An engineered lantipeptide synthetase as a general leader peptide-dependent kinase

Gabrielle N. Thibodeaux and Wilfred A. van der Donk[a]

Electronic Supplementary Information (ESI)

Materials and methods

Biological reagents and materials

Chemically competent *Escherichia coli* DH5α cells (New England Biolabs) were used for propagation and isolation of all plasmids. Chemically competent *E. coli* Rosetta 2 cells (Novagen) were used to over-express all proteins. All restriction enzymes, DNA T4 ligase, and Factor Xa were obtained from New England Biolabs unless otherwise stated. Phusion and Pfu Turbo polymerase were acquired from Finnzymes and Agilent, respectively. All primers were synthesized and purchased from Integrated DNA Technologies. DNA sequencing was performed by ACGT, Inc. All DNA was purified by Miniprep or PCR clean up kits (Qiagen). HisTrap 1 mL and 5 mL columns were purchased from GE Healthcare. SDS-PAGE gels were obtained from BioRad. All proteins were concentrated using Millipore Amicons. All proteins were dialyzed using cassettes from Pierce.

General Mass Spectrometry Data collection

MALDI-ToF mass spectra were recorded on a Voyager DE-STR Biospectrometer Workstation in positive mode. α-Cyano-3-hydroxycinnamic acid was used as the matrix for MALDI-ToF experiments. Liquid chromatography-mass spectrometry (LC-MS) was performed on a Waters Synapt system consisting of an Acquity Ultra Performance Liquid Chromatography system using a C8 (1.7 μM, 1.0 x 100 mm) column coupled to a ESI quadrupole ToF mass spectrometer operating in positive mode.

Cloning of ProcM mutants

ProcM mutants were generated with the primers listed in ESI Table 3 using the Quikchange® Site-Directed Mutagenesis Kit (Agilent) as directed. The template utilized in the site-directed mutagenesis was ProcM:pET28b or ProcA 2.8(MCSI)-ProcM(MCSII):pRSFDuet, which both were previously reported.[1, 2] The inserts of purified mutant plasmids were sequenced for verification of the desired mutation and to ensure that no other mutations were present.

Over-expression and purification of ProcM and its mutants (R510M and T516A)

Overnight cultures of *E. coli* Rosetta 2 cells transformed with ProcM or ProcM mutant constructs (in pET28b) were diluted 1:1000 in 3 L of fresh Luria-Bertani (LB) media supplemented with 50 μg/mL kanamycin and grown at 37 °C to an OD_{600} of 0.7. The cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for an additional 20 h at 18 °C. Cells were harvested by centrifugation (6000 x g for 20 min at 4 °C). Pellets were stored at -80 °C and re-suspended in start buffer (1 M NaCl, 20 mM Tris, 10% glycerol, pH 8.0) for protein isolation. The cell suspension was passed through a Avestin C3 homogenizer 6-10 times at 5,000-7,500 psi to achieve lysis. The lysed cells were cleared through centrifugation (12,000 x g for 40 min at 4 °C). The supernatant was subsequently applied to a HisTrap HP column (GE Healthcare) after the column was equilibrated with 5 column volumes of start buffer. The column was washed with 10 column volumes of wash buffer (1 M NaCl, 20 mM Tris, 30 mM imidazole, pH 8.0). ProcM was eluted from the column by applying a linear gradient from wash to elution buffer (1 M NaCl, 20 mM Tris, 200 mM imidazole, pH 8.0) over 60 min. The purification was carried out at 4 °C employing an ÄKTA FPLC (GE Healthcare). The flow rate was set to 1 mL/min and the absorbance at 280 nm was monitored. Fractions were examined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing highly pure ProcM were pooled, concentrated by Amicon, and total protein concentration was measured by absorbance at 280 nm (ProcM calculated extinction coefficient 143,110 M^{-1} cm^{-1}; ExPASy ProtParam tool).

Cloning, over-expression and purification of ProcM substrates

The desired peptides listed in Table S2 were amplified with two rounds of PCR with the primers listed in Table S3. For the remainder of this ESI, a shorthand notation will be used to refer to the engineered ProcA substrates. For example, 1.7/AKT will represent a peptide containing the ProcA 1.7 leader sequence followed by the AKT substrate peptide sequence. For some peptides. a Factor Xa cleavage site was introduced between leader peptide and substrate peptide, and these peptides are indicated as 1.7/Xa/AKT. For the construction of 1.7/Xa/AKT, the a DNA fragment corresponding to ProcA1.7 leader sequence, Factor Xa protease site and the *N*-terminal five amino acids of AKT were amplified with GenPep15NdeIFWD and ProcA1.7 LS Xa AKT REV primers using ProcA 1.7:pET15 as template.[2] The PCR product was purified and used as a template and forward primer in the second PCR reaction along with AKT XhoI

primer. The desired PCR product was purified. All other genes encoding the desired peptides were constructed in a similar manner with their corresponding primers in Table S3.

The expression vector (pET15a) and all final PCR products were successively digested with NdeI and XhoI restriction enzymes at 37 °C for 3 h and overnight, respectively. The digested products were purified with the Qiagen PCR cleanup kit and were ligated using Quick T4 DNA ligase for 5 min at room temperature. *E. coli* DH5α cells were transformed with the ligation mixture and the desired sequences of resultant plasmids were verified by sequencing.

