Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information

Expanding the Genetic Code with a Lysine Derivative Bearing an Enzymatically Removable Phenylacetyl Group

Marie Reille-Seroussi, Susanne V. Mayer, Wolfgang Dörner, Kathrin Lang, and Henning D. Mootz*

Supplementary Figures

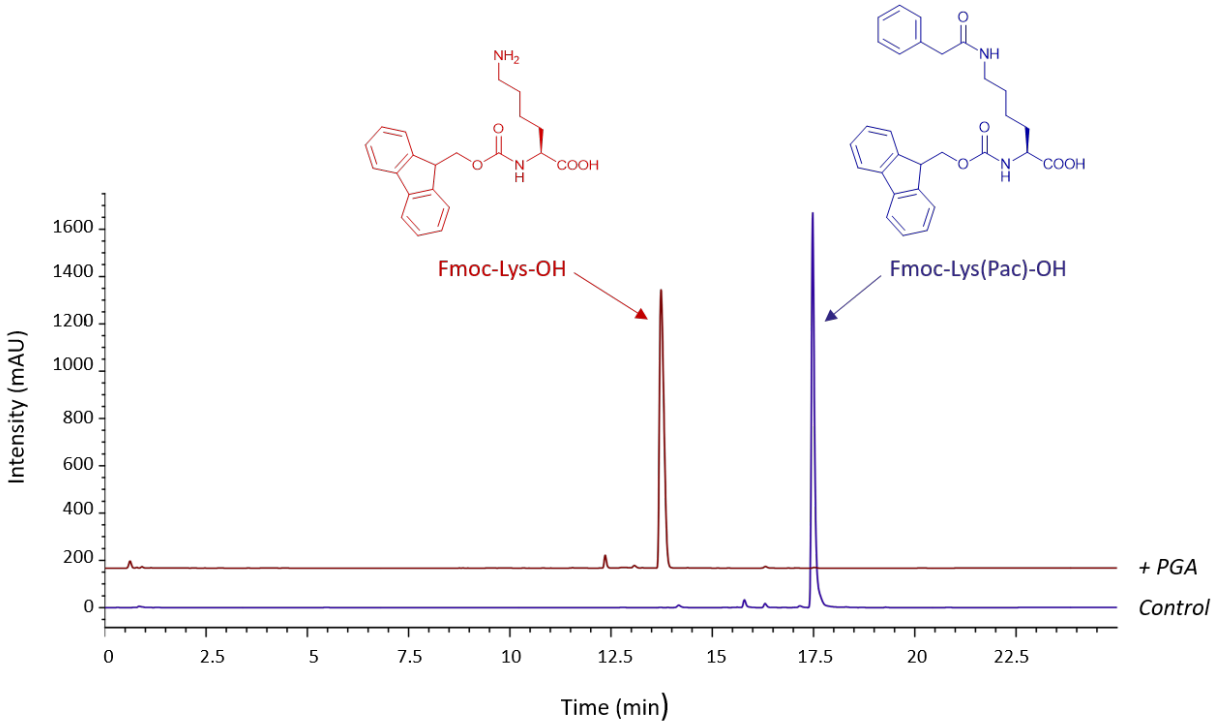
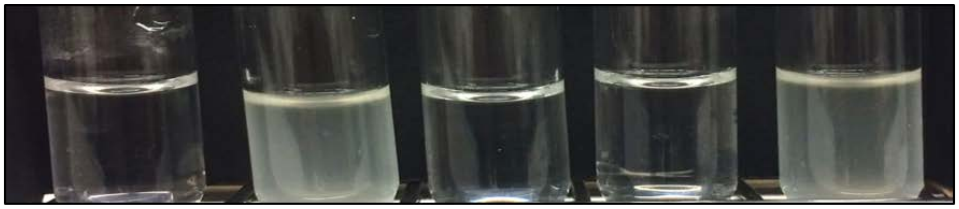
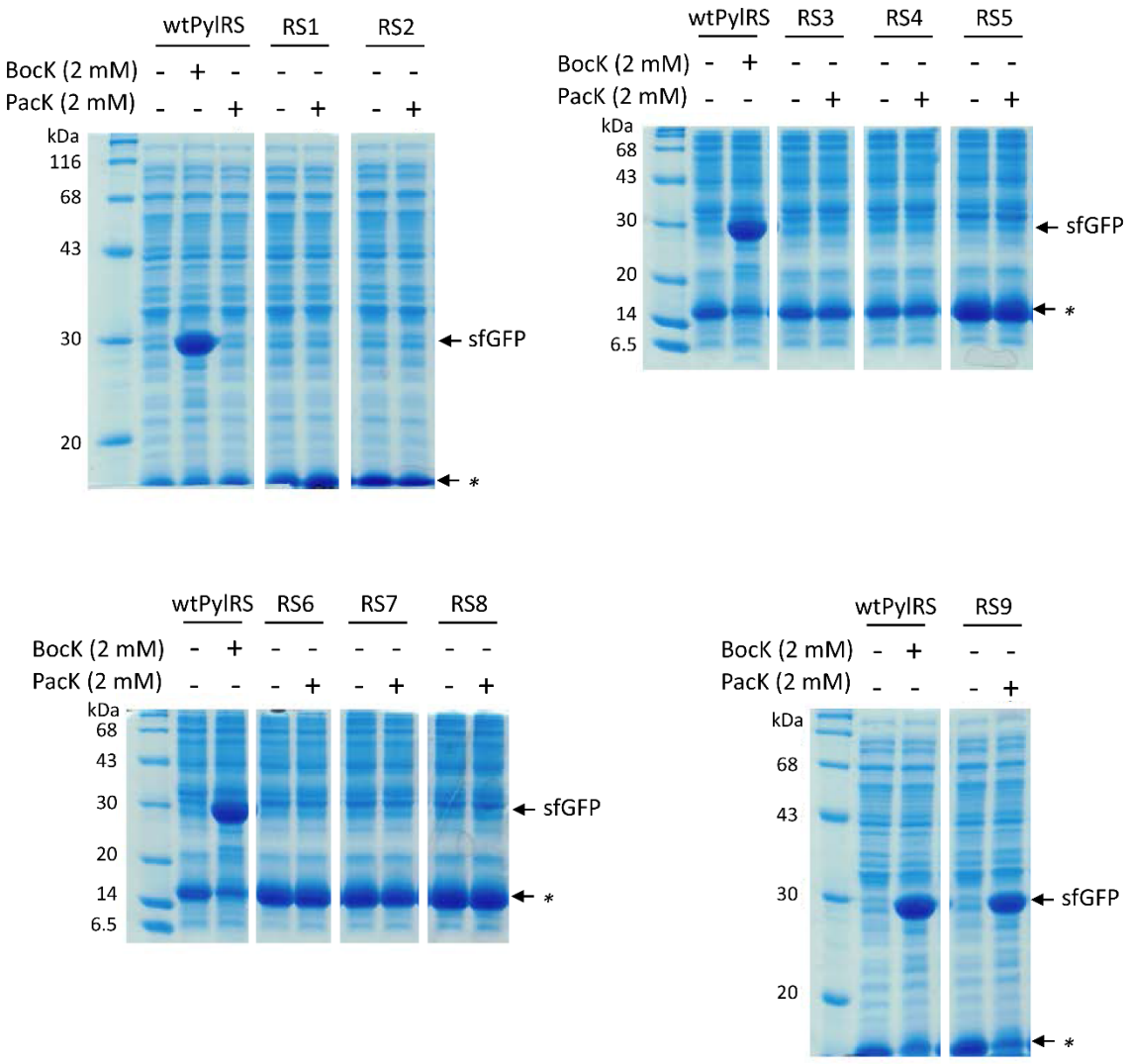


Figure S1 Enzymatic deprotection of the Fmoc-L-Lys(Pac)-OH by PGA. The reaction was performed at 37°C for 10 minutes using 0.5 U of soluble PGA. The shown HPLC analysis was performed using absorption at 280 nm and confirmed by MS analysis (data not shown).



	Minimal Media	Minimal Media + K (50µg/mL)	Minimal Media + PacK (100µg/mL)	Minimal Media + PacK (50µg/mL)	Minimal Media + PacK (50µg/mL) +K (50µg/mL)
OD (T0h)	0,04	0,06	0,04	0,04	0,06
OD (T7h)	0,04	0,18	0,04	0,04	0,13
OD (T21h)	0,03	0,69	0,04	0,04	0,52

Figure S2 Study of the metabolic stability of PacK in lysine auxotrophic *E. coli* strain JW2806-1. *E. coli* cells were cultivated at 37°C in minimal media either supplemented with Lys-OH and/or Lys(Pac)-OH or without these supplements. Shown the samples after 21h of culture growth and the respective OD600 values.



RS1 : L266M, L270I, Y271F, L274A, C313F
RS2 : C313S, Y349F
RS3 : C313T, Y349F
RS4 : L266V, L270I, Y271F, L274A, C313F
RS5 : L270I, Y271L, L274A, C313F
RS6 : L266M, Y271L, L274A, C313F
RS7 : L266M, Y271L, C313S
RS8 : L266M, L270I, Y271L, L274A, C313F
RS9 : Y271M, L274A, C313A

Figure S3 Screening of a series of existing PylRS mutants from *Methanosarcina* species for their capability to incorporate PacK (2mM) into a sfGFP(N150TAG) in *E. coli*. Incorporation of Nε-(tert-butoxycarbonyl)-L-lysine (BocK) with MbPylRS wt was chosen as a reference and is presented in all Coomassie-stained SDS-PAGE. * denotes the truncated sfGFP protein.

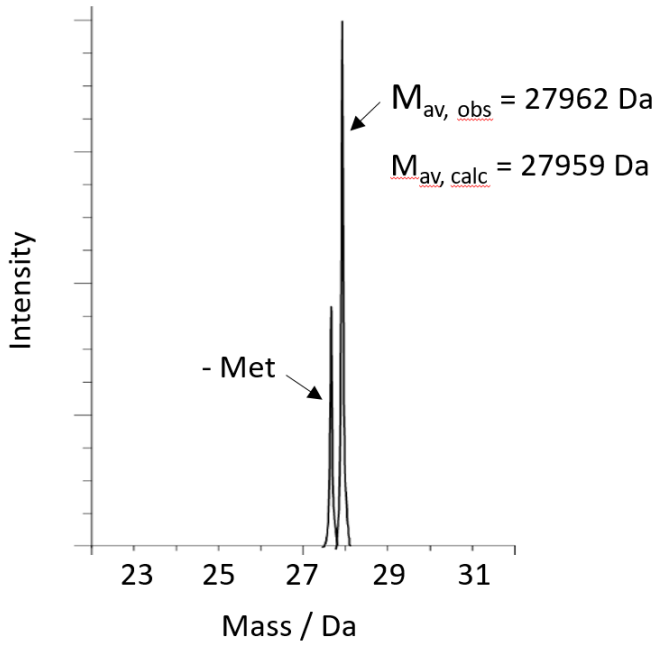


Figure S4 ESI-MS characterization of the purified sfGFP(N150PacK). The minor peak annotated on the spectrum corresponds to sfGFP(N150PacK) in which the N-terminal methionine has been cleaved.

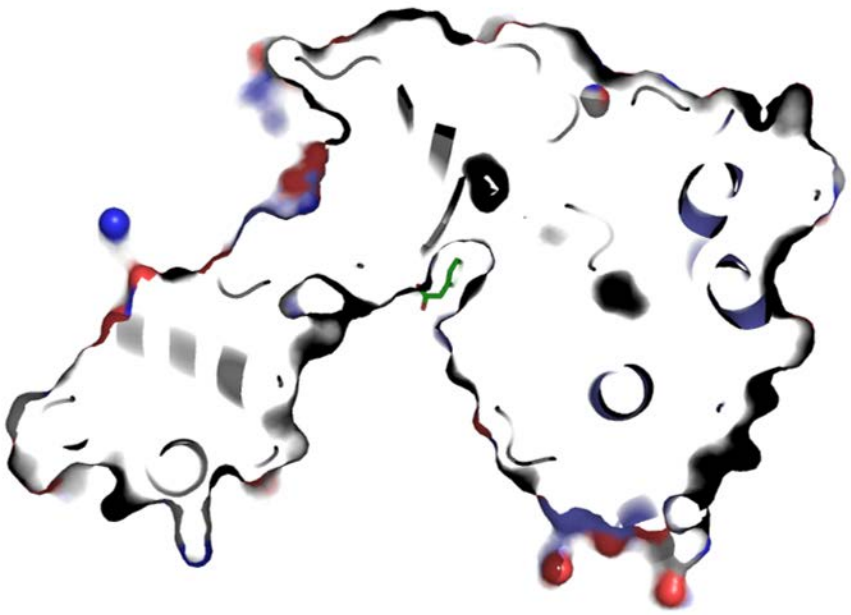


Figure S5 PGA crystal structure. Shown is a slice through the PGA enzyme. Bound phenylacetic acid (PAC) is shown in stick representation (green) at center of the protein and indicates the location of the active site. Note the deep and narrow channel leading to the active site. The illustration was generated using Pymol based on pdb file 1PNL.[1]

