Fluorescent probes reveal a minimal ligase recognition motif in the prokaryotic ubiquitin-like protein from *Mycobacterium tuberculosis*

Denis Smirnov, Abhinav Dhall, Kalkena Sivanesam, Rudy J. Sharar, and Champak Chatterjee*

Department of Chemistry, University of Washington, Seattle, WA 98195.

Email: chatterjee@chem.washington.edu

General Methods

1,5-diaminopentane, 1,4-diaminobutane, *N*-Boc-ethylenediamine, Lys, Gly, Ala and the amine-reactive fluorescein-5-isothiocyanate (FITC isomer 1) were purchased from Sigma-Aldrich Chemical Company (Milwaukee, WI). Nα-5-FAM-L-Lys (Probe **1**) was purchased from AnaSpec (Fremont, CA). Nε-Boc-L-Lys-OH and Nα-Boc-L-Lys-OH were purchased from ChemImpex International (Wood Dale, IL). The mixture of activated fluorophores 5(and 6)-carboxymethylrhodamine succinimidyl ester (5(6)-TAMRA, SE) was purchased from Invitrogen (Grand Island, NY). All other commonly used chemical reagents and solvents were purchased from Sigma-Aldrich Chemical Company (Milwaukee, WI) or Fischer Scientific (Pittsburgh, PA). Chemically competent DH5α and BL21(DE3), cells were purchased from Novagen (Madison, WI). The pET3a vector was purchased from EMD Millipore (Billerica, MA), restriction enzymes, and T4 DNA ligase were obtained from New England BioLabs (Ipswitch, MA). Primer synthesis and gene sequencing were performed by Integrated DNA Technologies (Coralville, IA) and Genewiz (South Plainfeld, NJ), respectively. Site-directed mutagenesis was performed with the QuikChange II XL site-directed mutagenesis kit from Agilent (Santa Clara, CA). Centrifugal filtration units were from Sartorius (Goettingen, Germany) and Slide-A-Lyzer dialysis cassettes were from Pierce (Rockford, IL). Plasmid mini-prep, PCR purification and gel extraction kits were purchased from Qiagen (Valencia, CA.) Cells were sonicated using a Sonifier Cell Disruptor, Model W185 from Heat Systems-Ultrasonics (Long Island, NY). Sterile filtration units, 0.45 μm pore-size, were purchased from EMD Millipore or Fisher Scientific. Size exclusion and nickel affinity chromatography was performed on an AKTA FPLC system from GE Healthcare (Waukesha, WI) equipped with a P-920 pump and UPC-900 monitor. SDS-PAGE gels and Tris-HCl running buffers were prepared as per Sambrook and Russell.[1] Protein molecular weight standards were obtained from Bio-Rad (Hercules, CA). Analytical reversed-phase HPLC (RP-HPLC) was performed on a Varian ProStar instrument with a Vydac C18 or Vydac C4 column (5 micron, 4 x 150 mm), employing 0.1% trifluoroacetic acid (TFA) in water (HPLC buffer A), and 90% acetonitrile, 0.1% TFA in water (HPLC buffer B), as the mobile phases. Typical analytical gradients were 0-73% HPLC buffer B over 30 min at a flow rate of 1 mL/min. Preparative scale purifications were conducted on a Vydac C18 or Vydac C4 preparative column (15-20 micron, 50 x 250 mm) at a flow rate of 9 mL/min. LC-ESI-MS analysis was conducted on a Bruker Esquire LC-ion trap spectrometer (Billerica, MA) in conjunction with an HP1100 series HPLC (Palo Alto, CA). MALDI-TOF MS data was acquired on a Bruker Autoflex II instrument. The full-length PafA-His_6 protein was analyzed on an Agilent 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC-MS instrument at the UW School of Pharmacy Mass Spectroscopy Center. All purified protein substrates were analyzed by C18 or C4 analytical RP-HPLC and ESI-MS or MALDI. Amino acid analysis was undertaken at the UC Davis Proteomics Core Facility. α-^{32}P-ATP was purchased from PerkinElmer Life Sciences (Waltham, MA), and was provided at a specific activity of 11 TBq (3000 Ci)/mmol. Polyethyleneimine (PEI) TLC plates were purchased from Fischer Scientific (Pittsburg, PA). In-gel fluorescence measurements and phosphorimaging of radioactive TLCs were conducted on a GE Typhoon 9000 Gel Imaging Scanner.

ImageJ software (http://rsb.info.nih.gov/ij/) was employed to quantify gels and TLCs.[2] Kinetic data was plotted using Kaleidagraph from Synergy software (Reading, PA).

Cloning of PupE and PafA from *Mycobacterium tuberculosis (Mtb)*

The *pup* and *pafA* genes were obtained by PCR from *Mtb* H37Rv genomic DNA kindly provided by the Sherman laboratory at the Seattle Biomedical Research Institute. PafA was cloned with a C-terminal hexahistidine (His_6) tag between the NdeI and BamHI restriction sites in the pET3a vector. Full-length PupE (Pup Q64E) and its site-directed mutants were cloned with an N-terminal His_6-tag between the NdeI and BamHI restriction sites in the pET3a vector. The full-length Ubiquitin (Ub)-PupE fusion protein and chimeras that were N-terminally truncated in Pup were cloned between the NdeI and BamHI restriction sites of the pET15b vector. This added an N-terminal His_6-tag to all Ub-PupE fusions. In order to obtain sufficient quantities of non-His_6-tagged PupE and its mutant and truncated variants for kinetic analysis in pupylation assays, an additional Ub-ENLYFQG-PupE fusion protein construct was generated by the introduction of the Tobacco Etch Virus (TEV) protease cleavage site (ENLYFQ/G) in existing Ub-PupE constructs by QuikChange (QC) mutagenesis. Removal of the Ub-ENLYFQ polypeptide by proteolysis with TEV protease resulted in the production of non-His_6-tagged Pup with an N-terminal Ala2Gly mutation (G-PupE). The correct sequences of all clones were confirmed by gene sequencing (Genewiz). All primers used in cloning experiments are tabulated below.

DNA Sequence (5'-to-3')	Restriction Site	Primer Name
GGGAATTC**CATATG**CAGCGTCGAATCATGGGCATCGAA	NdeI	PafA forward
GCCCGC**GGATCC**TTAGTGATGATGATGATGATGCATGCT CGCGATCAGCCGCTTAAC	BamHI	PafA reverse
GGGAATTC**CATATG**CATCATCATCATCATCACGCGCAAG AGCAGACCAAGCGTGGC	NdeI	PupE forward
GCCCGC**GGATCC**TTATTCTCCGCCCTTTTGGACGTATGC	BamHI	PupE reverse
GGGAATTC**CATATG**CAGATCTTCGTGAAGACTCTGACT	NdeI	Ub forward
ACCGCCACGCTTGGTCTGCTCTTGCGCCATACCACCTCTG AGACGGAG		Ub-PupE fusion
CAGATCGTCGGTCTCCTCGGTCAGCTTACCACCTCTGAG ACGGAGTAC		Ub-PupE34 fusion
GTCGATTTCGTCGAGCAGATCGTCGGTACCACCTCTGAG ACGGAGTAC		Ub-PupE29 fusion
GTCGTCGATTTCGTCGAGCAGATCGTCACCACCTCTGAG ACGGAGTAC		Ub-PupE28 fusion
GACGTCGTCGATTTCGTCGAGCAGATCACCACCTCTGAG ACGGAGTAC		Ub-PupE27 fusion
GAGGACGTCGTCGATTTCGTCGAGCAGACCACCTCTGAG ACGGAGTAC		Ub-PupE26 fusion
CTCGAGGACGTCGTCGATTTCGTCGAGACCACCTCTGAG ACGGAGTAC		Ub-PupE25 fusion
CTCCTCGAGGACGTCGTCGATTTCGTCACCACCTCTGAG ACGGAGTAC		Ub-PupE24 fusion
CTGGTACTCCGTCTCAGAGGTGGTGAAAACCTGTACTTC CAGGGTCAAGAGCAGACCAAGCGTGGCGGT		Ub-TEV-PupE QC forward
ACCGCCACGCTTGGTCTGCTCTTGACCCTGGAAGTACAG GTTTTCACCACCTCTGAGACGGAGTACCAG		Ub-TEV-Pup QC reverse
CTGGTACTCCGTCTCAGAGGTGGTGAAAACCTGTACTTC		Ub-TEV-Pup26

CAGGGTCTGCTCGACGAAATCGACGACGTC		QC forward
GACGTCGTCGATTTCGTCGAGCAGACCCTGGAAGTACAG GTTTTCACCACCTCTGAGACGGAGTACCAG		Ub-TEV-Pup26 QC reverse
GGGAATTC**CATATG**GTGAGCAAGGGCGAGGAGCTG	NdeI	EGFP forward
GACGTCGTCGATTTCGTCGAGCAGCTTGTACAGCTCGTC CATGCC		EGFP-Pup26 fusion
GTCCGCGCATACGTCCAAAAGGGCGGACAGTAAGGATC CGGCTGCTAACAAAGCCCG		Pup E64Q QC forward
CGGGCTTTGTTAGCAGCCGGATCCTTACTGTCCGCCCTTT TGGACGTATGCGCGGAC		Pup E64Q QC reverse
GAAAAGCTGACCGAGGAGACCGACGATGCGCTCGACGA AATCGACGACGTCCTCGAG		Pup L39A QC forward
CTCGAGGACGTCGTCGATTTCGTCGAGCGCATCGTCGGT CTCCTCGGTCAGCTTTTC		Pup L39A QC reverse
GAGGAGACCGACGATCTGGCCGACGAAATCGACGACGT C		Pup L40A QC forward
GACGTCGTCGATTTCGTCGGCCAGATCGTCGGTCTCCTC		Pup L40A QC reverse
AGCTGACCGAGGAGACCGACGATGCGGCGGACGAAATC GACGACGTCCTCGAG		Pup L39AL40A QC forward
CTCGAGGACGTCGTCGATTTCGTCCGCCGCATCGTCGGT CTCCTCGGTCAGCTT		Pup L39AL40A QC reverse
GAGGAGACCGACGATCTGCTCGACGAAGCGGACGACGT CCTCGAGGAGAACGCCGAGGAC		Pup I43A QC forward
GTCCTCGGCGTTCTCCTCGAGGACGTCGTCCGCTTCGTCG AGCAGATCGTCGGTCTCCTC		Pup I43A QC reverse
GATCTGCTCGACGAAATCGACGACGCGGCGGAGGAGAA CGCCGAGGACTTCGTC		Pup V46AL47A QC forward
GACGAAGTCCTCGGCGTTCTCCTCCGCCGCGTCGTCGAT TTCGTCGAGCAGATC		Pup V46AL47A QC reverse

General Method for Expression and Ni^{2+}-affinity Purification of His_6-PupE and PafA-His_6 proteins.

