Identification of a novel cobamide remodeling enzyme in the beneficial human gut bacterium *Akkermansia muciniphila*

Kenny C. Mok[a], Olga M. Sokolovskaya[a], Alexa M. Nicolas[a], Zachary F. Hallberg[a], Adam Deutschbauer[b], Hans K. Carlson[b], and Michiko E. Taga[a]#

[a]Department of Plant & Microbial Biology, University of California, Berkeley, Berkeley, CA

[b]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA

#Correspondence to taga@berkeley.edu

Running title: Cobamide remodeling in a beneficial gut bacterium

Abstract word count: 233

Text word count: 6,114

Abstract

The beneficial human gut bacterium *Akkermansia muciniphila* provides metabolites to other members of the gut microbiota by breaking down host mucin, but most of its other metabolic functions have not been investigated. *A. muciniphila* is known to use cobamides, the vitamin B_{12} family of cofactors with structural diversity in the lower ligand, though the specific cobamides it can use have not been examined. We found that growth of *A. muciniphila* strain Muc^T was nearly identical with each of seven cobamides tested, in contrast to nearly all bacteria that have been studied. Unexpectedly, this promiscuity is due to cobamide remodeling – the removal and replacement of the lower ligand – despite the absence of the canonical remodeling enzyme CbiZ in *A. muciniphila*. We identified a novel enzyme, CbiR, that is capable of initiating the remodeling process by hydrolyzing the phosphoribosyl bond in the nucleotide loop of cobamides. CbiR does not share homology with other cobamide remodeling enzymes or B_{12}-binding domains, and instead is a member of the AP endonuclease 2 enzyme superfamily. We speculate that CbiR enables bacteria to repurpose cobamides they otherwise cannot use in order to grow under a cobamide-requiring condition; this function was confirmed by heterologous expression of *cbiR* in *E. coli*. Homologs of CbiR are found in over 200 microbial taxa across 22 phyla, suggesting that many bacteria may use CbiR to gain access to the diverse cobamides present in their environment.

Importance

Cobamides, the vitamin B_{12} family of cobalt-containing cofactors, are required for metabolism in all domains of life, including most bacteria. Cobamides have structural variability in the lower ligand, and selectivity for particular cobamides has been observed in most organisms studied to date. Here, we discover that the beneficial human gut bacterium *Akkermansia muciniphila* can use a diverse range of cobamides due to its ability to change the cobamide structure via "cobamide remodeling". We identify and characterize the novel enzyme CbiR that is necessary for initiating the cobamide remodeling process. The discovery of this enzyme has implications not only for understanding the ecological role of *A. muciniphila* in the gut, but for other bacteria that carry this enzyme as well.

Introduction

The human gut microbiota is composed of diverse communities of microbes that play important roles in human health (1-4). Disruption of the composition of the microbiota, known as dysbiosis, is associated with numerous disease states (5-9). While the immense complexity and interindividual variability of the microbiota have made it challenging to identify the specific functions of most community members, particular taxa are starting to be linked to health and disease (10-12), with the bacterium *Akkermansia muciniphila* recently emerging as a beneficial microbe due to its distinctive metabolic capabilities (13, 14).

A. muciniphila is thought to benefit the host by inducing mucus production, improving gut barrier function, and stimulating a positive inflammatory response (15-24). *A. muciniphila* is one of few bacteria capable of using mucin, the main component of mucus, as a sole carbon, nitrogen, and energy source (25). Mucin degradation products released by *A. muciniphila* are used as carbon sources by butyrate-producing bacteria and likely other bacteria, and for this reason *A. muciniphila* is thought to be a keystone species in the gut (26, 27). In addition to providing metabolites to neighboring microbes, in coculture *A. muciniphila* can use a cobamide cofactor, pseudocobalamin (pCbl, Fig. 1A), provided by *Eubacterium hallii* for the production of propionate (26). Both butyrate and propionate positively affect host metabolism and immune function (28-30).

Cobamides are a family of cobalt-containing corrinoid cofactors that include vitamin B_{12} (cobalamin, Cbl), an essential micronutrient for humans. Cobamides are required by organisms in all domains of life, but are synthesized only by a subset of prokaryotes (31-33). While some strains of *A. muciniphila* were shown or predicted to produce cobamides *de novo*, the type strain, Muc^T, is incapable of *de novo* cobamide production (34). Instead, strain Muc^T and most other *A. muciniphila* strains are predicted to be capable of cobinamide (Cbi, Fig. 1A) salvaging (33, 34), a process in which a cobamide is synthesized from the late precursor Cbi (35). Thus, the four cobamide-dependent metabolic pathways present in *A. muciniphila* function in most strains, including Muc^T, only when a cobamide or a late precursor such as Cbi is provided by another organism. Several other human gut bacteria have similarly been found to use cobamide cofactors but are unable to produce them *de novo*, including *Bacteroides fragilis*, *Bacteroides thetaiotaomicron*, *Bacteroides vulgatus*, *Clostridioides difficile*, *Enterococcus faecalis*, *Escherichia coli*, and *Parabacteroides distasonis* (36-40). In addition to these specific examples, genomic analysis suggests that dependence on cobamide-producing microbes is widespread in the gut and other environments: 58% of human gut bacteria and 49% of all sequenced bacteria are predicted to use cobamides but lack the capacity to produce them (33).

A feature that sets cobamides apart from other enzyme cofactors is that different microbes produce structurally distinct cobamides (41). This variability is mostly limited to the lower ligand, which can be benzimidazolyl, purinyl, or phenolyl bases (Fig. 1C). Individual cobamide-producing bacteria typically synthesize only one type of cobamide, but microbial communities have been found to contain four to eight different cobamides or cobamide precursors (42-45). A study of 20 human subjects showed that the human gut is dominated by the purinyl class of cobamides, with

benzimidazolyl and phenolyl cobamides and Cbi also present (42). The structural diversity in cobamides impacts growth and metabolism, as most organisms studied to date are selective in their cobamide use (39, 46-54). For example, the human gut bacterium *B. thetaiotaomicron* can use benzimidazolyl and purinyl, but not phenolyl, cobamides (37); *Dehalococcoides mccartyi* and most eukaryotic algae are selective for particular benzimidazolyl cobamides (55-57); and *Sporomusa ovata* requires phenolyl cobamides (58). Thus, microbes that depend on cobamides produced by others may struggle to grow in environments lacking their preferred cobamides. However, some organisms have evolved mechanisms of acquiring the specific cobamides that function in their metabolism. For example, bacterial cobamide uptake can be somewhat selective, as shown in a study in *B. thetaiotaomicron* (37). Another strategy used by some microbes is cobamide remodeling, the removal and replacement of the lower ligand.

Cobamide remodeling was first described in *Rhodobacter sphaeroides* (59), but has also been observed in the bacteria *D. mccartyi* and *Vibrio cholerae* and the algae *Pavlova lutheri* and *Chlamydomonas reinhardtii* (45, 48, 55, 57). In each case, cobamide remodeling enables the organism to repurpose a cobamide that poorly supports growth. In *R. sphaeroides*, the cobamide remodeling process is initiated by the enzyme CbiZ, which hydrolyzes the amide bond adjacent to the aminopropanol linker (Fig. 1A) (59); in subsequent steps, cobamide biosynthesis is completed with a different lower ligand via the activity of six gene products, most of which are also required for Cbi salvaging. *In vitro*, *R. sphaeroides* CbiZ hydrolyzes pCbl but not Cbl (59). This specificity is thought to drive the conversion of pCbl, a cofactor that *R. sphaeroides* cannot use, into Cbl, which functions in its metabolism. *D. mccartyi* also has homologs of *cbiZ* (57), while cobamide remodeling in *V. cholerae* was recently shown to involve the cobamide biosynthesis enzyme CobS (48). The genes required for cobamide remodeling in algae have not been identified. Nevertheless, *A. muciniphila* does not encode a homolog of *cbiZ*, and therefore we assessed the cobamide selectivity of *A. muciniphila* strain Muc^T to understand the cobamide metabolism of this beneficial gut bacterium.

Here we show that *A. muciniphila* strain Muc^T is able to grow equivalently when provided a variety of cobamides. We found that this lack of selectivity is due to the unexpected ability of *A. muciniphila* to remodel cobamides. We identified a previously uncharacterized phosphodiesterase in *A. muciniphila* that we named CbiR, which initiates the remodeling process by hydrolyzing cobamides. Heterologous expression in *E. coli* shows that CbiR expands access to a cobamide that does not otherwise support growth. Homologs of CbiR are present in the genomes of microbes in diverse habitats from 22 phyla, and phylogenetic analysis establishes CbiR as a new, distinct clade within the AP endonuclease 2 superfamily. These observations enhance the understanding of the metabolic roles of *A. muciniphila* and improve our ability to predict cobamide-dependent physiology in other bacteria.

Results

A. muciniphila strain Muc^T salvages Cbi to produce pCbl

A. muciniphila strain Muc^T lacks most of the genes required for cobamide synthesis and does not produce cobamides *de novo* (34), but it is predicted to be capable of Cbi salvaging (33). To test this prediction, we extracted corrinoids from *A. muciniphila* cultured with and without Cbi and analyzed the corrinoid composition of the samples by high-performance liquid chromatography (HPLC). When cultured without Cbi, no corrinoids were detected in the extractions (Fig. 2A). However, when Cbi was added to the growth medium, a cobamide with the same retention time and nearly identical UV-Vis spectrum to pCbl was detected by HPLC (Fig. 2A). Mass spectrometry (MS) analysis corroborated that this cobamide is pCbl (Fig. S1).

A. muciniphila strain Muc^T does not show cobamide selectivity

Having established that *A. muciniphila* strain Muc^T cannot synthesize cobamides without the addition of a precursor, we next examined which cobamides it is capable of using by measuring growth in the presence of various cobamides under a cobamide-requiring condition. A homolog of the cobamide-dependent methionine synthase MetH is encoded in the genome of *A. muciniphila*. Because *A. muciniphila* lacks a homolog of the cobamide-independent methionine synthase MetE, growth in methionine-deplete medium is expected to require cobamide addition. We found this to be the case, as the addition of Cbi or any of the seven cobamides tested was necessary to support growth of *A. muciniphila* (Fig. 2B). Surprisingly, however, *A. muciniphila* shows essentially no cobamide selectivity, with less than twofold variation in the cobamide concentrations resulting in half-maximal growth (EC_{50}) (Fig. 2B).

A. muciniphila strain Muc^T remodels cobamides to pCbl

The ability of all of the tested cobamides to support nearly identical growth of *A. muciniphila* could be due to promiscuity in its cobamide-dependent methionine synthase. Alternatively, *A. muciniphila* could remodel cobamides, despite the absence of a homolog of *cbiZ* in its genome. If cobamide remodeling occurs in *A. muciniphila*, exogenously supplied cobamides will be altered by the bacterium. Therefore, we extracted corrinoids from *A. muciniphila* cultures supplemented with Cbl, [Cre]Cba or [MeAde]Cba to determine whether the added cobamides could be recovered. HPLC analysis revealed that none of the Cbl or [Cre]Cba, and only half of the [MeAde]Cba, remained in the extractions. This loss of the added cobamide coincided with the appearance of a new cobamide that co-eluted with pCbl (Fig. 2C). MS analysis confirmed that this cobamide is indeed pCbl (Fig. S2). These results demonstrate that *A. muciniphila* remodels cobamides to pCbl.

Identification and characterization of a novel cobamide remodeling enzyme in *A. muciniphila*

The identification of cobamide remodeling activity despite the absence of a *cbiZ* homolog in the genome suggested that a novel enzyme capable of hydrolyzing cobamides is present in *A. muciniphila*. We reasoned that the gene encoding this enzyme could be located near the cobamide biosynthesis and salvaging genes *cobDQ*, *cbiB*, *cobT*, *cobS*, and *cobU,* some or all of which would be required for completion of the remodeling process. These five genes are found at a single locus in the *A. muciniphila* genome that also contains an ORF with unknown function, annotated as

Amuc_1679 (Fig. 3A). Amuc_1679 is predicted to encode a protein with a conserved $(\beta/\alpha)_8$ TIM barrel domain from the AP endonuclease 2 superfamily (pfam01261). This superfamily is composed of several enzymes including endonuclease IV, which hydrolyzes phosphodiester bonds at apurinic or apyrimidinic (AP) sites in DNA (60). The proximity of Amuc_1679 to genes involved in cobamide biosynthesis and the presence of a phosphodiester bond connecting the lower ligand to the aminopropanol linker suggested that Amuc_1679 might play a role in cobamide biology in *A. muciniphila*. Further, homologs of this gene in other bacteria are also found in loci containing similar cobamide biosynthesis enzymes (Fig. 3A, Fig. S3).

To determine whether Amuc_1679 encodes an enzyme that can hydrolyze cobamides, we overexpressed and purified Amuc_1679 with N-terminal hexahistidine (His_6) and maltose-binding protein (MBP) tags for analysis of its activity *in vitro* (Fig. S4A). First, we tested whether a new product was formed when the protein was incubated with coenzyme B_{12} (AdoCbl), an active cofactor form of Cbl. We observed complete conversion of AdoCbl to a new corrinoid compound in reactions performed under anaerobic conditions (Fig. 3B). MS analysis showed that two reaction products, cobinamide-phosphate (Cbi-P) and α-ribazole, were formed, indicating hydrolysis of the phosphoribosyl bond of AdoCbl (Fig. 3C, D). Notably, Amuc_1679 targets a bond distinct from the enzyme CbiZ (Fig. 1A) (59). In keeping with the tradition of naming cobamide biosynthesis and remodeling enzymes with the "Cbi" prefix, we henceforth refer to Amuc_1679 as CbiR.

We were able to monitor CbiR activity continuously by measuring the rate of decrease in absorbance at 534 nm (A_{534}), as the reaction is characterized by a change in the UV-Vis spectrum that reflects the loss of AdoCbl and formation of AdoCbi-P (Fig. 3E). With this method, we found that the reaction proceeded only in the absence of oxygen, and additionally that the reaction requires the reducing agent DTT and is inhibited by the metal chelator EDTA (Fig. S4B). Using the same method, we determined the reaction kinetics of His_6-MBP-CbiR under steady-state conditions at a range of AdoCbl concentrations (Fig. 3F). Based on a fit to the Michaelis-Menten model, the reaction of His_6-MBP-CbiR with AdoCbl exhibited a K_M and k_{cat} for AdoCbl of 194 μM and 6.5 min^{-1}, respectively. Similarly, His_6-MBP-CbiR hydrolyzes MeCbl, the active cofactor form used by MetH and other methyltransferases, to MeCbi-P (Fig. S4C), with comparable kinetic parameters (Fig. S4D), indicating that AdoCbl and MeCbl are equally suitable substrates for His_6-MBP-CbiR.

A. muciniphila CbiR can hydrolyze several different cobamides *in vitro*

To determine the substrate selectivity of CbiR, His_6-MBP-CbiR activity was measured *in vitro* with seven different cobamides. Adenosylated cobamides with purinyl or phenolyl lower ligands do not show UV-Vis spectra distinct from AdoCbi-P under the reaction conditions, and thus activity was measured by HPLC. Each cobamide was completely converted to AdoCbi-P following an 18 h incubation, demonstrating that all of the cobamides are substrates for CbiR (Fig. 4A). The specific activities of His_6-MBP-CbiR with each cobamide are similar, with 4-fold differences among the benzimidazolyl and purinyl cobamides and slightly higher activity with phenolyl cobamides (Fig. 4B). These specific activities are similar to, though slightly lower than that

previously reported for CbiZ with Ado-pCbl (70 nmol/mg/min, (59)), albeit under somewhat different reaction conditions.

Expression of *cbiR* in *E. coli* enables expanded cobamide use

Given that the product AdoCbi-P can be used as a precursor for construction of a different cobamide, we hypothesize that CbiR activity enables bacteria to remodel cobamides, and therefore to gain access to cobamides in the environment that they otherwise may not be able to use. Because methods for targeted inactivation of genes in *A. muciniphila* have not been established, we used engineered *E. coli* strains to test this hypothesis. Like *A. muciniphila* strain Muc^T, *E. coli* MG1655 cannot synthesize cobamides *de novo*, but its genome has the cobamide biosynthesis genes *cobT*, *cobS*, *cobU*, and *cobC* that should allow *E. coli* to convert AdoCbi-P into a cobamide (61). We first tested whether *A. muciniphila* CbiR is functional in *E. coli*. Indeed, expression of *cbiR* on a plasmid in a Δ*cobTSU* Δ*cobC* background results in the loss of added Cbl, pCbl, [MeAde]Cba, and [Cre]Cba and the formation of two new corrinoid compounds (Fig. 5A). One of the products co-elutes with AdoCbi-P (Fig. 5A), and MS analysis confirmed that the dominant ion matches the m/z expected for Cbi-P (Fig. S5A). The second product has an m/z consistent with Cbi (Fig. S5B), which likely forms intracellularly by hydrolysis of the phosphate group of AdoCbi-P. Neither product was detected in an *E. coli* strain containing the empty vector (Fig. 5A, dashed lines). Therefore, the activity of CbiR that we observed *in vitro* can be recapitulated in aerobically cultured *E. coli*.

Catabolism of ethanolamine in *E. coli* requires the cobamide-dependent enzyme ethanolamine ammonia lyase, which is capable of using Cbl as a cofactor but is not functional with [Cre]Cba (62). We took advantage of this selectivity to design a cobamide remodeling-dependent growth assay in *E. coli*. In minimal medium supplemented with [Cre]Cba and 5,6-dimethylbenzimidazole (DMB, the lower ligand of Cbl), with ethanolamine as the sole nitrogen source, *E. coli* should be able to grow only if it can remodel [Cre]Cba to Cbl. Cbl, as expected, promotes growth of *E. coli* under this condition regardless of whether *cbiR* is present (Fig. 5B). In contrast, when [Cre]Cba is added, growth is observed only in the strain expressing *cbiR*, suggesting that CbiR activity enables *E. coli* to convert [Cre]Cba into Cbl (Fig. 5B). A cobamide with a retention time and m/z matching that of Cbl was detected in a corrinoid extraction of *E. coli* grown with [Cre]Cba and DMB when expressing CbiR, confirming that cobamide remodeling to Cbl occurred (Fig. S6). These results demonstrate that expression of CbiR expands the range of cobamides accessible to *E. coli*, and suggests that cobamide remodeling may serve a similar purpose in *A. muciniphila*.

CbiR is a member of the AP endonuclease 2 superfamily and is found in diverse bacteria

Analysis of the sequence of CbiR revealed that it is not similar to CbiZ or *V. cholerae* CobS, the other enzymes known to have cobamide remodeling activity. Instead, CbiR is a member of the AP endonuclease 2 superfamily, which includes the enzymes endonuclease IV, 2-keto-myo-inositol dehydratase, xylose isomerase, and other sugar isomerases and epimerases. A phylogenetic tree of this superfamily shows that the CbiR homologs identified by a BLAST search that are encoded in

genomic loci containing cobamide biosynthesis genes form a single, distinct clade within the superfamily (Fig. 6A). Some of the biochemically characterized enzymes in the superfamily require metal cofactors for activity, and between one and three metal ions are found in nearly all X-ray crystal structures of enzymes from the superfamily (63-77). A metal cofactor may also be required for CbiR function; in addition to the inhibition by the metal chelator EDTA (Fig. S4B), CbiR homologs contain conserved His, Asp, and Glu residues that, in the characterized members of the superfamily, are involved in metal coordination (Fig. 6B). Furthermore, single mutations in many of these conserved residues in CbiR eliminated most or all of its AdoCbl hydrolysis activity when expressed in *E. coli* (Fig. 6C). These results demonstrate that CbiR shares both sequence and functional features common to the AP endonuclease 2 superfamily, and represents a new function within the superfamily.

Finally, we investigated the prevalence of CbiR across sequenced organisms by examining genomes with *cbiR* homologs. The *cbiR* gene commonly occurs in the *Akkermansia* genus, as a search of the 191 available genomes in the NCBI database found that 184 have *cbiR*. Additionally, 282 homologs of *A. muciniphila* CbiR with Expect values below 10^{-3} were identified by BLAST in the genomes of 275 bacterial and 1 archaeal taxa from diverse habitats including aquatic environments, sewage, digesters, oil spills, bioreactors, soil, and human and animal hosts (Table S1). While 76% are found in the phyla Chlorobi, Chloroflexi, and Proteobacteria, CbiR homologs are also found in 19 other phyla including six candidate phyla and two candidate divisions (Fig. S3, Table S1), with relatively few in the PVC superphylum, to which *A. muciniphila* belongs. Similar to *A. muciniphila*, 80% of the *cbiR* homologs are located adjacent to genes involved in cobamide biosynthesis (Table S1), suggesting that they, too, function in cobamide remodeling. It is therefore likely that cobamide remodeling initiated by CbiR occurs in diverse bacteria and environments.