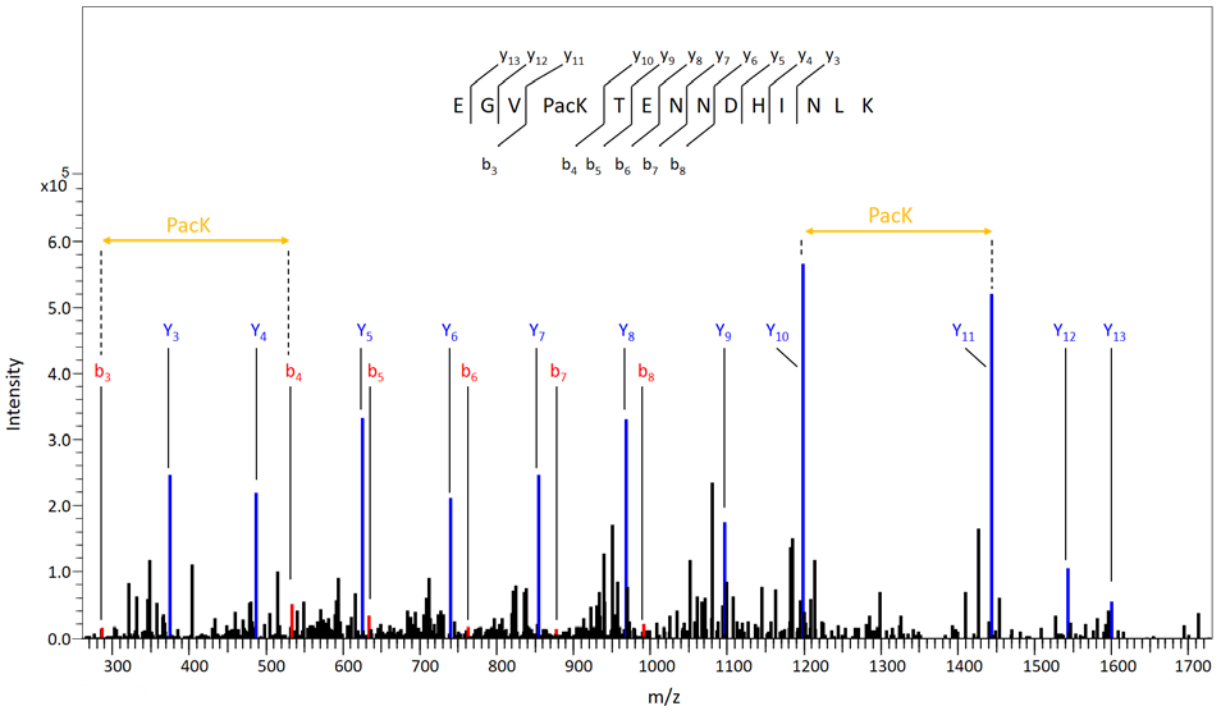


Figure S6 MS/MS-analysis of a tryptic fragment derived from the purified SUMO-2(K11PacK). The observed fragments confirmed the incorporation of PacK at position 11. Precursor ion: M_{obs} = 576.9582 Da; M_{calc} = 576.9566 Da (not shown).

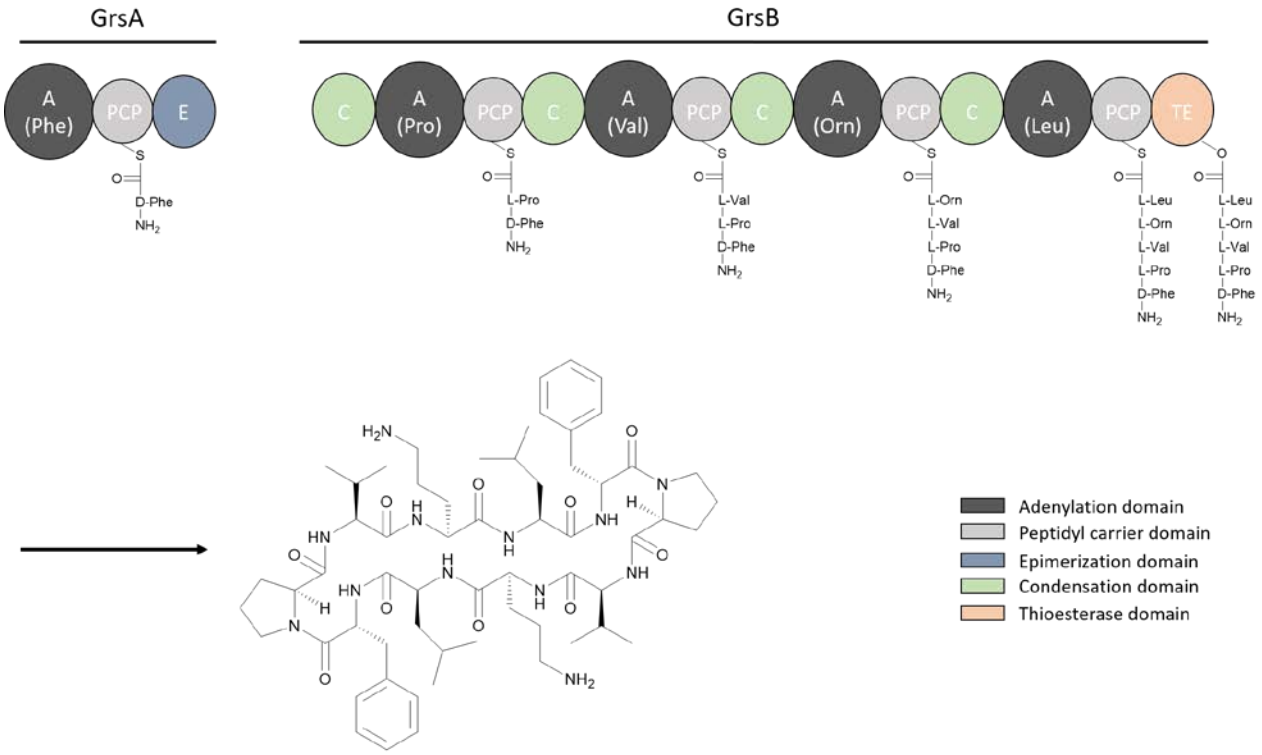


Figure S7 NRPS template and assembly-line model for gramicidin S. A pentamer is assembled on the five modules of the two synthetases and dimerized in an iterative manner.[2]

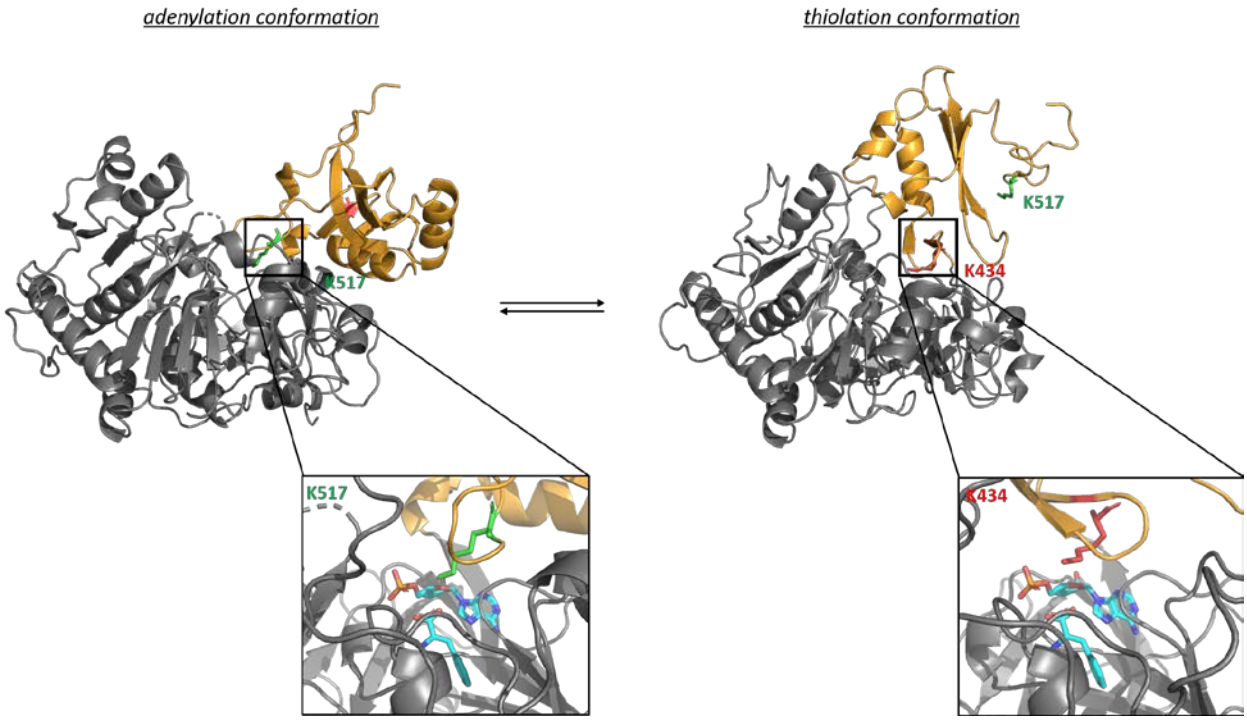


Figure S8 Localization of the two lysine residues K434 and K517 of GrsA in the domain alternation mechanism. Shown are the structures of the A-domain in adenylation (PDB 1AMU)[3] and thiolation conformation (model created by the phyre2 server[4] and based on PDB 4DG9).[5] The large N-terminal A^N domain is colored in grey and the small C-terminal A^C domain is colored in orange. The side chains of K434 and K517 are highlighted in red and green, respectively. The close views of the active site additionally show the bound L-Phe residue and AMP from the 1AMU structure in cyan stick representation.

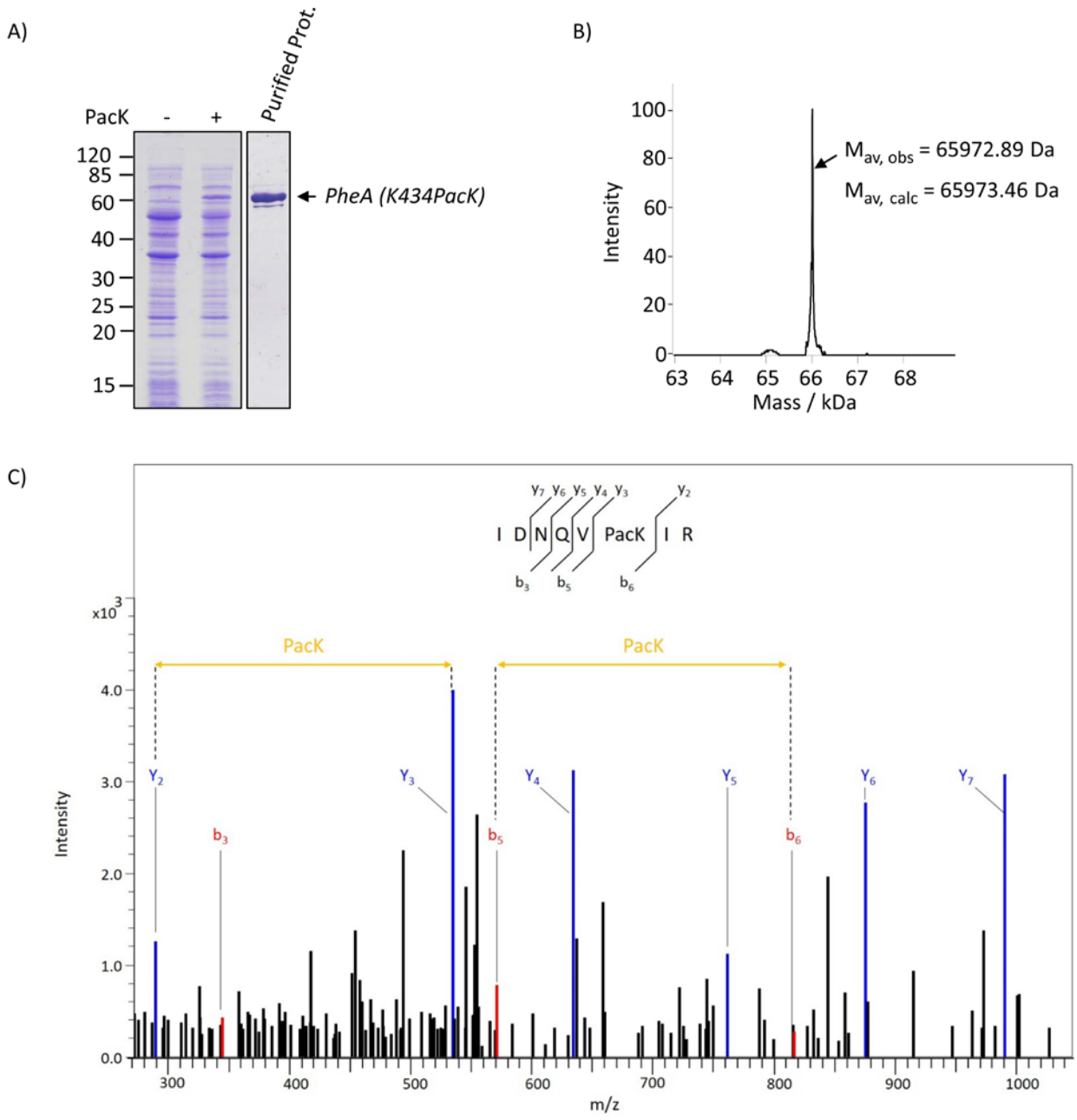


Figure S9 Preparation and analysis of the GrsA A domain with K434PacK. (A) Analysis of whole cell extracts and the purified fraction using a coomassie-stained SDS-PAGE for PheA(K434PacK)-His_6. (B) ESI-MS analysis. (C) MS/MS-analysis of a tryptic fragment derived from the purified PheA(K434PacK)-His_6. Precursor ion: M_{obs} = 552.3134 Da; M_{calc} = 552.3140 Da (not shown).

