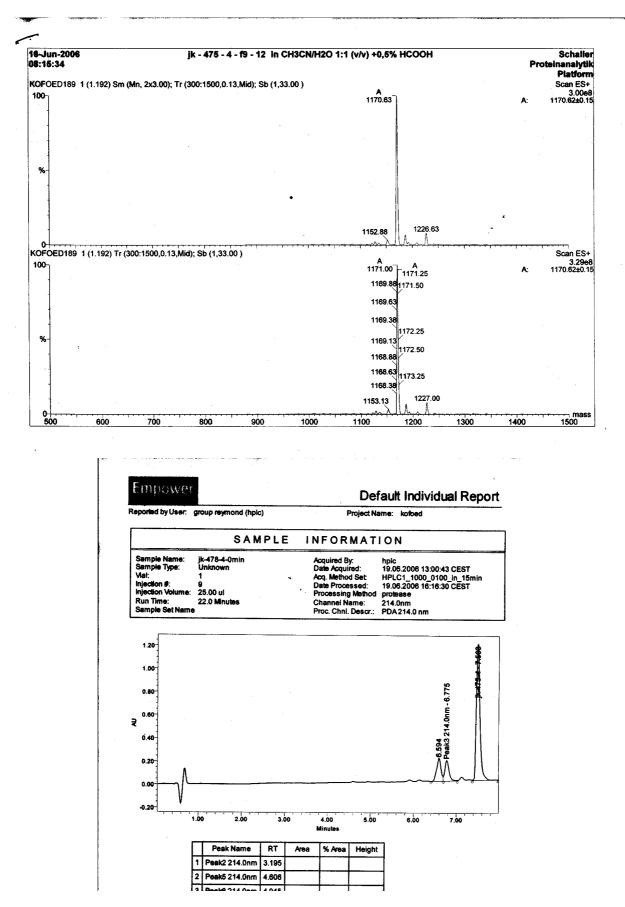
4

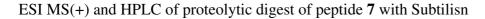
Ac-Thr-Phe-Glu-Arg-Arg-Met-Tyr-Gly-Bla-NH₂ (7). From NovaGelTM (160 mg, 0.63 mmol g⁻¹), 7 was obtained as colorless foamy solid after preparative HPLC purification (31 mg, 22%, as TFA-salt); anal. RP-HPLC (80% A, 20% D to 20% A, 80% D in 15 min): $t_{\rm R}$ = 4.45 min; ESI MS(+): calcd for C₅₁H₇₈N₁₆O₁₄S: 1170.56, found: 1170.63.

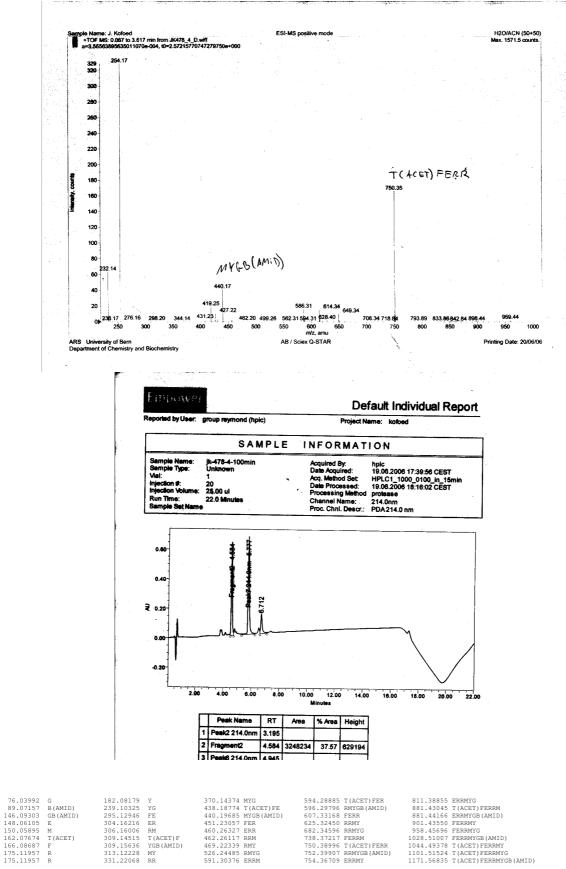


BT	RT (STD)	PW(50%)	Area	Height	A	⁵ eak Name
min	min	min	mAU*min	mAU	pmol	-еак мате
1,23	1,22	0,091	4,13	36,82	267,0	EDTA
2,24	2,25	0,064	1,81	26,25	134,7	Glu
3,76	3,76	0,071	2,25	27,23	148,7	Gly
4,06	4,06	0,070	8,55	107,34	573,4	B-Ala
4,35	4,35	0,066	3,86	53,02	275,8	Arg
4,81	4,81	0,069	1,49	19,24	116,8	Thr
7,31	7,31	0,069	1,78	23,44	120,7	Tyr
	0.04	0,080	0,24	2,77	13,9	Val
8,32	8,34				01.0	Mot
		0,071	0,40	5,21	31,6	Met
8,32	8,74		0,40 2,72	5,21 25,34	31,6 336,6	NH3
8,32 8,74	8,74 11,21	0,071				

