CobT and BzaC catalyze the regiospecific activation and methylation of the 5-hydroxybenzimidazole lower ligand in anaerobic cobamide biosynthesis

Vitamin B₁₂ and other cobamides are essential cofactors required by many organisms and are synthesized by a subset of prokaryotes via distinct aerobic and anaerobic routes. The aerobic biosynthesis of 5,6-dimethylbenzimidazole (DMB), the lower ligand of vitamin B₁₂, involves five reactions catalyzed by the bza operon gene products, namely the hydroxybenzimidazole synthase BzaAB/BzaF, phosphoribosyltransferase CobT, and three methyltransferases, BzaC, BzaD, and BzaE, that conduct three distinct methylation steps. Of these, the methyltransferases that contribute to benzimidazole lower ligand diversity in cobamides remain to be characterized, and the precise role of the bza operon protein CobT is unclear. In this study, we used the bza operon from the anaerobic bacterium Moorella thermaacetica (comprising bzaA-bzaB-cobT-bzaC) to examine the role of CobT and investigate the activity of the first methyltransferase, BzaC. We studied the phosphoribosylation catalyzed by MtCobT and found that it regiospecifically activates 5-hydroxybenzimidazole (5-OHBza) to form the 5-OHBza-ribotide (5-OHBza-RP) isomer as the sole product. Next, we characterized the domains of MtBzaC and reconstituted its methyltransferase activity with the MtCobT product 5-OHBza-RP and its ribosome derivative 5-OHBza-R. Unexpectedly, we found that 5-OHBza-R is the most favored MtBzaC substrate. Our results collectively explain the long-standing observation that the attachment of the lower ligand in anaerobic cobamide biosynthesis is regiospecific. In conclusion, we validate the MtBzaC as a SAM:hydroxybenzimidazole-ribose methyltransferase (HBIR-OMT). Finally, we propose a new pathway for the synthesis and activation of the benzimidazolyl lower ligand in anaerobic cobamide biosynthesis.

Vitamin B₁₂ (cobalamin) is an essential micronutrient required by humans and several other organisms to mediate biochemical reactions involving methyl transfers and radical-based rearrangements in primary metabolism (1–3). Cobalamin is a member of the cobamide cofactors family—members of this family are characterized by a tetrapyrrolic corrin ring and a central cobalt ion axially coordinated to an upper and a lower ligand (Fig. 1A). The lower ligand is covalently attached to the corrin ring via a nucleotide loop and is typically a benzimidazole, purine, or phenol derivative, thereby contributing to the diversity of the naturally occurring cobamide cofactors (1, 4, 5). The cobamide biosynthesis pathway utilizes over 30 enzymes for the synthesis and assembly of its structural components (6, 7). In the overall pathway, the corrin ring and lower ligand are synthesized separately and then attached together via the nucleotide loop to produce the cobamide 1. Cobalamin contains 5,6-dimethylbenzimidazole (DMB) 6 as the lower ligand, and humans exclusively use cobalamin as a cofactor for the synthesis of succinyl-CoA and methionine (5, 8). On the other hand, microbes synthesize and utilize several other benzimidazole derivatives, such as benzimidazole (Bza), 5-hydroxybenzimidazole (5-OHBza 3), 5-methoxybenzimidazole (5-OMeBza 4), 5-methoxy-6-methyl-benzimidazole (5-OMe-6-MeBza 5), and 5-methylbenzimidazole (5-MeBza) as lower ligands in cobamides (4, 9–12).

Cobamides can be synthesized via two distinct routes, one that requires molecular oxygen and another that is oxygen-sensitive (6, 7). Unlike the pathways to synthesize the corrin ring and the nucleotide loop that consist of some comparable steps, the lower ligand biosynthesis is completely different in the two routes (7, 13, 14). The aerobic biosynthesis of DMB 6 is catalyzed by the enzyme BluB, which orchestrates a complex fragmentation of reduced flavin mononucleotide in the presence of molecular oxygen (15–17). Then the phosphoribosyltransferase CobT introduces DMB 6 into the cobalamin biosynthesis pathway by forming α-ribazole phosphate (DMB-RP 7) via a nucleophilic substitution reaction (18). In contrast, the anaerobic biosynthesis of DMB 6 uses a modular approach involving the gene products of the bzaA-bzaB-cobT-bzaC-bzaD-bzaE operon (Fig. 1, B and C) (12). Previously, BzaF (the gene product of bzaF, the single gene homolog of bzaA and bzaB) has been shown to produce 5-OHBza 3 from 5-aminoimidazole ribotide (AIR 2), an intermediate in the purine biosynthesis pathway (19, 20) (Fig. 1C). The next steps are predicted to be as follows. BzaC methylates 5-OHBza 3 to produce 5-OMeBza 4, which is followed by a second methylation by BzaD to produce 5-OMe-6-MeBza 5, and then a final methylation by BzaE produces DMB 6. Lastly, CobT in the bza operon is predicted to activate DMB 6 in a manner similar to the aerobic pathway (Fig. 1C). Interestingly, all of the benzimidazole derivatives found as intermediates on the pathway to synthesize DMB 6 also exist as lower ligands in naturally occurring cobamides. Recent studies on anaerobic cobamide producers show that their genome typically contain none, one, two, or all three of the bza operon.
methyltransferases, which determines the cobamide they produce (Fig. 1B) (12, 21). For example, *Geobacter sulfurreducens* has only *bzaF* and *cobT* and produces 5-hydroxybenzimidazoyl-cobamide ([5-OHBza]Cba, Factor III) (12), *Moorella thermoacetica* has *bzaA-bzaB-cobT-bzaC* and produces 5-methoxybenzimidazoyl-cobamide ([5-OMeBza]Cba, Factor IIIm) (14, 22, 23), and *Eubacterium limosum* and *Acetobacterium woodii* contain *bzaA-bzaB-cobT-bzaC-bzaD-bzaE* and produce cobalamin (4, 12, 14). This leads to two important observations: (i) the variations in the *bza* operon found in microbes, mainly via the three methyltransferases *BzaC*, *BzaD*, and *BzaE*, appear to contribute to the diversity of benzimidazoyl cobamides, and (ii) among the *bza* operons characterized so far, a *cobT* gene is often found to succeed the *bzaA-bzaB* or *bzaF* genes.

Cobamides are ascribed an important role in shaping microbial communities (24, 25). Thus, development of computational approaches to identify *bza* operon genes in metagenomic data sets to predict cobamide diversity in a microbial community is starting to gain interest (21, 26–28). To confidently use *bza* gene sequences as a proxy for the cobamides produced in the community, the enzymatic activity and role of each enzyme in the *bza* operon needs to be experimentally elucidated. Until now, *in vitro* activity of only *BzaF* as a 5-hydroxybenzimidazole synthase has been established (19, 20). In this study, we undertake the characterization of the next two enzymes: CobT, a phosphoribosyltransferase, and *BzaC*, a SAM-dependent methyltransferase, in the organism *M. thermoacetica*. We choose *M. thermoacetica* as it has been reported to produce only one cobamide, [5-OMeBza]Cba, and it contains a subset of the *bza* operon with *bzaA-bzaB-cobT-bzaC* genes that, when heterologously expressed, produce this cobamide (12, 23). First, we report the *in vitro* characterization of the CobT homolog found within the *bza* operon and show that it regiospecifically phosphoribosylates 5-OHBza to produce 5-OHBza-RP. Next, we reconstitute the activity of a previously uncharacterized enzyme *BzaC*, the first methyltransferase in the *bza* operon, and establish its activity as a SAM-dependent methyltransferase. Further, we find that *MtBzaC* shows the highest activity with 5-hydroxybenzimidazole riboside (5-OHBza-R) formed by the dephosphorylation of the *MtCobT* product. Cumulatively, our results establish *BzaC* as a SAM-dependent methyltransferase (HBIR-OMT), provide a plausible role for the co-occurrence of the gene *cobT* within the *bza* operon, and allow us to explain