All proteins were expressed in *E. coli* BL21(DE3) cells from isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible plasmids. Cells were grown in 3 L baffled-flasks at 37°C until they reached an OD_{600} ~0.4–0.5, then cooled down to 25°C on the bench-top. Protein expression was induced with 0.3 mM IPTG. The induced culture was grown at 25°C for an additional 2 h for PafA-His_6 and 5 h for His_6-PupE, His_6-Ub-PupE, and His_6-Ub-TEV-PupE and all mutants/truncants thereof. Cells were harvested by centrifugation at 6000x*g* for 20 min and resuspended in a buffer consisting of 50 mM Tris, 200 mM NaCl, pH 7.4 for PupE, and 50 mM Tris, 300 mM NaCl, 1mM β-mercaptoethanol, 10% (v/v) glycerol, pH 7.4 for PafA. Cells were lysed by sonication and the insoluble cellular debris cleared by centrifugation at 20,000x*g* for 30 min, followed by filtration of the supernatant through a 0.45 μm PVDF syringe-filter. All proteins were initially purified by Ni^{2+}-affinity chromatography, either by gravity flow over HisPur™ Ni-NTA resin (Thermo Scientific, Waltham, MA) or using a HiTrap™ IMAC Fast Flow column (GE Healthcare) using the AKTA FPLC system by following the manufacturers instructions for protein elution with imidazole. Protein containing fractions were identified by 10% or 15% SDS-PAGE, combined and concentrated before further purification.

Purification of PafA-His_6

Concentrated Ni^{2+}-column eluates containing PafA-His_6 were further purified by size-exclusion chromatography on a Superdex 200 10/300 GL size-exclusion column (GE Healthcare) attached to the AKTA FPLC system. The elution buffer contained 50 mM Tris, 300 mM NaCl, 1 mM EDTA, 1 mM 1,4-dithio-D-threitol (DTT), and 10% (v/v) glycerol at pH 7.5 and was used at a flow-rate of 0.4 mL/min. Fractions containing pure PafA-His_6 (R_t ~16 min) were identified by 10% SDS-PAGE, combined and concentrated by 10,000 M.W. cut-off centrifugal filters (Sartorius). Stock PafA-His_6 concentrations were determined by comparison with known BSA standards after 10% SDS-PAGE and coomassie-staining.

Purification of His_6-PupE, His_6-PupE(L39A), His_6-PupE(L40A), His_6-PupE(L39A,L40A), His_6-PupE(I43A), and His_6-PupE(V46A,L47A)

Ni^{2+}-column eluate fractions containing His_6-PupE or the mutants His_6-PupE(L39A), His_6-PupE(L40A), His_6-PupE(L39A,L40A), His_6-PupE(I43A), and His_6-PupE(V46A,L47A) were further purified by C18 preparative RP-HPLC using a gradient of 30-55% buffer B over 60 min. Fractions containing pure proteins were identified by C18 or C4 analytical RP-HPLC using a gradient of 0-73% buffer B over 30 min, and the identity of each protein was confirmed by ESI-MS or MALDI-TOF-MS. The calculated extinction coefficient (ε) of 1490 M^{-1} cm^{-1} (ExPASy ProtParam)[3] was used to determine protein concentration from Abs_{280} measurements.

Purification of G-PupE, G-PupE(39-64) and G-PupE(L39A,L40A)

Ni^{2+}-column eluates containing His_6-Ub-TEV-PupE, His_6-Ub-TEV-PupE(39-64) or His_6-Ub-TEV-PupE(L39A,L40) were dialyzed against water overnight to remove imidazole. The His_6-Ub-tag was subsequently removed by proteolysis with TEV protease in a buffer containing 50 mM Tris-Cl, 1 mM EDTA, 10 mM DTT and 10 mM Cys, pH 6.9. The resulting G-PupE, G-PupE(39-64)and G-PupE(L39A,L40A) polypeptides were purified on a Vydac C18 preparative column and characterized by C4 analytical RP-HPLC and MALDI-TOF MS. The pure proteins were resuspended in sterile distilled water and their concentration was determined by amino acid analysis prior to measuring the kinetics of their phosphorylation by PafA.

Synthesis of fluorescent amines

Synthesis of Nα-FITC-L-Lys (2). Nε-Boc-L-Lys (41.26 mg, 0.168 mmol, 1.3 equiv) was dissolved in a mixture of 10 mL aqueous $NaHCO_3$ (0.1 M, pH 9.0) and 2 mL acetonitrile, on ice. A solution of fluorescein-5-isothiocyanate (50 mg, 0.128 mmol, 1.0 equiv) in 2.5 mL acetonitrile was added and the resultant mixture stirred for 18 h at 25 °C with the occlusion of light. At the end of this period, the mixture was neutralized by the drop wise addition of 1 M HCl, which resulted in the formation of a yellow precipitate containing the crude Boc-protected product. The precipitate was filtered on a Büchner funnel and the Boc- group subsequently removed by treatment with 3 mL of a 95:5 (v/v) mixture of TFA:H_2O for 2 h at 25 °C. The TFA was removed by rotary evaporation and the crude product **2** diluted in 20 mL MilliQ water. Pure product **2** (33.19 mg, 48% overall yield) was obtained after purification by C18 preparative HPLC. ^{1}H NMR (500 MHz, MeOD): δ 8.38 (s, 1H, H_{ar}), 7.97 (d, 1H, H_{ar}), 7.25 (d, 1H, H_{ar}), 7.05 (d, 2H, H_{ar}), 6.9 (s, 2H, H_{ar}), 6.78 (d, 2H, H_{ar}), 5.09 (m, 1H, NHC*H*), 3.07 (m, 2H, NHCH_2), 2.10-1.9 (m, 2H, CHCH_2), 1.8-1.7 (m, 4H, CH_2CH_2). ESI-MS m/z calculated for $C_{27}H_{25}N_3O_7S$ $[M+H]^+$ 536.14 Da, observed 536.30 Da.

Synthesis of 5-((5-aminopentyl)thioureidyl)fluorescein (3). 1,5-diaminopentane (39.50 mg, 0.387 mmol, 3.0 equiv) was dissolved in 10 ml of aqueous $NaHCO_3$ (0.1 M, pH 9.14) on ice. A solution of fluorescein-5-isothiocyanate (50 mg, 0.128 mmol, 1.0 equiv) in 2.5 ml of acetonitrile was added and the resultant mixture stirred for 12 h at 25 °C. The crude reaction mixture showed several products that were

inseparable by C18 RP-HPLC. Therefore the reaction mixture was further subjected to amine protection with Boc_2O. Briefly, solid $NaHCO_3$ (130 mg, 1.548 mmol, 12.0 equiv) was added to the pre-cooled reaction mixture on ice with vigorous stirring. An ice-cold solution of Boc_2O (253.30 mg, 1.161 mmol, 9.0 equiv) in *para*-dioxane was added to the stirred reaction mixture. The reaction was allowed to proceed for 1 hour at 0 °C and then warmed to room temperature overnight with constant stirring. The reaction mixture was diluted with 15 mL H_2O and the aqueous layer was extracted with 2 x 25 mL EtOAc. The organic layer was then back extracted twice with 10 mL saturated aqueous $NaHCO_3$. The combined aqueous layers were neutralized by the drop wise addition of 1 M HCl and extracted with 3 x 50 mL EtOAc. The solvent was dried over $MgSO_4$ and evaporated under reduced pressure. The product mixture was purified by C18 preparative HPLC on 25-60% B gradient over 60 minutes. Finally, the Boc-protecting group was removed from the lyophilized solid by the addition of 2 mL of 1:1 (v/v) TFA:DCM and stirring for 1 hour. Solvents were removed under reduced pressure and the final product washed with 3 x 3mL of DCM prior to drying *in vacuo*. This yielded 6.80 mg of compound **3** in 11.1% overall yield. ^{1}H NMR (500 MHz, CD_3OD): δ 8.1283(s, 1H, H_{ar}), 7.66(d, 1H, H_{ar}), 7.05(d, 1H, H_{ar}), 6.79(d, 2H, H_{ar}), 6.71(s, 2H, H_{ar}), 6.58(d, 2H, H_{ar}) 3.16 (d, 2H, NHCH_2), 2.80(t, 2H, CH_2NH$_2$), 1.57(t, 4H, CH_2), 1.11(d, 2H, CH_2). ESI-MS m/z calculated for $C_{26}H_{25}N_3O_5S$ $[M+H]^+$ 492.56 Da, observed 492.20 Da.

Synthesis of 5-((4-aminobutyl)thioureidyl)fluorescein (4). 1,4-diaminobutane (34.10 mg, 0.387 mmol, 3.0 equiv) was dissolved in a mixture of 10 mL aqueous $NaHCO_3$ (0.1 M, pH 9.0) and 2.5 mL acetonitrile on ice. A solution of fluorescein-5-isothiocyanate (50 mg, 0.128 mmol, 1.0 equiv) in 2.5 mL acetonitrile was added to the solution containing the amine and stirred for 18 h in the absence of light. At the end of this period, the mixture was neutralized by drop wise addition of 1 mL, 1 M HCl. The solution was diluted using 20 mL water and purified by C18 preparative HPLC to yield 41.80 mg of pure FITC-1,4-diaminobutane in 68% overall yield. ^{1}H NMR (500 MHz, D_2O): δ 7.56 (d, 1H, H_{ar}), 7.55 (s, 1H, H_{ar}), 6.97 (d, 1H, H_{ar}), 6.53 (d, 2H, H_{ar}), 6.41 (d, 2H, H_{ar}), 6.42 (s, 2H, H_{ar}), 2.99 (t, 2H, NHCH_2), 1.63 (m, 2H, CH_2NH$_2$), 1.44 (m, 4H, CH_2CH_2). ESI-MS m/z calculated for $C_{25}H_{23}N_3O_5S$ $[M+H]^+$ 478.14 Da, found 478.20 Da.

Synthesis of 5-((2-aminoethyl)thioureidyl)fluorescein (5). *tert*-butyl(2-aminoethyl)carbamate (50 mg, 0.314 mmol, 1.2 equiv) was dissolved in a mixture of 10 mL aqueous $NaHCO_3$ (0.1 M, pH 9.0) and 2.5 mL acetonitrile on ice. A solution of fluorescein-5-isothiocyanate (100 mg, 0.257 mmol, 1.0 equiv) in 2.5 mL acetonitrile was added to the solution containing the amine and stirred for 18 h in the absence of light. At the end of this period, the mixture was neutralized by the drop wise addition of 1 M HCl. The solution was diluted using 20 mL water and Boc-protected product purified by C18 preparative HPLC. After lyophilization to remove solvents, the Boc- group was removed by treatment with a mixture of 95:5 TFA:H_2O for 2 h at 25 °C. The deprotected crude was further purified by C18 preparative HPLC to yield 98.90 mg of pure **5** in 89% overall yield. ^{1}H NMR (500 MHz, MeOD): δ 8.20 (s, 1H, H_{ar}), 7.81 (d, 1H, H_{ar}), 7.24 (d, 1H, H_{ar}), 6.76 (s, 2H, H_{ar}), 6.75 (d, 2H, H_{ar}), 6.62 (d, 2H, H_{ar}), 4.01 (m, 2H, NHCH_2), 3.28 (m, 2H, CH_2CH_2). ESI-MS m/z calculated for $C_{23}H_{19}N_3O_5S$ $[M+H]^+$ 450.11 Da, found 450.8 Da.