E. coli Rosetta 2 cells were transformed with pET15a expression vectors containing genes encoding various peptides. A single colony was used to inoculate 120 mL of LB media supplemented with 75 μg/mL of ampicillin and incubated at 37 °C with shaking for 14 h. The pre-culture was harvested at 6000 x g at 4 °C for 10 min, the dry pellet was re-suspended in terrific broth (TB) and used to re-inoculate 6 L of TB medium supplemented with 75 μg/mL of ampicillin. The culture was grown at 37 °C to an OD_{600} of 0.7. The cells were induced with 1 mM IPTG and incubated for an additional 3 h at 37 °C with shaking. Cells were harvested by centrifugation (6000 x g for 10 min at 4 °C). Pellets were re-suspended in ProcA start buffer (20 mM NaH_2PO_4, 500 mM NaCl, 0.5 mM imidazole, 20% glycerol, pH 7.5) and were subjected to sonication cycles of 4 s bursts, 9.9 s rest for a total of 15 min. Lysed cells were centrifuged at 14,000 x g for 30 min at 4 °C and the supernatant was discarded. The remaining pellet was washed with LanA start buffer, sonicated, centrifuged and the supernatant (soluble wash fraction) was discarded similar to the previous step. The pellet was washed and then resuspended in 60 mL of LanA buffer 1 (20 mM NaH_2PO_4, 500 mM NaCl, 0.5 mM imidazole, 6 M guanidine, pH 7.5), sonicated, and centrifuged again. The supernatant was applied to a HisTrap column, which was equilibrated with 5 column volumes (CV) of LanA buffer. The bound His-peptide was washed with 5 CV of LanA buffer 1 followed by 5 CV of LanA buffer 2 (20 mM NaH_2PO_4, 300 mM NaCl, 30 mM imidazole, 4 M guanidine, pH 7.5). The peptides were eluted from the column with LanA elution buffer (20 mM NaH_2PO_4, 100 mM NaCl, 1 M imidazole, 4 M guanidine, pH 7.5) and stored at room temperature for 24-48 h. The Ni-NTA purified peptides were desalted by three different methods depending on scale. The traditional desalting method performed involved preparative reversed-phase HPLC (Waters Delta 600) employing a Waters C4 PrepPak cartridge. The purification was performed at room temperature by applying a linear gradient of 2% solvent A (80% acetonitrile and 0.1% TFA in water) to 100% solvent A over 45 min with the second solvent 0.1% TFA in water (solvent D). The flow rate was set to 8 mL/min and the absorbance at 220 nm was monitored. Fractions containing the desired peptide were pooled and lyophilized (Labconco). The product was analyzed by MALDI-ToF MS and stored at −20 °C. This desalting method was performed with high expressing peptides or cultures over 2 L. A more rapid procedure for relatively high peptide amounts involved precipitation. In this case, the Ni-NTA purified peptides were dialyzed for 2 h at room temperature against sterile deionized water that is 200-500 times the volume of the Ni-NTA eluent. The peptides were moved to fresh sterile deionized water for another 2 h and the water was replaced again and the peptide solution was dialyzed overnight at room temperature. The following morning the precipitated peptide was isolated, solubilized in acetonitrile, and lyophilized. The product was stored at −20 °C after the correct mass was confirmed by MALDI-ToF MS. Peptides that expressed in low quantities or obtained from small cultures were desalted by using C4 solid phase extraction columns (Grace Vydac) as directed by the product manual.

ProcM in vitro phosphorylation assays

Purified lyophilized ProcA peptides (20 μM) were dissolved in water and DMSO (DMSO final concentration <5%) because of solubility issues and were incubated with ProcM T516A (2 μM) in activity buffer (50 mM HEPES, 10 mM $MgCl_2$, 5 mM ATP, 1 mM TCEP, pH 7.5) for 5-20 h at room temperature. The reaction was quenched by addition of TFA to a final concentration of 0.5%. The reaction mixtures were subjected to C4 SPE columns to remove ProcM T516A. The C4 SPE eluent was lyophilized and re-suspended in 50 mM Tris, pH 8.0 and Factor Xa, trypsin, LysC or GluC was added and the solution was incubated for an additional 6 h to overnight at room temperature before analysis by MALDI-ToF MS.

Cloning of substrates into ProcM R510M or T516A (MCSII)-pRSFDuet

The genes encoding peptides (Proc-AKT, Proc-CaM, Proc-PKCμ and Proc-PKC) were amplified in the same manner as described above with the primers listed in Table S3. The only difference was that the resultant digested PCR product was ligated into multiple cloning site I (MCSI) of pRSFDuet vectors that already contained ProcM R510M or ProcM T516A in MCSII.

Co-expression and purification of ProcA 2.8 (or ProcM substrates) from ProcM mutants (MCSII)-pRSFDuet

Co-expression of ProcA 2.8 and ProcM (or ProcM mutants) was completed under the same conditions as over-expression of ProcM protein. Purification of the modified ProcA 2.8 (or other modified substrate peptides) was facilitated by a His tag as described above in the purification of ProcA peptides. The resultant peptide was treated with protease and analyzed by MALDI-ToF MS.

