Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

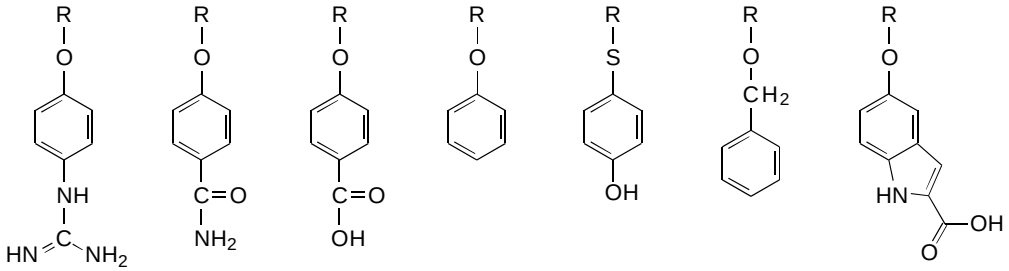
Supporting Information

69451 Weinheim, Germany

D-Amino Acid Specific Proteases and Native all-L-Proteins: A Convenient Combination for Semisynthesis

Nicole Wehofsky, Andreas Pech, Sandra Liebscher, Stephanie Schmidt, Hidenobu Komeda, Yasuhisa Asano and Frank Bordusa

Supplemental schemes and figures:



Scheme S1. Structures of substrate mimetics used. (1) 4-guanidinophenyl (OGp), (2) 4-carboxamidophenyl (OCap), (3) 4-carboxyphenyl (OCp), (4) phenyl (OPh), (5) 4-hydroxythiophenyl (SPhOH), (6) benzyl (OBzl), (7) 2-carboxy-1*H*-indol-5-yl (OInd), R = Boc-alanyl.

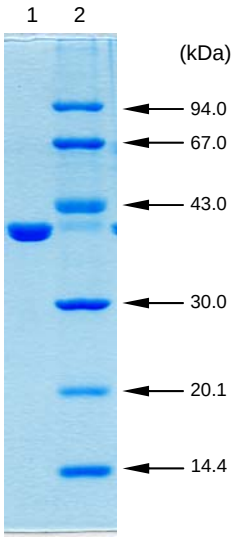


Figure S1: SDS-polyacrylamide electrophoresis of the alkaline D-peptidase (ADP). *lane 1:* purified ADP (10 µg, Mw: 37,952 Da); *lane 2:* molecular weight marker.

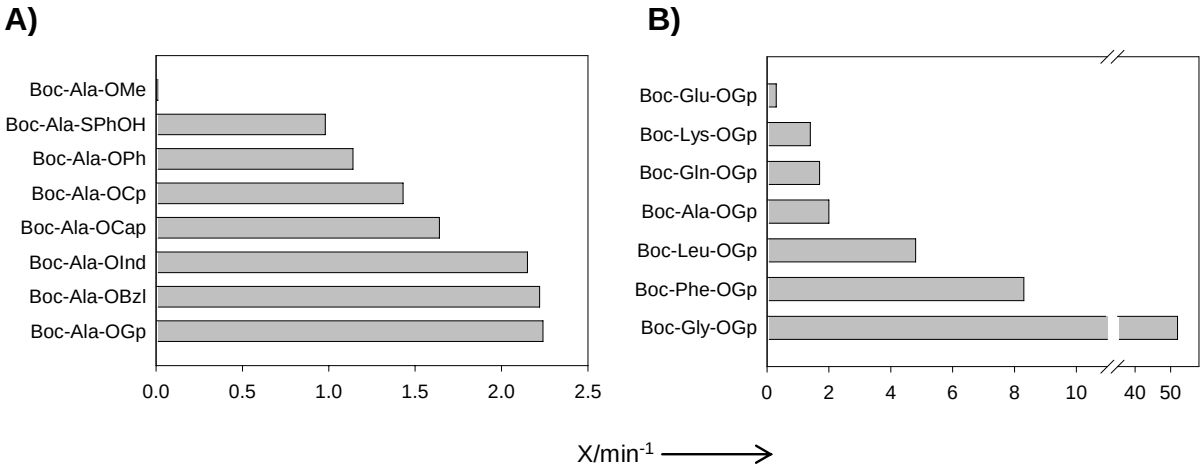


Figure S2. Initial rates of hydrolysis of Boc-Ala-O/SR esters (A) and Boc-Xaa-OGp esters (B) catalysed by the alkaline D-peptidase. Conditions: 30 °C, 0.1 M HEPES buffer pH 8.0, 10% (v/v) DMF, [substrate] = 2 mM, [enzyme] = 1.0×10^{-7} M – 1.0×10^{-5} M, *X* = hydrolysis rate. Errors are less than 5% (+/- 2.5%).