Synthesis of Nε-FITC-L-Lys (6). Nα-Boc-L-Lys (50 mg, 0.203 mmol, 1.6 equiv) was dissolved in a mixture of 10 mL aqueous $NaHCO_3$ (0.1 M, pH 9.0) and 2 mL acetonitrile, on ice. A solution of fluorescein-5-isothiocyanate (50 mg, 0.128 mmol, 1.0 equiv) in 2.5 mL acetonitrile was added and the resultant mixture stirred for 18 h at 25 °C with the occlusion of light. At the end of this period, the mixture was neutralized by the drop wise addition of 1 M HCl, which resulted in the formation of a yellow precipitate containing the crude Boc-protected product. The precipitate was filtered on a Büchner funnel and the Boc- group subsequently removed by treatment with 3 mL of a 95:5 (v/v) mixture of TFA:H_2O for 2 h at 25 °C. The TFA was removed by rotary evaporation and the crude product **6** diluted in 20 mL MilliQ water. Pure product **6** (28 mg, 41% overall yield) was obtained after purification by C18 preparative HPLC. ^{1}H NMR (500 MHz, MeOD): δ 8.17 (s, 1H, H_{ar}), 7.77 (dd, 1H, H_{ar}), 7.20 (d, 1H, H_{ar}), 6.76 (s, 1H, H_{ar}), 6.75 (s, 1H, H_{ar}), 6.74 (s, 1H, H_{ar}), 6.62 (s, 1H, H_{ar}), 6.60 (s, 1H, H_{ar}), 4.03 (t, 1H, CH),

3.70 (br. s, 2H, NHCH_2)), 2.07 (m, 1H, CH_2CH_2), 1.98 (m, 1H, CH_2CH_2), 1.78 (m, 1H, CH_2CH_2), 1.7-1.5 (m, 3H, CH_2CH_2). ESI-MS m/z calculated for $C_{27}H_{25}N_3O_7S$ $[M+H]^+$ 536.14 Da, observed 536.10 Da.

Synthesis of Nα-TAMRA-L-Lys (7). A mixture of 5(and 6)-carboxymethylrhodamine succinimidyl ester (5(6)-TAMRA, SE) (10 mg, 0.0189 mmol, 1 equiv) was dissolved in 0.5 ml of acetonitrile and added to a stirred solution of Nα-Boc-L-Lys (10 mg, 0.0406 mmol, 2.1 equiv) dissolved in 2 ml of 0.1M aqueous $NaHCO_3$, pH 8.31, on ice. The mixture was allowed to warm up to room temperature overnight with stirring, protected from light. The basic reaction mixture was neutralized with the dropwise addition of 1M HCl. The desired product was obtained by solvent extraction with 3 x 50 mL EtOAc. The combined organic layers were washed with brine and dried over $MgSO_4$. The solvent was evaporated and the resultant solid dried *in vacuo* overnight. The mixture of Boc-protected products was dissolved in 4 mL of a 1:1 mixture of TFA:DCM and stirred for 1 h. The volatile solvents were evaporated under reduced pressure and the final product was purified by C18 preparative HPLC employing a gradient of 0-73% B over 60 min to yield 4.10 mg of **7** with an overall yield of 38.8%. ESI-MS m/z calculated for $C_{31}H_{34}N_4O_6$ $[M+H]^+$ 559.25 Da, observed 559.10 Da.

General Protocol for Assays with Purified Substrates

Pupylation assays were typically undertaken in a 20 μL volume with 20 μM *wild-type,* mutant, or truncated forms of PupE in the presence of 0.5-1.0 μM PafA, 20-40 μM ATP, and 500-1000 μM of each probe for 10 min-1 h at 37 °C, unless indicated otherwise. The reaction buffer consisted of 50 mM Tris-Cl, 300 mM NaCl, 1 mM DTT, 20 mM $MgCl_2$, and 10% (v/v) glycerol at pH 7.4. The assays were quenched by the addition of 6x SDS gel-loading buffer, and analyzed by resolving on 15% SDS-PAGE. Fluorescent gels were gently nutated in water for 30 min to remove unconjugated probes from the gel prior to imaging. The PupE-fluorophore conjugate was subsequently observed by in-gel fluorescence emission at 521 nm for probe **1** and the FITC-derivatives **2**-**6** and 583 nm for the TAMRA derivative, **7**, using a Typhoon FLA 9000 scanner (GE Healthcare). Following fluorescence imaging, the gels were stained in Coomassie Brilliant Blue to quantify protein loading.

General Protocol for Dose-response and Time-point Assays with Purified Substrates

Pupylation assays were undertaken with 10 μM *wild-type* or mutant forms of PupE in the presence of 1 μM PafA, 20 μM ATP, and 25, 50, 100 and 250 μM of each probe in a 10 μL total volume (for dose-response assays), or 250 μM of each probe in a 50 μL total volume (for time-point assays). The reaction buffer consisted of 50 mM Tris-Cl, 300 mM NaCl, 1 mM DTT, 20 mM $MgCl_2$, and 10% (v/v) glycerol at pH 7.4. Assays were undertaken for 30 min at 37 °C for dose-response studies and quenched by the addition of gel-loading buffer. For the time-point assays conducted at 37 °C, 10 μL aliquots were removed at 30 s, 60 s, 1 min, 2 min, 5 min, and 15 min and quenched by mixing with 5 μL of 6x SDS-PAGE loading buffer. Assay products were analyzed by resolving on 15% SDS-PAGE. Fluorescent gels were gently nutated in water for 30-60 min with frequent washes to remove unconjugated probes from the gel prior to imaging. The PupE-fluorophore conjugate was subsequently observed by in-gel fluorescence emission at 521 nm for FITC-derivatives using a Typhoon FLA 9000 scanner (GE Healthcare). Following fluorescence imaging, the gels were stained in Coomassie Brilliant Blue to quantify protein loading.

General Protocol for Assays with Cellular Lysates

Briefly, separate *E. coli* BL21 (DE3) cell cultures bearing the desired substrates and PafA plasmids were induced to overexpress the target protein as described above. The cells were harvested, resuspended in a reaction buffer containing 50 mM Tris-Cl, 300 mM NaCl, 1 mM DTT, 20 mM $MgCl_2$, and 10% (v/v) glycerol at pH 7.4 and lysed by sonication. Lysates were cleared of insoluble debris by filtration through a 0.45 μM filter and stock aliquots prepared at a total protein concentration of 2 mg/mL (determined by the

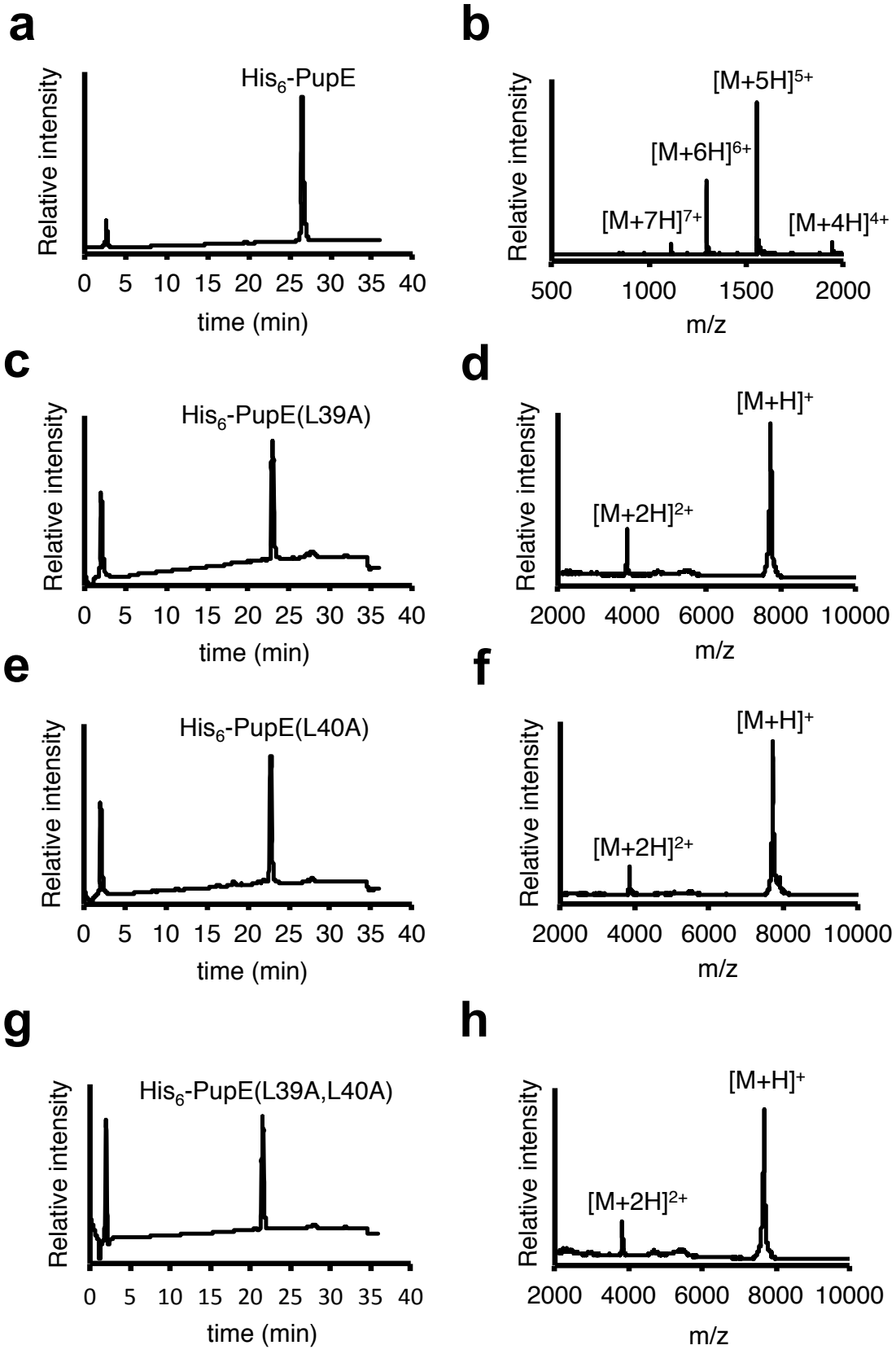
Bradford protein assay from Thermo Scientific) in the reaction buffer. Lysates were added to each pupylation assay at a final concentration of 0.5 mg/mL (10 μg total cellular proteins per 20 μL assay volume). Substrates that demonstrated reduced or no activity in lysate assays were further purified by means of their His_6-tags and re-tested with pure PafA to confirm and quantitate the results obtained with lysates.

