Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supplementary Information

Design of a Substrate-tailored Peptiligase Variant for the Efficient Synthesis of Thymosin-α_1

Marcel Schmidt,[a,b] Ana Toplak,[a] Henriëtte J. Rozeboom,[c] Hein J. Wijma,[c] Peter J.L.M. Quaedflieg,[a] Jan H. van Maarseveen,[b] Dick B. Janssen[c] and Timo Nuijens[a]*

[a] EnzyPep B.V., Brightlands Campus, Urmonderbaan 22, 6167 RD Geleen, The Netherlands. Phone: +31-464760675; E-mail: timo@enzypep.com.

[b] Van 't Hoff Institute of Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.

[c] Groningen Biomolecular Science and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.

Contents

1 Materials and Methods

If not denoted otherwise, chemicals were obtained from the following commercial sources in the highest grade purity available and used without further purification: Iris Biotech GmbH (Marktredwitz, Germany), Sigma Aldrich (St. Louis, MO, USA), Merck Millipore (Darmstadt, Germany), Biosolve (Valkenswaard, The Netherlands).

Water used was purified using a MilliQ desalting system (Merck Millipore, Darmstadt, Germany).

1.1 Peptide Purification

Crude peptide products were purified using preparative reversed-phase HPLC. Preparative HPLC was carried out on a Varian PrepStar solvent system configured with a Varian ProStar 330 PDA detector (Varian, Palo Alto, CA, USA).

The purification process was monitored at a wavelength of λ= 220 nm. Products were purified using a self-packed Luna Prep C18 column (10 µm, 200x 51 mm, Phenomenex, Torrance, CA, USA) eluting with a water acetonitrile gradient, with a flow rate of 45 mL/min. As mobile phase a binary mixture of A (MQ-water + 0.05 vol% TFA) and B (ACN + 0.05 vol% TFA) was used by default. The gradient was chosen individually for each peptide. Samples were prepared by dissolving the crude peptide in ACN/ MQ-water and filtration through a 0.22 µm filter unit.

Ostensible product fractions were analyzed by analytical HPLC-MS (1.2). Product containing fractions were pooled according to their purity, frozen in liquid N_2 and lyophilized (-90 °C). Products were stored at -20 °C for further use.

1.2 HPLC-MS Analysis

The peptide purity was assessed using an Agilent 1260 Infinity HPLC system coupled with an Agilent 6130 quadrupole mass spectrometer (Agilent, Santa Clara, CA, USA) to determine the peptide mass. Separation was performed using ReProSil-Pur C18 5µm, 4.6 x 250 mm (Dr. Maisch, Ammerbuch, Germany) column, eluting with 0.05% MSA in a water acetonitrile gradient, with a flow rate of 1 mL min^{-1}. As mobile phase a binary mixture of A (water + 0.05 vol% MSA) and B (ACN + 0.05 vol% MSA) was used. A linear gradient from 5 to 98 % B in 21 minutes, followed by isocratic 95% B for 5 min was used by default. For optimal results an appropriate gradient was chosen for each sample individually (see Table S 1). The purity of peptides was determined

by automatically integrating product and impurity peaks of the relevant HPLC spectrum (λ= 220 nm).

HPLC-MS-MS analysis was performed by DSM Resolve (Geleen, The Netherlands).

Table S 1. Individual conditions for HPLC-MS analysis of peptide segments.

Peptide	**Gradient**
Ac-Thm(1-14)-OCam-Leu-OH	12 to 30 vol% ACN/ MQ-water (v/v; 0.05 vol% MSA) in 21 min; flow 1 mL min^{-1}; 40 °C column temperature
H-Thm(15-28)-OH	5 to 14 vol% ACN/ MQ-water (v/v; 0.05 vol% MSA) in 30 min; flow 1.5 mL min^{-1}; 20 °C column temperature
Ac-Thm(1-28)-OH*	12 to 30 vol% ACN/ MQ-water (v/v; 0.05 vol% MSA) in 21 min; flow 1 mL min^{-1}; 40 °C column temperature

** Method also used for thymoligase library screenings.*

2 X-ray Crystal Structure Determination.

Thymoligase was further purified after the His-tag purification by gel filtration using a Superdex 200 HR10/30 column (GE Healthcare), equilibrated with 20 mM Hepes buffer, pH 7.5, containing 150 mM NaCl. The enzyme eluted at a molecular mass of 20 kDa. Thymoligase fractions were pooled and concentrated to 17.5 mg^{-1} mL^{-1} using a Vivaspin-10K filter unit (Sartorius). Dynamic light scattering experiments were performed using a DynaPro Nanostar instrument (Wyatt Technology Corporation, Santa Barbara, CA, USA) at 20 °C. Dynamic light scattering data were processed and analyzed using Dynamics software and an apparent molecular mass of 41kDa was deduced with a polydispersity of 12%.

Crystallization trials were performed in 96-well MRC2 plates (Molecular Dimensions Ltd., Newmarket, UK), using a Mosquito crystallization robot (TTP Labtech Ltd. Melbourn, UK) with commercially available screening matrices (PACT, JCSG+ and Wizard I&II, from Molecular Dimensions). Droplets containing reservoir solution (75-125 nL) and protein solution (125-75 nL) were incubated against 50 µL of each reservoir solution at 21°C. Plate-like thymoligase crystals were obtained from 1.0 M ammonium sulfate, 0.1 M Bis-Tris pH 5.5 and 1% PEG 3350.

Prior to data collection, single crystals were harvested and flash-cooled in liquid nitrogen with addition of 25% glycerol as cryoprotectant. A native dataset to 2.25 Å was collected on beamline P11 at Petra III (EMBL, DESY, Hamburg, Germany) at 100 K. Intensity data were processed using XDS[1] and the CCP4 package.[2] The space group was $P2_1$, with unit cell dimensions of a = 47.5, b = 40.1, c = 64.3 Å and β = 106.0°. With one monomer of 27.4 kDa in the asymmetric unit, the V_M

is 2.1 Å^3 Da^{-13}, with a calculated solvent content of 43%. A summary of data collection statistics is given in Table S 2.

Molecular replacement was performed using PHASER[4] with the thymoligase model obtained by YASARA. ARP/wARP[5] was used for automatic building, and the model was refined using REFMAC5[6]. COOT[7] was used for manual rebuilding and map inspection. In Fo–Fc maps, clear electron density was present for a nonapeptide in the active site which was derived from the propeptide of thymoligase.

The quality of the model was analyzed using MolProbity.[8] Figures were prepared using PYMOL (Schrödinger LLC). Atomic coordinates and experimental structure factor amplitudes for thymoligase have been deposited in the RCSB Protein Data Bank under accession code 5OX2.

Table S 2. Crystallographic data and refinement statistics. Values in parentheses are for the highest resolution shell.

Data collection	
Spacegroup	$P2_1$
Unit cell dimensions (Å, °)	a = 47.5, b = 40.1, c = 64.3, β = 106.0
Resolution range (Å)	45.6 - 2.24
N° total measurements	36626 (3300)
N° unique reflections	10355 (950)
R_{pim} (%)	7.1 (35.0)
Completeness (%)	91.6 (92.1)
Average I/σ	7.8 (2.7)
Multiplicity	3.5 (3.5)
Refinement	
Contents of A.U.	
protein	2 chains, residues 1 – 266 and residues 1 - 9
waters	52
other	4 sulfate ions
R/ R_{free}	18.8 / 23.0
Geometry	
r.m.s.d. bonds (Å)	0.009
r.m.s.d. angles (°)	1.4
Ramachandran plot (%) (favored/allowed/disallowed)	97.8 / 1.5 / 0.7