Discussion

Cobamides are considered to be important modulators of mammalian gut ecosystems because they are involved in several metabolic pathways, their production is limited to a subset of prokaryotes, and their diverse structures are differentially accessible to different microbes (54, 78, 79). *A. muciniphila* has been shown to have positive effects on host metabolism, gut barrier function, and the inflammatory response (15-24), yet knowledge of its metabolic and ecological roles in the gut remains incomplete. Previous studies showed that *A. muciniphila* strain Muc^T is unable to produce cobamides *de novo* (34), but can use pCbl produced by *E. hallii* or externally supplied Cbl for propionate production (26, 34). Here, while investigating the cobamide metabolism of *A. muciniphila* strain Muc^T, we uncovered a novel cobamide remodeling activity and identified and characterized an enzyme capable of initiating this process, CbiR. This discovery adds new complexity to the understanding of the roles of *A. muciniphila* in the gut. Not only does *A. muciniphila* degrade mucin to provide nutrients to the gut microbiota (26, 27), but it is also capable of altering cobamide structure, potentially changing the cobamide composition of its environment.

As a member of the AP endonuclease 2 superfamily, CbiR likely contains a $(\beta/\alpha)_8$ TIM barrel domain (63-77), unlike the structures predicted for the CbiZ and CobS protein families (80). Thus, not only does CbiR catalyze a unique reaction, but it is also distinct from the other cobamide remodeling enzymes in sequence and likely in structure. Intriguingly, while CbiR differs in sequence from B_{12}-binding domains in cobamide-dependent enzymes, the substrate-binding domains of many cobamide-dependent enzymes are comprised of a $(\beta/\alpha)_8$ TIM barrel structure, with the C-terminal face interacting with the cobamide cofactor (81-91). Given that CbiR is predicted to have a similar fold, it is possible that cobamide binding in CbiR and in these cobamide-dependent enzymes shares common features. The yet to be discovered enzyme responsible for remodeling in algae may also be unique, as neither a *P. lutheri* transcriptome (92) nor the *C. reinhardtii* genome contains homologs of CbiR or CbiZ. It therefore appears that cobamide remodeling mechanisms have independently evolved multiple times. Together with the multiple pathways that exist for cobamide biosynthesis, transport, and precursor salvaging (31, 33, 35, 93-99), the addition of CbiR to the growing list of enzymes involved in cobamide metabolism highlights the importance of cobamide physiology in the evolution of bacteria.

CbiR has unexpectedly promiscuous activity, hydrolyzing cobamides irrespective of their lower ligand structure. This differs from *R. sphaeroides* CbiZ, which does not hydrolyze Cbl *in vitro* (59), and *V. cholerae* CobS, which remodels neither Cbl nor [Cre]Cba (48). In these cases, the cobamide remodeling pathway does not act on a cobamide(s) that can function in its organism's metabolism. In contrast, *A. muciniphila* CbiR readily hydrolyzes pCbl, which functions as a cofactor for methionine synthesis and propionate metabolism in *A. muciniphila* and is the product of Cbi salvaging and cobamide remodeling in the bacterium itself. Thus, it is unclear how *A. muciniphila* prevents CbiR from continuing to hydrolyze pCbl after it is formed via cobamide remodeling. It is possible that pCbl is sequestered intracellularly by binding to MetH or other cobamide-dependent enzymes. Alternatively, CbiR activity could be coupled to cobamide uptake, as has been suggested for CbiZ (59, 100). Indeed, similar to some *cbiZ* homologs, 25% of *cbiR* homologs are located adjacent to genes for putative transport proteins, including in *A. muciniphila* strain Muc^T (Fig. S3). Remodeling in *D. mccartyi* strain 195 shows similar substrate promiscuity to *A. muciniphila* in the ability to act on numerous, structurally diverse cobamides (57), but the molecular basis of this promiscuity is unclear because its genome carries seven *cbiZ* homologs, none of which has been biochemically characterized. Aside from its activity on pCbl, the broad substrate range of CbiR may benefit *A. muciniphila* by enabling the bacterium to utilize a greater number of the cobamides present in its environment.

The discovery that *A. muciniphila* remodels cobamides leads us to reexamine its ecological roles in the gut. CbiR is found in all of the 75 recently sequenced *A. muciniphila* strains from the human and mouse gut, including the 26 strains that contain the *de novo* cobamide biosynthesis pathway (34, 101). Thus, like cobamide-dependent metabolism (34), cobamide remodeling appears to be nearly universal in *A. muciniphila*. Further, the role of *A. muciniphila* in the gut may be flexible, ranging from producing cobamides *de novo* to remodeling cobamides produced by other microbes, depending on which strains inhabit an individual. Notably, the end product of cobamide remodeling in *A. muciniphila*, pCbl, was the third most abundant corrinoid detected in the human

gut in a study of human subjects residing at a single geographic location (42). Interestingly, that study also presented evidence that cobamide remodeling occurs in the human gut, as individuals supplemented with high levels of Cbl showed transiently increased levels of Cbi and the specific purinyl and phenolyl cobamides that were present in the gut prior to Cbl supplementation. It is possible that *A. muciniphila* is involved in this remodeling activity and contributes to the pool of pCbl in the gut. This, in turn, could modulate the growth or metabolism of other cobamide-requiring bacteria that rely on particular cobamides for their metabolic needs. CbiR may therefore not only expand access to the cobamides available to *A. muciniphila*, but also affect those accessible to other bacteria in the gut. Further, homologs of CbiR are found in at least 276 other microbial taxa and may function similarly in these microbes that inhabit diverse environments. The addition of CbiR to the cobamide remodeling enzymes that have been characterized to date – CbiZ, certain CobS homologs, and the enzyme(s) responsible for remodeling in algae – suggests that cobamide remodeling is more widespread than previously thought.

Materials and Methods

Media and growth conditions

Akkermansia muciniphila strain Muc^T (DSM 22959, ATCC BAA-835) was cultivated at 37°C in a vinyl anaerobic chamber (Coy Laboratory Products Inc) under an atmosphere of approximately 10% CO_2, 3% H_2, and 87% N_2. A synthetic version of a basal mucin-based medium, in which mucin was replaced by soy peptone (16 g/L), L-threonine (4 g/L), glucose (2 g/L) and N-acetylglucosamine (2 g/L), was supplemented with 1% noble agar and used as a solid medium (21). This synthetic medium also contained L-methionine (125 mg/L) and omitted rumen fluid. M8 defined medium developed by Tramontano *et al.* (102) was used for liquid culturing. We found that the concentration of mucin in this medium (0.5%) was able to abrogate the requirement of methionine addition to the medium for *A. muciniphila* growth. Lowering the concentration to 0.25% resulted in methionine-deplete conditions for *A. muciniphila* and supplementation with methionine or cobamides restored robust growth. This mucin concentration was used for the MetH-dependent growth assays. However, batch to batch variations were seen with mucin such that media that supported robust growth while remaining methionine-deplete could not always be achieved. Cobamides were omitted from all growth media except when specified. For MetH-dependent growth assays, methionine was omitted from M8 medium.

Escherichia coli was cultured at 37°C with aeration in LB medium for cloning, protein expression, and assessing CbiR hydrolytic activity. Ethanolamine-based growth experiments used medium from Scarlett and Turner (38), with B_{12} omitted. Media were supplemented with antibiotics at the following concentrations when necessary: kanamycin, 25 mg/L (pETmini); ampicillin, 100 mg/L (pET-His$_6$-MBP) and chloramphenicol, 20 mg/L (pLysS).

For all *in vivo* experiments involving corrinoids, culture media were supplemented with cyanylated cobamides or $(CN)_2$Cbi, which are adenosylated following uptake into cells.

Genetic and molecular cloning techniques

The entire *A. muciniphila cbiR* open reading frame, except the start codon, was cloned into a modified pET16b vector (103) with N-terminal His_6 and MBP tags added for protein purification. For analysis of CbiR activity in *E. coli*, a minimized 3 kb derivative of pET28a (pETmini) containing the Kan^R marker, pBR322 origin, and *rop* gene was used. A constitutive promoter (BBa_J23100, iGEM) and RBS (BBa_B0034) were inserted into the vector for expression in *E. coli* MG1655-based strains (complete sequence: TTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGCGAATTCATACGACTCACTATAA AAGAGGAGAAA) and *A. muciniphila cbiR* was cloned downstream. Site-directed mutations were introduced into *cbiR* by PCR. All cloning was done by Gibson Assembly with *E. coli* XL1-Blue cells (104).

Construction of the *E. coli* MG1655 Δ*cobTSU* Δ*cobC* strain was accomplished using λ red-based recombination (105) and phage P1 transduction (106). An MG1655 Δ*cobTSU*::Kan^R operon deletion was constructed by λ red-based recombination. The Δ*cobC*::Kan^R allele was transduced into MG1655 via P1 transduction from *E. coli* strain JW0633-1, which was obtained from the Keio collection (107). Kan^R cassettes were removed by recombination of the flanking FRT sites as described (105).

Chemical reagents

Porcine gastric mucin was purchased from MilliporeSigma (M1778). AdoCbl (coenzyme B_{12}), MeCbl, CNCbl, and $(CN)_2$Cbi were purchased from MilliporeSigma.

Cobamide synthesis, adenosylation, and quantification

All other cyanylated cobamides used in the study were purified from bacterial cultures and cobamides were adenosylated and purified as previously described (50, 108). Cyanylated and adenosylated cobamides were quantified as previously described (50). MeCbl was quantified using an extinction coefficient of $\varepsilon_{519} = 8.7\ mM^{-1}\ cm^{-1}$ (109). AdoCbi-P and MeCbi-P were quantified using the dicyanylated corrinoid extinction coefficient $\varepsilon_{580} = 10.1\ mM^{-1}\ cm^{-1}$ following conversion to $(CN)_2$Cbi-P by incubation with 10 mM KCN in the presence of light (110).

A. muciniphila MetH-dependent growth assay

A. muciniphila was pre-cultured for 48 hours in M8 medium supplemented with 125 mg/L methionine. Cells were pelleted by centrifugation, washed twice with PBS, and diluted into 80 µL M8 medium to an OD_{600} of 0.02 in a 384-well plate (Nunc) with varying concentrations of cobamides and Cbi. The wells were sealed (ThermalSeal RTS™, Excel Scientific) and the plate was incubated at 37°C in a BioTek Epoch 2 microplate reader. OD_{600} was measured at regular intervals during growth.

E. coli ethanolamine-dependent growth assay

E. coli was pre-cultured 16 h in ethanolamine medium supplemented with 0.02% ammonium chloride. Cells were pelleted by centrifugation, washed three times with 0.85% NaCl, and diluted to an OD_{600} of 0.025 in 200 µL ethanolamine medium with the specified cobamide additions in a 96-well plate (Corning). The wells were sealed (Breathe-Easy, Diversified Biotech) and OD_{600} was monitored at 37°C in a BioTek Synergy 2 microplate reader with shaking.

Protein expression and purification

His_6-MBP-CbiR was expressed in *E. coli* Rosetta(DE3) pLysS. Cells were grown to an OD_{600} of 0.4 at 37°C and expression was induced with 1 mM IPTG for 6 h at 30°C. Cells were lysed by sonication in 20 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.4, with 0.5 mM PMSF, 1 µg/mL leupeptin, 1 µg/mL pepstatin, and 1 mg/mL lysozyme. The protein was purified from the clarified lysate using HisPur Ni-NTA resin (Thermo Scientific) and eluted with 250 mM imidazole. Purified protein was dialyzed into 25 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 10% glycerol and stored at -80°C.

His_6-MBP-CbiR *in vitro* reactions

Due to light sensitivity of the compounds, all work involving adenosylated cobamides or MeCbl was performed in the dark or under red or dim white light. Unless specified, the *in vitro* reactions were performed anaerobically at 37°C in a vinyl anaerobic chamber with an atmosphere as described above. The components of the reactions were 50 mM Tris buffer, 0.3 µM purified His_6-MBP-CbiR, 1 mM DTT, and variable concentrations of a cobamide. To prepare the Tris buffer, Tris base was dissolved and equilibrated within the anaerobic chamber. Prior to each experiment, the pH was adjusted with NaOH to account for acidification by the CO_2 present in the atmosphere of the chamber. The pH was adjusted to 8.8-8.9 at room temperature (approximately 24°C), corresponding to a predicted pH of 8.45-8.55 at 37°C. Protein concentration was determined by absorbance at 280 nm (A_{280}).

A BioTek Epoch 2 microplate reader and half-area UV-Star® 96-well microplates (Greiner Bio-One) were used for assays monitoring the reaction by absorbance. For these assays, separate 2X solutions of AdoCbl and His_6-MBP-CbiR were prepared in 50 mM Tris buffer and 1 mM DTT. A frozen aliquot of His_6-MBP-CbiR was thawed inside the anaerobic chamber prior to dilution. The 2X AdoCbl solution was pre-incubated at 37°C for 60 min, while the 2X CbiR solution was pre-incubated at 37°C for 20 min. 60 µL each of 2X AdoCbl and 2X His_6-MBP-CbiR were then mixed in a 96-well plate, with 100 µL transferred to a new well prior to measurements in the plate reader. Assays with MeCbl were prepared similarly.

Absorbances over time at 534 and 527 nm were used to monitor the conversion of AdoCbl to AdoCbi-P and MeCbl to MeCbi-P, respectively. To enable conversion of A_{534} into moles of AdoCbl and A_{527} into moles of MeCbl, the extinction coefficients of AdoCbl, MeCbl, and purified

AdoCbi-P and MeCbi-P were determined in 50 mM Tris buffer, pH 8.8, 1 mM DTT: ε_{534} (AdoCbl) = 7.8 mM^{-1} cm^{-1}, ε_{527} (MeCbl) = 8.0 mM^{-1} cm^{-1}, ε_{534} (AdoCbi-P) = 1.3 mM^{-1} cm^{-1}, ε_{527} (MeCbi-P) = 2.7 mM^{-1} cm^{-1}.

For reactions with adenosylated cobamides with different lower ligands monitored by HPLC, cobamides were mixed at 60 µM with 50 mM Tris buffer and 1 mM DTT and equilibrated to 37°C. His_6-MBP-CbiR was equilibrated to 37°C at 0.6 µM in 50 mM Tris buffer and 1 mM DTT. Each cobamide and His_6-MBP-CbiR were mixed in equal volume and incubated at 37°C. At three different timepoints, 100 µL of the reaction mix was removed and mixed with 5 µL of 600 mM EDTA to quench the reaction. The protein was removed from samples using Nanosep 10K centrifugal devices (Pall) prior to injection onto the HPLC. AdoCbi-P levels in the samples were quantified by HPLC by comparing to a standard curve generated with known quantities of purified AdoCbi-P. For reactions involving incubations of 4-18 h, initial equilibration at 37°C was not performed.

Corrinoid extraction

Cbi salvaging and cobamide remodeling were assessed in *A. muciniphila* by cultivating in M8 medium supplemented with 10 nM Cbi or cobamide, respectively, for 72 h. CbiR cobamide hydrolytic activity with different cobamides in *E. coli* was monitored using the MG1655 Δ*cobTSU* Δ*cobC* strain cultivated in LB medium supplemented with 10 nM cobamide for 16 h. Cobamide remodeling in *E. coli* MG1655 was assessed by cultivating in ethanolamine medium supplemented with 100 nM [Cre]Cba and 1 µM DMB for 94 h. CbiR mutants were analyzed in *E. coli* MG1655 Δ*cobTSU* Δ*cobC* by culturing in LB medium supplemented with 75 nM Cbl for 20 h.

Cyanation of corrinoids extracted from cells for Figures 2A, 2C, S1, and S2 was performed as previously described (57), with 5,000 corrinoid molar equivalents of KCN added. For extractions of adenosylated corrinoids (Figures 5A, 6C, S5, and S6), cell lysis was performed similarly, with KCN omitted; following removal of cellular debris by centrifugation, deionized water was added to the supernatant to decrease the methanol concentration to 10%. Solid phase extraction of cyanylated and adenosylated corrinoids with Sep-Pak C18 cartridges (Waters) was performed as described (37). Samples were dried, resuspended in 200 µL deionized water (pH 7), and filtered with 0.45 µm filters (Millex-HV, Millipore) or Nanosep 10K centrifugal devices prior to analysis by HPLC. For extractions involving adenosylated cobamides, all steps were performed in the dark or under red or dim white light.

HPLC and MS analysis

Corrinoids were analyzed on an Agilent 1200 series HPLC equipped with a diode array detector. For experiments in Figures 2A, 2C, 3B, 4A, and 4B, an Agilent Zorbax SB-Aq column (5 µm, 4.6 x 150 mm) was used as previously described (method 2, (58)). For experiments in Figures 5A, 6C, and S6A, an Agilent Zorbax Eclipse Plus C18 column (5 µm, 9.4 x 250 mm) was used with the following method: Solvent A, 0.1% formic acid in deionized water; Solvent B, 0.1% formic acid

in methanol; 2 mL/min at 30°C; 18% solvent B for 2.5 min followed by a linear gradient of 18 to 60% solvent B over 28.5 min.

An Agilent 1260 series fraction collector was used for HPLC purification of corrinoids and CbiR reaction products. The purification of CN-pCbl from *A. muciniphila* was performed using the Zorbax SB-Aq column with the method described above. Purification of AdoCbi-P and α-ribazole from the hydrolysis of AdoCbl by His_6-MBP-CbiR was performed in two steps. AdoCbi-P and α-ribazole were first separated and collected on a Zorbax Eclipse XDB-C18 column (5 µm, 4.6 x 150 mm) using the following method: Solvent A, 10 mM ammonium acetate pH 6.5; Solvent B, 100% methanol; 1 mL/min at 30°C; 0% B for 2 min followed by a linear gradient of 0 to 15% solvent B over 1.5 min, 15 to 35% over 6.5 min, 35 to 70% over 2 min, and 70 to 100% over 2 min. Each compound was subsequently run and collected on the Zorbax SB-Aq column with the method described above. The purification of MeCbi-P from the *in vitro* hydrolysis of MeCbl was performed using the Zorbax SB-Aq column with the method above. The purification of adenosylated hydrolysis products of CbiR from *E. coli* was performed with two rounds of collection; AdoCbi-P and AdoCbi were first separated and collected on the Zorbax Eclipse Plus C18 column with the method above, and then each compound was run and collected on the Zorbax SB-Aq column using the method described. AdoCbl remodeled from [Cre]Cba in *E. coli* was purified using the Zorbax Eclipse Plus C18 column with the method described above. Collected compounds were de-salted with Sep-Pak C18 cartridges.

MS analysis was performed on a Bruker Linear Iontrap Quadrupole coupled to a Fourier Transform Ion Cyclotron (LTQ-FT) mass spectrometer at the QB3/Chemistry MS Facility (UC Berkeley). Prior to MS analysis, the purified adenosylated and methylated corrinoids were exposed to light to remove the adenosyl and methyl upper ligands, respectively.

Phylogenetic analysis

282 homologs of *A. muciniphila* strain Muc^T CbiR were identified by BLAST with Expect values lower than 10^{-3} (Table S1; sequences from other strains of *A. muciniphila* are excluded). A subset of 203 sequences with Expect values lower than 10^{-14} and whose encoding genes were located adjacent to and in the same orientation as a cobamide biosynthesis gene(s) were chosen for phylogenetic analysis with the AP endonuclease 2 superfamily (pfam 01261) (Table S1). Sequences were clustered at 0.95 using CD-HIT to reduce the CbiR homolog sequence set by removing subspecies sequence diversity (111). This final set of 178 sequences and *A. muciniphila* CbiR were used to infer a phylogenetic tree with experimentally characterized members of the AP endonuclease 2 superfamily (Table S2). The sequences were aligned with MAFFT (112) and positions with 95% or greater gaps were removed by trimAl (113). A maximum likelihood tree presented in Figure 6A was inferred from this alignment using IQ-TREE v1.6.12 (114) with 1,500 ultrafast bootstraps and visualized in iTOL (115).