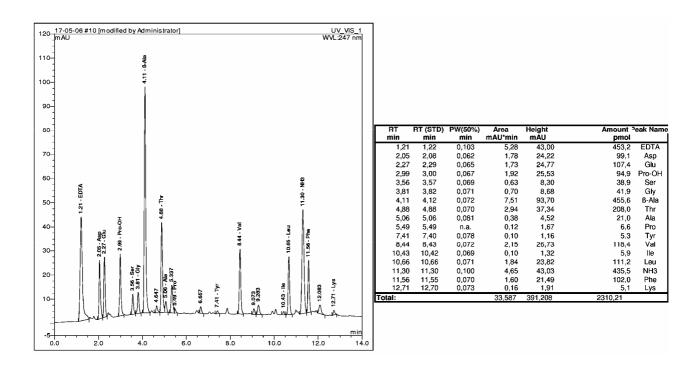


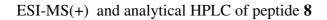


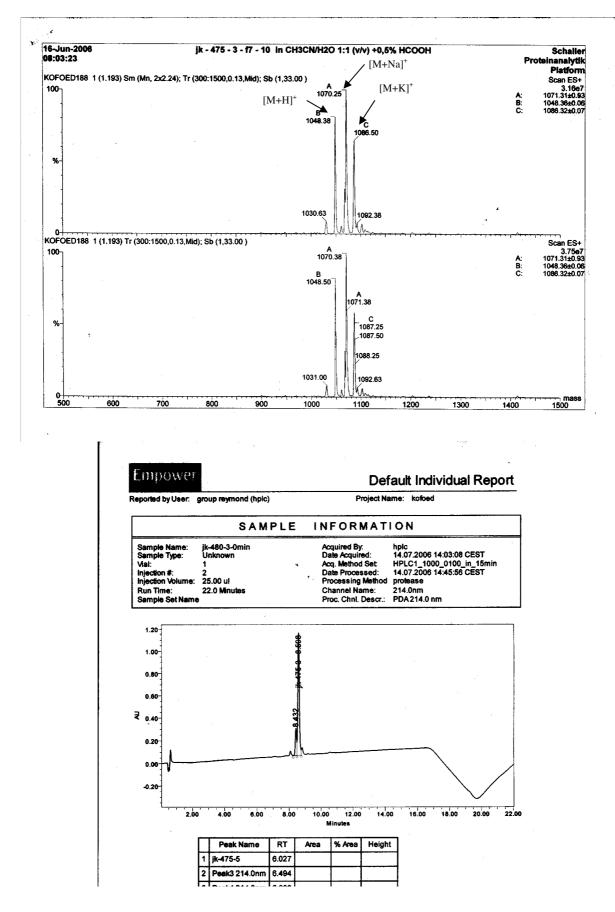


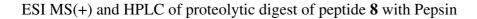


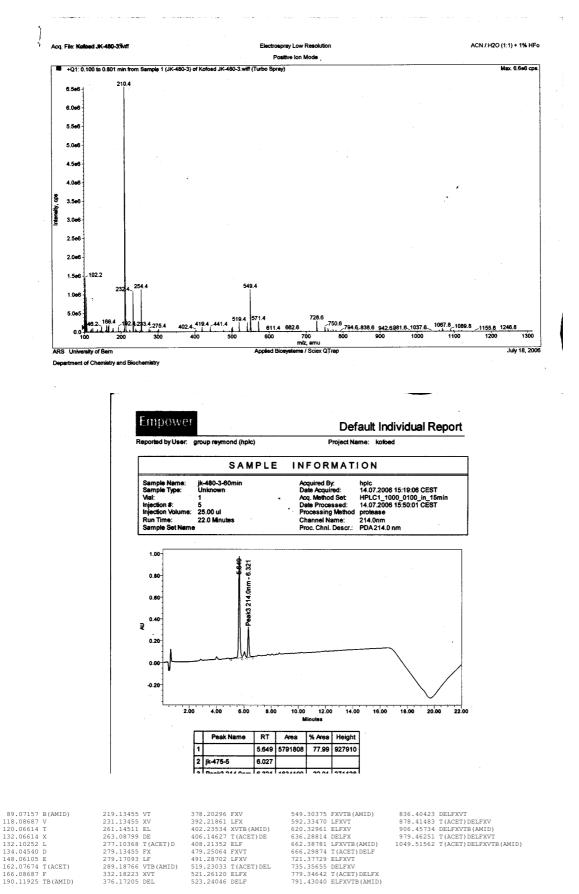
Ac-Thr-Asp-Glu-Leu-Phe-Hyp-Val-Thr-Bla-NH₂ (8). From NovaGelTM (160 mg, 0.63 mmol g⁻¹), 8 was obtained as colorless foamy solid after preparative HPLC purification (53 mg, 50%, as TFA-salt); anal. RP-HPLC (80% A, 20% D to 20% A, 80% D in 15 min): $t_{\rm R} = 6.41$ min; ESI MS(+): calcd for C₄₇H₇₂N₁₀O₁₇S: 1048.51, found: 1048.38.