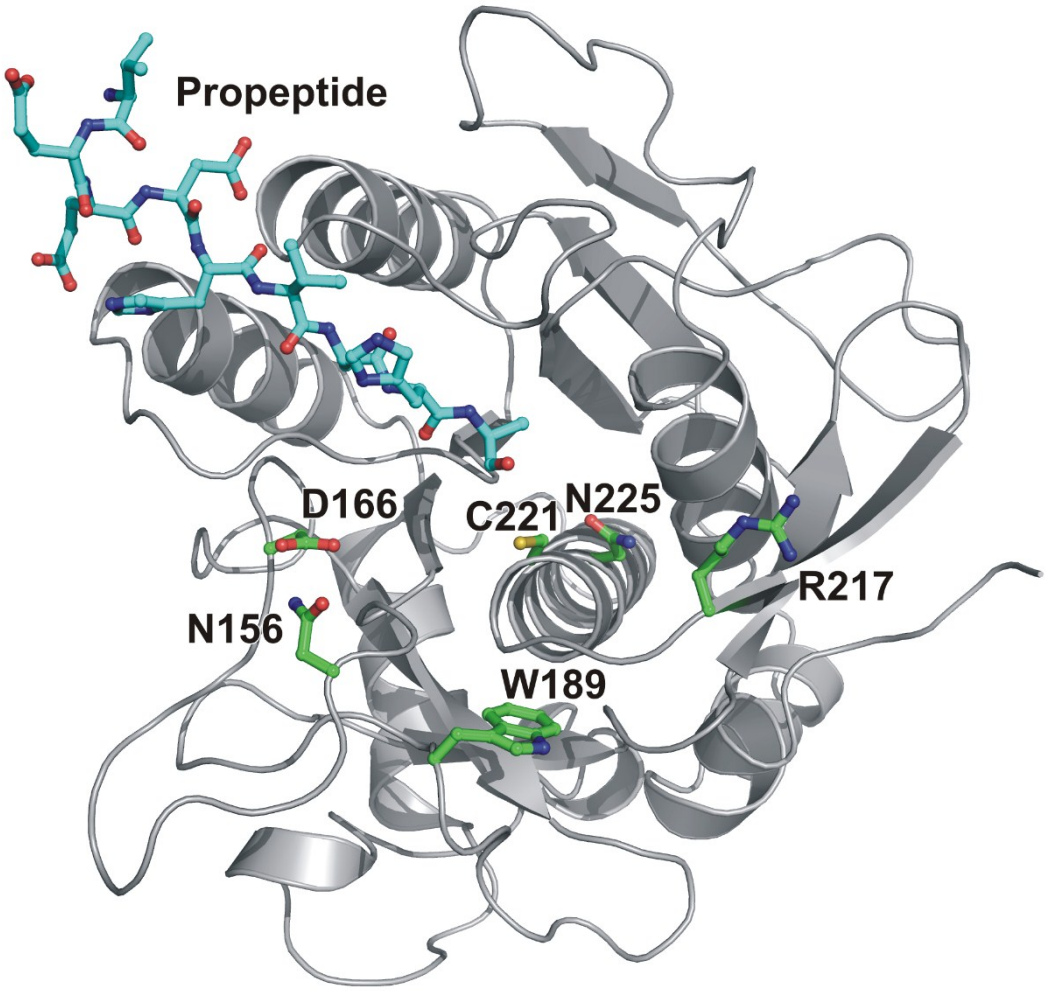

Figure S 1. Cartoon representation of thymoligase (pdb code: 5OX2) colored in gray. N156 and D166 at S1, the catalytic residue Cys221, N225 and R217 at S1', and W189 at S2' are shown in green sticks. The propeptide (VEEDHVAHA) in the substrate-binding region is shown in cyan sticks.

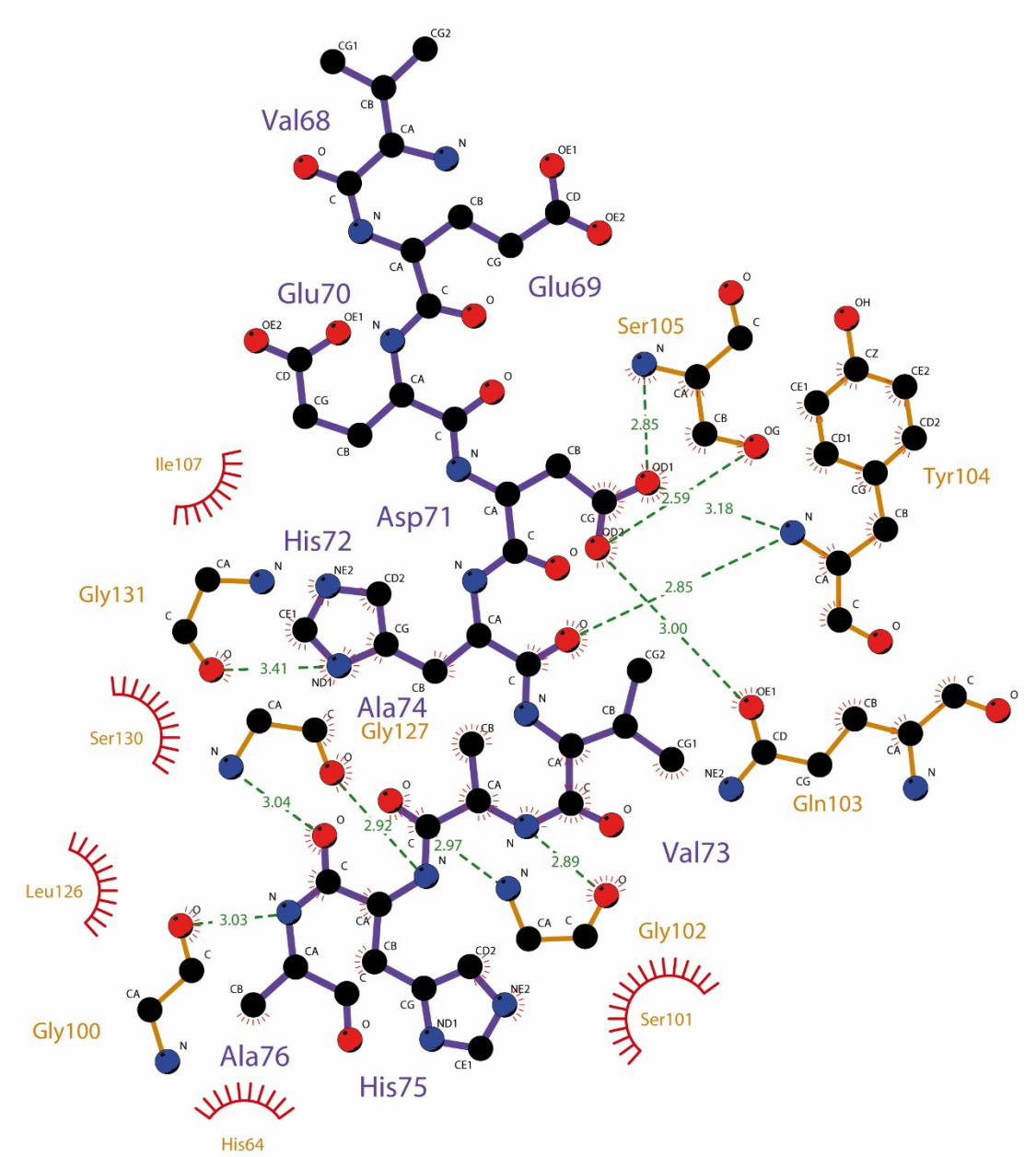

Figure S 2. A LigPlot+ diagram[9] of the peptide binding site of thymoligase with orange bonds and residue names and the nonapeptide (VEEDHVAHA) (residues 68'-76', chain P) with purple bonds and residue names. The green dashed lines, with distances, represent the hydrogen bonds between nonapeptide and thymoligase. Hydrophobic contacts made with the nonapeptide are indicated by the spoked arcs.

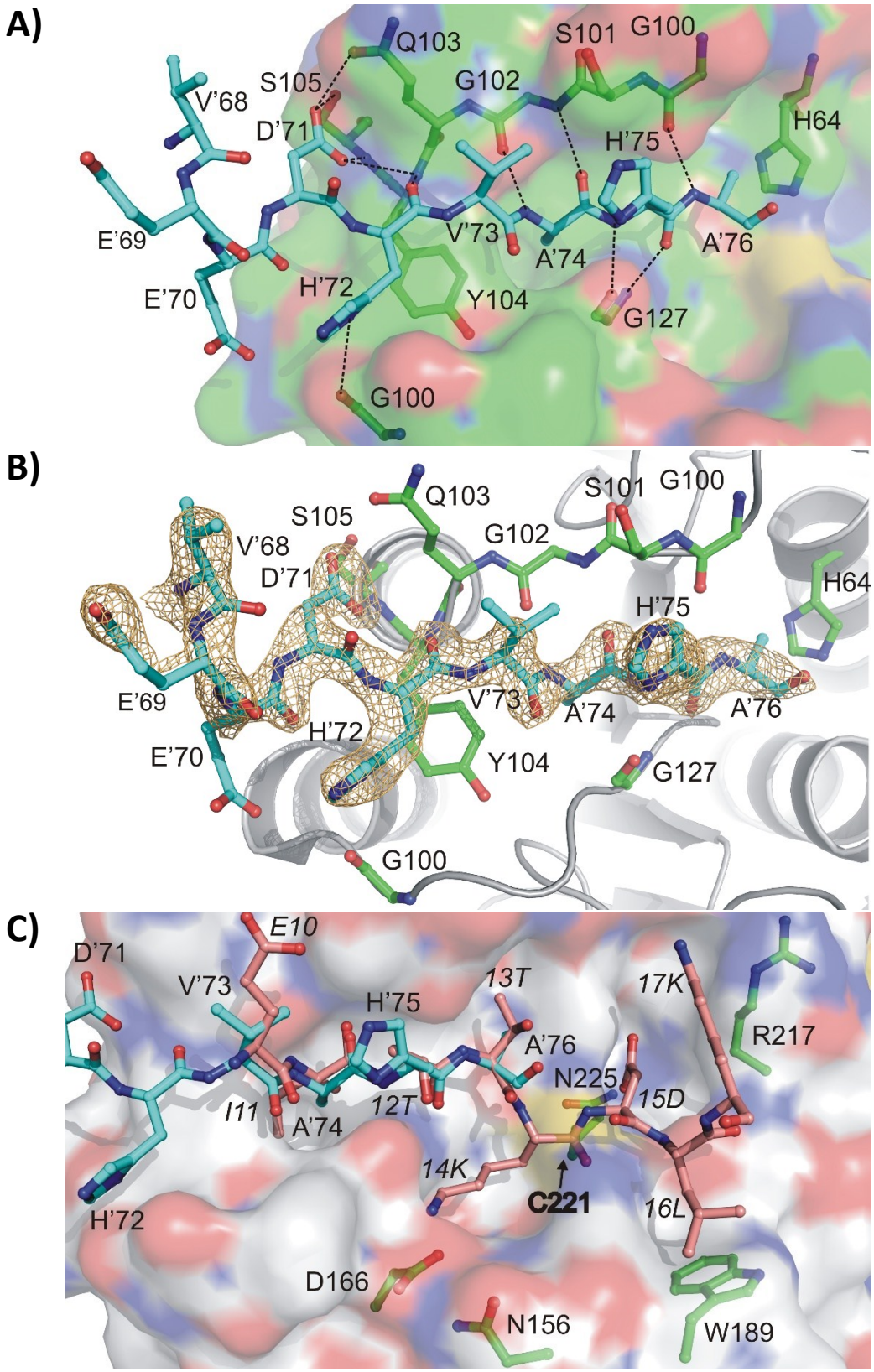

Figure S 3. Peptide binding sites in thymoligase. A) The residues names with prime belong to the nonapeptide in cyan sticks and the others are the hydrogen bonded residues from the protein in green. Hydrogen bond interactions between thymoligase and nonapeptide are indicated with black dashed lines. B) Fo-Fc omit electron density for the nonapeptide in cyan sticks is shown at 3σ in light orange. C) Modeling of thymosin-α_1. N156 and D166, the catalytic residue C221, N225, R217, and W189 are shown in green sticks. Residues from thymosin-α_1 *E10-17K* are shown in salmon sticks and italic letters. The nonapeptide is shown in cyan.