---

**Figure 1. Methyltransferases of the bza operon contribute to the diversity of cobamide lower ligands.**

A, general structure of cobamides 1 with the benzimidazole class of lower ligand shown in the gray box. B, the table shows the lower ligands reported to be produced by the cobamide-producing anaerobes *G. sulfurreducens, Desulfuromonas acetoxidicans, M. thermoacetica, Clostridium formicoaceticum, and E. limosum* (4). Naturally occurring benzimidazole lower ligands generally differ by substitutions at the C5 and C6 positions of the benzimidazole as indicated with R1 and R2. In *C. formicoaceticum*, both 5-methoxy, 6-methylbenzimidazole (5-OMeBza,6-MeBza 5) and 5,6-dimethylbenzimidazole (DMB 6) are found as lower ligands. The benzimidazole derivative produced can be attributed to the presence of the combination of *bza* genes in each of these organisms. C, predicted pathway for biosynthesis of DMB using the *bza* operon (12). Distinct subsets of the *bza* operon are found in organisms as shown in B, yielding a variety of benzimidazolyl lower ligands. In the pathway, 5-hydroxybenzimidazole (5-OHBza 3) is synthesized from the purine biosynthesis intermediate 5-aminoimidazoleribotide (AIR 2) by *BzaAB* or their single gene homolog *BzaF* and is predicted to undergo subsequent methylations by the *bzaC*, *bzaD*, and *bzaE* gene products (12, 19, 20). The *cobT* gene encodes for a phosphoribosyltransferase, which is predicted to activate the final lower ligand produced by the operon for formation of the cobamide 1.
previous observations of regiospecific attachment of lower ligands in the biosynthesis of benzimidazolyl cobamides.

Results

Bioinformatic analysis of the *M. thermoacetae* bza operon

*MtCobT* confirms the presence of conserved catalytic residues

The phosphoribosyltransferase *MtCobT* introduces the free lower ligands into the cobamide biosynthesis pathway by activating it via a nucleophilic substitution reaction. The reaction leads to the formation of a unique α-glycosidic bond between a heteroatom (nitrogen or oxygen) of the lower ligand and C1’ of a ribose phosphate ring (Fig. 1C) (7, 18, 29, 30). Incidentally, most of the previously characterized *MtCobT* homologs are from aerobic microorganisms, and no *MtCobT* homologs from anaerobes harboring the bza operon have been characterized yet (18, 31–35).

Comparison of protein sequences of the *M. thermoacetae bza operon* (*MtCobT*) with previously characterized *MtCobT* homologs from *Salmonella enterica* (*ScoB*,) *Sporomusa ovata* (*SoArsA, SoArsB*), *Sinorhizobium meliloti* (called *SmtCobU*), and *Methanocaldococcus jannaschii* (*MjtCobT*) showed that *MtCobT* shares 20–44% identity with each of these homologs. The sequence alignment shows that *MtCobT* possesses the conserved catalytic Glu124 residue along with the known consensus sequences Met-Arg-Leu-Glu-Gly-$(X)$ of the active site. Comparison with the *ScoB* sequence shows that *Leu*31–*Ser*37, *Pro*82, and *Val*86 from *MjtCobT* align well with the residues known to interact with benzimidazole substrates. Also, Gly179–Thr187 of *MjtCobT* align with the binding site for the phosphoribosyl donor, nicotinic acid mononucleotide (32, 36, 37) (Fig. 2A and Fig. S1A). The conservation of these key residues indicates that the *MtCobT* homolog encoded by the *M. thermoacetae bza operon* is likely a functional phosphoribosyltransferase.

Bioinformatic analysis of the BzaC protein sequences reveals two major domain architectures among BzaC homologs

Previous studies have demonstrated that the heterologous expression of bzaC genes from obligate anaerobes *E. limosum* and *M. thermoacetae* in *Escherichia coli* results in the production of [5-OMeBza]Cba (12). We conducted a bioinformatic search to compare the BzaC enzyme with previously characterized methyltransferases in the literature and to infer its predicted function. We observed that *MtBzaC* and *EBzaC* share limited sequence identity with few class I methyltransferases. Among the previously characterized enzymes, 3-hydroxy-5-methyl-1-naphthoate *O*-methyltransferase (AziB2) shares 27.8% identity with *MtBzaC* and 22.3% identity with *EBzaC*, acetylserotonin *O*-methyltransferase (ASMT) shares 24.7 and 20.7% identity, and 4-amino-4-de-(dimethylamino)-anhydrotetracycline-N,N-dimethyl-*O*-methyltransferase (OxyT) shares 25.8 and 18.6% identity, respectively (38–40). As per a conserved domain search, both *MtBzaC* and *EBzaC* sequences contain a dimerization domain belonging to Pfam16864 at the N terminus followed by a methyltransferase domain, belonging to the superfamily cl17173 (Fig. 2B). The dimerization domain of Pfam16864 is also found in other methyltransferase subfamilies and is linked to protein dimerization (39). The second domain in all of the BzaC homologs we examined is a member of the class I methyltransferases, known to methylate DNA, protein, and small molecules such as catechols, ubiquinone, and flavones among a wide range of substrates. All members of this class employ SAM as a methyl donor (41, 42). A sequence alignment of BzaC with other class I methyltransferases with previously solved crystal structures highlighted the conserved consensus sequence Gly-$X$-Gly-$X$-Gly characteristic of SAM-binding sites (43) (Fig. 2C and Fig. S1B).

Additionally, *EBzaC* contains a 145-amino acid C-terminal domain of unknown function called DUF2284 belonging to Pfam10050 (Fig. 2B). The structural, functional, and physiological relevance of this domain is still unclear. A recent *in silico* survey for occurrence and distribution of ~3400 domains of unknown functions reported the presence of DUF2284 to be limited to bacteria and archaea (44). We also found that along with being associated with BzaCs, DUF2284 is also present as a separate protein encoded within other genomic contexts (data not shown). A sequence alignment of the DUF2284 domain of BzaC sequences shows two conserved cysteine-rich consensus of $C_2X_2C_3X$ and $C_2X_3C_2X_3C$, which resemble the consensus sequence for Fe-S cluster–binding motifs (Fig. S1B, asterisks) (45, 46). Overall, the analysis of the protein sequences shows that there exist two main types of BzaC enzymes distinguished by the presence or absence of the DUF2284 domain. The phylogenetic analysis of the dimerization and methyltransferase domains of BzaC homologs (Fig. 2D and Fig. S1C) shows that the sequences are taxonomically conserved. Also, we observe that the overall distribution of the BzaC homologs shows that the ones that lack DUF2284 are scattered among the clades of BzaC sequences that possess DUF2284. This suggests that the dimerization and methyltransferase domains are conserved across BzaC homologs from various taxonomies, and the presence of the DUF2284 is independent of these two domains. We therefore predict that the role of DUF2284 in the activity of certain BzaCs to be beyond the methyltransferase reaction.

