Supplementary Information for:

Improved pyrrolysine biosynthesis through continuous directed evolution of the complete pathway

Joanne M. L. Ho[1,*], Corwin A. Miller[1,*], Jacob R. Mattia[1], and Matthew R. Bennett[1, 2,†]

[1] Department of Biosciences, Rice University MS-140, 6100 Main St., Houston, TX 77005, USA

[2] Department of Bioengineering, Rice University MS-140, 6100 Main St. Houston, TX 77005, USA

* These authors contributed equally.

† Correspondence: matthew.bennett@rice.edu

Table of Contents

Materials & Methods

Alt-PANCE of *pylBCD* pathway. One round of Alt-PANCE was performed each day; each round consisted of two passages—one mutagenic (16h) and one selective (8h). For mutagenic passages, *E. coli* strain S1059 cells carrying plasmid MP6 were used as the host; 10 mM L-arabinose was added to the culture to induce MP6 immediately following the addition of phage. Activity of MP6 was tested by separate serial spotting of mutagenic host cells onto glucose and arabinose containing agar (see Figure S2A). For selective passages, *E. coli* strain S1030 cells carrying an accessory plasmid were used as the host (for specific accessory plasmids used at each transfer, see Table S1). Test outgrowths were performed at higher stringency conditions after each selective passage. Nε-Boc-L-Lysine (BocK) (Chem-Impex, Cat. No. 00363) was added during initial selective passages to tune selection strength (see Table S1). Ramps in stringency condition of S1030/AP are as shown in Figure S2C. Three lineages (Lineages A–C; Figures 2 and S3) were independently evolved. *E. coli* host cell and bacteriophage culture conditions for each phage passage were performed as previously described by Suzuki *et al*[1]. Phage produced following each passage were quantified by phage titer.

Molecular cloning. The *M. acetivorans* Pyl biosynthetic pathway (NCBI locus tags MA_RS00810, MA_RS00815, and MA_RS00820) was codon-optimized for expression in *E. coli*, and a ribosome binding site (SD4) was added upstream of *pylC* as the natural ribosome binding site lies within the coding region of *pylB* (see Supplementary Data for DNA sequences). This sequence was synthesized as an IDT gBlock, and the synthetic DNA was subsequently cloned into pCDF. All other DNA fragments were prepared by PCR. Selection phage (SP) vectors were cloned by inserting DNA into M13 bacteriophage, replacing *gIII* as previously described[2]. Combinatorial mutants were prepared via Gibson assembly and QuikChange mutagenesis; all other vectors were cloned using Gibson assembly. We deposited a set of plasmids, along with maps and sequences, in Addgene (Table S6). S1059, S1030, selection phage scaffold, and mutagenesis plasmids were generous gifts from the David Liu lab at Harvard University. The aaRS variants were generous gifts from the Dieter Söll lab at Yale University.

Fluorescence-based *pylBCD* activity assay. We performed a fluorescence-based readthrough assay as previously reported[3]. We prepared BL21 (DE3) and C321.ΔA.exp strains containing a reporter vector encoding constitutively expressed PylRS, $tRNA^{Pyl}$, and AraC, as well as sfGFP.3TAG (N39O, N135O, and Y151O) under the control of the Para promoter. This strain was transformed with a producer vector encoding either the wild-type *pylBCD* variant or an engineered variant under the control of the PT7/Lac promoter. Cells were grown to OD600 0.2—0.5, expression of the biosynthetic cassette was induced with 10 μM IPTG for 2h, and expression of the reporter sfGFP.3TAG protein was induced with 1.3 mM L-arabinose for 3h–18h. Fluorescence intensity (excitation 488 nm, emission 509 nm) was normalized by optical density (A_{600}) and the best producer was identified. Experiments were performed on biological triplicates and background correction was performed by subtracting the Fl/OD value of the strain transformed with an empty producer vector. Concentrations of IPTG and L-arabinose used in the above protocol were found to result in the highest amount of fluorescent signal after empirically testing additional concentrations.

Luminescence-based *pylBCD* activity assay. We performed a luminescence-based readthrough assay as previously reported[4]. We prepared strain C321.ΔA.exp cells containing a reporter vector encoding constitutively expressed *chPylS* and pylT, as well as *luxCDABE* wherein *luxB* contained an amber stop codon after the initiator methionine. This strain was transformed with a producer vector encoding a pylBCD natural or combinatorial variant under the control of the Para promoter. Cells were grown to OD600 0.2–0.5, and expression of the biosynthetic cassette was induced with 1.3mM L-arabinose for 3h–18h. Luminescence intensity was normalized by A600. Experiments were

performed on biological triplicates and background correction was performed by subtracting the LUMI/OD value of the strain transformed with an empty producer vector.

Purification and LC-MS/MS of the Pyl-containing sfGFP. Purification of the Pyl-containing sfGFP variants was enabled via a C-terminal His-tag. Reporter protein was expressed for 24h at 25°C in BL21 (DE3) encoding the reporter vector and best producer vector. Protein was purified using a Ni-NTA column, run on an SDS-PAGE gel, and the excised 28 kDa band was in-gel digested with chymotrypsin and analyzed via LC-MS/MS (Yale Proteomics Center) as previously described[5].

Light microscopy of *E. coli* cells. *E. coli* containing pCDF plasmids encoding the SUMO-tagged or untagged *pylBCD* operon variants were grown in LB to OD 0.5 and induced with 10 μM IPTG. After 3h, bright field imaging was performed with the Nikon Ti-E inverted microscope using the Nikon Plan Apo 100x Oil Objective, 1.40 NA and Photometrics CoolSNAP HQ2 CCD camera. Images were processed using the Nikon Elements 4.51 software.

Purification of PylB variants. Protein purification was performed using a procedure adapted from the methods described by Quitterer *et al*[6]. *E. coli* strain BL21 (DE3) cells containing either plasmids JH767, JH768v2, or JH769v2 – which encode different *pylB* variants containing an N-terminal His-tag followed by a TEV protease cleavage site (see Table S6)- were initially grown overnight in 20 mL of 2xYT media containing kanamycin at 37ºC. Cultures were subsequently diluted into 800 mL of 2xYT media and grown at 25ºC until OD600 reached 0.4-0.6. Cells were then induced by adding 500 μM IPTG, and were grown overnight at 25ºC. Cultures were next pelleted, and resuspended in 25-40 mL of PylB Lysis Buffer, containing 100 mM Tris hydrochloride (pH 8.0), 500 mM NaCl, 5 mM sodium dithionite, and 15 mM imidazole hydrochloride. Cells were lysed by sonication, and then centrifuged at 17,000 x g for 1 hour at 4ºC. Gravity columns containing Nickel-NTA resin were equilibrated with PylB Lysis Buffer, and clarified protein supernatants were applied to separate columns. Samples were eluted with PylB Elution Buffer, containing 100 mM Tris hydrochloride (pH 8.0), 500 mM NaCl, 5 mM sodium dithionite, and 500 mM imidazole hydrochloride. SDS-PAGE was used to confirm purification of His-tag containing PylB from each eluate and to determine protein purity (see Figure S12).

For each sample, buffer exchange into PylB Lysis buffer was subsequently performed using Amicon Ultra-15 centrifugal filtration units (3,000 Da nominal molecular weight limit). Protein concentrations were determined by measuring absorbance at 280 nm, and samples were diluted to 0.4 mg/mL. Samples were next digested with TEV protease (provided by New England Biolabs, cat# P8112S) by incubating at 30ºC for 2 hours. Undigested PylB and residual protease were subsequently removed from each sample by reverse purification using Nickel-NTA resin. SDS-PAGE was subsequently used to determine the purity of the cleaved PylB samples. To prepare samples under low salt conditions, aliquots of each sample were taken following purification samples and were dialyze against Low Salt Lysis Buffer, containing 100 mM Tris hydrochloride (pH 8.0), 20 mM NaCl, 5 mM sodium dithionite, and 15 mM imidazole hydrochloride.

Quantification of PylB protein yields. For quantitative yield assessment of PylB variants, purification of uncleaved His-tagged variants was performed as described above (without performing TEV protease digestion). Each variant was purified in biological triplicate, beginning from separate colonies. Purifications were performed using a 100 mL expression volume of 2xYT. Total protein concentrations were determined by measuring absorbance at 280 nm, and PylB purity was determined using SDS-PAGE.

PylB solubility assays. Solubiilty assays were performed using a procedure adapted from Kramer *et al*[7], using cleaved PylB samples (without His tags) purified as described above. Each protein sample was initially concentrated to a uniform concentration of 2.6 mg/mL. Separate aliquots of each protein

sample were next mixed with solutions of either ammonium sulfate or PEG-8000 in a one-to-one ratio. For ammonium sulfate, solutions were prepared at concentrations ranging from 0.5M to 1.5M; for PEG-8000, solutions were prepared at concentrations ranging from 5% to 15% (w/v). After thorough mixing, the solutions were allowed to equilibrate at room temperature for 10 minutes. Solutions were subsequently centrifuged at 17,000 x g for 10 minutes, and the resulting supernatant was transferred into fresh tube. Concentration of each supernatant was determined by measuring absorbance at 280 nm. This protein concentration corresponds to the solubility of each protein at the corresponding concentration of precipitant.

Analysis of PylB solubility. Protein solubility data was analyzed as previously described[7], using solubility measured at different precipitant concentrations as described above. The following formula was used to describe the relationship between precipitant concentration and protein solubility:

$$\text{Log } S = \text{Log } S_0 - \beta \text{ [Precipitant]}$$

In this expression, S is the measured solubility at a given concentration of precipitant, S_0 is the solubility in the absence of precipitant, and β is the solubility constant, denoting the dependence of solubility on precipitant concentration for a given protein. Log protein concentration measurements were plotted against precipitant concentrations, and β values as well as uncertainties were determined for each PylB variant using linear regression function in Microsoft Excel. Points lying outside the linear region of the plot were excluded from the regression analysis. Here we define aggregation propensity as negative β values ($-\beta$).

