Article

Enhanced Solubilization of Class B Radical S-adenosylmethionine Methylases by Improved Cobalamin Uptake in *Escherichia coli*

Nicholas Lanz, Anthony J. Blaszczyk, Erin McCarthy, Bo Wang, Roy Wang, Brianne Jones, and Squire J. Booker

Biochemistry, **Just Accepted Manuscript** • DOI: 10.1021/acs.biochem.7b01205 • Publication Date (Web): 03 Jan 2018

Enhanced Solubilization of Class B Radical *S*-adenosylmethionine Methylases by Improved Cobalamin Uptake in *Escherichia coli*

Nicholas D. Lanz, [†, §, #] Anthony J. Blaszczyk, [†, #] Erin L. McCarthy, [†] Bo Wang, [‡] Roy X. Wang, [‡] Brianne S. Jones, [‡] and Squire J. Booker[†, ‡, ⊥, *]

Departments of [†]Biochemistry and Molecular Biology and of [‡]Chemistry, and the [⊥]Howard Hughes Medical Institute, The Pennsylvania State University, University Park, PA 16802

[§]Present address. Laboratory of Bacterial Polysaccharides, Center for Biologics Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, 20993

[#]These two individuals contributed equally

Correspondence and requests for materials should be addressed to S. J. B. (squire@psu.edu)

ABSTRACT

The methylation of unactivated carbon and phosphorus centers is a burgeoning area of biological chemistry, especially given that such reactions constitute key steps in the biosynthesis of numerous enzyme cofactors, antibiotics, and other natural products of clinical value. These kinetically challenging reactions are catalyzed exclusively by enzymes in the radical *S*-adenosylmethionine (SAM) superfamily and have been grouped into four classes: A, B, C, and D. Class B radical SAM (RS) methylases require a cobalamin cofactor in addition to the [4Fe-4S] cluster that is characteristic of RS enzymes. However, their poor solubility upon overexpression and their generally poor turnover has hampered detailed *in vitro* studies of these enzymes. It has been suggested that improper folding, possibly caused by insufficient cobalamin during their overproduction in *Escherichia coli*, leads to formation of inclusion bodies. Herein, we report our efforts to improve the overproduction of class B RS methylases in a soluble form by engineering a strain of *E. coli* to take in more cobalamin. We cloned five genes (*btuC*, *btuE*, *btuD*, *btuF*, and *btuB*) that encode proteins that are responsible for cobalamin uptake and transport in *E. coli* and coexpressed these genes with those that encode TsrM, Fom3, PhpK, and ThnK, four class B RS methylases that suffer from poor solubility during overproduction. This strategy markedly enhances cobalamin uptake into the cytoplasm and improves the solubility of the target enzymes significantly.

INTRODUCTION

Radical *S*-adenosylmethionine (SAM) methylases constitute one of the largest subfamilies of the radical SAM (RS) superfamily.[1-3] These enzymes methylate unactivated carbon and phosphorus centers, many of which are found in a growing number of natural products that display antibiotic or antineoplastic activity, as well as enzyme cofactors and other cellular metabolites, such as bacteriochlorophyll (enzyme cofactor),[4] pactamycin (antitumor agent and antibiotic),[5] mitomycin C (antitumor agent),[6] moenomycin A (phosphoglycolipid antibiotic),[7] fosfomycin (broad-spectrum antibiotic),[8-11] thienamycin (β-lactam antibiotic),[12] gentamicin (aminoglycoside antibiotic),[13] clorobiocin (aminocoumarin antibiotic),[14] fortimicin (aminoglycoside antibiotic),[15, 16] thiostrepton (thiopeptide antibiotic)[17] polytheonamide

(secondary metabolite),[18, 19] quinomycin (peptide antibiotic),[20] and chondrochloren (antibiotic)[21] (**Figure 1**). They have been organized into four classes based on cofactor requirement, architecture, and mechanism of action.[2, 3] The best studied of the RS methylases are RlmN and Cfr, which use a cysteine dyad to methylate C2 and C8 of adenosine 2503 in ribosomal RNA or C2 of adenosine 37 in specific transfer RNAs (tRNAs).[22-27] Class B RS methylases require a cobalamin cofactor, and methylate both sp^2- and sp^3-hybridized carbon centers as well as phosphinate phosphorus atoms. Class C enzymes uniquely methylate, methenylate, or cyclopropanate sp^2-hydridized carbon centers that are typically found in complex natural products.[28-30] Class D RS methylases are the least characterized. Unlike class A, B, and C enzymes, which transfer a C1 unit from SAM, class D enzymes employ an alternative source of the C1 unit, hypothesized to be a derivative of methylenetetrahydrofolate. A fifth class of RS methylases has also recently been suggested. These class E enzymes appear to methylate a sulfido ion in a metallocluster intermediate during the maturation of the M cluster of nitrogenase.[31]

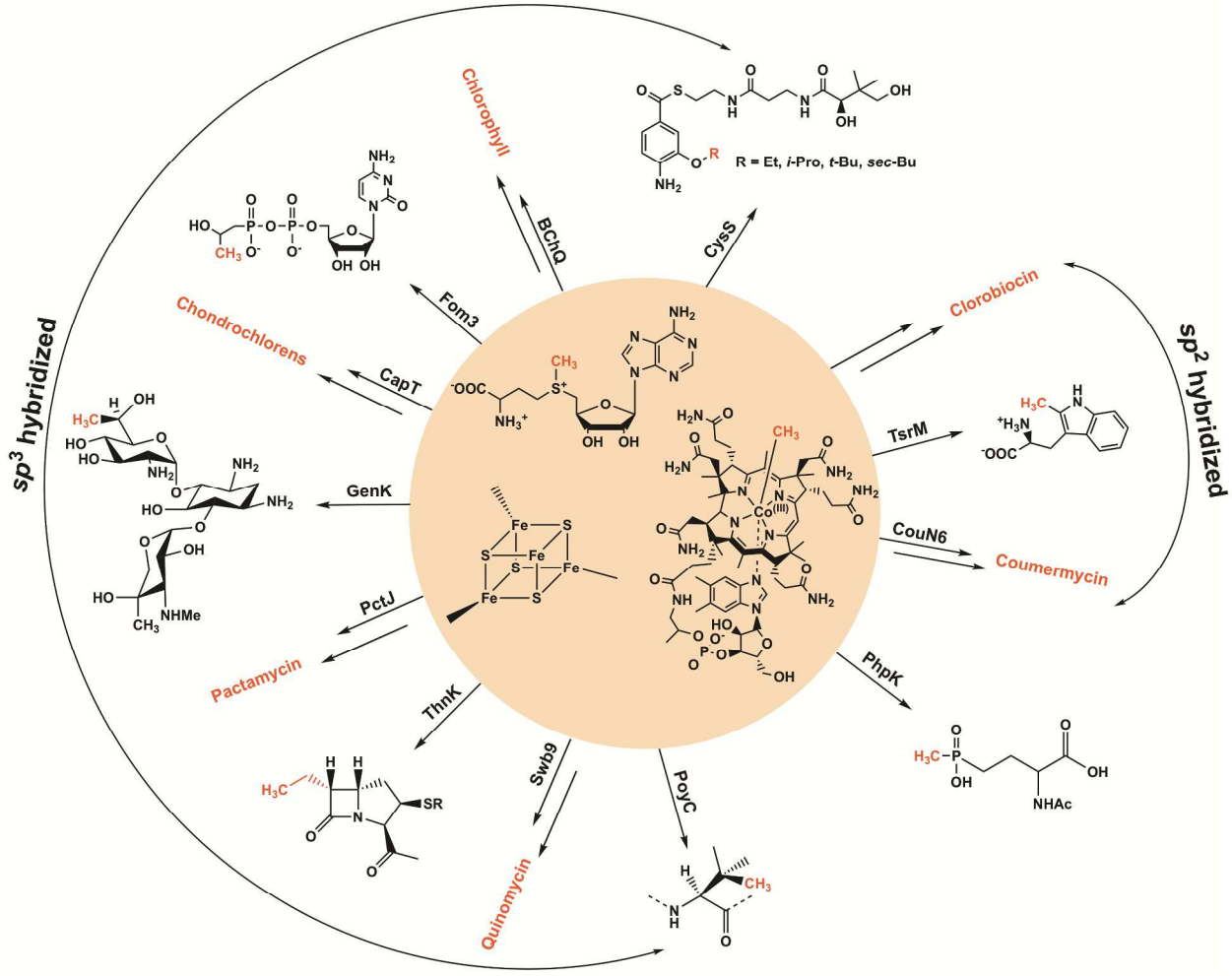


Figure 1. Various natural products that contain methyl moieties installed by cobalamin-dependent RS methylases. Compounds with structures shown represent products of class B RS methylases that have been studied *in vitro*. One arrow indicates the actual product from an *in vitro* reaction, while two arrows indicate the final natural product whose biosynthesis involves a class B RS methylation step.

Class B enzymes are the most abundant RS methylases and appear to be the most versatile. Moreover, their importance in the biosynthesis of numerous natural products of clinical value has led to efforts to isolate and characterize them and their mechanisms of catalysis. However, they have been resistant to robust mechanistic analysis because of poor solubility upon overproduction and/or low activity.[2] Three of the first class B RS methylases to be the subject of *in vitro* reports—PhpK from *Kitasatospora*

phosalacinea, GenK from *Micromonospora echinospora*, and Fom3 from *Streptomyces wedmorensis*—were all purified from inclusion bodies and then refolded and reconstituted with iron, sulfide, and methylcobalamin (MeCbl) to generate their iron-sulfur (Fe-S) and cobalamin cofactors. Characterization of the enzymes by UV-vis and electron paramagnetic resonance (EPR) spectroscopies, when conducted, showed them to exhibit limited cofactor incorporation, and their activities were poor if at all existent. A fourth class B methylase, TsrM from *Streptomyces laurentii*, was the first to be isolated from the soluble fraction of *E. coli* crude extract.[32] Although yields of pure protein were not reported, the enzyme was of modest quality as judged by its activity and UV-visible spectrum. In subsequent studies of the enzyme, a construct encoding TsrM fused to the protein SUMO was used for overproduction in *E. coli* cultured in a specialized medium containing ethanolamine as the sole source of nitrogen and carbon, conditions that have been suggested to enhance cobalamin uptake in *E. coli*.[33] Using this strategy, low milligram quantities of homogeneous enzyme could be purified from 32 L of *E. coli* culture, allowing for spectroscopic studies of the enzyme to be conducted.[34] A similar strategy proved successful for purifying soluble ThnK[35]—a class B RS methylase involved in the biosynthesis of thienomycin—and PoyC, a class B RS methylase involved in the biosynthesis of polytheonamide.[18, 36]

In this work, we extend our efforts to isolate soluble class B RS methylases in larger yields by co-expressing RS methylase genes along with *btuC*, *btuE*, *btuD*, *btuF*, and *btuB* from *E. coli*, which, except for *btuE*, encode proteins that are involved in cobalamin uptake and trafficking. These genes were cloned into a pBAD42 vector under the control of an arabinose-inducible P_{BAD} promoter, and the resulting pBAD42-BtuCEDFB vector was used to transform *E. coli* containing pDB1282—which harbors genes involved in Fe-S cluster biosynthesis—and a pET-based vector containing the target RS methylase. Upon expression of the *btu* genes, a substantial increase in the intracellular cobalamin concentration—particularly that of hydroxocobalamin (OHCbl) and adenosylcobalamin (AdoCbl) (**Figure 2)**—is observed. Furthermore, coexpression of either the *tsrM*, *fom3*, *thnK*, or *phpK* genes with the *btu* genes leads to a marked increase in soluble purified proteins.

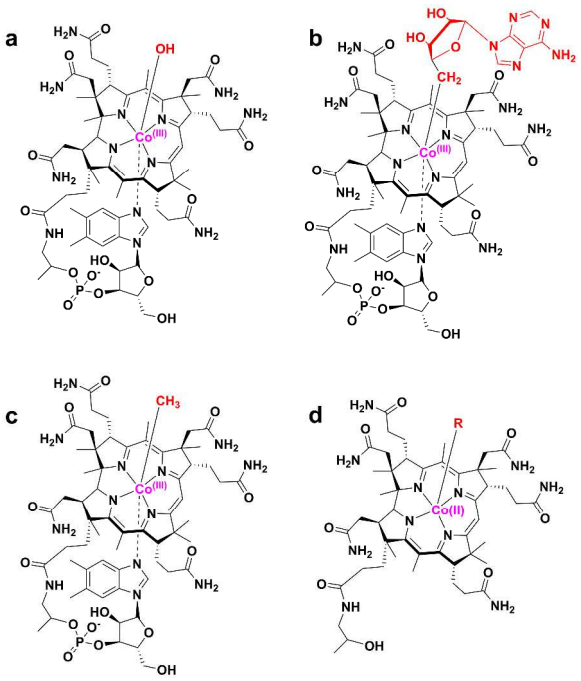


Figure 2. Structures of (a) hydroxocobalamin, (b) adenosylcobalamin, (c) methylcobalamin, and (d) cobinamide.

Materials and Methods

Materials

DNA primers were obtained from IDT DNA (Coralville, IA). Enzymes and other reagents for cloning and nucleic acid manipulation were obtained from New England Biolabs (Ipswich, MA). OHCbl, MeCbl, and AdoCbl were purchased from Sigma-Aldrich (St. Louis, MO). Ferric chloride was obtained from EMD Biosciences (Gibbstown, NJ). Bradford reagent for protein concentration determination as well as the bovine serum albumin (BSA) standard was purchased from Pierce, Thermo Fisher Scientific (Rockford, IL). Dithiothreitol (DTT), L-(+)-arabinose, isopropyl β-D-1-thiogalactopyranoside (IPTG), ampicillin, spectinomycin, and kanamycin were purchased from Gold Biotechnology (St. Louis, MO). SUMO express protease was purchased from Lucigen (Middleton, WI). All other materials and reagents have been previously reported,[34, 37] or were of the highest available quality.

General Methods. The polymerase chain reaction (PCR) was conducted using either a Stratagene Robocycler Gradient 40 or a BioRad S1000 thermocycler. DNA isolation kits were purchased from Macherey-Nagel (Bethlehem, PA). The pRham N-His SUMO-TsrM (pRham-SUMO-TsrM) construct was created using the Expresso Rhamnose SUMO cloning system (Lucigen; Middleton, WI). The P1vir, Δ*btuR BW25113* strain, pBAD42, and pCP20 were purchased from the Yale Coli Genetic Stock Center. A Cary 50 spectrometer from Varian was used to record UV-visible spectra using the WinUV software package to control the system. All anaerobic experiments were conducted in a Coy (Grass Lakes, MI) anaerobic chamber.