*MtCobT* shows regiospecificity for activation of 5-OHBza

We cloned, expressed, and purified *MtCobT* and *EcCobT* proteins of sizes 39.7 and 39.1 kDa, respectively (Fig. S2A, lanes 1–4). We then reconstituted the *in vitro* activity of *MtCobT* with benzimidazole substrates predicted to be synthesized by the *M. thermoacetae bza operon*, 5-OHBza 3 and 5-OMeBza 4. Previous studies have established the role of CobT in the regioselective formation of cobamide isomers and its contribution to cobamide structure and function (33, 47, 48). Homologs of CobT show varying regioselectivity in the formation of isomers arising from attachment of the C1’ of the ribose with either of the two nitrogens of asymmetric benzimidazolyl lower ligands, such as 5-OHBza 3 and 5-OMeBza 4 (47) (Fig. 3A and Fig. S4A). With 5-OHBza 3 as substrate (synthesized as described in the supporting Methods, Fig. S3), *MtCobT* shows a single peak at 9.5 min (Fig. 3B). To confirm that the observed product peak is of a single isomer and to further confirm the identity of the isomer produced, the *MtCobT* reaction with 5-
OHBza 3 had to be compared with the reaction of a CobT that produced both isomer products. After studying the reactions of a few homologs, we found that E. coli CobT (EcCobT) yields two product peaks that elute at 9.6- and 10.4-min retention time. The peak at 9.6 min exhibited the distinct absorbance spectrum of 5-OHBza-RP 8, and the peak at 10.4 min showed that of 6-OHBza-RP 9 with identical mass spectra on LC–MS with a fragment of 135 m/z corresponding to 5-OHBza 3 (Fig. 3, C and D), confirming them as isomers. The retention time and the UV-visible spectra of the single product formed by MtCobT coincided with the peak for 5-OHBza-RP 8, and its identity was further confirmed by LC–MS.

On the contrary, MtCobT upon reaction with 5-OMeBza 4 formed two isomers, and both peaks exhibit absorbance spectra similar to reported spectra for 5-OMeBza-RP 12 and 6-OMeBza-RP 13 isomers (Fig. S4, B and C) (47). The mass spectra of both peaks are identical with a 149 m/z peak confirming the fragment for 5-OMeBza 4 in the two isomeric products (Fig. S4D). Interestingly, we observed that EcCobT specifically forms a single isomer 5-OMeBza-RP 12 with 5-OMeBza 4 as the substrate (Fig. S4B). From our results, the regiospecificity exhibited by MtCobT for phosphoribosylation of 5-OHBza 3 and the lack of regiospecificity for phosphoribosylation of 5-OMeBza 4 may imply a mechanism sensitive to the methylation state of the benzimidazole substrate, which is likely to be novel and unique to this class of CobT homologs.

Biochemical analysis of the dimerization domain and methyltransferase domain of MtBzaC

Next, we cloned, heterologously overexpressed, and purified the MtBzaC protein with the expected molecular mass of 41.9 kDa (Fig. S2, A (lanes 5 and 6) and B). Size-exclusion chromatography with the purified enzyme showed a peak corresponding to the dimer size and some higher oligomers of the protein (Fig. 4A, solid trace). Next, a construct with dimerization domain–deleted MtBzaCΔDiD was created, which yielded a soluble protein corresponding to the mass 31.9 kDa (Fig. S2, A...
Characterization of CobT and BzaC from the bza operon

**Editors' Pick:**

Figure 3. Characterization of the phosphoribosyltransferase activity of MtCobT with 5-OHBza. A, CobT reaction scheme for the formation of α-riboside phosphate (RP) products. The two expected products for an asymmetric substrate such as 5-OHBza are 5-OHBz-RP and 6-OHBz-RP. B, HPLC trace for the in vitro reconstitution of E. coli CobT (EcCobT) and MtCobT with 5-OHBza as the substrate. EcCobT produces both 5-OHBz-RP and 6-OHBz-RP isomers, whereas MtCobT specifically produces only 5-OHBz-RP. C, characteristic UV-visible spectra of the two isomers validate the identities of the products (47). D, the mass spectra of peaks at 9.6 and 10.9 min in the LC trace corresponding to the reported α-riboside phosphate isomers for 5-OHBza show identical mass as expected. The fragment of 135.055 corresponds to the 5-OHBza base.

In vitro reconstitution and optimization of methyltransferase activity of MtBzaC with 5-OHBza as substrate

The gene product of MtBzaC has been previously demonstrated to show methylation activity under both aerobic and anaerobic conditions when heterologously expressed in *E. coli* K-12 str. MG1655 (12). Thus, we attempted to reconstitute the in vitro activity of MtBzaC under aerobic conditions with the predicted substrate 5-OHBza and the methyl donor SAM to improve the product yield. The end point reaction with MtBzaC and 5-OHBza as a substrate for 72 h at 25 °C showed a new peak at 19.8 min on the HPLC (Fig. S5A) and 16.5 min on LC–MS, with retention time and mass identical to that of a standard of 5-OMeBza, the expected product (Fig. 5A). The MtBzaCADid construct yielded a *K*~d~ value of 400.8 μM with SAM, implying that deletion of the dimerization domain does not impair the affinity of the methyltransferase domain toward the cofactor (Fig. 4B, triangle).

**Figure 4.** Validation of the predicted dimerization domain and SAM-binding domain of MtBzaC. A, size-exclusion chromatogram of purified MtBzaC and MtBzaCADid. MtBzaC shows peaks corresponding to masses of a monomer (~82.8 kDa) and some higher oligomer masses near the void volume. MtBzaCADid shows a prominent peak for a monomer (~37.1 kDa) and a minor peak for a trimer mass (~134.8 kDa). B, the intrinsic protein fluorescence of MtBzaC and MtBzaCADid was monitored in the presence of increasing amounts of SAM and SAH. Fitting the points to Equation 1 (see “Experimental procedures”) gave the following binding constants: *K*~d~ SAM = 505.5 ± 87.1 μM, *K*~d~ SAH = 404.7 ± 46.4 μM with MtBzaC, and *K*~d~ SAM = 400.8 ± 46.9 μM with MtBzaCADid (lanes 7 and 8) and C. When subjected to size-exclusion chromatography, MtBzaCADid eluted as a major peak corresponding to the expected size of its monomer and few peaks for aggregates and higher oligomers (Fig. 4A, dashed trace). This confirms that the N-terminal dimerization domain of MtBzaC is functional and is involved in oligomerization of the enzyme.

Next, using the phenomenon of change in the intrinsic fluorescence of the protein upon ligand binding, we investigated the ability of MtBzaC to bind SAM. MtBzaC contains 4 tryptophan and 11 tyrosine residues and shows a significant intrinsic fluorescence with an excitation maximum at 280 nm and corresponding emission maxima at 328 nm (Fig. S5A). When SAM was added to the enzyme in increasing concentrations, we observed a steady decrease in the protein fluorescence, which saturated at a concentration of ~1.0 mM SAM (Fig. S5A). A plot of normalized fluorescent intensity (*F~/*F~0~) at 328 nm against the concentration of SAM fitted to Equation 1 (see “Experimental procedures”) and Ref. 49) yielded the dissociation constant (*K*~d~) of 505.5 μM for SAM (Fig. 4B, squares). SAH 15, which is a known inhibitor for class I methyltransferases, also shows a similar trend, with a *K*~d~ value of 404.7 μM, suggesting that SAH 15 is likely an inhibitor for MtBzaC (Fig. 4B, circle) (50). The MtBzaCADid construct yielded a *K*~d~ value of 400.8 μM with SAM 14, implying that deletion of the dimerization domain does not impair the affinity of the methyltransferase domain toward the cofactor (Fig. 4B, triangle).