Acknowledgements

We thank Amanda Shelton, Amrita Hazra, Kristopher Kennedy, Per Malkus, Raphael Valdivia, and Sebastian Gude for helpful discussions and Kristopher Kennedy and Sebastian Gude for critical reading of the manuscript. We are grateful to Kristopher Kennedy and Sebastian Gude for construction of the pETmini plasmid and *E. coli* MG1655 Δ*cobTSU* Δ*cobC* strain, Nina Kirmiz and Gilberto Flores for reagents, and Rita Nichiporuk at the UC Berkeley QB3 Mass Spectrometry Facility for analysis of MS samples. This work was funded by National Institutes of Health Grant R01GM114535.

References

1. Cani PD, Van Hul M, Lefort C, Depommier C, Rastelli M, Everard A. 2019. Microbial regulation of organismal energy homeostasis. Nat Metab 1:34-46.
2. Cryan JF, O'Riordan KJ, Cowan CSM, Sandhu KV, Bastiaanssen TFS, Boehme M, Codagnone MG, Cussotto S, Fulling C, Golubeva AV, Guzzetta KE, Jaggar M, Long-Smith CM, Lyte JM, Martin JA, Molinero-Perez A, Moloney G, Morelli E, Morillas E, O'Connor R, Cruz-Pereira JS, Peterson VL, Rea K, Ritz NL, Sherwin E, Spichak S, Teichman EM, van de Wouw M, Ventura-Silva AP, Wallace-Fitzsimons SE, Hyland N, Clarke G, Dinan TG. 2019. The Microbiota-Gut-Brain Axis. Physiol Rev 99:1877-2013.
3. Hooper LV, Littman DR, Macpherson AJ. 2012. Interactions between the microbiota and the immune system. Science 336:1268-73.
4. Libertucci J, Young VB. 2019. The role of the microbiota in infectious diseases. Nat Microbiol 4:35-45.
5. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148:1258-70.
6. Dinan TG, Cryan JF. 2017. The Microbiome-Gut-Brain Axis in Health and Disease. Gastroenterol Clin North Am 46:77-89.
7. Durack J, Lynch SV. 2019. The gut microbiome: Relationships with disease and opportunities for therapy. J Exp Med 216:20-40.
8. Honda K, Littman DR. 2016. The microbiota in adaptive immune homeostasis and disease. Nature 535:75-84.
9. Kho ZY, Lal SK. 2018. The Human Gut Microbiome - A Potential Controller of Wellness and Disease. Front Microbiol 9:1835.
10. Miquel S, Martín R, Rossi O, Bermúdez-Humarán LG, Chatel JM, Sokol H, Thomas M, Wells JM, Langella P. 2013. *Faecalibacterium prausnitzii* and human intestinal health. Curr Opin Microbiol 16:255-61.
11. O'Callaghan A, van Sinderen D. 2016. Bifidobacteria and Their Role as Members of the Human Gut Microbiota. Front Microbiol 7:925.
12. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G, Grangette C, Vasquez N, Pochart P, Trugnan G, Thomas G, Blottière HM, Doré J, Marteau P, Seksik P, Langella P. 2008. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A 105:16731-6.
13. Derrien M, Belzer C, de Vos WM. 2017. *Akkermansia muciniphila* and its role in regulating host functions. Microb Pathog 106:171-181.
14. Zhai Q, Feng S, Arjan N, Chen W. 2019. A next generation probiotic, *Akkermansia muciniphila*. Crit Rev Food Sci Nutr 59:3227-3236.

15. Alam A, Leoni G, Quiros M, Wu H, Desai C, Nishio H, Jones RM, Nusrat A, Neish AS. 2016. The microenvironment of injured murine gut elicits a local pro-restitutive microbiota. Nat Microbiol 1:15021.
16. Derrien M, Van Baarlen P, Hooiveld G, Norin E, Müller M, de Vos WM. 2011. Modulation of Mucosal Immune Response, Tolerance, and Proliferation in Mice Colonized by the Mucin-Degrader *Akkermansia muciniphila*. Front Microbiol 2:166.
17. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, Guiot Y, Derrien M, Muccioli GG, Delzenne NM, de Vos WM, Cani PD. 2013. Cross-talk between *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. Proc Natl Acad Sci U S A 110:9066-71.
18. Grander C, Adolph TE, Wieser V, Lowe P, Wrzosek L, Gyongyosi B, Ward DV, Grabherr F, Gerner RR, Pfister A, Enrich B, Ciocan D, Macheiner S, Mayr L, Drach M, Moser P, Moschen AR, Perlemuter G, Szabo G, Cassard AM, Tilg H. 2018. Recovery of ethanol-induced *Akkermansia muciniphila* depletion ameliorates alcoholic liver disease. Gut 67:891-901.
19. Li J, Lin S, Vanhoutte PM, Woo CW, Xu A. 2016. *Akkermansia Muciniphila* Protects Against Atherosclerosis by Preventing Metabolic Endotoxemia-Induced Inflammation in *Apoe*-/- Mice. Circulation 133:2434-46.
20. Ottman N, Reunanen J, Meijerink M, Pietilä TE, Kainulainen V, Klievink J, Huuskonen L, Aalvink S, Skurnik M, Boeren S, Satokari R, Mercenier A, Palva A, Smidt H, de Vos WM, Belzer C. 2017. Pili-like proteins of *Akkermansia muciniphila* modulate host immune responses and gut barrier function. PLoS One 12:e0173004.
21. Plovier H, Everard A, Druart C, Depommier C, Van Hul M, Geurts L, Chilloux J, Ottman N, Duparc T, Lichtenstein L, Myridakis A, Delzenne NM, Klievink J, Bhattacharjee A, van der Ark KC, Aalvink S, Martinez LO, Dumas ME, Maiter D, Loumaye A, Hermans MP, Thissen JP, Belzer C, de Vos WM, Cani PD. 2017. A purified membrane protein from *Akkermansia muciniphila* or the pasteurized bacterium improves metabolism in obese and diabetic mice. Nat Med 23:107-113.
22. Reunanen J, Kainulainen V, Huuskonen L, Ottman N, Belzer C, Huhtinen H, de Vos WM, Satokari R. 2015. *Akkermansia muciniphila* Adheres to Enterocytes and Strengthens the Integrity of the Epithelial Cell Layer. Appl Environ Microbiol 81:3655-62.
23. Shin NR, Lee JC, Lee HY, Kim MS, Whon TW, Lee MS, Bae JW. 2014. An increase in the *Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. Gut 63:727-35.
24. Wu W, Lv L, Shi D, Ye J, Fang D, Guo F, Li Y, He X, Li L. 2017. Protective Effect of *Akkermansia muciniphila* against Immune-Mediated Liver Injury in a Mouse Model. Front Microbiol 8:1804.
25. Derrien M, Vaughan EE, Plugge CM, de Vos WM. 2004. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. Int J Syst Evol Microbiol 54:1469-1476.
26. Belzer C, Chia LW, Aalvink S, Chamlagain B, Piironen V, Knol J, de Vos WM. 2017. Microbial Metabolic Networks at the Mucus Layer Lead to Diet-Independent Butyrate and Vitamin B_{12} Production by Intestinal Symbionts. mBio 8:e00770-17.
27. Chia LW, Hornung BVH, Aalvink S, Schaap PJ, de Vos WM, Knol J, Belzer C. 2018. Deciphering the trophic interaction between *Akkermansia muciniphila* and the butyrogenic

gut commensal *Anaerostipes caccae* using a metatranscriptomic approach. Antonie Van Leeuwenhoek 111:859-873.

28. Corrêa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MA. 2016. Regulation of immune cell function by short-chain fatty acids. Clin Transl Immunology 5:e73.
29. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. 2016. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. Cell 165:1332-1345.
30. Louis P, Flint HJ. 2017. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol 19:29-41.
31. Roth JR, Lawrence JG, Bobik TA. 1996. Cobalamin (coenzyme B_{12}): Synthesis and biological significance. Annual Review of Microbiology 50:137-181.
32. Banerjee R, Ragsdale SW. 2003. The many faces of vitamin B_{12}: catalysis by cobalamin-dependent enzymes. Annu Rev Biochem 72:209-47.
33. Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR, Taga ME. 2019. Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. ISME J 13:789-804.
34. Kirmiz N, Galindo K, Cross KL, Luna E, Rhoades N, Podar M, Flores GE. 2020. Comparative Genomics Guides Elucidation of Vitamin B_{12} Biosynthesis in Novel Human-Associated *Akkermansia* Strains. Appl Environ Microbiol 86:e02117–e02119.
35. Escalante-Semerena JC. 2007. Conversion of cobinamide into adenosylcobamide in bacteria and archaea. J Bacteriol 189:4555-60.
36. Baker KA, Perego M. 2011. Transcription antitermination by a phosphorylated response regulator and cobalamin-dependent termination at a B_{12} riboswitch contribute to ethanolamine utilization in *Enterococcus faecalis*. J Bacteriol 193:2575-86.
37. Degnan PH, Barry NA, Mok KC, Taga ME, Goodman AL. 2014. Human gut microbes use multiple transporters to distinguish vitamin B_{12} analogs and compete in the gut. Cell Host Microbe 15:47-57.
38. Scarlett FA, Turner JM. 1976. Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B_{12}-dependent ethanolamine ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*. J Gen Microbiol 95:173-6.
39. Shelton AN, Lyu X, Taga ME. 2020. Flexible Cobamide Metabolism in *Clostridioides* (*Clostridium*) *difficile* 630 D*erm*. J Bacteriol 202:e00584-19.
40. Varel VH, Bryant MP. 1974. Nutritional features of *Bacteroides fragilis* subsp. *fragilis*. Appl Microbiol 28:251-7.
41. Renz P. 1999. Biosynthesis of the 5,6-dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids, p 557-566. *In* Banerjee R (ed), Chemistry and Biochemistry of B12. John Wiley & Sons, Inc., New York.
42. Allen RH, Stabler SP. 2008. Identification and quantitation of cobalamin and cobalamin analogues in human feces. American Journal of Clinical Nutrition 87:1324-1335.
43. Girard CL, Berthiaume R, Stabler SP, Allen RH. 2009. Identification of cobalamin and cobalamin analogues along the gastrointestinal tract of dairy cows. Arch Anim Nutr 63:379-88.
44. Girard CL, Santschi DE, Stabler SP, Allen RH. 2009. Apparent ruminal synthesis and intestinal disappearance of vitamin B_{12} and its analogs in dairy cows. J Dairy Sci 92:4524-9.

45. Men Y, Seth EC, Yi S, Crofts TS, Allen RH, Taga ME, Alvarez-Cohen L. 2015. Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities. Environ Microbiol 17:4873-84.
46. Keller S, Kunze C, Bommer M, Paetz C, Menezes RC, Svatoš A, Dobbek H, Schubert T. 2018. Selective utilization of benzimidazolyl-norcobamides as cofactors by the tetrachloroethene reductive dehalogenase of *Sulfurospirillum multivorans*. J Bacteriol 200:e00584-17.
47. Ma AT, Beld J, Brahamsha B. 2017. An Amoebal Grazer of Cyanobacteria Requires Cobalamin Produced by Heterotrophic Bacteria. Appl Environ Microbiol 83:e00035-00017.
48. Ma AT, Tyrell B, Beld J. 2020. Specificity of cobamide remodeling, uptake and utilization in *Vibrio cholerae*. Mol Microbiol 113:89-102.
49. Schubert T, von Reuss SH, Kunze C, Paetz C, Kruse S, Brand-Schon P, Nelly AM, Nuske J, Diekert G. 2019. Guided cobamide biosynthesis for heterologous production of reductive dehalogenases. Microb Biotechnol 12:346-359.
50. Sokolovskaya OM, Mok KC, Park JD, Tran JLA, Quanstrom KA, Taga ME. 2019. Cofactor Selectivity in Methylmalonyl Coenzyme A Mutase, a Model Cobamide-Dependent Enzyme. mBio 10:e01303-19.
51. Tanioka Y, Miyamoto E, Yabuta Y, Ohnishi K, Fujita T, Yamaji R, Misono H, Shigeoka S, Nakano Y, Inui H, Watanabe F. 2010. Methyladeninylcobamide functions as the cofactor of methionine synthase in a Cyanobacterium, *Spirulina platensis* NIES-39. FEBS Lett 584:3223-6.
52. Watanabe F, Nakano Y, Stupperich E. 1992. Different corrinoid specificities for cell growth and cobalamin uptake in *Euglena gracilis* Z. J Gen Microbiol 138:1807-1813.
53. Yan J, Bi M, Bourdon AK, Farmer AT, Wang PH, Molenda O, Quaile AT, Jiang N, Yang Y, Yin Y, Şimşir B, Campagna SR, Edwards EA, Löffler FE. 2018. Purinyl-cobamide is a native prosthetic group of reductive dehalogenases. Nat Chem Biol 14:8-14.
54. Sokolovskaya OM, Shelton AN, Taga ME. 2020. Sharing vitamins: Cobamides unveil microbial interactions. Science 369:eaba0165.
55. Helliwell KE, Lawrence AD, Holzer A, Kudahl UJ, Sasso S, Kräutler B, Scanlan DJ, Warren MJ, Smith AG. 2016. Cyanobacteria and Eukaryotic Algae Use Different Chemical Variants of Vitamin B_{12}. Curr Biol 26:999-1008.
56. Yan J, Şimşir B, Farmer AT, Bi M, Yang Y, Campagna SR, Löffler FE. 2016. The corrinoid cofactor of reductive dehalogenases affects dechlorination rates and extents in organohalide-respiring *Dehalococcoides mccartyi*. ISME J 10:1092-101.
57. Yi S, Seth EC, Men YJ, Stabler SP, Allen RH, Alvarez-Cohen L, Taga ME. 2012. Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in *Dehalococcoides mccartyi*. Appl Environ Microbiol 78:7745-52.
58. Mok KC, Taga ME. 2013. Growth inhibition of *Sporomusa ovata* by incorporation of benzimidazole bases into cobamides. J Bacteriol 195:1902-11.
59. Gray MJ, Escalante-Semerena JC. 2009. The cobinamide amidohydrolase (cobyric acid-forming) CbiZ enzyme: a critical activity of the cobamide remodelling system of *Rhodobacter sphaeroides*. Mol Microbiol 74:1198-210.
60. Daley JM, Zakaria C, Ramotar D. 2010. The endonuclease IV family of apurinic/apyrimidinic endonucleases. Mutat Res 705:217-27.

746 61. Lawrence JG, Roth JR. 1995. The cobalamin (coenzyme B_{12}) biosynthetic genes of
747 *Escherichia coli*. J Bacteriol 177:6371-80.
748 62. Poppe L, Bothe H, Bröker G, Buckel W, Stupperich E, Rétey J. 2000. Elucidation of the
749 coenzyme binding mode of further B_{12}-dependent enzymes using a base-off analogue of
750 coenzyme B_{12}. J Mol Catal 10:345-350.
751 63. Asano R, Ishikawa H, Nakane S, Nakagawa N, Kuramitsu S, Masui R. 2011. An additional
752 C-terminal loop in endonuclease IV, an apurinic/apyrimidinic endonuclease, controls
753 binding affinity to DNA. Acta Crystallogr D Biol Crystallogr 67:149-55.
754 64. Chan HC, Zhu Y, Hu Y, Ko TP, Huang CH, Ren F, Chen CC, Ma Y, Guo RT, Sun Y.
755 2012. Crystal structures of D-psicose 3-epimerase from *Clostridium cellulolyticum* H10
756 and its complex with ketohexose sugars. Protein Cell 3:123-31.
757 65. Collyer CA, Henrick K, Blow DM. 1990. Mechanism for aldose-ketose interconversion by
758 D-xylose isomerase involving ring opening followed by a 1,2-hydride shift. J Mol Biol
759 212:211-35.
760 66. Hosfield DJ, Guan Y, Haas BJ, Cunningham RP, Tainer JA. 1999. Structure of the DNA
761 repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic
762 sites and three-metal-ion catalysis. Cell 98:397-408.
763 67. Jenkins J, Janin J, Rey F, Chiadmi M, van Tilbeurgh H, Lasters I, De Maeyer M, Van Belle
764 D, Wodak SJ, Lauwereys M, et al. 1992. Protein engineering of xylose (glucose) isomerase
765 from *Actinoplanes missouriensis*. 1. Crystallography and site-directed mutagenesis of
766 metal binding sites. Biochemistry 31:5449-58.
767 68. Kim K, Kim HJ, Oh DK, Cha SS, Rhee S. 2006. Crystal structure of D-psicose 3-epimerase
768 from *Agrobacterium tumefaciens* and its complex with true substrate D-fructose: a pivotal
769 role of metal in catalysis, an active site for the non-phosphorylated substrate, and its
770 conformational changes. J Mol Biol 361:920-31.
771 69. Lavie A, Allen KN, Petsko GA, Ringe D. 1994. X-ray Crystallographic Structures of D-
772 Xylose Isomerase-Substrate Complexes Position the Substrate and Provide Evidence for
773 Metal Movement during Catalysis. Biochemistry 33:5469-5480.
774 70. Shi R, Pineda M, Ajamian E, Cui Q, Matte A, Cygler M. 2008. Structure of L-xylulose-5-
775 Phosphate 3-epimerase (UlaE) from the anaerobic L-ascorbate utilization pathway of
776 *Escherichia coli*: identification of a novel phosphate binding motif within a TIM barrel
777 fold. J Bacteriol 190:8137-44.
778 71. Tomanicek SJ, Hughes RC, Ng JD, Coates L. 2010. Structure of the endonuclease IV
779 homologue from *Thermotoga maritima* in the presence of active-site divalent metal ions.
780 Acta Crystallogr Sect F Struct Biol Cryst Commun 66:1003-12.
781 72. Uechi K, Sakuraba H, Yoshihara A, Morimoto K, Takata G. 2013. Structural insight into
782 L-ribulose 3-epimerase from *Mesorhizobium loti*. Acta Crystallogr D Biol Crystallogr
783 69:2330-9.
784 73. Whitlow M, Howard AJ, Finzel BC, Poulos TL, Winborne E, Gilliland GL. 1991. A metal-
785 mediated hydride shift mechanism for xylose isomerase based on the 1.6 A *Streptomyces*
786 *rubiginosus* structures with xylitol and D-xylose. Proteins 9:153-73.
787 74. Yoshida H, Yamada M, Nishitani T, Takada G, Izumori K, Kamitori S. 2007. Crystal
788 structures of D-tagatose 3-epimerase from *Pseudomonas cichorii* and its complexes with
789 D-tagatose and D-fructose. J Mol Biol 374:443-53.
790 75. Yoshida H, Yoshihara A, Gullapalli PK, Ohtani K, Akimitsu K, Izumori K, Kamitori S.
791 2018. X-ray structure of *Arthrobacter globiformis* M30 ketose 3-epimerase for the

production of D-allulose from D-fructose. Acta Crystallogr F Struct Biol Commun 74:669-676.

76. Zhang RG, Dementieva I, Duke N, Collart F, Quaite-Randall E, Alkire R, Dieckman L, Maltsev N, Korolev O, Joachimiak A. 2002. Crystal structure of *Bacillus subtilis* IolI shows endonuclase IV fold with altered Zn binding. Proteins 48:423-6.

77. Zhang W, Xu Y, Yan M, Li S, Wang H, Yang H, Zhou W, Rao Z. 2018. Crystal structure of the apurinic/apyrimidinic endonuclease IV from *Mycobacterium tuberculosis*. Biochem Biophys Res Commun 498:111-118.

78. Degnan PH, Taga ME, Goodman AL. 2014. Vitamin B_{12} as a modulator of gut microbial ecology. Cell Metab 20:769-778.

79. Rowley CA, Kendall MM. 2019. To B_{12} or not to B_{12}: Five questions on the role of cobalamin in host-microbial interactions. PLoS Pathog 15:e1007479.

80. Ovchinnikov S, Park H, Varghese N, Huang PS, Pavlopoulos GA, Kim DE, Kamisetty H, Kyrpides NC, Baker D. 2017. Protein structure determination using metagenome sequence data. Science 355:294-298.

81. Berkovitch F, Behshad E, Tang KH, Enns EA, Frey PA, Drennan CL. 2004. A locking mechanism preventing radical damage in the absence of substrate, as revealed by the x-ray structure of lysine 5,6-aminomutase. Proc Natl Acad Sci U S A 101:15870-5.

82. Doukov T, Seravalli J, Stezowski JJ, Ragsdale SW. 2000. Crystal structure of a methyltetrahydrofolate- and corrinoid-dependent methyltransferase. Structure 8:817-30.

83. Evans JC, Huddler DP, Hilgers MT, Romanchuk G, Matthews RG, Ludwig ML. 2004. Structures of the N-terminal modules imply large domain motions during catalysis by methionine synthase. Proc Natl Acad Sci U S A 101:3729-36.