Biosynthetic production and extraction of Pyl. Procedures for expression and extraction of Pyl were adapted from previously described protocols[8, 9]. *E. coli* strain BL21(DE3) cells were transformed with either plasmid JH572 (encoding the empty pBAD vector), JH577 (encoding *pylBCD* variant 3f2), or plasmid JH579 (encoding variant JM10.1). Cells were inoculated into 100-500 mL of LB broth supplemented with 1 mM L-lysine, and grown at 37ºC. Upon reaching OD600 = 0.3, cultures were induced with 1.3 mM L-arabinose, and subsequently grown at 37ºC for periods ranging from 4-18 hours. Cells were subsequently pelleted and washed with distilled water. Cells were subsequently extracted with either 2 mL of methanol per gram of wet cell weight at room temperature, or with 2 mL of 66% aqueous methanol at 70ºC. Extractions were incubated for 30 minutes at the tested temperatures. Extracts were then centrifuged at 13,000 x g for 5 minutes, and the supernatant was retained. For a subset of samples an additional filtration step was then performed, wherein samples were passed through an Amicon Ultra-15 Centrifugal Filter unit (3,000 NMWL), retaining the flow-through; samples were diluted with distilled water to 60% methanol prior to filtration. For another subset of samples an ethyl acetate extraction was then performed, wherein ethyl acetate was added to each sample until reaching a total concentration of 80%; samples were then centrifuged at 13,000 x g for 5 minutes, and the supernatant was retained. Extracts were dried during centrifugation under vacuum prior to analysis.

Differential scanning fluorimetry (DSF) assays of PylB variants. Procedures for protein DSF assays (also known as the Thermofluor assay) were adapted from previously described protocols[10]. Briefly, 50 µL reaction mixtures were prepared in white 96-well PCR plates (Bio-Rad Laboratories, cat # HSP9655), with each mixture containing purified PylB samples diluted to 5 µM as well as SYPRO orange dye (Sigma-Aldrich cat. # S5692) at 20x working concentration. Sealed plates were run on a Bio-Rad CFX Connect Real-Time PCR machine, heating from 25ºC to 95ºC. Samples were heated at a rate of 0.5ºC per 30 seconds, and fluorescent readings were taken at every 30 second interval. All protein samples were run in triplicate, including separate low salt and high salt negative controls which contained only buffer and dye (no protein). To plot melting curves, averages of each protein

sample were calculated and the background signal from the negative controls was subtracted. Melting temperature (Tm) values were identified by determining the temperature at which the first derivative of the fluorescent signal reached its lowest minimum value.

Thin-layer chromatography (TLC) analysis of Pyl extracts. Procedures used were adapted from methods previously described for TLC separation of amino acids (protocols are publicly available from Millipore-Sigma). Dried extract samples were resuspended in methanol and spotted onto Cellulose 400 TLC plates (100 µM, 20x20 cm). Quantity of extracts spotted were empirically optimized by trial and error. Plates were run twice in the same dimension using a 1-butanol/acetone/glacial acetic acid/water (35%/35%/7%/23%) mobile phase. Plates were dried for 1 hour after running. To visualize amino acids, plates were uniformly sprayed with ninhydrin (0.125% w/v, dissolved in acetone) and developed by incubating at 70ºC for 5-10 minutes. Developed samples were visualized under white light.

Cellular fluorescence analysis of Pyl extracts. *E. coli* strain BL21 (DE3) and strain C321.ΔA.exp cells were separately transformed with plasmid JH528, encoding arabinose-inducible sfGFP with a single amber codon as well as constitutively expressed *pylS* and *pylT* (see Table S6). Separately, dried extract samples were resuspended in distilled water. Cells cultures were grown to OD600 0.2—0.5 in LB media, and subsequently mixed with an equal volume of resuspended extract. Cultures were simultaneously induced with 1.3 mM L-arabinose, and grown for 3-18h. Fluorescence measurements were performed during cell growth, as described in the ‘Fluorescence-based *pylBCD* activity assay’ section above.

Supplementary Data

>WT *M. acetivorans pylB* sequence

ttgatcaaaaaaatggcaaccgaggaccttgacaggttcggagagaaaataattgaaggctttcaattgtctgatgatgacctgagggctcttct
ctccctcgaatccgaagaagagctggaaaagctatactatgtagcccgaaaggtcagaaaccactattttggtaacagggtgtttcttaactgct
ttatctatttctctacctactgtaagaaccagtgcgctttttgctactataactgcaaaaacgaaatcaaccgctaccggctgagcagagaggaa
ataaaagagacctgcaaagccttaaaaggagcaggttttcacatgatcgacctgaccatgggagaagacccttattactatgaggagccgg
accgcttcgttgagcttgtccagatggtaaagaacgaactcgggctccccattatgatttccccgggagtgatggacaacgccacccttctcaa
agccagggaaaaggggcaaatttcttcgccctgtaccaggaaacctatgaccaggaactttacgaaaaactccgggtaggccagtccttta
aagggagatataacacccgccgctttgcaaaggaacaggggtactgcatagaagacgggattctcacgggtgttgggaacgatattgagtc
aacaatcaagtccctgaggggaatgaaatcaaacgatcctgatatggtccgggttatgacgttcctgcctcaggaaggtaccccactcgaga
ggttcaaagagaattcaaacctttcggaactgaaaataattgcaatcctcaggctcatattcccgaaatgcctgatcccggcatcccttgacctc
gaaggcatcgacggcatggttcaccgcttaaatgcaggcgcaaacattgtcacttccatcctcccctcggattcaaagttagaaggcgttgcca
actatgaccgcgaccttgaagagcgggatcgggatataaaaagcgtaatcaaaaggcttgaaggcatggggatggaacctgctccgcagg
cagaattcgaaactgtcctgaggtgctag

>WT *M. acetivorans pylC* sequence (including upstream intergenic region)

cttgaaaaccatatgccttataggcgggaaactccagggcttcgaagcggcttacctgtctaagaaagccgacatgaaagtagttgtgataga
taaaaaccctcaggcgcttataaggaactatgccgacgagttccactgttttgatgtaataaaagagcctgaaaaacttctcgagatctcaaaa
aacgttgacgccatactccctgtaaatgaaaaccttgaatgtataaagttcctgagttccgtaaaagaggaattctcctgccctgtcctcttcgattt
cgaggcttaccggatcagcagggacaagaaaaaatcaaaagaatacttcaggtccataggaatcccgaccccgcaggacaaaccaggc
aggccgccttattttgtaaaacccccctgtgaaagcagcagtgtgggagcaaggataatttatgacgaagaggagcttggagagcttgagcc
gggaatgctggtcgaagaatatgtggagggggaagtgatttcccttgaagtcataggagacgggacccattttgcagtggtaaaggaaaccc
ttatacacatcgacagaacctacgactgccacatggtaacccctcttccttctgacctttccttcagggagatatcctattcccttgcagcaaacct
gcccttaaaagggatcatggacgtggaagcaatttccggcccccagggattgaaagtaatcgaaatcgattcccgcttcccgagccagaccc
caactgcagtctactattcttcgggagttaacctcatagaactcctgttccgcgccttcggagaaggtgtcgaagaggtcaaaaccctcctgaa
gacaggtactgcatttacgaacacctcatgcttgcagaaaacggagccctcatcccggtgggagagcaggtgctctctatgggaaacgatta
cggcaaatattatgaagagcccggcattgagattttcctgtgcagaggagaggaccctgtgttcaccctggtcttctggggcaaagataggga
agaagccgagaataaaaagagaacaggactcttgatcctgaaagaccgcttcggagcggctgtgtaa

>WT *M. acetivorans pylD* sequence (including upstream intergenic region)

aatggcaaaaatcggataagaataaaaatttgcccaaagtaaaaatttgtttaaaacagaaaaaggagataaaaaatggcacttttaacccc
ggaagacctggaaaacattaacagacagcttcatgaagctgattctactgtacggagagttacagggcttgacattaaaggagtctgcaaag
acttctacggcacgaccccatgctgtgaaaaggtaggcattgtacctgtgacctcgggaaacgggataataggaaatttctcggaatccctgc
gtgcaattaccgaatatttcggatttgacagctttgttacgactatgcctgatgtaagcgggtactatgaggcagtaagaaacggcgcccagatt
atcctcatggcggatgacaataccttccttgcccataacctcaaaaacgggaaaatagccaataaccagccctgcacaggcataatttatgcg
gaaatagcctccagatacatgaaagccgattcaaaagacgtacttgccgttggcctgggaaaggtgggttttccgggagcagcgcacctcgt
aaataagggcttcagggtctatggctatgatgctgataaagcccttcttgagaaaaccgtttctgaccttggaattactccttttgaacccgaaaac
ccgagaagattttccattattttcgaggcaactccctgtgcggacacggttcccgaatccgtaatttcggataactgcgtaatttccacccccgga
atcccctgcgcaatttccagggagcttcaagaaaagtacggagtggaacttataatggaacctctcgggatagggacggcatccatgctgtatt
ctatcctgtaa

>Codon optimized *M. acetivorans pylB* sequence

atgattaagaagatggcgactgaagatctggatcgttttggcgaaaaaattattgaaggatttcagttatctgacgacgatttacgcgctttgctta
gcctggaatcggaagaagaattggaaaagctgtattacgtcgcacgtaaagtacgcaaccattactttggaaatcgcgtattcctgaactgtttt
atctacttcagcacatactgcaaaaaccaatgtgcgttttgctattataactgcaaaaacgagattaaccgctatcgtttgtcccgcgaagaaatc
aaagagacttgcaaggcccttaaaggtgcgggtttccatatgattgatctgactatgggggaggatccctattattacgaggaacctgatcgctt
cgtggagctggtccaaatggttaagaatgagcttggacttccaatcatgatttcgccaggagtaatggataatgctacccttttaaaagcacgcg
aaaagggagctaatttcttcgccttatatcaagaaacctatgatcaagagttgtacgagaagttgcgcgtaggtcagtcgttcaaggggcgcta
caatactcgtcgttttgctaaagaacaagggtattgcattgaagacggtatcttaacgggagtcgggaacgatatcgagtcaacgatcaaatctt
tacgcggcatgaaatcaaatgatcccgacatggttcgcgtgatgacttttcttaccccaggaaggaactccccttgagcgttttaaggaaaatagt

aacctgagcgaactgaaaatcattgcgatcttgcgtctgatctttccgaagtgcttaatccccgcgtcgttggacctggaagggatcgacgggat
ggtacaccgtcttaacgcaggtgctaatatcgtgactagcatccttcctagtgattcaaaacttgagggagttgctaactacgatcgtgacctgga
agagcgcgaccgcgacattaaaagcgtcatcaagcgccttgagggaatgggcatggaaccagcgccacaggcggaattcgagactgtac
tgcgctgctaa