Cloning of the btuCEDFB Operon. The genes encoding the cobalamin transport system of *E. coli* were cloned into a pBAD42 plasmid under the control of the P_{BAD} promoter using semi-nested restriction sites to allow for the genes to be inserted sequentially. Three of the five genes, *btuC*, *btuE*, and *btuD*, are located in an operon in the *E. coli* genome. The remaining genes, *btuF* and *btuB*, are found in unrelated operons. The *btuCED* operon was amplified from *E. coli* K12 genomic DNA using the PCR. The forward primer, 5′-CGC GGC GTC GAA TTC ATG CTG ACA CTT GCC CGC CAA CAA CAG CG-3′, contained a nine base-pair GC clamp, an *Eco*RI restriction site (underlined), and the first 29 base pairs of *btuC*. The reverse primer, 5′- CGC GGC GTC TCT AGA CGG TAG GAC **CTC GAG** TCA GAT GGT CGA AAT CAG CAT TCT GTG ACC-3′, contained a nine base-pair GC clamp, an *Xba*I restriction site (underlined), an arbitrary nine base-pair sequence, an *Xho*I restriction site (bolded), and the last 30 base-pairs of *btuD*. The PCR product was digested with *Eco*RI and *Xba*I and ligated into the pBAD42 vector digested with the same enzymes.

The *btuF* gene was amplified using a forward primer, 5′-CGC GGC GTC CTC GAG ATG GCT AAG TCA CTG TTC AGG GCG CTG GTC-3′, containing a nine base-pair GC clamp, an *Xho*I restriction site (underlined), and the first 30 base-pairs of *btuF*. The reverse primer, 5′-CGC GGC GTC TCT AGA CGG TAG GAC **GGT ACC** TCA ATC TAC CTG TGA AAG CGC ATT ACA GAG-3′,

contained a nine base-pair GC clamp, an *Xba*I restriction site (underlined), an arbitrary nine base-pair sequence, a *Kpn*I restriction site (bolded), and the last 30 base-pairs of the *btuF* gene. The PCR product was digested with *Xho*I and *Xba*I and ligated into the pBAD-*btuCED* construct after it was digested with the same enzymes.

The *btuB* gene was amplified using a forward primer, 5′-CGC GGC GTC GGT ACC ATG ATT AAA AAA GCT TCG CTG CTG ACG GCG -3′, containing a nine base-pair GC clamp, a *Kpn*I restriction site (underlined), and the first 30 base-pairs of *btuB*. The reverse primer, 5′- CGC GGC GTC TCT AGA TTA GAA GGT GTA GCT GCC AGA CAA GGT G -3′, contained a nine base-pair GC clamp, an *Xba*I restriction site (underlined), and the last 28 base-pairs of the *btuB* gene. The PCR product was digested with *Kpn*I and *Xba*I and ligated into the pBAD-*btuCED-btuF* construct after it was digested with the same enzymes.

Spacing the Genes for Expression as an Operon. The resulting pBAD-*btuCED-btuF-btuB* construct was modified by sequential site-directed mutagenesis to remove the interstitial *Xho*I and *Kpn*I restriction sites using the primer, 5′-GGT CAC AGA ATG CTG ATT TCG ACC ATC TGA TG GCT AAG TCA CTG TTC AGG GCG CTG GTC-3′, and its reverse complement to remove the *Xho*I site, and the primer, 5′-C TGT AAT GCG CTT TCA CAG GTA GAT TGA TG ATT AAA AAA GCT TCG CTG CTG ACG GCG-3′, and its reverse complement to remove the *Kpn*I site. The stop codon of the 5′ gene and the start codon of the 3′ gene overlap by one base-pair in both cases (underlined) for optimal operon expression.[38, 39]

Optimizing Ribosome Binding Sites. To increase the translation efficiency of the encoded genes, the associated ribosome binding sites (rbs) were optimized by site-directed mutagenesis. Two ribosome binding sites are present within the construct, one encoded by the pBAD42 vector 5′ of the *btuC* gene and a second in the >60 base-pair intergenic region between *btuC* and *btuE*. Both rbs were converted to 5′-

TAAGGAGGT–X_5– ATG-3′, where the final three bases are the start codon of the following gene. The resulting plasmid was designated pBAD42-BtuCEDFB.

Overexpression of the btuCEDFB Operon. The *btuCEDFB* operon was expressed in the BL21 (DE3) strain of *E. coli* in M9-ethanolamine medium, M9-minimal medium, or LB medium, each of which was supplemented with 2.5 μM hydroxocobalamin. The M9-ethanolamine media utilizes ethanolamine as the sole carbon and nitrogen source and, contrary to M9-minimal medium, is supplemented with OHCbl, trace metals, and all 20 amino acids, but lacks glucose and ammonium chloride. Cultures were incubated at 37 °C until they reached an OD_{600} of 0.3, at which point gene expression was induced with 0.2% (w/v) arabinose. The cultures were incubated at 37 °C for an additional 4 h and harvested by centrifugation. To test the ability of btuCEDFB to increase cobalamin uptake, growth studies were also conducted with an empty pBAD42 vector under the same conditions described above.

Cloning of the phpK and fom3 Genes. The *phpK* gene from *Kitasatospora phosalacinea* and the *fom3* gene from *Streptomyces fradiae* were codon-optimized for expression in *E. coli* and synthesized by GeneArt (ThermoFisher, Carlsbad, CA). The genes were delivered in a pMK-RQ-Bb vector between a 5′ *Bsa*I restriction site and a 3′ *Xho*I restriction site. The genes were excised from the pNMK-RQ-Bb vector with *Bsa*I and *Xho*I and then subcloned into a pSUMO vector digested with the same enzymes. The resulting constructs encode PhpK or Fom3 fused to the C-terminus of SUMO, which contains an N-terminal hexahistidine tag. The correct DNA sequences were verified and designated pSUMO-PhpK and pSUMO-Fom3.

Analysis of Operon Overexpression. Harvested cells were introduced into an anaerobic chamber, resuspended in buffer containing 50 mM HEPES, pH 7.5, 300 mM KCl, 10% glycerol, and 1 mM β-mercaptoethanol, and lysed by sonication. The lysate was centrifuged at 25,000 × *g* for 1 h at 4 °C. All subsequent studies were conducted in dim light in an anaerobic chamber to avoid photolysis of

cobalamin. The supernatant was treated with an equal volume of 200 mM H_2SO_4 to precipitate the protein component and release the cobalamin cofactor. The samples were centrifuged again for 15 min at 18,000 × *g* before analysis by liquid chromatography with tandem mass spectrometry (LC-MS/MS).

LC-MS/MS Analysis of Cobalamin Content. LC-MS/MS analysis of the cobalamin forms in *E. coli* was performed using an Agilent 3470 QQQ MS attached to an Agilent 1290 Infinity II LC system. Samples were separated on an Agilent Technologies Zorbax Extend-C18 RRHD column (2.1 mm × 50 mm, 1.8 μm particle size) equilibrated in 95% solvent A (0.1% formic acid) and 5% solvent B (acetonitrile). An isocratic gradient was maintained for 0.4 min, which was followed by a gradient of 5–20% solvent B from 0.4 to 0.8 min. Another gradient from 20-42% solvent B from 0.8 to 1.6 min was applied before returning to 5% solvent B from 1.6 to 1.9 min. A flow rate of 0.35 mL/min was maintained throughout the chromatographic procedure. The column was allowed to re-equilibrate for 0.5 min under initial conditions before subsequent sample injections. OHCbl elutes at approximately 1.16 min under these conditions, while AdoCbl and MeCbl elute at 1.51 and 1.71 min, respectively.

OHCbl, MeCbl, and AdoCbl were detected by MS/MS multiple-reaction monitoring (MRM) using Agilent Jet Spray electrospray-ionization in positive mode ($AJSESI^+$). OHCbl was detected in the +2 charge-state at an *m/z* of 664.9, which corresponds to the loss of a water molecule. Transitions of 664.9→635.8 and 664.9→147.0 were used to quantify hydroxocobalamin. Transitions of 673.0→665.6 and 673.0→147.0 for MeCbl, and 790.6→665.6 and 790.6→147.0 for AdoCbl were used to quantify those molecules. The $AJSESI^+$ source was equilibrated with a nitrogen gas temperature of 300 °C and flow rate of 5.0 L/min, a nebulizer pressure of 45 PSI, and a capillary voltage of 3500 V. The optimal fragmentor voltages for OHCbl, MeCbl, and AdoCbl were determined to be 120 V, 104 V, and 108 V, respectively.

	Transition (*m/z*)	**Collision Energy (V)**
Hydroxocobalamin	664.9→635.8	11

	664.9→147.0	41
Methylcobalamin	673.0→665.6	17
	673.0→147.0	57
Adenosylcobalamin	790.6→665.6	17
	790.6→147.0	75

Standards (~10 μM) of OHCbl, MeCbl, and AdoCbl were prepared in water, and their concentrations were determined by boiling them in 100 mM potassium cyanide and measuring the resulting absorbance at 367 nm (ε_{367} = 30,800 M^{-1} cm^{-1}).[40] A standard curve for analysis by LC–MS/MS contained OHCbl, MeCbl, and AdoCbl at concentrations of 31 nM – 64 μM in two-fold increments, and 100 μM tyrosine, which was used as the internal standard (IS) for quantitative analysis.

Overexpression and Purification of Class B RS Methylases in the Presence of the btuCEDFB Operon. Genes encoding TsrM and Fom3 were expressed as previously described,[34] with the following two amendments: the BL21 (DE3) strain of *E. coli* was used rather than the RosettaBlue (DE3) pLysS strain, and the pBAD42-BtuCEDFB plasmid, which confers resistance to spectinomycin, was maintained in the strain by supplementation of the media with 50 μg/mL spectinomycin. Purification of the SUMO-tagged enzyme was performed as previously described.[34, 41] Briefly, harvested cells were resuspended and lysed inside an anaerobic chamber. Target proteins were purified by an initial immobilized metal affinity chromatography (IMAC) step using Ni-NTA, followed by cleavage of the SUMO tag using the Ulp1 protease overnight on ice. The cleavage reaction was reapplied to a Ni-NTA column to remove the hexahistidine-tagged Ulp1 protease, the SUMO tag, and any non-specifically bound contaminating proteins. The flow-through containing the native protein was collected and analyzed for purity by SDS-PAGE. The concentration was determined by the Bradford assay using bovine serum albumin, fraction V (BSA) as a standard, as previously described.[42]

Insertion of Gly Linker into pSUMO-PhpK. A codon for glycine was inserted between the gene encoding the SUMO tag and the *phpK* gene using site-directed mutagenesis to allow for efficient cleavage

of the resulting fusion protein by Ulp1. The forward primer, 5′-CGC GAA CAG ATT GGA GGT GGA ATG AAA CAT TGC-3′, and its reverse complement were used to insert the glycine linker codon (underlined) into the pSUMO-PhpK construct.

Overproduction and Purification of PhpK. *E. coli* BL21 (DE3) was transformed sequentially with the plasmids pSUMO-Gly-PhpK, pDB1282, and pBAD42-BtuCEDFB. A starter culture was inoculated from a single colony and incubated for 15 h at 37 °C while shaking at 250 rpm. Overexpression of the *phpK* gene was conducted at 37 °C in 16 L of auto-induction media supplemented with OHCbl (1.3 μM).[43] Each 6 L flask of culture media was inoculated with 1 mL of starter culture. At an OD_{600} of ~0.3, 50 μM $FeCl_3$ and 150 μM cysteine were added to the media and expression of the genes encoded on pDB1282 and pBAD42-BtuCEDFB was induced with 0.2% (w/v) of L-arabinose. Once cultures reached an OD_{600} of ~1.0, flasks were placed on ice for 20 min. Cultures were then incubated at 18 °C for 16 h before the cells were collected by centrifugation at 7,000 × *g* for 12 min. Purification of PhpK was performed as described above for TsrM and Fom3.

Overexpression and Purification of ThnK. *E. coli* BL21 (DE3) cells containing pDB1282 were transformed with the plasmid pET24b-ThnK (gift from the lab of Dr. Craig Townsend, Johns Hopkins University) and pBAD42-BtuCEDFB. Overexpression of the *thnK* gene was conducted in ethanolamine media, as described above. Purification of ThnK was performed as described above. However, because it was not overproduced as a fusion protein with SUMO, it was not treated with the SUMO protease. Rather, ThnK was purified further by size-exclusion chromatography using an ÅkTA FPLC system fitted with an S200 column and housed in an anaerobic chamber.

Determination of Protein Concentration. The concentrations of Fom3 and PhpK were determined by the Bradford assay using BSA as a standard. An amino acid analysis was conducted to determine correction factors for Fom3 and PhpK, as has previously been described.[44] It was determined that the

Bradford assay overestimates the concentration of Fom3 and Phpk by factors of 1.18 and 1.17, respectively, while a correction factor of 1.39 has previously been reported for TsrM.[34]

Determination of Stoichiometry of Bound Cobalamin. The concentration of protein-bound cobalamin was determined using a previously described procedure.[40] Each protein sample was boiled in 0.1 M potassium cyanide for 5 min to convert all forms of cobalamin to dicyanocobalamin. The concentration of dicyanocobalamin was calculated using the known extinction coefficient at 367 nm (ε_{367} = 30,800 M^{-1} cm^{-1}).