3 SDS PAGE Gels of Peptiligase Library Productions

Below SDS PAGE gels from the production and purification of different peptiligase variants are shown. SDS-PAGE was performed using a NuPAGE® electrophoresis system (Thermo Scientific, Waltham, MA, USA) following the manufacturer´s instructions. NuPAGE BisTris Mini Gels (4-12 %) were used under reducing conditions and samples were prepared according to the manufacturers instructions; 10 µL of each sample were applied to the gel.

As a protein size marker Novex Sharp Pre-Stained Protein Standard (Thermo Scientific, Waltham, MA, USA) was used.

Abbreviations:

M **=** Novex Sharp Pre-Stained Protein Standard Marker

Med = Supernatant of bacterial cell culture medium after centrifugation

FT **=** Flow-Through during IMAC purification (non column bound material)

E = Eluate of IMAC purification

3.1 Peptiligase Library I

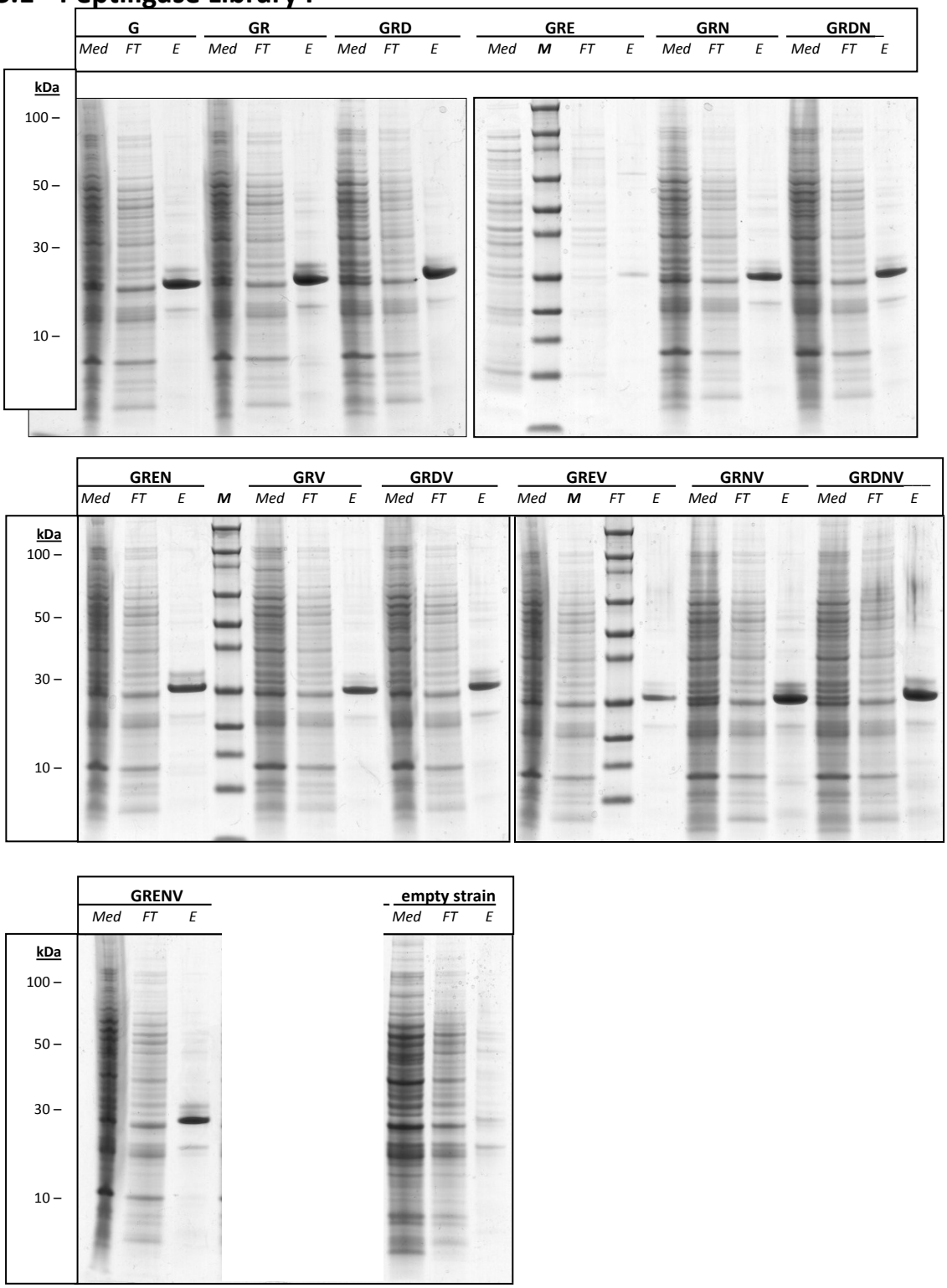

Figure S 4. SDS gels from the production and purification of the peptiligase library I. Mutations were introduced into pBE-S peptiligase: *G* – M222G, *R* – L217R, *D* – S166D, *E* – S166E, *N - A225N, V* – I107V, *P* – M222P, *H* – L217H. For each peptiligase variant the following samples were taken during the production and purification process and analyzed via SDS-PAGE: *Med* – bacterial culture supernatant, *FT* – flow-through after incubation of the culture supernatant with the IMAC column material, *E* – elution fraction of each peptiligase variant, respectively. SDS-PAGE was performed using a NuPAGE® electrophoresis system (Thermo Scientific, Waltham, MA, USA) following the manufacturer´s instructions. Premade NuPAGE BisTris Mini Gels (4-12 %) were used under reducing conditions and samples were prepared according to the manufacturer's instructions; 10 µL of each sample were applied to the gel. As a protein size marker Novex Sharp Pre-Stained Protein Standard (Thermo Scientific, Waltham, MA, USA) was used.