>Codon optimized *M. acetivorans pylC* sequence (including upstream intergenic region)
taaggaggaaaaatgaaaaccatctgcctgattggtggtaaattgcagggttttgaggccgcgtatctgtctaaaaaagccgatatgaaagta
gtggttatcgacaagaacccccaagctcttatccgcaattacgcagacgagtttcattgttttgatgttatcaaggaaccagagaaacttcttgaa
atcagcaaaaatgtagacgctatcttgcccgtcaacgaaaacttagaatgtattaaattcttaagtagtgtaaaagaggaattttcttgcccagta
ctgttcgattttgaggcttaccgcatctcgcgcgataagaagaaatcgaaagagtatttccgttccatcggcattccgacaccgcaagacaaac
ctggtcgtcctccatacttcgttaagccaccgtgcgaaagctcgtcggttggtgctcgcattatctacgatgaagaagaactgggagaattggag
cccggcatgctggtagaagagtatgtggaaggtgaggttatttcacttgaagtaatcggggacggtacccactttgcagttgttaaagagacact
tattcacatcgaccgcacctatgactgccatatggtaactcccctgccatctgatttgtctttccgtgagatcagctattctttagcggccaacctgcc
tttgaaggggatcatggacgttgaagcaatttcgggcccgcaaggtttgaaggtcattgaaattgattcacgtttcccctcccagacccccacag
ccgtgtattattcgtcgggagtgaatttaattgagctgttgtttcgtgcctttggagaaggggtggaggaagttaaaacccttccagaggaccgcta
ctgtatctacgagcacctgatgcttgcagaaaatggggccctgatcccagtaggtgagcaggttctgtcgatggggaatgactatgggaaata
ctatgaagaaccaggaattgaaatctttttatgtcgcggagaggaccccgtgtttacgcttgtcttttggggaaaggaccgtgaggaagccgaa
aacaaaaagcgcactggtttgctgattttgaaggaccgcttcggggcagccgtttaa

>Codon optimized *M. acetivorans pylD* sequence (including upstream intergenic region)
aatggcaaaaatcggataagaataaaaatttgcccaaagtaaaatttgtttaaaacagaaaaaggagataaaaaatggctttgttgactcccg
aggacctggagaacattaaccgtcaacttcacgaagccgattctacggtacgtcgtgtgactgggttagatatcaagggagtttgtaaagatttc
tatgggactactccatgctgtgagaaggtcggtatcgtcccggtaactagtggcaacggcatcattggcaacttttctgaaagtcttcgtgcgatc
actgagtactttggatttgactcattcgtgacgaccatgccagacgttagtggctactatgaggcggtccgtaatggagcacaaattattcttatgg
cagatgataatacgtttcttgcacacaatttaaagaacggcaagatcgctaataatcaaccgtgtaccggaattatctacgcggagatcgcctc
gcgctatatgaaagccgattccaaagatgtgcttgccgtcgggcttggaaaagtggggtttccaggcgctgcacatttggttaataaggggtttc
gtgtgtatgggtacgacgcggataaggcgttgttagaaaagactgtatcagaccttgggatcacacctttcgagcctgagaatccccgtcgcttc
tcaatcatttttgaggcgaccccgtgtgctgatacagtaccagagtctgtcatcagcgacaattgtgtcatctctacgcccggtatcccttgcgcga
tcagccgtgaacttcaagaaaaatacggggtggagttgattatggaacctctgggtattggaactgcttctatgctgtatagcattttgtaa

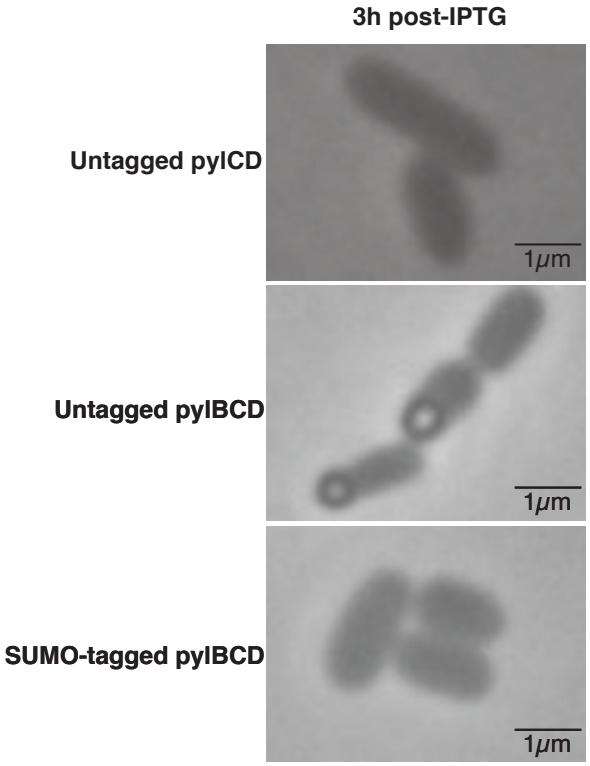


Figure S1. Microscopy of *E. coli* cells expressing *pylBCD*. Cells expressing WT *M. acetivorans* genes *pylC* and *pylD* from plasmid JH60 are shown in upper panel; cells expressing the WT *M. acetivorans pylBCD* pathway from plasmid JH123 are shown in the middle panel; cells expressing *pylBCD* following the addition of a SUMO tag to the N-terminus of *pylB* (SUMO-*pylBCD*) from plasmid JH126 are shown in lower panel. While inclusion bodies are visible in cells expressing the WT *pylBCD* pathway, they are not observed in the absence of *pylB* or in cells expressing SUMO-*pylBCD*.

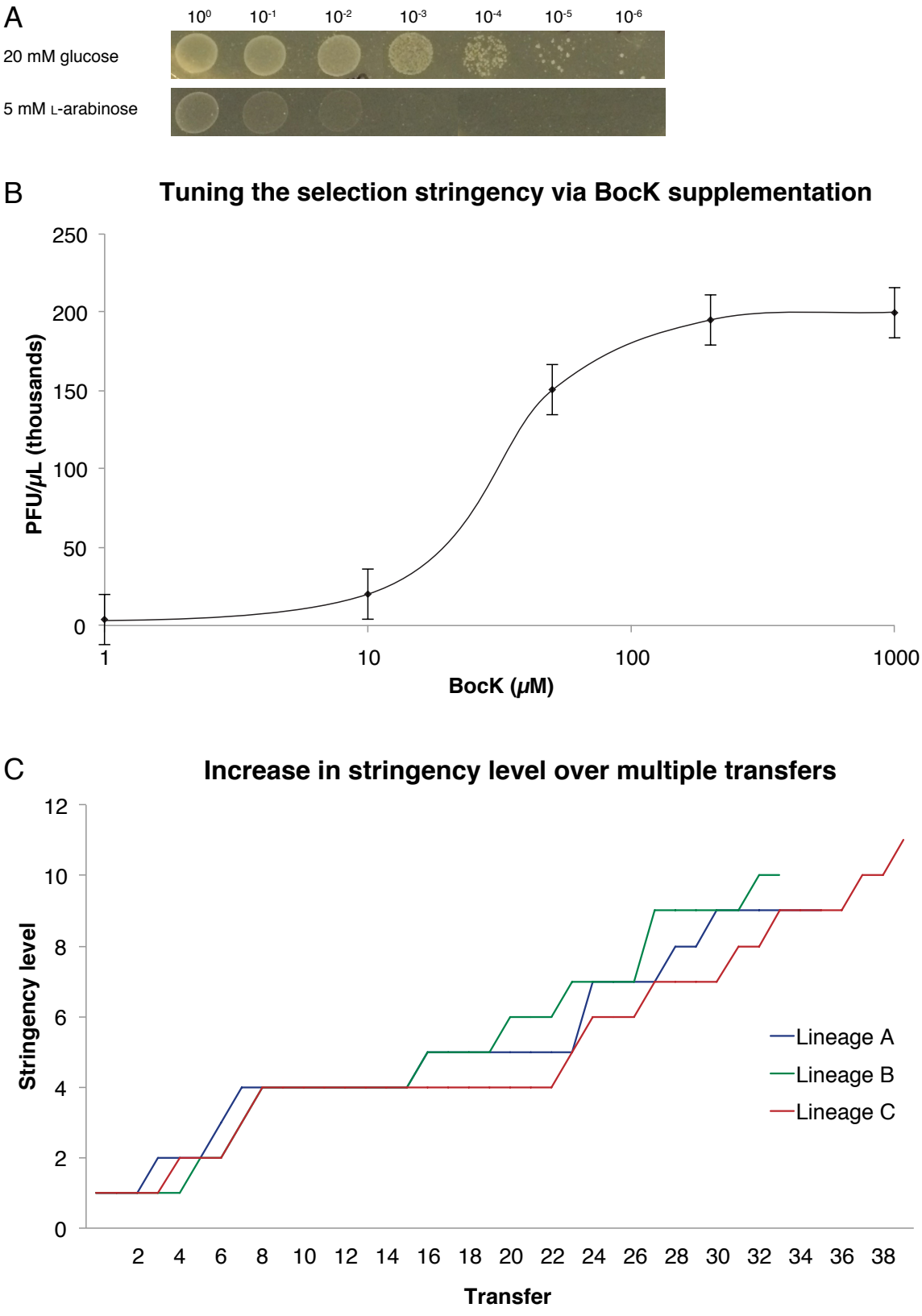


Figure S2. Alt-PANCE of *pylBCD* evolution details. (A) Glucose-repression and arabinose-activation of mutagenesis in PANCE host cells. (B) BocK supplementation curve—Alt-PANCE was initiated at the BocK concentration that provided 80% maximum PFU/μL. (C) Stringency ramping curve for S1030/AP, corresponding to increasing stringency levels 1–11 over the course of 34–40 rounds of Alt-PANCE.

Figure S3. Evolutionary timeline. Mutations shown were identified at different time points during Alt-PANCE of *pylBCD*. Mutations were identified by sequencing isolated SP.BCD plaques during evolution.