In vitro reconstitution and optimization of methyltransferase activity of MtBzaC with 5-OHBza as substrate

The gene product of MtBzaC has been previously demonstrated to show methylation activity under both aerobic and anaerobic conditions when heterologously expressed in *E. coli* K-12 str. MG1655 (12). Thus, we attempted to reconstitute the in vitro activity of MtBzaC under aerobic conditions with the predicted substrate 5-OHBza and the methyl donor SAM to improve the product yield. The end point reaction with MtBzaC and 5-OHBza as a substrate for 72 h at 25 °C showed a new peak at 19.8 min on the HPLC (Fig. S5A) and 16.5 min on LC–MS, with retention time and mass identical to that of a standard of 5-OMeBza, the expected product (Fig. 5A). The MtBzaCADid construct yielded a *K*~d~ value of 400.8 μM with SAM, implying that deletion of the dimerization domain does not impair the affinity of the methyltransferase domain toward the cofactor (Fig. 4B, triangle).

In vitro reconstitution and optimization of methyltransferase activity of MtBzaC with 5-OHBza as substrate

The gene product of MtBzaC has been previously demonstrated to show methylation activity under both aerobic and anaerobic conditions when heterologously expressed in *E. coli* K-12 str. MG1655 (12). Thus, we attempted to reconstitute the in vitro activity of MtBzaC under aerobic conditions with the predicted substrate 5-OHBza and the methyl donor SAM 14 (Fig. 5A). The end point reaction with MtBzaC and 5-OHBza as a substrate for 72 h at 25 °C showed a new peak at 19.8 min on the HPLC (Fig. S5B) and 16.5 min on LC–MS, with retention time and mass identical to that of a standard of 5-OMeBza, the expected product (Fig. 5A). The UV-visible (Fig. 5D) and NMR spectra (Fig. 5E) of the product obtained confirmed its identity as 5-OMeBza, thus verifying that MtBzaC is a functional methyltransferase catalyzing the methylation of 5-OHBza using SAM as the methyl donor. Moreover, the reconstitution assays with the MtBzaCADid mutant did not result in product formation, which indicates that the dimerized or oligomerized form of MtBzaC is likely essential for the methylation (Fig. S5C). Notably, the net conversion of 500 μM 5-OHBza to 5-OMeBza product by MtBzaC, quantified using standard curves (Fig. S6, A–C) was ~13–17 μM (3.1–4.4% conversion), indicating a limited activity of MtBzaC under in vitro

---

**Figure 4.** Validation of the predicted dimerization domain and SAM-binding domain of MtBzaC. A, size-exclusion chromatogram of purified MtBzaC and MtBzaCADid. MtBzaC shows peaks corresponding to masses of a monomer (~82.8 kDa) and some higher oligomer masses near the void volume. MtBzaCADid shows a prominent peak for a monomer (~37.1 kDa) and a minor peak for a trimer mass (~134.8 kDa). B, the intrinsic protein fluorescence of MtBzaC and MtBzaCADid was monitored in the presence of increasing amounts of SAM and SAH. Fitting the points to Equation 1 (see “Experimental procedures”) gave the following binding constants: *K*~d~ SAM = 505.5 ± 87.1 μM, *K*~d~ SAH = 404.7 ± 46.4 μM with MtBzaC, and *K*~d~ SAM = 400.8 ± 46.9 μM with MtBzaCADid.
Figure 5. *In vitro* reconstitution and optimization of methyltransferase activity of MtBzaC activity with 5-OHBza. A, predicted methyltransferase reaction of BzaC-5-hydroxybenzimidazole (5-OHBza 3) is methylated to 5-methoxybenzimidazole (5-OMeBza 4) using SAM 14 as the methyl donor. B, the LC–MS extracted ion chromatogram (EIC) for the reconstitution of the methyltransferase activity of MtBzaC. The reaction yields a product peak corresponding to the retention time of a commercially obtained standard of 5-OMeBza 4. C, the mass spectrum of the product peak shows a mass/charge of 149.069 corresponding to the mass/charge of a standard sample of 5-OMeBza 4. D, the product shows UV–visible spectra identical to a standard of 5-OMeBza 4 with maxima at 250 and 280 nm. E, 1H NMR of the purified product shows peaks for Hα (8.17, s), Hβ (7.28, 7.27, d, J = 1.96), Hγ (7.049, 7.043, 7.027, 7.021, dd, J = 8.8, 2.4), Hδ (7.63, 7.61, d, J = 8.8), Hζ (3.83, s), identical to the 1H NMR of the standard sample of 5-OMeBza 4. F, optimization of the MtBzaC reaction to increase the net yield of 5-OMeBza 4. In the first set, 1 mM commercially available SAM was used (white square), and in the second set, SAM was synthesized *in situ* by SAM synthetase (EcMetK) from 1 mM methionine and 2 mM ATP (white circle). Methyltransferase inhibitors SAH 14 and MTA (marked with an asterisk) were removed by the addition of methylthioadenosine nucleosidase (EcMTAN) in the MtBzaC reaction in the third and fourth sets (filled squares, filled circles). The activity of MtBzaC increased 5.78 ± 0.49 and 4.75 ± 0.69 folds with commercial and *in situ* synthesized SAM, respectively (average of n = 2, S.E.) with the addition of MTAN.

conditions. To examine the biochemical factors affecting the methyltransferase activity of MtBzaC, the reaction was reconstituted under a host of different conditions. The activity of MtBzaC is mostly unaffected by metal ions, DTT, and crowding agent BSA (Fig. S7, A–C). Reconstitution of the activity under an anoxic environment improved the product yield by only 1.34-fold, which implies that MtBzaC may not be affected by oxygen (Fig. S7D). Next, a linear increase in the enzyme concentration resulted in a linear increase in the concentration of product formed (Fig. S7E). This implies that the enzyme MtBzaC may be the limiting factor in the reaction because of irreversible inhibition by a reaction by-product or the absence of other factors involved in the reaction. HPLC traces of purified enzyme and commercially bought SAM 14 shows the presence of degradation products of SAM 14, namely SAH 15 and 5-methylthioadenosine (MTA), which are known inhibitors of class I methyltransferases (50, 51). Thus, we reconstituted the reaction using SAM 14 synthesized *in situ* by SAM synthetase (EcMetK), resulting in a 1.44 ± 0.10-fold increase in MtBzaC activity (Fig. 5F, filled squares and filled circles). Next, the enzyme methylthioadenosine nucleosidase (EcMTAN) that hydrolyzes the potential inhibitors MTA and SAH 15 was added to the reaction to remove accumulated MTA (Fig. S5D) (51, 52). Co-incubation with EcMTAN increased the activity of MtBzaC 5.78 ± 0.49-fold with commercial SAM 14 and 4.75 ± 0.69-fold with *in situ* synthesized SAM 14, along with an increase in the optimum reaction temperature from 25 to 37 °C (Fig. 5F (squares and circles) and Fig. S7F). Thus, we conclude that SAH 15 and MTA formed due to spontaneous degradation of SAM 14 inhibit the methyltransferase activity of MtBzaC.

**Investigating the repertoire of possible substrates for BzaC**

While commencing this study, we predicted three possible ways in which the 5-OHBza 3 base can be methylated by MtBzaC: (i) 5-OHBza 3 can be methylated by MtBzaC to yield 5-OMeBza 4, which is then phosphoribosylated by MtCoBT; (ii) the 5-OHBza 3 is first activated to form 5-OHBza-RP 8 by MtCoBT, which is then methylated by MtBzaC to form 5-OMeBza-Ribotide (5-OMeBza-RP 12); or (iii) the 5-OHBza 3 is first activated and attached as the lower ligand to form [5-OHBza]Cba, which is later methylated by MtBzaC to finally yield [5-OMeBza]Cba. The absence of any known cobamide-binding site in the MtBzaC primary sequence or in its domains eliminates the third possibility right away. Because we have
extensively tested possibility (i), we moved on to test possibility (ii), where 5-OHBza
3 is first activated by CobT and then methylated by BzaC.