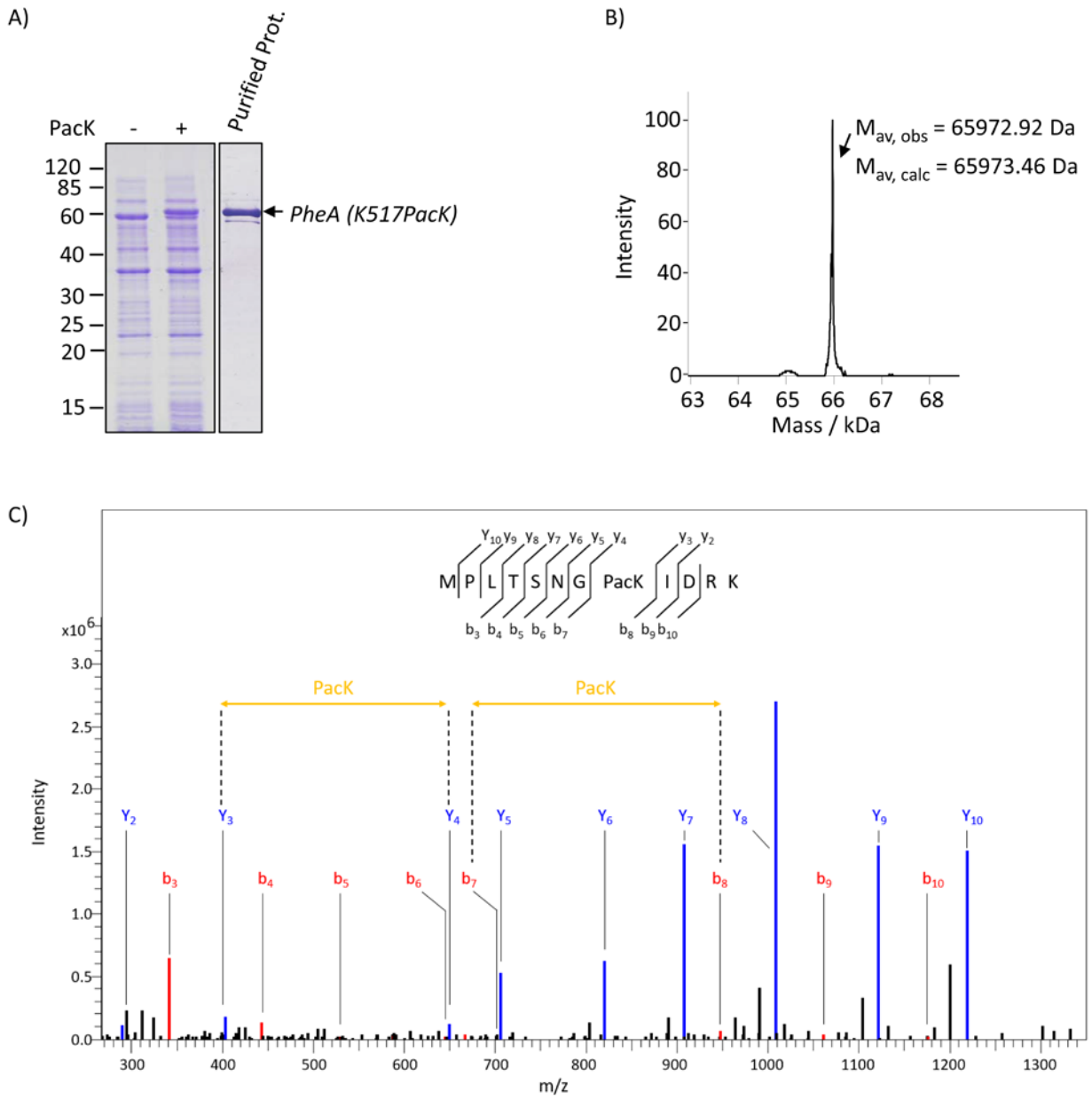


Figure S10 Preparation and analysis of the GrsA A domain with K517PacK. (A) Analysis of whole cell extracts and the purified fraction using a coomassie-stained SDS-PAGE for PheA(K517PacK)-His_6. (B) ESI-MS analysis. (C) MS/MS-analysis of a tryptic fragment derived from the purified PheA(K517PacK)-His_6. Precursor ion: M_{obs} = 498.5987 Da; M_{calc} = 498.5975 Da (not shown).

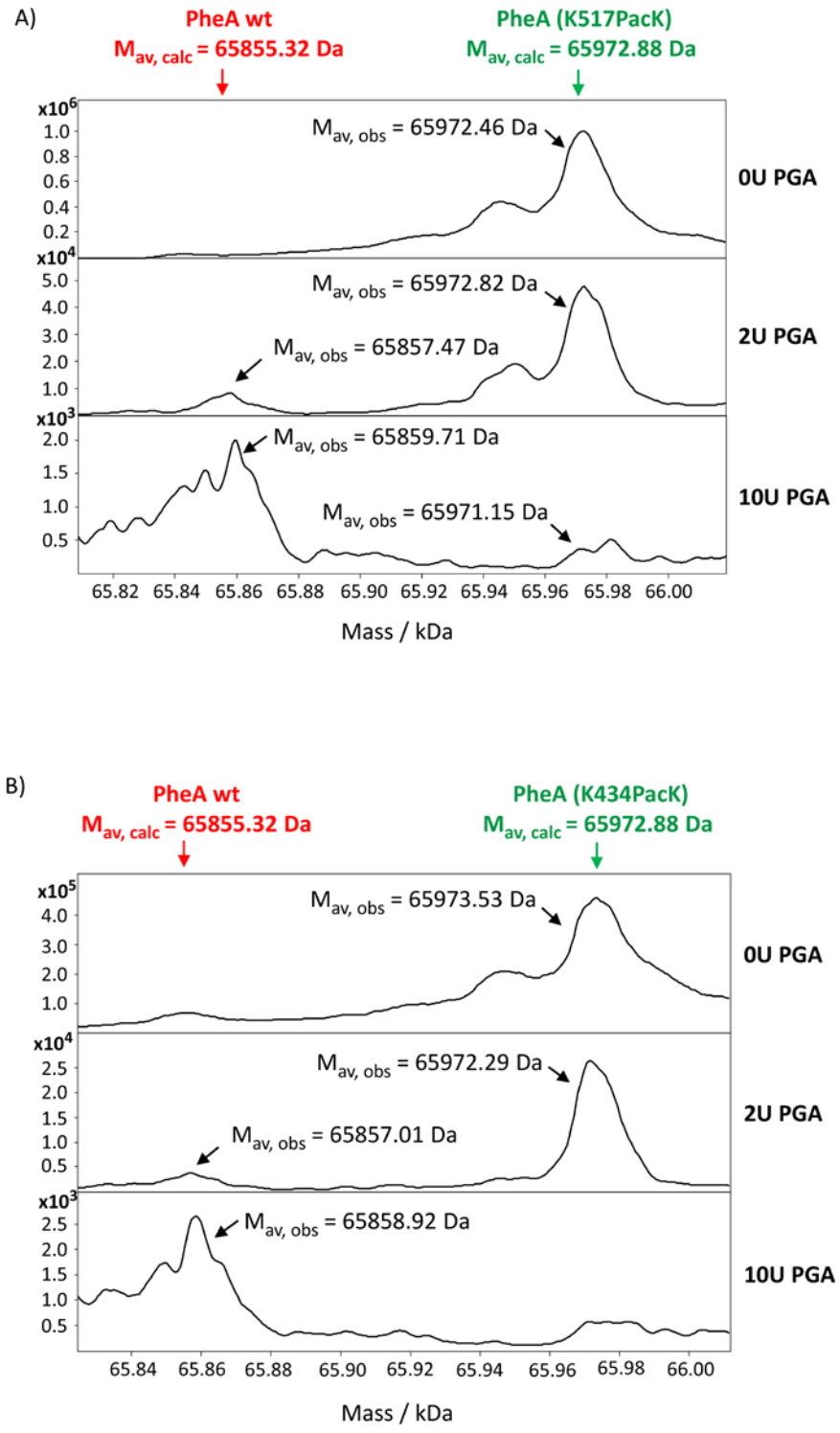


Figure S11 PGA deprotection of PacK incorporated in the PheA domain of GrsA. ESI-MS analysis of PGA-mediated deprotection of (A) the PheA(K517PacK) and (B) the PheA(K434PacK). Deprotection was performed overnight at 25°C. Note that incubation with PGA always resulted in diminished sensitivity of the MS-analyses.

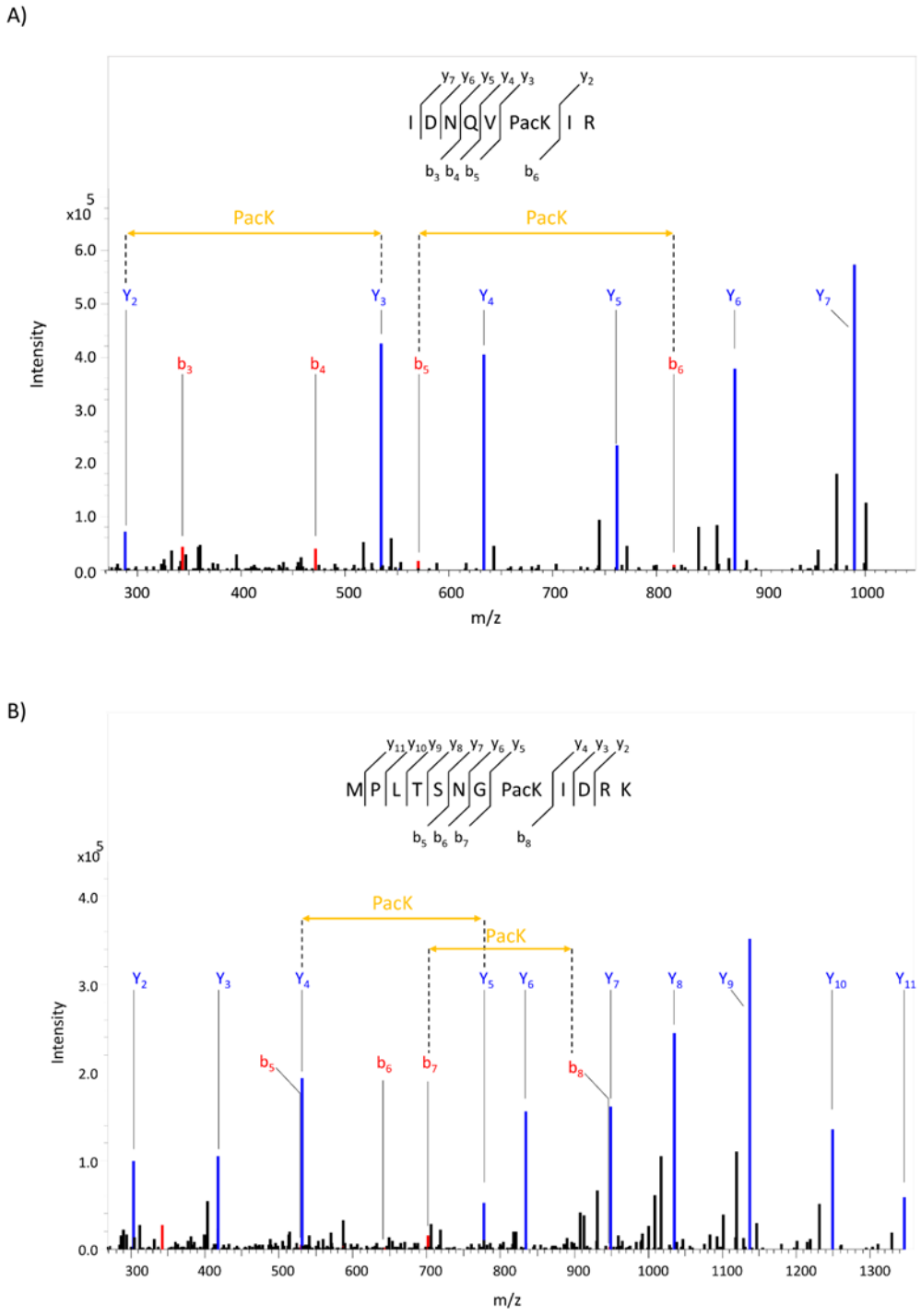


Figure S12 Analysis of PacK incorporation in GrsA mutants. (A) MS/MS-analysis of a tryptic fragment derived from purified GrsA(K434PacK). Precursor ion: M_{obs} = 552.3152 Da; M_{calc} = 552.3140 Da (Not shown). (B) MS/MS-analysis of a tryptic fragment derived from purified GrsA(K517PacK). Precursor ion: M_{obs} = 493.2670 Da; M_{calc} = 493.2659 Da (not shown).