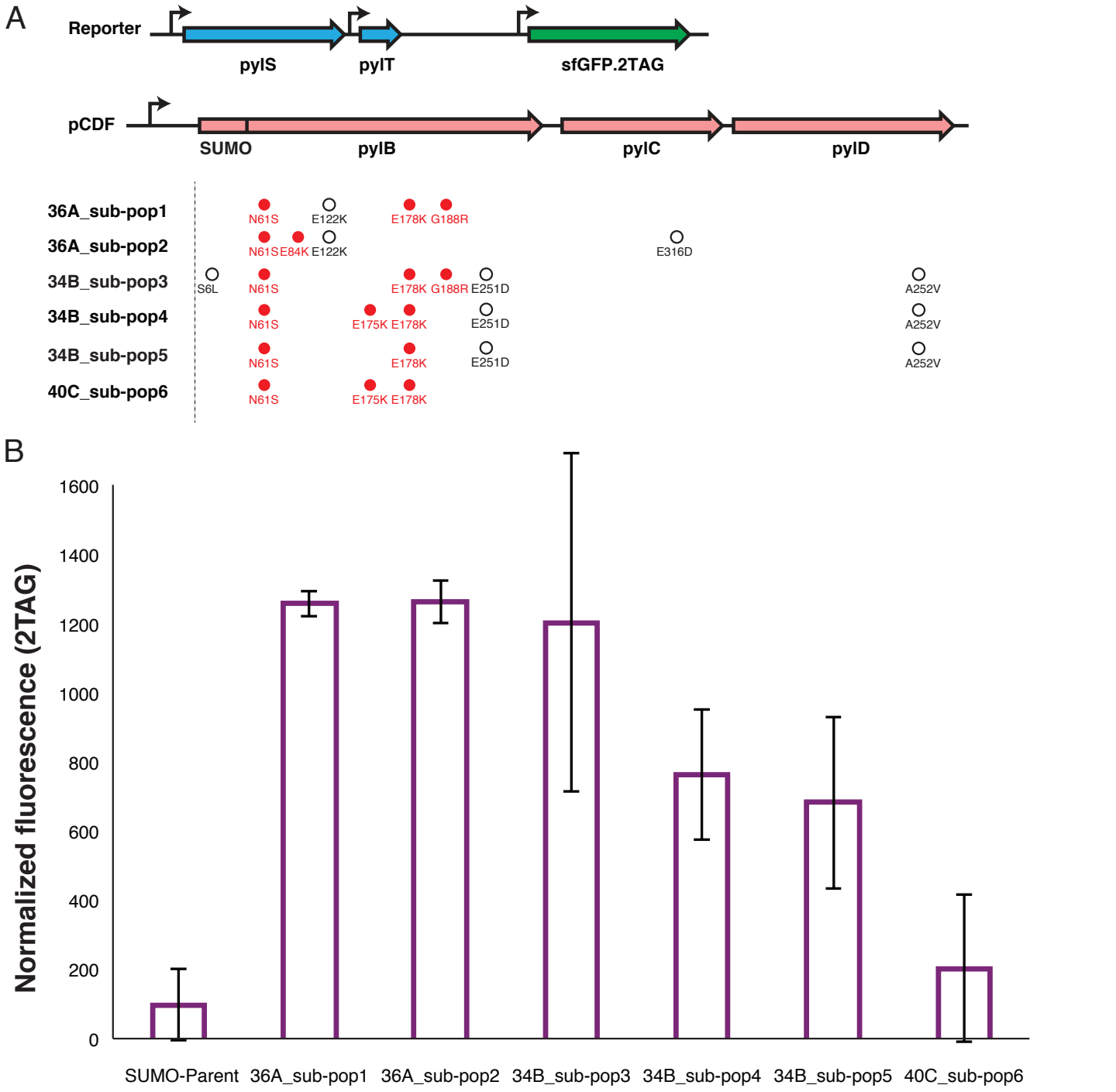


Figure S4. Fluorescence-based plate reader assay with six evolved variants. (A) Operon maps are shown for the reporter plasmid 2VecJH.Cam.sfGFP.2TAG and the pCDF expression vector containing *pylBCD* variants (see Methods). Mutations within each variant tested are shown to the right of each variant name. Each mutation is shown below the corresponding gene in which it is found (e.g., *pylB*, *pylC*, or *pylD*). (B) Normalized fluorescence following induction in *E. coli* strain BL21 (DE3) cells containing the plasmids described above (see Methods). Fluoresence mediated by the ancestral variant (SUMO-Parent) is shown on the far left, while fluoresence mediated by various evolved *pylBCD* variants are shown to the right. Samples were tested in biological triplicate; data shown represents mean values ± s.d.

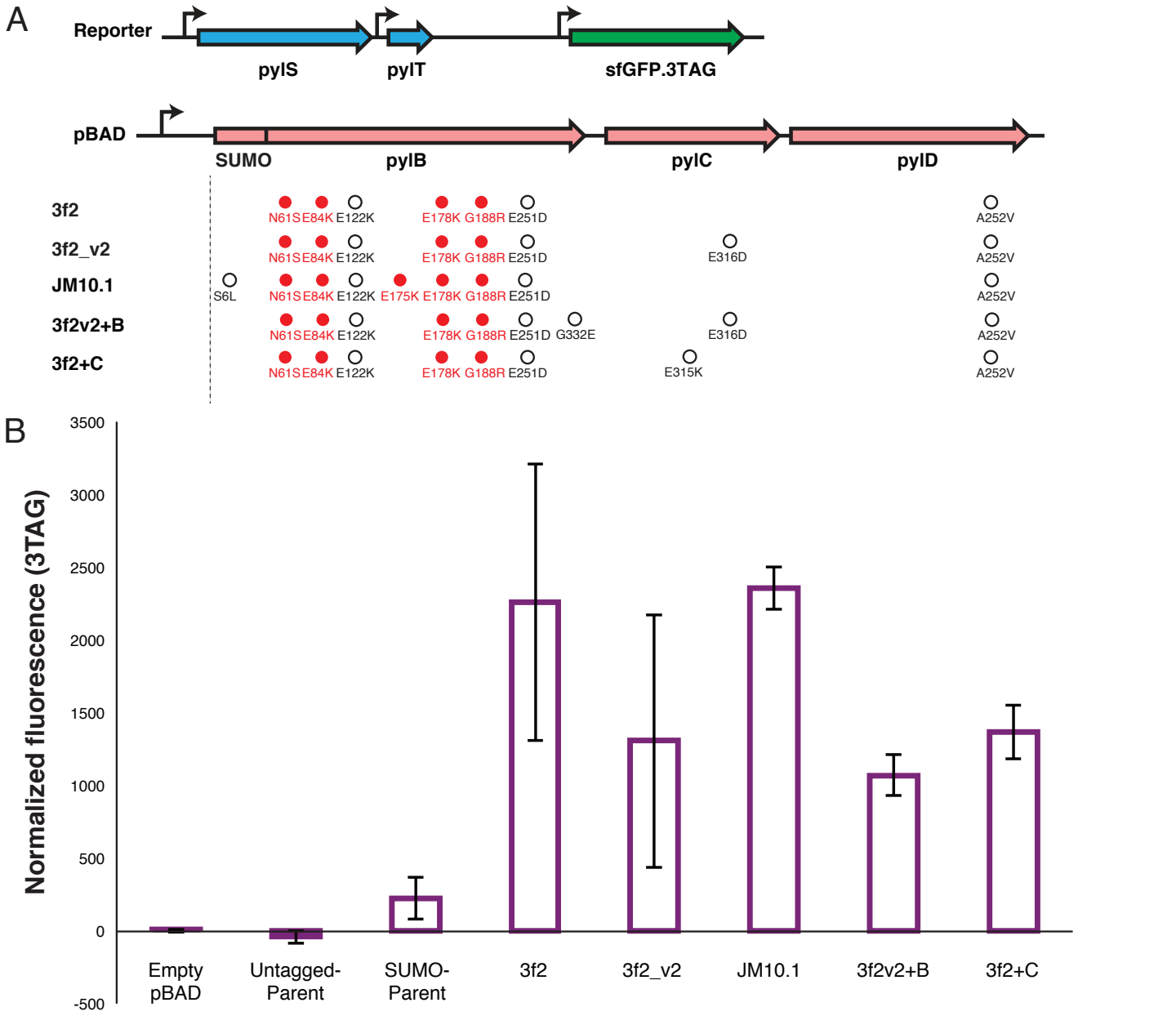


Figure S5. Fluorescence-based plate reader assay with five combinatorial variants. (A) Operon maps are shown for the reporter plasmid 2VecJH.Cam.sfGFP.3TAG and the pBAD expression vector containing *pylBCD* variants (see Methods). Mutations within each variant tested are shown to the right of each variant name. Each mutation is shown below the corresponding gene in which it is found (e.g., *pylB*, *pylC*, or *pylD*). (B) Normalized fluorescence following induction in *E. coli* strain C321.ΔA.exp cells containing the plasmids described above (see Methods). Fluoresence mediated by the empty pBAD vector, the WT *pylBCD* pathway (termed 'untagged parent'), and the SUMO-tagged ancestral variant (SUMO-Parent) are shown on the left, while fluoresence mediated by various evolved *pylBCD* variants are shown to the left. Samples were tested in biological triplicate; data shown represents mean values ± s.d.