General Protocol for Assays with Free Amino Acids

Pupylation assays were undertaken with 20 μM PupE, 0.5 μM PafA, 20 μM ATP, and 100 mM of each amino acid overnight at 37 °C in a reaction buffer consisting of 50 mM Tris-Cl, 300 mM NaCl, 1 mM DTT, 20 mM $MgCl_2$, and 10% (v/v) glycerol at pH 7.4. The samples were analyzed by LC-MS by separation on a C18 analytical column employing a gradient of 0-73%B (B= 99.9% acetonitrile, 0.1% acetic acid, A= 90% H_2O, 9.9% acetonitrile, 0.1% acetic acid) over 30 min followed by ESI-MS on a Bruker Esquire LC-ion trap spectrometer.

Pupylation assays with α-^{32}P-ATP

All assays were carried out at 25 °C in 20 μL reaction volumes, in a buffer containing 50 mM Tris, 300 mM NaCl, 1 mM DTT, 20 mM $MgCl_2$, and 10% (v/v) glycerol, pH 7.4. Assays contained 20 μM of each Pup substrate and 4 μM of ATP containing α-^{32}P-ATP at 166 mCi/mmol, and were initiated by the addition of 1 μM PafA-His_6. At 7, 15, 30, 45, 60, 90, 120, 180, 240 and 480 s time-points, 2 μL aliquots were withdrawn and immediately quenched by dilution into an equal volume of 5 M aqueous formic acid. Half-microliter aliquots of the quenched samples were spotted on 10 x 20 cm PEI Cellulose TLC plates, 1.5 cm from the bottom edge, and allowed to dry. The plates were developed in a buffer consisting of 1 M aqueous formic acid and 0.5 M LiCl and dried under a stream of air. A phosphor screen (GE Health Sciences) was exposed to the dried TLC plates for 1.5 h and imaged using a Typhoon FLA 9000 scanner. Images were analyzed by ImageJ to determine the ratio of [ADP] produced to total nucleotide present in the assay [ATP+ADP].[4] The data was plotted using Kaleidagraph and fit to the equation $[ADP]_t = [ADP]_0 + [ADP]_f(1-e^{-kt})$ in order to obtain the apparent rate constants for phosphorylation of *wild-type* and mutant Pup.



Continued on the next page

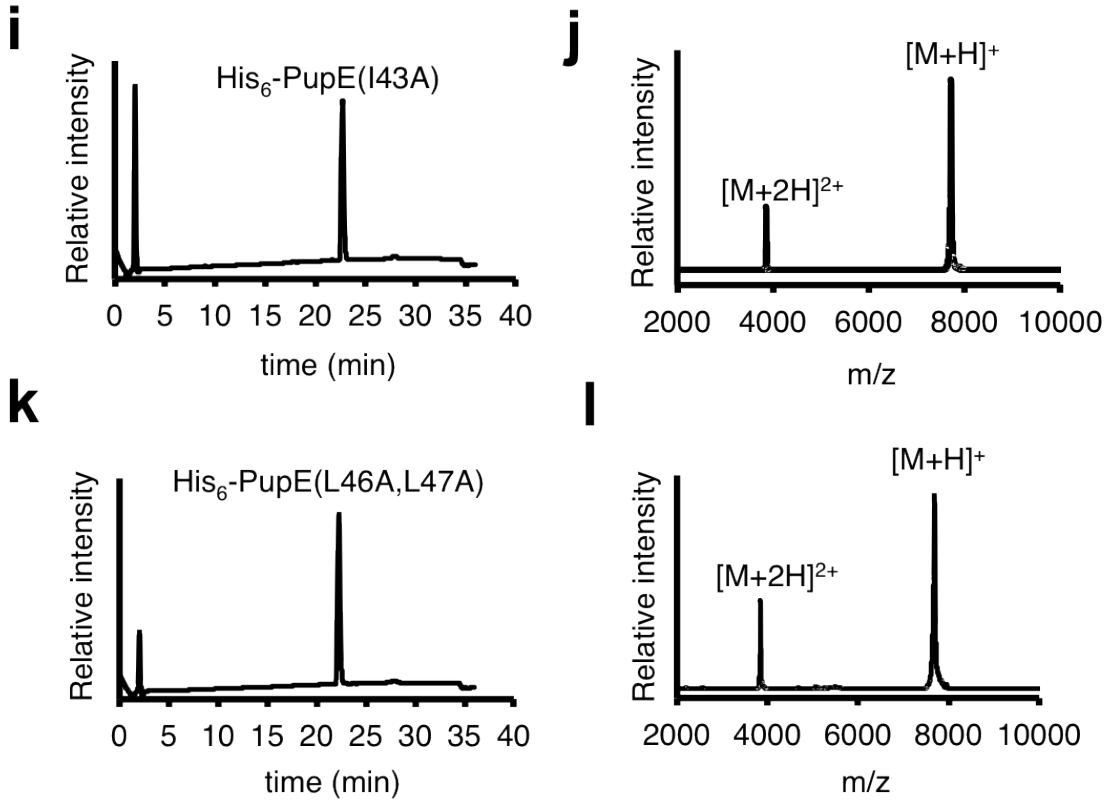


Figure S1. Characterization of N-terminally His_6-tagged wild-type and mutant *Mtb* PupE. (**a**) C18 analytical RP-HPLC chromatogram of purified His_6-PupE. (**b**) ESI-MS of purified His_6-PupE. Calculated, $[M+H]^+$ 7,769.1 Da, observed 7,768.9 ± 4 Da. (**c**) C4 analytical RP-HPLC chromatogram of purified His_6-PupE(L39A). (**d**) MALDI-TOF-MS of purified His_6-PupE(L39A). Calculated, $[M+H]^+$ 7,727.0 Da, observed 7,725.6 ± 4 Da. (**e**) C4 analytical RP-HPLC chromatogram of purified His_6-PupE(L40A). (**f**) MALDI-TOF-MS of purified His_6-PupE(L40A). Calculated, $[M+H]^+$ 7,727.0 Da, observed 7,724.0 ± 4 Da. (**g**) C4 analytical RP-HPLC chromatogram of purified His_6-PupE(L39A,L40A). (**h**) MALDI-TOF-MS of purified His_6-PupE(L39A,L40A). Calculated, $[M+H]^+$ 7,684.9 Da, observed 7,682.6 ± 4 Da. (**i**) C4 analytical RP-HPLC chromatogram of purified His_6-PupE(I43A). (**j**) MALDI-TOF-MS of purified His_6-PupE(I43A). Calculated, $[M+H]^+$ 7,727.0 Da, observed 7,723.2 ± 4 Da. (**k**) C4 analytical RP-HPLC chromatogram of purified His_6-PupE(V46A,L47A). (**l**) MALDI-TOF-MS of purified His_6-PupE(V46A,L47A). Calculated, $[M+H]^+$ 7,699.0 Da, observed 7,693.1 ± 4 Da. All RP-HPLC gradients were 0-73% B, 30 min.

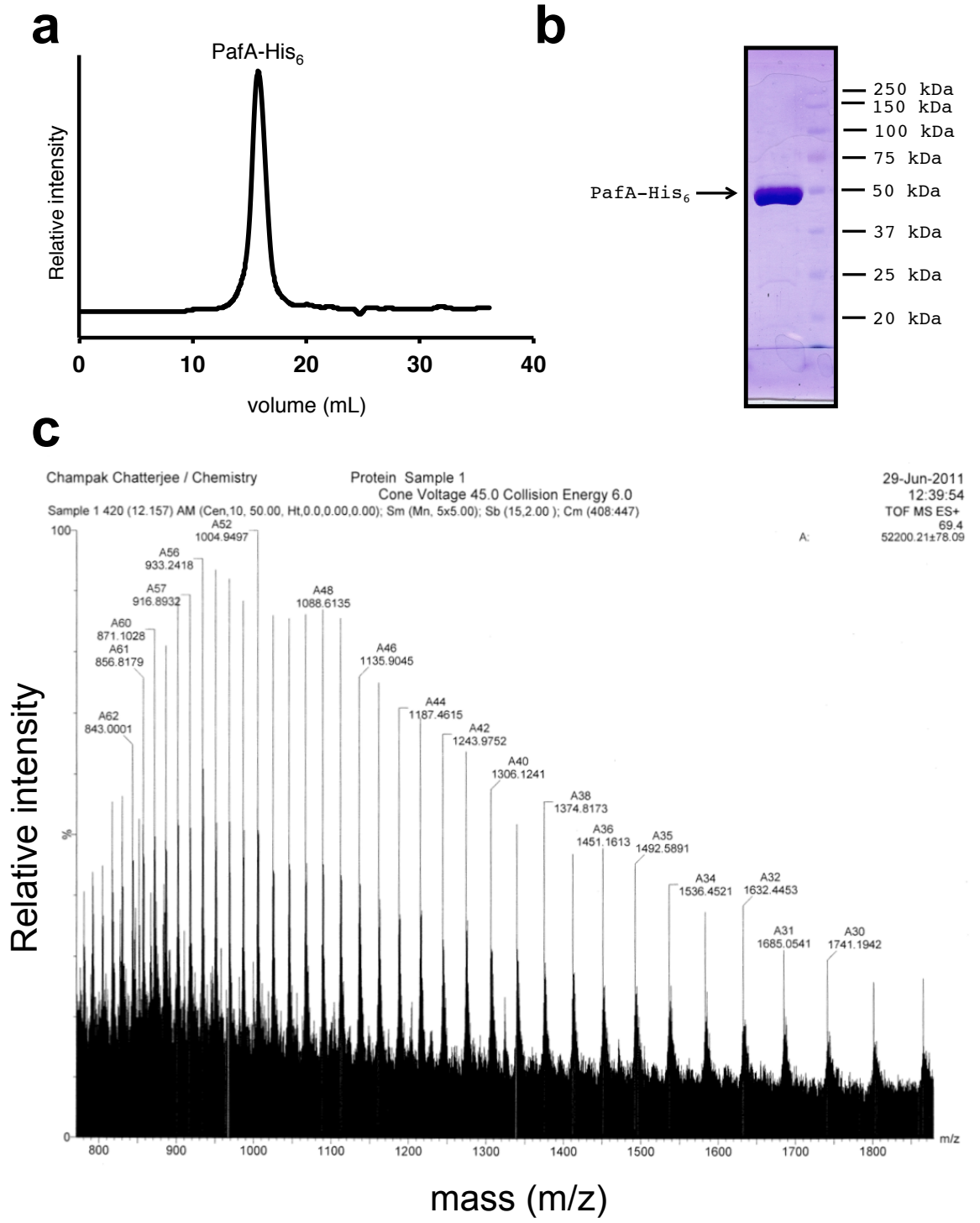


Figure S2. Purification of C-terminally His_6-tagged *Mtb* PafA. (**a**) FPLC chromatogram of pure PafA-His_6 visualized by Abs 280 nm during gel filtration on S-200 column. (**b**) Coomassie-stained 10% SDS-PAGE gel of pure PafA-His_6. (**c**) ESI-time-of-flight-MS of PafA-His_6 after initial desalting by LC. Calculated, $[M+H]^+$ 52,207.9 Da, observed 52,200.21 ± 78 Da.