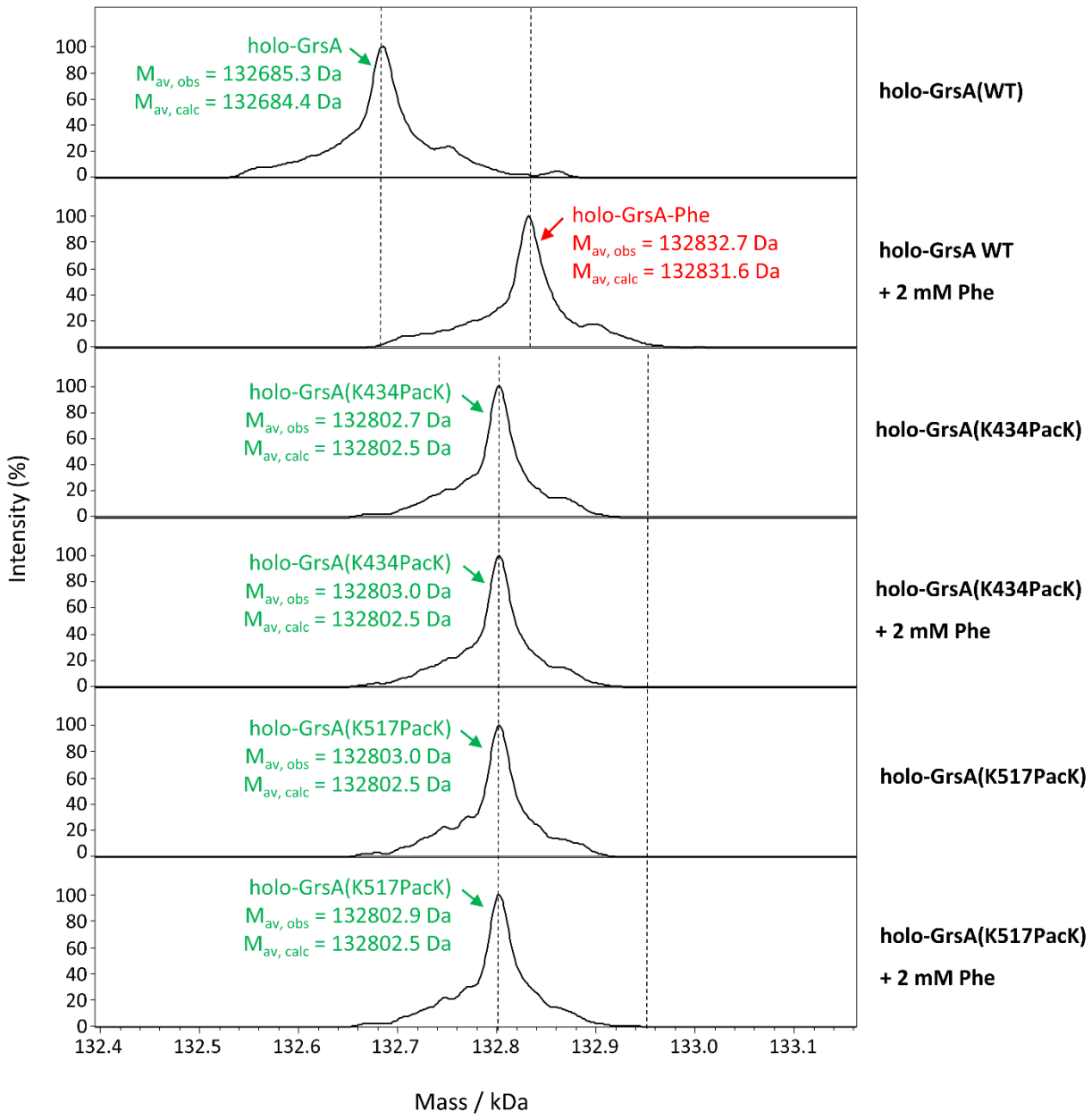


Figure S13 MS-analysis of the thioester formation of GrsA wild-type and GrsA(K434PacK) and GrsA(K517PacK) mutants. First, the proteins were modified with the 4'-phosphopantetheine moiety to obtain holo-GrsA WT, holo-GrsA(K434PacK) and holo-GrsA(K517PacK). Then, the reaction of Phe-thioester formation was performed at 25°C for 30 min using one of the GrsA proteins (each at 5 µM) with ATP (2 mM) and L-Phe (2 mM). Attachment of phenylalanine was detected only for GrsA wild-type.

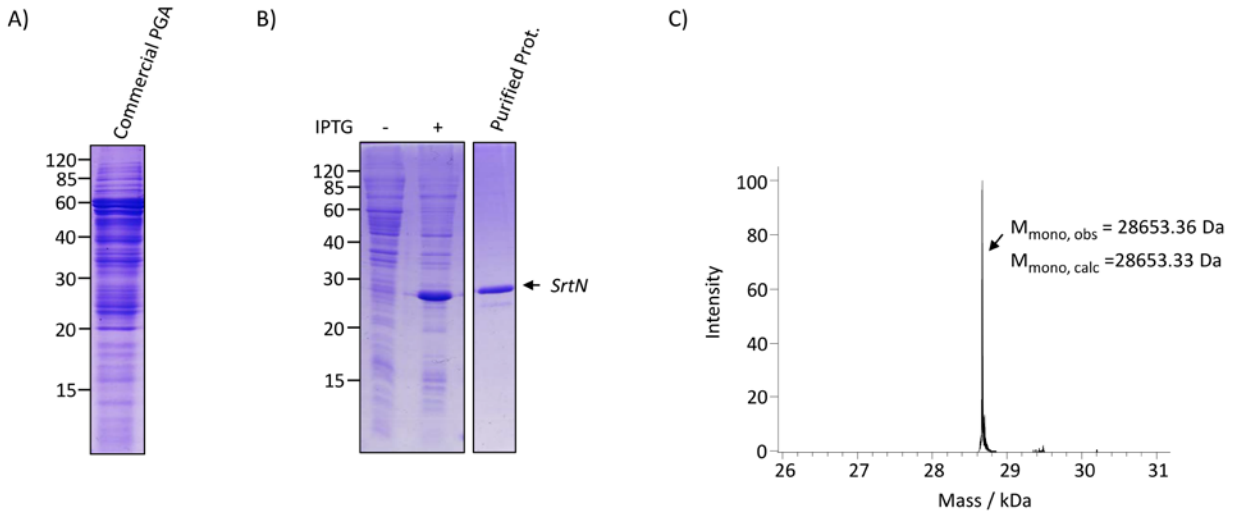


Figure S14 Analysis of PGA and SrtN. (A) Coomassie-stained SDS-PAGE of the commercially obtained PGA. The alpha and beta subunit have calculated masses of 23 kDa and 62 kDa. (B) Preparation and purification of SrtN. Shown is a coomassie-stained SDS-PAGE analysis of *E. coli* whole cell extracts and the purified protein. (C) ESI-MS analysis of purified SrtN.

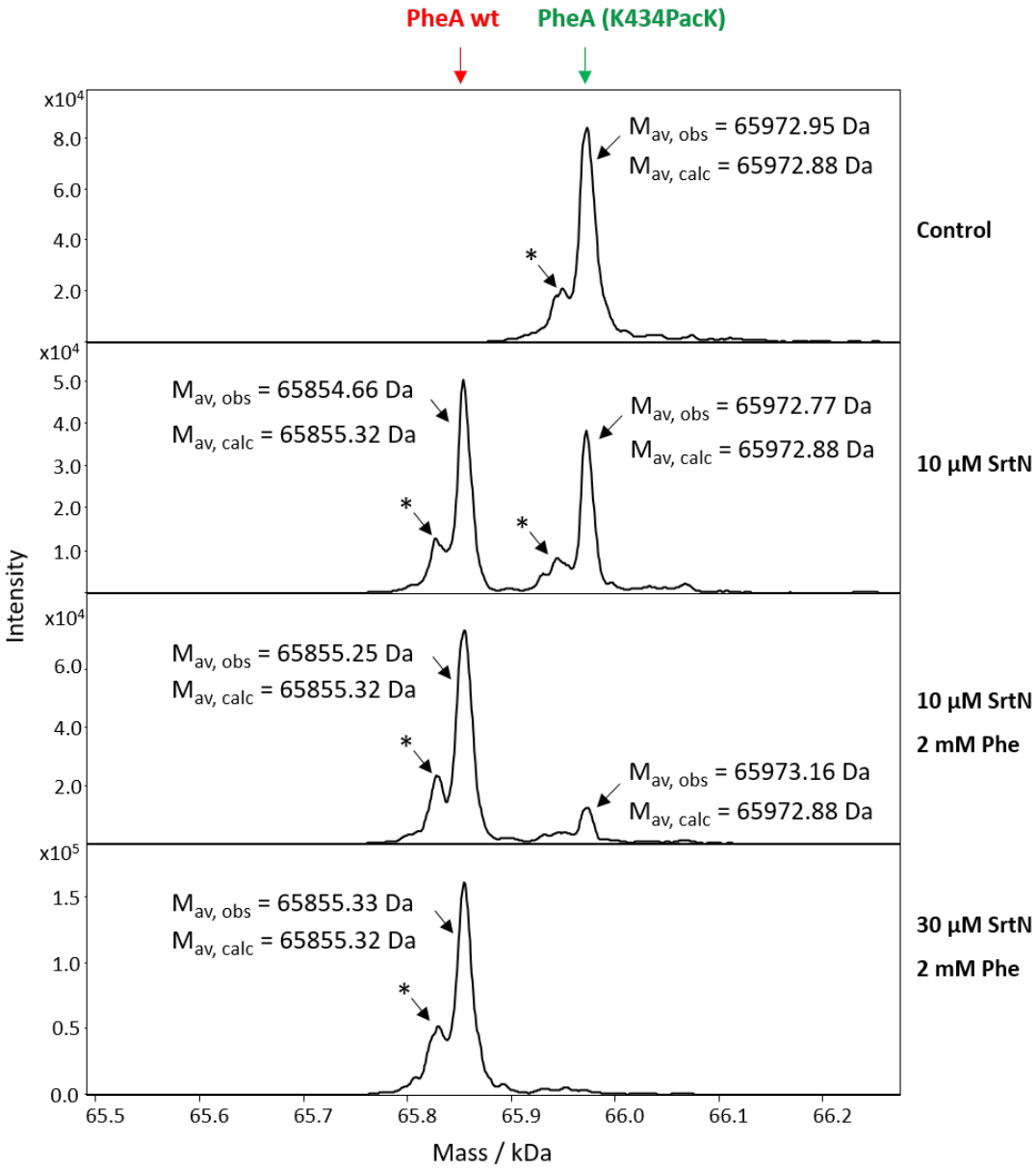


Figure S15 MS analysis of the enzymatic deprotection of PheA(K434PacK) by SrtN. The deprotection PheA(K434PacK)(10 µM) was performed ON at 25°C with different concentrations of SrtN. 2 mM phenylalanine was additionally used in some samples with the purpose to partially trigger the A conformation and thereby increase the exposition of the lysine K434 outside of the active site.[6] * denotes a non-identified impurity present in the sample (Δ = -27 Da).

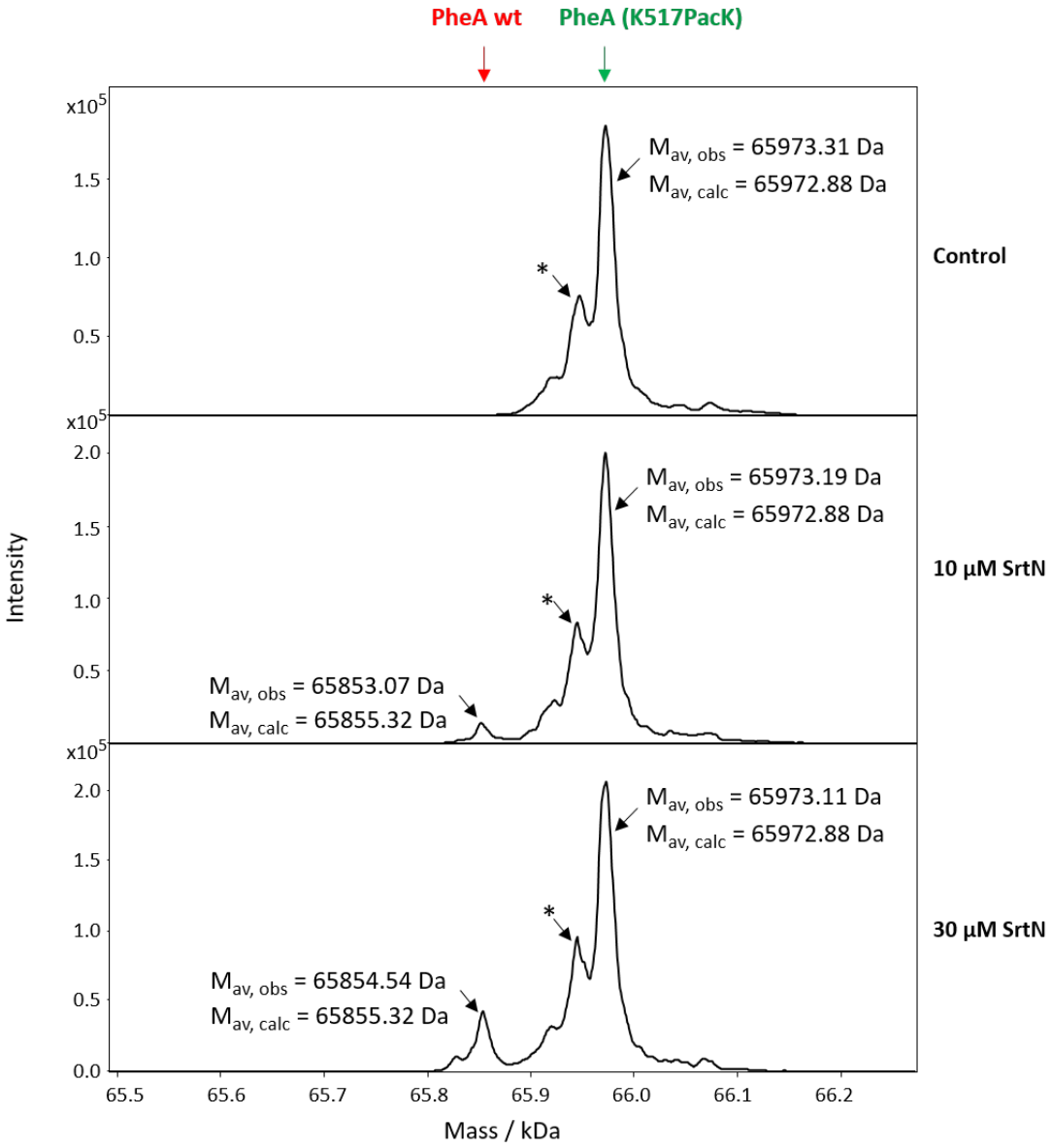


Figure S16 MS analysis of the enzymatic deprotection of PheA(K517PacK) by SrtN. The deprotection of A-domain(K517PacK) (10 µM) was performed ON at 25°C with different concentrations. * denotes a non-identified impurity present in the sample (Δ = -27Da).