3.2 Peptiligase Library II

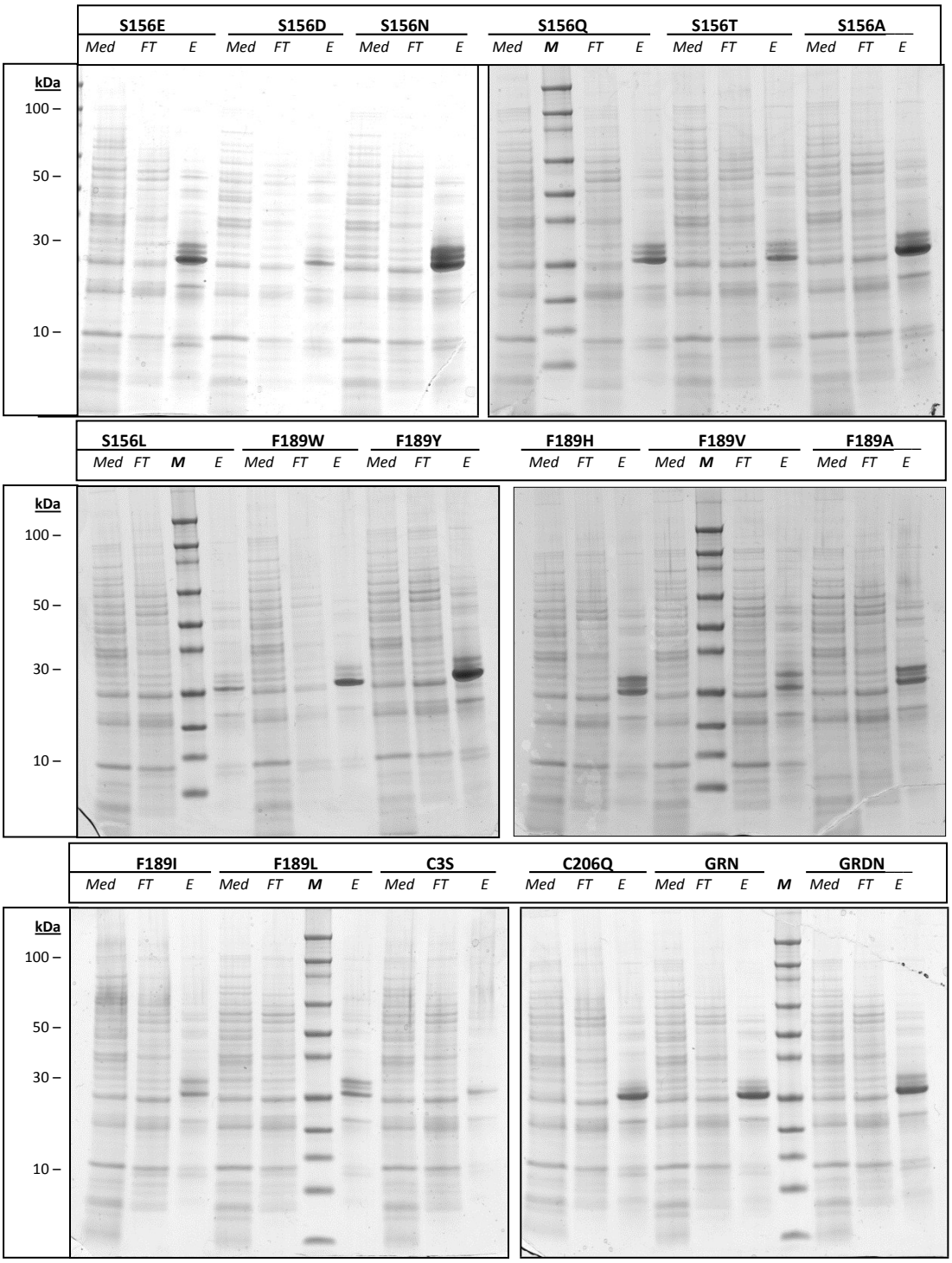

Figure S 5. SDS gels from the production and purification of the peptiligase library II. Mutations were incorporated into pBE-S peptiligase GRDN: *G* – M222G, *R* – L217R, *D* – S166D, *N* - A225N. The mutations introduced are given. For each peptiligase variant the following samples were taken during the production and purification process and analyzed via SDS-PAGE: *Med* – bacterial culture supernatant, *FT* – flow-through after incubation of the culture supernatant with the IMAC column material, *E* – elution fraction of each thymoligase variant, respectively. SDS-PAGE was performed using a NuPAGE® electrophoresis system (Thermo Scientific, Waltham, MA, USA) following the manufacturer´s instructions. Premade NuPAGE BisTris Mini Gels (4-12 %) were used under reducing conditions and samples were prepared according to the manufacturer's instructions; 10 µL of each sample were applied to the gel. As a protein size marker Novex Sharp Pre-Stained Protein Standard (Thermo Scientific, Waltham, MA, USA) was used.

3.3 Peptiligase Library III

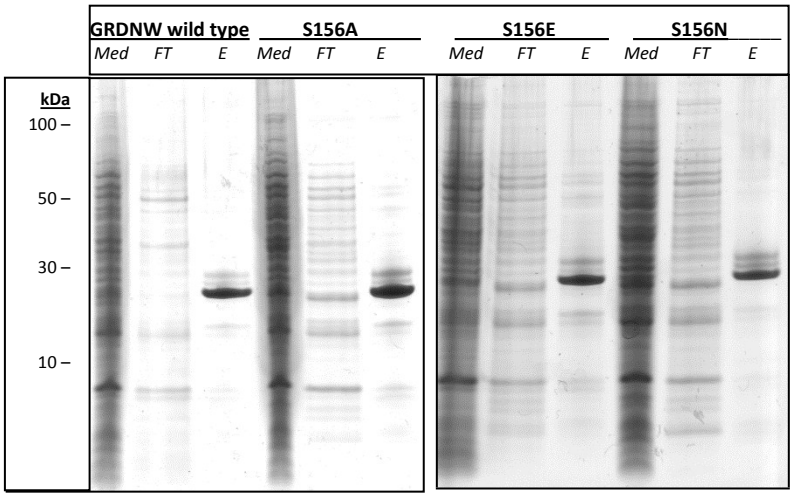

Figure S 6. SDS gels from the production and purification of the peptiligase library IIII. Mutations were incorporated into pBE-S peptiligase GRDNW: *G* – M222G, *R* – L217R, *D* – S166D, *N* - A225N, *and W* – F189W. The mutations introduced are given. For each peptiligase variant the following samples were taken during the production and purification process and analyzed via SDS-PAGE: *Med* – bacterial culture supernatant, *FT* – flow-through after incubation of the culture supernatant with the IMAC column material, *E* – elution fraction of each thymoligase variant, respectively. SDS-PAGE was performed using a NuPAGE® electrophoresis system (Thermo Scientific, Waltham, MA, USA) following the manufacturer´s instructions. Premade NuPAGE BisTris Mini Gels (4-12 %) were used under reducing conditions and samples were prepared according to the manufacturer's instructions; 10 µL of each sample were applied to the gel. As a protein size marker Novex Sharp Pre-Stained Protein Standard (Thermo Scientific, Waltham, MA, USA) was used.

4 Enzyme Testing Results

4.1 Peptiligase Library III

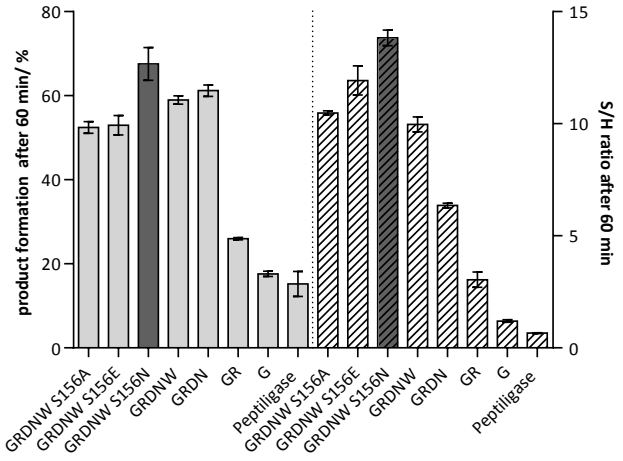

Figure S 7. Result of the testing of peptiligase library III – product formation rates and S/H ratios after 60 min are shown. Besides the new thymoligase variants results of the former best performing thymoligase variants (GRDNW, GRDN, GR, G) as well as the peptiligase wild type are shown. Mutations S156*X* (*X*= A,E,N) were incorporated into peptiligase GRDNW (M222G, L217R, S166D, A225N, F189W). Errors are depicted as the SEM and the arithmetic mean is shown; each measurement was performed in triplicates – however, due to problems with ineffective quenching only results of a double measurement are shown (data of 1/3 measurements were rejected); after 2-3 days of sample storage substantial amounts of acidic chemical background hydrolysis of the Cam-ester is observed. Values were corrected for chemical background hydrolysis by subtracting the hydrolysis of blank Ac-Thm(1-14)-OCam-Leu-OH/ % (without enzyme) at *t*= 2.5 min from each value measured in time. The screening was performed in 1 M potassium phosphate buffer (200 µL total volume, pH 8.0, 3.5 mM TCEP, room temperature (20 °C)) using 2 µg of enzyme, crude Ac-Thm(1-14)-OCam-Leu-OH (7.15 mM) and crude H-Thm(15 28)-OH (8.55 mM). Samples were quenched with a 125-fold excess of ACN/ H_2O (1:2 (v/v); 0.2% MSA (v/v)) and subsequently analyzed with analytical HPLC (*λ*= 220 nm). The results of the best thymoligase variant is highlighted in dark grey (GRDW S156N).

4.2 Diverse

In addition to the peptiligase variants testing results, the coupling reaction was investigated with different purity grades of thymoligase GRDN: clarified cell culture medium and cell culture medium exchanged to buffer (25 mM Tricine, 500 mM NaCl, pH 7.5). Surprisingly, in both cases the coupling reaction worked extremely well and conclusively, enzymes secreted by the protease deficient strain *B. subtilis* GX4935 did not interfere with the enzyme activity and the [14+14] thymosin-α_1 coupling reaction.

Moreover, we demonstrated that the product formation completely relies on the specific activity of thymoligase. As a control isolates from the empty strain *B. subtilis* GX4935 did not exhibit any synthesis nor hydrolysis activity in the [14+14] thymosin-α_1 coupling assay. Secreted proteins were isolated and purified using standard protocols.

Additionally no product formation was observed when using the ester mimic Ac-Thm(1-14)-Gly-Leu-OH as substrate implying that no enzymatic transpeptidation occurs under the given conditions.

5 Calculation of Net Product Content/ Response Factor

For the assessment of the net product content of peptide samples, the corresponding peptide was purified to HPLC purity > 95% via RP-HPLC. A dilution series of suitable concentrations was prepared and analyzed with analytical HPLC (λ= 220 nm) using a ReProSil-Pur C18 5µm, 4.6 x 250 mm (Dr. Maisch, Ammerbuch, Germany) column. As mobile phase a binary mixture of A (water+ 0.05 vol% MSA) and B (ACN + 0.05 vol% MSA) was used by default. The elution conditions were chosen individually for each peptide (Table S 1).

The signal was monitored at λ= 220 nm and the peptide peak was integrated using Agilent ChemStation software (Agilent, Santa Clara, CA, USA). The peak area was plotted against the peptide concentration and a calibration curve was obtained via linear regression of the data points. Standardly 25 µL of each sample was injected. Linear regression analysis was performed using GraphPad Prism 6 (GraphPad Software, La Jolla, CA, USA) and the net product content was calculated using the calibration curves ($y= mx+ b$).

Table S 3. Calibration curve data for the calculation of the net product content of certain peptide samples. The equations of the regression curves ($y= mx+ b$) are given for peptides Ac-Thm(1-14)OCam-Leu-OH, H-Thm(15-28)-OH and H-Thm(1-28)-OH. *y* defines the calculated/ measured integrated peak area, *m* is the slope, *x* the concentration of the peptide solution and *b* describes the y-intercept. By solving the equation towards *x* the effective peptide concentration was calculated. The net product content (%) was calculated by dividing the effective peptide concentration by the theoretical concentration.

Peptide	Slope m	y-intercept	Injection volume/ µL
Ac-Thm(1-14)-OCam-Leu-OH	5727	4.71	25
H-Thm(15-28)-OH	4885	-0.33	25

Linear regression curves for calculating the net product content as well as standard errors of the mean (SEM) are shown in the diagrams below. Linear regression was performed using GraphPad Prism 5. Statistical data of the regression are also given below.