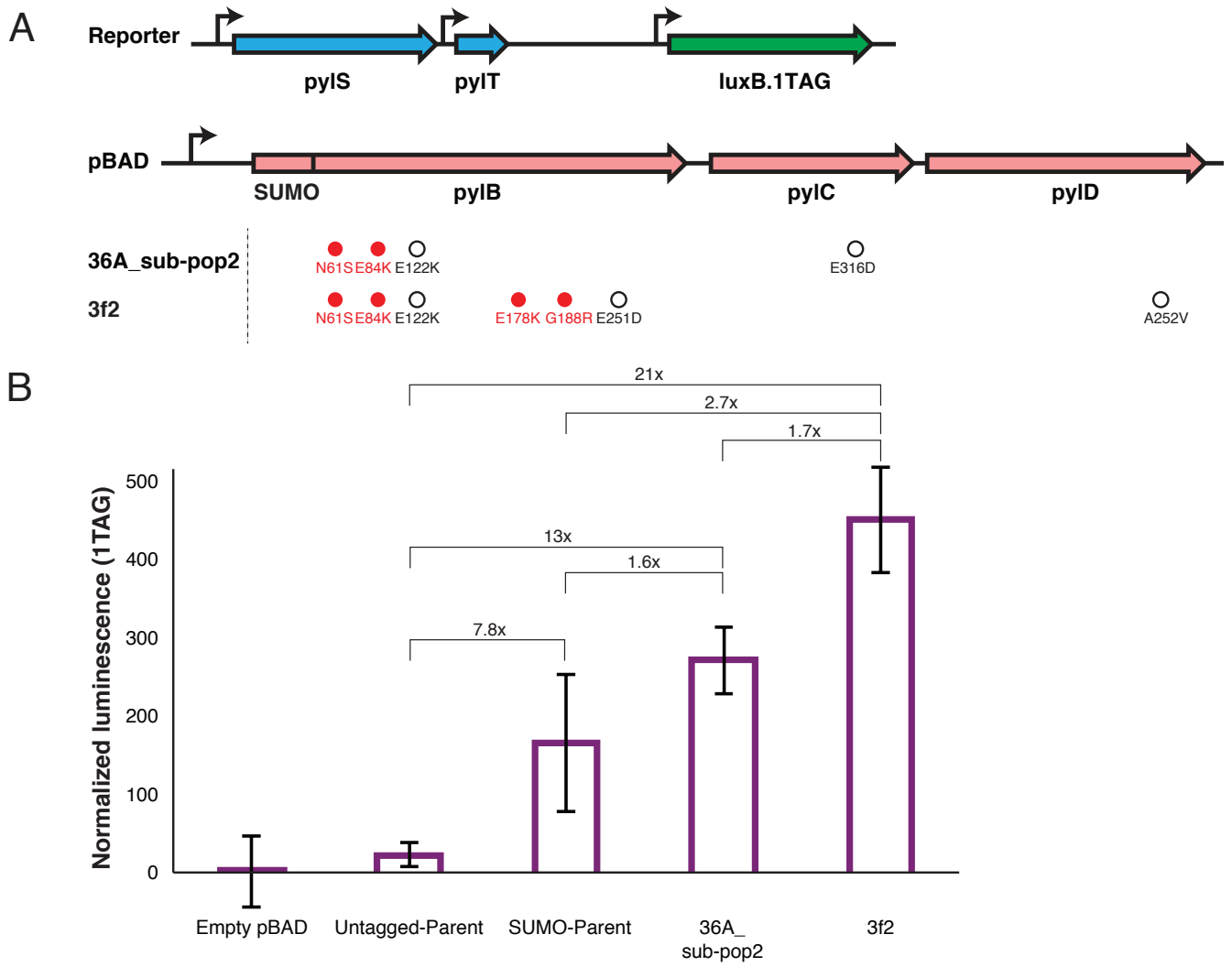


Figure S6. Luminescence-based plate reader assay with the evolved and combinatorial variants. (A) Operon maps are shown for the reporter plasmid 2VecJH.Lux.1TAG and the pBAD expression vector containing *pylBCD* variants (see Methods). Mutations within each variant tested are shown to the right of each variant name. Each mutation is shown below the corresponding gene in which it is found (e.g., *pylB*, *pylC*, or *pylD*). (B) Normalized luminescence following induction in *E. coli* strain C321.ΔA.exp cells containing the plasmids described above (see Methods). Luminescence mediated by the empty pBAD vector, the WT *pylBCD* pathway (termed 'untagged parent'), and the SUMO-tagged ancestral variant (SUMO-Parent) are shown on the left, while luminescence mediated by various evolved *pylBCD* variants are shown to the left. Luminescence results reveal the increasing biosynthetic yield of the pathway after the Sumo tag is appended to *pylB* (7.8-fold), after PANCE evolution (13-fold), and after screening different combinations of the emergent mutations (21-fold). Samples were tested in biological triplicate; data shown represents mean values ± s.d.

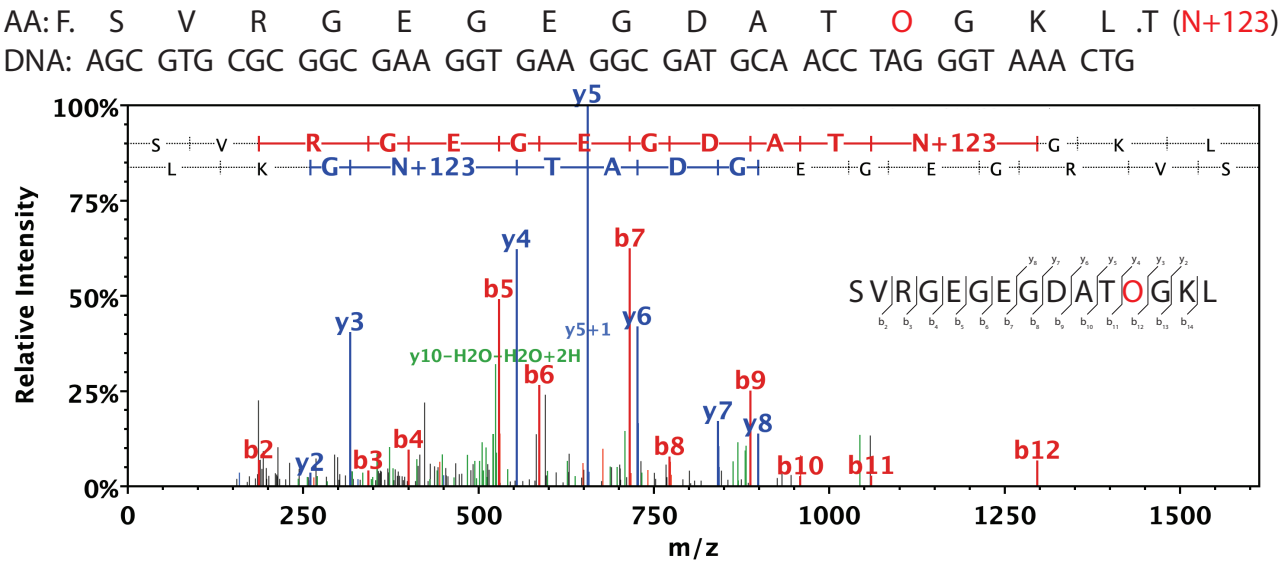


Figure S7. LC-MS/MS spectrum of Pyl-sfGFP confirming Pyl insertion at position N39O. We performed searches for molecular mass shifts indicative of Pyl incorporation (+123.19 g/mol for N to O) and potential desmethyl-pyrrolysine (dmPyl) incorporation (+109.16 g/mol for N to dmPyl). Prior work by Gaston *et al* (2011)[8] showed that cells expressing PylC and PylD (lacking PylB) grown in media supplemented with D-ornithine will produce a side product, dmPyl. LC-MS/MS did not detect any dmPyl incorporation, which was expected since dmPyl requires supplementation of media with D-ornithine and was not performed here. This observation indicates that the evolved pathway produced only Pyl, as intended and expected. Detailed fragmentation information in Table S4.

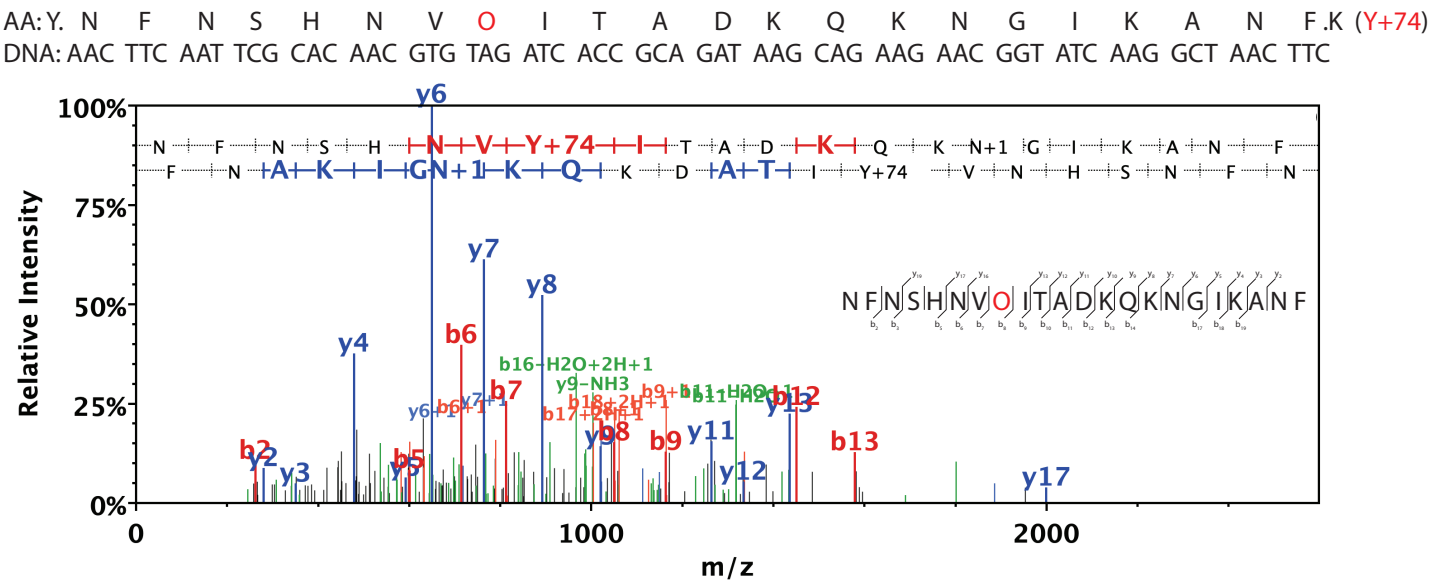


Figure S8. LC-MS/MS spectrum of Pyl-sfGFP confirming Pyl insertion at Y151O. We performed searches for molecular mass shifts indicative of Pyl incorporation (+74.12 g/mol for Y to O) and potential dmPyl incorporation (+60.09 g/mol for Y to dmPyl). LC-MS/MS did not detect any dmPyl incorporation, which was expected since dmPyl requires supplementation of media with D-ornithine and was not performed here. This observation indicates that the evolved pathway produced only Pyl, as intended and expected. Detailed fragmentation information in Table S5.