ΔbtuR BL21(DE3) Strain Construction. The *E. coli BW25113* strain with a kanamycin cassette replacing the *btuR* gene (Δ*btuR*:*kan*) was obtained from the Keio collection. P1 transduction was performed using established methods to move the genetic marker from the BW25113 strain to *E. coli* BL21 (DE3).[45] Briefly, the host strain (*BW25113* Δ*btuR*:*kan*) was cultured in LB media overnight at 37 °C with shaking at 250 rpm for ~14 h. The cells (1 mL) were then used to inoculate 200 mL of LB media supplemented with 0.2% glucose and 5 mM $CaCl_2$. 1 mL of a 1:100 dilution of P1*vir*, obtained from the Yale Coli Genetic Stock Center, was added to the culture (100 mL) and incubated for 4 h at 37 °C with shaking at 250 rpm. The OD_{600} was monitored every 30 min until the culture cleared, as evidenced by visible cell debris and a significant drop in the optical density. The culture was then treated with 5% (v/v) chloroform, vortexed, and centrifuged for 30 min at 3,500 × *g* at 4 °C. The supernatant was removed, carefully avoiding cell debris. The new lysate was stored at 4 °C supplemented with 5% (v/v) chloroform. To perform the transduction, the recipient strain, *E. coli* BL21(DE3), was first cultured at 37 °C with shaking at 250 rpm overnight. The following day, 200 μL of culture was aliquoted into four microcentrifuge tubes and centrifuged at 14,000 × *g* for 5 min. The supernatant was removed and the cells were resuspended in 100 μL λ-Ca++ buffer (10 mM Tris-HCl, pH 7.9, 20 mM $MgSO_4$, and 5 mM $CaCl_2$). Varying amounts (0 μL, 10 μL, 50 μL and 100 μL) of P1*vir* lysate were added to each of the tubes, which

were then incubated for 30 min at 37 °C without shaking. LB media (1 mL) was added to each of the tubes and supplemented with 100 mM sodium citrate. The samples were then incubated for an additional hour at 37 °C and then plated on LB + Kan (25 μg/mL) agar with an additional 100 μL of 1 M citrate spread on top of the plate. The plate was incubated at 37 °C overnight. The resulting colonies were purified by re-streaking, and then tested for sensitivity to P1 phage. The T7 tester phage from Novagen was used to confirm the presence of the DE3 lysogen in the strain according to the manufacturer's instructions. Following transduction, chemically competent Δ*btuR*:*kan* BL21(DE3) cells were transformed with the pCP20 plasmid—which encodes a temperature-sensitive origin of replication and the FLP recombinase—in order to excise the kanamycin cassette.[46] A culture was inoculated with a single colony and incubated overnight at 43 °C and 180 rpm to induce FLP recombinase expression and select for loss of pCP20. A 10^6 dilution of the culture was plated to select for individual colonies and incubated at 30 °C overnight to prevent partial loss of plasmids from colonies containing cells that did not lose the pCP20 plasmid. The following day, individual colonies were patched on three different plates in the following order: LB + kanamycin (25 μg/mL), LB + ampicillin (100 μg/mL), and LB to select for colonies that had lost the kanamycin cassette as well as the pCP20 plasmid, yielding the Δ*btuR* BL21(DE3) strain.

Cloning of the tsrM Gene into the pRham-SUMO Vector. The *tsrM* gene from *Streptomyces laurentii* was amplified from a plasmid encoding TsrM by the PCR using the following forward and reverse primers, respectively: 5′- CGC GAA CAG ATT GGA GGT ATG CTC CGT AAA GGC ACC GTC GC-3′ and 5′- GTG GCG GCC GCT CTA TTA CCG GAC GGC CTC CGC GAG CTC-3′. Following purification by gel-electrophoresis, Ecloni 10G chemically competent cells (Lucigen) were co-transformed with the pre-processed pRham-SUMO vector and the purified PCR product. The sequence of the resulting *tsrM-SUMO* construct was verified by DNA sequencing at the Penn State University Nucleic Acid Facility. Chemically competent Δ*btuR* BL21 (DE3) cells were co-transformed with

pDB1282 (Amp^R) encoding Fe-S cluster biosynthesis machinery, pBAD42-BtuCEDFB ($Spec^R$), encoding cobalamin uptake machinery, and the pRham N-His SUMO-*tsrM* construct (Kan^R). A list of all the strains used in this study can be found in **Table S1**.

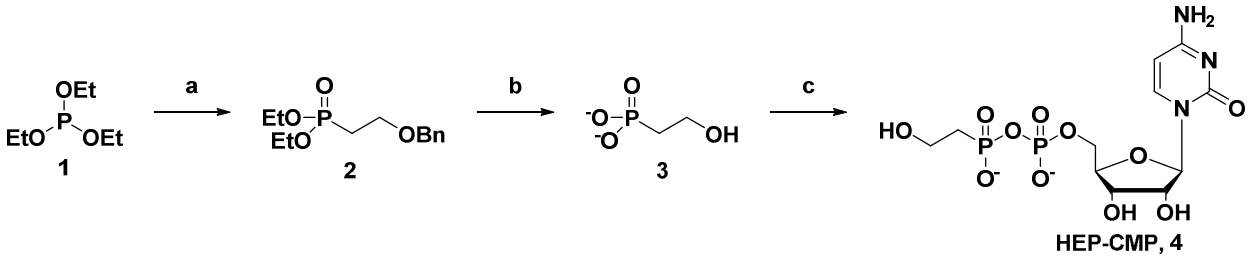
Overexpression of the tsrM Gene in the ΔbtuR BL21 Strain of E. coli. The Δ*btuR* BL21 strain of *E. coli* was co-transformed with pRham N-His SUMO-*tsrM* together with the pDB1282 and pBAD42-BtuCEDFB plasmids. Bacterial cultures were incubated at 37 °C in 16 L of LB media supplemented with 1.3 μM OHCbl and allowed to shake at 180 rpm. At an OD_{600} of ~0.3, expression of the genes encoded on the pDB1282 and pBAD42-BtuCEDFB plasmids was induced with 0.2% (w/v) L-arabinose, and the cultures were supplemented with 25 μM $FeCl_3$ and 75 μM cysteine. At an OD_{600} of ~0.6, the cultures were placed on ice for 20 min prior to induction of expression with 0.2% (w/v) L-rhamnose, and were supplemented with an additional 25 μM $FeCl_3$ and 75 μM cysteine. Cultures were incubated for 16 h at 20 °C prior to harvesting cells by centrifugation at 7,000 × *g* for 12 min.

TsrM and Fom3 Activity Determinations. All reactions were conducted under anaerobic conditions and contained the following in a final volume of 170 μL: 20-50 mM HEPES, pH 7.5, 10-100 mM KCl, 5% glycerol, 1-2 mM SAM, 1-2 mM substrate, 100 μM tyrosine, enzyme, and a reducing agent. In reactions containing TsrM (1 μM), the flavodoxin reducing system (25 μM flavodoxin, 10 μM flavodoxin reductase, and 1 mM NADPH) was used, while in reactions containing Fom3 (9 μM), methyl viologen (1 mM) and NADPH (2 mM) were used as reductant. All reactions were quenched by addition of H_2SO_4 to a final concentration of 100-200 mM. Standard curves for product quantification by LC-MS/MS contained 100 μM tyrosine and varying concentrations of substrate (HEP-CMP or Trp), SAH, SAM, and product (2-methyltryptophan or HPP-CMP).

Quantification of Product Formation by LC-MS/MS. Quantification of 5'-dA, SAH, and MeTrp was performed using methods similar to those previously described.[34, 41] Quenched reaction mixtures were separated on an Agilent Zorbax Extended-C18 RRHD column (2.1 mm × 50 mm, 1.8 μm particle size).

In the Fom3 reaction, HPP-CMP was separated using a Zorbax SB-Aq RRHD column (2.1 mm × 50 mm, 1.8 μm particle size) equilibrated in 99% solvent A (40 mM ammonium acetate, pH 6.0, 5% methanol) and 1% solvent B (methanol). A gradient of 1-5% solvent B was applied from 0.4 to 0.7 min, followed by a gradient from 5-60% solvent B from 0.7 to 1.4 min, which was followed by a gradient from 60%-1% solvent B from 1.4 to 2.1 min. The column was allowed to re-equilibrate for 0.5 min under initial conditions before subsequent sample injections. HPP-CMP was detected in the +1 charge state using AJSESI$^+$ with MRM, and the 446.1 → 112 and 446.1 → 94.9 transitions were used to quantify the molecule.

Synthesis of HEP-CMP.



Scheme 1. Steps involved in the synthesis of HEP-CMP.

Reactions and conditions: a) ((2-bromoethoxy)methyl)benzene, 170 ^oC, 2 h; b) 6 M HCl, 110 ^oC, overnight; c) cytidine monophosphate, CDI, TEA, DMF, then **3**, DMF, rt, 2 d.

step a:

The neat mixture of triethylphosphite (1.7 g, 10.2 mmol, 1 equiv) and ((2-bromoethoxy)methyl)benzene (2.2 g, 10.2 mmol, 1 equiv) was heated at 170 ^oC for 2 h. After cooling to room temperature, the reaction mixture was purified by silica gel flash chromatography (hexanes : ethyl acetate = 1:1) to give 2.0 g of compound **2** as a colorless syrup in 71% yield. ^1H NMR (500 MHz, $CDCl_3$) δ 7.38 – 7.27 (m, 5H), 4.53 (s, 2H), 4.14 – 4.04 (m, 4H), 3.76 (dt, J = 12.0, 7.4 Hz, 2H), 2.16 (dt, J = 18.6, 7.4 Hz, 2H), 1.31 (t, J = 7.1 Hz, 6H) (**Figure S1**); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 137.88, 128.42, 127.74, 73.03, 64.21, 61.68, 61.63, 27.63, 26.52, 16.42, 16.37 (**Figure S2**); ^{31}P NMR (202 MHz, $CDCl_3$) δ = 28.62 (**Figure S3**).

step b:

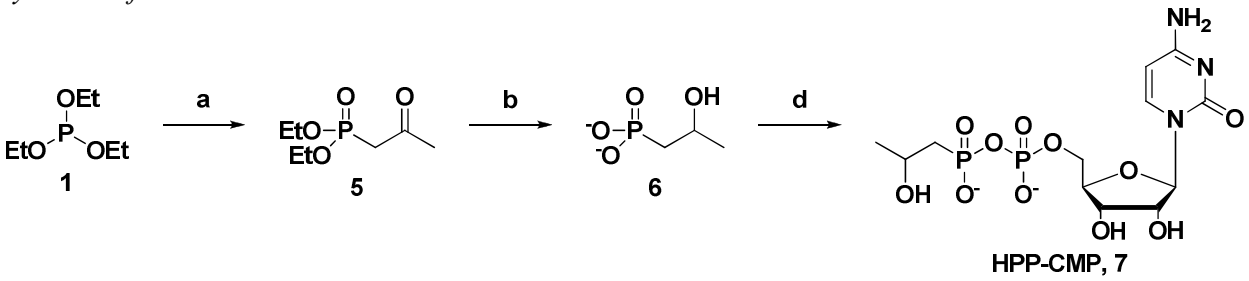
Compound **2** (1.0 g, 3.7 mmol, 1 equiv) was suspended in HCl (6 M, 30 mL), and the resulting mixture was refluxed at 110 ^{o}C overnight. After cooling to room temperature, the reaction mixture was washed with diethyl ether (30 mL), and the aqueous phase was dried *in vacuo*. The resulting residue was re-suspended in water (10 mL), and the pH was adjusted to 8 with triethylamine. The resulting solution was lyophilized to give 725 mg of compound **3** as white wax in 87% yield. Compound 3 was pure enough for further reaction, and no further purification was performed. ^{1}H NMR (500 MHz, D_2O) δ 3.78 – 3.66 (m, 2H), 3.09 (q, J = 7.3 Hz, 5H), 1.95 – 1.81 (m, 2H), 1.17 (t, J = 7.3 Hz, 8H) (**Figure S4**); ^{13}C NMR (126 MHz, D_2O) δ = 56.84, 46.61, 31.36, 30.33, 8.19 (**Figure S5**); ^{31}P NMR (202 MHz, D_2O) δ = 23.39 (**Figure S6**).

step c:

To a stirred suspension of cytidine 5'-monophosphate (CMP, 161.6 mg, 0.5 mmol, 1 equiv) in the co-solvent of DMF (10 mL) and triethylamine (TEA, 10 mL) was added carbonyldiimidazole (CDI, 243.2 mg, 1.5 mmol, 3 equiv) in one portion, and the resulting mixture was stirred at room temperature for 2 h. Another portion of CDI (81.1 mg, 0.5 mmol, 1 equiv) was added, and the reaction mixture was stirred until it became clear. Although clarity is an indication of reaction completeness, it was nonetheless analyzed by thin layer chromatography (TLC) using iPrOH:NH_4OH:H_2O = 6:3:1 as solvent. Additional portions of CDI were added if CMP was not completely consumed. After completion, the solvent was removed under nitrogen flow, and the resulting syrup was re-dissolved in methanol and water (1:1) containing 5% triethylamine. The solvent was dried *in vacuo* completely, and the resulting gum was re-suspended in DMF (10 mL) followed by addition of compound **3** (62.0 mg, 0.5 mmol, 1 equiv). The resulting reaction mixture was stirred vigorously at room temperature for 2 days. After completion, the reaction was quenched with water (30 mL), and the resulting mixture was extracted with EtOAc (20 mL) ten times to remove the DMF from the aqueous phase. The aqueous phase containing the product was purified over a Macro-Prep DEAE column with a linear elution of aqueous ammonium bicarbonate solution (from 10 mM to 300 mM). Fractions containing product were combined and concentrated *in vacuo*. Repeated evaporation on a rotary evaporator in a water bath at 50 ^{o}C was performed to remove the

residual ammonium bicarbonate to give 110.0 mg of HEP-CMP as a white powder in 51% yield. ^{1}H NMR (500 MHz, D_2O) δ 7.99 (d, *J* = 7.7 Hz, 1H), 6.14 (br s, 1H), 5.90 (d, *J* = 3.8 Hz, 1H), 4.31 – 4.23 (m, 2H), 4.23 – 4.16 (m, 2H), 4.14 – 4.08 (m, 1H), 3.76 (q, *J* = 16.1, 8.5 Hz, 2H), 2.08 – 1.96 (m, 2H) (**Figure S7**); ^{13}C NMR (126 MHz, D_2O) δ = 163.41, 154.06, 142.45, 96.17, 89.40, 82.90, 82.83, 74.25, 69.12, 64.33, 64.30, 57.07, 32.07, 31.01; ^{31}P NMR (202 MHz, D_2O) δ = 14.39, -11.34 (**Figure S8**); ^{31}P NMR (202 MHz, $CDCl_3$) δ = 28.62 (**Figure S9**).

Synthesis of HPP-CMP.



Scheme 2. Steps involved in the synthesis of HPP-CMP.

Reactions and conditions: a) chloroacetone, NaI, acetonitrile, rt, overnight; b) $NaBH_4$, EtOH, 0 ^{o}C, 1 h; then 6 M HCl, 110 ^{o}C, overnight; c) cytidine monophosphate, CDI, TEA, DMF, then **6**, DMF, rt, 2 d.

step a:

To a stirred solution of triethylphosphite (1.7 g, 10.2 mmol, 1 equiv) and sodium iodide (152.9 mg, 1.0 mmol, 0.1 equiv) in acetonitrile (100 mL) was added chloroacetone (4.1 mL, 51.0 mmol, 5 equiv), and the resulting brown reaction mixture was stirred at room temperature overnight. After quenching by addition of water (100 mL), the reaction mixture was extracted with EtOAc (200 mL), and the organic phase was washed thoroughly with saturated sodium thiosulfate, water and brine. The organic phase was concentrated *in vacuo*, and the resulting residue was purified by silica gel flash chromatography (hexanes : ethyl acetate = 1:1) to give 1.0 g of compound **5** as a yellowish syrup in 51% yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.17 – 4.05 (m, 4H), 3.05 (d, *J* = 22.8 Hz, 2H), 2.28 (s, 3H), 1.30 (t, *J* = 7.1 Hz, 6H) (**Figure S10**); ^{13}C NMR (126 MHz, $CDCl_3$) δ = 199.89, 199.84, 62.54, 62.49, 43.81, 42.80, 31.31, 16.27, 16.22 (**Figure S11**); ^{31}P NMR (202 MHz, $CDCl_3$) δ = 19.64 (**Figure S12**).

step b:

To an ice-water cooled solution of compound 5 (971.0 mg, 5.0 mmol, 1 equiv) in methanol (25 mL) was added sodium borohydride (189.2 mg, 5.0 mmol, 1 equiv). The reaction was stirred for 1 h and then quenched by addition of water (50 mL). The resulting mixture was extracted with EtOAc (100 mL), and the organic layer was washed thoroughly with water and brine. The organic phase was concentrated *in vacuo*, and the resulting residue was used for the next step without further purification.