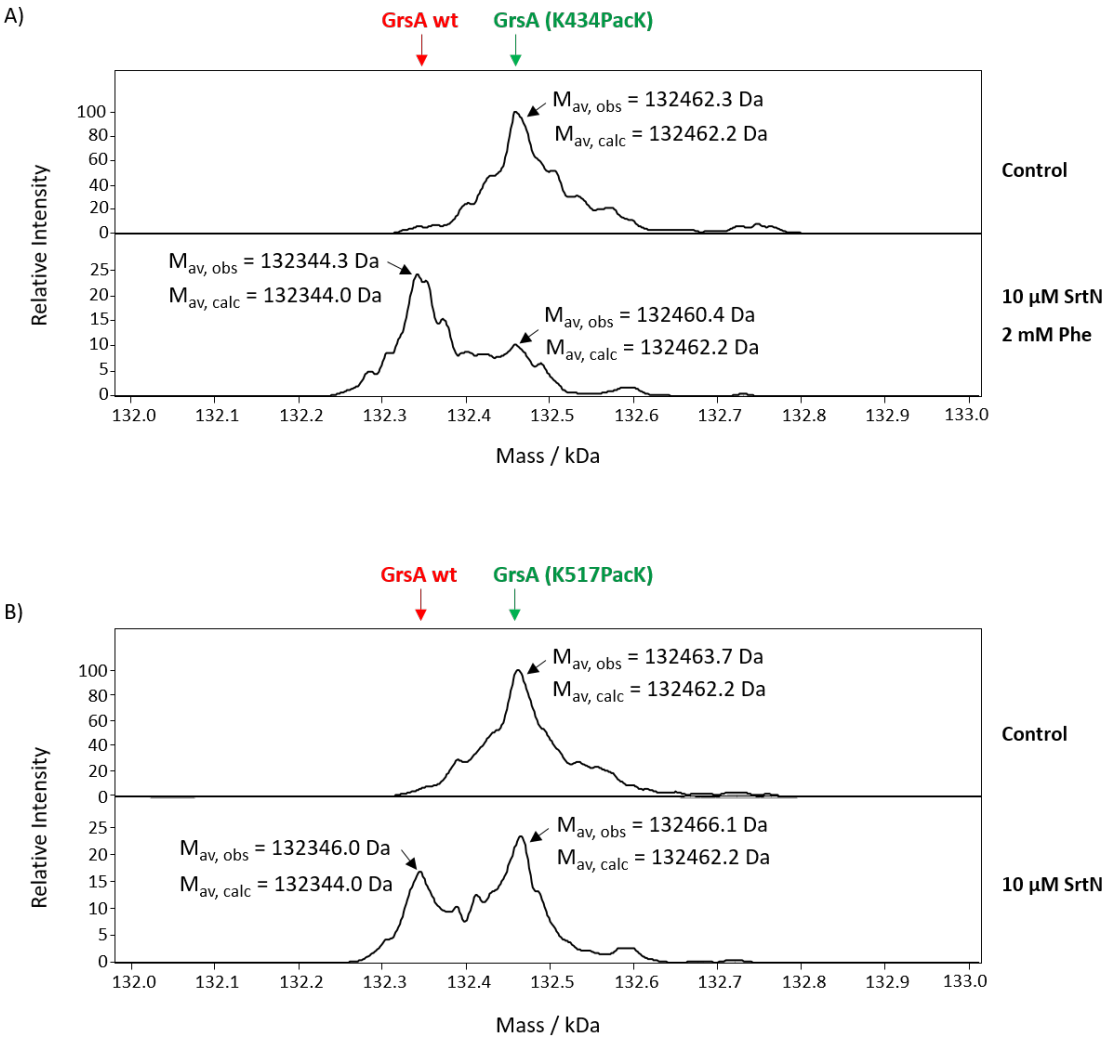
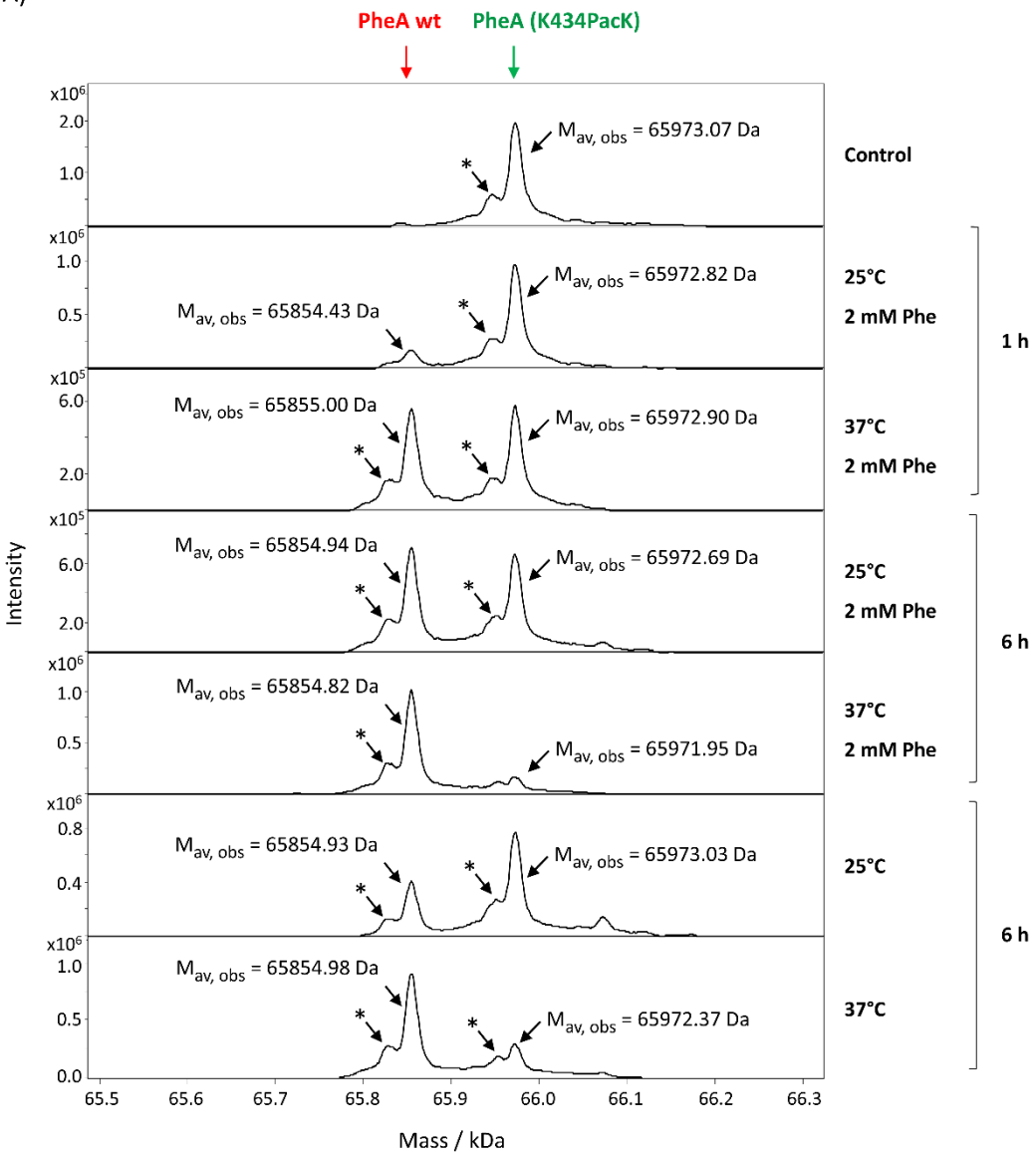


Figure S17 MS analysis of the enzymatic deprotection of GrsA mutants by SrtN. (a) Deprotection of GrsA(K434PacK) (3 µM) was performed ON at 25°C. (b) Deprotection of GrsA(K517PacK) (3 µM) was performed ON at 25°C.

A)


PheA wt
PheA (K434PacK)
Intensity
Mass / kDa
65.5
65.6
65.7
65.8
65.9
66.0
66.1
66.2
66.3
x10^6
2.0
1.0
M av, obs = 65973.07 Da
Control
1.0
0.5
M av, obs = 65854.43 Da
M av, obs = 65972.82 Da
25°C
2 mM Phe
1 h
x10^5
6.0
2.0
M av, obs = 65855.00 Da
M av, obs = 65972.90 Da
37°C
2 mM Phe
x10^5
6.0
2.0
M av, obs = 65854.94 Da
M av, obs = 65972.69 Da
25°C
2 mM Phe
6 h
x10^6
1.0
0.5
M av, obs = 65854.82 Da
M av, obs = 65971.95 Da
37°C
2 mM Phe
x10^6
0.8
0.4
M av, obs = 65854.93 Da
M av, obs = 65973.03 Da
25°C
6 h
x10^6
1.0
0.5
0.0
M av, obs = 65854.98 Da
M av, obs = 65972.37 Da
37°C