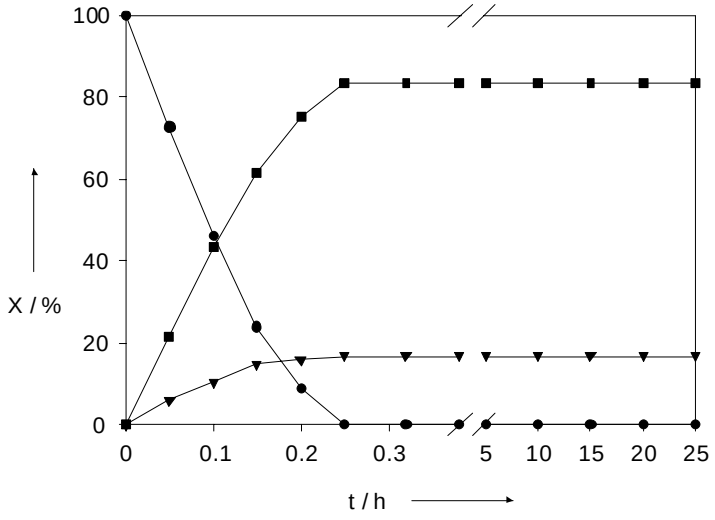


Figure S3. Representative time-course of ADP-catalysed peptide fragment ligations on example of Bz-Gly-OGp and LIVDAVLEPVKAAGAY. Conditions: 30 °C, 0.1 M HEPES buffer pH 8.0, 10% (v/v) DMF, [acyl donor] = 2 mM, [acyl acceptor] = 10 mM, [enzyme] = 8.7×10^{-6} M. (–●–), Bz-Gly-OGp; (–▼–), Bz-Gly-OH; (–■–) Bz-Gly-Leu-Ile-Val-Asp-Ala-Val-Leu-Glu-Pro-Val-Lys-Ala-Ala-Gly-Ala-Tyr-OH. *X* = relative concentration.