Table S1: Calculated and experimental m/z values for phosphorylation substrates. [a][M+4H]$^{4+}$, [b][M+2H]$^{2+}$ and [c][M+3H]$^{3+}$

Peptide	Figure #	calculated m/z	observed m/z
ProcA 2.8 unmodified	1	1269.36[a]	1268.59[a]
ProcA 2.8 + 2 PO_3^{2-}	1	1289.36[a]	1288.56[a]
ProcA 2.8 + 3 PO_3^{2-}	1	1309.35[a]	1308.56[a]
Proc-PKCμ unmodified	2a	1614.88	1614.89
Proc-PKCμ + PO_3^{2-}	2a	1694.88	1694.86
Proc-H1 unmodified	2b	1252.77	1252.84
Proc-H1 + PO_3^{2-}	2b	1332.74	1332.77
ProcA 1.1 unmodified	SI 1	901.24[a]	900.33[a]
ProcA 1.1 + 2 PO_3^{2-}	SI 1	921.24[a]	920.33[a]
ProcA 1.1 + 3 PO_3^{2-}	SI 1	941.24[a]	940.33[a]
S19 unmodified	SI 2a	1309.53[b]	1309.93[b]
S19 + PO_3^{2-}	SI 2a	1349.52[b]	1349.57[b]
S6 unmodified	SI 2b	2619.77	2619.98
S6 + PO_3^{2-}	SI 2b	2699.73	2699.94
S14 unmodified	SI 2c	4891.15	4890.7
S14 + PO_3^{2-}	SI 2c	4971.12	4971.68
S1 unmodified	SI 2d	4891.15	4891.8
S1 + PO_3^{2-}	SI 2d	4971.12	4971.42
RW unmodified	SI 3a	4100.52	n.d.
RW + PO32-	SI 3a	4180.42	4180.6
RLT unmodified	SI 3b	1324.62	n.d.
RLT + PO_3^{2-}	SI 3b	702.38[b]	701.56[b]
Proc-CaM unmodified	SI 4a	4346.74	4346.45
Proc-CaM + PO_3^{2-}	SI 4a	4425.44	4425.58
Proc-PKCμ unmodified	SI4b	4602.13	4602.47
Proc-PKCμ + PO_3^{2-}	SI 4b	4682.09	4682.28
Proc-LR unmodified	SI 4c	2441.72	2441.58
Proc-LR + PO_3^{2-}	SI 4c	2521.69	2521.36
Proc-PKG unmodified	SI 4d	1116.28	n.d.
Proc-PKG + PO_3^{2-}	SI 4d	1195.62	1194.26
Proc-AKT unmodified	SI 4e	1360.65	n.d.
Proc-AKT + PO_3^{2-}	SI 4e	1440.62	1440.99
Proc-AKT + 2 PO_3^{2-}	SI 4e	1520.58	1520
ProcA 2.8 unmodified	SI 5a	1269.36[a]	1268.19[a]
ProcA 2.8 + 2 PO_3^{2-}	SI 5a	1289.36[a]	1288.18[a]
ProcA 2.8 + 3 PO_3^{2-}	SI 5a	1309.35[a]	1308.16[a]
Proc-CaM unmodified	SI 5b	1359.65	1359.87
Proc-CaM + PO_3^{2-}	SI 5b	1439.62	1439.86
Proc-PKC unmodified	SI 6b	1701.01	1702.4
Proc-PKC + PO_3^{2-}	SI 6b	1780.98	1781.88

Table S2: ProcM kinase substrate sequences. Kinase products were digested at the *C*-terminus of the red amino acid using LysC or trypsin for cleavage after K and GluC for cleavage after E, and Factor Xa for cleavage after IEGR. Phosphorylations took place at the italic residues within the green region.

	Leader peptides	Phosphorylated peptide
In vitro		
ProcA 2.8	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGGAACHNHAP***S***MPP***S***YWEGEC	
ProcA 1.1	MSEEQLKAFIAKVQADTSLQEQLKAEGADVVAIAKAAGFSITTEDLEKEHRQTLSDDDLEGVAGGFFCVQG***T***ANRF***T***INVC	
S19	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGKYHHYNCYNFNLFNNYNNN***S***Y	
S6	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGKYHHYN***S***YNFNLFNNYNNNCY	
S1	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGG***S***YNFNLFNNYNNNCY	
S14	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGGCYNFNLFNNYNNN***S***Y	
RW	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGGRWVR***S***ALLI	
RLT	MSEEQLKAFIAKVQADTSLQEQLKVEGADVVAIAKASGFAITTEDLKAHQANSQKNLSDAELEGVAGGIEGRRLIK***T***FAYV	
Proc-AKT	MSEEQLKAFIAKVQADTSLQEQLKVEGADVVAIAKASGFAITTEDLKAHQANSQKNLSDAELEGVAGGIEGRARKRERTY***S***FGHHA	
Proc-CaM	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGGMHRQE***T***VDCLK	
Proc-PKCμ	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGGAALVRQM***S***VAFFFK	
Proc-LR	MSEEQLKAFIAKVQADTSLQEQLKVEGADVVAIAKASGFAITTEDLKAHQANSQKNLSDAELEGVAGGIEGRLRRA***S***VA	
Proc-PKG	MSEEQLKAFIAKVQADTSLQEQLKVEGADVVAIAKASGFAITTEDLKAHQANSQKNLSDAELEGVAGGIEGRRKR***S***RAE	
In vivo		
Proc-CaM	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGKMHRQE***T***VDCLK	
Proc-PKCμ	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGKAALVRQM***S***VAFFFK	
Proc-H1	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGEGGGPA***T***PKKAKKL	
Proc-PKC	MSEEQLKAFLTKVQADTSLQEQLKIEGADVVAIAKAAGFSITTEDLNSHRQNLSDDELEGVAGKRRGR***T***GRGRRGIFR	