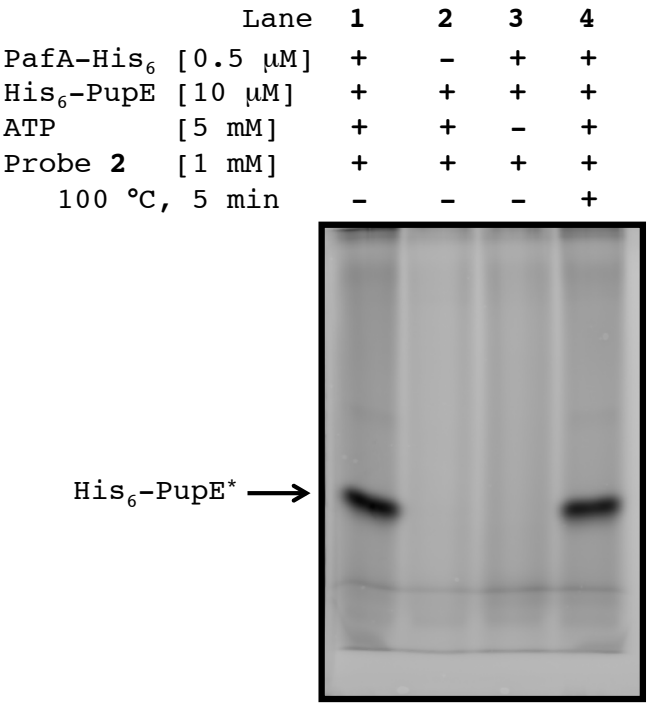


Figure S3. PafA-His$_6$ catalyzed labeling of His$_6$-PupE with Probe 2. In-gel fluorescence after 15% SDS-PAGE, showing PafA and ATP dependent labeling of His$_6$-PupE with probe 2. In order to test the covalent linkage between His$_6$-PupE and probe **2**, assay products in lane 4 were boiled for 5 min prior to loading on SDS-PAGE. His$_6$-PupE* indicates the fluorescent His$_6$-PupE-**2** conjugate.

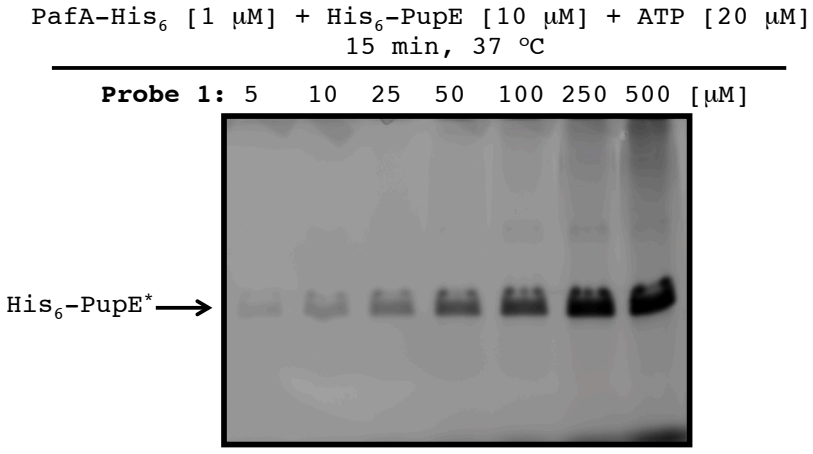
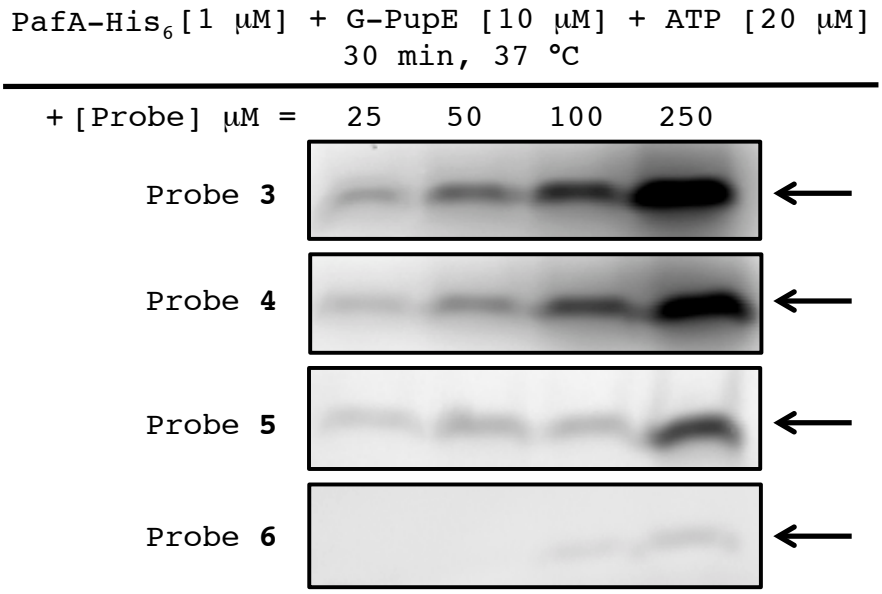


Figure S4. Detection of PafA-His$_6$ activity at varying concentrations of Probe 1. 15% SDS-PAGE of pupylation assays containing PafA-His$_6$, His$_6$-PupE, ATP and probe **1** for 15 minutes at 37°C. His$_6$-PupE* indicates the fluorescent His$_6$-PupE-**1** conjugate.

a



b

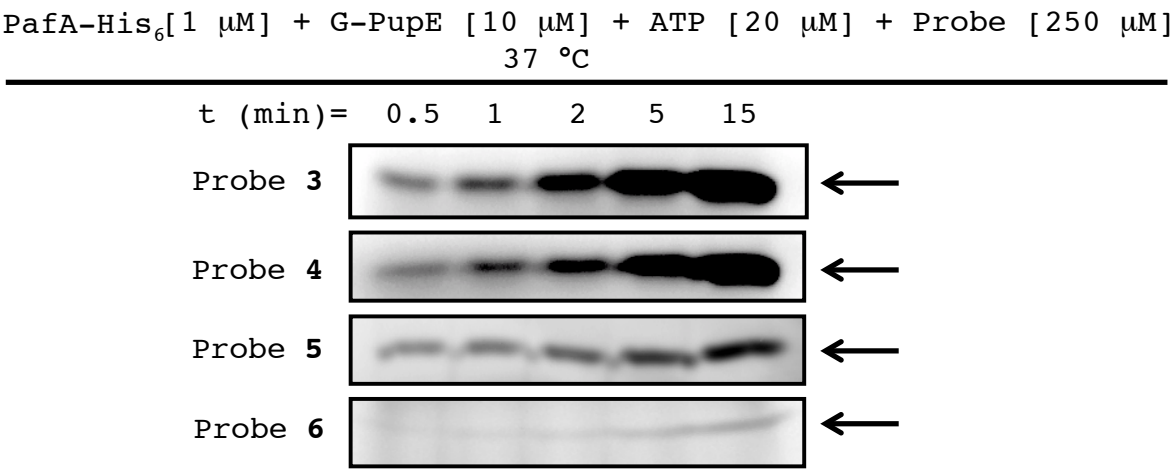


Figure S5. Concentration and time-dependence of G-PupE labeling with probes 3-6. (a) In-gel fluorescence from 15% SDS-PAGE gels of pupylation assays containing PafA-His_6, G-PupE, ATP and the indicated probes **3**-**6** at concentrations of 25, 50, 100, and 250 μM. Arrows indicate the probe-labeled G-PupE. (**b**) In-gel fluorescence from 15% SDS-PAGE gels showing the time-course of pupylation assays containing PafA-His_6, G-PupE, ATP and the indicated probes **3**-**6** at 250 μM. The gels in (**b**) were slightly overexposed in order to better visualize differences in the early time-points. Arrows indicate probe-labeled G-PupE.

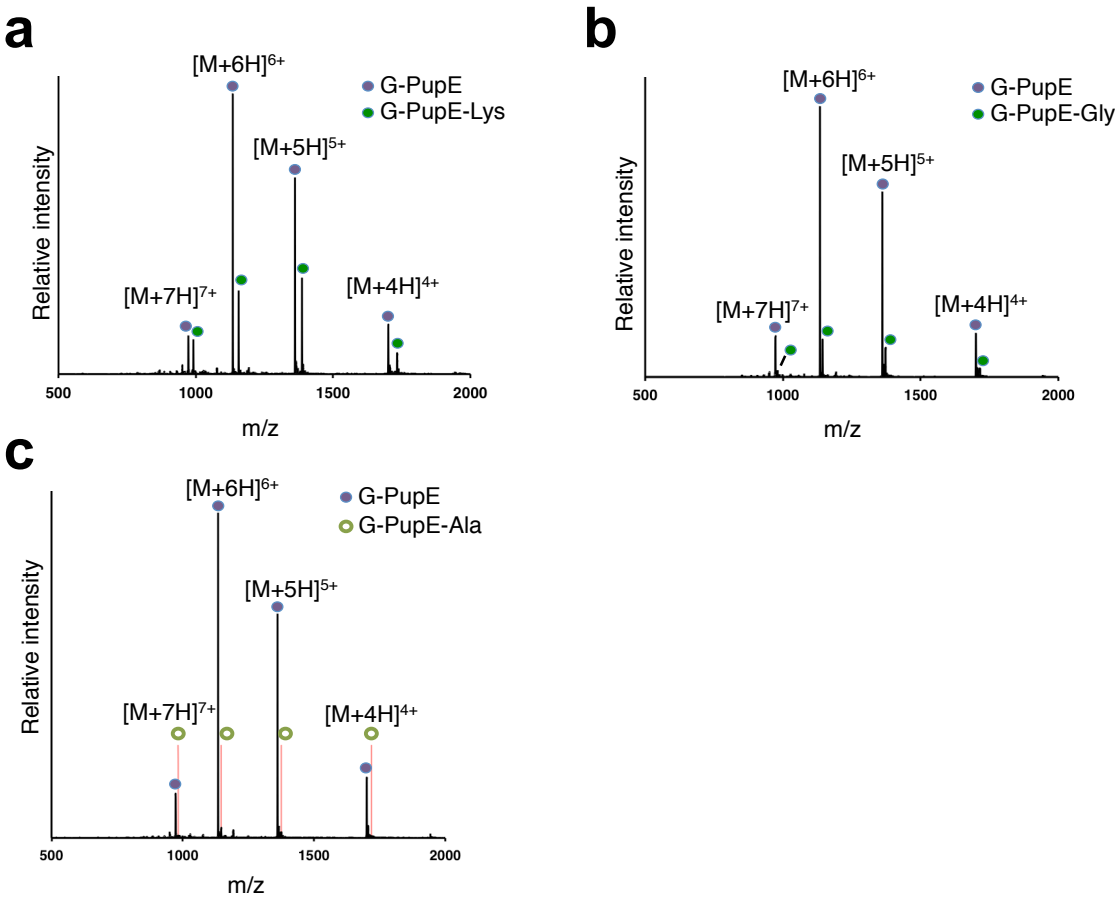


Figure S6. LC-MS analysis of PafA-His$_6$ activity with Lys, Gly, and Ala as substrates. (**a**) ESI-MS of assay products upon reacting 20 μM G-PupE with 0.5 μM PafA-His$_6$, 20 μM ATP and 100 mM Lys for 12 h at 37 °C. Calculated, $[M+H]^+$ 6,929.2 Da, observed 6,929.9 ± 1.4 Da. (**b**) ESI-MS of assay products upon reacting 20 μM G-PupE with 0.5 μM PafA-His$_6$, 20 μM ATP and 100 mM Gly for 12 h at 37 °C. Calculated, $[M+H]^+$ 6,858.1 Da, observed 6,858.5 ± 1.8 Da. (**c**) ESI-MS of assay products upon reacting 20 μM G-PupE with 0.5 μM PafA-His$_6$, 20 μM ATP and 100 mM Ala for 12 h at 37 °C. Calculated, $[M+H]^+$ 6,872.1 Da, observed 6,800.5 ± 1.2 Da. Peaks corresponding to a G-Pup-Ala conjugate are modeled in red and labeled with a hollow green circle.