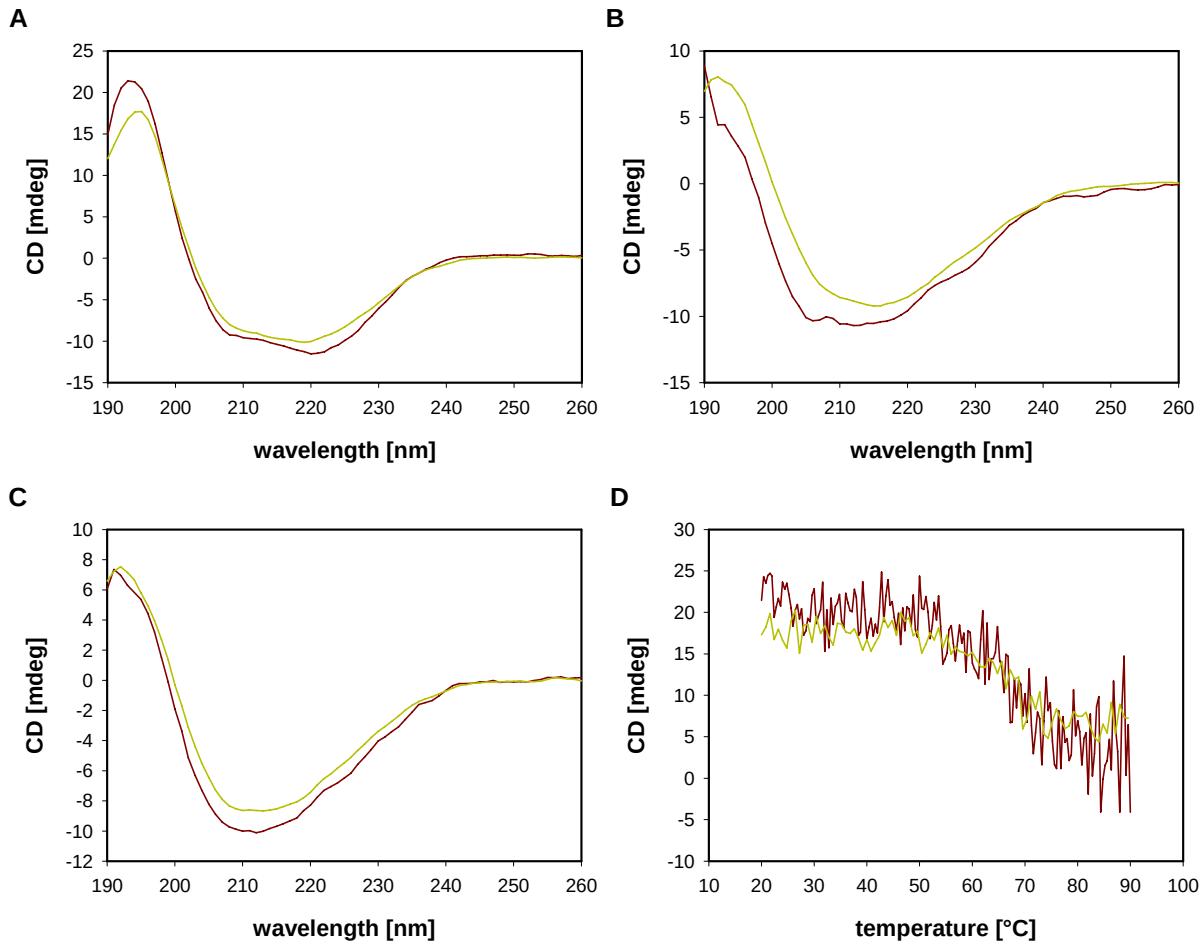


Figure S3. CD-spectroscopical analysis of synthesized and recombinant *E. coli* Parvulin 10. A) temperature 20 °C; B) temperature 90 °C; C) temperature 20 °C after heating to 90 °C; D) course of signal change at 195 nm from 20 to 90 °C (— recombinant *E. coli* parvulin 10; — synthetic *E. coli* parvulin 10). Conditions: 10 mM potassium phosphate, pH 7.8, 10 µM protein.

Experimental Section

Materials. The alkaline D-peptidase according to [1] has been expressed in *E.coli* BL21 (DE3) with a *C*-terminal His_6-tag fusion. Purification of the construct was performed via Ni-NTA (Qiagen) and final FPLC-MonoQ column chromatography (*cf.* Figure S1). The enzyme was stored in 10 mM Tris-HCl (pH 8.0) containing 10% (v/v) glycerol, 0.05 mM $MgSO_4$, and 100 mM NaCl in a concentration of 5.9 mg/ml. Specific activity of the enzyme was found to be 5.3 U/mg.

Wild-type parvulin 10 from *E. coli* was obtained from Prof. G. Fischer (Max-Planck Research Unit for Enzymology of Protein Folding, Halle (Saale), Germany) produced as described in [2].

All further substances (excluding peptides, amino acid esters, and peptide esters) including the materials for the synthesis of peptides, amino acid and peptide esters were products of Bachem, Fluka, Merck, Aldrich or Novabiochem, respectively. If not otherwise stated, all reagents were of the highest available commercial purity. Solvents were purified and dried by usual methods.

Chemical syntheses of amino acid- and tripeptide-derived substrate mimetics. Boc-Ala-OBzl, Boc-Ala-OCap, Boc-Ala-OCp, Boc-Ala-OGp, Boc-Ala-OInd, Boc-Ala-OPh, Boc-Ala-SPhOH, and Boc-Xaa-OGp esters were prepared according to our previously described protocols by coupling of Boc-Ala-OH or site-chain protected Boc-Xaa-OH with the respective alcoholic component using the DCC or TBTU activation method, respectively.[3] Substrate mimetics of the structure Bz-$(Xaa)_n$-OGp were prepared from the appropriate Boc-$(Xaa)_n$-OGp(Z,Z) precursors by *N*-terminal deprotection, subsequent benzoylation and final catalytic hydrogenation following the procedure described by Thormann et al.[4] The identity and purity of all final products were checked by analytical HPLC analysis at a wavelength of 220 nm, NMR, thermospray mass spectrometry, and elementary analysis. Satisfactory analytical data were found in all cases ($\pm$ 0.4% for C, H, N).