The product from the previous step was suspended in HCl (6 M, 30 mL), and the resulting mixture was refluxed at 110 ^{o}C overnight. After cooling to room temperature, the reaction mixture was washed with diethyl ether (30 mL), and the aqueous phase was dried *in vacuo*. The resulting residue was re-suspended in water (10 mL), and the pH was adjusted to 8 with triethylamine. The resulting solution was lyophilized to give 420 mg of compound **6** as white wax in 61% yield. Compound **6** was pure enough for further reaction, and no further purification was performed. ^{1}H NMR (500 MHz, D_2O) δ 4.09 – 3.93 (m, 1H), 3.07 (q, J = 7.3 Hz, 4H), 1.95 – 1.68 (m, 2H), 1.18 – 1.13 (m, 9H) (**Figure S13**); ^{13}C NMR (126 MHz, D_2O) δ = 63.82, 46.60, 37.28, 36.23, 23.23, 23.15, 8.20 (**Figure S14**); ^{31}P NMR (202 MHz, D_2O) δ = 24.08 (**Figure S15**).

step c:

To a stirred suspension of cytidine 5'-monophosphate (CMP, 161.6 mg, 0.5 mmol, 1 equiv) in the co-solvent of DMF (10 mL) and triethylamine (TEA, 10 mL) was added carbonyldiimidazole (CDI, 243.2 mg, 1.5 mmol, 3 equiv) in one portion, and the resulting mixture was stirred at room temperature for 2 h. Another portion of CDI (81.1 mg, 0.5 mmol, 1 equiv) was added, and the reaction mixture was stirred until it became clear. Although clarity is an indication of reaction completeness, the reaction was analyzed by thin layer chromatography (TLC) using iPrOH:NH_4OH:H_2O = 6:3:1 as solvent. Another potion of CDI was added if CMP was not consumed completely. After completion, the solvent was removed under nitrogen flow, and the resulting syrup was re-dissolved in methanol and water (1:1) containing 5% triethylamine. The solvent was dried *in vacuo* completely, and the resulting gum was re-suspended in DMF (10 mL) followed by addition of compound **6** (69.0 mg, 0.5 mmol, 1 equiv). The

resulting reaction mixture was stirred vigorously at room temperature for 2 days. After completion, the reaction was quenched by water (30 mL), and the resulting mixture was extracted with EtOAc (20 mL) ten times to remove the DMF from the aqueous phase. The aqueous phase containing the product was purified over a Macro-Prep DEAE column with a linear elution of aqueous ammonium bicarbonate solution (from 10 mM to 300 mM). Fractions containing product were combined and concentrated *in vacuo*. Repeated evaporation on a rotary evaporator in a water bath at 50 ^{o}C was performed to remove the residual ammonium bicarbonate to give 88.0 mg of HPP-CMP as a white powder in 40% yield. ^{1}H NMR (500 MHz, D_2O) δ 8.14 (d, J = 8.0 Hz, 1H), 6.23 (d, J = 8.0 Hz, 1H), 5.86 (d, J = 2.8 Hz, 1H), 4.31 – 4.18 (m, 4H), 4.15 – 4.04 (m, 2H), 2.07 – 1.87 (m, 2H), 1.21 (d, J = 6.2 Hz, 3H) (**Figure S16**); ^{13}C NMR (126 MHz, D_2O) δ = 159.08, 148.41, 144.12, 95.21, 89.57, 83.25, 83.18, 74.31, 68.94, 64.16, 63.89, 37.90, 36.83, 23.10, 23.03 (**Figure S17**); ^{31}P NMR (202 MHz, D_2O) δ = 15.50, -11.47 (**Figure S18**).

RESULTS

Strategy for Overproduction of Class B RS Methylases in Soluble Form. In our previous studies of TsrM, the *tsrM* gene, cloned into a pSUMO vector, was expressed alongside genes on plasmid pDB1282, which encode proteins involved in Fe-S cluster biosynthesis and trafficking.[44] The chosen pSUMO vector confers kanamycin resistance, and expression of the *tsrM* gene is controlled by an IPTG-inducible T7 promoter. By contrast, plasmid pDB1282 confers ampicillin resistance, and expression of the *isc* gene cluster on this plasmid is controlled by an arabinose-inducible (P_{BAD}) promoter. During gene expression, *E. coli* was cultured in a specialized medium supplemented with ethanolamine as the sole carbon and nitrogen source, as well as OHCbl. Ethanolamine is metabolized in *E. coli* to acetaldehyde and ammonia by ethanolamine ammonia lyase (EAL), an AdoCbl-dependent enzyme.[47] Although *E. coli* cannot synthesize AdoCbl *de novo*, it contains the genes necessary to generate it from cobinamide, an incomplete corrinoid that lacks the lower axial nucleotide loop (**Figure 2d**). When cultured on ethanolamine as the sole source of carbon and nitrogen, *E. coli* upregulates the expression of EAL, which has been suggested to enhance the intracellular levels of cobalamin.[33] When TsrM is overproduced in this fashion, pure

protein is obtained in reasonable (0.5 mg L^{-1})—but still challenging—yields. Whereas, when TsrM is overproduced in media lacking ethanolamine, only meager amounts of protein are obtained, and its homogeneity is 50% or less. Consistent with cobalamin accessibility being one of the limiting factors in obtaining soluble TsrM, the as-isolated protein contains a stoichiometric amount of bound cobalamin.[34]

In an effort to improve upon the low yields of soluble TsrM, we devised a strategy to increase the cellular concentrations of cobalamin, with the expectation that the immediate availability of the cofactor might facilitate proper intracellular folding during and/or after protein synthesis. The *E. coli* genome encodes two other cobalamin-dependent proteins in addition to EAL, but the organism does not require the cofactor for growth. *E. coli* preferentially uses a MeCbl-dependent methionine synthase (MetH) to generate methionine from homocysteine,[48] but can also catalyze the reaction in the absence of cobalamin using a cobalamin-independent methionine synthase.[49] It also contains QueG, which uses cobalamin to catalyze the final step in the biosynthesis of queuosine, a modified nucleoside found in the anticodon stem loop of some tRNAs.[50, 51]

Cobalamin uptake in *E. coli* is a regulated process that minimally involves four dedicated gene products: BtuC, BtuD, BtuF, and BtuB (**Figure 3**).[52] BtuB is a high-affinity cobalamin transporter located in the outer membrane of *E. coli*, which acts in concert with TonB to transport various cobalamin derivatives into the periplasmic space of the bacterium in an energy-dependent process.[53-55] Transport across the cytoplasmic membrane is mediated by BtuC and BtuD, which form a membrane-bound $\alpha_2\beta_2$ complex that functions as an ABC transporter.[56-59] Lastly, BtuF is a periplasmic cobalamin-binding protein that delivers cobalamin to the BtuCD complex.[60] Regulation of the cobalamin uptake system is mediated by AdoCbl-dependent feedback inhibition of *btuB* gene expression. When intracellular concentrations of AdoCbl are sufficiently high, AdoCbl binds to *btuB* messenger RNA (mRNA), including a 5′-leader sequence, inhibiting translation by preventing mRNA binding to the ribosome.[61]

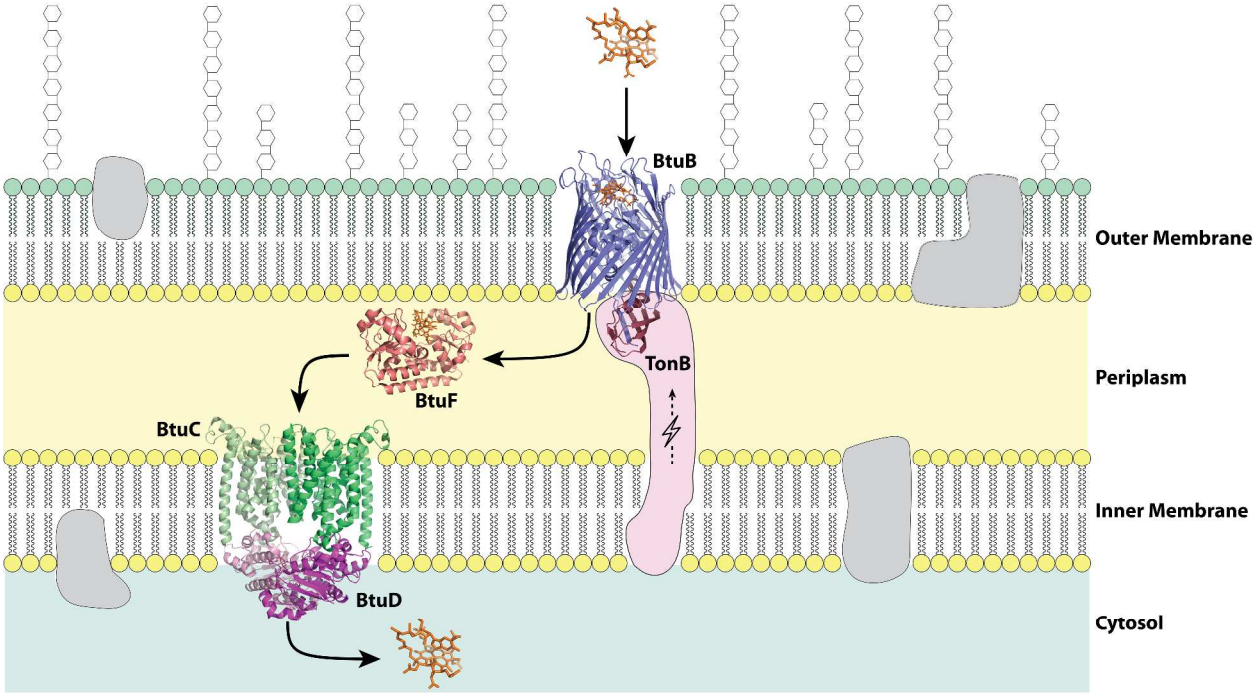


Figure 3. Cobalamin uptake system in *E. coli.* Cobalamin is transported across the outer membrane by the TonB-dependent outer membrane protein BtuB. BtuF, a periplasmic membrane transporter, then delivers the cobalamin to BtuC and BtuD, which subsequently transport cobalamin across the inner membrane into the cytosol.

We designed a plasmid that incorporates the *btuC*, *btuE*, *btuD*, *btuF* and *btuB* genes to allow for their overexpression. The *btuC* and *btuD* genes are part of a *btuCED* operon in *E. coli*, although the *btuE* gene product does not play a role in cobalamin uptake and trafficking. In fact, it is predicted to be a peroxidase based on its primary structure.[62] The *btuB* and *btuF* genes are found elsewhere in the *E. coli* chromosome. To facilitate cloning, the *btuE* gene was not excised from the *btuCED* operon. The pBAD42 plasmid was chosen as a compatible vector for use with the expression system that we typically employ to overproduce RS enzymes. This plasmid encodes resistance to spectinomycin, which allows it to be maintained in the same host with the kanamycin resistance-encoding pET-based vectors that we use for gene expression of RS enzymes, the ampicillin resistance-encoding pDB1282 vector for expression of the *isc* operon, and the chloramphenicol resistance-encoding plasmids that are present in some expression strains for the

expression of low-copy tRNAs. The multiple cloning site of pBAD42 is downstream of the P_{BAD} promoter, which is also the relevant promoter in the pDB1282 plasmid. This strategy allows for temporal control of expression of the cobalamin uptake genes with respect to expression of the RS gene of interest. Typically, expression of the *isc* operon is induced at an OD_{600} of 0.3 by addition of arabinose, while expression of the RS gene is induced by addition of IPTG at an OD_{600} of approximately 0.6, such that the Fe-S cluster assembly proteins are present prior to expression of the target RS gene.[44] Likewise, the cobalamin uptake genes are under the control of the same P_{BAD} promoter and are induced simultaneously with the genes in the *isc* operon.

The *btuCED*, *btuF*, and *btuB* genes were cloned from *E. coli* genomic DNA and ligated into the pBAD42 vector sequentially using semi-nested restriction sites (**Figure 4**). To improve translation of the *btuCEDFB* mRNAs, the spacing between genes and their respective ribosome binding sites (RBS) was further optimized. The interstitial sequences between *btuD* and *btuF* and between *btuF* and *btuB* were removed by site-directed mutagenesis such that the stop codon of the upstream gene and the start codon of the downstream gene overlap by one base pair. Because the *btuCED* fragment was cloned as an operon, the spacing between these genes was not altered. The *btuE* and *btuD* genes overlap by one base pair in the *E. coli* genome; however, there is a >60 base-pair intergenic region between *btuC* and *btuE*, necessitating an additional RBS. Neither the RBS encoded by the vector nor the RBS encoded in the intergenic region between *btuC* and *btuE* were optimal for translation in *E. coli*. Therefore, the sequence of each RBS was changed by site-directed mutagenesis to 5'- TAAGGAGGT $-X_5$-ATG-3', where the last three bases are the start codon of the gene. The final optimized plasmid was termed pBAD42-BtuCEDFB.