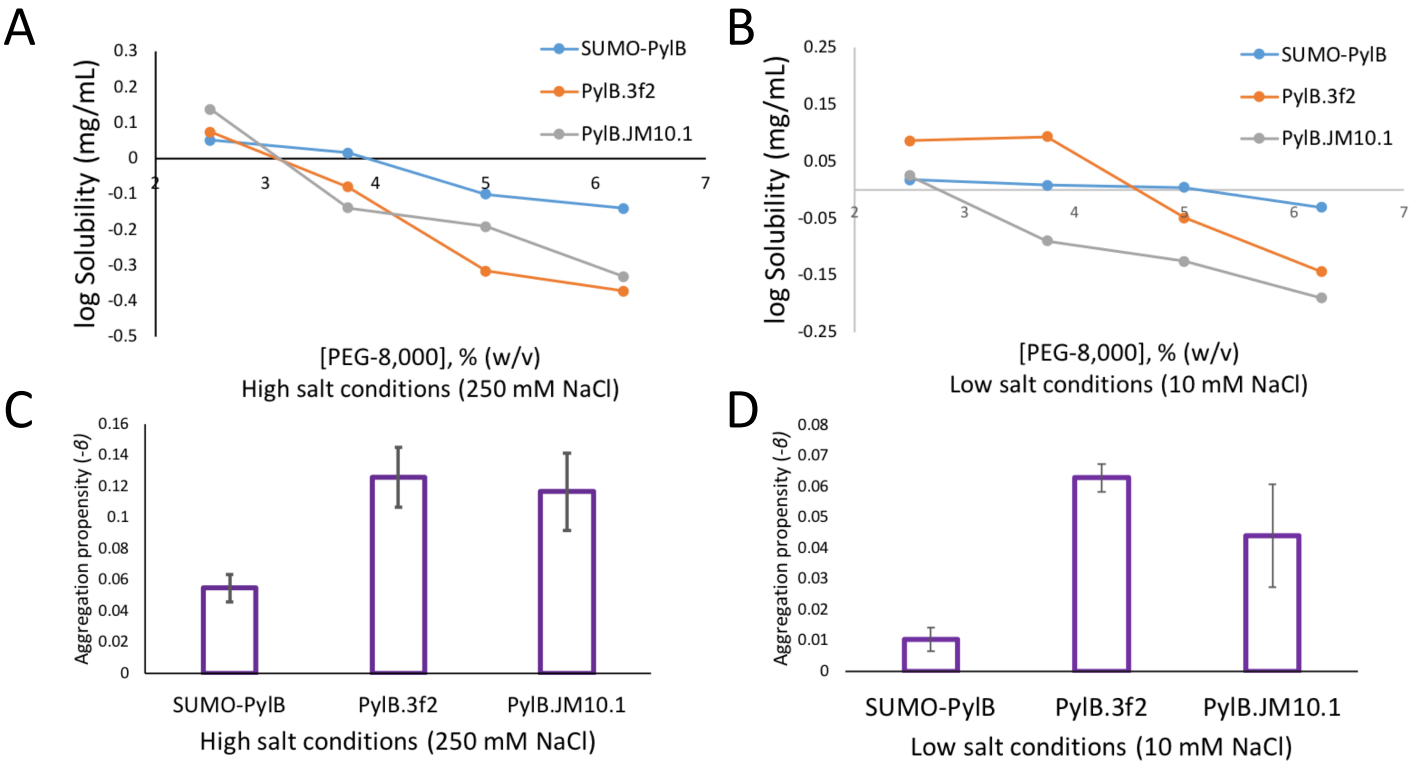


Figure S9. PylB variant solubility assays. Purified and His-tag cleaved protein samples of PylB variants were exposed to varying concentrations of the precipitant PEG-8000, and maximum protein solubility for each sample was plotted against the corresponding precipitant concentration (see Methods). Protein samples used were prepared at ≥90% purity. Solubility assays were separately conducted under (A) high salt conditions (250 mM NaCl) and (B) low salt conditions (10 mM NaCl). Aggregation propensity (-β) is shown for each PylB protein variant following both (C) high salt assays and (D) low salt assays. Higher aggregation propensity indicates reduced protein solubility. Aggregation propensity (-β) values as well as uncertainties were calculated for each PylB variant using linear regression in Microsoft Excel (see Methods). Data collected from samples tested against four concentrations of PEG-8,000 were included in the analysis.

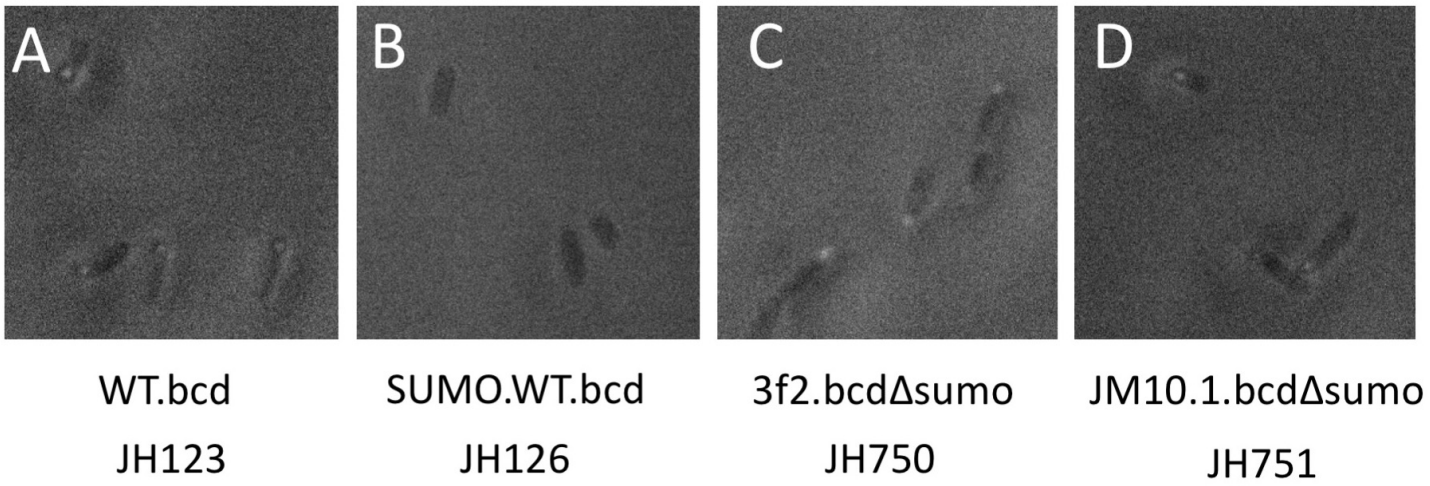
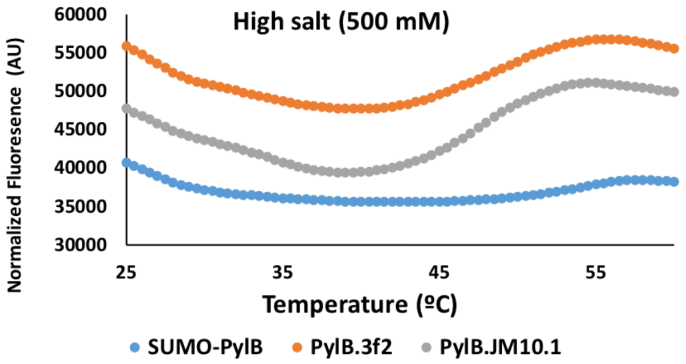


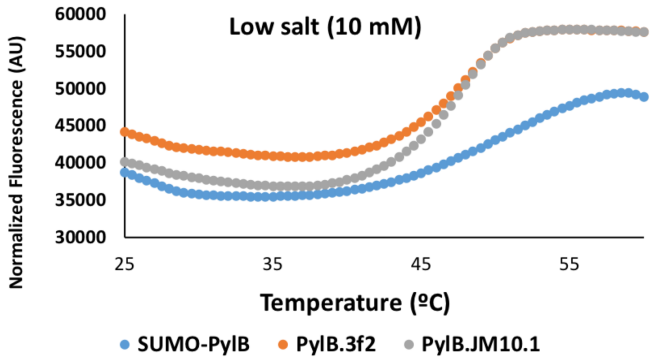
Figure S10. Microscopy of *E. coli* cells expressing mutant *pylBCD* variants. Cells expressing different variants of the *pylBCD* pathway. The corresponding plasmid name expressing each variant is shown below each image. Cells shown are expressing (A) WT *M. acetivorans pylBCD*, (B) SUMO-*pylBCD*, (C) evolved variant 3f2 with the SUMO tag removed, and (D) evolved variant JM10.1 with the SUMO tag removed. Inclusion bodies can be observed in all samples without SUMO tags (panels A, C, and D) but are not observed in panel B.

A



High salt samples	Tm (ºC)
SUMO-PylB	54.6 ± 0.5
PylB.3f2	48.0 ± 3.4
PylB.JM10.1	48.8 ± 1.8

B



Low salt samples	Tm (ºC)
SUMO-PylB	50.6 ± 0.3
PylB.3f2	48.4 ± 0.5
PylB.JM10.1	48.1 ± 0.6

Figure S11. Differential scanning fluorimetry (DSF) assays of PylB variants. Purified PylB samples were gradually heated, and a fluorescent probe was used to evaluate the extent of protein denaturation at each temperature (see Methods). Curves shown are averaged across 3 replicates; background subtraction was also performed using data collected using negative control samples that did not contain protein (see Methods). Melting temperature (Tm) values shown indicate the transition midpoint for each sample. Samples were tested under both low and high salt conditions. Across both salt conditions, the ancestral SUMO-PylB variant showed a less pronounced transition but a higher Tm value compared to the evolved variants.

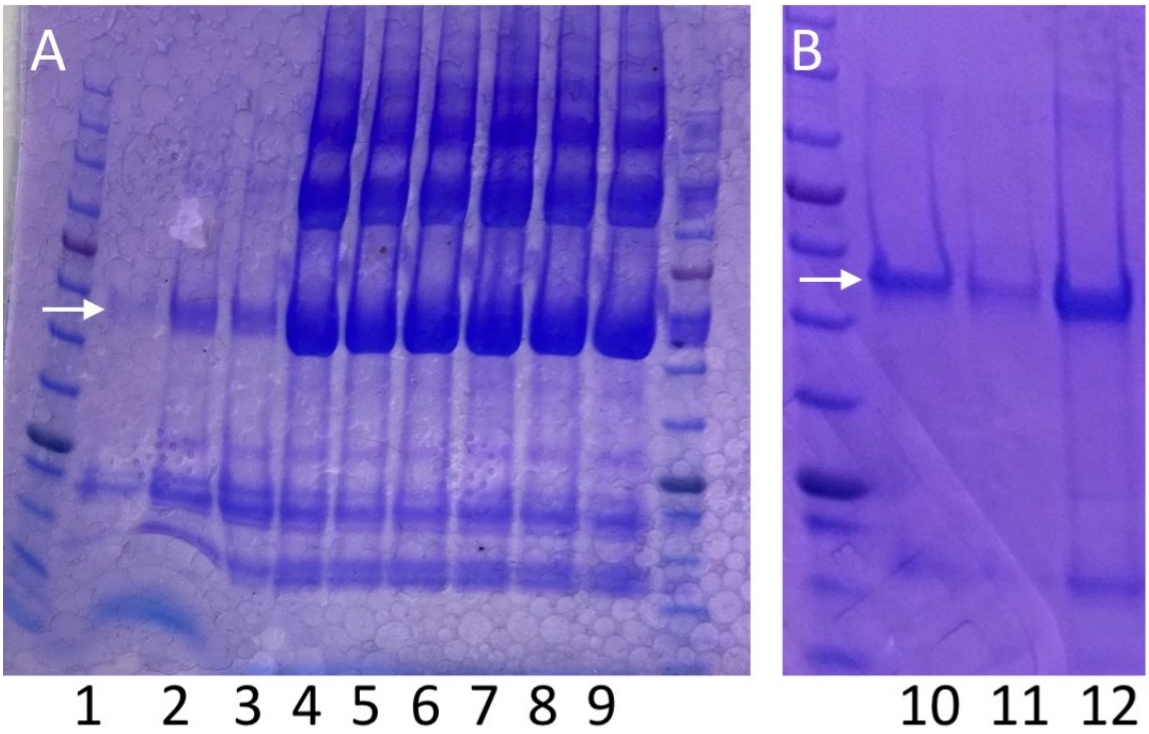


Figure S12. SDS-PAGE analysis of purified PylB samples. (A) Three PylB mutants containing N-terminal histidine tags were each purified in triplicate under similar conditions (see Methods). Labeled samples 1-3 correspond to variant SUMO-PylB; samples 4-6 correspond to variant PylB.3f2 samples; Samples 7-9 correspond to variant PylB.JM10.1 samples. PylB was estimated at ~37% purity for each sample containing histidine tags. (B) Histidine tags were cleaves from each purified protein, and samples were further purified by reverse Ni-NTA chromatography. Sample 10 corresponds to pooled variant SUMO-PylB samples; sample 11 corresponds to pooled variant PylB.3f2 samples; sample 12 corresponds to pooled variant PylB.JM10.1 samples. PylB was estimated at ~90% purity for each sample following cleavage of histidine tags. White arrows indicates the expected position of PylB (~52 kDa) on each gel. BLUEstain^{TM} 2 Protein ladder (GoldBio) was run alongside samples on each gel.