5-OHBza-RP 8 and 6-OHBza-RP 9 were enzymatically synthesized using 5-OHBza
3 and MtCobT and EcCobT, respectively (see supporting Methods and Fig. S8 (A
and B)). The methyltransferase activity of MtBzaC was individually tested with 5-OHBza-RP 8
and 6-OHBza-RP 9 as substrates and SAM 14 as methyl donor under reaction conditions
optimized previously. LC–MS analysis of these reactions showed that MtBzaC does not methylate
either of the two phosphoribosylated substrates (Fig. S9, A and B). Instead, the LC–MS chromatogram
of the end point reaction of MtBzaC with 5-OHBza-RP 8 for 48 h repeatedly (n = 3) showed new peaks for 5-hydroxybenzimidazole riboside (5-OHBza-R 16) and 5-methoxybenzimidazole riboside (5-OmeBza-R 17) at 11.2 and 14.6 min, respectively (Fig. S9A). The riboside derivatives 5-OHBza-R 16 and 5-
OmeBza-R 17 can be formed by spontaneous dephosphorylation of 5-OHBza-RP 8 and 5-OmeBza-RP 12, respectively. Thus, the formation of 5-OmeBza-R 17 could occur either if 5-OHBza-R 16 was being methylated by MtBzaC or if 5-
OmeBza-RP 12 formed in the reaction was spontaneously dephosphorylating to 5-OmeBza-R 17. To investigate this observation, we enzymatically synthesized 5-OHBza-R 16 (Fig. S8C) and 6-hydroxybenzimidazole riboside (6-OHBza-R 18) and tested the methyltransferase activity of MtBzaC with both of the riboside isomers (Fig. 6A). Surprisingly, we found that MtBzaC methylates the two isomers to form 5-OmeBza-R 17 and 6-OmeBza-R 19 under both aerobic and anoxic conditions. The MtBzaC activity was 5.05 times higher with 5-OHBza-R 16
and 1.22 times with 6-OHBza-R 18 than the activity with 5-OHBza 3 (Fig. 6B, standard curves used are shown in Fig. S6 (C

Figure 6. Methylation of CobT products by MtBzaC. A, HPLC-fluorescence chromatogram for MtBzaC reaction with 5-OHBza 3 and the two novel substrates, 5-OHBza-R 16 and 6-OHBza-R 18. All three substrates are methylated to form 5-OmeBza 4, 5-OmeBza-R 17, and 6-OmeBza-R 19, respectively. B, quantitation of methylated product formed by MtBzaC with 5-OHBza 3, 5-OHBza-R 16, 6-OHBza-R 18, 5-OmeBza-RP 8, and 6-OmeBza-RP 9 as the substrates. Relative,
5-OHBza-R 16 is the most preferred substrate, whereas no detectable amount of riboside phosphate products (i.e. 5-OmeBza-RP 12 and 6-OmeBza-RP 13) were found. C, LC–MS EIC of the MtBzaC and 5-OHBza-R 16 reaction. The EIC corresponding to 5-OmeBza-R 17 (black trace) shows a peak at 14.5 min in the
reaction. The mass spectrum of the product peak confirms its identity as 5-OmeBza-R 16 with a fragment of m/z 149.0702 corresponding to 5-OmeBza 4 base as shown in the inset. D, LC–MS EIC of the MtBzaC and 6-OHBza-R 18 reaction. The EIC corresponding to 6-OmeBza-R 19 (black trace) shows a peak at 14.5 min in the reaction. The mass spectrum of the product peak confirms its identity as 6-OmeBza-R 19 with a fragment of m/z 149.0702 corresponding to 5-OmeBza 4 base as shown in the inset.
and D)). LC–MS analysis of individual reactions confirmed the product peaks as 5-OHbzA-R 17 and 6-OHbzA-R 19 (Fig. 6, C and D). In summary, of all of the substrates presented to MtBzAC under in vitro conditions, 5-OHbzA-R 16 is preferentially methylated (Fig. 6B and Fig. S9 A and B)). Thus, we conclude that 5-OHbzA-R 16 is the likely physiological substrate of MtBzAC and that activation of the lower ligand precedes its methylation in the benzimidazole biosynthesis pathway.

Discussion

The bza operon, which codes for the genes bzaA–bzaB–cobT–bzaC–bzaD–bzaE, is implicated in the anaerobic biosynthesis of DMB. According to the previously proposed pathway, the formation of 5-OHbzA 3 by the gene products of bzaA–bzaB (or its single gene homolog bzaF) is followed by three methylation reactions catalyzed by gene products of bzaC, bzaD, and bzaE, each catalyzing a unique methylation to yield DMB (12). Finally, the gene product of cobT is proposed to activate DMB for the formation of cobalamin. Until now, only the formation of 5-OHbzA 3 was characterized by in vitro studies in the literature (12, 19, 20). In this study, we characterize the in vitro reactions catalyzed by the next two gene products CobT and BzaC using the M. thermoacetica bza operon, which contains the genes bzaA–bzaB–cobT–bzaC (a naturally occurring subset of bza operon). Our explorations reveal that, contrary to the previously proposed pathway, the activation of the 5-OHbzA precedes its methylation (i.e. 5-OHbzA is first phosphoribosylated and then dephosphorylated to form 5-OHbzA-R, which is methylated by BzaC to yield 5-OHbzA-R).

We analyzed the in vitro reaction of purified MtCobT and showed that, similar to previously characterized CobT homologs, it is a functional phosphoribosyltransferase (18, 33). Remarkably, MtCobT displays a regio-specific attachment of 5-OHbzA 3, producing only the 5-OHbzA-R 8 isomer (Fig. 3B). All of the previously characterized CobT homologs lack such regio-specificity when tested for phosphoribosylation of 5-OHbzA 3. (35, 47, 48). Notably, MtCobT does not exhibit regio-specificity for activation of 5-OHbzA 4 and produces both 5-OHbzA-R 12 and 6-OHbzA-R 13, which sets it apart from previously reported CobT homologs from S. melliloti (SmCobT) and S. enterica (SeCobT) (Fig. 3B) (47). Thus, CobT homologs located within the bza operon may have different reactivity as compared with CobT enzymes found in alternate gene neighborhoods, and we are investigating this further. We also show that the E.coli CobT homolog (EcCobT) is similar to SeCobT and activates 5-OHbzA 4 regiospecifically but lacks regio-specificity with 5-OHbzA 3 as substrate. This property of regio-specificity of CobT homologs has been shown to play a significant role in determining the variety of cobamides and norcobamides produced by various microorganisms (47, 48).

Next, we characterized MtBzAC, which has a dimerization domain and a methyltransferase domain with a SAM-binding site and, hence, is predicted to utilize SAM 14 to methylate 5-OHbzA 3 to form 5-OHbzA 4. The successful reconstitution of its activity with the predicted substrate 5-OHbzA 3 and SAM 14 as the methyl donor agrees with previous studies in M. thermoacetica using 14C-labeled methionine (23). However, the net methylation activity was low even under a wide range of optimization methods that were tested (Fig. S5F and Fig. S7 A–F).