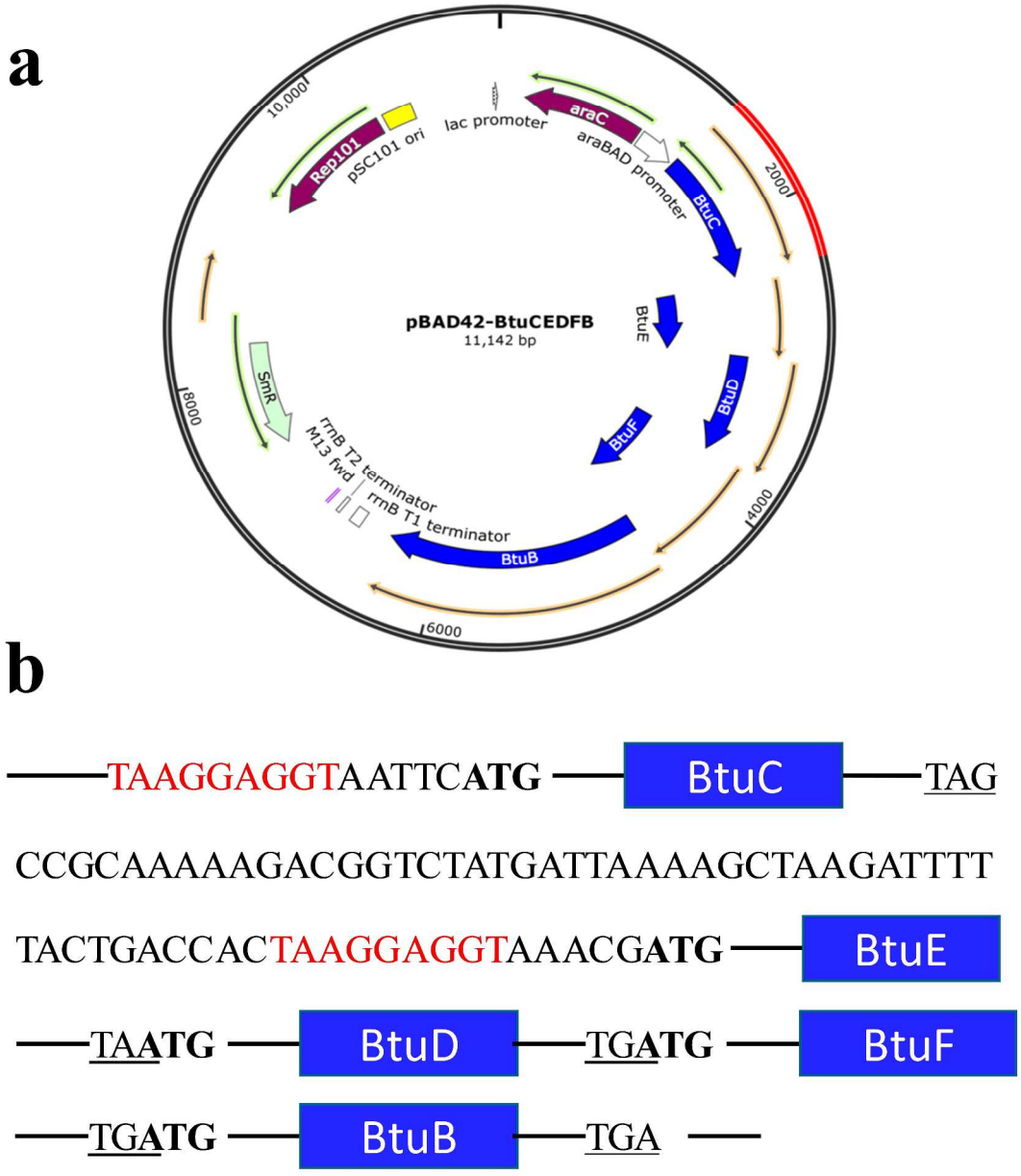


Figure 4. The pBAD42-BtuCEDFB vector. (a) A plasmid map of the vector. Expression of *btuCEDFB* is driven by an arabinose inducible araBAD promoter. (b) The cloning regions of the *btuCEDFB* operon. The optimized ribosome binding sites are shown in red, while the start (bold) and stop (underlined) codons are also displayed.

Analysis of Cobalamin Content in E. coli by LC-MS/MS. LC-MS/MS analysis was used to compare the concentration of cobalamin in *E. coli* harboring the pBAD42-BtuCEDFB plasmid to the concentration of cobalamin in *E. coli* containing an empty pBAD42 vector. Gene expression was carried out in *E. coli* cultured in M9 minimal media supplemented with OHCbl, LB media supplemented with OHCbl, or the M9-ethanolamine media previously used to improve cobalamin uptake. Because OHCbl is converted into MeCbl or AdoCbl *in vivo,*[63] it was necessary to determine not only if induction of the *btuCEDFB* operon increases the cobalamin content of *E. coli* cells, but also to determine what forms of cobalamin are present in the cell. In **Figure 5**, relative intracellular concentrations of OHCbl, AdoCbl, and MeCbl are displayed for *E. coli* cultured in different growth media and in the absence or presence of pBAD42-BtuCEDFB. In M9 minimal media and in LB media, induction of the *btuCEDFB* operon results in a 13- and 19-fold increase in total (sum of all forms) cobalamin concentration, respectively, as compared to the vector-only control, while a 54-fold increase is observed for *E. coli* cultured in M9-ethanolamine media. In addition to their increased cobalamin content detected by LC-MS/MS, cell pellets of *E. coli* containing the pBAD42-BtuCEDFB plasmid are noticeably pinker in color (**Figure S19**). The substantial enhancement in cobalamin uptake indicates that overexpression of the operon elicits the intended outcome, increasing the intracellular cobalamin concentration for enhanced cobalamin incorporation into TsrM or other cobalamin-containing proteins.

All three forms of cobalamin are detected in *E. coli* extracts (**Figure 5**); however, OHCbl and AdoCbl predominate. By contrast, relatively small amounts of MeCbl are observed under each of the conditions. When the expression of the *btuCEDFB* operon is induced in *E. coli* cultured in M9 minimal media, approximately 52% of the cobalamin is found as OHCbl, while <1% and 47% is found as MeCbl and AdoCbl, respectively. However, when expression is induced in *E. coli* cultured in M9-ethanolamine media, 63% of the cobalamin is found as AdoCbl, and <1% and 37% are found as MeCbl and OHCbl, respectively. Similarly, 19% of total cobalamin is found as OHCbl, 81% as AdoCbl, and <1% as MeCbl when the *btuCEDFB* operon is induced in *E. coli* cultured in LB media (**Figure 5**). These results clearly

indicate that *E. coli* cells harboring the pBAD-BtuCEDFB plasmid uptake significantly more cobalamin than those that do not have the plasmid and that OHCbl and AdoCbl are the major forms of cobalamin found in the cell.

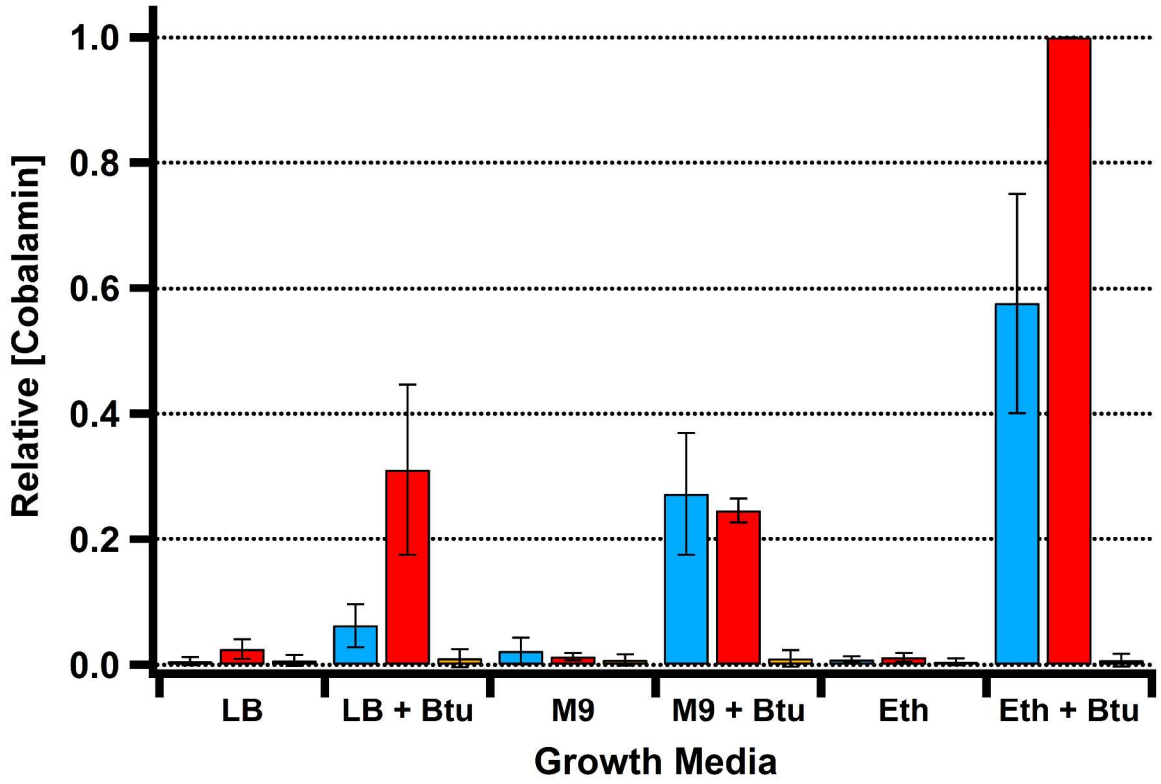


Figure 5. Relative concentrations of MeCbl (orange), OHCbl (blue), and AdoCbl (red) in *E. coli* cultured in LB, M9, and M9-ethanolamine (Eth) media with and without pBAD42-BtuCEDFB normalized to the amount of AdoCbl in *E. coli* cultured in M9-ethanolamine media with the pBAD42-BtuCEDFB plasmid.

pBAD42-BtuCEDFB-Dependent Enhancement of TsrM Solubility. TsrM was purified to homogeneity from soluble crude lysate in a previous study, but the strategy employed yielded meager amounts of protein (~0.5 mg/L) that was only soluble when overproduced in the RosettaBlue (DE3) pLysS strain of *E. coli* and when cultured in M9-ethanolamine media supplemented with cobalamin.[34, 41] When a similar strategy was applied to TsrM overproduced in the *E. coli* BL21 (DE3) strain rather than the RosettaBlue (DE3) strain, only ~0.5 mg/L soluble protein that was approximately 25% pure was obtained (**Figure**

S20). To determine if additional cobalamin in the cell could increase the amount of soluble TsrM, TsrM was overproduced as a fusion protein with SUMO as previously described,[34] but under conditions in which proteins encoded by the pBAD42-BtuCEDFB plasmid are present. TsrM was subsequently purified as previously described, and a yield of ~3.7 mg/L of >95% homogenous protein was collected upon removal of the SUMO tag using Ulp1 protease (**Figure S21**). Thus, co-expression of *tsrM* with the genes on pBAD42-BtuCEDFB in BL21 (DE3) yields a 30-fold increase in soluble protein as compared to expression in the same strain without induction of the *btuCEDFB* operon, and more than a 7-fold increase in yield as compared to expression in the RosettaBlue (DE3) strain.

pBAD42-BtuCEDFB-Dependent Enhancement of Fom3 Solubility. To further test the efficacy of the pBAD42-BtuCEDFB construct in enhancing class B RS methylase solubility, we set out to purify Fom3 from *Streptomyces wedmorensis*, which has previously been reported to be insoluble upon overproduction in *E. coli*.[64] In a previous study, this protein was purified from inclusion bodies and resolubilized and refolded before being reconstituted with its Fe-S and cobalamin cofactors.[64] As a control, a SUMO-tagged Fom3 was first overproduced in the absence of the pBAD42-BtuCEDFB plasmid; however, as has been previously reported, no soluble protein was obtained from the crude lysate. However, when Fom3 was overproduced in the presence of pBAD42-BtuCEDFB, it could be purified to ≥95% homogeneity using the same protocol as that described for TsrM (**Figure S22**).[34] This strategy yields 3.0 mg/L of soluble Fom3 as determined by Bradford analysis.

Forms of Cobalamin Bound to TsrM and Fom3. After purification of Fom3 and TsrM, each protein was treated with potassium cyanide to form dicyanocobalamin to determine the molar ratio of bound cobalamin to protein. Using the extinction coefficient of dicyanocobalamin (ε_{367} = 30,800 M^{-1} cm^{-1}),[40] it was determined that TsrM and Fom3 contain 1.03 and 0.86 equivalents of cobalamin per polypeptide, respectively (**Figure S23**). Additionally, the relative ratios of the different forms of cobalamin present in TsrM and in Fom3 were determined by LC-MS/MS analysis. Fom3 contains 0.85 mol of MeCbl, 0.15

mol of OHCbl, and 0.00 mol of AdoCbl per polypeptide (**Figure 6a**). Consistent with previous studies,[37] TsrM contains 0.00 mol of MeCbl, 0.63 mol of OHCbl, and 0.40 mol of AdoCbl per polypeptide (**Figure 6b**). The lack of MeCbl in TsrM may derive from the ready availability of tryptophan, its substrate, during gene expression, while the substrate for Fom3 is not expected to be present in *E. coli*. In line with this argument, when TsrM is overproduced in *E. coli*, 2-methyltryptophan (MeTrp) is observed in cell extracts.

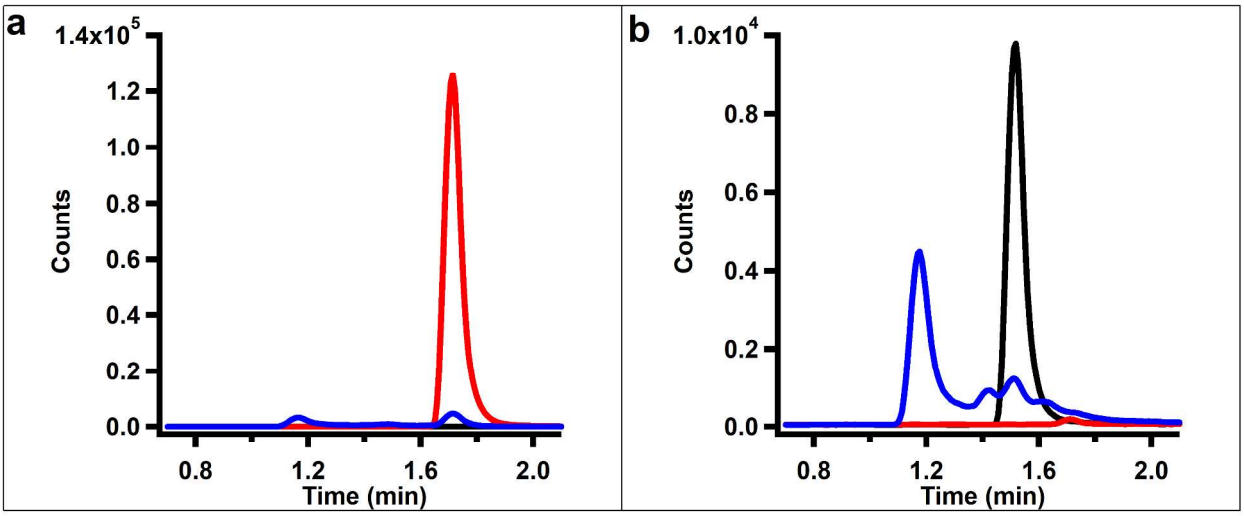


Figure 6. LC-MS/MS analysis of the cobalamin content of (a) as-isolated Fom3 (20 μM) and (b) as-isolated TsrM. (a) As-isolated Fom3 contains 18% OHCbl (blue line), 0% AdoCbl (black line) and 82% MeCbl (red line). (b) TsrM contains 61% OHCbl (blue line), 39% AdoCbl (black line) and <1% MeCbl (red line).