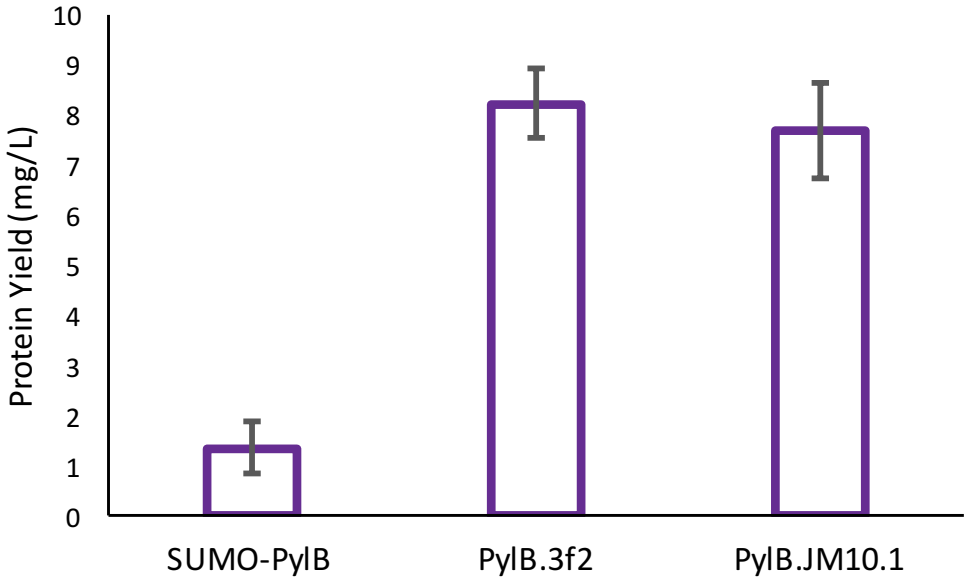


Figure S13. Protein yields of PylB variants. Protein variants were overexpressed in *E. coli* strain BL21 (DE3) cells following IPTG induction (see Methods). Yields of His-tag containing proteins purified by metal affinity chromatography are shown in mg of PylB protein produced per liter of cell culture, adjusted for protein purity. PylB variants 3f2 and JM10.1 produced 6.0-fold and 5.6-fold greater protein yields, respectively, compared to SUMO-PylB. Protein samples were purified from biological triplicate cultures; data shown represents mean values ± s.d.

Level	[BocK]/ μM	PylRS variant	Genotype of *gIII*	PylRS level	Plasmid No.
1	200	Chimeric PylRS	S12O	High	JH61
2	50	Chimeric PylRS	S12O	High	JH61
3	25	Chimeric PylRS	S12O	High	JH61
4	10	Chimeric PylRS	S12O	High	JH61
5	0	Chimeric PylRS	S12O	High	JH61
6	0	Chimeric PylRS	S12O	Low	JH19
7	0	Chimeric PylRS	S12O, Y166O	Low	JH76
8	0	Chimeric PylRS	S12O, P83O, Y166O	Low	JH77
9	0	32A-N-terminal	S12O	High	JH190
10	0	32A-N-terminal	S12O, Y166O	High	JH191
11	0	32A	S12O	Low	JH202

Table S1. Evolutionary selection conditions. Selection stringency was gradually increased by withdrawal of BocK, use of PylRS variants with lower affinities for Pyl, incorporation of additional stop codons in gIII, and lowering the expression level of PylRS. Chimeric PylRS was previously prepared by fusion of the *Methanosarcina barkeri* PylRS N-terminal and *M. mazei* PylRS C-terminal domains, as previously reported[5]. Variant 32A and variant 32A-N-terminal are chimeric PylRS mutants with reduced recognition of Pyl[1].

	SUMO		PylB							PylC		PylD				
	S6L	T26I	N61S	E84K	E122K	E175K	E178K	G188R	E251D	E315K	E316D	A19V	T42I	T54I	K142N	A252V
Line A plaque 1			N61S	E84K	E122K											
Line A plaque 2			N61S		E122K		E178K	G188R								
Line A plaque 3			N61S	E84K	E122K											
Line A plaque 4			N61S	E84K	E122K											
Line A plaque 5			N61S	E84K	E122K											
Line A plaque 10			N61S	E84K	E122K											
Line A plaque 11			N61S	E84K	E122K						E316D					
Line A plaque 12			N61S	E84K	E122K						E316D					
Line A plaque 13			N61S	E84K	E122K						E316D					
Line A plaque 15			N61S	E84K	E122K											
Line B plaque 2	S6L		N61S				E178K	G188R	E251D							A252V
Line B plaque 3			N61S				E178K		E251D							A252V
Line B plaque 4			N61S			E175K	E178K		E251D							A252V
Line B plaque 5			N61S					G188R	E251D							A252V
Line B plaque 6			N61S				E178K									A252V
Line B plaque 7			N61S				E178K		E251D							A252V
Line B plaque 8	S6L		N61S						E251D							A252V
Line B plaque 9							E178K			E315K						A252V
Line B plaque 10				E84K			E178K		E251D							A252V
Line B plaque 11			N61T					G188R	E251D			A19V				A252V
Line B plaque 14			N61S					G188R								A252V
Line B plaque 15			N61S				E178K		E251D							A252V
Line C plaque 1			N61S			E175K	E178K									
Line C plaque 2			N61S			E175K	E178K							T54I	K142N	
Line C plaque 3			N61S			E175K	E178K									
Line C plaque 4			N61S				E178K	G188R								
Line C plaque 5		T26I	N61S				E178K	G188R								
Line C plaque 6			N61S			E175K	E178K									
Line C plaque 7			N61S			E175K	E178K									
Line C plaque 9			N61S			E175K	E178K									
Line C plaque 10			N61S			E175K	E178K									
Line C plaque 11			N61S				E178K	G188R								
Line C plaque 12			N61S				E178K	G188R					T42I			
Line C plaque 14		T26I	N61S				E178K	G188R								
Line C plaque 15			N61S			E175K	E178K									

Table S2. List of mutations in the phage plaques sequenced at the endpoints of each lineage (A–C). Corresponding lineage and plaque number are listed on the left. Gene within which each mutation was identified are listed at the top of the table. Mutation under the 'SUMO' label were identified within the SUMO tag appended to the N-terminus of *pylB*. Sub-populations with representative clusters of mutations were recloned into expression vectors for further analyses.

Mutation		**S6L**	N61S	E84K	E122K	E175K	E178K	G188R	E251D	G332E	E315K	E316D	A252V	N61S	E84K
Old codon		TCG	AAC	GAG	GAG	GAG	GAG	GGG	GAA	GGA	GAA	GAA	GCT	AAC	GAG
New codon		TTG	AGC	AAG	AAG	AAG	AAG	CGG	GAC	GAA	AAA	GAT	GTT	AGC	AAG
Transition (TS) or Transversion (TV)		TS	TS	TS	TS	TS	TS	TV	TV	TS	TS	TV	TS	TS	TS
Mutant Type	**Name**	**SUMO**	**PylB**								**PylC**		**PylD**	**pBAD**	**pCDF**
Control	Empty	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	JH572	JH334
Wild-type	Untagged	n/a												JH573	JH20
PANCE parent	Tagged													JH462	JH126
Evolved (A2)	36A_subpop1		N61S		E122K		E178K	G188R						JH463	JH219
Evolved (A11)	36A_subpop2		N61S	E84K	E122K							E316D		JH574	JH220
Evolved (B2)	34B_subpop3	S6L	N61S				E178K	G188R	E251D				A252V	JH575	JH221
Evolved (B4)	34B_subpop4		N61S			E175K	E178K		E251D				A252V	JH576	JH222
Evolved (B15)	34B_subpop5		N61S				E178K		E251D				A252V	N/A	JH223
Evolved (C10)	40C_subpop6		N61S			E175K	E178K							N/A	JH224
Combinatorial	3f2		N61S	E84K	E122K		E178K	G188R	E251D				A252V	JH577	JH338
Combinatorial	3f2_v2		N61S	E84K	E122K		E178K	G188R	E251D			E316D	A252V	JH578	JH341
Combinatorial	JM10.1	S6L	N61S	E84K	E122K	E175K	E178K	G188R	E251D				A252V	JH579	JH405
Combinatorial	3f2v2+B		N61S	E84K	E122K		E178K	G188R	E251D	G332E		E316D	A252V	JH580	JH351
Combinatorial	3f2+C		N61S	E84K	E122K		E178K	G188R	E251D		E315K		A252V	JH581	JH359

Table S3. Mutations found within the *pylBCD* cassette in the evolved variants (sub-populations 1–6) and combinatorial variants (3f2, 3f2v2, JM10.1, 3f2v2+B, and 3f2+C). Nucleotide sequences corresponding to each mutation are shown at the top of the table. On the far right, plasmid names for both pBAD and pCDF expression vectors containing the corresponding *pylBCD* variant are shown. Genotypes of evolved variants emerged in different plaques isolated following ALT-PANCE; combinatorial variants were cloned by rationally combining mutations identified in different plaques (see Text). A separate numbering scheme was used for the single mutation observed within the SUMO tag region of PylB (mutation S6L). For subsequent mutations in PylB, the first residue following the SUMO tag was designated residue '1'. As the start codon is included within the SUMO tag region, subsequent PylB residues are labeled with one number lower compared to the corresponding residue in WT PylB (eg., residue N61 here corresponds to residue N62 within the WT protein sequence).

B	B Ions	B+2H	B-NH3	B-H2O	AA	Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1					S					15
2	187.11				V					14
3	343.21				R					13
4	400.23				G					12
5	529.27	265.14	512.25	511.26	E					11
6	586.29				G					10
7	715.34			697.33	E					9
8	772.36				G	898.5	449.75		880.49	8
9	887.39	444.2		869.37	D	841.48				7
10	958.42				A	726.45			708.44	6
11	1,059.47	530.24	1,042.44		T	655.41	328.21		637.4	5
12	1,296.62	648.81			O (N+123)	554.37				4
13		677.32			G	317.22	159.11			3
14		741.37			K	260.2		243.17		2
15					L					1

Table S4. Fragmentation table for LC-MS/MS spectra of sfGFP.2TAG with Pyl insertion confirmed at N39O. Masses of each B and Y ion are shown for each fragment. Masses of ions with an additional H^+, NH_3, and/or H_2O group are shown when detected. Amino acid residues corresponding to each fragment mass are indicated; O is used to indicate Pyl residues, with the amino acid and mass shift for the residue contained in the WT sfGFP sequence given in parenthesis.