84. Hagemeier CH, Krer M, Thauer RK, Warkentin E, Ermler U. 2006. Insight into the mechanism of biological methanol activation based on the crystal structure of the methanol-cobalamin methyltransferase complex. Proc Natl Acad Sci U S A 103:18917-22.

85. Mancia F, Keep NH, Nakagawa A, Leadlay PF, McSweeney S, Rasmussen B, Bösecke, P., Diat O, Evans PR. 1996. How coenzyme B_{12} radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. Structure 4:339-50.

86. Reitzer R, Gruber K, Jogl G, Wagner UG, Bothe H, Buckel W, Kratky C. 1999. Glutamate mutase from *Clostridium cochlearium*: the structure of a coenzyme B_{12}-dependent enzyme provides new mechanistic insights. Structure 7:891-902.

87. Shibata N, Masuda J, Tobimatsu T, Toraya T, Suto K, Morimoto Y, Yasuoka N. 1999. A new mode of B_{12} binding and the direct participation of a potassium ion in enzyme catalysis: X-ray structure of diol dehydratase. Structure 7:997-1008.

88. Shibata N, Tamagaki H, Hieda N, Akita K, Komori H, Shomura Y, Terawaki S, Mori K, Yasuoka N, Higuchi Y, Toraya T. 2010. Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B_{12} analogs and substrates. J Biol Chem 285:26484-93.

89. Svetlitchnaia T, Svetlitchnyi V, Meyer O, Dobbek H. 2006. Structural insights into methyltransfer reactions of a corrinoid iron-sulfur protein involved in acetyl-CoA synthesis. Proc Natl Acad Sci U S A 103:14331-6.

90. Wolthers KR, Levy C, Scrutton NS, Leys D. 2010. Large-scale domain dynamics and adenosylcobalamin reorientation orchestrate radical catalysis in ornithine 4,5-aminomutase. J Biol Chem 285:13942-50.

91. Yamanishi M, Yunoki M, Tobimatsu T, Sato H, Matsui J, Dokiya A, Iuchi Y, Oe K, Suto K, Shibata N, Morimoto Y, Yasuoka N, Toraya T. 2002. The crystal structure of coenzyme

B_{12}-dependent glycerol dehydratase in complex with cobalamin and propane-1,2-diol. Eur J Biochem 269:4484-94.

92. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ, Beszteri B, Bidle KD, Cameron CT, Campbell L, Caron DA, Cattolico RA, Collier JL, Coyne K, Davy SK, Deschamps P, Dyhrman ST, Edvardsen B, Gates RD, Gobler CJ, Greenwood SJ, Guida SM, Jacobi JL, Jakobsen KS, James ER, Jenkins B, John U, Johnson MD, Juhl AR, Kamp A, Katz LA, Kiene R, Kudryavtsev A, Leander BS, Lin S, Lovejoy C, Lynn D, Marchetti A, McManus G, Nedelcu AM, Menden-Deuer S, Miceli C, Mock T, Montresor M, Moran MA, Murray S, et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. PLoS Biol 12:e1001889.

93. Campbell GR, Taga ME, Mistry K, Lloret J, Anderson PJ, Roth JR, Walker GC. 2006. *Sinorhizobium meliloti bluB* is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B_{12}. Proc Natl Acad Sci U S A 103:4634-9.

94. Gopinath K, Venclovas C, Ioerger TR, Sacchettini JC, McKinney JD, Mizrahi V, Warner DF. 2013. A vitamin B_{12} transporter in *Mycobacterium tuberculosis*. Open Biol 3:120175.

95. Gray MJ, Escalante-Semerena JC. 2010. A new pathway for the synthesis of a-ribazole-phosphate in *Listeria innocua*. Mol Microbiol 77:1429-38.

96. Hazra AB, Han AW, Mehta AP, Mok KC, Osadchiy V, Begley TP, Taga ME. 2015. Anaerobic biosynthesis of the lower ligand of vitamin B_{12}. Proc Natl Acad Sci U S A 112:10792-7.

97. Rempel S, Gati C, Nijland M, Thangaratnarajah C, Karyolaimos A, de Gier JW, Guskov A, Slotboom DJ. 2020. A mycobacterial ABC transporter mediates the uptake of hydrophilic compounds. Nature 580:409-412.

98. Santos JA, Rempel S, Mous ST, Pereira CT, Ter Beek J, de Gier JW, Guskov A, Slotboom DJ. 2018. Functional and structural characterization of an ECF-type ABC transporter for vitamin B_{12}. Elife 7:e35828.

99. Wexler AG, Schofield WB, Degnan PH, Folta-Stogniew E, Barry NA, Goodman AL. 2018. Human gut *Bacteroides* capture vitamin B_{12} via cell surface-exposed lipoproteins. Elife 7:e37138.

100. Gray MJ, Tavares NK, Escalante-Semerena JC. 2008. The genome of *Rhodobacter sphaeroides* strain 2.4.1 encodes functional cobinamide salvaging systems of archaeal and bacterial origins. Mol Microbiol 70:824-36.

101. Guo X, Li S, Zhang J, Wu F, Li X, Wu D, Zhang M, Ou Z, Jie Z, Yan Q, Li P, Yi J, Peng Y. 2017. Genome sequencing of 39 *Akkermansia muciniphila* isolates reveals its population structure, genomic and functional diverisity, and global distribution in mammalian gut microbiotas. BMC Genomics 18:800.

102. Tramontano M, Andrejev S, Pruteanu M, Klünemann M, Kuhn M, Galardini M, Jouhten P, Zelezniak A, Zeller G, Bork P, Typas A, Patil KR. 2018. Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies. Nat Microbiol 3:514-522.

103. Kranzusch PJ, Lee ASY, Wilson SC, Solovykh MS, Vance RE, Berger JM, Doudna JA. 2014. Structure-guided reprogramming of human cGAS dinucleotide linkage specificity. Cell 158:1011-1021.

882 104. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. 2009.
883 Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods
884 6:343-5.
885 105. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in
886 *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640-5.
887 106. Silhavy TJ, Berman ML, Enquist LW. 1984. Experiments with Gene Fusions. Cold Spring
888 Harbor Laboratory Press, Cold Spring Harbor, NY).
889 107. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M,
890 Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene
891 knockout mutants: the Keio collection. Mol Syst Biol 2:2006 0008.
892 108. Crofts TS, Hazra AB, Tran JL, Sokolovskaya OM, Osadchiy V, Ad O, Pelton J, Bauer S,
893 Taga ME. 2014. Regiospecific formation of cobamide isomers is directed by CobT.
894 Biochemistry 53:7805-15.
895 109. Hogenkamp HPC. 1975. The chemistry of cobalamins and related compounds, p 55. *In*
896 Babior BM (ed), Cobalamin Biochemistry and Pathophysiology. John Wiley and Sons, Inc.
897 110. Stupperich E, Eisinger HJ, Kräutler B. 1988. Diversity of corrinoids in acetogenic bacteria.
898 *p*-cresolylcobamide from *Sporomusa ovata*, 5-methoxy-6-methylbenzimidazolylcobamide
899 from *Clostridium formicoaceticum* and vitamin B_{12} from *Acetobacterium woodii*. Eur J
900 Biochem 172:459-64.
901 111. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-
902 generation sequencing data. Bioinformatics 28:3150-2.
903 112. Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence
904 alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160-1166.
905 113. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
906 alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-3.
907 114. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective
908 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol
909 32:268-74.
910 115. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display
911 and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-5.

912

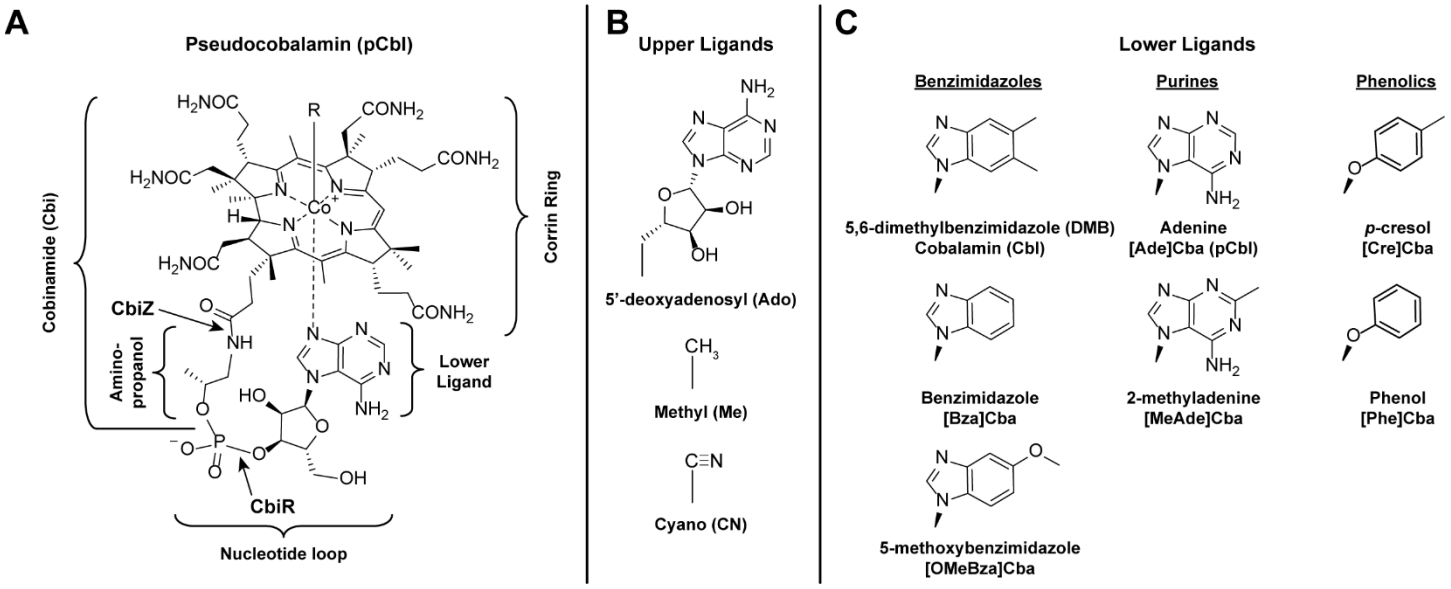

Figure 1. Cobamide structures.

A. Structure of pCbl. All cobamides are composed of a corrin ring containing a central cobalt ion and an upper (R) and lower ligand. In pCbl, the lower ligand is adenine. The lower ligand, together with the ribose and phosphate moieties, comprise the nucleotide loop, which is covalently attached to the corrin ring via an aminopropanol linker. The bonds hydrolyzed by the CbiZ amidohydrolase and the CbiR phosphodiesterase are indicated with arrows. The part of the molecule comprising Cbi is shown. Cobamides and their corrin-containing biosynthetic precursors and degradation products are together known as corrinoids.

B. Upper ligands (R) in cobamides, the catalytic center of the cofactor; prefixes used in the text to denote the upper ligand are shown in parentheses.

C. The three chemical classes of lower ligands present in cobamides. The structures of the seven cobamide lower ligands used in this study are shown. Names of the lower ligand base and abbreviations used for the corresponding cobamides are given below the structures.

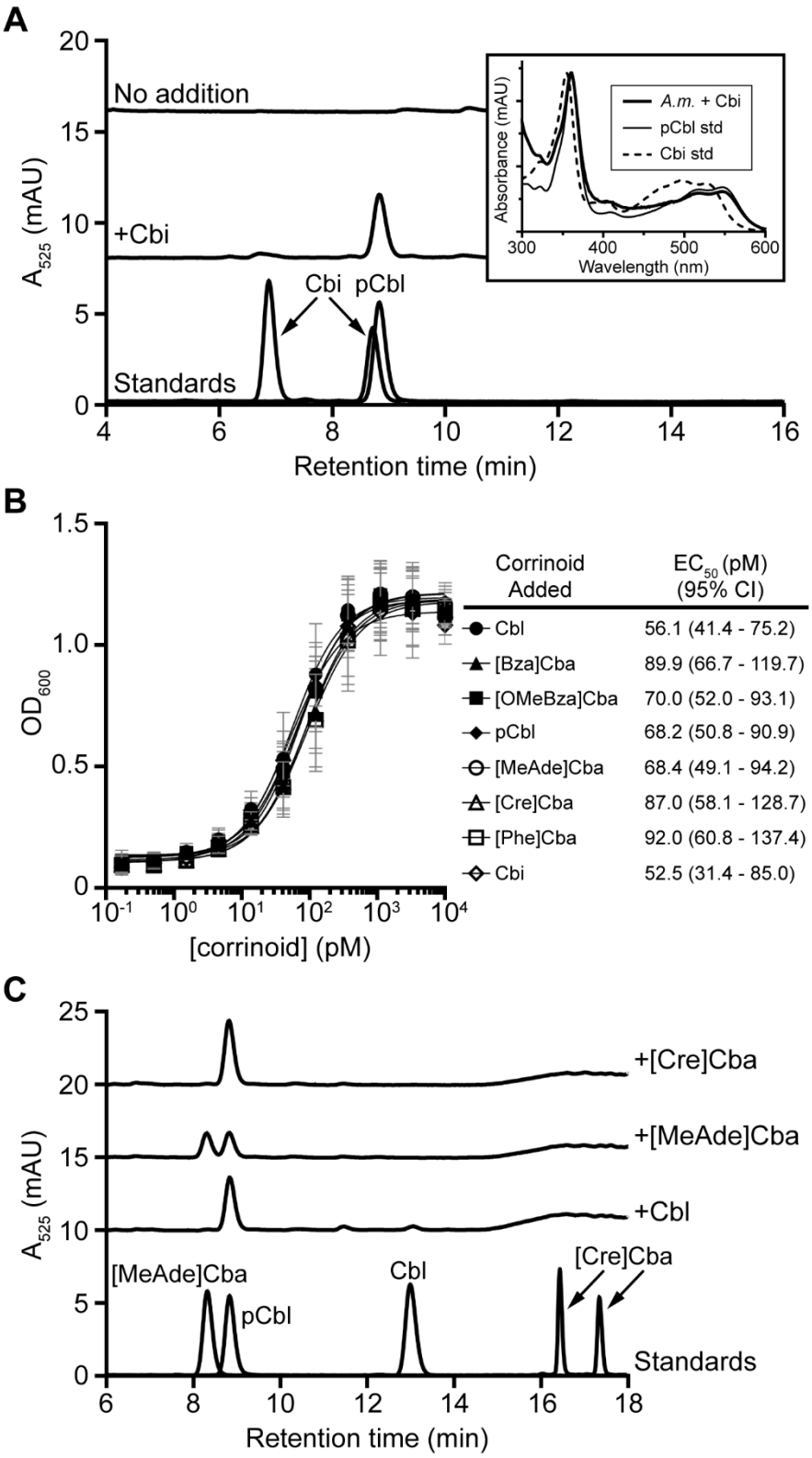

Figure 2. *A. muciniphila* strain Muc^T can salvage Cbi and remodel cobamides.

A. HPLC analysis of corrinoid extractions from *A. muciniphila* grown with or without 10 nM Cbi for 72 h shows that *A. muciniphila* can salvage Cbi. Standards for Cbi and pCbl are shown at the bottom. Comparison of UV-Vis spectra (inset) of the HPLC peaks at 8.8 min shows that the corrinoid produced by *A. muciniphila* (*A.m.*) grown with Cbi (thick line) is similar to a pCbl standard (std) (thin line) and not a Cbi standard (dashed line). Spectra were normalized to each other based on their maxima to aid comparison.

B. *A. muciniphila* growth under methionine-deplete conditions. OD_{600} of saturated cultures is shown after 29 h of growth with the indicated concentrations of each corrinoid. EC_{50} values and their 95% confidence intervals for each corrinoid are given in the table. Data points and error bars represent the mean and standard deviation, respectively, of three biological replicates. The results are representative of three independent experiments.

C. HPLC analysis of corrinoid extractions from *A. muciniphila* grown with 10 nM Cbl, [MeAde]Cba, or [Cre]Cba for 72 h shows that *A. muciniphila* remodels cobamides to pCbl. Cobamide standards are shown at the bottom.

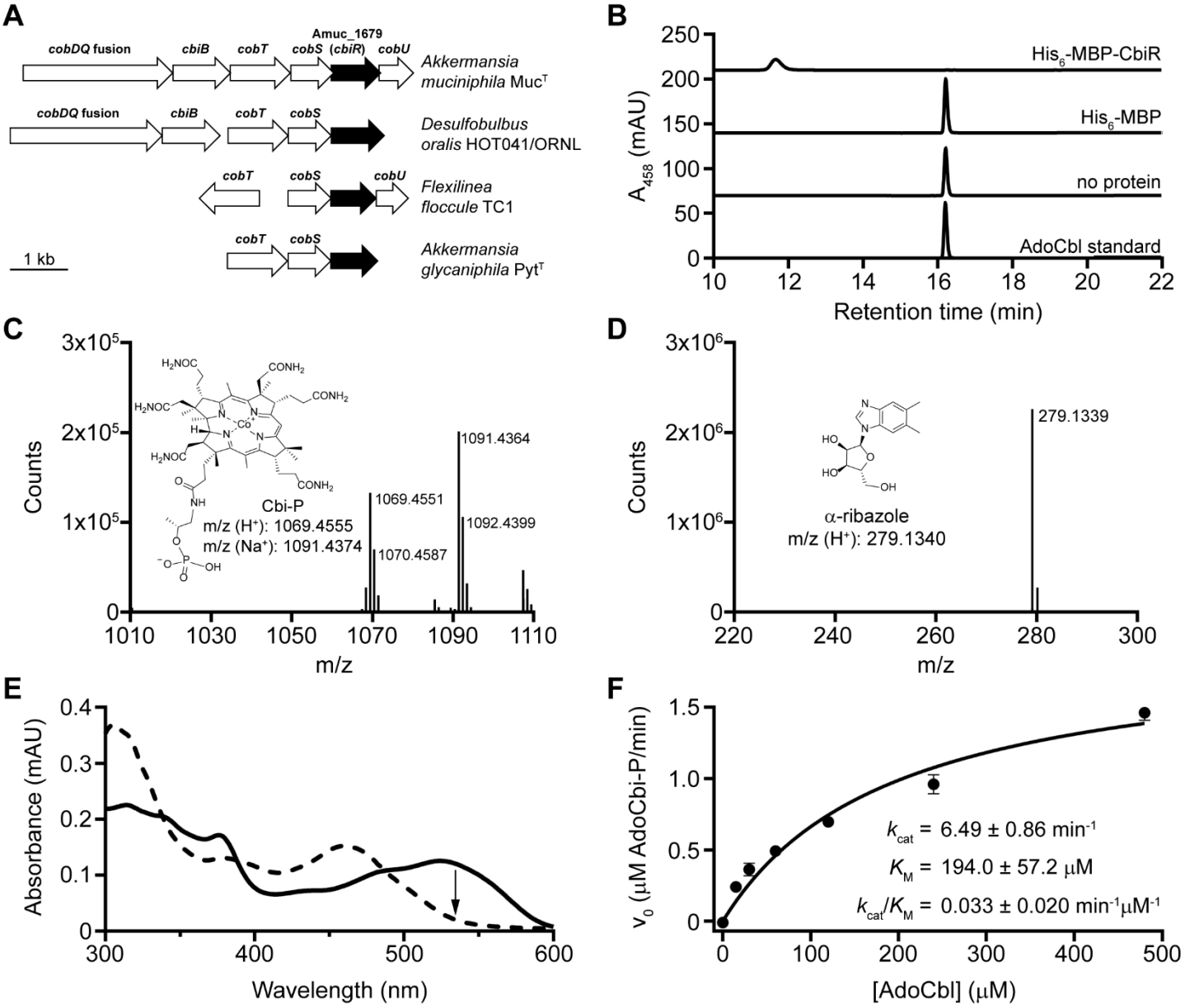

Figure 3. Purified CbiR hydrolyzes AdoCbl to form AdoCbi-P and α-ribazole *in vitro*.

A. *A. muciniphila* Amuc_1679 (*cbiR*) and homologs in other bacteria (black arrows) are located near cobamide biosynthesis genes (white arrows). An expanded genomic comparison is shown in Figure S3.

B. Purified CbiR converts AdoCbl to another corrinoid compound *in vitro* anaerobically. HPLC analysis of reactions containing 10 μM AdoCbl incubated for 4 h with 0.1 μM His$_6$-MBP-CbiR, 0.1 μM His$_6$-MBP, or no protein is shown. An AdoCbl standard is shown at bottom.