5.1 Ac-Thm(1-14)-OCam-Leu-OH

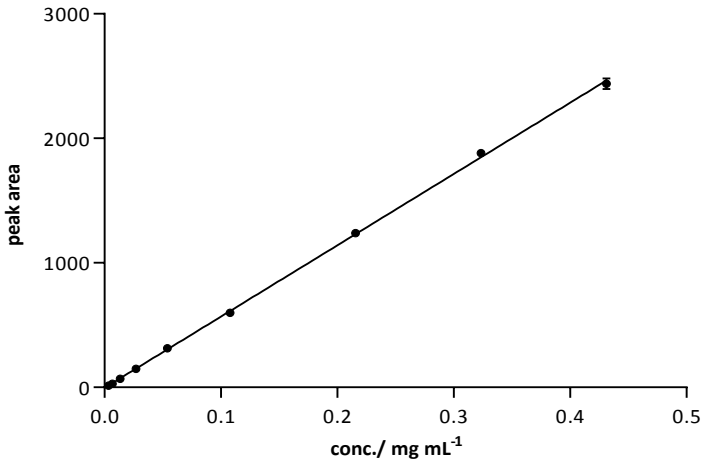

Figure S 8. Linear regression curve for the calculation of the net product content of Ac-Thm(1-14)-OCam-Leu-OH.

Table S 4. Statistical Data of the linear regression curve for the calculation of the net product content of Ac-Thm(1-14)-OCam-Leu-OH.

Best-fit values	
Slope	5727 ± 38,24
Y-intercept when X=0.0	-4,716 ± 7,569
X-intercept when Y=0.0	0,0008234
1/slope	0,0001746
95% Confidence Intervals	
Slope	5648 to 5806
Y-intercept when X=0.0	-20,31 to 10,88
X-intercept when Y=0.0	-0,001917 to 0,003514
Goodness of Fit	
R square	0,9989
Sy.x	29,41
Is slope significantly non-zero?	
F	22434
DFn, DFd	1,000, 25,00
P value	< 0,0001
Deviation from zero?	Significant
Data	
Number of X values	9
Maximum number of Y replicates	3
Total number of values	27
Number of missing values	0

5.2 H-Thm(15-28)-OH

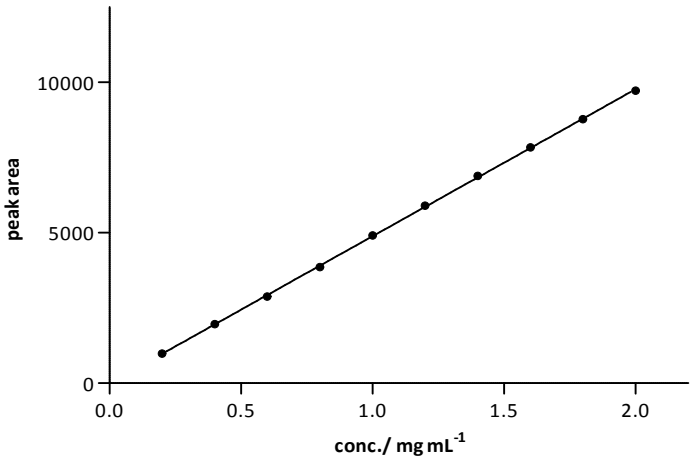

Figure S 9. Linear regression curve for the calculation of the net product content of H-Thm(15-28)-OH.

Table S 5. Statistical Data of the linear regression curve for the calculation of the net product content of H-Thm(15-28)-OH.

Best-fit values	
Slope	4885 ± 28,02
Y-intercept when X=0.0	-0,3334 ± 34,77
X-intercept when Y=0.0	6,825e-005
1/slope	0,0002047
95% Confidence Intervals	
Slope	4826 to 4944
Y-intercept when X=0.0	-73,39 to 72,72
X-intercept when Y=0.0	-0,01505 to 0,01486
Goodness of Fit	
R square	0,9994
Sy.x	71,99
Is slope significantly non-zero?	
F	30395
DFn, DFd	1,000, 18,00
P value	< 0,0001
Deviation from zero?	Significant
Data	
Number of X values	10
Maximum number of Y replicates	2
Total number of values	20
Number of missing values	0

6 Peptides – HPLC-MS Data

Peptides were analyzed using an Agilent 1260 Infinity HPLC system coupled with an Agilent 6130 quadrupole mass spectrometer (Agilent, Santa Clara, CA, USA) to determine peptide mass. Separation was performed using ReProSil-Pur C18 5µm, 4.6 x 250 mm (Dr. Maisch, Ammerbuch, Germany) column.

6.1 Ac-Thm (1-14)-OCamLeu-OH

Ac-Thm(1-14)-OCam-Leu-OH:

Ac-^{1}Ser-^{2}Asp-^{3}Ala-^{4}Ala-^{5}Val-^{6}Asp-^{7}Thr-^{8}Ser-^{9}Ser-^{10}Glu-^{11}Ile-^{12}Thr-^{13}Thr-^{14}Lys-OCam-Leu-OH

Mol. weight: 1637.71 g mol^{-1} ($C_{67}H_{112}N_{16}O_{31}$)

5.0 mmol scale synthesis crude:

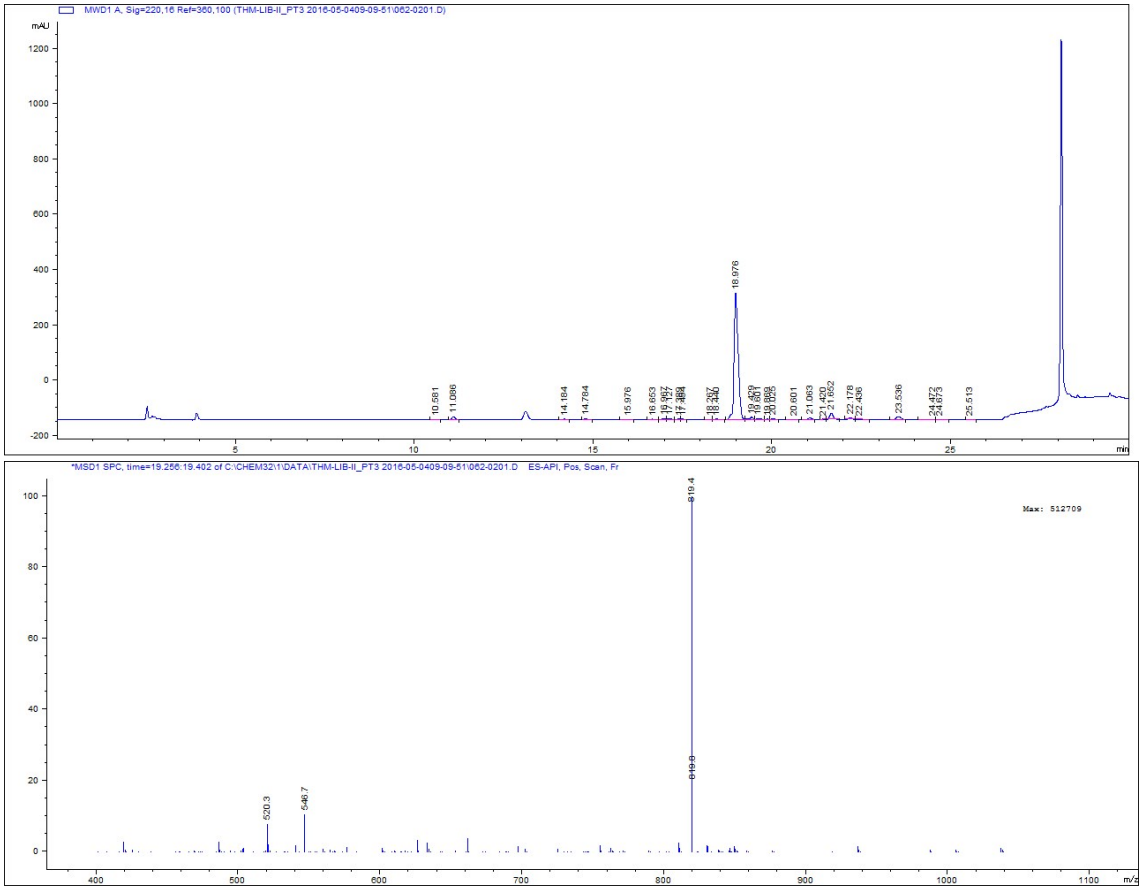

Figure S 10. HPLC spectrum (λ= 220 nm) (up) of crude Ac-Thm(1-14)-OCam-Leu-OH synthesized at 5.0 mmol scale and corresponding mass spectrum of the product peak (down). A gradient from 12 to 30 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 21 min was used. The flow was 1.5 mL min^{-1} and the column temperature was set to 40 °C.