Procedure for the preparation of peptides. LGSVKASAYK, LIVDAVLEPVKAAGAY, and KHSICPSGKRGGDLGEFRQGQMVPAFDKVVFSCPVLEPTGPLHTQFGYHIIKVLYRN were prepared with a fully automated peptide synthesiser (Applied Biosystems, USA) using standard Fmoc chemistry and 2-chlorotrityl chloride resin. After simultaneous release of the peptides from the resin and side-chain deprotection by TFA/TIS/water treatment, the peptides were precipitated with dry diethylether and purified by preparative HPLC.

LGSVKASAYK × 3 TFA: Anal. calcd. for $C_{52}H_{81}F_9N_{12}O_{20}$: C 45.75, H 5.98, N 12.31. Found: C 46.29, H 5.91, N 12.40. MS (MALDI-ToF), *m/z* calcd. for $C_{46}H_{78}N_{12}O_{14}$ (1022.6): 1023.9 $[M+H]^+$.

LIVDAVLEPVKAAGAY × 2 TFA: Anal. calcd. for $C_{80}H_{127}F_6N_{17}O_{26}$: C 51.74, H 6.89, N 12.82. Found: C 51.83, H 6.91, N 12.90. MS (MALDI-ToF), *m/z* calcd. for $C_{76}H_{125}N_{17}O_{22}$ (1627.9): 1629.4 $[M+H]^+$.

KHSICPSGKRGGDLGEFRQGQMVPAFDKVVFSCPVLEPTGPLHTQFGYHIIKVLYRN × 11 TFA: Anal. calcd. for $C_{310}H_{455}F_{33}N_{80}O_{99}S_3$: C 48.93, H 6.03, N 14.73. Found: C 49.02, H 6.07, N 14.70. MS (MALDI-ToF), *m/z* calcd. for $C_{288}H_{444}N_{80}O_{77}S_3$ (6351.2): 6353.8 $[M+H]^+$.

Synthesis protocol for Z-AKTAAALHILVKEEKLALDLLEQIKNGADFGKLAK-OGp and Bz-AYLDAYVKAG-OGp. The peptide 4-guanidinophenyl esters Bz-AYLDAYVKAG-OGp and the 35-mer Z-AKTAAALHILVKEEKLALDLLEQIKNGADFGKLAK-OGp were synthesised with a fully automated peptide synthesiser (Applied Biosystems, USA) using Kenner's alkanesulfonamide safety-catch linker approach.[5] The first amino acid, i.e. Fmoc-Ala-OH (relates to the second last moiety within the aforementioned sequences), was loaded to 4-sulfamylbenzoyl aminomethyl resin by one PyBOP/DIEA coupling step resulting in a loading of 0.35 mmol/g resin. All remaining amino acids were coupled by stepwise solid-phase peptide synthesis using HCTU/DIEA activation protocols. For the preparation of Bz-AYLDAYVKAG-OGp, at the end of synthesis the N^{α}-amino group of the resin-bound peptide was deprotected by treatment with DBU/piperidine/DMF and subsequently benzoylated using activated benzoic acid. In contrast, the automated synthesis of the resin-linked 35-mer peptide ends up with the coupling of Z-Ala-OH. Alkylation of the linker´s sulfonamide functionality has been achieved with iodoacetonitrile according to the procedure described by Backes et al.[6] leading to the respective activated *N*,*N*-

cyanomethylacylalkane sulfonamide ester intermediates. The protected peptide esters were liberated from the resin by adding a 5-fold excess of the tosylate salts of the self-prepared H-Lys(Boc)-OGp(Boc,Boc) and H-Gly-OGp(Boc,Boc), respectively. Neutralisation of the amino acid ester tosylates was achieved by adding appropriate equivalents of NMM. Simultaneous deprotection of trifunctional peptide side-chains and guanidino functionalities by a TFA/TIS/water treatment and purification of the crude products by preparative HPLC resulted in the respective final peptide esters. The identity and purity of the amino acid 4-guanidinophenyl esters and synthesis products were checked by analytical HPLC, NMR (for Xaa-OGp(Boc,Boc) , elementary analysis, and mass spectrometry. In all cases satisfactory analytical data have been found.