C. The corrinoid product of His$_6$-MBP-CbiR was purified by HPLC, exposed to light to remove the adenosyl upper ligand, and analyzed by MS. The structure and predicted m/z for Cbi-P are shown for comparison.

D. The second product of the *in vitro* reaction with His$_6$-MBP-CbiR and AdoCbl was purified by HPLC and analyzed by MS. The structure and predicted m/z for α-ribazole are shown for comparison.

E. Comparison of the UV-Vis spectra before (solid line) and after completion (dashed line) of the reaction of His$_6$-MBP-CbiR with 30 μM AdoCbl shows a decrease in absorbance at 534 nm (arrow).

F. Michaelis-Menten kinetic analysis of His_6-MBP-CbiR with AdoCbl. Reactions contained 0.3 μM His_6-MBP-CbiR. Points and error bars represent the mean and standard deviation, respectively. Kinetic constants were determined from two independent experiments, each with three technical replicates.

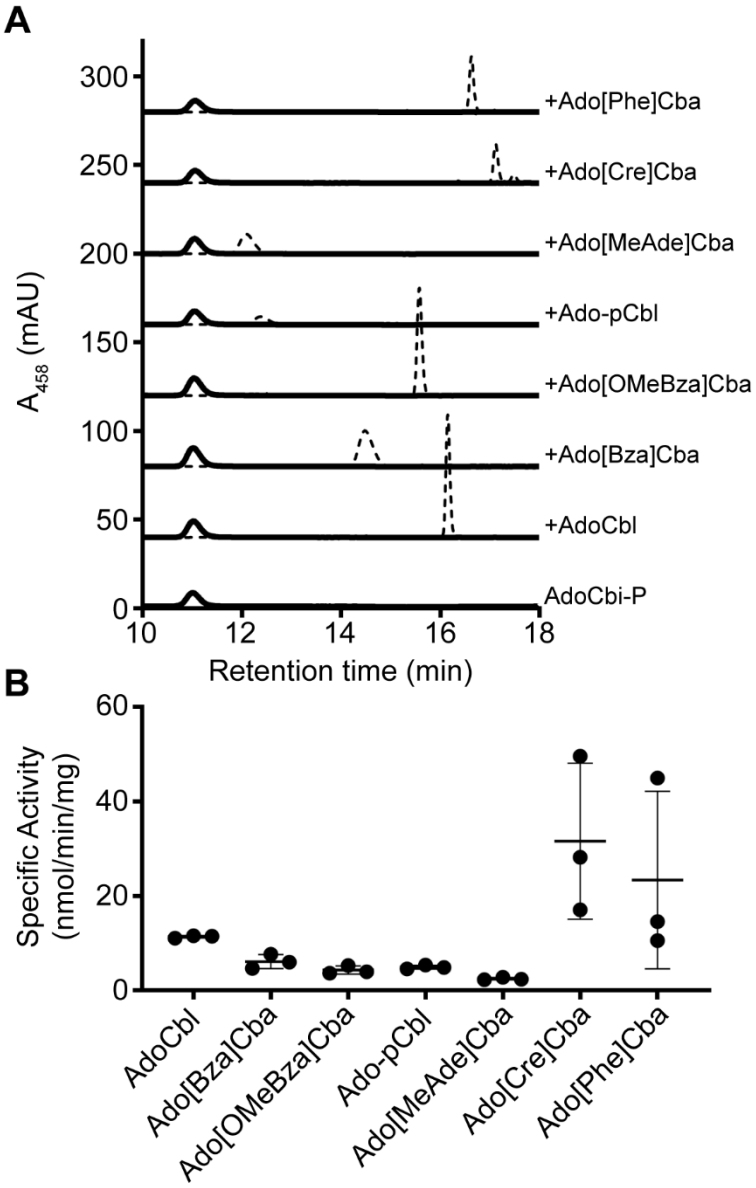

Figure 4. CbiR hydrolyzes many cobamides to form AdoCbi-P

A. HPLC analysis of *in vitro* reactions with different cobamides (10 μM), quenched after 18 h, is shown for reactions containing 0.1 μM His_6-MBP-CbiR (solid lines) or without enzyme (dashed lines). A sample of purified AdoCbi-P is shown at the bottom.

B. Specific activity of His_6-MBP-CbiR with different cobamide substrates. 0.3 μM His_6-MBP-CbiR was incubated with 30 μM of each cobamide individually and the rate of AdoCbi-P production was determined based on HPLC measurements at three time points. The lines represent the mean and standard deviation for three independent experiments.

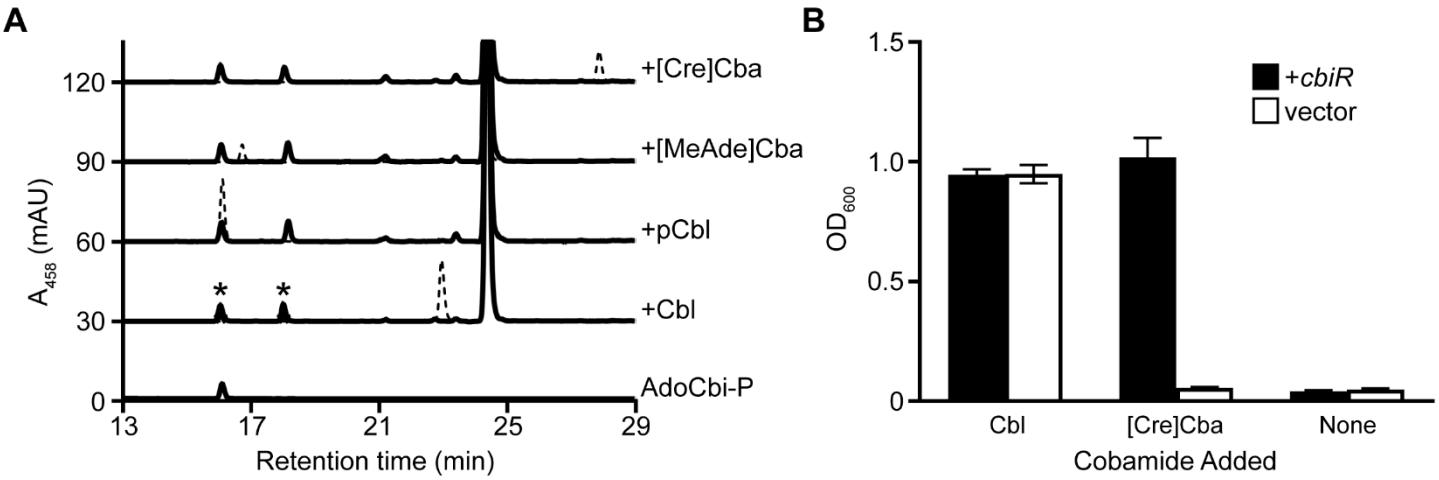

Figure 5. CbiR mediates cobamide remodeling in *E. coli*

A. CbiR hydrolyzes cobamides in *E. coli*. *A. muciniphila cbiR* was expressed in an *E. coli* strain with the *cobTSU* operon and *cobC* gene deleted to prevent modification of cobamide hydrolysis products. Corrinoid extractions of *E. coli* strains carrying pETmini-*cbiR* (solid lines) or the pETmini empty vector (dashed lines), grown with 10 nM Cbl, pCbl, [MeAde]Cba, or [Cre]Cba were analyzed by HPLC. A sample of purified AdoCbi-P is shown at the bottom. Corrinoids labeled with asterisks were purified for MS analysis (Fig. S5). The large peak at 24.5 min corresponds to a flavin that is present in all of the corrinoid extractions.

B. Expression of *A. muciniphila cbiR* enables growth of *E. coli* on ethanolamine. Wild type *E. coli* MG1655 harboring pETmini-*cbiR* (black bars) or the pETmini empty vector (white bars) was cultured in minimal medium containing ethanolamine as the sole nitrogen source and 1 μM DMB. Cultures were supplemented with 100 nM Cbl, [Cre]Cba, or no cobamide and OD_{600} measurements are shown after 72 h of growth. Bars and error bars represent the mean and standard deviation, respectively, of three biological replicates.

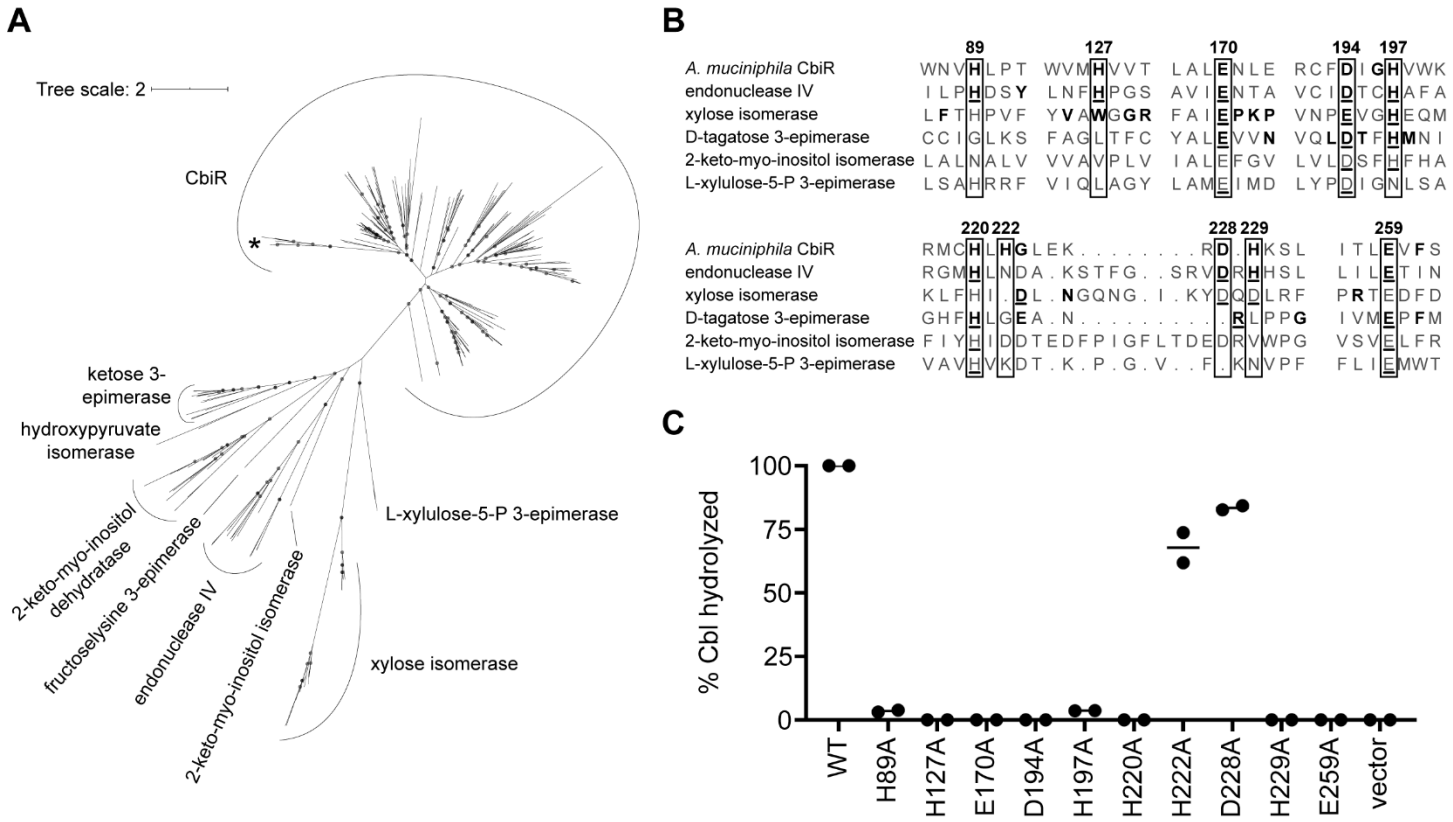

Figure 6. CbiR is a newly described member of the AP endonuclease 2 superfamily.

A. Maximum likelihood tree of CbiR homologs and members of the AP endonuclease 2 superfamily that have been experimentally characterized. *A. muciniphila* CbiR is indicated by an asterisk. CbiR homologs included in the tree were identified in a BLAST search queried with *A. muciniphila* CbiR with E values lower than 10^{-14} and encoded adjacent to and in the same orientation as one or more cobamide biosynthesis gene (Table S1). Superfamily member sequences are listed in Table S2. Gray circles overlaid on tree nodes represent bootstrap values of >95%. The scale bar corresponds to the average number of substitutions per site across the alignment.

B. Sequence alignment of regions containing highly conserved His, Asp, and Glu residues in *A. muciniphila* CbiR and representative sequences of biochemically and structurally characterized enzyme classes in the AP endonuclease 2 superfamily (*E. coli* endonuclease IV, *Streptomyces rubiginosus* xylose isomerase, *Pseudomonas cichorii* D-tagatose 3-epimerase, *Bacillus subtilis* 2-keto-myo-inositol isomerase, *E. coli* L-xylulose-5-P 3-epimerase). Numbers correspond to positions in *A. muciniphila* CbiR. For CbiR, endonuclease IV, xylose isomerase, and D-tagatose (ketose) 3-epimerase, bolded residues represent conserved amino acids in the enzyme classes. Underlined residues indicate amino acids in the X-ray crystal structures that interact with the metal cofactor(s), and with the substrate in the case of *P. cichorii* D-tagatose 3-epimerase.

C. Mutational analysis of *A. muciniphila* CbiR. Corrinoids were extracted from cultures of *E. coli* Δ*cobTSU* Δ*cobC* strains carrying pETmini-*cbiR* (WT), pETmini-*cbiR* with the specified alanine mutations, or the pETmini empty vector grown for 20 h with 75 nM Cbl and analyzed by HPLC. The y-axis represents the combined amount of AdoCbi-P and AdoCbi present out of the total adenosylated corrinoids extracted. Minimal amounts of CNCbl were detected in the mutant extractions and were excluded from the analysis. The total intracellular corrinoid content was similar between samples except for WT and D228A, which had 4- and 2.5-fold higher levels, respectively. Lines show the mean of the two independent experiments.

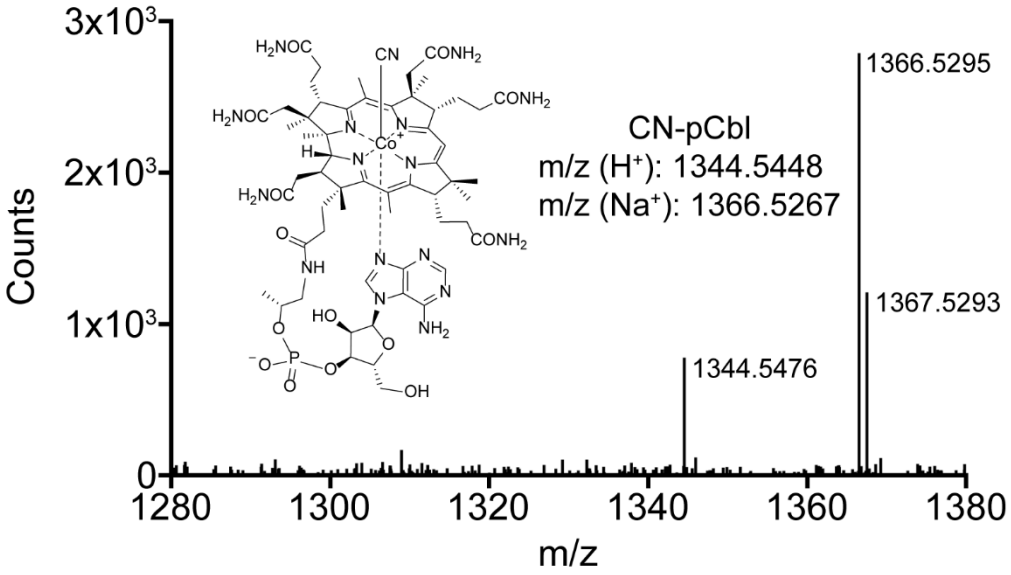

Figure S1. The corrinoid produced by *A. muciniphila* when grown with Cbi was purified by HPLC and analyzed by MS. The structure and predicted m/z for CN-pCbl are shown for comparison.

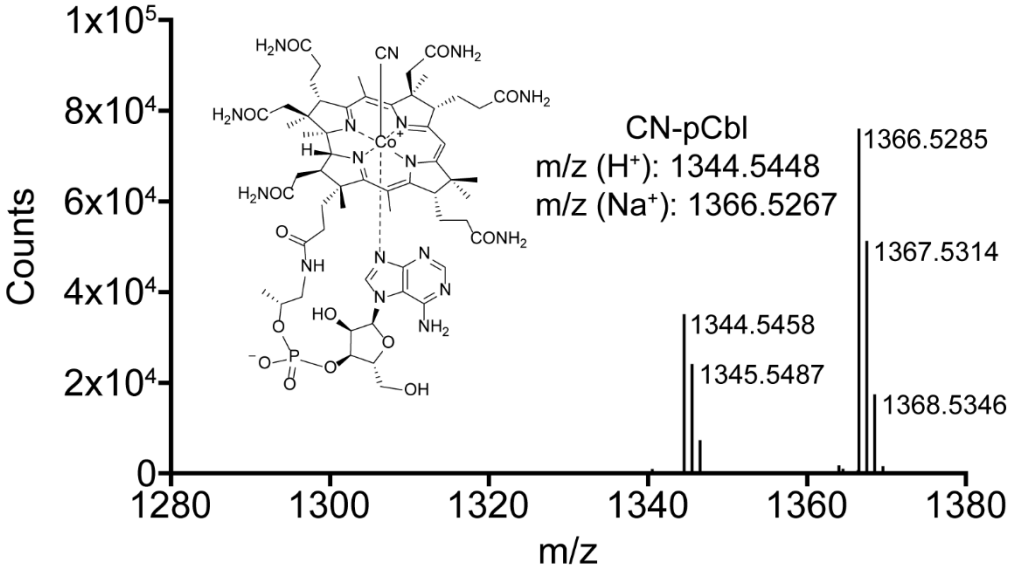

Figure S2. The corrinoid produced by *A. muciniphila* when grown with Cbl was purified by HPLC and analyzed by MS. The structure and predicted m/z for CN-pCbl are shown for comparison.

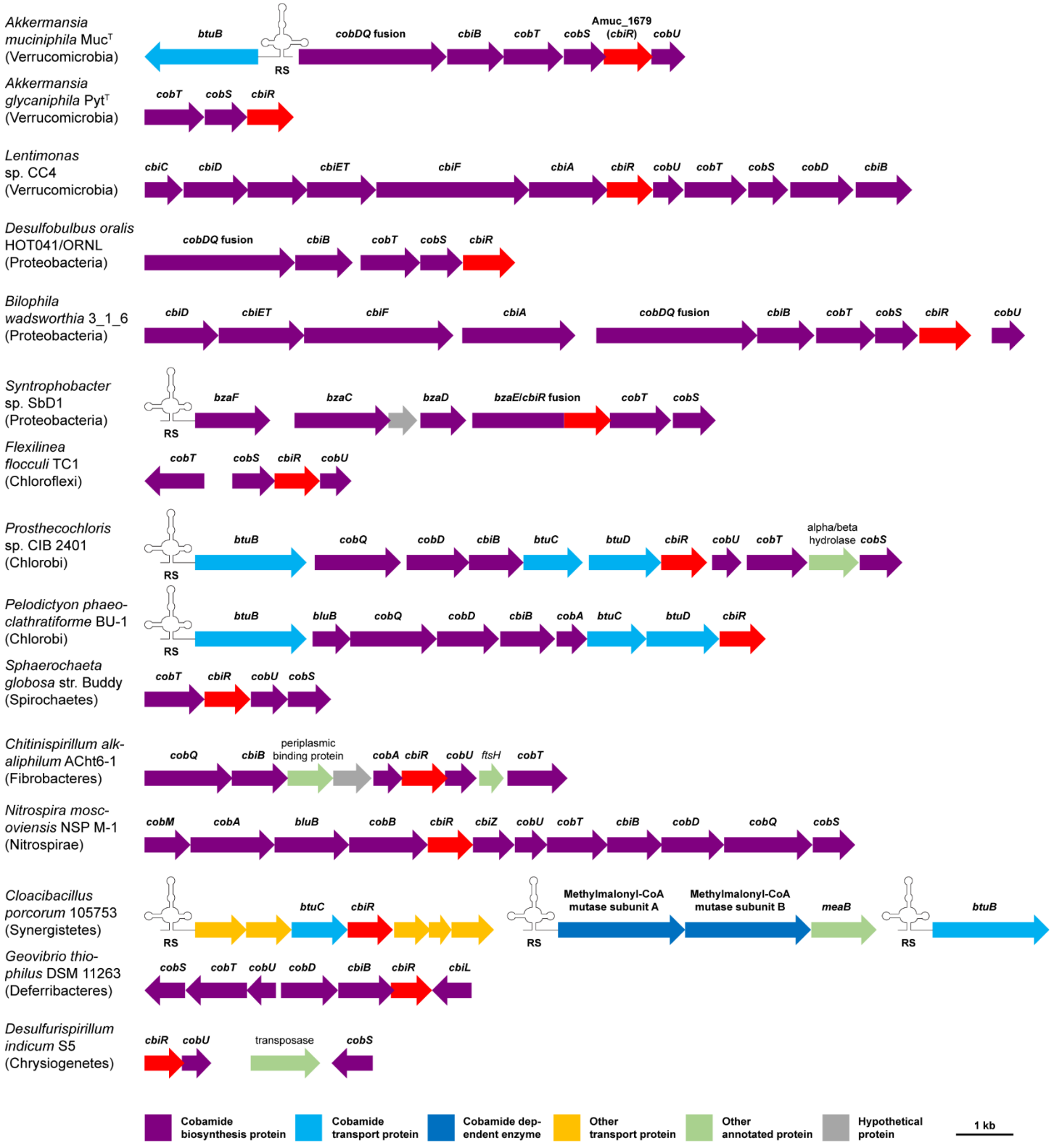

Figure S3. Expanded list of homologs of Amuc_1679 (*cbiR*, red arrows). Species and strain names are given, with phylum names in parentheses. RS denotes a predicted cobalamin riboswitch. The lengths and positions of the ORFs are drawn to scale.