Table S3: ProcM mutagenesis primers; bold font indicates site of mutagenesis.	
Primer name	Sequence (5' ➔ 3')
R510M FWD	tgcaggattgccacgc**atg**attgtgttgcgagcc
R510M REV	ggctcgcaacacaat**cat**gcgtggcaatcctgca
T516A FWD	gattgtgttgcgagcc**gcg**cgtgtgtatttcac
T516A REV	gtgaaatacacacg**cgc**ggctcgcaacacaatc

Table S4: ProcM substrate primers for pET28; bold font indicates restriction enzyme site.	
Primer name	Sequence (5' ➔ 3')
GenPep15NdeIFWD	ggaattc**catatg**atgtcagaagaacaa
S19 REV XhoI	ccg**ctcgag**ttagtaactgttattatt
S6 REV XhoI	ccg**ctcgag**ttagtagcagttattatt
S1 REV XhoI	same as S19 REV XhoI
S14 REV XhoI	same as S6 REV XhoI
ProcA2.8 LS RW REV	gctgcgcacccagcgtcccccagccacacc
RW XhoI	ccg**ctcgag**ttagatcagcagcgcgctgcgcacccagcg
ProcA1.7 LS Xa RLT REV	tttaatcaggcgacgaccttcgattcccccagccaca
RLT XhoI	ccg**ctcgag**ttacacatacgcaaaggttttaatcaggcgacgaccttcgat
ProcA1.7 LS Xa AKT REV	gcgtttgcgcgcacgaccttcgattcccccagccaca
AKT XhoI	ccg**ctcgag**ttacgcgtgatggccaaagctataggtgcgttcgcgtttgcgcgc
ProcA2.8 LS CaM REV	ttcctggcgatgcattcccccagccacacc
CaM XhoI	ccg**ctcgag**ttatttcaggcaatccacggtttcctggcgatgcat
ProcA2.8 LS PKCμ REV	gcgcaccagcgccgctttcccagccacacc
PKCμ XhoI	ccg**ctcgag**ttatttaaagaaaaacgccacgctcatctggcgcaccagcgccgc
ProcA1.7 LS Xa LR REV	cgcacggcgcagacgaccttcgattcccccagccaca
LR XhoI	ccg**ctcgag**ttacgcctcgctcgcacggcgcagacgacc
ProcA1.7 LS Xa PKG REV	gctgcgtttgcgacgaccttcgattcccccagccaca
PKG XhoI	ccg**ctcgag**ttattccgcgcggctgcgtttcgcacgaccttcgat

Table S5: ProcM substrate primers used to insert substrate genes into MCSI of pRSFDuet MCSI-ProcM R510M MCSII; bold font indicates restriction enzyme site.

Primer name	Sequence (5' ➔ 3')
GenPepDuetEcoRIFWD	gga**gaattc**atgtcagaagaacaa
ProcA2.8 LS CaM REV	ttcctggcgatgcattttcccagccacacc
CaM NotI	ccgattg**cggccgc**ttatttcaggcaatccacggtttcctggcgatgcat
ProcA2.8 LS PKCμ REV	gcgcaccagcgccgctttcccagccacacc
PKCμ NotI	ccgatt**gcggccgc**ttatttaaagaaaaacgccacgctcatctggcgcaccagcgccgc
ProcA2.8 LS PKC REV	ggtgcgaccgcggcgtttcccagccacacc
PKC NotI	ccgatt**gcggccgc**ttagcgaaaaatgccgcggcgaccgcggccggtgcgaccgcggcg
ProcA2.8 LS AKT REV	ttcgcgtttgcgcgctttcccagccacacc
AKT NotI	ccgatt**gcggccgc**ttacgcgtgatggccaaagctataggtgcgttcgcgtttgcgcgc

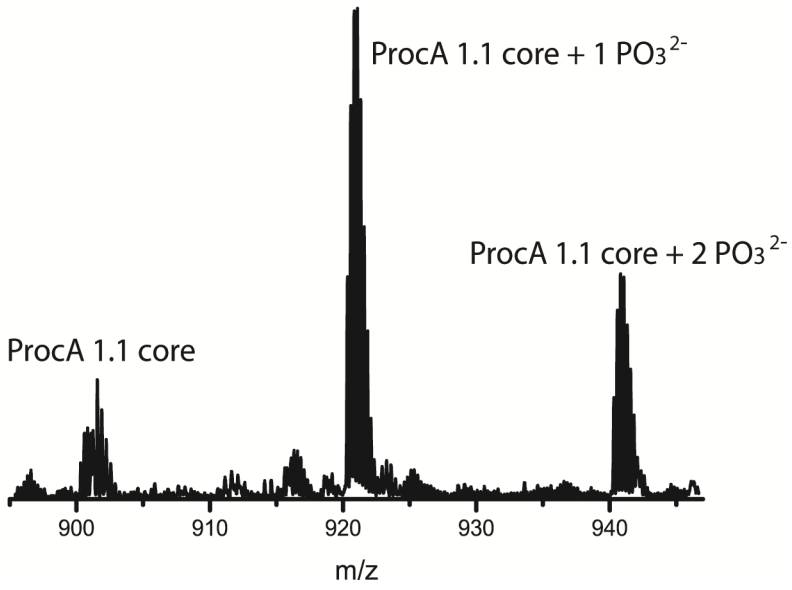

Figure S1: LC-ESI MS spectra of ProcA 1.1 after incubation with ProcM T516A and treatment with LysC protease.