The 4-guanidinophenyl esters H-Lys(Boc)-OGp(Boc,Boc) and H-Gly-OGp(Boc,Boc) were prepared by condensation of the appropriate *Z*-protected amino acid and 4-[*N´,N´´*-bis(Boc)guanidino]phenol using the TBTU/DIEA activation protocol. A final catalytic hydrogenation results in the respective N^{α}-deprotected derivatives which were further analysed as described.

H-Gly-OGp(Boc,Boc) × Tos × H_2O: ^{1}H NMR (300 MHz, DMSO): δ 1.46/1.48 (s/s, 18H, C_4H_9, Boc), 2.32 (s, 3H, CH_3, Tos), 4.16 (m, 2H, CH_2, Gly), 7.14/7.61 (m/m, 4H, C_6H_4, OGp), 7.20/7.51 (m/m, 4H, C_6H_4, Tos), 8.39 (m, 2H, NH_2, Gly), 10.03 (s, 1H, NH, OGp), 11.39 (s, 1H, NH, OGp). Anal. calcd. for $C_{26}H_{38}N_4O_{10}S$: C 52.16, H 6.40, N 9.36. Found: C 52.28, H 6.41, N 9.29. MS (ESI), *m/z* calcd. for $C_{19}H_{28}N_4O_6$ (408.2): 409.3 $[M+H]^+$.

H-Lys(Boc)-OGp(Boc,Boc) × Tos × H_2O: ^{1}H NMR (300 MHz, DMSO): δ 1.38/1.41/1.52 (s/s/s, 27H, C_4H_9, Boc), 1.41-1.80 (m, 6H, $C_{\beta\text{-}\delta}H_2$, Lys), 2.29 (s, 3H, CH_3, Tos), 2.73-2.89 (m, 2H, $C_{\varepsilon}H_2$, Lys), 4.11 (m, 1H, $C_{\alpha}H$, Lys), 6.79 (m, 1H, $N_{\varepsilon}H$, Lys) 7.11/7.60 (m/m, 4H, C_6H_4, OGp), 7.19/7.50 (m/m, 4H, C_6H_4, Tos), 8.61 (t, 2H, NH_2, Lys), 9.99 (s, 1H, NH, OGp), 11.37 (s, 1H, NH, OGp); Anal. calcd. for $C_{35}H_{55}N_5O_{12}S$: C 54.60, H 7.20, N 9.10. Found: C 54.78, H 7.14, N 9.19. MS (ESI), *m/z* calcd. for $C_{28}H_{45}N_5O_8$ (579.3): 580.2 $[M+H]^+$.

Bz-AYLDAYVKAG-OGp × 2 TFA: Anal. calcd. for $C_{68}H_{88}F_6N_{14}O_{20}$: C 53.19, H 5.78, N 12.77. Found: C 53.69, H 5.81, N 12.90. MS (MALDI-ToF), *m/z* calcd. for $C_{64}H_{86}N_{14}O_{16}$ (1306.6): 1308.2 $[M+H]^+$.

Z-AKTAAALHILVKEEKLALDLLEQIKNGADFGKLAK-OGp × 8 TFA: Anal. calcd. for $C_{202}H_{312}F_{24}N_{48}O_{67}$: C 49.10, H 6.36, N 13.61. Found: C 48.89, H 6.31, N 13.70. MS (MALDI-ToF), *m/z* calcd. for $C_{186}H_{304}N_{48}O_{51}$ (4026.3): 4028.1 $[M+H]^+$.