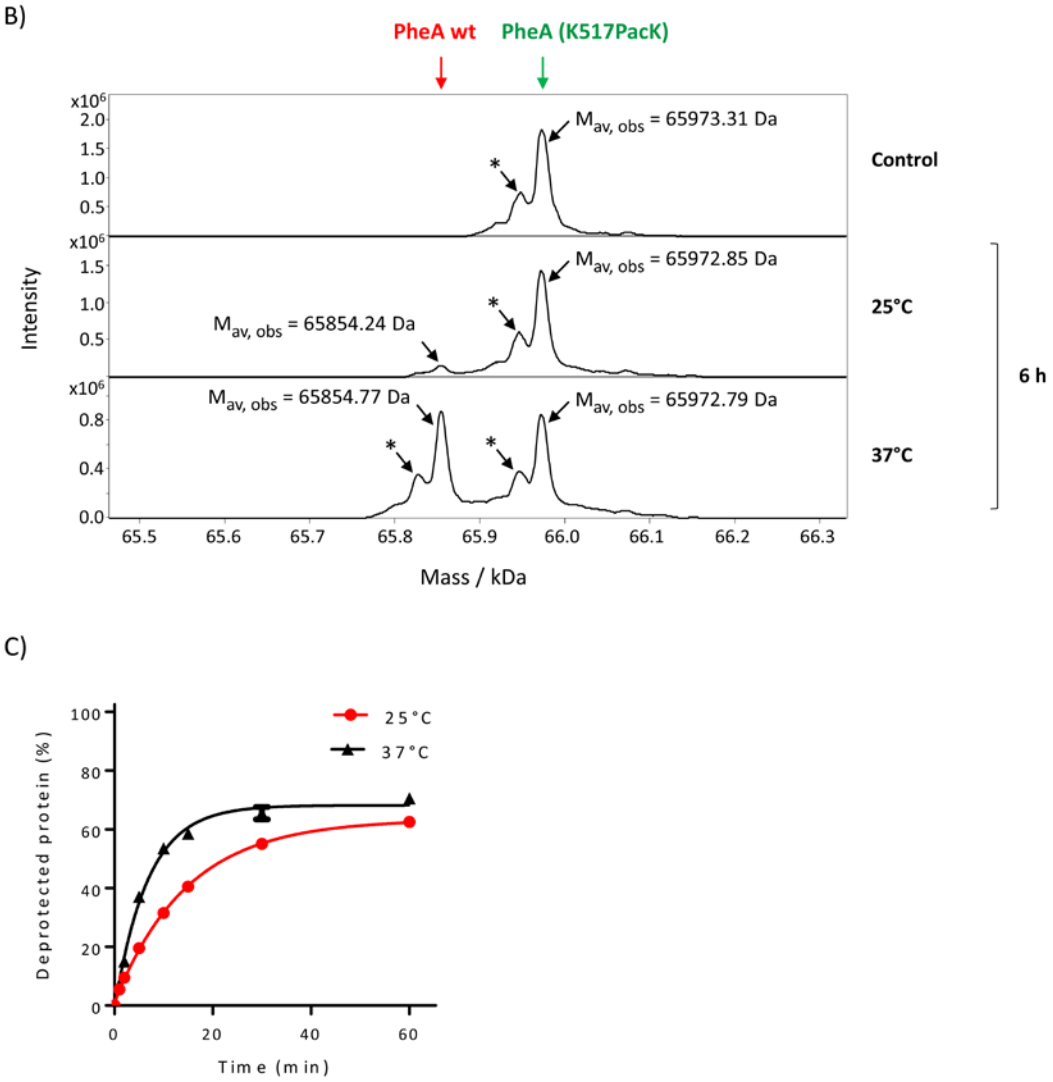


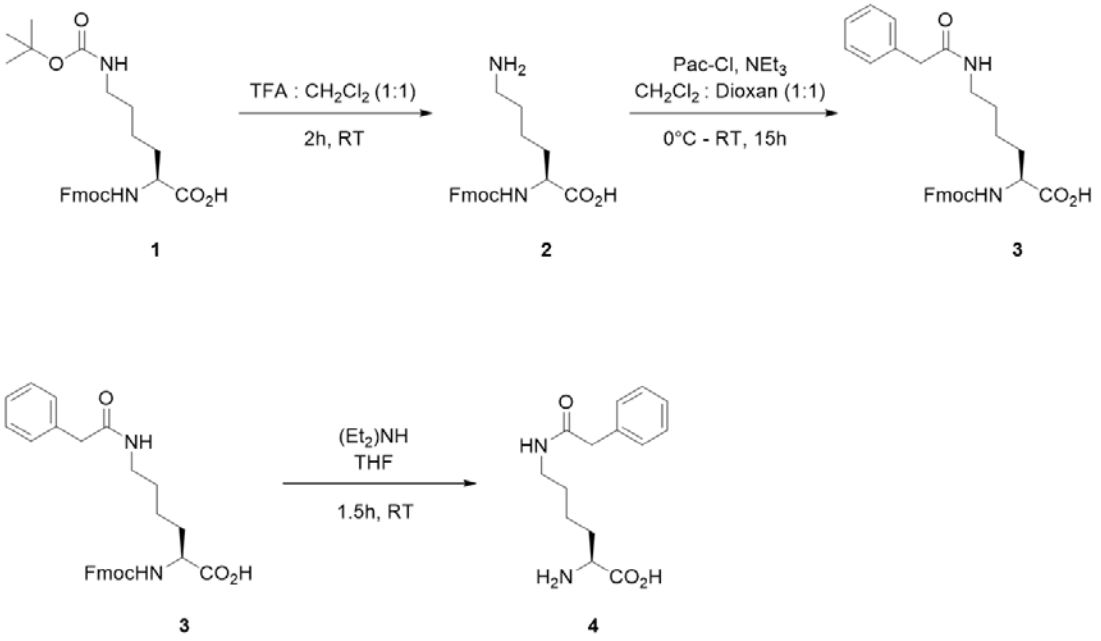
Figure S18 Influence of the temperature on SrtN-mediated PacK deprotection. A) MS-analysis of the enzymatic deprotection of PheA(K434PacK) by SrtN. The deprotection was performed at 25°C and 37°C for a period of 1h and 6h using an equimolar amount of PheA(K434PacK)(10 µM) and SrtN (10 µM). 2 mM phenylalanine was additionally used in some samples with the purpose to partially trigger the A conformation and thereby increase the exposition of the lysine K434 outside of the active site. PheA(K434PacK) : Mav, calc = 65972.88. B) MS-analysis of the enzymatic deprotection of PheA(K517PacK) by SrtN. The deprotection reaction was performed at 25°C and 37°C for a period of 6h using an equimolar amount of PheA(K517PacK)(10 µM) and SrtN (10 µM). * denotes a non-identified impurity present in the sample (Δ = -27 Da). PheA(K517PacK) : Mav, calc = 65972.88. C) MS-analysis of the enzymatic deprotection of SUMO(K11PacK) by SrtN. Time course analysis of the deprotection at 25°C and 37°C using a sub-equimolar amount of enzyme ([Sumo(K11PacK)] = 5 µM; [SrtN] = 0.5 µM). The rate constants for the reaction at 25°C and 37°C were 0.065 min^{-1} and 0.137 min^{-1}, respectively. A lower concentration of SrtN than the one used in the Figure 4 (main text) was used by purpose to better observe the difference of reaction speed obtained with the two temperatures. The results of two experiments are presented, error bars represent standard deviations.

Material and Methods

General

Solvents and standard chemical reagents were purchased from Sigma, Acros Organics, TCI, Alfa Aesar, Iris or Merck and were used without further purification. NMR spectra were recorded on a Bruker spectrometer (300 or 400 MHz). Chemical shifts (δ), reported in ppm, were internally referenced to residual protonated solvent signals. The abbreviations used for signal multiplicity are the following: s (singlet), d (doublet), t (triplet), m (multiplet or massif), br (broad signals). Coupling constants (J) are reported in Hertz (Hz). Oligonucleotides were purchased from Biolegio. Restriction enzymes were from Thermo Scientific. Plasmids were verified by DNA sequencing by Seqlab (Göttingen, Germany). If not further specified, buffer reagents, antibiotics, media components and other compounds used for cell culture or biochemical assays were purchased from Carl Roth, Applichem, Sigma Aldrich, Fluka or Jena Bioscience. Ni-NTA agarose beads were ordered from Cube Biotech and Chitin Resin from New England Biolabs. Error bars represent standard deviations from at least three independent experiments.

Chemical Synthesis



Scheme S1 Synthetic route for N6-(2-phenylacetyl)-L-lysine (PacK, 4).

(((9H-fluoren-9-yl)methoxy)carbonyl)-L-lysine (2)

2 was synthesized as previously described[7] and obtained as a white powder (1.9 g, 89%).

^{1}H NMR (400 MHz, DMSO-*d6*): δ (ppm) = 1.34–1.43 (m, 2 H, CH2), 1.51–1.77 (m, 4 H, 2CH2), 2.76–2.80 (t, J = 7.6, 2 H, CH2), 3.90–3.96 (m, 1 H, CH), 4.21–4.26 (m, 1 H, CH), 4.29–4.34 (m, 2 H, CH2), 7.33 (t, J = 7.5 Hz, 2 H, 2CH), 7.42 (t, J = 7.5 Hz, 2 H, 2CH), 7.60-7.62 (m, 1 H, CH), 7.71–7.74 (dd, J = 7.5 and 3.4 Hz, 1 H, CH), 7.83 (br. , 1 H, CONH), 7.89 (d, J = 7.5 Hz, 2 H, 2CH)
^{13}C NMR (100 MHz, DMSO-*d6*): δ (ppm) = 23.02 (CH2), 26.98 (CH2), 30.68 (CH2), 39.03 (CH2), 47.12 (CH), 54.13 (CH), 66.03 (CH2), 120.57 (2CH), 125.71 (2CH), 127.53 (2CH), 128.10 (2CH), 141.17 (2C), 144.24 (C), 144.29 (C), 156.67 (CONH), 174.34 (COOH)
MS (ESI+): $[M+H]^+$ m/z_{obs} = 369.2 and m/z_{calc} = 369.4

N^2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N^6-(2-phenylacetyl)-L-lysine (3)

3 was synthesized by adaptation of a previously reported protocol.[8] Fmoc-Lys-OH (**2**) (1.3 g, 2.6 mmol) was dissolved in a mixture CH_2Cl_2 / 1,4-dioxane (1:1, 25 ml) and cooled to 0 °C under nitrogen. Triethylamine (1.5 ml, 10.6 mmol) and phenylacetyl chloride (0.7 ml, 5.3 mmol) was added drop-wise. The reaction mixture was stirred at room temperature overnight. Solvents were concentrated under vacuum. The residue was partitioned between EtOAc and KHSO4 1M. The organic phase was washed with 10% NaCl, dried over anhydrous sodium sulfate, filtered and concentrated under reduced pressure. Purification by chromatography on a silica gel column (EtOAc/Cyclohexan/HCOOH, 80:20:1) yielded to Fmoc-L-Lys(Pac)-OH (**3**) as a white powder (700 mg, 54%).

^{1}H NMR (300 MHz, CDCl3): δ (ppm) = 1.25–1.50 (m, 4H, 2CH2), 1.56–1.90 (m, 2H, CH2), 3.19–3.24 (m, 2H, CH2), 3.61 (s, 2H, CH2), 4.16-4.50 (m, 4H, 2CH+CH2), 7.21-7.35 (m, 9H, CH), 7.53-7.65 (m, 2H, CH), 7.75-7.78 (m, 2H, CH)
MS (ESI+): $[M+H]^+$ m/z_{obs} = 487.2 and m/z_{calc} = 487.6

N^6-(2-phenylacetyl)-L-lysine (L-Lys(Pac)-OH, 4)

Fmoc-L-Lys(Pac)-OH (**3**; 500 mg, 1.0 mmol) was dissolved in a solution THF / 20% Diethylamine (20 ml) and stirred at room temperature for 2 hours. After completion, the reaction mixture was concentrated under reduced pressure. The residue was washed with acetonitrile and concentrated under reduced pressure. Finally, the product was precipitated in cold diethyl ether to yield to the L-Lys(Pac)-OH (**4**) as a white powder (272 mg, 73%).

Remark: To improve the solubility, the TFA salt can be prepared by stirring L-Lys(Pac)-OH (**4**) in a solution TFA/CH_2Cl_2.

^{1}H NMR (300 MHz, D_2O): δ (ppm) = 1.20–1.28 (m, 2H, CH2), 1.37–1.47 (m, 2H, CH2), 1.61-1.77 (m, 2H, CH2), 3.09 (t, J = 6.8, 2H, CH2), 3.47 (s, 2H, CH2), 3.59 (t, J = 6.2, 1H, CH), 7.19-7.30 (m, 5H, CH)
^{13}C NMR (75 MHz, D_2O-*d6*): δ (ppm) = 21.54 (CH2), 27.80 (CH2), 29.45 (CH2), 38.86 (CH2), 42.43 (CH2), 54.19 (CH), 127.25 (CH), 128.89 (2CH), 129.01 (2CH), 135.06 (C), 172.70 (CONH), 174.58 (COOH)
MS (ESI+): $[M+H]^+$ m/z_{obs} = 265.2 and m/z_{calc} = 265.3

Biochemical methods:

Table S1 Expression plasmids and amino acid sequences of proteins used in this study

Protein	Expression plasmid	Vector (antibiotic resistance)	Sequence (UAA underlined)	Ref.
SUMO-2(K11PacK)-Int-CBD : TEV site-His_6-SUMO2(K11PacK ; Q98R)-GyrA Intein-CBD	pPM01	pGEX-4T-1 (Amp)	The sequence of SUMO after IMPACT purification is highlighted in bold: **MGENLYFQGGSSHHHHHHGADEKPKEGVPacKTENNDHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFRQQTGG**CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ	this work
GrsA- WT: SBP-GrsA-His_6	pGV196	pET28a (Kan)	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREPGASMLNSSKSILIHAQNKNGTHEEEQYLFAVNNTKAEYPRDKTIHQLFEEQVSKRPNNVAIVCENEQLTYHELNVKANQLARIFIEKGIGKDTLVGIMMEKSIDLFIGILAVLKAGGAYVPIDIEYPKERIQYILDDSQARMLLTQKHLVHLIHNIQFNGQVEIFEEDTIKIREGTNLHVPSKSTDLAYVIYTSGTTGNPKGTMLEHKGISNLKVFFENSLNVTEKDRIGQFASISFDASVWEMFMALLTGASLYIILKDTINDFVKFEQYINQKEITVITLPPTYVVHLDPERILSIQTLITAGSATSPSLVNKWKEKVTYINAYGPTETTICATTWVATKETIGHSVPIGAPIQNTQIYIVDENLQLKSVGEAGELCIGGEGLARGYWKRPELTSQKFVDNPFVPGEKLYKTGDQARWLSDGNIEYLGRIDNQVKIRGHRVELEEVESILLKHMYISETAVSVHKDHQEQPYLCAYFVSEKHIPLEQLRQFSSEELPTYMIPSYFIQLDKMPLTSNGKIDRKQLPEPDLTFGMRVDYEAPRNEIEETLVTIWQDVLGIEKIGIKDNFYALGGDSIKAIQVAARLHSYQLKLETKDLLKYPTIDQLVHYIKDSKRRSEQGIVEGEIGLTPIQHWFFEQQFTNMHHWNQSYMLYRPNGFDKEILLRVFNKIVEHHDALRMIYKHHNGKIVQINRGLEGTLFDFYTFDLTANDNEQQVICEESARLQNSINLEVGPLVKIALFHTQNGDHLFMAIHHLVVDGISWRILFEDLATAYEQAMHQQTIALPEKTDSFKDWSIELEKYANSELFLEEAEYWHHLNYYTENVQIKKDYVTMNNKQKNIRYVGMELTIEETEKLLKNVNKAYRTEINDILLTALGFALKEWADIDKIVINLEGHGREEILEQMNIARTVGWFTSQYPVVLDMQKSDDLSYQIKLMKENLRRIPNKGIGYEIFKYLTTEYLRPVLPFTLKPEINFNYLGQFDTDVKTELFTRSPYSMGNSLGPDGKNNLSPEGESYFVLNINGFIEEGKLHITFSYNEQQYKEDTIQQLSRSYKQHLLAIIEHCVQKEDTELTPSDFSFKELELEEMDDIFDLLADSLTGSRSHHHHHH	ref[9]