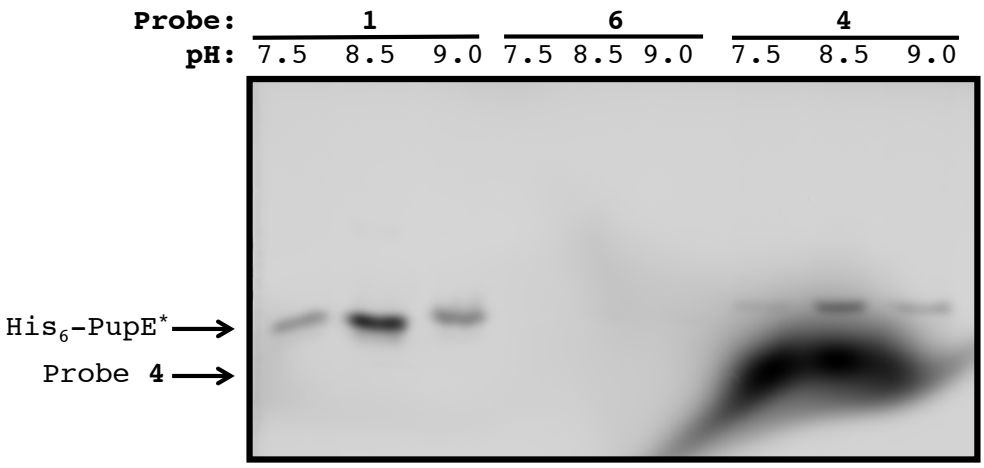


Figure S7. Labeling of His_6-Pup at different pHs. Assays containing 0.5 μM PafA-His_6, 20 μM His_6-PupE, 20 μM ATP, and 500 μM of probe **1, 4** or **6** at the indicated pH of 7.5, 8.5 or 9.0 were incubated at 37 °C for 10 minutes. His_6-$PupE^*$ indicates the fluorescent His_6-PupE-probe conjugate. Excess probe **4** in the gel is indicated with an arrow.

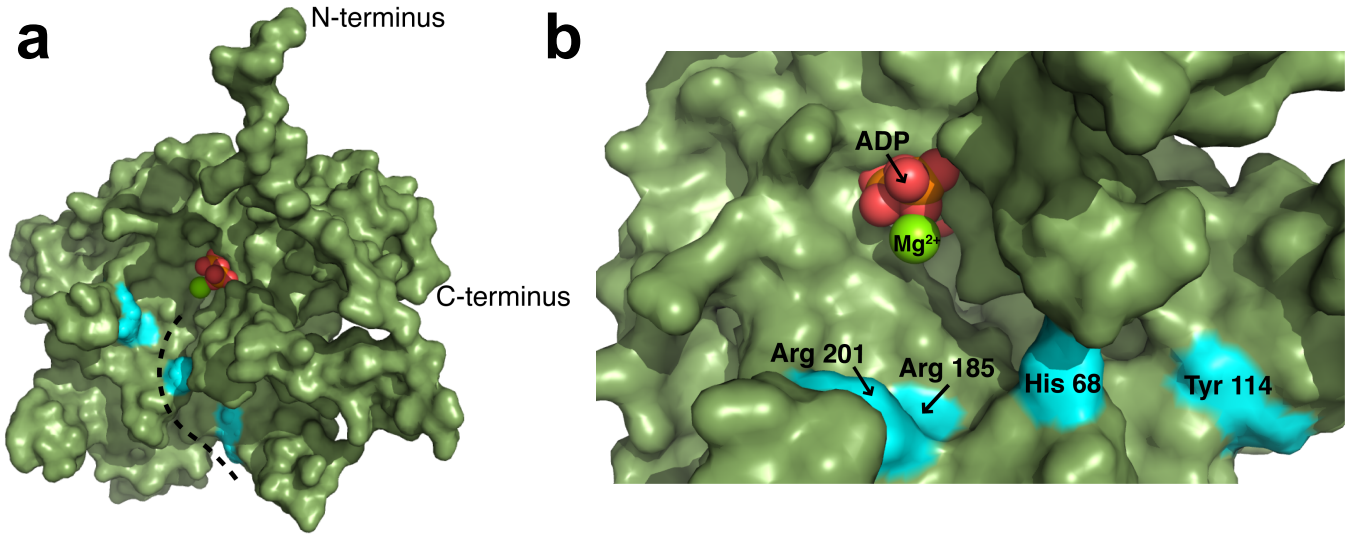


Figure S8. Structure of PafA from *Corynebacterium glutamicum*. (**a**) The structure of PafA from *C. glutamicum* (PDB code 4B0T) rendered in surface representation with PyMol software (DeLano Scientific LLC). The bound ADP and Mg^{2+} ions are shown in space-filling representation. A dotted line indicates the putative Pup-binding groove.[5] (**b**) Close-up of the groove surrounding the ADP-binding site in PafA. Residues in PafA that may be involved in interacting with the Pup polypeptide are shown in cyan.[5]

Figure S9. Sequence alignment of Pup homologues in *actinomycetes*. The sequences of Pup homologues from different *actinomycete* strains were aligned with ClustalW. Residues highlighted in yellow are completely conserved and residues highlighted in pale blue are strongly conserved. Residue numbering is based on the *Mtb* sequence. NCBI accession numbers are, *Mycobacterium tuberculosis H37Rv* (NP_216627.1), *Mycobacterium smegmatis str. MC*^{2} *155* (YP_888186.1), *Mycobacterium marinum M* (YP_001851376.1), *Mycobacterium leprae TN* (NP_301949.1), *Corynebacterium diphtheriae INCA 402* (YP_005127660.1), *Tsukamurella paurometabola DSM 20162* (YP_003647108.1), *Gordonia bronchialis DSM 43247* (YP_003273596.1), *Segniliparus rugosus ATCC BAA-974* (ZP_07965045.1), *Nocardia farcinica IFM 10152* (YP_119384.1), *Amycolicicoccus subflavus DQS3-9A1* (AEF40457.1), *Rodococcus equi 103S* (YP_004007057.1), *Streptomyces lividans TK24* (ZP_06532019.1), *Streptomyces coelicolor A3(2)* (NP625921.1), *Catenulispora acidiphila DSM 44928* (YP_003113184.1), *Acidothermus cellulolyticus 11B* (YP_872945), and *Corynebacterium bovis DSM 20582* (ZP_08516251.1).

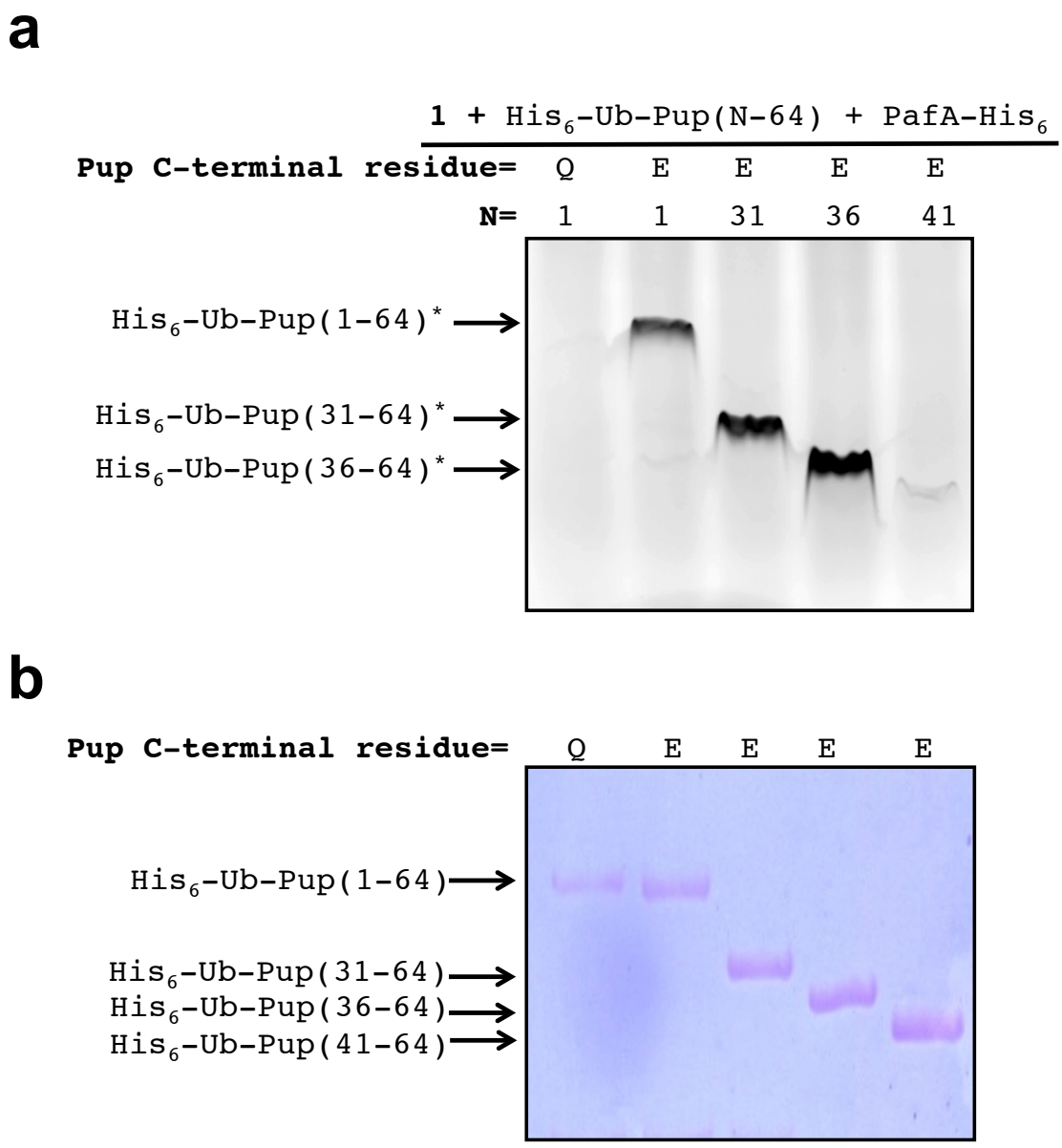


Figure S10. Identification of a minimal pupylation sequence. (**a**) The indicated full-length *wild-type* His_6-Ub-Pup, full-length His_6-Ub-PupE, and His_6-Ub-PupE truncants were directly assayed in cell lysates (2 mg/mL total protein) containing PafA-His_6 and probe **1**. Asterisks indicate the probe-labeled His_6-Ub-Pup fusions. (**b**) Coomassie stained 15% SDS-PAGE gel of protein containing fractions obtained after Ni^{++}-affinity purification of fusion proteins from the cell lysates used in (**a**). Arrows indicate the full-length and truncated His_6-Ub-Pup fusion proteins.