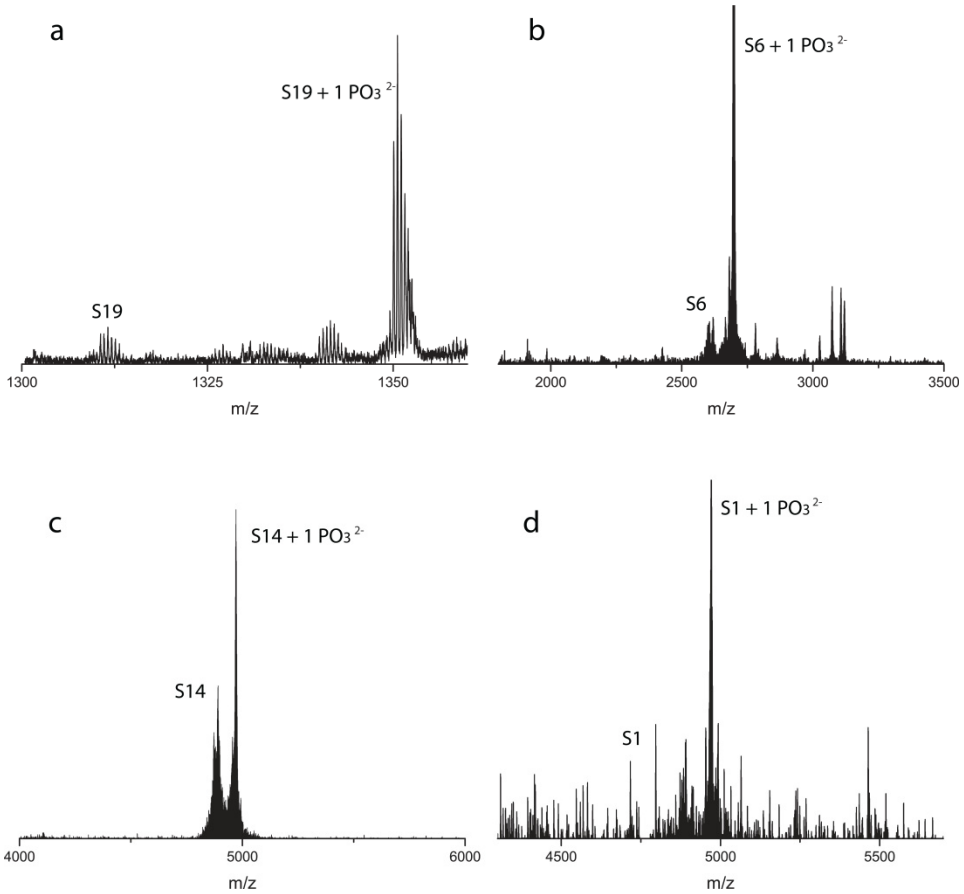

Figure S2: MALDI-ToF MS spectra of substrates (A) S19, (B) S6, (C) S14, and (D) S1 after incubation with ProcM T516A mutant and subsquent treatment with LysC protease.

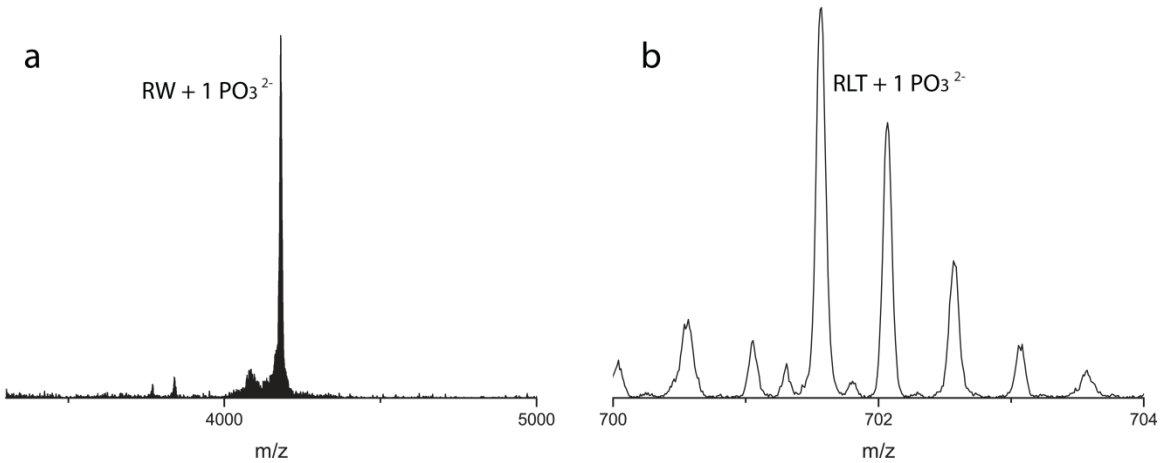

Figure S3: (A) MALDI-ToF MS spectra of RW and (B) LC-ESI MS spectra of RLT after incubation with ProcM T516A mutant and treatment with LysC protease.