GrsA (K434PacK): SBP-GrsA(K434PacK)-His_6	pED123	pET28a (Kan)	GrsA-WT with K434PacK mutation	this work
GrsA (K517PacK): SBP-GrsA(K517PacK)-His_6	pED124	pET28a (Kan)	GrsA-WT with either K517PacK or K517NPocK mutations	this work
GrsA (K517A): SBP-GrsA(K517A)-His_6	pED139	pET28a (Kan)	GrsA-WT with K517A mutation	this work
A-domain (K434PacK): SBP-PheA(K434PacK)-His_6	pMR28	pET28a (Kan)	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQG QREPGASMLNSSKSILIHAQNKNGTHEEEQYLFAVN NTKAEYPRDKTIHQLFEEQVSKRPNNVAIVCENEQL TYHELNVKANQLARIFIEKGIGKDTLVGIMMEKSIDL FIGILAVLKAGGAYVPIDIEYPKERIQYILDDSQARML LTQKHLVHLIHNIQFNGQVEIFEEDTIKIREGTNLHVP SKSTDLAYVIYTSGTTGNPKGTMLEHKGISNLKVFFE NSLNVTEKDRIGQFASISFDASVWEMFMALLTGASL YIILKDTINDFVKFEQYINQKEITVITLPPTYVVHLDPE RILSIQTLITAGSATSPSLVNKWKEKVTYINAYGPTET TICATTWVATKETIGHSVPIGAPIQNTQIYIVDENLQL KSVGEAGELCIGGEGLARGYWKRPELTSQKFVDNPF VPGEKLYKTGDQARWLSDGNIEYLGRIDNQVPacKI RGHRVELEEVESILLKHMYISETAVSVHKDHQEQPYL CAYFVSEKHIPLEQLRQFSSEELPTYMIPSYFIQLDKM PLTSNGKIDRKQLPEPDLTFGSRSHHHHHH	this work
A-domain (K517PacK): SBP-PheA(K517PacK)-His_6	pMR29	pET28a (Kan)	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQG QREPGASMLNSSKSILIHAQNKNGTHEEEQYLFAVN NTKAEYPRDKTIHQLFEEQVSKRPNNVAIVCENEQL TYHELNVKANQLARIFIEKGIGKDTLVGIMMEKSIDL FIGILAVLKAGGAYVPIDIEYPKERIQYILDDSQARML LTQKHLVHLIHNIQFNGQVEIFEEDTIKIREGTNLHVP SKSTDLAYVIYTSGTTGNPKGTMLEHKGISNLKVFFE NSLNVTEKDRIGQFASISFDASVWEMFMALLTGASL YIILKDTINDFVKFEQYINQKEITVITLPPTYVVHLDPE RILSIQTLITAGSATSPSLVNKWKEKVTYINAYGPTET TICATTWVATKETIGHSVPIGAPIQNTQIYIVDENLQL KSVGEAGELCIGGEGLARGYWKRPELTSQKFVDNPF VPGEKLYKTGDQARWLSDGNIEYLGRIDNQVKIRG HRVELEEVESILLKHMYISETAVSVHKDHQEQPYLCA YFVSEKHIPLEQLRQFSSEELPTYMIPSYFIQLDKMPL TSNGPacKIDRKQLPEPDLTFGSRSHHHHHH	this work
sfGFP(N150PacK)-His_6	psfGFP (N150PacK)	pBAD (Amp)	MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGE GDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCF SRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTY KTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEY NFNSHPacKVYITADKQKNGIKANFKIRHNVEDGSV QLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDP NEKRDHMVLLEFVTAAGITHGMDELYKGSHHHHH H	Ref[10]
SrtN	pQE61-yhdz	pQE61 (Amp)	METFKSILHEAQRIVVLTGAGMSTESGIPDFRSAGGI WTEDASRMEAMSLDYFLSYPRLFWPKFKELFQMK MSGSFEPNEGHLLLAELEKQGKQVDIFTQNIDGLHK KAGSRHVYELHGSIQTAACPACGARYDLPHLLEREV	Ref[11]

			PECTAAGNNGDICGTVLKTDVVLFGDAVMHFDTLY EKLDQADLLLVIGTSLEVAPARFVPEDASLIPGMKKV IINLEPTYCDSLFDMVIHQKIGEFARSLGMKKTSRSH HHHHH	

Production and Purification of Proteins

All constructs were expressed in Escherichia coli BL21 (DE3) gold. To this end, cells were co-transformed with the appropriate plasmids.

Expression strains for sfGFP(N150PacK), SUMO-2(K11PacK), GrsA(K434PacK), GrsA(K517PacK), PheA(K434PacK) and PheA(K517PacK) contained pBrCnKRS[10] and either psfGFP(N150) or pPM01, pED123, pED124, pMR28 and pMR29, respectively. LB medium supplemented with the appropriate antibiotics was inoculated with an overnight culture and the cells were grown at 37 °C until the OD 600 reached 0.7. Then, 2 mM PacK (final concentration) was added to the culture. 15 min later, the expression was induced with 0.2 % L-(+)-arabinose (w/v) for sfGFP(N150PacK) and with 0.2 % L-(+)-arabinose (w/v) and 400 µM IPTG for all the other proteins (SUMO-2(K11PacK), GrsA(K434PacK), GrsA(K517PacK), PheA(K434PacK) and PheA(K517PacK)). The cells were cultivated for additional 4 h at 28 °C. Cells were harvested by centrifugation (4000 g, 30 min, 4 °C). The pellet was resuspended in CBD lysis buffer for SUMO(K11PacK) (20 mM HEPES, 500 mM NaCl, 1 mM EDTA, pH 8.0) or Ni-NTA buffer (50 mM HEPES, 100 mM NaCl, pH 8.0) for all the other proteins (GrsA(K434PacK), GrsA(K517PacK), PheA(K434PacK), PheA(K517PacK)). The cell suspension was lysed by three passages through an emulsifier (EmulsiFlex®-C5, Avestin). Insoluble cell debris was removed by centrifugation (12000 g, 30 min, 4 °C). Proteins were then purified either using Ni-NTA or the IMPACT strategy.

SUMO-2(K11PacK) was isolated with a yield of ~2.5 mg/L. The yields of purified PheA mutants K517PacK and K434PacK were 55.0 mg/L and 35.0 mg/mL, respectively. GrsA(K517PacK) and GrsA(K434PacK) were obtained with yields of 8.5 and 7.8 mg/L, respectively.

CBD-GyrA Intein-tagged SUMO construct was purified using the IMPACT purification strategy. Chitin beads, equilibrated with CBD buffer, were incubated with the clarified cell lysate for 2 h at 4 °C. The beads were then washed with 30 column volumes (CV). On-column thiolysis was initiated by the addition of CBD buffer containing 100 mM DTT to get the free C-terminal carboxylic acid. After incubation for 48h at 4 °C under slight agitation, the solution was incubated with fresh chitin beads for additional 24h. The cleavage efficiency was analyzed by SDS-PAGE. If necessary, the solution was passed again over fresh chitin beads. The protein solution was dialyzed two time against a phosphate dialysis buffer (20 mM phosphate, 150 mM NaCl, pH 7.5) and then with 10% (v/v) glycerol.

His_6-tagged sfGFP, GrsA and PheA proteins were purified over a Ni-NTA-agarose column equilibrated with Ni-NTA buffer supplemented with 20 mM imidazole. The beads were washed with 20 CV of Ni-NTA buffer (+20 mM imidazole) and 10 CV of Ni-NTA buffer (+40 mM imidazole). Elution of the proteins was performed by the addition of seven times 0.5 CV of Ni-NTA buffer containing 250 mM of imidazole. Purified fractions were identified by SDS-PAGE, pooled and dialyzed two times again NRPS assay buffer (50 mM HEPES, 100 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, pH 7.0) and then with 10% (v/v) glycerol.

For the production of GrsA wild-type and GrsA(K517A), expression and purification was performed using *E. coli* BL21(DE3) strains containing the plasmid pGV196 and pED139, respectively, and as described above for the other GrsA constructs, except for that no unnatural amino acid was added.

Recombinant production of SrtN from *Bacillus subtilis* was performed by adaptation of a previously reported protocol.[11] M15 (pREP4) cells were transformed with the pQE61-Yhdz expression vector. The cells were grown overnight at 37 °C in LB medium supplemented with Ampicillin (100 µg/mL), Kanamycin (50µg/mL) and $ZnCl_2$ (25 µM). Then, the cells were diluted (1:100) in 600 mL of fresh LB medium (Amp 100 µg/mL) and incubated at 37 °C until the OD 600 reached 0.7. The expression was then induced with 500 µM IPTG. The cells were cultivated for additional 4 h at 37 °C and harvested by centrifugation (4000 g, 30 min, 4 °C). The pellet was resuspended in Ni-NTA buffer (50 mM HEPES, 100 mM NaCl, pH 8.0). The cells were lysed by three passages through an Avestin C5 emulsifier. Insoluble cell debris were removed by centrifugation (12000 g, 30 min, 4 °C). The protein was purified over a Ni-NTA-agarose column using the same protocol as described earlier. Purified fractions were identified by SDS-PAGE, pooled and dialyzed two times against Sirtuin assay buffer (50 mM NaH_2PO_4, 100 mM NaCl, 25 µM $ZnCl_2$, pH 8.0) and finally with 50% (v/v) glycerol.

In all cases, protein concentration was determined using calculated absorbances at 280 nm. Purified aliquots were stored at -80°C after being flash frozen in liquid nitrogen.

Enzymatic deprotection using penicillin G acylase (PGA)

Fmoc-Lys(Pac)-OH **3** (1 mM) was dissolved in phosphate buffer (100 mM phosphate, pH 7.5). Soluble PGA (Sigma, 0.5U) was added and the reaction proceeded at 37 °C for 10 min. The reaction mixture was then loaded on a Vivaspin 500® column (5000 MWCO, Sartorius) and centrifuged to separate the protein from the small organic compounds. The flow-through containing the small organic compound was then acidified and analyzed by LC-MS using a gradient 5 % to 90 % B in 15 min with a flow rate of 0.4 mL/min.

Pac-protected proteins were diluted in a phosphate buffer (50 mM phosphate, 300 mM NaCl, pH 7.5) or HEPES buffer (50 mM HEPES, 100 mM NaCl, pH 7.5) to reach a final concentration of 5 to 20 µM. 1 to 10 U of soluble PGA (Sigma) was added to the solution and the reaction proceeded at 25 °C. After 15h reaction, the proteins were purified using Ni-NTA-agarose beads and analyzed by LC-MS to determine the deprotection yield.

Enzymatic deprotection using *Bacillus subtilis* SrtN

This assays was performed by adaptation of a previously reported protocol.[11] Pac-protected proteins were diluted in Sirtuin assay buffer (50 mM NaH_2PO_4, 100 mM NaCl, 50 µM $ZnCl_2$, pH 8.0) to reach a final concentration of 5 to 20 µM. Then the desired amount of purified SrtN enzyme and NAD^+ (2 mM) were added to the solution. The reaction was incubated at 25 °C. For time course experiments, aliquots were removed and quenched by addition of nicotinamide (30 mM). Deprotection yields were monitored by LC-MS.

DKP dipeptide formation assay

The reaction was performed as previously described.[9] As a preliminary step, GrsA and TycB1 proteins were modified with the 4'-phosphopantetheine moiety. Proteins were incubated in NRPS assay buffer supplemented with 10 mM $MgCl_2$, 2 mM TCEP, 50 equiv. CoA-SH and 0.02 equiv. Sfp for 15 h at 0 °C. To form the dipeptide D-Phe-L-Pro-diketopiperazine (DKP), GrsA and TycB1 (wild-type or mutants) were incubated in NRPS assay buffer containing 20 mM $MgCl_2$, 1 mM L -Pro, 1 mM L -Phe and 5 mM ATP. The final concentrations for the GrsA derivatives and TycB1 were respectively 0.5 µM and 5 µM, respectively. The reaction was performed at 37 °C for 45 min. 100 µL of the reaction mixture was then mixed with 400 µL butanol/chloroform (4:1, v/v) and 200 µL H_2O to quench the reaction and extract the formed DKP product. Samples were vortexed, and the phases separated by centrifugation (RT, 14,000 g, 2 min). The aqueous phase was extracted a second time with 400 µL butanol/chloroform. The organic phases were pooled, washed two times with 300 µL H_2O and dried in a speed vac. The residue was solubilized with 95% H_2O, 5% acetonitrile and 0.1% TFA. The product was then analyzed by analytical LC-MS using a gradient 5 % to 80 % B in 17 min with a flow rate of 0.4 mL/min. Quantification was done by integrating the chromatograms at 210 nm. Two normalization steps were realized to determine the percentage of DKP formation. In a first step, the concentration of the GrsA-mutants used for the DKP assay was corrected by densitometric analysis of SDS gels using the software ImageJ. For this, a standard curve of GrsA-wt was constructed by plotting the concentration determined by absorption at 280 nm against the corresponding band intensity (curve obtained using at least three repeats). Using this curve and densitometric analysis of the bands obtained for the GrsA mutants, the concentrations of these proteins were corrected. In a second step, the correlation between the concentration of GrsA-wt and the percentage of DKP product formed was determined. By this, the amount of formed DKP by the GrsA-mutants could be directly compared to the one formed by GrsA-wt at the same concentration. Therefore, the activity could be corrected. The percentage of formed DKP by GrsA-wt was fitted with an exponential function (Y = 100*(1-e(-k*x)); R2 = 0,834).