B	B Ions	B+2H	B-NH3	B-H2O	AA	Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1					N					22
2	262.12		245.09		F					21
3			359.13		N					20
4					S		1,112.10			19
5	600.25		583.23		H					18
6	714.3			696.28	N	1,999.10				17
7	813.36				V	1,885.05				16
8	1,050.51			1,032.50	O (Y+74)					15
9	1,163.60	582.3	1,146.57		I					14
10		632.83		1,246.63	T	1,435.75	718.38		1,417.74	13
11	1,335.68			1,317.67	A	1,334.71				12
12	1,450.71			1,432.70	D	1,263.67				11
13	1,578.80	789.9			K	1,148.64		1,131.62		10
14			1,689.83		Q	1,020.55		1,003.52	1,002.54	9
15					K	892.49			874.48	8
16					N	764.39		747.37		7
17		1,004.01			G	649.37				6
18		1,060.55			I	592.35				5
19		1,124.60			K	479.26				4
20					A	351.17				3
21					N	280.13		263.1		2
22					F					1

Table S5. Fragmentation table for LC-MS/MS spectra of sfGFP.2TAG with Pyl insertion confirmed at Y151O. Masses of each B and Y ion are shown for each fragment. Masses of ions with an additional H^+, NH_3, and/or H_2O group are shown when detected. Amino acid residues corresponding to each fragment mass are indicated; O is used to indicate Pyl residues, with the amino acid and mass shift for the residue contained in the WT sfGFP sequence given in parenthesis.

Plasmid	Purpose	Insert	Mutation	Resistance	Origin
JH19	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.1TAG, luxAB; Plpp pylS; PproK pylT	gIII(P29*)	Spectinomycin	ColE1
JH60	Microscopy inclusion body assay	pCDF.CD	PylCD partial operon (PylB deleted to assess impact on inclusion body formation)	Spectinomycin	pBR322
JH61	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.1TAG, luxAB; Plpp [eG SD8] pylS; PproK pylT	gIII(P29*); [strong promoter eG, strong RBS SD8] pylS	Spectinomycin	ColE1
JH76	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.2TAG, luxAB; Plpp pylS; PproK pylT	gIII(P29*, Y184*)	Spectinomycin	ColE1
JH77	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.3TAG, luxAB; Plpp pylS; PproK pylT	gIII(P29*, P83*, Y184*)	Spectinomycin	ColE1
JH123	Microscopy inclusion body assay	pCDF.B.(SD4)CD	PylBCD operon (RBS inserted for PylC)	Spectinomycin	pBR322
JH126	Microscopy inclusion body assay	pCDF.SUMO-B.(SD4).CD	PylBCD operon (SUMO tag inserted for PylB; RBS inserted for PylC)	Spectinomycin	pBR322
JH190	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.1TAG, luxAB; Plpp [eG SD8] 32A-Nter; PproK pylT	gIII(P29*); [strong promoter eG, strong RBS SD8] 32A-Nter pylS variant	Spectinomycin	ColE1
JH191	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.2TAG, luxAB; Plpp [eG SD8] 32A-Nter; PproK pylT	gIII(P29*, Y184*); [strong promoter eG, strong RBS SD8] 32A-Nter pylS variant	Spectinomycin	ColE1
JH202	Accessory Plasmid for PACE of pylBCD variants	Ppsp gIII.1TAG, luxAB; Plpp 32A; PproK pylT	gIII(P29*); 32A pylS variant	Spectinomycin	ColE1
JH223	pylBCD variant in pCDF backbone	PT7/lac (pylBCD 34B_sub-pop5); LacI	PylB(N61S, E178K, E251D) PylD(A252V)	Spectinomycin	pBR322
JH224	pylBCD variant in pCDF backbone	PT7/lac (pylBCD 40C_sub-pop6); LacI	PylB(N61S, E175K, E178K)	Spectinomycin	pBR322
JH291	Reporter plasmid for pylBCD variant in pCDF backbone	Para sfGFP.2TAG; Plpp pylS; PproK pylT; AraC	sfGFP(N39*., Y151*)	Ampicillin	p15A
JH462	pylBCD variant in pBAD backbone	Para (Sumo-tagged pylBCD); AraC	SUMO tag fused to N-terminus of PylB	Ampicillin	pBR322
JH463	pylBCD variant in pBAD backbone	Para (pylBCD 36A_sub-pop1); AraC	PylB(N61S, E122K, E178K, G188R)	Ampicillin	pBR322
JH572	Negative control pBAD backbone	pBAD backbone	No insert	Ampicillin	pBR322
JH573	pylBCD variant in pBAD backbone	Para (Untagged pylBCD); AraC	Codon-optimized Methanosarcina acetivorans pylBCD	Ampicillin	pBR322
JH574	pylBCD variant in pBAD backbone	Para (pylBCD 36A_sub-pop2); AraC	PylB(N61S, E84K, E122K) PylC(E316D)	Ampicillin	pBR322
JH575	pylBCD variant in pBAD backbone	Para (pylBCD 34B_sub-pop3); AraC	Sumo(S6L) PylB(N61S, E178K, G188R, E251D) PylD(A252V)	Ampicillin	pBR322
JH576	pylBCD variant in pBAD backbone	Para (pylBCD 34B_sub-pop4); AraC	PylB(N61S, E175K, E178K, E251D) PylD(A252V)	Ampicillin	pBR322
JH577	pylBCD variant in pBAD backbone	Para (pylBCD 3f2); AraC	PylB(N61S, E84K, E122K, E178K, G188R, E251D) PylD(A252V)	Ampicillin	pBR322
JH578	pylBCD variant in pBAD backbone	Para (pylBCD 3f2_v2); AraC	PylB(N61S, E84K, E122K, E178K, G188R, E251D) PylC(E316D) PylD(A252V)	Ampicillin	pBR322
JH579	pylBCD variant in pBAD backbone	Para (pylBCD JM10.1); AraC	Sumo(S6L) PylB(N61S, E84K, E122K, E175K, E178K, G188R, E251D) PylD(A252V)	Ampicillin	pBR322
JH580	pylBCD variant in pBAD backbone	Para (pylBCD 3f2v2+B); AraC	PylB(N61S, E84K, E122K, E178K, G188R, E251D, G332E) PylC(E316D) PylD(A252V)	Ampicillin	pBR322
JH581	pylBCD variant in pBAD backbone	Para (pylBCD 3f2+C); AraC	PylB(N61S, E84K, E122K, E178K, G188R, E251D) PylC(E315K) PylD(A252V)	Ampicillin	pBR322
JH528	Reporter plasmid for pylBCD variant in pBAD backbone	Para sfGFP.1TAG; Plpp pylS; PproK pylT; AraC	sfGFP(N39*)	Tetracycline	p15A
JH530	Reporter plasmid for pylBCD variant in pBAD backbone	Para sfGFP.3TAG; Plpp pylS; PproK pylT; AraC; Tet	sfGFP(N39*, N135*, Y151*)	Tetracycline	p15A
JH643	Reporter plasmid for pylBCD variant in pBAD backbone	Para sfGFP.3TAG; Plpp pylS; PproK pylT; AraC; Cam	sfGFP(N39*, N135*, Y151*)	Chloramphenicol	p15A
JH645	Reporter plasmid for pylBCD variant in pBAD backbone	Pem7 luxCDA(B.1TAG)E; Plpp pylS; PproK pylT	Inserted amber codon immediately after initiator methionine of luxB	Chloramphenicol	p15A
JH750	Microscopy inclusion body assay	pCDF.3f2, deleted sumo	Evolved PylBCD operon; SUMO tag deleted from PylB [PylB(N61S, E84K, E122K, E178K, G188R, E251D) PylD(A252V)]	Spectinomycin	pBR322
JH751	Microscopy inclusion body assay	pCDF.JM10.1, deleted sumo	Evolved PylBCD operon; SUMO tag deleted from PylB [PylB(N61S, E84K, E122K, E175K, E178K, G188R, E251D) PylD(A252V)]	Spectinomycin	pBR322
JH767	PylB overexpression and purification	PT7/Lac-[His6-TEV-SUMO-PylB.WT]	pET28a_LacI_PT7/Lac-[His6-TEV-SUMO-PylB]	Kanamycin	pBR322
JH768v2	PylB overexpression and purification	PT7/Lac-[His6-TEV-SUMO-PylB.3f2]	pET28a_LacI_PT7/Lac-[His6-TEV-SUMO-PylB(N61S, E84K, E122K, E178K, G188R, E251D)]	Kanamycin	pBR322
JH769v2	PylB overexpression and purification	PT7/Lac-[His6-TEV-SUMO-PylB.JM10.1]	pET28a_LacI_PT7/Lac-[His6-TEV-SUMO-PylB(N61S, E84K, E122K, E175K, E178K, G188R, E251D)]	Kanamycin	pBR322

Table S6. Description of plasmids deposited in Addgene.

Supplementary References

1. Suzuki, T. et al. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat Chem Biol* **13**, 1261-+ (2017).
2. Esvelt, K.M., Carlson, J.C. & Liu, D.R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499-U550 (2011).
3. Fan, C.G., Ho, J.M.L., Chirathivat, N., Soll, D. & Wang, Y.S. Exploring the Substrate Range of Wild-Type Aminoacyl-tRNA Synthetases. *Chembiochem* **15**, 1805-1809 (2014).
4. Bryson, D.I. et al. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat. Chem. Biol.* **13**, 1253-1260 (2017).
5. Ho, J.M. et al. Efficient Reassignment of a Frequent Serine Codon in Wild-Type Escherichia coli. *Acs Synth Biol* **5**, 163-171 (2016).
6. Quitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis. *Angew Chem Int Edit* **51**, 1339-1342 (2012).
7. Kramer, R.M., Shende, V.R., Motl, N., Pace, C.N. & Scholtz, J.M. Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility. *Biophys J* **102**, 1907-1915 (2012).
8. Gaston, M.A., Zhang, L., Green-Church, K.B. & Krzycki, J.A. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature* **471**, 647-650 (2011).
9. Longstaff, D.G. et al. A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1021-1026 (2007).
10. Vollrath, F., Hawkins, N., Porter, D., Holland, C. & Boulet-Audet, M. Differential Scanning Fluorimetry provides high throughput data on silk protein transitions. *Sci Rep* **4**, 5625 (2014).