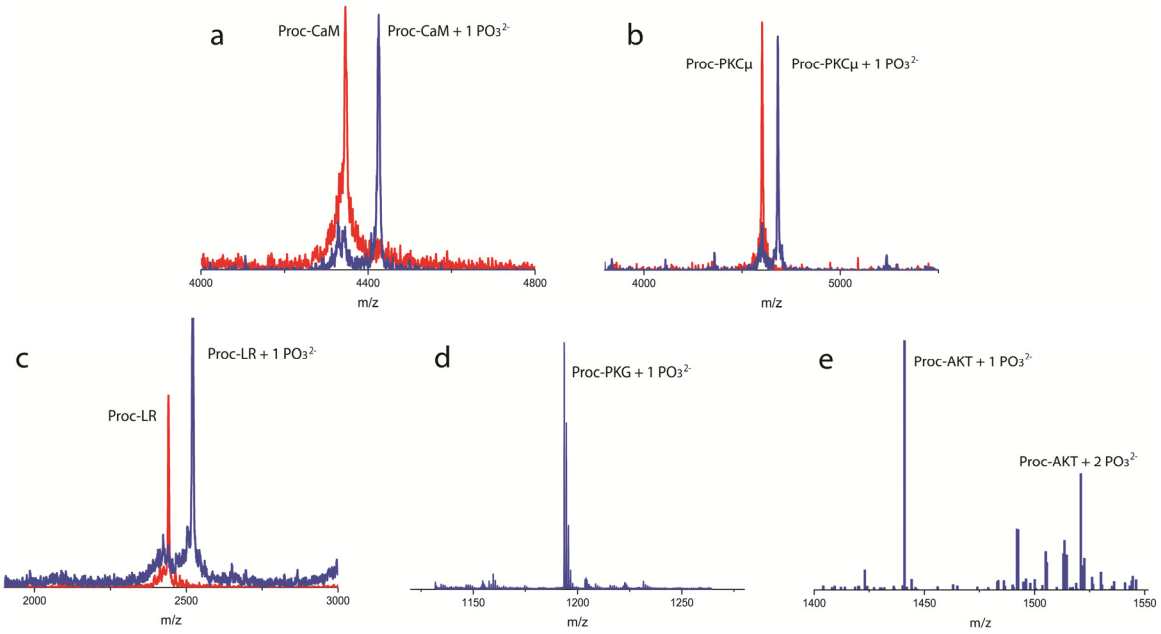

Figure S4: MALDI-ToF MS spectra of substrates (A) Proc-CaM, (B) Proc-PKCμ, (C) Proc-LR, (D) Proc-PKG, and (E) Proc-AKT after incubation with ProcM-T516A and subsquent treatment with LysC protease.

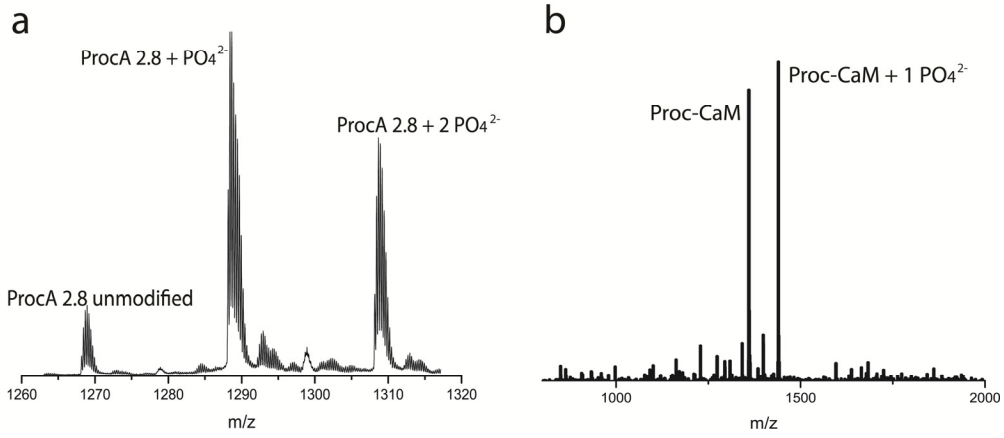

Figure S5: LC-ESI MS mass spectra of ProcA 2.8 (A) and Proc-CaM (B) co-expressed with ProcM R510M in *E. coli*, purified by Ni^{2+}-NTA affinity chromatography and treated with LysC protease.