Enzyme kinetics (ADP). Hydrolysis reactions with the substrates listed in Fig. 2S were performed at 30 °C using an assay mixture containing 0.1 M HEPES buffer pH 8.0 and 10% (v/v) DMF to realise complete solubility of the ester derivatives. The final substrate concentration was 2.0 mM and the enzyme concentrations were in the range of 0.1 and 10 μM. After thermal equilibration of assay mixtures, the reactions were initiated by addition of aqueous enzyme stocks. The rate of reactions was analysed by analytical HPLC (254 nm) determining the disappearance of substrate esters. For this purpose, aliquots were withdrawn at defined time intervals and diluted with a quenching solution containing 50% (v/v) methanol and 5% (v/v) trifluoroacetic acid. Experiments without

enzyme were performed to determine the extent of non-enzymatic ester hydrolysis which was generally less than 5%. All values reported are the average of at least three independent experiments.

Enzymatic dipeptide syntheses. Enzymatic coupling reactions of amino acid-derived esters and amides were performed at 30 °C in 0.1 M HEPES buffer pH 8.0 containing 10% (v/v) DMF as cosolvent. Stock solutions of acyl donor esters (20 mM) were prepared in DMF. Acyl acceptor components (stock solution: 40 mM) were dissolved in 0.2 M HEPES buffer pH 8.0 and subsequently readjusted to a pH value of 8.0 by adding appropriate equivalents of NaOH. The final concentrations of acyl donor and acyl acceptor were 2 mM and 20 mM, respectively. The latter was calculated as free, N^{α}-unprotected species according to the Henderson-Hasselbalch equation $[HN]_0 = [N]_0/(1 + 10^{pK-pH})$. The p*K* values of the α-amino group of the acyl acceptor components were determined by inflection point titration on a Video titrator VIT 90 (Radiometer, Denmark). After mixing and thermal equilibration of assay mixtures at 30 °C, the reactions were initiated by the addition of enzyme stocks resulting in ADP concentrations between 0.3 and 44 µM. After defined time intervals, the reactions were quenched by addition of trifluoroacetic acid (5% (v/v) solution) and further analysed by HPLC (254 nm) as described. Reaction times of 15 min led to complete ester consumption. To control for spontaneous reactions, parallel reactions without enzyme were analysed in all cases. On the basis of these controls, non-enzymatic peptide synthesis could be entirely ruled out and the extent of spontaneous ester hydrolysis was found to be less than 5%. The identity of the formed peptide products was confirmed by thermospray mass spectrometry. The data reported are the average of at least three independent reactions with errors less than 5% (+/- 2.5%).

Enzymatic coupling of elongated peptides. ADP-catalysed ligation reactions of elongated peptides were performed at 30 °C in 0.1 M HEPES buffer pH 8.0 containing 10% (v/v) DMF as cosolvent. Stock solutions of acyl donor esters (20 mM) were prepared in DMF. Acyl acceptor components (stock solution: 20 mM) were dissolved in 0.2 M HEPES buffer pH 8.0 and subsequently readjusted to a pH value of 8.0 by adding appropriate equivalents of NaOH. If not otherwise stated, the final concentrations of acyl donor and acyl acceptor were 2 mM and 10 mM, respectively. After mixing and thermal equilibration of assay mixtures at 30 °C, the reactions were initiated by addition of enzyme stocks resulting in ADP concentrations between 5.5 and 8.7 µM. After defined time intervals the reactions were stopped by addition of trifluoroacetic acid (5% (v/v) solution) and further analysed as described. Reaction times of 15 to 30 min led to complete ester consumption. To control for spontaneous reactions, parallel reactions without enzyme were analysed in all cases. On the basis of these controls, non-enzymatic peptide synthesis could be entirely ruled out and the extent of spontaneous ester hydrolysis was found to be less than 5%. The identity of the formed peptide products was established by thermospray and MALDI-ToF mass spectrometry. The data reported are the average of at least three independent reactions with errors less than 5% (+/- 2.5%).