Finally, when we tested with 5-OHbzA–derived ribotide and riboside as alternate substrates, MtBzAC showed the highest activity with 5-OHbzA-R 16 (Fig. 6, B and C). Hence, contrary to its previously proposed role in the benzimidazole biosynthesis pathway, our in vitro studies show that MtBzAC is a SAM-dependent HBIR-OMT.

We hypothesize that if the in vitro substrate preference we find for MtBzAC has to be achieved in vivo, activation of 5-OHbzA 3 by MtCobT followed by a dephosphorylation step must precede the methylation by MtBzAC. There exists precedence for this possibility in the study of the cobamide biosynthesis pathway—the enzyme CobC has been demonstrated to dephosphorylate the α-ribazole phosphate (DMB-RP 7) to form α-ribazole (DMB-R 20) (53) (Fig. 7A, Route 1). Alternately, the dephosphorylation can occur nonenzymatically (54). Further, the enzyme cobalamin synthase encoded by the gene cobS, which is responsible for attaching the activated lower ligand to a cobinamide backbone, has been reported to use both DMB-RP 7 and DMB-R 20 as substrates (55) (Fig. 7A, Route 1 and Route 2), even though DMB-R is a less preferred substrate for CobS, as illustrated through studies in S. enterica (56). We find homologs for CobS and CobC in the M. thermoacetica genome (Fig. 7B), and further characterization of these and other CobS and CobC homologs from obligate anaerobes harboring a bza operon will be required to validate these findings. Additionally, DMB-R 20 is shown to occur in cobamide-producing organisms—it has been shown to be utilized by certain microorganisms such as Listeria innocua and Geobacillus kaustophilus and fluxed into the cobalamin biosynthesis pathway (54, 57).

Combining the in vitro activities for MtCobT and MtBzAC in this study and previous literature reports of cobamides made by anaerobic microbes, we lay out the following. Based on the pathway proposed currently in the literature, where activation of lower ligand by MtCobT occurs after the methylation of 5-OHbzA to form 5-OHbzA, the pathway would likely yield both [5-OHbzA]Cba and [6-OHbzA]Cba isomers as MtCobT produces both isomeric phosphoribosylated products. Instead, our findings show that MtCobT phosphoribosylates 5-OHbzA 3 to produce a single ribosylated isomer, 5-OHbzA-RP, which is later dephosphorylated and methylated. Therefore, we expect the final cobamide produced to be [5-OHbzA]Cba exclusively. Our findings are corroborated by previous literature that shows M. thermoacetica naturally produces only [5-OHbzA]Cba (10, 23). Also, the heterologous expression of the M. thermoacetica bza operon in E. coli showed the formation of [5-OHbzA]Cba exclusively as compared with an E. coli control with added 5-OHbzA, where both [5-OHbzA]Cba and [6-OHbzA]Cba were formed (12). Finally, results from our study show that despite 6-OHbzA-R also being an activated form of 5-OHbzA 3, MtBzAC shows poor activity with this isomer as compared with 5-OHbzA-R. All of this taken together strongly puts forth 5-OHbzA-R 16 as an intermediate in the bza operon pathway (Fig. 7C).
Finally, our studies shed light on a long-standing unexplained observation in anaerobic DMB biosynthesis (14, 58). Extensive labeling studies conducted to understand how DMB is synthesized in anaerobic organisms had shown that the two nitrogens in DMB are derived from glycine and glutamine (59, 60). Intriguingly, even though DMB is a symmetric molecule, the nucleotide loop is attached specifically through the nitrogen atom derived from glutamine (58, 60). This puzzling result was justified by the hypothesis that an asymmetric intermediate from benzimidazole biosynthesis must undergo regiospecific phosphoribosylation, and the resulting intermediate would be a substrate for subsequent methylations (12, 14, 61).

Because the M. thermoacetica bza operon produces the asymmetric benzimidazoles 5-OHBza and 5-OMeBza, the activities of the MtCobT and MtBzaC, provided significant insights into this long-standing puzzle. The co-occurrence of the cobT gene with the bzaA-bzaB or bzaF genes and our finding that MtCobT produces one regiospecific product that is then methylated by MtBzaC provides evidence for the convergence of biosynthesis of the benzimidazole lower ligand with its final incorporation into the cobamide. Also, these insights are important for enzymatic characterization of the subsequent methyltransferase bzaD and bzaE gene products, which may utilize activated lower ligand or cobamide intermediates as substrates.

**Conclusions**

In this study, we have characterized the activity of a new methyltransferase, BzaC, and established it to be a SAM-dependent 5-OHBza-R (also called α-ribazole) as well as adenosylcobamide 5-phosphate (AdoCb-5'-P 21). The cobalamin 5-phosphate synthase enzyme CobS is reported to attach both α-DMB-ribazole and α-DMB-riboside 20 with activated cobinamide (AdoCbl-GDP) to form AdoCbl-5'-P 21 and AdoCbl 22, respectively. B, the genes encoding for enzymes CobS and CobC (dark gray) are present along with all of the genes required for cobamide biosynthesis in M. thermoacetica (named as described in a review by Mattes et al. (7)). C, based on our studies, we propose this pathway for biosynthesis and activation of the lower ligand in M. thermoacetica. The 5-OHBza 3 synthesized by BzaAB undergoes a regiospecific activation by MtCobT to form 5-OHBza-RP 8, which is dephosphorylated either by MtCobC (53) or nonenzymatically (54) to yield the riboside 5-OHBza-R 16. The 5-OHBza-R 16 is methylated by MtBzaC to yield 5-OMeBza-R 17, which can further be attached to the cobinamide backbone to form [5-OMeBza]Cba, the native cobamide produced by M. thermoacetica.

**Experimental procedures**

**Chemicals**

All medium components and antibiotics for bacterial culture were obtained from HiMedia, SRL chemicals, and TCI Chemicals. Enzymes used for cloning were purchased from DSS
Takara Biosciences and New England Biolabs. DNA and plasmid purification kits were obtained from Agilent and Qiagen. HPLC solvents were obtained from SD Fine Chemicals. 5-OMeBza 4, SAM 14, SAH 15, and LC–MS grade methanol were purchased from Merck–Sigma Aldrich. 5-OHBza 3 and its ribotide and riboside derivatives (5-OHBza-RP 8, 6-OHBzazRP 9, 5-OHBzazR 16, and 6-OHBzazR 18) were synthesized as described in the supporting Methods.

Bioinformatic studies

The protein sequences of CobT homologs from M. thermoacetica, E. coli K-12 str. MG1655, S. enterica str. typhimurium, S. melliloti, S. ovata, and M. jannaschii were obtained from NCBI GenBank™ (62). The FASTA sequences were aligned, and the percentage identity matrix was obtained using the MUSCLE alignment tool hosted at RRID:SCR_011812 (63). The resulting alignment file was visualized using Boxshade (RRID:SCR_007165). The sequences of BzaC homologs were identified by a protein BLAST (RRID:SCR_004870) (64) search using BzaC protein sequences from E. limosum, M. thermoacetica, Geobacter lovleyi, Desulfotomaculum kuznetsovii, Thermicola potens, and Syntrophacectis schinkii against the non-redundant protein sequence database at NCBI. To increase the pool of sequences, BLAST searches were conducted against each bacterial phylum and then each class when more than 100 sequences returned from a phylum. Duplicates were removed, and sequences were retrieved using Batch entrez (RRID:SCR_016634). This led to a list of 833 sequences comprising predicted O-methyltransferases and hypothetical proteins. This data set was refined for finding candidate BzaC sequences based on scores from HMM search HMM models (21) and gene neighborhood search. The corresponding nucleotide sequence for each protein sequence was used to trace the gene neighborhood in the GenBank™ (62) database. Because most of the bza operon gene products are annotated as hypothetical proteins, the identities of the neighboring genes were verified using conserved domain search (RRID:SCR_018729) (65). The O-methyltransferases that were encoded by a gene downstream to a cobalamin riboswitch or to other genes of the bza operon were considered as BzaC homologs. The resulting set of 90 unique BzaC-like protein sequences was subjected to phylogenetic analysis.