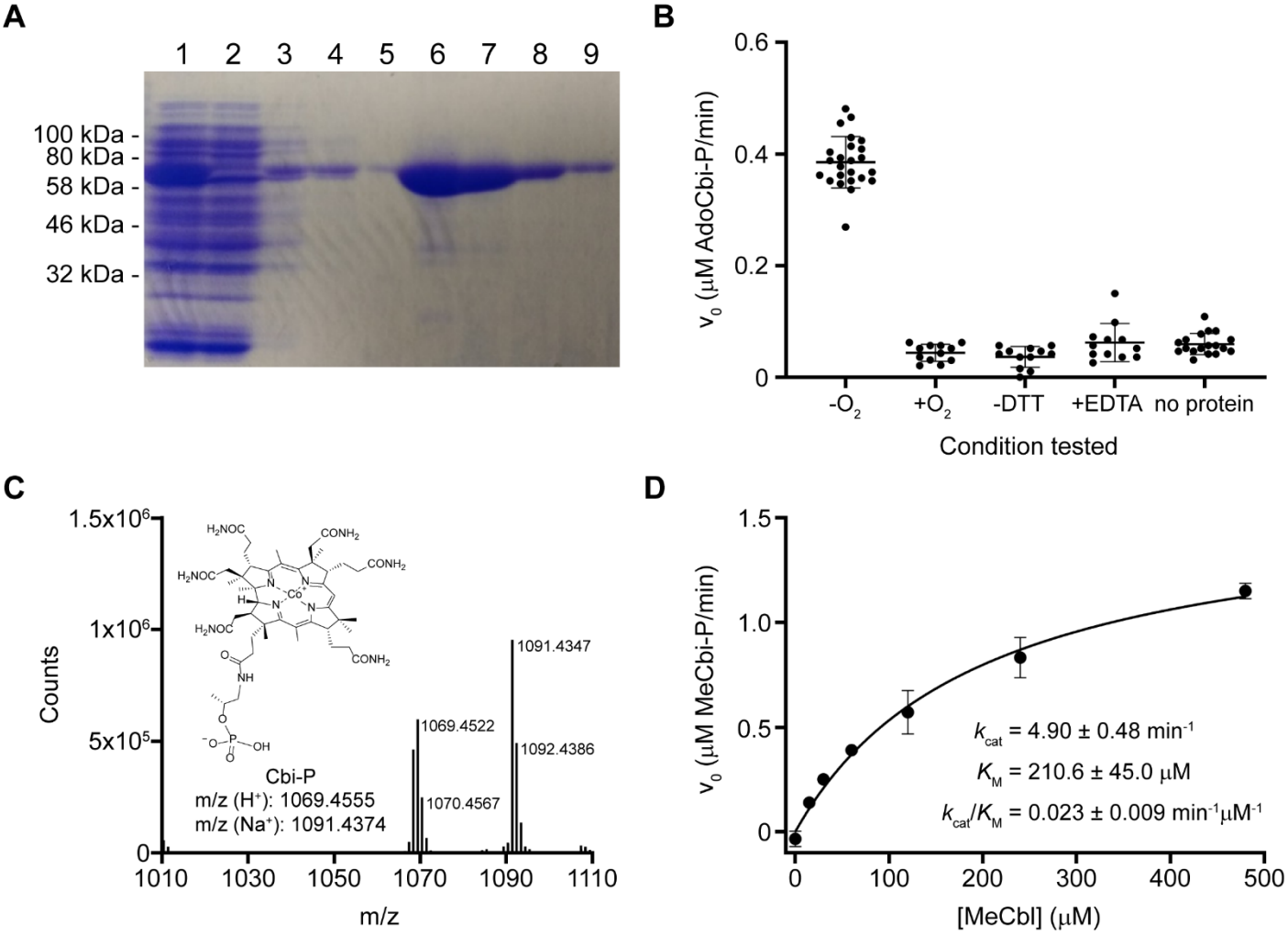

Figure S4. Biochemical characterization of His_6-MBP-CbiR.

A. SDS-PAGE gel of purification of His_6-MBP-CbiR with Ni-NTA resin. Lane 1: Cell lysate, Lane 2: Flowthrough, Lanes 3-5: Wash fractions, Lanes 6-9: Elution fractions. His_6-MBP-CbiR has a predicted molecular weight of 77 kDa.

B. *In vitro* characterization of His_6-MBP-CbiR. The reaction rates were determined for 0.3 μM His_6-MBP-CbiR incubated with 30 μM AdoCbl. The standard reaction mixture used throughout the manuscript (labeled as $-O_2$) contained 50 mM Tris, pH 8.45-8.55 and 1 mM DTT and was performed in an anaerobic chamber. Activity was also measured in ambient O_2 ($+O_2$), with DTT omitted (-DTT), with 3 μM EDTA added (+EDTA), and in the absence of His_6-MBP-CbiR (no protein). The reaction was monitored by measuring A_{534} over time. Lines and error bars show the mean and standard deviation, respectively.

C. The corrinoid product of His_6-MBP-CbiR incubated with MeCbl was purified by HPLC, exposed to light to remove the methyl upper ligand, and analyzed by MS. The structure and predicted m/z for Cbi-P are shown for comparison.

D. Michaelis-Menten kinetic analysis of His_6-MBP-CbiR with MeCbl. His_6-MBP-CbiR was tested at 0.3 μM and the reaction was monitored by measuring the decrease in absorbance at 527 nm (A_{527}). The experiment was performed twice with similar results. Data from a representative experiment with three technical replicates are shown. Kinetic constants were calculated using data from all six replicates.

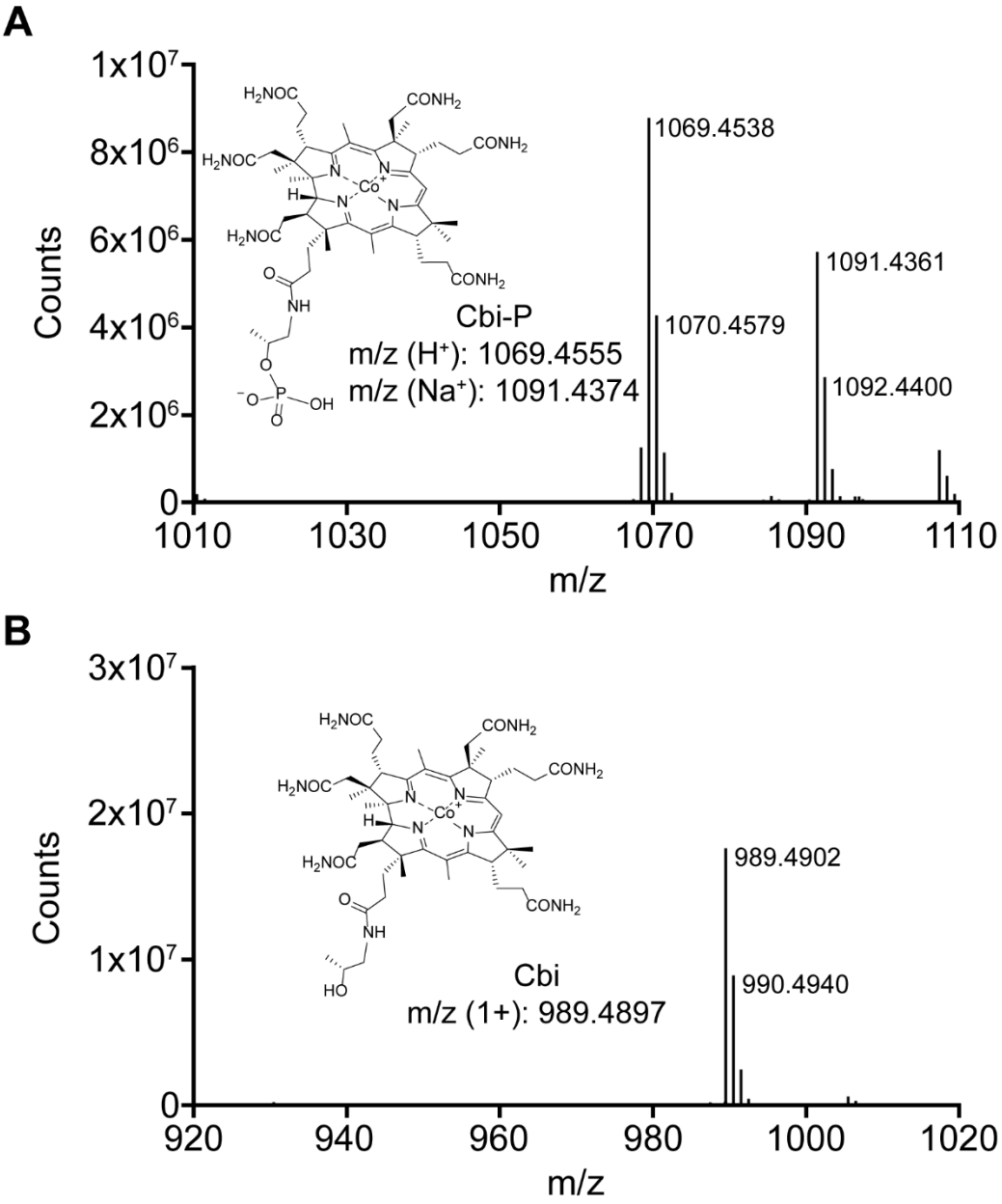

Figure S5. MS analysis of CbiR hydrolysis products extracted from *E. coli*.

MS analysis of the corrinoid products labeled with asterisks in Fig. 5A are shown for the peaks at A. 16 min and B. 18 min. The corrinoids were purified by HPLC and exposed to light to remove the adenosyl upper ligand prior to analysis by MS. The structures and predicted m/z for Cbi-P and Cbi are shown for comparison.

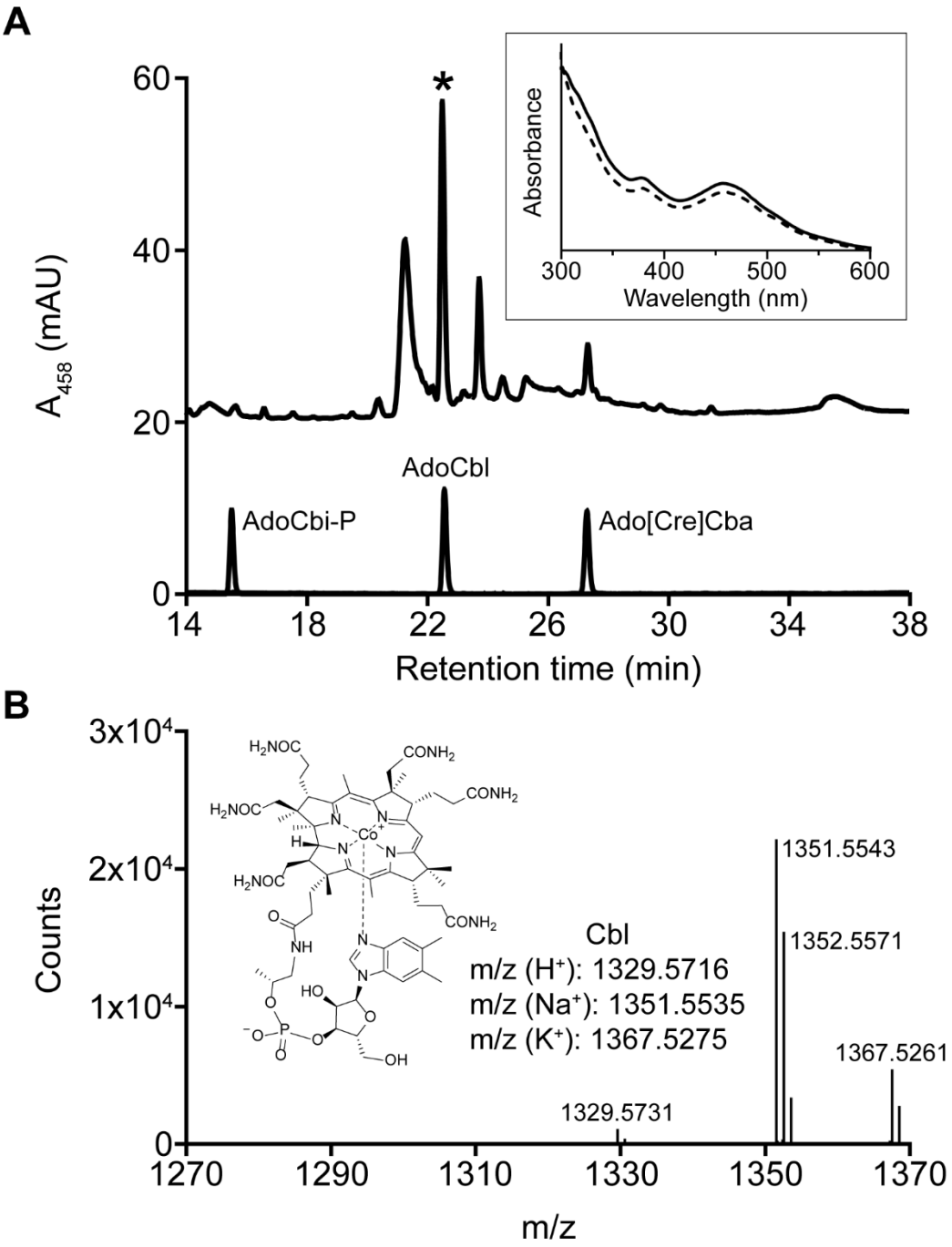

Figure S6. Cobamide remodeling of [Cre]Cba to Cbl in *E. coli* expressing *cbiR*.

A. Remodeling of [Cre]Cba to Cbl in *E. coli*. Wild type strain MG1655 containing pETmini-*cbiR* was grown in 1L minimal medium with ethanolamine supplemented with 100 nM [Cre]Cba and 1 μM DMB. Corrinoids were extracted and analyzed by HPLC. AdoCbl and Ado[Cre]Cba standards and purified AdoCbi-P are shown at the bottom. Comparison of the UV-Vis spectra of AdoCbl (dashed line) and the starred peak (solid line) are shown in the inset; spectra were normalized to each other at the local maxima at 458 nm to aid comparison. The spectrum of AdoCbl differs from that in Fig. 3E due to the acidic HPLC conditions.

B. MS analysis of the peak labeled with an asterisk in panel A. The corrinoid was purified by HPLC and exposed to light to remove the adenosyl upper ligand prior to analysis by MS. The structure and predicted m/z for Cbl are shown for comparison.

Table S1. Homologs of A. muciniphila CbiR

Organism	Phylum	Isolation Source	Accession number	E value	% Identity to *A. muciniphila* MucT CbiR	Adjacent to and in same orientation as cobamide biosynthesis gene(s)?
Akkermansia muciniphila strain MucT	Verrucomicrobia	human feces	WP_081429195.1		(100)	Y
Akkermansia glycaniphila strain Pyt	Verrucomicrobia	reticulated python feces	OCA02218.1	2.E-115	62	Y
Akkermansia sp. strain BIOML-A47	Verrucomicrobia	human feces	KAA3306648.1	2.E-174	84	Y
Akkermansia sp. strain BIOML-A60	Verrucomicrobia	human feces	KAA3165711.1	1.E-175	85	Y
Akkermansia sp. strain KLE1797	Verrucomicrobia	human feces	KXT49684.1	7.E-173	87	Y
Aminomonas paucivorans strain DSM 12260	Synergistetes	anaerobic lagoon of a dairy wastewater treatment plant (Columbia)	EFQ22869.1	9.E-15	28	N
anaerobic bacterium MO-CFX2	Chloroflexi	marine sediment (Shimokita Peninsula, Japan)	WP_162909697.1	5.E-27	31	Y
Anaerolineaceae bacterium isolate AS05jafATM_106	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	HHV05068.1	5.E-35	35	Y
Anaerolineaceae bacterium oral taxon 439 strain W11661	Chloroflexi	human subgingival dental plaque	AOH42763.1	7.E-35	35	Y
Anaerolineae bacterium isolate AS06rmzACSIP_331	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLF10836.1	4.E-24	30	Y
Anaerolineae bacterium isolate HyVt-475	Chloroflexi	rock sample from hydrothermal vent (Lau Basin, Tonga)	HFD38495.1	6.E-23	30	Y
Anaerolineae bacterium isolate Nak57	Chloroflexi	sulfidic hot spring (Azumino, Japan)	PWH15486.1	3.E-37	34	Y
Anaerolineae bacterium isolate SpSt-822	Chloroflexi	mud, sediment (Yellowstone National Park, USA)	HGX29340.1	2.E-25	30	Y
Anoxybacter fermentans strain DY22613	Firmicutes	deep-sea hydrothermal vent	AZR72222.1	1.E-04	25	N
Ardenticatena maritima strain 110S	Chloroflexi	hydrothermal field (Yamagawa, Japan)	GAP62401.1	3.E-16	30	Y
Bellilinea caldifistulae isolate SpSt-556	Chloroflexi	hot spring sediment (British Columbia, Canada)	HGS87922.1	2.E-30	33	Y
Bilophila sp. strain 4_1_30	Proteobacteria	human feces	EGW44733.1	5.E-66	41	Y
Bilophila wadsworthia strain 3_1_6	Proteobacteria	human feces	EFV46162.1	2.E-64	40	Y
Burkholderiales bacterium isolate AS22ysBPME_216	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMA29311.1	1.E-14	28	Y
Caldilinea aerophila isolate SpSt-289	Chloroflexi	hot spring sediment (British Columbia, Canada)	HDX32185.1	2.E-24	28	Y
Caldilinea aerophila strain DSM 14535	Chloroflexi	sulfur-turf sample in a hot spring (Takayama, Japan)	BAL99387.1	4.E-24	28	N
Calditerrivibrio nitroreducens strain DSM 19672	Deferribacteres	terrestrial hot spring (Lake Nojiri, Japan)	ADR19788.1	3.E-05	26	Y
candidate division KSB1 bacterium isolate B7_G2	Candidate division KSB1	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RKY75023.1	6.E-16	26	Y
candidate division KSB1 bacterium isolate B39_G15	Candidate division KSB1	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RKY84749.1	7.E-16	26	Y