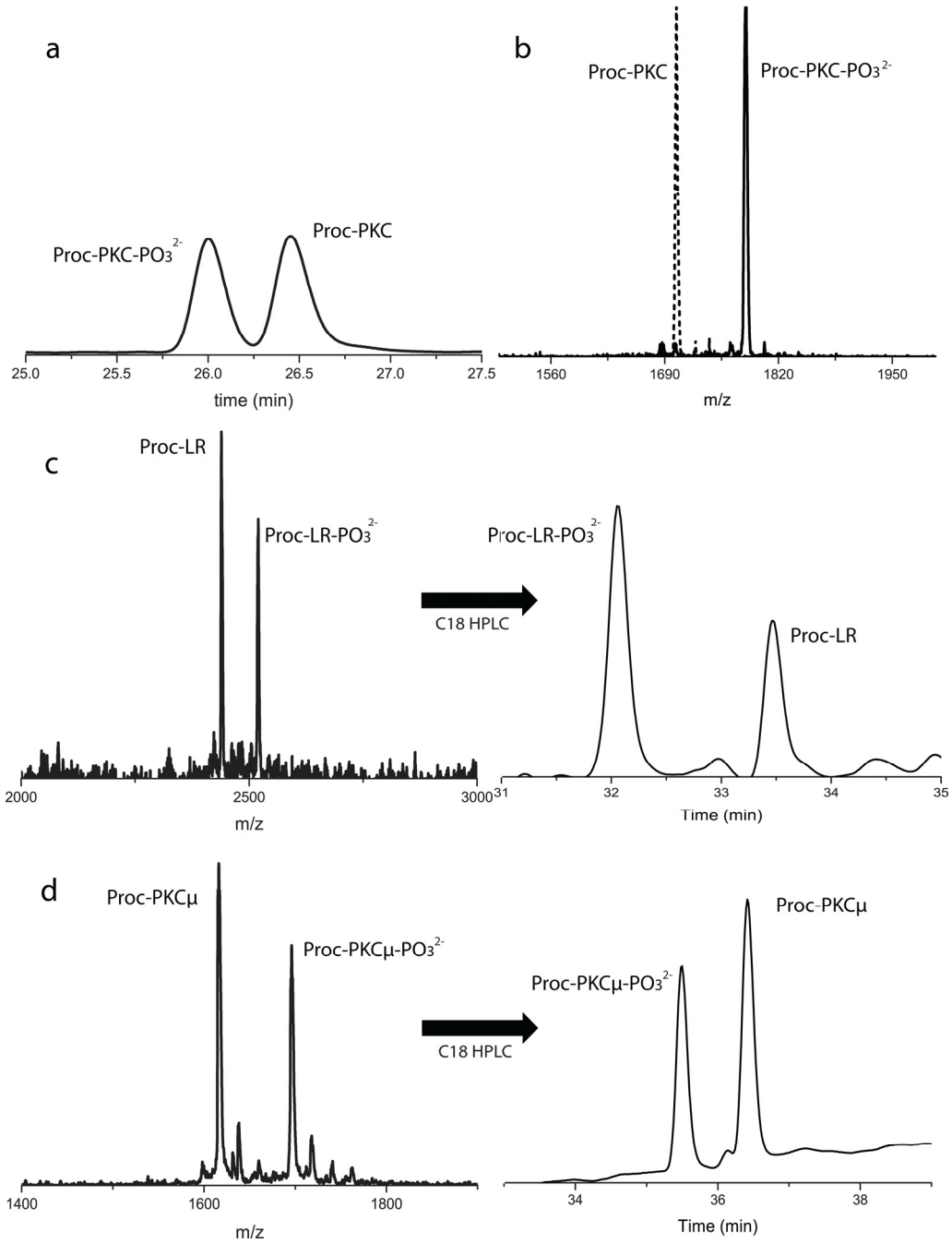

Figure S6: (A) HPLC trace showing the PKC peptide and its phosphorylated product. (B) MALDI-MS spectra of purified peptides. (C) MALDI-MS spectrum of Proc-LR treated with ProcM T516A and LysC protease. The same sample was also analyzed by reverse-phase (C18) HPLC. (D) MALDI-MS spectrum of Proc-PKCμ co-expressed with ProcM R510M in *E. coli*, purified by Ni^{2+}-NTA affinity chromatography and treated with LysC protease. The same sample was also analyzed by reverse-phase (C18) HPLC. These representative HPLC traces and MALDI-MS spectra show that for the peptides in this study, the MALDI spectra are reasonable indicators of the abundance of phosphorylated and unphosphorylated peptides.

Analysis of phosphorylated products with calf intestine phosphatase. To assure that the mass additions upon treatment with ProcM mutants indeed correspond to phosphorylations, the product peptides were subvjected to two enzymatic assays. Treatment with calf intestine phosphatase resulted in the conversion of the products back to their staring peptides, as expected for phosphorylated peptides (Figure S7). Furthermore, treatment of the product peptides with wild type ProcM resulted in peptides with masses consistent with phosphate elimination to generate dehydro amino acids (dehydroalanine from Ser, dehydrobutyrine from Thr; Figure S8). Collectively, these observations can only be explained if the products were phosphorylated on Ser/Thr.