To eliminate the possibility of the DUF2284 sequence biasing the phylogenetic distribution because it constitutes about one-third of the sequence, the analysis was conducted by aligning only the sequences corresponding to the dimerization and methyltransferase domains. The sequences of the first two domains of BzaC homologs were aligned using MUSCLE (RRID:SCR_011812) (63). The alignment was visualized and converted to phylip4.0 format using BioEdit (66). The phylip4.0 file was uploaded on the CIPRES Science gateway server (RRID:SCR_008439) (67), and a maximum likelihood tree was constructed using rapid bootstrapping RaXML-HPC2 (68) on XSEDE (69). The results were extracted as a Newick tree and uploaded on the iToL (RRID:SCR_018174) (70) server to visualize, analyze, and annotate the tree.

**Construction of plasmids and overexpression and purification of the recombinant proteins**

All of the plasmids used in this study were constructed using standard molecular biology techniques and restriction-free cloning (71), as described in the supporting information. Purified plasmids were transformed into competent E. coli BL21 (DE3) cells (72). All protein purifications were conducted using standard immobilized metal affinity chromatography using nickel-nitrilotriacetic acid and as described in the supporting information.

**Size-exclusion chromatography**

Size-exclusion chromatography was performed on an analytical GE Sephadex-200 column calibrated for molecular markers from 600 to 29 kDa (73). The column was equilibrated with buffer containing 50 mM Tris-Cl, pH 8.0, containing 150 mM NaCl and 0.025% β-mercaptoethanol. Protein freshly eluted from nickel-nitrilotriacetic acid chromatography was loaded onto the equilibrated column. To elute the fractions, 30 ml of the buffer was run through the column at 0.25 ml/min, and absorbance at 280 nm was recorded to monitor the protein elution profile.

**Intrinsic fluorescence assays**

To study protein-ligand binding, the phenomenon of fluorescence quenching of the protein was used. Purified MtBzaC shows fluorescence emission maxima at 328 nm upon an excitation at 280 nm. Henceforth, for all calculations, the relative fluorescence intensity at 328 nm was used. To calculate the $K_d$ value for SAM 14 or SAH 15 binding to MtBzaC, we used the following procedure. In a 96-well fluorescence plate, 100 μl of 2× protein (10 μM) was added in each of 9 wells and then 2× solution of 9 dilutions of the ligand in the range of 0–2.5 mM was added. The plate was incubated in the plate reader at 25 °C for 10 min, and then fluorescence emission was recorded from 300 to 600 nm upon excitation at 280 nm. The fold-change in fluorescence was plotted against the final concentration of ligand in each well. The plot was fit to the following equation as described previously (49) using GraphPad Prism version 6.

$$ F = F_0 + \frac{\Delta F[S]}{[S] + K_d} $$

(Eq. 1)

Where $F$ represents fluorescence for protein + ligand, $F_0$ is fluorescence for protein, $\Delta F = F_{\max} - F$, [S] is concentration of ligand, and $K_d$ is the dissociation constant.

**Synthesis of the substrates**

The substrates 5-OHBza 3 and 5-OHBza-RP 8, 6-OHBza-RP 9, 5-OHBza-R 16, and 6-OHBza-R 18 were synthesized and purified using a method described previously (47) with some modifications to protocol (see supporting Methods and Figs. S3 and S8).

J. Biol. Chem. (2020) 295(31) 10522–10534 10531

EDITORS’ PICK: Characterization of CobT and BzaC from the bza operon

---

**Editors’ Pick**

Characterization of CobT and BzaC from the bza operon

---
**Enzymatic reactions and analysis**

Enzymatic reactions with MtCobT (UniProt accession number A0A0K1TPX5) and EcCobT (A0A037Y2M3) were set up as described previously (35). Briefly, 10 μM purified enzyme, 2 mM NMN 10, and 500 μM 5-OHBza 3 or 5-OMeBza 4 were added in 50 mM Tris buffer, pH 8.0, and 1 mM MgCl₂. Reactions were incubated at 25 °C for 48 h and quenched with 1% formic acid. They were analyzed by reverse-phase HPLC method-2 and LC–MS method-1 as described in the supporting Methods.

For enzymatic reactions with MtBzaC (UniProt number A0A0K1Y1W0), for all end point enzyme assays with 5-OHBza 3, freshly purified His-tagged protein was used, and reactions were set up under either ambient laboratory conditions or anoxic conditions maintained in a glove bag (95% N₂, 5 %H₂) as specified in each experiment. A typical reaction contained the following: 50 mM Tris-Cl, pH 8.0, with 46–50 μM MtBzaC protein, 500 μM substrate, and 1000 μM SAM 14, 10 mM MgCl₂, and 0.025% β-mercaptoethanol. As indicated in select experiments, 20 μM EcMTAN was additionally present.

For the MtBzaC assay in the presence of SAM synthetase (EcMetK) and MTA nucleosidase (EcMTAN), reactions were set up as follows. 1 mM methionine, 2 mM ATP, and 1 mM 5-OHBza 3 were added to 50 mM Tris-Cl, pH 8.0, buffer containing 10 mM MgCl₂, 50 mM KCl, and 0.025% β-mercaptoethanol. SAM synthesis was initiated by the addition of 2 μM EcMetK and after incubation at 25 °C for 30 min, 50 μM MtBzaC was added to the reaction mix. As indicated in select experiments, 20 μM EcMTAN was added to the reaction 5 min prior to the addition of MtBzaC.

After an incubation of 48 h at 25 or 37 °C, the reaction was quenched by the addition of 1% formic acid and centrifuged at 14,000 rpm at 4 °C to remove precipitated proteins, and the supernatant was used for reverse-phase HPLC and LC–MS analysis as described in the supporting information.

**Data availability**

All data are contained within the article.

**Acknowledgments**—We thank Olga Sokolovskaya and Amanda Shelton for providing experimental and bioinformatics insights. We thank Mohan Reddy for help with size-exclusion chromatography and MALDI-TOF analysis and the mass spectrometry and NMR facilities and their managers at IISER Pune. We thank Aindrila Mukhopadhyay, Ruchi Anand, and Thomas Pucadyil for helpful discussions. Contributions by research assistant Bhushan S. Kulkarni are also acknowledged. We also thank Michiko E. Taga, Jennifer Bridwell-Rabb, and Mrutyunjay Nair for critical reading of the manuscript.

**Author contributions**—Y. M. and A. B. H. conceptualization; Y. M. and S. S. data curation; Y. M., S. S., and A. B. H. formal analysis; Y. M., S. S., P. M. D., and A. B. H. investigation; Y. M., S. S., P. M. D., and M. B. S. methodology; Y. M. and A. B. H. writing-original draft; A. B. H. supervision; A. B. H. funding acquisition; A. B. H. project administration.