candidate division KSB1 bacterium isolate CS2-K093	Candidate division KSB1	pelagic sediment (Gulf of Kutch, India)	NIV91984.1	2.E-25	33	Y
Candidatus Abyssubacteria bacterium isolate SURF_5	Candidatus Abyssubacteria	deeply circulating subsurface aquifer fluids (Lead, SD, USA)	RJP21339.1	7.E-07	27	N
Candidatus Acidulodesulfobacterium acidiphilum isolate AP4	Proteobacteria	acid mine drainage (Guangzhou, China)	RZV40568.1	3.E-05	31	N
Candidatus Acidulodesulfobacterium ferriphilum isolate AP3	Proteobacteria	acid mine drainage (Guangzhou, China)	RZD14164.1	2.E-04	27	N
Candidatus Fermentibacter daniensis isolate Ran_1	Candidatus Fermentibacteria	enrichment reactor seeded with activated sludge (Randers, Denmark)	KZD16791.1	8.E-25	31	Y
Candidatus Handelsmanbacteria bacterium isolate RIFCSPLOWO2_12_FULL_64_10	Candidatus Handelsmanbacteria	Rifle well FP-101 (Rifle, CO, USA)	OGG49119.1	3.E-28	34	Y
Candidatus Latescibacteria bacterium isolate UBA9260	Candidatus Latescibacteria	marine	HAA78539.1	2.E-28	32	Y
Candidatus Methylomirabilis limnetica strain Zug	candidate division NC10	freshwater (Lake Zug, Switzerland)	PTL35634.1	4.E-21	32	Y
Candidatus Methylomirabilis oxyfera isolate SB1	candidate division NC10	biomass on anode of bioreactor after 21 days (Ooijpolder, Netherlands)	KAB2960242.1	8.E-18	30	Y
Candidatus Nitrosotenuis sp. isolate GW928_bin.18	Thaumarchaeota	groundwater (Oak Ridge, TN, USA)	TBR21516.1	4.E-16	28	Y
Candidatus Raymondbacteria bacterium isolate RifOxyA12_full_50_37	Candidatus Raymondbacteria	Rifle well CD01 (Rifle, CO, USA)	OGJ89256.1	8.E-16	29	Y
Candidatus Thermofonsia Clade 1 bacterium isolate CP2_2F	Chloroflexi	alkaline, sulfidic hot spring Cone Pool 2 (Azumino, Japan)	PJF30363.1	4.E-42	36	Y
Candidatus Woesebacteria bacterium isolate B4_G12	Candidatus Woesebacteria	deep-sea hydrothermal vent sediments (Guaymas Basin, Mexico)	RLC29799.1	1.E-22	30	Y
Chitinispirillum alkaliphilum strain ACht6-1	Fibrobacteres	hypersaline alkaline lake (Wadi al Natrun, Egypt)	KMQ49640.1	1.E-19	29	Y
Chlorobaculum limnaeum strain DSM1677	Chlorobi	Lake Kinnevet, Israel	AOS83552.1	3.E-22	32	Y
Chlorobaculum parvum strain NCIB 8327 DSM 263	Chlorobi	freshwater	ACF11581.1	3.E-19	31	Y
Chlorobaculum sp. strain 24CR	Chlorobi	water (Carmel River, CA, USA)	RXK88351.1	3.E-21	31	Y
Chlorobaculum tepidum strain TLS	Chlorobi	high-sulfide hot spring (Rotorua, New Zealand)	AAM72179.1	1.E-17	29	Y
Chlorobaculum thiosulfatiphilum strain DSM 249	Chlorobi	tassajara hot spring (CA, USA)	TNJ40258.1	2.E-22	31	Y
Chlorobiaceae bacterium isolate June25_Bin_2	Chlorobi	freshwater lake (Trout Bog Lake, WI, USA)	NMW21650.1	4.E-17	27	Y
Chlorobiaceae bacterium isolate Oct13_Bin_1	Chlorobi	freshwater lake (Trout Bog Lake, WI, USA)	NMW18963.1	6.E-16	28	Y
Chlorobium chlorochromatii strain CaD3	Chlorobi	eutrophic lake (Brandenburg, Germany)	ABB28338.1	6.E-16	30	N
Chlorobium ferrooxidans strain DSM 13031	Chlorobi	ditch sediment (Konstanz, Germany)	EAT58859.1	6.E-17	27	Y
Chlorobium limicola strain DSM 245	Chlorobi	hot spring (CA, USA)	ACD90141.1	5.E-17	29	Y
Chlorobium limicola strain Frasassi	Chlorobi	sediment-water interface in an artificial aquarium in the Frasassi cave system (Italy)	KUL32430.1	3.E-16	29	N
Chlorobium phaeobacteroides strain BS1	Chlorobi	chemocline of the Black Sea	ACE04425.1	3.E-17	29	N
Chlorobium phaeobacteroides strain DSM 266	Chlorobi	Lake Blankvann (Norway)	ABL65163.1	3.E-24	33	Y

Chlorobium phaeovibrioides strain BrKhr17	Chlorobi	Lake Bolshye Khruslomeny, water chemocline zone	RTY40038.1	7.E-18	31	Y
Chlorobium phaeovibrioides strain DSM 265	Chlorobi	saline intertidal flat	ABP36835.1	7.E-20	31	Y
Chlorobium phaeovibrioides strain PhvTcv-s14	Chlorobi	water chemocline zone (Russia)	QEQ57063.1	6.E-18	31	Y
Chlorobium sp. isolate L227-2013-22	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU55211.1	2.E-21	31	Y
Chlorobium sp. isolate L227-2013-55	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU57894.1	8.E-19	29	N
Chlorobium sp. isolate L227-2013-56	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU51735.1	9.E-26	30	Y
Chlorobium sp. isolate L227-S-6D	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU81745.1	6.E-17	29	N
Chlorobium sp. isolate L304-S-6D	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU83342.1	2.E-21	31	Y
Chlorobium sp. isolate L442-64	Chlorobi	freshwater lake (Kenora, Ontario, Canada)	TLU54122.1	4.E-18	29	N
Chlorobium sp. isolate UBA8843	Chlorobi	synthetic microbial community	HCD35528.1	9.E-20	31	Y
Chlorobium sp. strain BLA1	Chlorobi	Brownie Lake (Minneapolis, MN, USA)	NHQ60379.1	2.E-17	28	Y
Chlorobium sp. strain KB01	Chlorobi	meromictic lake (Kabuno Bay, Congo)	WP_076791818.1	4.E-13	27	Y
Chlorobium sp. strain N1	Chlorobi	sediment (Norsminde Fjord, Denmark)	TCD47456.1	8.E-21	30	Y
Chloroflexi bacterium isolate AS06rmzACSIP_450	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLW71729.1	7.E-33	32	Y
Chloroflexi bacterium isolate AS15tlH2ME_173	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLA79515.1	6.E-35	34	Y
Chloroflexi bacterium isolate AS26fmACSIPLY_11	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	HHY58046.1	2.E-22	30	Y
Chloroflexi bacterium isolate AS27yjCOA_1	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMB61392.1	1.E-36	31	Y
Chloroflexi bacterium isolate AS27yjCOA_123	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMB86957.1	4.E-33	33	Y
Chloroflexi bacterium isolate AS27yjCOA_124	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMB61392.1	1.E-36	31	Y
Chloroflexi bacterium isolate AS27yjCOA_4	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMC52453.1	1.E-23	33	Y
Chloroflexi bacterium isolate AS27yjCOA_56	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMC45596.1	4.E-35	32	Y
Chloroflexi bacterium isolate B46_G1	Chloroflexi	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC71331.1	5.E-39	35	Y
Chloroflexi bacterium isolate CFX1	Chloroflexi	marine ANAMMOX bioreactor	KAA3663928.1	4.E-35	34	N
Chloroflexi bacterium isolate CR04	Chloroflexi	marine sediment (Pacific Ocean)	TDI86467.1	2.E-33	33	Y
Chloroflexi bacterium isolate HGW-Chloroflexi-1	Chloroflexi	groundwater (Horonobe, Japan)	PKO23290.1	1.E-33	34	Y
Chloroflexi bacterium isolate HGW-Chloroflexi-10	Chloroflexi	groundwater (Horonobe, Japan)	PKO13937.1	1.E-35	34	Y
Chloroflexi bacterium isolate M_MetaBat.114	Chloroflexi	deepsea hydrothermal sulfide chimney (Pacific Ocean)	NOZ27705.1	8.E-31	30	Y
Chloroflexi bacterium isolate metabat2.725	Chloroflexi	Prairie Pothole Region wetland sediment (Cottonwood Lake, ND, USA)	RPH61221.1	1.E-37	35	Y

Chloroflexi bacterium isolate RBG_16_54_18	Chloroflexi	Rifle well D04 (Rifle, CO, USA)	OGO29581.1	8.E-40	34	Y
Chloroflexi bacterium isolate SpSt-271	Chloroflexi	hot spring sediment (British Columbia, Canada)	HDY04783.1	4.E-35	35	Y
Chloroflexi bacterium isolate SpSt-326	Chloroflexi	hot spring sediment (British Columbia, Canada)	HDV36418.1	2.E-36	35	Y
Chloroflexi bacterium isolate SpSt-327	Chloroflexi	hot spring sediment (British Columbia, Canada)	HEG74209.1	5.E-38	36	Y
Chloroflexi bacterium isolate SpSt-422	Chloroflexi	hot spring sediment (British Columbia, Canada)	HFN11568.1	5.E-34	32	Y
Chloroflexi bacterium isolate SpSt-438	Chloroflexi	hot spring sediment (British Columbia, Canada)	HFM60814.1	7.E-33	34	Y
Chloroflexi bacterium isolate SpSt-474	Chloroflexi	hot spring sediment (British Columbia, Canada)	HGU24517.1	2.E-37	35	Y
Chloroflexi bacterium isolate SpSt-552	Chloroflexi	hot spring sediment (British Columbia, Canada)	HGQ02489.1	2.E-31	32	Y
Chloroflexi bacterium isolate SpSt-583	Chloroflexi	hot spring sediment (British Columbia, Canada)	HGT18656.1	3.E-33	34	Y
Chloroflexi bacterium isolate SpSt-600	Chloroflexi	hot spring sediment (Tengchong, China)	HGO05051.1	8.E-34	32	Y
Chloroflexi bacterium isolate SpSt-660	Chloroflexi	hot spring sediment (Tengchong, China)	HGL63430.1	3.E-33	32	Y
Chloroflexi bacterium isolate SpSt-996	Chloroflexi	mud, sediment (Yellowstone National Park, USA)	HFU29806.1	2.E-33	34	Y
Chloroflexi bacterium isolate UBA9854	Chloroflexi	anaerobic digester	HAJ37108.1	2.E-26	31	Y
Chloroflexi bacterium isolate UTCFX4	Chloroflexi	anaerobic digester filtrate from the belt filter press of CVWRF wastewater treatment plant (Salt Lake City, UT, USA)	OQY79586.1	2.E-23	31	Y
Cloacibacillus evryensis strain DSM 19522	Synergistetes	sewage sludge (Évry, France)	WP_034444262.1	6.E-15	26	N
Cloacibacillus porcorum strain 105753	Synergistetes	pig feces	NMF17534.1	3.E-14	26	N
Cloacibacillus porcorum strain CL-84	Synergistetes	pig cecal mucosa	ANZ46032.1	8.E-12	24	N
Cloacibacillus sp. strain An23	Synergistetes	red junglefowl cecum	OUO93278.1	9.E-16	27	N
Deferribacter desulfuricans strain SSM1	Deferribacteres	deep-sea hydrothermal vent (Suiyo Seamount, Japan)	BAI81501.1	9.E-05	24	Y
Dehalococcoidia bacterium isolate Baikal-deep-G109	Chloroflexi	Lake Baikal (Russia)	MSQ14392.1	3.E-13	26	N
Deltaproteobacteria bacterium isolate 1MN72D_58_314	Proteobacteria	groundwater (Munich, Germany)	TDB36245.1	2.E-12	28	Y
Deltaproteobacteria bacterium isolate AS06rmzACSIP_532	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLV24503.1	2.E-13	30	Y
Deltaproteobacteria bacterium isolate AS06rmzACSIP_589	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLX52857.1	3.E-38	35	Y
Deltaproteobacteria bacterium isolate AS4AglBPMA_32	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMC97429.1	3.E-38	35	Y
Deltaproteobacteria bacterium isolate B3_G2	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC06739.1	3.E-17	28	Y
Deltaproteobacteria bacterium isolate B13_G4	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC22860.1	2.E-24	30	Y
Deltaproteobacteria bacterium isolate B17_G16	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLB37252.1	7.E-12	26	Y
Deltaproteobacteria bacterium isolate B23_G16	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC12904.1	9.E-19	31	Y

Deltaproteobacteria bacterium isolate B46_G9	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLB18191.1	1.E-19	29	Y
Deltaproteobacteria bacterium isolate B125_G9	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC29483.1	6.E-19	26	Y
Deltaproteobacteria bacterium isolate B133_G9	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC20422.1	1.E-24	32	N
Deltaproteobacteria bacterium isolate B144_G9	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RLC25857.1	9.E-18	28	Y
Deltaproteobacteria bacterium isolate CG_4_8_14_3_um_filter_51_11	Proteobacteria	aquifer (Crystal Geyser, UT, USA)	PIX18886.1	2.E-18	27	N
Deltaproteobacteria bacterium isolate CG03_land_8_20_14_0_80_45_14	Proteobacteria	aquifer (Crystal Geyser, UT, USA)	PIV21324.1	1.E-22	27	Y
Deltaproteobacteria bacterium isolate DOLJORAL78_50_10	Proteobacteria	gingival sulcus from 5 year old male Dolphin_J	PIE73735.1	7.E-21	31	Y
Deltaproteobacteria bacterium isolate DOLJORAL78_53_22	Proteobacteria	gingival sulcus from 5 year old male Dolphin_J	PIE70991.1	4.E-07	25	N
Deltaproteobacteria bacterium isolate HGW-Deltaproteobacteria-1	Proteobacteria	groundwater (Horonobe, Japan)	PKN88290.1	3.E-30	32	Y
Deltaproteobacteria bacterium isolate HGW-Deltaproteobacteria-5	Proteobacteria	groundwater (Horonobe, Japan)	PKN10657.1	2.E-34	33	Y
Deltaproteobacteria bacterium isolate HGW-Deltaproteobacteria-6	Proteobacteria	groundwater (Horonobe, Japan)	PKN18563.1	3.E-36	35	Y
Deltaproteobacteria bacterium isolate HGW-Deltaproteobacteria-11	Proteobacteria	groundwater (Horonobe, Japan)	PKN59704.1	2.E-33	33	Y
Deltaproteobacteria bacterium isolate HGW-Deltaproteobacteria-19	Proteobacteria	groundwater (Horonobe, Japan)	PKN34680.1	1.E-21	32	Y
Deltaproteobacteria bacterium isolate INTA.AUR.1003	Proteobacteria	first pond of the AUR serial dairy industry stabilization pond system (Vila, Argentina)	NCC25723.1	1.E-14	28	Y
Deltaproteobacteria bacterium isolate J040	Proteobacteria	iron-rich hot spring (Shikinejima Island, Japan)	RMG60923.1	1.E-23	31	N
Deltaproteobacteria bacterium isolate MAG 48	Proteobacteria	hydrothermal vent (Pacific Ocean)	RTZ96695.1	3.E-15	27	Y
Deltaproteobacteria bacterium isolate RIFOXYC2_FULL_48_10	Proteobacteria	Rifle well CD01 (Rifle, CO, USA)	OGQ92707.1	3.E-22	29	Y
Deltaproteobacteria bacterium isolate SpSt-510	Proteobacteria	hot spring sediment (British Columbia, Canada)	HFH11051.1	2.E-20	30	Y
Deltaproteobacteria bacterium isolate SpSt-772	Proteobacteria	mud, sediment (Yellowstone National Park, USA)	HGV81120.1	2.E-16	27	Y
Deltaproteobacteria bacterium isolate SpSt-871	Proteobacteria	mud, sediment (Yellowstone National Park, USA)	HGZ73194.1	7.E-16	27	Y
Deltaproteobacteria bacterium isolate SpSt-1015	Proteobacteria	mud, sediment (Yellowstone National Park, USA)	HHH87581.1	9.E-20	30	Y
Deltaproteobacteria bacterium isolate UBA10529	Proteobacteria	oil sands and coal beds	HAR96709.1	2.E-15	28	Y
Desulfamplus magnetovallimortis strain BW-1	Proteobacteria	brackish spring (Death Valley, CA, USA)	SLM31755.1	7.E-20	25	Y
Desulfatibacillum aliphaticivorans strain AK-01	Proteobacteria	Arthur Kill, NJ/NY waterway (USA)	ACL02166.1	5.E-19	33	N
Desulfatibacillum aliphaticivorans strain DSM 15576	Proteobacteria	hydrocarbon-polluted marine sediment (Gulf of Fos, France)	WP_051327074.1	5.E-20	34	N

Desulfatibacillum alkenivorans strain DSM 16219	Proteobacteria	oil-polluted sediment of operation station of ballast water and tank water cleaning (Fos Harbor, France)	WP_083611154.1	1.E-17	32	N
Desulfatiglans anilini strain DSM 4660	Proteobacteria	marine sediment (North Sea coast, Germany)	WP_028322618.1	1.E-22	31	N
Desulfatitalea sp. isolate Site_C24	Proteobacteria	beach sand (Middle Park Beach, Australia)	NNK02546.1	1.E-12	28	Y
Desulfatitalea tepidiphila strain S28bF	Proteobacteria	marine sediment (Tokyo Bay, Japan)	WP_076750459.1	2.E-23	30	Y
Desulfobacter curvatus strain DSM 3379	Proteobacteria	marine mud (Venice, Italy)	WP_020586554.1	2.E-20	30	Y
Desulfobacter hydrogenophilus strain AcRS1	Proteobacteria	marine mud (Venice, Italy)	QBH12555.1	8.E-15	27	Y
Desulfobacter postgatei strain 2ac9	Proteobacteria	anaerobic sediment of brackish water ditch (Jadebusen, Germany)	EIM62640.1	1.E-21	30	Y
Desulfobacter sp. isolate UBA12168	Proteobacteria	sediment	HBT90065.1	2.E-21	30	Y
Desulfobacter vibrioformis strain DSM 8776	Proteobacteria	water-oil separation system on oil platform (North Sea, Norway)	WP_035235263.1	4.E-18	29	Y
Desulfobacteraceae bacterium isolate 4484_190.3	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	OPX36574.1	8.E-19	28	Y
Desulfobacteraceae bacterium isolate 4572_88	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	OQY55368.1	3.E-20	31	N
Desulfobacteraceae bacterium isolate 4572_123	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	OQY04121.1	5.E-26	30	Y
Desulfobacteraceae bacterium isolate CG2_30_51_40	Proteobacteria	groundwater (Crystal Geyser, UT, USA)	OIP43152.1	2.E-18	27	N
Desulfobacteraceae bacterium isolate DS_bin_10	Proteobacteria	beach sand (Middle Park Beach, Australia)	NNG02102.1	2.E-12	25	Y
Desulfobacteraceae bacterium isolate Eth-SRB1	Proteobacteria	anaerobic ethane-degrading enrichment culture inoculated with marine sediment (Gulf of Mexico)	RZB29647.1	5.E-18	27	Y
Desulfobacteraceae bacterium isolate Eth-SRB2	Proteobacteria	anaerobic ethane-degrading enrichment culture inoculated with marine sediment (Gulf of Mexico)	RZB36540.1	3.E-25	30	Y
Desulfobacteraceae bacterium isolate HyVt-13	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	HDI59857.1	1.E-10	29	N
Desulfobacteraceae bacterium isolate HyVt-208	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	HDL07956.1	2.E-22	28	Y
Desulfobacteraceae bacterium isolate IS3	Proteobacteria	microbial mat in sulfidic groundwater-fed sinkhole (Lake Huron, USA)	OQX21718.1	1.E-23	27	N
Desulfobacteraceae bacterium isolate IS3	Proteobacteria	microbial mat in sulfidic groundwater-fed sinkhole (Lake Huron, USA)	OQW99265.1	2.E-19	30	Y
Desulfobacteraceae bacterium isolate IS3	Proteobacteria	microbial mat in sulfidic groundwater-fed sinkhole (Lake Huron, USA)	OQX26173.1	2.E-18	30	Y
Desulfobacteraceae bacterium isolate Madre2	Proteobacteria	bioplastic (PHA) biofilm from water-sediment interface of coastal lagoon (Upper Laguna Madre, TX, USA)	THB81596.1	3.E-17	27	Y
Desulfobacteraceae bacterium isolate SURF_4	Proteobacteria	deeply circulating subsurface aquifer fluids (Lead, SD, USA)	RJQ76849.1	7.E-11	26	Y
Desulfobacteraceae bacterium isolate SURF_4	Proteobacteria	deeply circulating subsurface aquifer fluids (Lead, SD, USA)	RJQ65780.1	2.E-05	21	N