Semisynthesis of *E. coli* Par10. The ADP-catalysed ligation of the 35-mer peptide 4-guanidinophenyl ester with the 57-mer peptide fragment was performed at 30 °C in 0.1 M HEPES buffer pH 8.0 containing 20% (v/v) DMF as cosolvent in a total volume of 250 µl. Stock solutions of acyl donor esters (1 mM) were prepared in DMF. The acyl acceptor component (stock solution: 0.2 mM) was dissolved in 0.2 M HEPES buffer pH 8.0 and subsequently

readjusted to a pH value of 8.0 by adding appropriate equivalents of NaOH. The final concentrations of acyl donor (35-mer) and acyl acceptor (57-mer) were 0.2 mM and 0.1 mM, respectively. After mixing and thermal equilibration of assay mixtures at 30 °C, the reactions were initiated by the addition of enzyme stock resulting in an ADP concentration of 20 μM. After defined time intervals, the reactions were quenched by addition of trifluoroacetic acid (5% (v/v) solution) and further analysed by HPLC (220 nm) and SDS-PAGE as described. Reaction times of 30 min led to complete ester consumption. To control for spontaneous reactions, parallel reactions without enzyme were analysed in all cases. On the basis of these controls, non-enzymatic protein synthesis could be entirely ruled out. The identity of the Par10 synthesised was established by MALDI-ToF mass spectrometry, SDS-PAGE (16.5% Tris/Tricine gel), CD-measurements, and enzyme kinetic studies as reported. The yield of ligation was determined from SDS-Page by quantifying the band intensity densitometrically using the analysis package Multi-Analyst version 1.5 (Bio-Rad laboratories). In addition, quantitative HPLC analyses using recombinant Par 10 as the calibration standard were performed. The variations between the SDS-Page and HPLC analysis were found to be less than 5%. The yield reported represents the average between the two analysis methods based on three independent ligation reactions with a total error of less than 5% (+/- 2.5%).

Isolation, deprotection, purification, and refolding of *E. coli* Par10. The enzymatically formed N^{α}-Z-protected *E. coli* Par10 was isolated from the reaction mixture by RP-HPLC using gradient elution. The resulting pure fractions containing Par10 were concentrated and lyophilised overnight. After deprotection of the protein's *N*-terminus by TFA/thioanisole treatment over a time period of 3h at 25 °C according to the procedure of Kiso et al.[7] and a final chromatographic purification (HPLC) of the product, the fully deprotected lyophilised Par10 was re-suspended in a concentration of 4 mg/ml in 10 mM HEPES buffer pH 7.8 containing 1 mM EDTA, 5 mM DTT, and 4 M guanidinium hydrochloride for protein refolding. After 1 h incubation at room temperature, the refolding mixture was rapidly diluted 8-fold by adding pure 10 mM HEPES buffer pH 7.8. Finally, the resulting diluted protein mixture was dialysed against 10 mM HEPES buffer pH 7.8.

Determination of the k_{cat}/K_M-values of synthetic and *wt*-Par10. The specificity constants k_{cat}/K_M of the synthesised Par10 and its recombinantly expressed *wild-type* analogue have been measured by using the protease-coupled assay established by Fischer et al.[8] Accordingly, 33 μM chymotrypsin and 25 μM of the substrate Suc-Ala-Phe-Pro-Phe-4-NA were used. The stock solutions of chymotrypsin (0.2 mM) and the substrate (10 mM) were prepared in aqua bidest. and dimethyl sulfoxide, respectively. The reactions were carried out in 35 mM HEPES buffer pH 7.8, at 10 °C and followed spectrophotometrically at 390 nm (detection of liberated 4-nitroanilide) using a Hewlett Packard 8452A diode array UV/VIS spectrophotometer (USA). 500 data points were used to calculate the first-order rate constants. Three independent measurements were performed. The respective errors are given at the point of data appearance.

CD spectrometry. The isolated and lyophilised protein was dissolved in 10 mM phosphate buffer pH 7.8. CD spectra were recorded from 190 to 260 nm with 10 μM of the synthetic Par10. Similar measurements were performed with wt-Par10 obtained from recombinant expression. All measurements were done in a 1 mm cuvette at

20 °C with a J-710 spectropolarimeter (Jasco, Tokyo, Japan), averaging 20 transients with Jasco Software Version 1,33.00.