Activity of Fom3. Fom3 is one of the enzymes in a microbial biosynthetic pathway for production of fosfomycin from phosphoenolpyruvate. It was originally predicted to catalyze formation of 2-hydroxypropylphosphonate (2-HPP) from 2-hydroxyethylphosphonate (2-HEP), and indeed, one laboratory reported this activity with Fom3 purified from inclusion bodies.[9, 11, 64] However, when we incubate Fom3 (purified from the soluble fraction of *E. coli* lysate) with 2-HEP, SAM, reductant (flavodoxin reducing system, reduced methyl viologen, or dithionite), formation of neither 2-HPP nor *S*-

adenosylhomocysteine (SAH) is detected by LC-MS/MS. Recently, Sato et al. found that Fom3 actually catalyzes the C-methylation of cytidylyl-2-hydroxyethylphosphonate (HEP-CMP) to give cytidylyl-2-hydroxypropylphosphonate (HPP-CMP) (**Figure S24**).[65] When HEP-CMP is used in our Fom3 activity determinations, formation of HPP-CMP, SAH, and 5′-deoxyadenosine (5′-dA) are readily observed (**Figure 7**). Sato et al. also reported that the steady-state kinetic analysis of Fom3 required the addition of MeCbl to their reaction mixtures and suggested that MeCbl is a co-substrate in the reaction rather than a cofactor, meaning that it is regenerated outside of the Fom3 active site. As shown in **Figure 7**, Fom3, overproduced and isolated as described in this study, exhibits a k_{cat} of 1.81 h^{-1} in the absence of added MeCbl, which is 2.03 times greater than the turnover number reported by Sato et al (0.89 h^{-1}).[65]

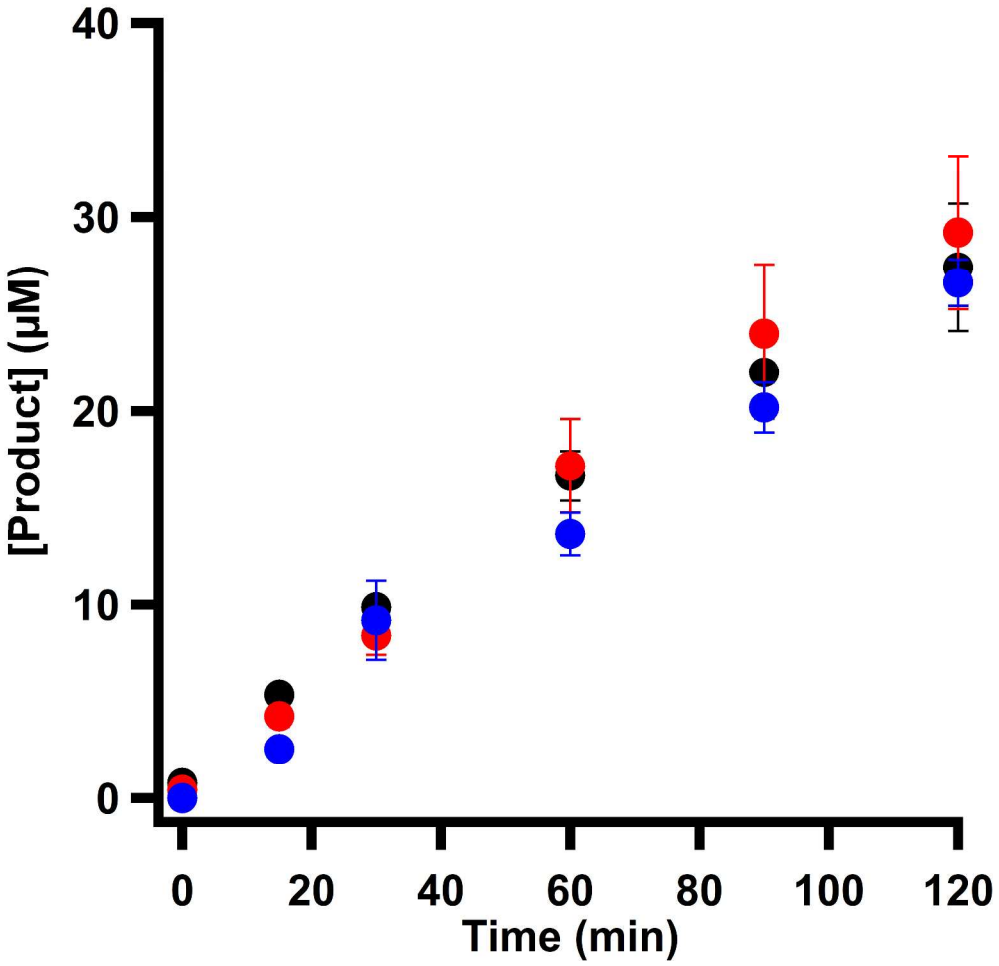


Figure 7. Time-dependent formation of HPP-CMP (black), SAH (red), and 5'-dA (blue) by Fom3 (9 μM) in the presence of 1.5 mM HEP-CMP, 1.5 mM SAM, 1 mM methyl viologen and 2 mM NADPH.

Purification of PhpK from Soluble Crude Lysate. Given the success with Fom3, we also set out to overproduce soluble SUMO-tagged PhpK and purify it. Although previous attempts to purify PhpK from soluble crude lysate were unsuccessful,[66] we were able to purify small amounts (~0.22 mg/L) of SUMO-tagged PhpK when overproduction of the fusion protein was conducted in *E. coli* cultured in M9-ethanolamine media. However, the Ulp1 protease is unable to excise the SUMO protein from the SUMO-PhpK construct unless a glycine residue is inserted between SUMO and PhpK, which is suggested by the SUMO cloning kit manual for constructs that are difficult to cleave. In this instance, the addition of the pBAD42-BtuCEDFB plasmid did not result in a higher yield of PhpK, even when it was overproduced in *E. coli* cultured in M9-ethanolamine media. In an effort to increase the yield of PhpK, overproduction of the protein was performed under autoinduction conditions, which uses a specialized media supplemented with glucose, glycerol, and lactose. Under these conditions, expression of the *phpK* gene is slowly induced by lactose when the glucose in the media is depleted.[43] Using this strategy, overexpression of the *phpK* gene alongside genes on the pBAD42-BtuCEDFB plasmid results in a four-fold higher yield of soluble PhpK (0.84 mg/L) as compared to overproduction in *E. coli* cultured in M9-ethanolamine media under the same conditions. However, no soluble protein is recovered when overproduction of PhpK is conducted in the absence of pBAD42-BtuCEDFB in autoinduction media. An SDS-PAGE analysis reveals native PhpK to be ~95% pure following excision of SUMO from the fusion construct (**Figure S25**). Cobalamin analysis of purified PhpK reveals that the enzyme contains 0.85 equivalents of the cofactor per polypeptide (**Figure 8a**). LC-MS/MS analysis of the cobalamin content of as-isolated PhpK shows the enzyme to contain 8%, 15%, and 77% AdoCbl, OHCbl, and MeCbl, respectively (**Figure 8b**).

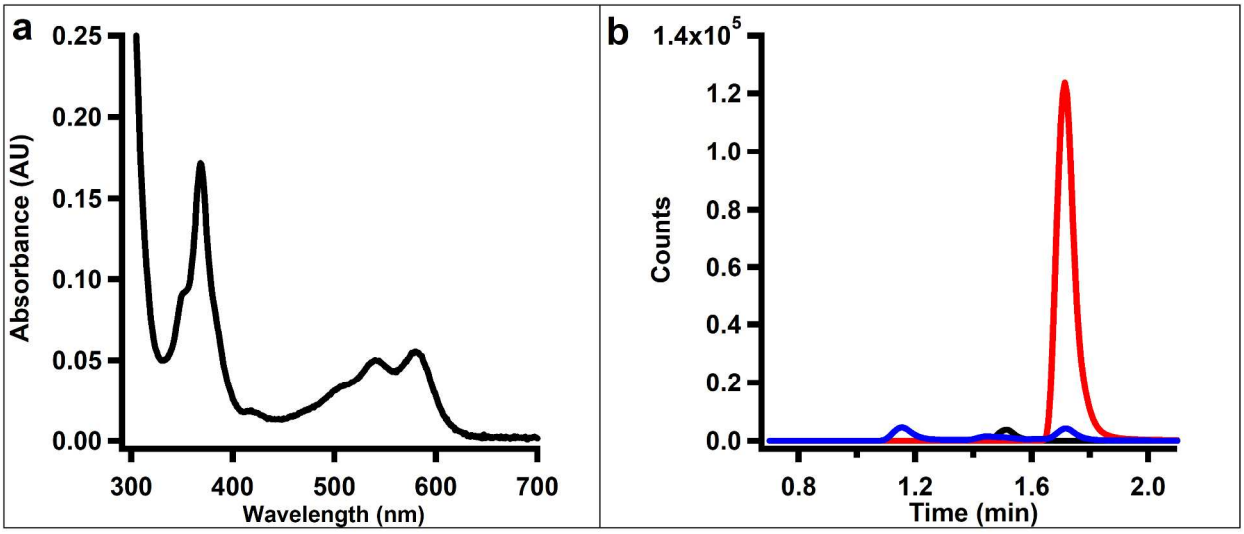


Figure 8. (a) UV-Vis spectrum of PhpK (6.5 μM) treated with potassium cyanide. (b) LC-MS/MS analysis of as-isolated PhpK, which contains 8% AdoCbl (black), 15% OHCbl (blue), and 77% MeCbl (red).

PhpK is found in a biosynthetic pathway for the production of L-phosphinothricin-L-Ala-L-Leu (phosalacine), an herbicidal antibiotic tripeptide produced by *K. phosalacinea*.[67] The enzyme was first purified from inclusion bodies and refolded and reconstituted with its Fe-S cluster and cobalamin cofactors. It was purported to catalyze the conversion of 2-acetylamino-4-hydroxyphosphinylbutanoate (NAcDMPT) to 2-acetylamino-4-hydroxymethylphosphinylbutanoate (NAcPT) by a single time point NMR assay (**Figure S24**); however, the formation of 5′-dA or SAH was not reported. After purifying soluble PhpK and determining its cofactor content, we assessed its ability to catalyze formation of the expected methylated product from NAcDMPT, NAcDMPT-L-Ala-L-Ala and NAcDMPT-L-Ala-L-Leu using either the flavodoxin reducing system, dithionite, or reduced methyl viologen as reductants. However, no product formation was detected, suggesting—as was observed for Fom3—that the identity of the *in vivo* substrate has not yet been elucidated.

Overproduction and Purification of TsrM from a ΔbtuR strain of E. coli. As detailed above, when TsrM is purified from *E. coli*, most of its bound cobalamin exists as AdoCbl, which affords an inactive state of the enzyme. In *E. coli*, formation of AdoCbl from hydroxocobalamin and ATP is catalyzed by BtuR in the presence of a suitable reductant. Therefore, to prevent formation of AdoCbl in our expression host, we generated a BL21(DE3) strain of *E. coli*, in which its *btuR* gene is replaced with a kanamycin cassette. This knock-out was constructed by moving the kanamycin cassette from *E. coli* BW25113 (Δ*btuR*:*kan*)—obtained from the Keio collection—into *E. coli* BL21(DE3) via P1 transduction. This mutant strain was transformed with pSUMO-TsrM; however, overproduction of TsrM was not observed. Moreover, the overproduction of other non-cobalamin-containing proteins was also unsuccessful in this strain. Tests for the DE3 lysogen showed that it is present in the BW25113 (Δ*btuR*:*kan*) strain, suggesting that the absence of the *btuR* gene in this strain has an effect on the T7 expression system. Therefore, an alternative expression strategy was adopted, wherein a gene encoding a SUMO-TsrM fusion protein was

cloned into a pRham-SUMO plasmid under the control of a rhamnose-inducible rhaP$_{BAD}$ promoter rather than an IPTG-inducible T7 promoter. Overexpression of the *SUMO-tsrM* fusion encoded on pRham-SUMO-TsrM was conducted in the Δ*btuR* strain of *E. coli* along with the pDB1282 and pBAD42-BtuCEDFB plasmids. Because AdoCbl is required for cell growth in M9 ethanolamine media, we chose to overproduce TsrM in LB media supplemented with cobalamin. Induction with 0.2% L-rhamnose results in high levels of expression of the *SUMO-tsrM* fusion (**Figure S26**). Purification of the SUMO-TsrM fusion protein was performed as previously described, which results in native TsrM (~3.5 mg/L) that is ~95% pure based on SDS-PAGE analysis.

Analysis of TsrM Overproduced in E. coli BW25113 ΔbtuR. After purifying TsrM to homogeneity, the stoichiometry of bound cobalamin is determined to be 1.1 per polypeptide (**Figure S27**). LC-MS/MS analysis was then used to identify the forms of cobalamin present in the purified enzyme, and osbserved no AdoCbl in TsrM that is overproduced in the Δ*btuR* BW25113 strain of *E. coli* (**Figure 9a**). Moreover, overproduction of TsrM in the Δ*btuR* strain of *E. coli* does not increase the proportion of MeCbl present in the enzyme. Instead, TsrM contains more OHCbl, cob(I)alamin, and cob(II)alamin. Because OHCbl is detected by MS/MS without its hydroxy group, cob(I)alamin and cob(II)alamin have the same *m/*z, but have retention times of 1.5 min and 1.7 min, respectively. Because we do not have standards for cob(I)alamin and cob(II)alamin, however, we cannot accurately quantify these peaks by LC-MS/MS.