**References**

1. Roth, J. R., Lawrence, J. G., and Bobik, T. A. (1996) Cobalamin (coenzyme B₁₂): synthesis and biological significance. *Annu. Rev. Microbiol.* 50, 137–181 CrossRef Medline
2. Banerjee, R., and Ragsdale, S. W. (2003) The many faces of vitamin B₁₂ catalysis by cobalamin-dependent enzymes. *Annu. Rev. Biochem.* 72, 209–247 CrossRef Medline
3. Giedyk, M., Goliszewska, K., and Gryko, D. (2015) Vitamin B₁₂ catalyzed reactions. *Chem. Soc. Rev.* 44, 3391–3404 CrossRef Medline
4. Supperich, E., Eisinger, H. J., and Schurr, S. (1990) Corrinoids in anaerobic bacteria. *FEMS Microbiol. Rev.* 87, 355–360 CrossRef
5. Allen, R. H., and Stabler, S. P. (2008) Identification and quantitation of cobalamin and cobalamin analogues in human feces. *Am. J. Clin. Nutr.* 87, 1324–1335 CrossRef Medline
6. Moore, S. J., and Warren, M. J. (2012) The anaerobic biosynthesis of vitamin B₁₂. *Biochem. Soc. Trans.* 40, 581–586 CrossRef Medline
7. Mattes, T. A., Escalante-Semerena, J. C., Deery, E., and Warren, M. J. (2017) Cobalamin biosynthesis and insertion. In *Encyclopedia of Inorganic and Bioinorganic Chemistry* (Scott, R.A., ed) pp. 1–24, John Wiley & Sons, Inc., New York
8. Supperich, E., and Nexo, E. (1991) Effect of the cobal-N coordination on the cobamide recognition by the human vitamin B₁₂ binding proteins intrinsic factor, transcobalamin and haptocorrin. *Eur. J. Biochem.* 199, 299–303 CrossRef Medline
9. Scherer, P., Hölßriegel, V., Krug, C., Bokel, M., and Renz, P. (1984) On the biosynthesis of 5-hydroxybenzimidazolylcobamide (vitamin B₁₂-factor III) in *Methanosarcina barkeri*. *Arch. Microbiol.* 138, 354–359 CrossRef
10. Wurm, R., Weyhenmeyer, R., and Renz, P. (1975) On the biosynthesis of 5-methoxybenzimidazole: precursor-function of 5-hydroxybenzimidazole, benzimidazole and riboflavin. *Eur. J. Biochem.* 56, 427–432 CrossRef Medline
11. Kräutler, B., Kohler, H. E., and Supperich, E. (1988) 5′-Methylbenzimidazolyl-cobamides are the corrinoids from some sulfate-reducing and sulfur-metabolizing bacteria. *Eur. J. Biochem.* 176, 461–469 CrossRef Medline
12. Hazra, A. B., Han, A. W., Mehta, A. P., Mok, K. C., Osadchiy, V., Begley, T. P., and Taga, M. E. (2015) Anaerobic biosynthesis of the lower ligand of vitamin B₁₂. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10792–10797 CrossRef Medline
50. Zhang, J., and Zheng, Y. G. (2016) SAM/SAH analogs as versatile tools for SAM-dependent methyltransferases. *ACS Chem. Biol.* **11**, 583–597

51. Siegrist, J., Aschwanden, S., Mordhorst, S., Thöny-Meyer, L., Richter, M., and Andexer, J. N. (2015) Regioselective O-methylation of catechols by using three-enzyme cascades. *ChemBioChem* **16**, 2576–2579

52. Cornell, K. A., Swarts, W. E., Barry, R. D., and Riscoe, M. K. (1996) Characterization of recombinant *Escherichia coli* 5′-methyladenosine/5′-adenosylhomocysteine nucleosidase analysis of enzymatic activity and substrate specificity. *Biochim. Biophys. Res. Commun.* **228**, 724–732

53. Maggio-Hall, L. A., and Escalante-Semerena, J. C. (1999) *In vitro* synthesis of the nucleotide loop of cobalamin by *Salmonella typhimurium* enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11798–11803

54. Mattes, T. A., and Escalante-Semerena, J. C. (2017) *Salmonella enterica* synthesizes 5,6-dimethylbenzimidazolyl-(DMB)-α-riboside. Why some Firmicutes do not require the canonical DMB activation system to synthesize adenosylcobalamin. *Mol. Microbiol.* **103**, 269–281

55. Maggio-Hall, L. A., Claas, K. R., and Escalante-Semerena, J. C. (2004) The last step in coenzyme B12 synthesis is localized to the cell membrane in bacteria and archaea. *Microbiology* **150**, 1385–1395

56. Zayas, C. L., and Escalante-Semerena, J. C. (2007) Reassessment of the late steps of coenzyme B12 synthesis in *Salmonella enterica*: evidence that dephosphorylation of adenosylcobalamin-5′-phosphate by the CobC phosphatase is the last step of the pathway. *J. Bacteriol.* **189**, 2210–2218

57. Gray, M. J., and Escalante-Semerena, J. C. (2010) A new pathway for the synthesis of α-ribazole-phosphate in *Listeria innocua*. *Mol. Microbiol.* **77**, 1429–1438

58. Lamm, L., Beckmann, G., and Renz, P. (1982) Biosynthesis of vitamin B12 in anaerobic bacteria. Mode of incorporation of glycine into the 5,6-dimethylbenzimidazole moiety in *Eubacterium limosum*. *Eur. J. Biochem.* **122**, 569–571

59. Lamm, L., Hörig, J. A., Renz, P., and Beckmann, G. (1980) Biosynthesis of vitamin B12 in anaerobic bacteria. Mode of incorporation of glycine into the 5,6-dimethylbenzimidazole moiety and some labelled substrates. *Eur. J. Biochem.* **109**, 115–118

60. Vogt, J. R. A., and Renz, P. (1988) Biosynthesis of vitamin B12 in anaerobic bacteria. Experiments with *Eubacterium limosum* on the origin of the amide groups of the corrin ring and of N-3 of the 5,6-dimethylbenzimidazole part. *Eur. J. Biochem.* **171**, 655–659

61. Schulze, B., Vogler, B., and Renz, P. (1998) Biosynthesis of vitamin B12 in anaerobic bacteria—experiments with *Eubacterium limosum* on the transformation of 5-hydroxy-6-methylbenzimidazole, its nucleoside, its corrinoid, and of 5-hydroxybenzimidazolylcobamide in vitamin B12. *Eur. J. Biochem.* **254**, 620–625

62. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2005) GenBank. *Nucleic Acids Res.* **33**, 34–38

63. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797

64. Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410

65. Marchler-Bauer, A., and Bryant, S. H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W327–W331

66. Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98

67. Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, 2010, pp. 1–8

68. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313

69. Towns, J., Cockerill, T., Dahan, M., Foster, L., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G., Roskies, R., Scott, J. R., and Wilkins-Diehr, N. (2014) XSEDE: accelerating scientific discovery. *Comput. Sci. Eng.* **16**, 62–74

70. Letunic, I., and Bork, P. (2016) Interactive tree of life (iTOI) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245

71. Van Den Ent, F., and Löwe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74

72. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130

73. Sandanaraj, B. S., Reddy, M. M., Bhandari, P. J., Kumar, S., and Aswal, V. K. (2018) Rational design of supramolecular dynamic protein assemblies by using a micelle-assisted activity-based protein-labeling technology. *Chem. Eur. J.* **24**, 16085–16096

74. Cheng, J., Sibley, C. D., Zabeh, R., and Finan, T. M. (2007) A *Sinorhizobium meliloti minE* mutant has an altered morphology and exhibits defects in legume symbiosis. *Microbiology* **153**, 375–387

75. Yang, J., and Zhang, Y. (2016) Protein structure and function prediction using I-TASSER. *Curr. Protoc. Bioinformatics* **52**, 5.8.1–5.8.15

76. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction, and analysis. *Nat. Protoc.* **10**, 845–858