0,5 mmol scale synthesis purified:

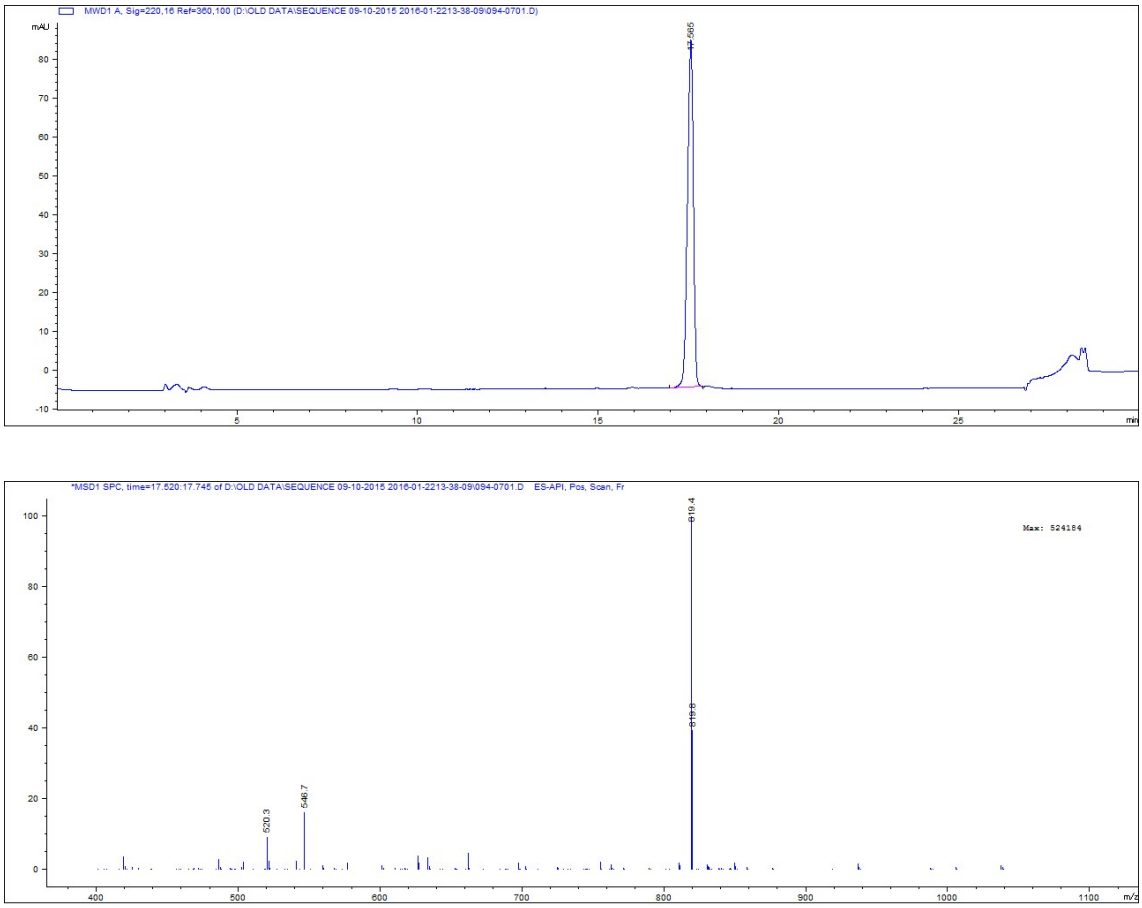

Figure S 11. HPLC spectrum (*λ*= 220 nm) (up) of purified Ac-Thm(1-14)-OCam-Leu-OH synthesized at 0,5 mmol scale and corresponding mass spectrum of the product peak (down). A gradient from 12 to 30 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 21 min was used. The flow was 1.5 mL min^{-1} and the column temperature was set to 40 °C.

6.2 H-Thm (15-28)-OH

H-Thm(15-28)-OH:

H-^{15}Asp-^{16}Leu-^{17}Lys-^{18}Glu-^{19}Lys-^{20}Lys-^{21}Glu-^{22}Val-^{23}Val-^{24}Glu-^{25}Glu-^{26}Ala-^{27}Glu-^{28}Asn-OH

Mol. weight: 1659.81 g mol^{-1} ($C_{70}H_{118}N_{18}O_{28}$)

5.0 mmol scale synthesis crude:

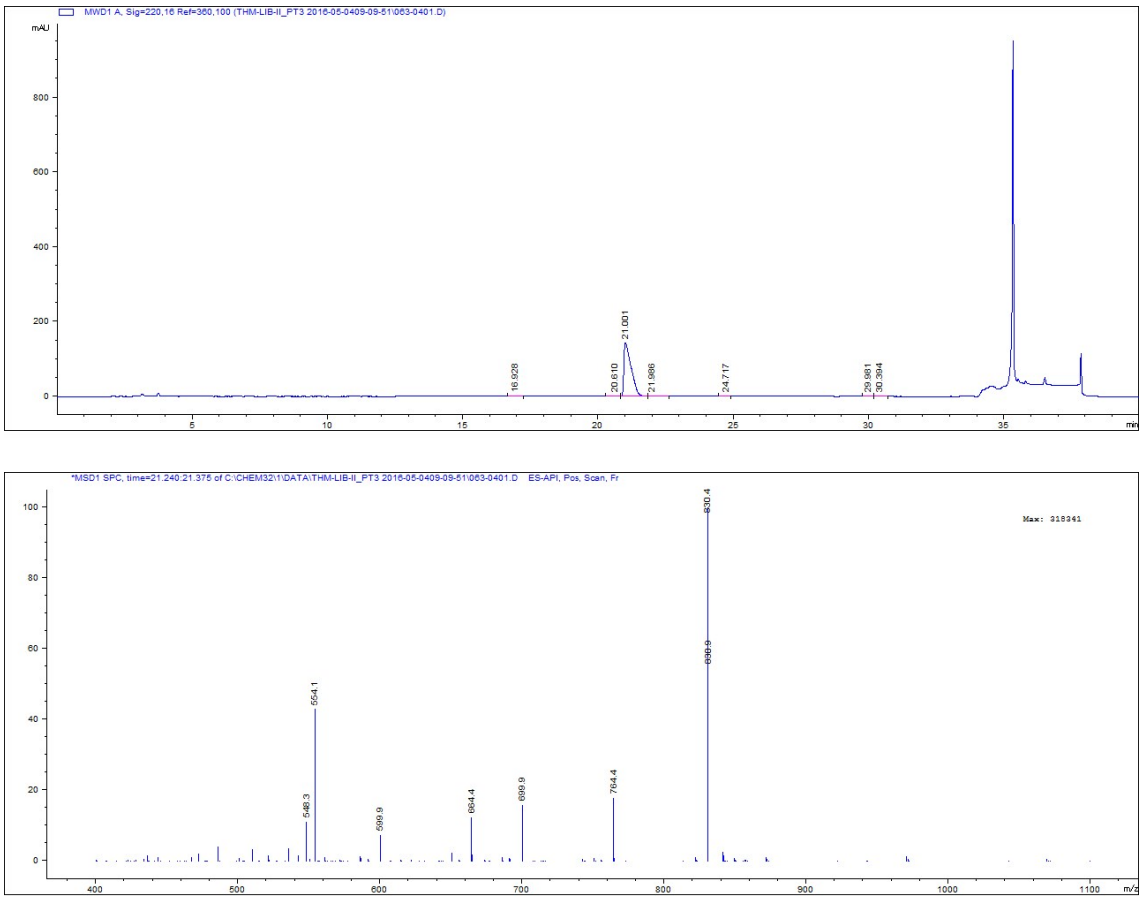

Figure S 12. HPLC spectrum (λ= 220 nm) of crude H-Thm(15-28)-OH synthesized at 5.0 mmol scale (up) and corresponding mass spectrum of the product peak (down). A gradient from 5 to 14 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 30 min was used. The flow was 1.5 mL min^{-1} and the column temperature was set to 20 °C.

0.5 mmol scale synthesis purified:

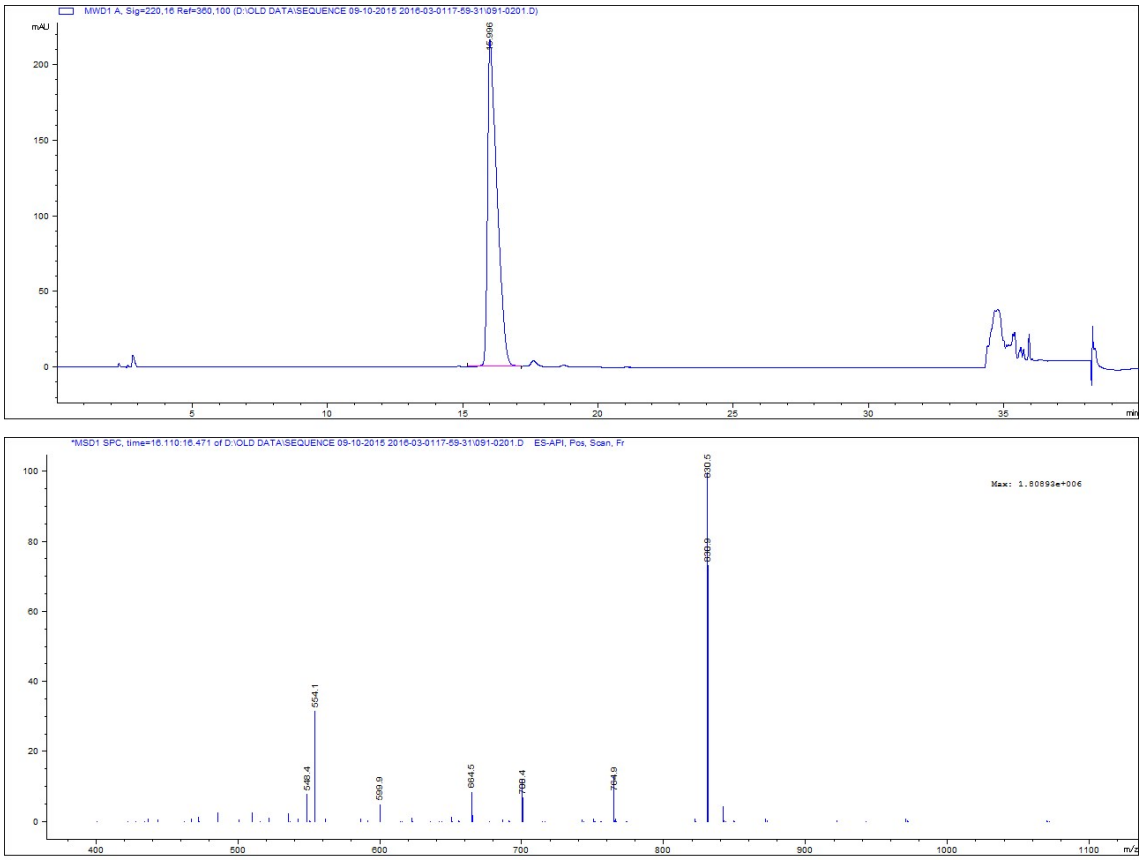

Figure S 13. HPLC spectrum (*λ*= 220 nm) of crude H-Thm(15-28)-OH synthesized at 5.0 mmol scale (up) and corresponding mass spectrum of the product peak (down). A gradient from 5 to 14 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 30 min was used. The flow was 1.5 mL min^{-1} and the column temperature was set to 20 °C.