The activity of TsrM, which catalyzes the methylation of Trp to form MeTrp (**Figure S24**), was determined by monitoring the time-dependent production of MeTrp by LC-MS/MS as previously described.[34, 37] Analysis of TsrM activity was conducted in a reaction containing SAM (1 mM), tryptophan (1 mM), and TsrM (1 μM) using the flavodoxin/flavodoxin reductase/NADPH reducing system. At saturating conditions of tryptophan and SAM, as-isolated TsrM lacking AdoCbl exhibits a k_{cat} of 17.1 min^{-1} (**Figure 9b**), which is greater than the previously reported value of 10 min^{-1} for reconstituted TsrM under the same conditions. The increased activity of TsrM lacking AdoCbl relative to the enzyme that contains a mixture of AdoCbl and OHCbl suggests that AdoCbl is not required for TsrM activity—and is in fact an inhibitor—and that TsrM adventitiously binds AdoCbl when overproduced in *E. coli.*

Under steady-state turnover conditions, TsrM lacking AdoCbl maintains approximately 80% of its cobalamin as MeCbl (**Figure S28**), which is significantly greater than the 29% observed during previous studies of TsrM that contained AdoCbl.[37]

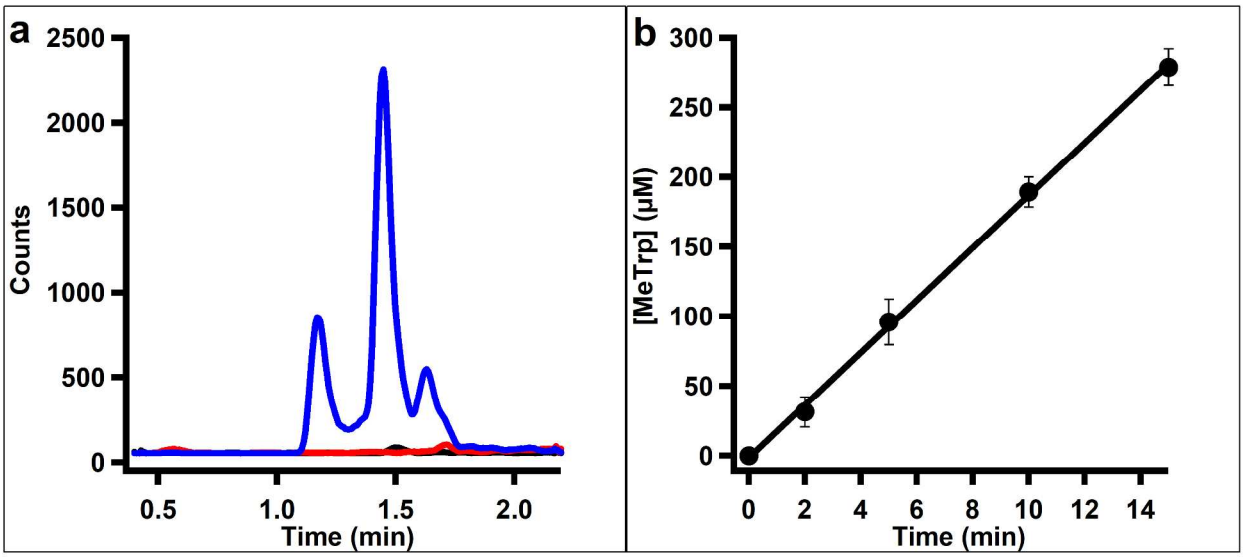


Figure 9. (a) LC-MS/MS analysis of TsrM overexpressed in the Δ*btuR* strain of *E. coli* BW25113 reveals no bound AdoCbl (black). (b) Time-dependent formation of MeTrp by TsrM (0.1 μM) in the presence of SAM (1 mM) and Trp (1 mM).

pBAD42-BtuCEDFB-Dependent Enhancement of ThnK Solubility. To determine if co-expression with pBAD42-BtuCEDFB results in higher yields of soluble ThnK, which has previously been purified from soluble crude lysate in relatively low yields when overproduced in the Rosetta 2(DE3) strain of *E. coli*,[35] we overexpressed ThnK in *E. coli* BL21 (DE3) with and without pBAD42-BtuCEDFB. Using purification methods similar to those described above, we do not recover any soluble ThnK when its gene is expressed in the absence of pBAD42-BtuCEDFB. However, we isolate ~300 mg of ThnK that is ~85% pure from 16 L of culture when its gene is co-expressed in the presence of the pBAD42-BtuCEDFB plasmid. Because the expression plasmid containing the *thnK* gene does not yield a SUMO-fusion protein, this batch of protein was further purified by size-exclusion chromatography using an S200

column, resulting in 113 mg of protein (7.1 mg/L) that is ~95% pure (**Figure S29**). The significant loss of protein following size-exclusion chromatography in this instance is due to conservative pooling of fractions in order assure a homogeneous final product.

Discussion

The poor solubility of some cobalamin-dependent RS enzymes upon their overproduction in *E. coli,* and the slight enhancement in their solubility that is observed when overproduction is conducted in *E. coli* cultured in M9 minimal media containing ethanolamine,[18, 34, 35] led us to formulate a strategy to further enhance the solubility of these enzymes during heterologous overproduction. This strategy was based on the premise that ethanolamine, when present in a specialized *E. coli* culture media, enhances both the expression of EAL and the uptake of cobalamin, a required cofactor for the enzyme. We therefore designed a plasmid, designated pBAD42-BtuCEDFB, that contains genes that encode proteins involved in cobalamin uptake and trafficking under the control of an arabinose-inducible P_{BAD} promoter. When this plasmid is introduced into *E. coli* cultured in media (LB, M9, or M9 + ethanolamine) containing arabinose, a dramatic increase (up to 50-fold) in the intracellular cobalamin concentration is observed. Unexpectedly, however, there is no significant increase in cobalamin content in *E. coli* that lack the plasmid and that are cultured in media containing ethanolamine as compared to *E. coli* that lack the plasmid and that are cultured in media lacking ethanolamine. This result suggests that the previously observed enhancement in protein yield in *E. coli* cultured in media containing ethanolamine is most likely not due to an increase in intracellular cobalamin concentration. Rather, it may derive from ethanolamine acting as a chemical chaperone that enhances *in vivo* protein folding or that stabilizes folded proteins during their extraction from inclusion bodies. For example, it is well known that arginine and other metabolites such as trehalose and glycine betaine, among others, can enhance *in vivo* protein folding.[68, 69] The enhancement in solubility might also be due to simple differences in metabolism between *E. coli* using ethanolamine as an energy source and *E. coli* using glucose.

AdoCbl and OHCbl are the most prominent forms of cobalamin in *E. coli* cultured in either of the three media with or without pBAD42-BtuCEDFB, while MeCbl is present only in minute quantities. However, when we analyzed Fom3 and PhpK that were overproduced in this expression system for cobalamin content, both contained MeCbl as the predominant form, which most likely derives from the conversion of OHCbl to MeCbl on the enzyme, as has been observed for the reaction catalyzed by TsrM. TsrM, however, was isolated with approximately 40% of its cobalamin as AdoCbl, a fraction that is similar to what we previously reported.[37] We do not envisage any role for AdoCbl in the mechanism of TsrM, or other class B RS methylases, and therefore suspect that TsrM adventitiously binds the molecule because of an active site that must accommodate both cobalamin and SAM, which transfers its methyl group to cob(I)alamin to render MeCbl.[32, 34, 37] To eliminate the pool of free AdoCbl and to produce a more homogenous preparation of TsrM with the active cofactor, we generated a Δ*btuR*:BW25113 strain of *E. coli*, which prevents the conversion of OHCbl to AdoCbl. When overproduced in the Δ*btuR* strain, TsrM is purified without bound AdoCbl, but still contains very little MeCbl. Activity analysis of TsrM purified from the Δ*btuR* strain shows it to exhibit almost two-fold more activity than previously isolated enzyme, which contains bound AdoCbl. This observation is consistent with the conclusion that AdoCbl-bound TsrM is inactive.

The major objective of this work was to assess whether enhanced intracellular cobalamin concentrations can facilitate the purification of class B RS methylases in a soluble form. With respect to TsrM, its overproduction in *E. coli* containing pBAD42-BtuCEDFB and cultured in M9+ethanolamine media led to a significant increase in the yield (from 0.5 mg/L to 3.7 mg/L) of protein purified from the soluble fraction of crude lysate. This yield cannot be attributed to the presence of ethanolamine in the *E. coli* growth media, because a similar yield and purity is obtained when gene expression is conducted in the absence of ethanolamine, as with *tsrM* cloned into the pRham-SUMO plasmid. Similar results were obtained with other class B RS methylases, such as Fom3 and PhpK, which have been resistant to purification from the soluble fraction of *E. coli* crude lysate upon heterologous expression of their genes. In these instances, yields increased from basically 0 to 3 mg/L and 0.84 mg/L of pure protein,

respectively. Moreover, all three proteins were purified with bound cobalamin in a near 1:1 stoichiometry per polypeptide, which eliminates the need to reconstitute the protein with additional cobalamin to measure enzymatic activity. As such, we were able to show that Fom3 is catalytic in the absence of exogenous MeCbl, which contrasts with the result obtained by Sato et al.[65] By contrast, although we are able to isolate pure PhpK containing a near-stoichiometric complement of cobalamin, we are not able to demonstrate the activity reported by Werner et al.[66] Currently, we are investigating other analogs of NAcDMPT as possible substrates. Lastly, when ThnK is overproduced in *E. coli* harboring pBAD42-BtuCEDFB, its yield (7.1 mg/L) is enhanced dramatically over that obtained in previous isolations.[35]

In summary, the strategy that we have developed for the overproduction of class B RS methylases affords yields that are sufficient for further mechanistic analysis and spectroscopic characterization of these enzymes. It must be mentioned, however, that not all cobalamin-dependent RS enzymes are difficult to purify from soluble crude lysate, as has been shown from *in vitro* studies of CysS[70] and OxsB.[71] Indeed, Sato et al. were able to purify Fom3 from *S. wedmorenesis*; however, their yields were not reported.[65] Moreover, not all class B RS methylases that are recalcitrant to purification from soluble crude lysate are readily rendered soluble by this strategy, as our efforts (though fairly minimal) to purify GenK were unsuccessful. Some of these recalcitrant proteins may be produced in insoluble forms for reasons other than, or in addition to, ready access to cobalamin, and other interventions might be needed to render them tractable for purification. Nevertheless, we expect that our strategy may allow the purification of other cobalamin-containing proteins that would otherwise be refractory to purification from the soluble fraction of crude lysate.

ASSOCIATED CONTENT

Supporting Information.

Additional figures (Figure S1 – Figure S29) and Table S1 are available free of charge on the ACS Publications website at DOI: http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*squire@psu.edu

ORCID
Squire J. Booker: 0000-0002-7211-5937

Present Addresses

[§]N.D.L. Laboratory of Bacterial Polysaccharides, Center for Biologics Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, 20993

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

[#]N.D.L and A.J.B. contributed equally

Funding Sources

This work was supported by NIH GM-122595, GM-103268, and AI-111419 to S. J. B, who is also an investigator of the Howard Hughes Medical Institute.

ACKNOWLEDGMENT

The authors thank the Huck Institute for Life Sciences Genomics Core for DNA sequencing. B. J. acknowledges the Penn State Summer Research Opportunity Program for funding.

ABBREVIATIONS

2-HEP, 2-hydroxyethylphosphonate; 2-HPP, 2-hydroxypropylphosphonate; 5'-dA, 5'-deoxyadenosine; AdoCbl, adenosylcobalamin; DTT, dithiothreitol; EAL, ethanolamine ammonia lyase; EPR, electron paramagnetic resonance; Fe-S, iron sulfur; HEP-CMP, cytidylyl-2-hydroxyethylphosphonate; HPP-CMP, cytidyl-2-hydroxypropylphosphonate; LC-MS/MS, liquid chromatography tandem mass spectrometry; MeCbl, methylcobalamin; NAcDMPT, 2-acetylamino-4-hydroxyphosphinylbutanoate; NAcPT, 2-acetylamino-4-hydroxymethylphosphinylbutanoate; OHCbl, hydroxocobalamin; PCR, polymerase chain reaction; rbs, ribose binding site; RS, radical SAM; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine; Trp, tryptophan.

REFERENCES

[1] Akiva, E., Brown, S., Almonacid, D. E., Barber, A. E., Custer, A. F., Hicks, M. A., Huang, C. C., Lauck, F., Mashiyama, S. T., Meng, E. C., Mischel, D., Morris, J. H., Ojha, S., Schnoes, A. M., Stryke, D., Yunes, J. M., Ferrin, T. E., Holliday, G. L., and Babbitt, P. C. (2014) The Structure–Function Linkage Database, *Nucleic Acids Res. 42*, D521-D530.

[2] Bauerle, M. R., Schwalm, E. L., and Booker, S. J. (2015) Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation, *J. Biol. Chem. 290*, 3995-4002.

[3] Zhang, Q., van der Donk, W. A., and Liu, W. (2012) Radical-mediated enzymatic methylation: a tale of two SAMS, *Acc. Chem. Res. 45*, 555-564.

[4] Gomez Macqueo Chew, A., and Bryant, D. A. (2007) Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity, *Ann. Rev. Microbiol. 61*, 113-129.

[5] Kudo, F., Kasama, Y., Hirayama, T., and Eguchi, T. (2007) Cloning of the pactamycin biosynthetic gene cluster and characterization of a crucial glycosyltransferase prior to a unique cyclopentane ring formation, *J. Antibiot. 60*, 492–503.

[6] Mao, Y., Varoglu, M., and Sherman, D. H. (1999) Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564, *Chem. Biol. 6*, 251–263.

[7] Ostash, B., Saghatelian, A., and Walker, S. (2007) A streamlined metabolic pathway for the biosynthesis of moenomycin A, *Chem. Biol. 14*, 257–267.

[8] Hidaka, T., Goda, M., Kuzuyama, T., Takei, N., Hidaka, M., and Seto, H. (1995) Cloning and nucleotide sequence of fosfomycin biosynthetic genes of Streptomyces wedmorensis, *Mol. Gen. Genet. 249*, 274-280.

[9] Kuzuyama, T., Hidaka, T., Kamigiri, K., Imai, S., and Seto, H. (1992) Studies on the biosynthesis of fosfomycin. 4. The biosynthetic origin of the methyl group of fosfomycin, *J. Antibiot. 45*, 1812-1814.

[10] Woodyer, R. D., Shao, Z., Thomas, P. M., Kelleher, N. L., Blodgett, J. A., Metcalf, W. W., van der Donk, W. A., and Zhao, H. (2006) Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster, *Chem. Biol. 13*, 1171-1182.

[11] Woodyer, R. D., Li, G., Zhao, H., and van der Donk, W. A. (2007) New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin, *Chem. Commun.*, 359-361.

[12] Nunez, L. E., Mendez, C., Brana, A. F., Blanco, G., and Salas, J. A. (2003) The biosynthetic gene cluster for the beta-lactam carbapenem thienamycin in *Steptomyces cattleya*, *Chem. Biol. 10*, 301–311.