LC-MS analysis

Small organic compounds were analyzed using an Agilent 1260 Infinity series system (Agilent Technologies, Waldbronn, Germany). Samples were diluted with 95% H_2O, 5% acetonitrile and 0.1% TFA, centrifuged (14000 rpm, 2 min) and loaded on a C18 column (ZORBAX SB-C18 RR HT, 3 x 50 mm, 1.8 µm, Agilent Technologies, Waldbronn, Germany) at a flow rate of 0.4 mL/min.

Mass analyses of intact proteins were performed using an UltiMate™ 3000 RS system (Thermo Fisher Scientific Inc., MA, USA) connected to a maXis II UHR-TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) with a standard ESI source (Apollo, Bruker Daltonik GmbH, Bremen, Germany). Proteins were reduced with 2 mM TCEP on ice during 15 minutes to avoid inhomogeneity issues. Then, samples were acidified using a 5% formic acid solution to reach a pH 2-3 and centrifuged (14000 rpm, 2 min). According to the protein concentration, an appropriate volume of the supernatant was loaded on a C4 column (Advance Bio RP-mAb C4, 2.1 mm x 50 mm, 3.5 µm, Agilent Technologies, Waldbronn, Germany) at a flow rate of 0.6 mL/min in 5% eluent B (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile). After a desalting period of 7 minutes at 5% B, a steep gradient was applied (5-60% B in 2 min). MS settings: capillary voltage 4500 V, end-plate offset 500 V, nebulizer 5.0 bar, dry gas 9.0 L/min, dry T=200°C, mass range m/z 300-3000. Data were analyzed with DataAnalysis 4.4 (Bruker Daltonik GmbH,

Bremen, Germany) and deconvolution was performed using the MaxEnt algorithm implemented in the software.

In-gel tryptic digest followed by LC-MS/MS analysis of proteins

In-gel tryptic digest was performed in the presence of ProteaseMAX™ (Promega, WI, USA) according to the instructions of the manufacturer with some modifications: Protein bands were cut from Coomassie-stained SDS-gels and destained at 40°C using a solution composed of 50% methanol, 50% water and 0.1% TFA. After two washing steps with water (2 x 400 µL), each band was incubated with 400 µL of acetonitrile for 20 minutes. The supernatant was removed and the band was dried for 15 minutes in a Speed Vac SC110 (Savant instruments Inc. NY, USA). Reduction was performed at 56°C for 20 minutes using 40 µL of a solution composed of 10 mM DTT in 100 mM ammonium bicarbonate buffer. After a washing step with 400 µL of ammonium bicarbonate buffer, the protein sample was alkylated in the dark at room temperature for 30 minutes, using 40 µL of a 55 mM 2-iodacetamide solution in 100 mM ammonium bicarbonate buffer. The band was subsequently washed twice with 400 µL water and dried in a Speed Vac. In the meantime, a 400 ng aliquot trypsin dissolved in 50 mM ammonium bicarbonate buffer was activated for 30 minutes at 30°C. Then, the dried band was pre-incubated at room temperature for 10 minutes with the activated trypsin before adding 40 µL of a solution of 0.01% ProteaseMAX™ (Promega, WI, USA) in 50 mM ammonium bicarbonate buffer. After 2 h incubation at 37°C, the supernatant was collected and acidified with a 5% formic acid solution to reach a pH 2-3. The resulting tryptic digest was then analyzed on a LC-MS consisting of an UltiMate™ 3000 RS LC nano system (Thermo Fisher Scientific Inc., MA, USA) connected to a maXis II UHR-TOF mass spectrometer with a nano-ESI source (Bruker Daltonik GmbH, Bremen, Germany). For all proteins, 5 µL of the digest solution were injected. The solution was loaded on a C18 trapping column (Acclaim PepMap100, 5 µm, 100 Å, ID 300 µm x L 5 mm, Thermo Scientific) at a flow rate of 20 µL/min in 2% eluent B (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile). After 5 minutes of washing at 2% B, a 50 minutes gradient (2% to 60% B, flow rate 300 nL/min) was applied for the separation on a C18 nano column (Acclaim™ PepMap™ 100 C18, 2 µm, 100 Å, ID 0.075 mm x L 250 mm, Thermo Fisher Scientific Inc., MA, USA). MS settings: capillary voltage 1.200 V, mass range: m/z 150-2200. CID was employed for data-dependent MS/MS. DataAnalysis 4.4 (Bruker Daltonik GmbH, Bremen, Germany) was used for chromatogram processing and ProteinScape 4.0.3 (Bruker Daltonik GmbH, Bremen, Germany) was used for in-house database search (Mascot 2.5, Matrix Science Ltd., London, UK) of focused databases and for the analysis of MSMS data.

Assay with lysine auxotroph *E. coli* cells

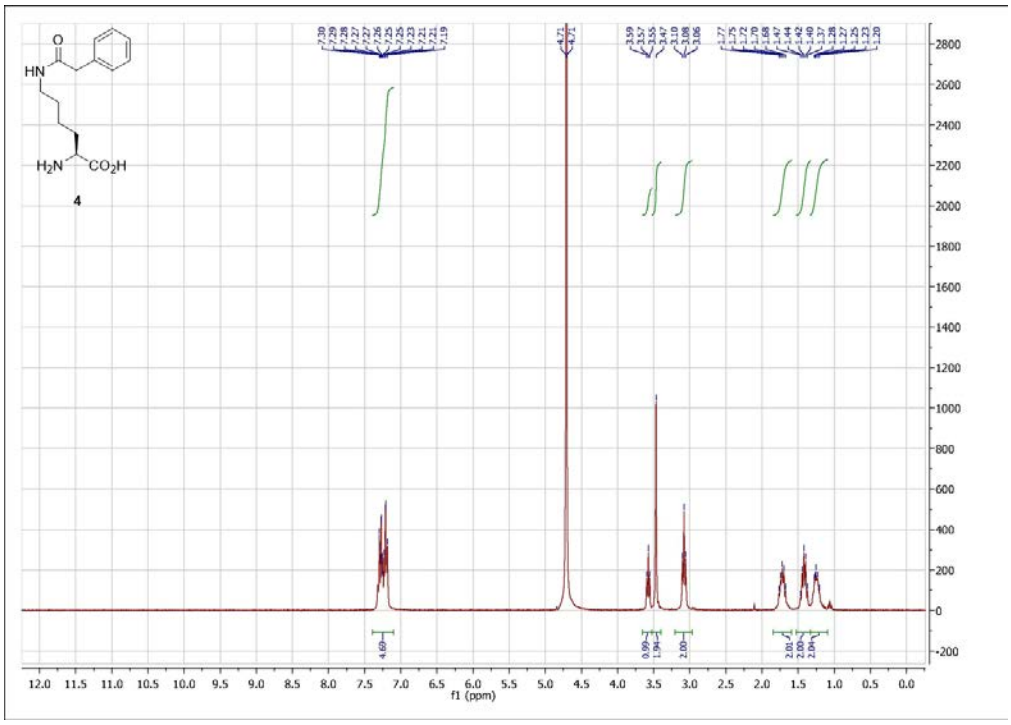
E. coli strain JW2806-1 (Yale *E. coli* center) was grown overnight at 37°C in LB medium supplemented with kanamycin (50 µg/mL) and lysine (50 µg/mL). Then, the cells were collected, centrifuged (4000 rpm, 15 min, 4 °C) and washed twice with minimal medium (33.7 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.55 mM NaCl, 100 nM $CaCl_2$, 1 mM $MgSO_4$, 10 ng/mL thiamine, 0.02 % NH_4Cl, 0.04 % Glucose, 22 nM Fe(III)Cl_3). Cells were then diluted 1:100 into minimal medium supplemented with kanamycin (50 µg/mL) in presence or absence of lysine (50 µg/mL), PacK (50-100 µg/mL) or a mixture of both. Cells were grown at 37°C and cell growth was determined by measuring OD600.

MS assay to analyze the thioester formation of GrsA mutants

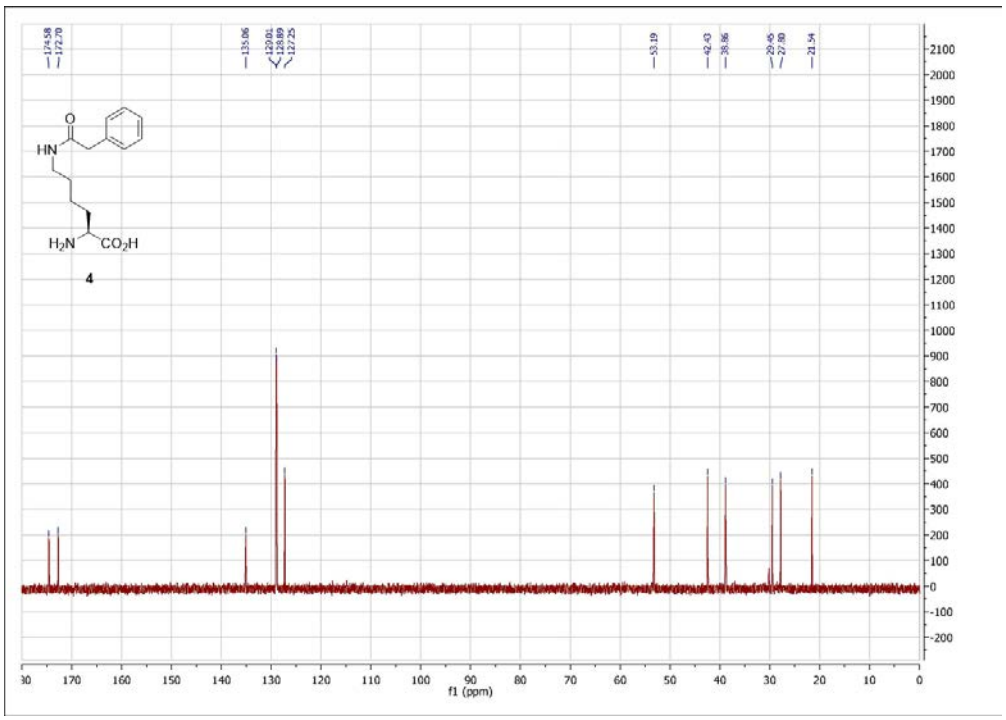
The thioester formation assay was realized by adaptation of a previously reported protocol.[6] First, GrsA wild-type, GrsA(K434PacK) and GrsA(K517PacK) were modified with the 4'-phosphopantetheine moiety to form the holo-proteins corresponding. For this, the proteins (5-10 μM) were incubated in NRPS assay buffer supplemented with 10 mM $MgCl_2$, 1 mM TCEP, 70 equiv. CoA-SH and 0.07 equiv. Sfp for 15 h at 0 °C. Then, reaction of thioester formation was performed at 25 °C. The holo-proteins were incubated with 2 mM ATP, 2 mM L-phenylalanine and 10 mM $MgCl_2$. After 30 minutes, the reaction mixtures were quenched by adding formic acid (final concentration of 1%). The samples were then centrifuged (14,500 rpm, 2 min) and the covalent attachment of phenylalanine was measured by LC-MS.

NMR-spectra

^{1}H NMR spectrum of PacK **4** (300 MHz, D_2O)



^{13}C NMR spectrum of PacK **4** (75 MHz, D_2O)



Supplementary references

1. H. J. Duggleby, S. P. Tolley, C. P. Hill, E. J. Dodson, G. Dodson and P. C. Moody, *Nature*, 1995, **373**, 264-268.
2. K. M. Hoyer, C. Mahlert and M. A. Marahiel, *Chem Biol*, 2007, **14**, 13-22.
3. E. Conti, T. Stachelhaus, M. A. Marahiel and P. Brick, *EMBO J*, 1997, **16**, 4174-4183.
4. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg, *Nat Protoc*, 2015, **10**, 845-858.
5. C. A. Mitchell, C. Shi, C. C. Aldrich and A. M. Gulick, *Biochemistry*, 2012, **51**, 3252-3263.
6. J. Alfermann, X. Sun, F. Mayerthaler, T. E. Morrell, E. Dehling, G. Volkmann, T. Komatsuzaki, H. Yang and H. D. Mootz, *Nat Chem Biol*, 2017, **13**, 1009-1015.
7. S. Vila, C. Camo, E. Figueras, E. Badosa, E. Montesinos, M. Planas and L. Feliu, *Eur J Org Chem*, 2014, **22**, 4785-4794.
8. L. Zakova, D. Zyka, J. Jezek, I. Hanclova, M. Sanda, A. M. Brzozowski and J. Jiracek, *J Pept Sci*, 2007, **13**, 334-341.
9. E. Dehling, G. Volkmann, J. C. Matern, W. Dorner, J. Alfermann, J. Diecker and H. D. Mootz, *J Mol Biol*, 2016, **428**, 4345-4360.
10. M. Cigler, T. G. Muller, D. Horn-Ghetko, M. K. von Wrisberg, M. Fottner, R. S. Goody, A. Itzen, M. P. Muller and K. Lang, *Angew Chem Int Ed Engl*, 2017, **56**, 15737-15741.
11. J. Seidel, C. Klockenbusch and D. Schwarzer, *Chembiochem*, 2016, **17**, 398-402.