6.3 Ac-Thm(1-28)-OH

H-Thm(1-28)-OH:

Ac-^{1}Ser-^{2}Asp-^{3}Ala-^{4}Ala-^{5}Val-^{6}Asp-^{7}Thr-^{8}Ser-^{9}Ser-^{10}Glu-^{11}Ile-^{12}Thr-^{13}Thr-^{14}Lys-^{15}Asp-^{16}Leu-^{17}Lys-^{18}Glu-^{19}Lys-^{20}Lys-^{21}Glu-^{22}Val-^{23}Val-^{24}Glu-^{25}Glu-^{26}Ala-^{27}Glu-^{28}Asn-OH

Mol. weight: 3108.32 g mol^{-1} ($C_{129}H_{215}N_{33}O_{55}$)

The crude spectrum of the CEPS of thymosin-α_1 using thymoligase is shown below:

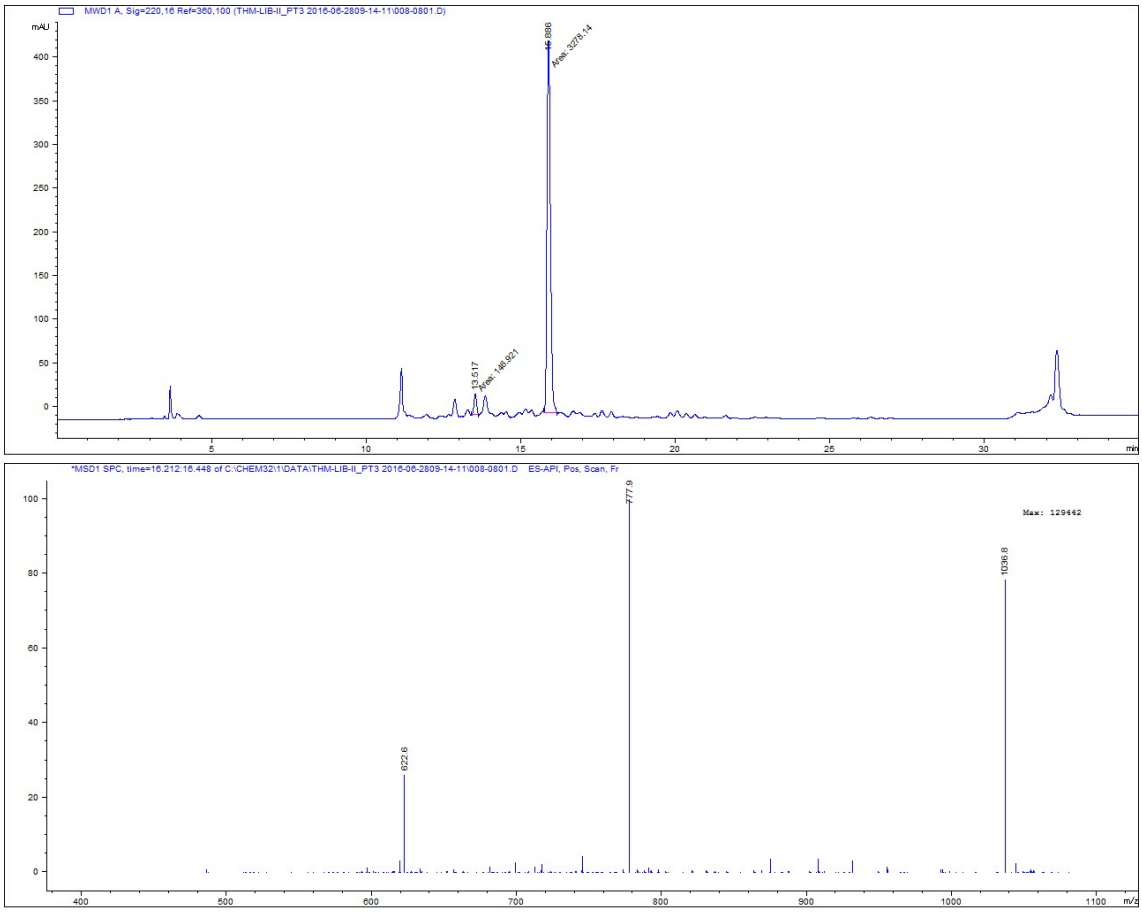

Figure S 14. HPLC spectrum (λ= 220 nm) of crude Ac-Thm(1-28)-OH (up) synthesized via CEPS using thymoligase GRDNWN via a [14+14]-mer strategy and corresponding mass spectrum of the product peak (down). A gradient from 12 to 30 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 21 min was used. The flow was 1.0 mL min^{-1} and the column temperature was set to 40 °C. CEPS was performed at room temperature (20°C); 1 g Cam-ester Ac-Thm(1-14)-OCam-Leu-OH with 1.2 eq. of the amine segment H-Thm(15-28)-OH were used. The Cam-ester was dosed over time (200 mg every 30 min) and the reaction was performed in 1 M potassium phosphate (pH 8.3, 3.5 mM TCEP) using thymoligase GRDNWN (2 mg).

The spectrum of purified thymosin-α_1 produced by CEPS using thymoligase is shown below:

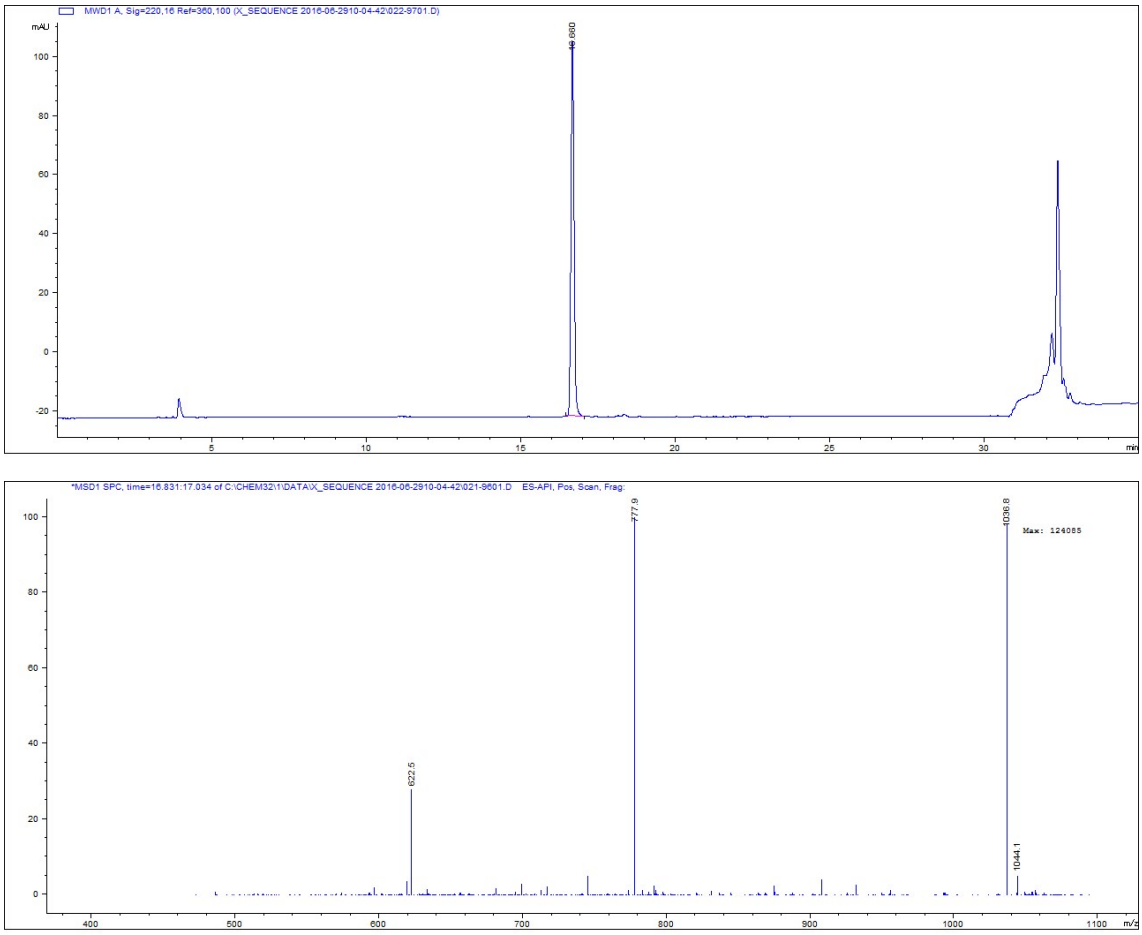

Figure S 15. HPLC spectrum (λ= 220 nm) of purified Ac-Thm(1-28)-OH (up) synthesized via CEPS using thymoligase GRDNWN via a [14+14]-mer strategy and corresponding mass spectrum of the product peak (down). A gradient from 12 to 30 vol % ACN/ MQ-water (v/v; 0.05 vol % MSA) in 21 min was used. The flow was 1.0 mL min^{-1} and the column temperature was set to 40 °C. CEPS was performed at room temperature (20°C); 1 g Cam-ester Ac-Thm(1-14)-OCam-Leu-OH with 1.2 eq. of the amine segment H-Thm(15-28)-OH were used. The Cam-ester was dosed over time (200 mg every 30 min) and the reaction was performed in 1 M potassium phosphate (pH 8.3, 3.5 mM TCEP) using thymoligase GRDNWN (approximately 2 mg). Purification was carried out via a single RP-HPLC run using a self-packed Luna Prep C18 column (10 µm, 200x 51 mm, Phenomenex, Torrance, CA, USA) eluting with a water acetonitrile gradient (16 % to 28 % ACN in 42 min), with a flow rate of 45 mL/min. As mobile phase a binary mixture of A (MQ-water + 0.05 vol % TFA) and B (ACN + 0.05 vol % TFA) was used by default.