Analyses. Enzymatic reactions were analysed under optimised conditions by analytical reversed phase HPLC using a LiChrospher C18 column (25 × 0.4 cm, 5 µm, Merck, Germany) or a Capcell PAK C8 column (25 × 0.4 cm, 5 µm, Shiseido, Japan). Samples were eluted with various mixtures of acetonitrile/water containing 0.1% (v/v) trifluoroacetic acid under gradient conditions at flow rates of 1.0 mL min^{-1}. Detection was carried out at 254 nm or 220 nm. Reaction rates and product yields were calculated from peak areas of acyl donor esters, hydrolysis and aminolysis products, respectively. Peptide purification was performed on a HPLC system with 220 nm UV detection using a preparative RP C8 column (25 × 2.5 cm, 7 µm, Merck, Germany) and gradient elution at a flow rate of 15 mL min^{-1}. Mass spectra were recorded for separated and lyophilised probes using MALDI-ToF (MALDI 5 V5.1.2, Kratos Kompakt, Japan) or ESI (Apex II/7 Tesla, Bruker-Daltonics, Germany) ionisation. NMR (GEMINI 300, Varian, USA) and elementary analysis (Analyser CHN-O Rapid, Heraeus, Germany) were used to verify the identity of some reactants.

Abbreviations.

Ac, acetyl; ADP, alkaline D-peptidase; Boc, *tert*-butyloxycarbonyl; Bz, benzoyl; DBU, 1,8-Diaza-bicyclo[5.4.0]-7-undecen; DCC, dicyclohexyl carbodiimide; DIEA, *N, N*-diisopropylethylamine; DMF, *N,N*-dimethyl-formamide; DTT, dithiothreitol, EDTA, ethylenediaminetetraacetic acid; Fmoc, fluoren-9-ylmethoxycarbonyl; HCTU, 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate; HEPES, *N*-[2-hydroxyethyl]piperazine-*N'*-[2-ethanesulfonic acid]; 4-NA; 4-nitroanilide; NMM, 4-methylmorpholine; OBzl, benzyl ester; OCap, 4-carboxamidophenyl ester; OCp, 4-carboxyphenyl ester; OGp, 4-guanidinophenyl ester; OInd, 2-carboxy-1*H*-indol-5-yl ester; OMe, methyl ester; OPh, phenyl ester; PyBOP, (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate; SPhOH, 4-hydroxythiophenyl ester; TBTU, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate; TCEP, tris(2-carboxyethyl) phosphine; TFA, trifluoroacetic acid; TIS, triisopropylsilane; Tos, *p*-toluenesulfonic acid; Tricine, *N*-(2-hydroxy-1,1-bis[hydroxymethyl]ethyl) glycine; TRIS, 2-Amino-2-(hydroxymethyl)propane-1,3-diol; Suc, succinyl;
Z, benzyloxycarbonyl.

References.

[1] Y. Asano, H. Ito, T. Dairi, Y. Kato, *J. Biol. Chem.* **1996**, *271*, 30256-30262.

[2] J.-U. Rahfeld, A. Schierhorn, K. Mann, G. Fischer, *FEBS Lett.* **1994**, *343*, 65-69.

[3] R. Günther, C. Elsner, S. Schmidt, H.-J. Hofmann, F. Bordusa, *Org. Biomol. Chem.* **2004**, *2*, 1442-1446.

[4] M. Thormann, S. Thust, H.-J. Hofmann, F. Bordusa, *Biochemistry* **1999**, *38*, 6056-6062.

[5] G. W. Kenner, J. R. McDermott, R. C. Sheppard, *Chem. Commun.* **1971**, 636-637.

[6] a) B. J. Backes, A. A. Virgilio, J. A. Ellman, *J. Am. Chem. Soc.* **1996**, *118*, 3055-3065; b) B. J. Backes, J. A. Ellman, *J. Org. Chem.* **1999**, *64*, 2322-2330.

[7] Y. Kiso, K. Ukawa, T. Akita, *J. C. S. Chem. Comm.* **1980**, 101-102.

[8] a) G. Fischer, H. Bang, C. Mech, *Biomed. Biochim. Acta* **1984**, *43*, 1101-1111; b) G. Fischer, B. Wittmann-Liebold, K. Lang, T. Kiefhaber, F. X. Schmid, *Nature* **1989**, *337*, 476-478.