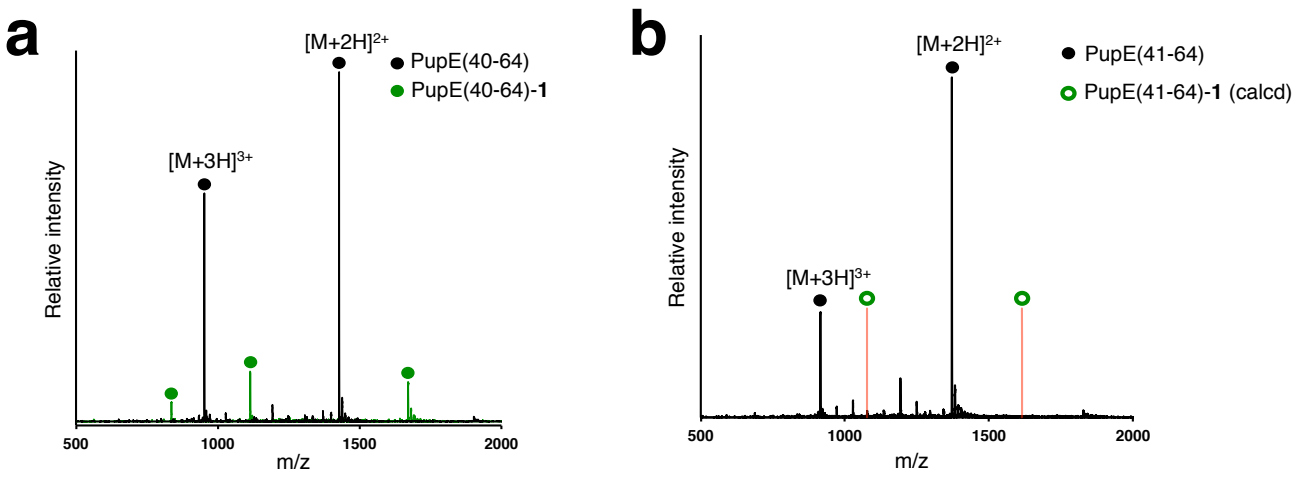


Figure S11. Pupylation assays with synthetic PupE(40-64) and PupE(41-64) peptides. (**a**) LC-ESI-MS of assay products from incubation of 20 μM PupE(40-64) with 0.5 μM PafA-His_6, 20 μM ATP and 1 mM probe **1** for 12 h at 37 °C. Calcd. PupE(40-64)-**1** $[M+H]^+$ 3,341.5 Da, obsd. $[M+H]^+$ 3,341.1 ± 1.4 Da. Unreacted PupE(40-64) calcd. $[M+H]^+$ 2,854.0 Da, obsd. 2854.0 ± 0.6 Da. (**b**) ESI-MS of assay products from incubation of 20 μM PupE(41-64) with 0.5 μM PafA-His_6, 20 μM ATP and 1 mM probe **1** for 12 h at 37 °C. Calcd. PupE(41-64)-**1** $[M+H]^+$ 3,227.3 Da, obsd. $[M+H]^+$ 2741.5 ± 1.4 Da which corresponds to the unmodified PupE(41-64) peptide. Peaks corresponding to a PupE(41-64)-**1** conjugate are modeled in red and labeled with a hollow green circle.

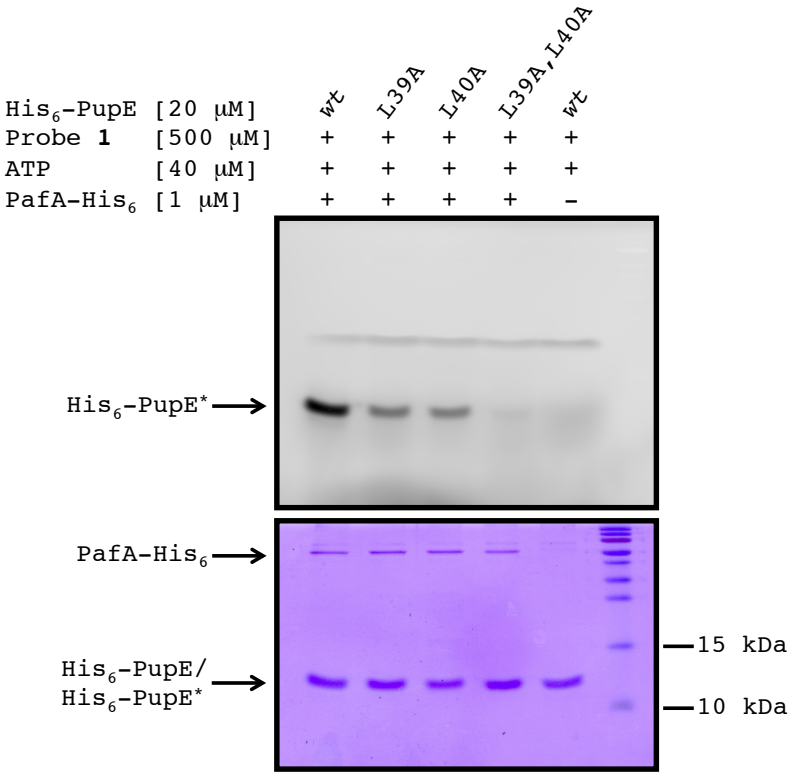


Figure S12. PafA-His$_6$ activity with His$_6$-PupE mutants. The top gel shows in-gel fluorescence after 15% SDS-PAGE of assay products with the indicated *wild-type* and mutant forms of His$_6$-PupE. The bottom gel shows coomassie staining of the top-gel to quantify the loading of each protein. His$_6$-PupE* indicates the position of the Pup-probe conjugate.

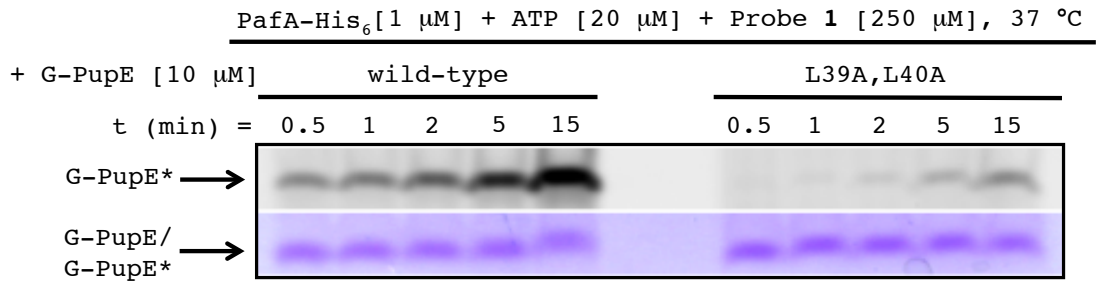


Figure S13. Time-course of PafA-His$_6$ activity with G-PupE and the G-PupE(L39A,L40A) mutant. The top gel shows in-gel fluorescence after 15% SDS-PAGE of assay products with the indicated *wild-type* and mutant forms of G-PupE. The bottom gel shows coomassie staining of the top-gel to verify equal substrate loading. G-PupE* indicates the position of the Pup-probe **1** conjugate.

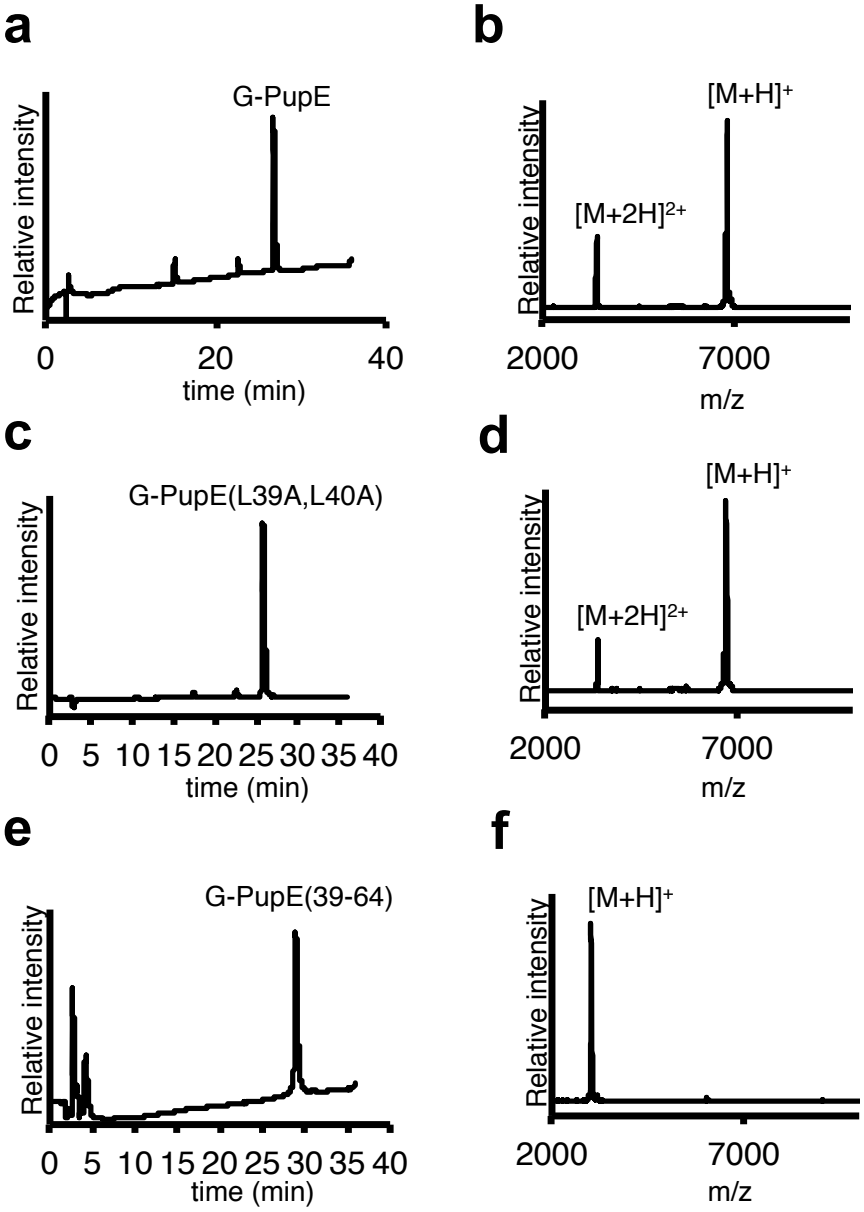


Figure S14. Characterization of tag-less wild-type and mutant *Mtb* PupE. (**a**) C18 analytical RP-HPLC chromatogram of purified G-PupE. (**b**) MALDI-TOF-MS of purified G-PupE. Calculated, $[M+H]^+$ 6,801.0 Da, observed 6,798.8 ± 3 Da. (**c**) C18 analytical RP-HPLC chromatogram of purified G-PupE(L39A,L40A). (**d**) MALDI-TOF-MS of purified G-PupE(L39A,L40A). Calculated, $[M+H]^+$ 6,716.9 Da, observed 6,714.7 ± 3 Da. (**e**) C18 analytical RP-HPLC chromatogram of purified G-PupE(39-64). (**f**) MALDI-TOF-MS of purified G-PupE(39-64). Calculated, $[M+H]^+$ 3,025.2 Da, observed 3,025.3 ± 2 Da. All RP-HPLC gradients were 0-73% B, 30 min.