[13] Kim, J.-Y., Suh, J.-W., Kang, S.-H., Phan, T. H., Park, S.-H., and Kwon, H.-J. (2008) Gene inactivation study of gntE reveals its role in the first step of pseudotrisaccharide modifications in gentamicin biosynthesis, *Biochem. Biophys. Res. Commun. 372*, 730–734.

[14] Anderle, C., Alt, S., Gulder, T., Bringmann, G., Kammerer, B., Gust, B., and Heide, L. (2007) Biosynthesis of clorobiocin: investigation of the transfer and methylation of the pyrrolyl-2-carboxyl moiety, *Arch. Microbiol. 187*, 227–237.

[15] Dairi, T., Ohta, T., Hashimoto, E., and Hasegawa, M. (1992) Self cloning in *Micromonospora olivasterospora* of fms genes for fortimicin A (astromicin) biosynthesis, *Mol. Gen. Genet. 232*, 262–270.

[16] Kuzuyama, T., Seki, T., Dairi, T., Hidaka, T., and Seto, H. (1995) Nucleotide sequence of fortimicin KL1 methyltransferase gene isolated from Micromonospora olivasterospora, and comparison of its deduced amino acid sequence with those of methyltransferases involved in the biosynthesis of bialaphos and fosfomycin, *J. Antibiot. 48*, 1191–1193.

[17] Kelly, W. L., Pan, L., and Li, C. (2009) Thiostrepton biosynthesis: prototype for a new family of bacteriocins, *J. Am. Chem. Soc. 131*, 4327–4334.

[18] Parent, A., Guillot, A., Benjdia, A., Chartier, G., Leprince, J., and Berteau, O. (2016) The B12-Radical SAM Enzyme PoyC Catalyzes Valine Cbeta-Methylation during Polytheonamide Biosynthesis, *J. Am. Chem. Soc. 138*, 15515-15518.

[19] Freeman, M. F., Gurgui, C., Helf, M. J., Morinaka, B. I., Uria, A. R., Oldham, N. J., Sahl, H. G., Matsunaga, S., and Piel, J. (2012) Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides, *Science (New York, N.Y.) 338*, 387-390.

[20] Watanabe, K., Hotta, K., Nakaya, M., Praseuth, A. P., Wang, C. C., Inada, D., Takahashi, K., Fukushi, E., Oguri, H., and Oikawa, H. (2009) Escherichia coli allows efficient modular incorporation of newly isolated quinomycin biosynthetic enzyme into echinomycin biosynthetic pathway for rational design and synthesis of potent antibiotic unnatural natural product, *J. Am. Chem. Soc. 131*, 9347-9353.

[21] Rachid, S., Scharfe, M., Blöcker, H., Weissman, K. J., and Müller, R. (2009) Unusual chemistry in the biosynthesis of the antibiotic chondrochlorens, *Chem. Biol. 16*, 70-81.

[22] Yan, F., LaMarre, J. M., Röhrich, R., Wiesner, J., Jomaa, H., Mankin, A. S., and Galoníc Fujimori, D. (2010) RlmN and Cfr are radical SAM enzymes involved in methylation of ribosomal RNA, *J. Am. Chem. Soc. 132*, 3953-3964.

[23] Grove, T. L., Benner, J. S., Radle, M. I., Ahlum, J. H., Landgraf, B. J., Krebs, C., and Booker, S. J. (2011) A radically different mechanism for S-adenosylmethionine-dependent methyltansferases, *Science 332*, 604–607.

[24] Grove, T. L., Livada, J., Schwalm, E. L., Green, M. T., Booker, S. J., and Silakov, A. (2013) A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr, *Nat. Chem. Biol. 9*, 422–427.

[25] Schwalm, E. L., Grove, T. L., Booker, S. J., and Boal, A. K. (2016) Crystallographic capture of a radical S-adenosylmethionine enzyme in the act of modifying tRNA, *Science 352*, 309-312.
[26] Silakov, A., Grove, T. L., Radle, M. I., Bauerle, M. R., Green, M. T., Rosenzweig, A. C., Boal, A. K., and Booker, S. J. (2014) Characterization of a cross-linked protein-nucleic acid substrate radical in the reaction catalyzed by RlmN, *J. Am. Chem. Soc. 136*, 8221–8228.
[27] Benítez-Páez, A., Villarroya, M., and Armengod, M.-E. (2012) The *Escherichia coli* RlmN methyltransferase is a dual-specificity enzyme that modifies both rRNA and tRNA and controls translational accuracy, *RNA 18*, 1–13.
[28] LaMattina, J. W., Wang, B., Badding, E. D., Gadsby, L. K., Grove, T. L., and Booker, S. J. (2017) NosN, a Radical S-Adenosylmethionine Methylase, Catalyzes Both C1 Transfer and Formation of the Ester Linkage of the Side-Ring System during the Biosynthesis of Nosiheptide, *J. Am. Chem. Soc. 139*, 17438–17445.
[29] LaMattina, J. W., Nix, D. B., and Lanzilotta, W. N. (2016) Radical new paradigm for heme degradation in Escherichia coli O157:H7, *Proc. Natl. Acad. Sci. 113*, 12138-12143.
[30] Ding, W., Li, Y., Zhao, J., Ji, X., Mo, T., Qianzhu, H., Tu, T., Deng, Z., Yu, Y., Chen, F., and Zhang, Q. (2017) The Catalytic Mechanism of the Class C Radical S-Adenosylmethionine Methyltransferase NosN, *Angew Chem. Int. Ed. Engl. 56*, 3857-3861.
[31] Hu, Y., and Ribbe, M. W. (2016) Maturation of nitrogenase cofactor-the role of a class E radical SAM methyltransferase NifB, *Curr. Opin. Chem. Biol. 31*, 188-194.
[32] Pierre, S., Guillot, A., Benjdia, A., Sandström, C., Langella, P., and Berteau, O. (2012) Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes, *Nat. Chem. Biol. 8*, 957-959.
[33] Bandarian, V., and Matthews, R. G. (2004) Measurement of energetics of conformational change in cobalamin-dependent methionine synthase, *Methods Enzymol. 380*, 152-169.
[34] Blaszczyk, A. J., Silakov, A., Zhang, B., Maiocco, S. J., Lanz, N. D., Kelly, W. L., Elliott, S. J., Krebs, C., and Booker, S. J. (2016) Spectroscopic and Electrochemical Characterization of the Iron-Sulfur and Cobalamin Cofactors of TsrM, an Unusual Radical S-Adenosylmethionine Methylase, *J. Am. Chem. Soc. 138*, 3416-3426.
[35] Marous, D. R., Lloyd, E. P., Buller, A. R., Moshos, K. A., Grove, T. L., Blaszczyk, A. J., Booker, S. J., and Townsend, C. A. (2015) Consecutive radical S-adenosylmethionine methylations form the ethyl side chain in thienamycin biosynthesis, *Proc. Natl. Acad. Sci. 112*, 10354-10358.
[36] Freeman, M. F., Vagstad, A. L., and Piel, J. (2016) Polytheonamide biosynthesis showcasing the metabolic potential of sponge-associated uncultivated 'Entotheonella' bacteria, *Curr. Opin. Chem. Biol. 31*, 8-14.
[37] Blaszczyk, A. J., Wang, B., Silakov, A., Ho, J. V., and Booker, S. J. (2017) Efficient Methylation of C2 in L-Tryptophan by the Cobalamin-dependent Radical S-Adenosylmethionine Methylase TsrM Requires an Unmodified N1 Amine, *J. Biol. Chem. 292*, 15456-15467.
[38] Adhin, M. R., and van Duin, J. (1990) Scanning model for translational reinitiation in eubacteria, *J. Mol. Biol. 213*, 811-818.
[39] Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes, *Gene 234*, 187-208.
[40] Ljungdahl, L. G., LeGall, J., and Lee, J. P. (1973) Isolation of a protein containing tightly bound 5-methoxybenzimidazolylcobamide (factor 3m) from Clostridium thermoaceticum, *Biochemistry 12*, 1802-1808.
[41] Blaszczyk, A. J., Wang, R. X., and Booker, S. J. (2017) TsrM as a Model for Purifying and Characterizing Cobalamin-Dependent Radical S-Adenosylmethionine Methylases, *Methods in Enzymol. 595*, 303-329.
[42] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248-254.
[43] Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein. Expr. Purif. 41*, 207-234.

[44] Lanz, N. D., Grove, T. L., Gogonea, C. B., Lee, K. H., Krebs, C., and Booker, S. J. (2012) RlmN and AtsB as models for the overproduction and characterization of radical SAM proteins, *Methods Enzymol. 516*, 125-152.
[45] Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press,U.S., New York, United States.
[46] Cherepanov, P. P., and Wackernagel, W. (1995) Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant, *Gene 158*, 9-14.
[47] Babior, B. M., Moss, T. H., Orme-Johnson, W. H., and Beinert, H. (1974) The mechanism of action of ethanolamine ammonia-lyase, a B-12-dependent enzyme. The participation of paramagnetic species in the catalytic deamination of 2-aminopropanol, *J. Biol. Chem. 249*, 4537-4544.
[48] Davis, B. D., and Mingioli, E. S. (1950) Mutants of Escherichia coli requiring methionine or vitamin B12, *J. Bacteriol. 60*, 17-28.
[49] Kung, H.-F., Spears, C., Greene, R. C., and Weissbach, H. (1972) Regulation of the terminal reactions in methionine biosynthesis by vitamin B12 and methionine, *Arch. Biochem. Biophys. 150*, 23-31.
[50] Miles, Z. D., Myers, W. K., Kincannon, W. M., Britt, R. D., and Bandarian, V. (2015) Biochemical and Spectroscopic Studies of Epoxyqueuosine Reductase: A Novel Iron-Sulfur Cluster- and Cobalamin-Containing Protein Involved in the Biosynthesis of Queuosine, *Biochemistry 54*, 4927-4935.
[51] Bridwell-Rabb, J., and Drennan, C. L. (2017) Vitamin B12 in the spotlight again, *Curr. Opin. Chem. Biol. 37*, 63-70.
[52] Roth, J. R., Lawrence, J. G., and Bobik, T. A. (1996) Cobalamin (Coenzyme B12): Synthesis and biological significance, *Annu. Rev. Microbiol. 50*, 137–181.
[53] Gudmundsdottir, A., Bell, P. E., Lundrigan, M. D., Bradbeer, C., and Kadner, R. J. (1989) Point mutations in a conserved region (TonB box) of Escherichia coli outer membrane protein BtuB affect vitamin B12 transport, *J. Bacteriol. 171*, 6526-6533.
[54] Mann, B. J., Holroyd, C. D., Bradbeer, C., and Kadner, R. J. (1986) Reduced activity of TonB-dependent functions in strains of Escherichia coli, *FEMS Microbiology Letters 33*, 255-260.
[55] Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: Structure of the BtuB:TonB complex, *Science 312*, 1396–1399.
[56] Locher, K. P., Lee, A. T., and Rees, D. C. (2001) The E. coli BtuCD structure: A framework for ABC transporter architecture and mechanism, *Science 296*.
[57] Korkhov, V. M., Mireku, S. A., Veprintsev, D. B., and Locher, K. P. (2014) Structure of AMP-PMP-bound BtuCD and mechanism of ATP-powered vitamin B12 transport by BtuCD–F, *Nat. Struct. Mol. Biol. 21*, 1097–1099.
[58] DeVeaux, L. C., and Kadner, R. J. (1985) Transport of vitamin B12 in Escherichia coli: cloning of the btuCD region, *J. Bacteriol. 162*, 888-896.
[59] de Veaux, L. C., Clevenson, D. S., Bradbeer, C., and Kadner, R. J. (1986) Identification of the btuCED polypeptides and evidence for their role in vitamin B12 transport in Escherichia coli, *J. Bacteriol. 167*, 920-927.
[60] Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Köster, W., Mohanty, A. K., Wiener, M. C., and Kadner, R. J. (2002) Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*, *J. Bacteriol. 184*, 706–717.
[61] Nou, X., and Kadner, R. J. (2000) Adenosylcobalamin inhibits ribosome binding to *btuB* RNA, *Proc. Natl. Acad. Sci. USA 97*, 7190–7195.
[62] Rioux, C. R., and Kadner, R. J. (2989) Vitamin B12 transport in *Escherichi coli* K12 does not require the *btuE* gene of the *btuCED* operon, *Mol. Gen. Genet. 217*, 301–308.
[63] Lundrigan, M. D., and Kadner, R. J. (1989) Altered cobalamin metabolism in Escherichia coli btuR mutants affects btuB gene regulation, *J. Bacteriol. 171*, 154-161.

[64] Allen, K. D., and Wang, S. C. (2014) Initial characterization of Fom3 from Streptomyces wedmorensis: The methyltransferase in fosfomycin biosynthesis, *Arch. Biochem. Biophys. 543*, 67-73.
[65] Sato, S., Kudo, F., Kim, S.-Y., Kuzuyama, T., and Eguchi, T. (2017) Methylcobalamin-Dependent Radical SAM C-Methyltransferase Fom3 Recognizes Cytidylyl-2-hydroxyethylphosphonate and Catalyzes the Nonstereoselective C-Methylation in Fosfomycin Biosynthesis, *Biochemistry 56*, 3519-3522.
[66] Werner, W. J., Allen, K. D., Hu, K., Helms, G. L., Chen, B. S., and Wang, S. C. (2011) In vitro phosphinate methylation by PhpK from Kitasatospora phosalacinea, *Biochemistry 50*, 8986-8988.
[67] Hidaka, T., Hidaka, M., Kuzuyama, T., and Seto, H. (1995) Sequence of a P-methyltransferase-encoding gene isolated from a bialaphos-producing Streptomyces hygroscopicus, *Gene 158*, 149-150.
[68] Prasad, S., Khadatare, P. B., and Roy, I. (2011) Effect of chemical chaperones in improving the solubility of recombinant proteins in *Escherichia coli*, *Appl. Environ. Microbiol. 77*, 4603–4609.
[69] Leibly, D. J., Nguyen, T. N., Kao, L. T., Hewitt, S. N., Barrett, L. K., and van Voorhis, W. C. (2012) Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins, *PloS One 7*, e52482.
[70] Wang, Y., Schnell, B., Baumann, S., Muller, R., and Begley, T. P. (2017) Biosynthesis of Branched Alkoxy Groups: Iterative Methyl Group Alkylation by a Cobalamin-Dependent Radical SAM Enzyme, *J. Am. Chem. Soc. 139*, 1742-1745.
[71] Bridwell-Rabb, J., Zhong, A., Sun, H. G., Drennan, C. L., and Liu, H. W. (2017) A B12-dependent radical SAM enzyme involved in oxetanocin A biosynthesis, *Nature 544*, 322-326.

Insert Table of Contents artwork here