7 Enantiomeric Purity and Peptide Content of Ac-Thm(1-28-OH)

Analysis has been performed by C.A.T. GmbH (Tübingen, Germany).

Analysis number	thymosin-a1
9995014-1-2/16	

Results: A.0.3. /rv 140206

The listed amino acid(s) were identified via retention time and mass spectra.
The identity of the main component(s) comply.
The following impurity of the optical antipode was found:

Alanine	<0.10 % D-Enantiomer
Valine	<0.10 % D-Enantiomer
Threonine	>99.4 % L-Threonine
	<0.10 % D-Threonine
	0.16 % L-allo Threonine
	0.23 % D-allo Threonine
Isoleucine	>99.7 % L-Isoleucine
	<0.10 % D-Isoleucine
	<0.10 % L-allo-Isoleucine
	<0.10 % D-allo-Isoleucine
Leucine	0.16 % D-Enantiomer
Serine	<0.10 % D-Enantiomer
Aspartic acid	0.11 % D-Enantiomer
Glutamic acid	<0.10 % D-Enantiomer
Lysine	<0.10 % D-Enantiomer

Notes:

Method specific deviations or irregularities are not observed.

The method is generically validated. However it may not meet all requirements for the release of drug substances and drug products. It is to prove if substance specific validation is required.

Analysis number	thymosin-a1
9995014-1-1/16	0

Parameter of Analysis: **A.0.4.** **/rv 090113**

Sample weight [mg]:	1.043
Molecular weight [g/mol]:	3108.00
Time of hydrolysis [h]:	48
Conditions:	110°C o.Vak.

Results:

		Concentrations		Residues		
Amino acid		**mg/g**	**µmol/g**	**Theory**	**weighted**	**µmol/µmol**
Alanine	ALA	69.7	782.1	3	3	2.95
Valine	VAL	87.1	743.1	3	3	2.80
Threonine	THR	90.0	755.6	3		2.85
Isoleucine	ILE	33.1	252.3	1		0.95
Leucine	LEU	36.8	280.5	1	1	1.06
Serine	SER	78.2	744.2	3		2.81
Aspartic acid	ASP	148.0	1112.0	4	4	4.19
Glutamic acid	GLU	254.1	1726.9	6		6.51
Lysine	LYS	152.1	1040.4	4		3.92

Peptide content:	82.4 % or	265.2 µmol/g

Figure S 16. Enantiomeric Purity and peptide content of Ac-Thm(1-28-OH).

8 NMR Fmoc-Lys(Boc)-O-Glycolic acid

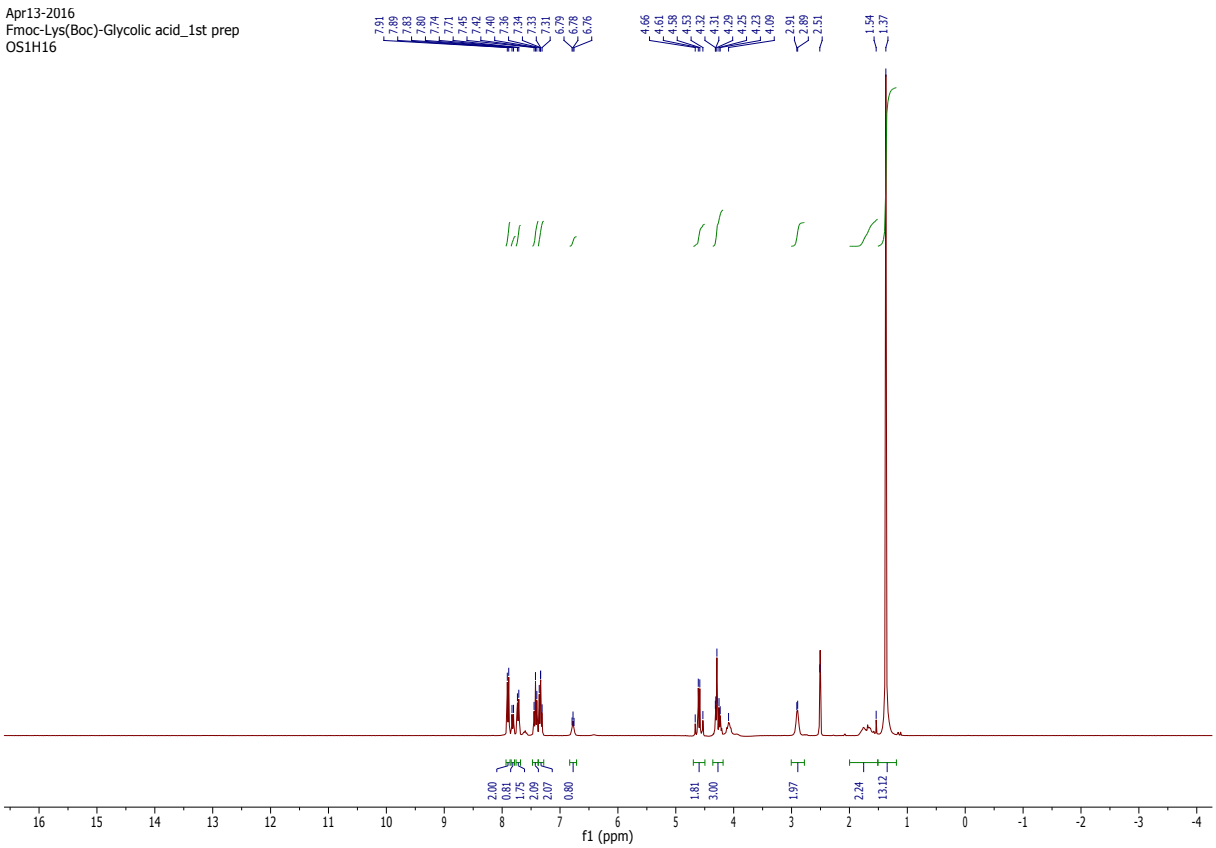

Figure S 17. NMR spectrum of Fmoc-Lys(Boc)-glycolic acid.

9 Reaction Optimization

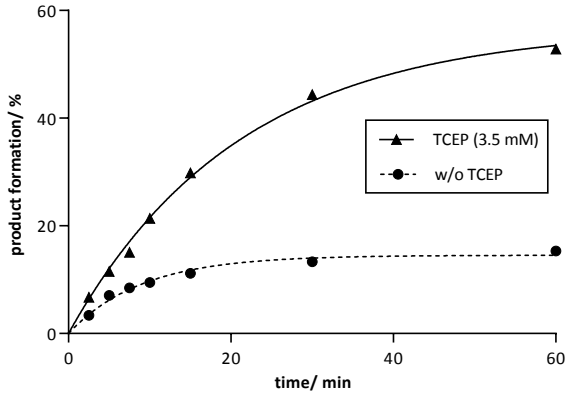

Figure S 18. Thymoligase GRDN progress curve with the use of TCEP (3.5 mM) and without the use of TCEP as a reducing agent. Values were corrected for chemical background hydrolysis by subtracting the hydrolysis of blank Ac-Thm(1-14)-OCam-Leu-OH/ % (without enzyme) at *t*= 2.5 min from each value measured in time. The reaction was performed in 1 M potassium phosphate buffer (200 µL total volume, pH 8.3, room temperature (20 °C)) using 2 µg of enzyme, crude Ac-Thm(1-14)-OCam-Leu-OH (7.15 mM) and crude H-Thm(15 28)-OH (8.55 mM). Samples were quenched with a 125-fold excess of ACN/ H_2O (1:2 (v/v); 0.2 % MSA (v/v)) and subsequently analyzed with analytical HPLC (*λ*= 220 nm). Data points were fitted by a one phase exponential association equation (least squares method) using GraphPad Prism 6.

10 References

1 W. Kabsch, *Acta Crystallogr. D. Biol. Crystallogr.*, 2010, **66**, 125–32.

2 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 235–242.

3 B. W. Matthews, *J. Mol. Biol.*, 1968, **33**, 491–7.

4 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

5 G. Langer, S. X. Cohen, V. S. Lamzin and A. Perrakis, *Nat. Protoc.*, 2008, **3**, 1171–1179.

6 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1997, **53**, 240–255.

7 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 486–501.

8 V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 12–21.

9 R. A. Laskowski and M. B. Swindells, *J. Chem. Inf. Model.*, 2011, **51**, 2778–2786.