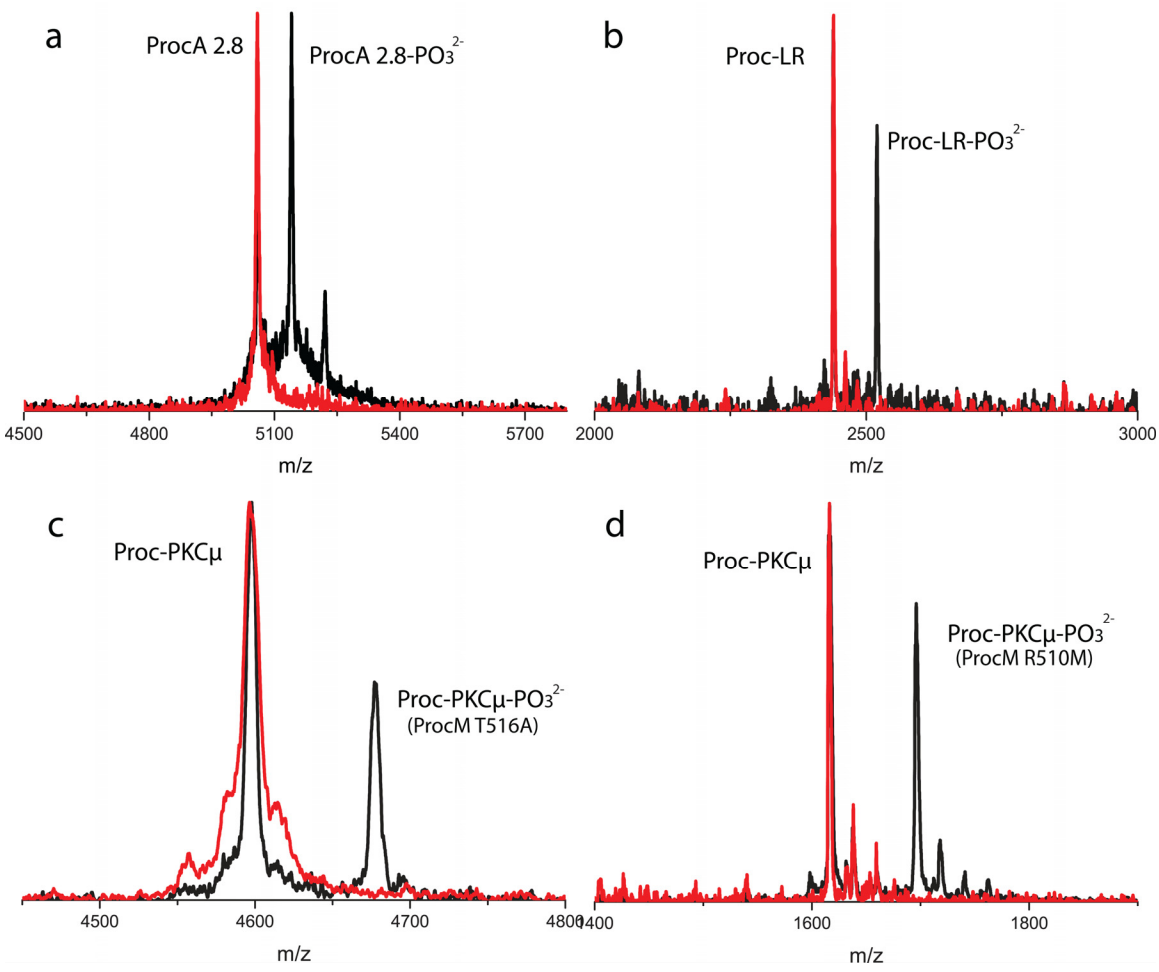

Figure S7: Phosphorylated products (black) were treated with calf intestine phosphatase (red) for 1 hour. MALDI-ToF MS spectra of (A) ProcA 2.8, (B) Proc-LR, (C) Proc-PKCμ and (D) Proc-PKCμ. Phosphorylated peptides were enzymatically synthesized by ProcM T516A mutant (except D, which was prepared by ProcM-R510M *E. coli*) and subsquently treated with LysC protease.

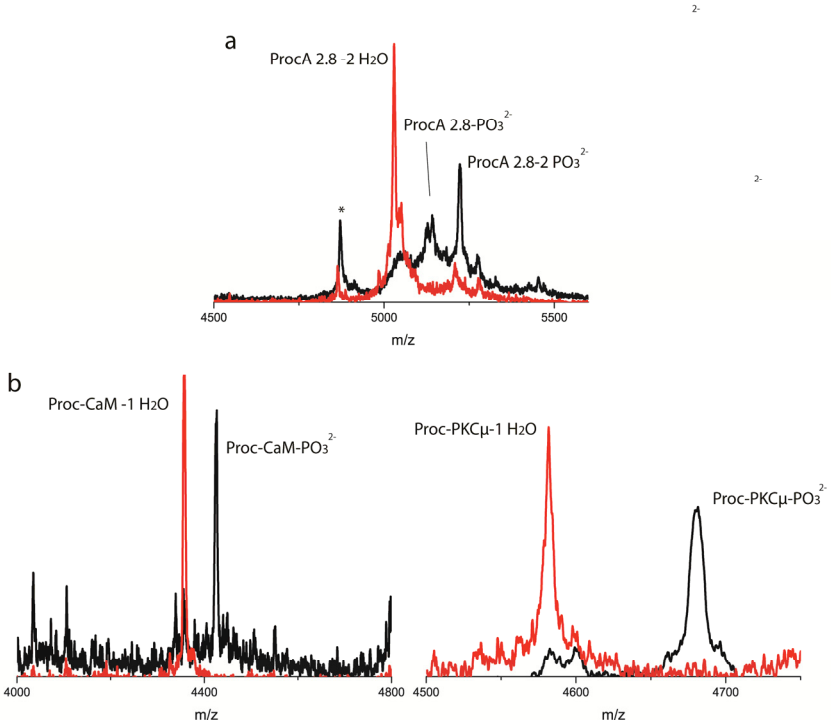

Figure S8: Phosphorylated products (black) were treated with wt ProcM (red) and subsquently treated with LysC protease. MALDI-ToF MS spectra of (A) ProcA 2.8, (B) Proc-CaM, and (D) Proc-PKCμ. Phosphorylated substrates were enzymatically synthesized with the ProcM T516A mutant.

References ESI

1. Y. Shi, X. Yang, N. Garg and W. A. van der Donk, *J Am Chem Soc*, 2011, **133**, 2338-2341.
2. B. Li, D. Sher, L. Kelly, Y. Shi, K. Huang, P. J. Knerr, I. Joewono, D. Rusch, S. W. Chisholm and W. A. van der Donk, *Proc Natl Acad Sci U S A*, 2010, **107**, 10430-10435.