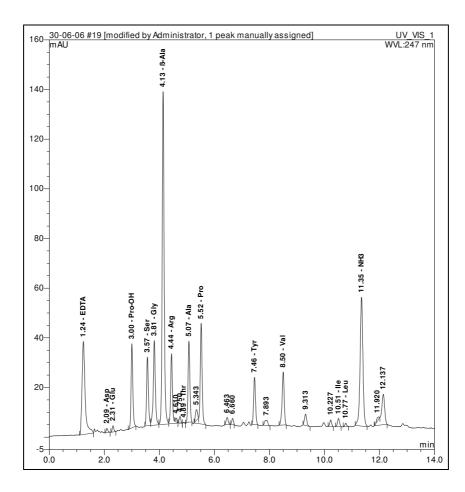






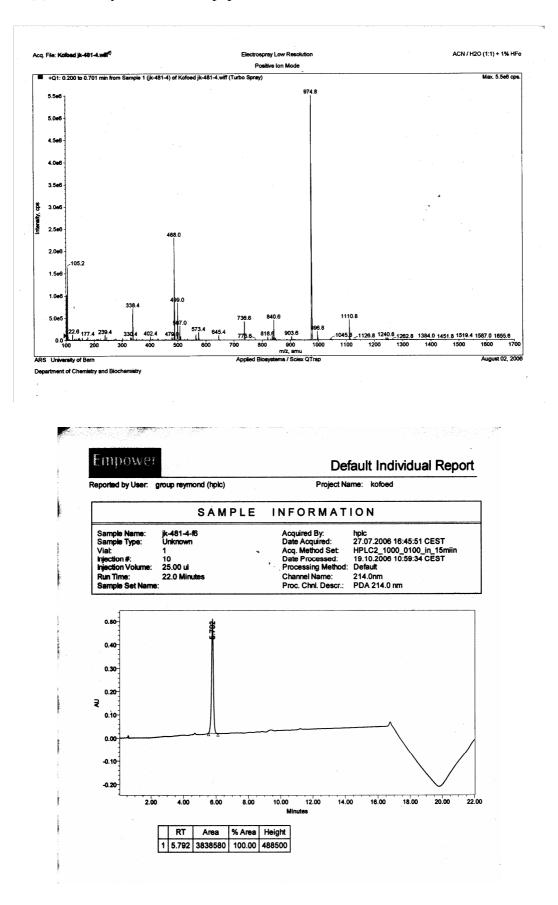
190.11925 TB(AMID)

Ac-Pro-Val-Tyr-Ser-Arg-Hyp-Ala-Gly-Bla-NH₂ (9). From NovaGelTM (160 mg, 0.63 mmol g⁻¹), 9 was obtained as colorless foamy solid after preparative HPLC purification (59 mg, 53%, as TFA-salt); anal. RP-HPLC (100% A, 0% D to 0% A, 100% D in 15 min): $t_{\rm R} = 5.79$ min; ESI MS(+): calcd for C₄₃H₆₇N₁₃O₁₃: [M+H]+ 974.5, found: 974.8.





ESI-MS(+) and analytical HPLC of peptide 9



Ac-Hyp-Val-Tyr-Arg-Pro-Ser-Ala-Gly-Bla-NH₂ (10). From NovaGelTM (160 mg, 0.63 mmol g⁻¹), 10 was obtained as colorless foamy solid after preparative HPLC purification (74 mg, 67%, as TFA-salt); anal. RP-HPLC (100% A, 0% D to 0% A, 100% D in 15 min): $t_{\rm R} = 5.76$ min; ESI MS(+): calcd for C₄₃H₆₇N₁₃O₁₃: [M+H]+ 974.5, found: 974.8.

ESI-MS(+) and analytical HPLC of peptide 10