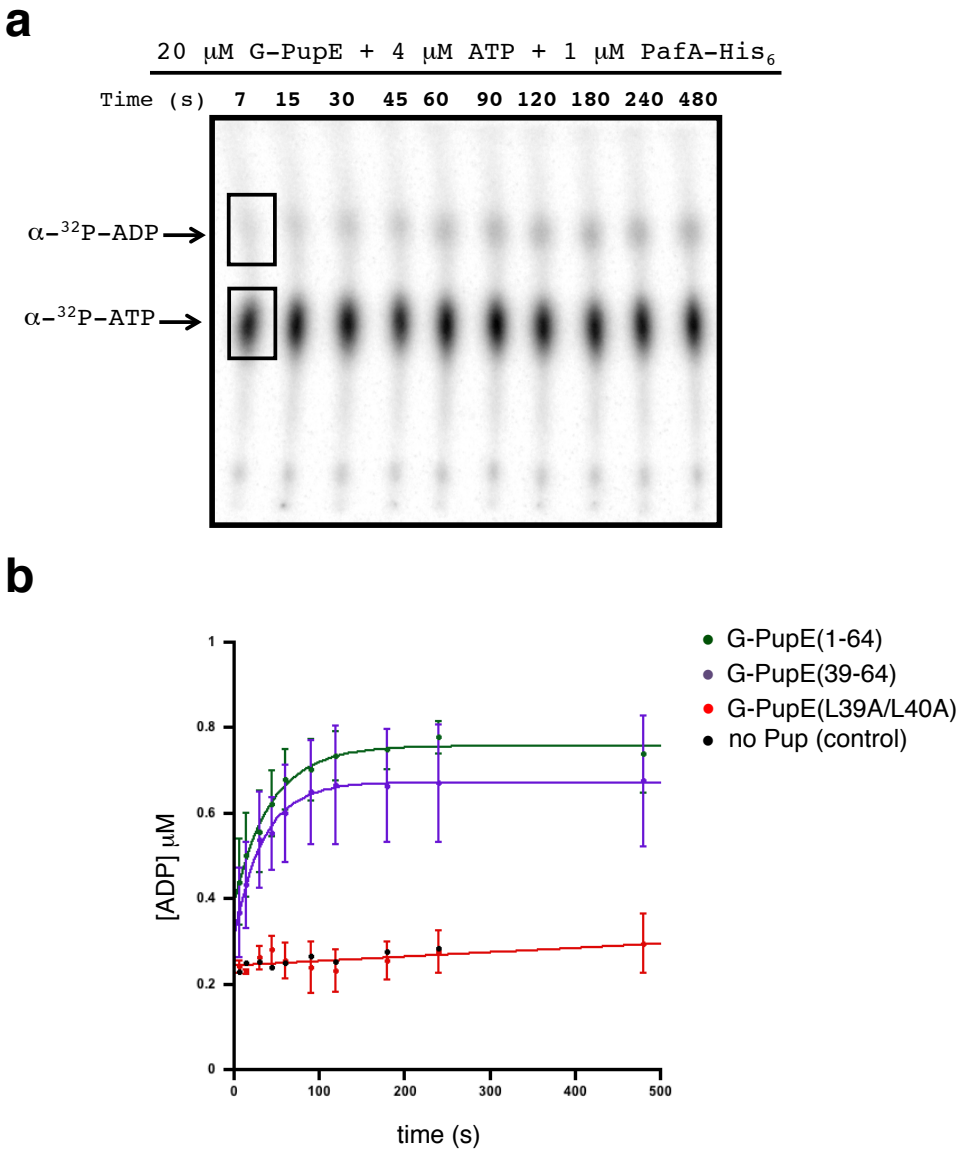


Figure S15. Kinetics of phosphorylation of *wild-type* and mutant G-PupE by PafA-His$_6$. (**a**) Representative TLC image of reaction aliquots at the indicated times from assays with PafA, *wild-type* G-PupE, and α-^{32}P-ATP. Reaction aliquots from identical assays with G-PupE(L39A,L40A), and G-PupE(39-64) were spotted on polyethyleneimine-modified cellulose TLC plates and resolved in 1 M formic acid, 0.5 M LiCl prior to phosphorimaging. Boxes indicate the regions selected for densitometric analysis with ImageJ software.[2] (**b**) Kinetics of phosphorylation of wild-type G-PupE (green spheres), G-PupE(39-64) (purple spheres), and G-PupE(L39A/L40A) (red spheres), all measurements were in triplicate. Background hydrolysis of α-^{32}P-ATP by PafA-His$_6$ in the absence of Pup is shown in black spheres. The data obtained was fit to the equation $[ADP]_t = [ADP]_0 + [ADP]_f(1\text{-}e^{-kt})$ using KaleidaGraph software. $[ADP]_0$ represents the initial amount of ADP in the reaction, $[ADP]_f$ is the final amount of ADP at the end of the reaction and **k** is the observed rate constant for the reaction. For G-PupE, k_{obs} = 1.37 ± 0.14 min^{-1}, G-PupE(39-64) k_{obs} = 1.71 ± 0.15 min^{-1}, G-PupE(L39A,L40A) k_{obs} ~0.006 min^{-1}.

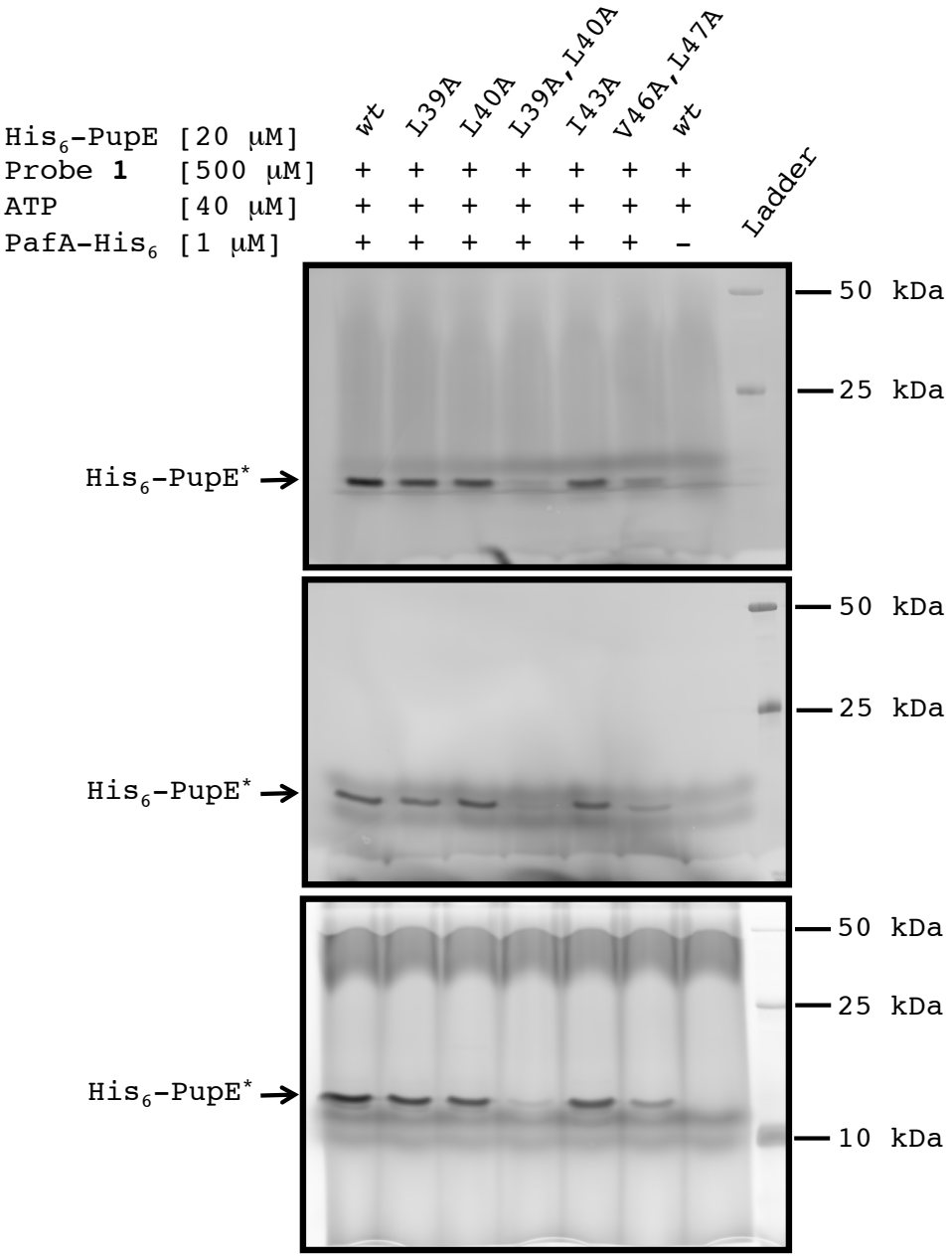


Figure S16. Importance of the C-terminal hydrophobic residues in His_6-PupE for PafA-His_6 activity. In-gel fluorescence from 15% SDS-PAGE gels of three independent assays with wild-type His_6-PupE (20 μM), or the indicated mutant, in the presence of ATP (40 μM), PafA-His_6 (1 μM), and probe **1** (500 μM) for 10 min at 37 °C. His_6-PupE* indicates the position of the fluorescent wild-type or mutant His_6-PupE-**1** conjugate.

References

(1) Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual, Third Edition*; Third ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, **2001**.

(2) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat. Methods* **2012**, *9*, 671.

(3) Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; Grosdidier, A.; Hernandez, C.; Ioannidis, V.; Kuznetsov, D.; Liechti, R.; Moretti, S.; Mostaguir, K.; Redaschi, N.; Rossier, G.; Xenarios, I.; Stockinger, H. *Nucleic Acids Res.* **2012**, *40*, W597.

(4) Guth, E.; Thommen, M.; Weber-Ban, E. *J. Biol. Chem.* **2011**, *286*, 4412.

(5) Ozcelik, D.; Barandun, J.; Schmitz, N.; Sutter, M.; Guth, E.; Damberger, F. F.; Allain, F. H.; Ban, N.; Weber-Ban, E. *Nat. Commun.* **2012**, *3*, 1014.