Desulfobacteraceae bacterium isolate SURF_7	Proteobacteria	deeply circulating subsurface aquifer fluids (Lead, SD, USA)	RJP79134.1	6.E-21	29	Y
Desulfobacteraceae bacterium isolate UBA8212	Proteobacteria	seawater	HCY85389.1	9.E-14	28	Y
Desulfobacteraceae bacterium isolate UBA8400	Proteobacteria	microbial mat (Lake Huron, MI, USA)	HAO21021.1	5.E-27	27	N
Desulfobacterales bacterium isolate DOLJORAL78_48_6	Proteobacteria	gingival sulcus from 5 year old male Dolphin_J	PIE60952.1	7.E-18	28	Y
Desulfobacterales bacterium isolate GLR701	Proteobacteria	Glendhu Ridge methane seep (Pacific Ocean)	NOQ19417.1	9.E-24	30	N
Desulfobacterales bacterium isolate HyVt-57	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	HGY11719.1	1.E-17	28	Y
Desulfobacterales bacterium isolate HyVt-57	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	HGY11834.1	4.E-15	28	Y
Desulfobacterales bacterium isolate HyVt-59	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	HHC24326.1	1.E-23	32	Y
Desulfobacterales bacterium isolate S7086C20	Proteobacteria	methane seep sediment (Pacific Ocean)	OEU45016.1	4.E-24	31	Y
Desulfobacterales bacterium isolate Site_B14	Proteobacteria	beach sand (Middle Park Beach, Australia)	NNK84523.1	2.E-23	29	Y
Desulfobacterales bacterium isolate SpSt-563	Proteobacteria	hot spring sediment (British Columbia, Canada)	HGP66397.1	6.E-19	30	Y
Desulfobacterales bacterium isolate UWMA-0216	Proteobacteria	Mid-Cayman Rise vent fluids (Atlantic Ocean)	HID59813.1	2.E-15	26	Y
Desulfobacterium autotrophicum strain HRM2	Proteobacteria	marine sediment (Mediterranean Sea)	ACN17620.1	3.E-19	30	Y
Desulfobacterium vacuolatum strain DSM 3385	Proteobacteria	marine mud (Venice, Italy)	SMC85487.1	7.E-17	29	Y
Desulfobacula phenolica strain DSM 3384	Proteobacteria	marine mud (Venice, Italy)	SDU45768.1	3.E-16	26	Y
Desulfobacula sp. isolate GWF2_41_7	Proteobacteria	Rifle well CD01 (Rifle, CO, USA)	OGR15363.1	5.E-19	27	Y
Desulfobacula sp. isolate SM1_1_2	Proteobacteria	stromatolite mat (Schoenmakerskop, South Africa)	NJM02770.1	1.E-17	28	N
Desulfobacula toluolica strain Tol2	Proteobacteria	marine mud (Woods Hole, MA, USA)	CCK82146.1	2.E-15	26	Y
Desulfobulbaceae bacterium isolate DB1	Proteobacteria	anode biofilm in microbial fuel cells	OKY74567.1	6.E-05	27	Y
Desulfobulbus oralis strain HOT041/ORNL	Proteobacteria	human subgingival sample	AVD70372.1	2.E-57	37	Y
Desulfobulbus oralis strain HOT041/ORNL	Proteobacteria	human subgingival sample	AVD71509.1	1.E-61	40	Y
Desulfococcus multivorans strain DSM 2059	Proteobacteria	sewage digester (Göttingen, Germany)	EPR35993.1	2.E-19	32	Y
Desulfococcus sp. isolate 4484_241	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	OQX63297.1	3.E-20	29	N
Desulfococcus sp. isolate 4484_242	Proteobacteria	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	OQX61744.1	2.E-16	28	N
Desulfoluna spongiiphila strain AA1	Proteobacteria	marine sponge Aplysina aerophoba (Mediterranean Sea)	SCX79058.1	2.E-06	27	N
Desulfoluna spongiiphila strain DBB	Proteobacteria	marine intertidal sediment (L'Escala, Spain)	VVS90461.1	8.E-06	26	N
Desulfomonile tiedjei isolate SpSt-769	Proteobacteria	mud, sediment (Yellowstone National Park, USA)	HGH61094.1	5.E-28	32	Y
Desulfomonile tiedjei strain DSM 6799	Proteobacteria	sewage sludge (Adrian, MI, USA)	AFM25414.1	1.E-36	35	Y
Desulfonatronospira thiodismutans strain ASO3-1	Proteobacteria	sediment from a highly alkaline saline soda lake on the Kulunda Steppe (Altai, Russia)	EFI33801.1	1.E-22	33	Y
Desulfonema ishimotonii strain Tokyo 01	Proteobacteria	marine sediment (Tokyo Bay, Japan)	GBC64034.1	3.E-21	30	N

Desulfosarcina alkanivorans strain PL12	Proteobacteria	oil spill of shallow marine sediment (Shuaiba, Kuwait)	BBO72287.1	2.E-19	30	Y
Desulfosarcina cetonica strain JCM 12296	Proteobacteria	flooded oil stratum of the Apsheron peninsula (Azerbaijan)	WP_054702774.1	2.E-13	28	N
Desulfosarcina ovata subsp. sediminis strain 28bB2T	Proteobacteria	tidal flat sediment (Tokyo Bay, Japan)	BBO80496.1	3.E-11	27	Y
Desulfosarcina sp. strain BuS5	Proteobacteria	sediment (Guaymas Basin, Mexico)	WP_027353905.1	9.E-15	25	Y
Desulfosarcina widdelii strain PP31	Proteobacteria	oil spill of shallow marine sediment (Shuaiba, Kuwait)	BBO74481.1	8.E-15	26	Y
Desulfospira joergensenii strain DSM 10085	Proteobacteria	marine surface sediment below sea grass (Arcachon Bay, France)	WP_022666032.1	2.E-19	28	Y
Desulfotignum balticum strain DSM 7044	Proteobacteria	marine mud (Saxild, Denmark)	WP_024336727.1	6.E-23	32	Y
Desulfotignum phosphitoxidans strain DSM 13687	Proteobacteria	marine sediment (Venice, Italy)	EMS79033.1	3.E-23	32	Y
Desulfovibrio sp. strain An276	Proteobacteria	red junglefowl cecum	OUO52973.1	3.E-60	35	Y
Desulfovibrionaceae bacterium isolate CIM:MAG 1040	Proteobacteria	human feces	PWM70546.1	1.E-65	41	Y
Desulfurispirillum indicum strain S5	Chrysiogenetes	river sediment (Chennai, India)	ADU66823.1	1.E-04	26	Y
Desulfurivibrio sp. isolate SURF_16	Proteobacteria	deeply circulating subsurface aquifer fluids (Lead, SD, USA)	RJX31147.1	8.E-11	28	Y
Desulfuromonadales bacterium isolate C00003093	Proteobacteria	sediment from site of active methane seepage at Hydrate Ridge South (Pacific Ocean)	OEU75150.1	2.E-18	28	Y
Fibrobacter sp. isolate AS06rmzACSIP_199	Fibrobacteres	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLG16780.1	3.E-22	29	Y
Fibrobacter sp. isolate AS06rmzACSIP_529	Fibrobacteres	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLD91845.1	1.E-21	27	N
Fibrobacter sp. isolate AS06rmzACSIP_543	Fibrobacteres	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLE01532.1	7.E-21	30	Y
Fibrobacter sp. isolate AS09scLD_71	Fibrobacteres	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLP03213.1	2.E-22	30	Y
Fibrobacter sp. isolate AS22ysBPME_152	Fibrobacteres	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLL13864.1	3.E-20	27	Y
Fibrobacteres bacterium isolate UBA10882	Fibrobacteres	wetland surface sediment	HAJ80221.1	6.E-20	30	Y
Firmicutes bacterium isolate AS23ysBPME_266	Firmicutes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMB11274.1	4.E-06	25	N
Flexilinea flocculi strain TC1	Chloroflexi	methanogenic granular sludge (Belgium)	GAP39404.1	5.E-43	34	Y
Fretibacterium sp. isolate AS21ysBPME_7	Synergistetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLL37650.1	1.E-16	32	N
Geovibrio thiophilus strain DSM 11263	Deferribacteres	drainage ditch (Konstanz, Germany)	QAR33452.1	4.E-08	28	Y
Ignavibacteriae bacterium isolate B3_G1	Chlorobi	deep-sea hydrothermal vent sediment (Guaymas Basin, Mexico)	RKY90554.1	4.E-22	29	Y
Lentimonas sp. CC4	Verrucomicrobia	seawater (Nahant, MA, USA)	CAA6679502.1	1.E-15	30	Y
Lentimonas sp. CC10	Verrucomicrobia	seawater (Nahant, MA, USA)	CAA6691588.1	3.E-15	29	Y

Lentisphaerae bacterium isolate AS06rmzACSIP_375	Lentisphaerae	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLE68485.1	5.E-15	29	Y
Lentisphaerae bacterium isolate RIFOXYA12_64_32	Lentisphaerae	Rifle well CD01 (Rifle, CO, USA)	OGV62430.1	2.E-28	33	Y
Leptolinea sp. isolate AS06rmzACSIP_266	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLF50870.1	2.E-41	35	Y
Leptolinea sp. isolate AS27yjCOA_114	Chloroflexi	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMB53094.1	2.E-36	34	Y
Leptolinea tardivitalis strain YMTK-2	Chloroflexi	wastewater from a factory producing shochu (Yamagawa Beach, Japan)	KPL70500.1	2.E-38	33	Y
Levilinea saccharolytica strain KIBI-1	Chloroflexi	sludge granules from sugar-processing plant (Niigata, Japan)	KPL75587.1	3.E-35	35	Y
Nitrolancea hollandica strain Lb	Chloroflexi	nitrifying bioreactor (Rotterdam, Netherlands)	CCF84771.1	8.E-06	24	N
Nitrospira moscoviensis strain NSP M-1	Nitrospirae	eroded iron pipe (Moscow, Russia)	ALA58029.1	2.E-16	29	Y
Olavius algarvensis Delta 1 endosymbiont	Proteobacteria	shallow marine sediment (Italy)	CAB1077770.1	1.E-21	29	Y
Olavius sp. associated proteobacterium Delta 1	Proteobacteria	shallow marine sediment (Italy)	CAB1061238.1	6.E-20	28	Y
Ornatilinea apprima strain P3M-1	Chloroflexi	microbial mat formed in wooden bath filled with hot water from a 2775 m-deep well (Siberia, Russia)	KPL78704.1	1.E-20	33	Y
Pelodictyon luteolum strain DSM 273	Chlorobi	meromictic lake (Norway)	ABB23996.1	2.E-15	29	Y
Pelodictyon phaeoclathratiforme strain BU-1	Chlorobi	monimolimnion of Buchensee (Germany)	ACF43535.1	1.E-22	32	Y
Pelolinea submarina strain DSM 23923	Chloroflexi	marine subsurface sediment (Pacific Ocean)	BBB47806.1	2.E-35	32	Y
Planctomycetes bacterium isolate GSL.Bin20	Planctomycetes	microbial mat (Bridger Bay, UT, USA)	NBB94753.1	5.E-22	30	Y
Prosthecochloris aestuarii isolate SpSt-1181	Chlorobi	hot spring sediment (Wilbur Springs, CA, USA)	HED30548.1	2.E-28	36	Y
Prosthecochloris aestuarii strain DSM 271	Chlorobi	hydrogen sulfide containing mud of brackish lagoon (Lake Sasyk-Sivash, Russia)	ACF46288.1	2.E-20	30	N
Prosthecochloris marina strain V1	Chlorobi	coastal area of the South China Sea	PWW83239.1	5.E-23	31	N
Prosthecochloris sp. isolate C10	Chlorobi	seawater lake chemocline (Dragon's Eye Lake, Croatia)	NEX14508.1	3.E-17	28	N
Prosthecochloris sp. strain CIB 2401	Chlorobi	coastal lagoon (Mallorca, Spain)	ANT65030.1	2.E-26	34	Y
Prosthecochloris sp. strain GSB1	Chlorobi	deep-sea hydrothermal vent (Pacific Ocean)	ASQ91644.1	2.E-26	33	N
Prosthecochloris sp. strain HL-130-GSB	Chlorobi	phototrophic microbial mat (Hot Lake, WA, USA)	ARM30833.1	5.E-28	36	Y
Prosthecochloris sp. strain ZM_2	Chlorobi	water from chemokine of meromictic lakes Green cape (Karelia, Russia)	RNA64855.1	8.E-24	34	Y
Prosthecochloris vibrioformis strain DSM 260	Chlorobi	rivermouth	TNJ36903.1	9.E-27	34	Y
Proteobacteria bacterium isolate DOLZORAL124_55_4	Proteobacteria	gingival sulcus from 29 year old lactating female Dolphin_Z	PID40410.1	2.E-14	28	Y
Proteobacteria bacterium isolate SpSt-1152	Proteobacteria	hot spring sediment (Wilbur Springs, CA, USA)	HEC99630.1	2.E-15	28	Y
Pseudodesulfovibrio sp. strain S3	Proteobacteria	Black Sea water (Bulgaria)	RWU02434.1	2.E-06	26	Y
Smithella sp. isolate AS06rmzACSIP_551	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLD79966.1	2.E-35	33	Y

Smithella sp. isolate AS17jrsBPGN_1	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLA41131.1	1.E-32	33	Y
Smithella sp. isolate AS4AglBPMA_11	Proteobacteria	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMC92510.1	5.E-35	34	Y
Smithella sp. isolate D17	Proteobacteria	produced water from an oilfield (Medicine Hat, Alberta, Canada)	KFZ45035.1	2.E-36	34	Y
Smithella sp. isolate F21	Proteobacteria	oil sands tailings pond (Medicine Hat, Alberta, Canada)	KFN39154.1	7.E-36	34	Y
Smithella sp. isolate SCADC	Proteobacteria	mature fine tailings from oil sands tailings pond (Medicine Hat, Alberta, Canada)	KFO68937.1	3.E-37	35	Y
Smithella sp. isolate SDB	Proteobacteria	sediment (San Diego Bay, CA, USA)	KQC09736.1	6.E-33	33	Y
Smithella sp. isolate UBA10513	Proteobacteria	oil sands and coal beds	HBI47682.1	1.E-38	34	Y
Sphaerobacter thermophilus strain DSM 20745	Chloroflexi	thermal treated municipal sewage sludge (Muenchen-Grosslappen, Germany)	ACZ37929.1	9.E-07	28	N
Sphaerochaeta dissipatitropha strain GLS2	Spirochaetes	arctic permafrost (Russia)	SMP46469.1	1.E-21	31	Y
Sphaerochaeta globosa strain Buddy	Spirochaetes	marine hot spring (Shiashkoten Island, Russia)	ADY12711.1	8.E-22	30	Y
Sphaerochaeta halotolerans strain 4-11	Spirochaetes	Production water from oilfield (Vostochno-Anzirskoe oilfield, Russia)	RFU94749.1	1.E-17	31	Y
Sphaerochaeta pleomorpha strain Grapes	Spirochaetes	Red Cedar River (Okemos, MI, USA)	AEV28959.1	1.E-18	30	Y
Sphaerochaeta sp. isolate MAG5	Spirochaetes	Red-pigmented microbial biofilm developed on the inner wall of a bioreactor (Jena, Germany)	TAH57281.1	2.E-17	30	Y
Sphaerochaeta sp. isolate UBA8956	Spirochaetes	oil sands and coal beds	HAF86241.1	4.E-12	26	Y
Sphaerochaeta sp. isolate UBA9948	Spirochaetes	oil sands and coal beds	HAP57835.1	9.E-20	31	Y
Sphaerochaeta sp. isolate UBA11053	Spirochaetes	bioreactor	HCU30833.1	2.E-17	30	Y
Sphaerochaeta sp. isolate UBA11175	Spirochaetes	oil sands and coal beds	HBO36282.1	7.E-20	31	Y
Spirochaetales bacterium isolate AS06rmzACSIP_457	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLE15601.1	7.E-20	30	Y
Spirochaetales bacterium isolate AS07pgkLD_77	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLY07479.1	1.E-14	30	Y
Spirochaetales bacterium isolate AS08sgBPME_324	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	HHT81641.1	2.E-17	29	Y
Spirochaetales bacterium isolate AS10tlH2TH_379	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLA97632.1	2.E-20	30	Y
Spirochaetales bacterium isolate AS22ysBPME_10	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLL25178.1	4.E-17	29	Y
Spirochaetales bacterium isolate AS22ysBPME_249	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NMA22686.1	3.E-15	30	Y
Spirochaetales bacterium isolate AS23ysBPME_118	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLK13860.1	8.E-18	29	Y
Spirochaetales bacterium isolate AS23ysBPME_120	Spirochaetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLK06812.1	1.E-15	29	Y
Spirochaetes bacterium isolate ADurb.Bin315	Spirochaetes	anaerobic digester (Urbana, IL, USA)	OQA43369.1	4.E-16	28	Y

Spirochaetes bacterium strain RBG_16_49_21	Spirochaetes	Rifle well D04 (Rifle, CO, USA)	OHD69784.1	2.E-12	27	N
Spirochaetia bacterium isolate INTA.CYC.017	Spirochaetes	first pond of the CYC serial dairy industry stabilization pond system (San Carlos Sur, Argentina)	NCC13313.1	4.E-17	28	Y
Spirochaetia bacterium isolate UBA12135	Spirochaetes	aquifer (Colorado River, CO, USA)	HBE03855.1	4.E-05	24	Y
Synergistaceae bacterium isolate AS07pgkLD_87	Synergistetes	anaerobic digestion of organic wastes under variable temperature conditions and feedstocks	NLV82264.1	5.E-11	26	N
Synergistes sp. strain 3_1_syn1	Synergistetes	human gut	EHL68284.1	7.E-15	26	N
Synergistetes bacterium isolate HGW-Synergistetes-1	Synergistetes	groundwater (Horonobe, Japan)	PKL03800.1	4.E-08	25	N
Syntrophaceae bacterium isolate UBA10514	Proteobacteria	oil sands and coal beds	HBH86494.1	7.E-19	30	Y
Syntrophaceae bacterium isolate UBA11395	Proteobacteria	oil sands and coal beds	HBJ74262.1	3.E-38	35	Y
Syntrophobacter sp. strain SbD1	Proteobacteria	peat soil (Weißenstadter Forst-Nord, Germany)	SPF48485.1	2.E-24	34	Y
Syntrophus sp. isolate UBA8958	Proteobacteria	oil sands and coal beds	HAJ26831.1	1.E-21	32	Y
Syntrophus sp. isolate UBA10520	Proteobacteria	oil sands and coal beds	HAR97668.1	2.E-18	31	Y
Thermaerobacter sp. strain PB12/4term	Firmicutes	Sediment (Lake Baikal, Russia)	QIA27848.1	4.E-04	28	N
uncultured Desulfatiglans sp. isolate IK1	Proteobacteria	asphalt lake (Pitch Lake, Trinidad and Tobago)	VBB43204.1	3.E-19	32	N
uncultured Desulfobacteraceae bacterium isolate CR-1	Proteobacteria	ectosymbiote of a magnetic protist of Mediterranean sea	VEN73776.1	2.E-15	28	Y
Uncultured Desulfobacterium sp.	Proteobacteria	contaminated aquifer (Stuttgart, Germany)	CBX28553.1	2.E-24	30	N
Verrucomicrobia bacterium isolate DP16D_bin.41	Verrucomicrobia	groundwater (Oak Ridge, TN, USA)	TAN38092.1	4.E-22	29	Y

Table S2. Biochemically and structurally characterized members of AP endonuclease 2 superfamily

Enzyme	Organism	Accession number	References
Endonuclease IV	Bacillus subtilis subsp. subtilis str. 168	P54476.1	Salas-Pacheco JM, Urtiz-Estrada N, Martínez-Cadena G, Yasbin RE, & Pedraza-Reyes M (2003) J Bacteriol 185(18):5380-5390
Endonuclease IV	Chlamydia pneumoniae AR39	AAF37910.1	Liu X & Liu J (2005) Biochim Biophys Acta 1753(2):217-225
Endonuclease IV	Escherichia coli str. K-12 substr. MG1655	AAC75220.1	Hosfield DJ, Guan Y, Haas BJ, Cunningham RP, & Tainer JA (1999) Cell 98(3):397-408; Ishchenko AA, Ide H, Ramotar D, Nevinsky G, & Saparbaev M (2004) Biochemistry 43(48):15210-15216
Endonuclease IV	Geobacillus kaustophilus HTA426	BAD76759.1	Asano R, et al. (2011) Acta Crystallogr D Biol Crystallogr 67(Pt 3):149-155
Endonuclease IV	Mycobacterium tuberculosis H37Rv	NP_215184.1	Zhang W, et al. (2018) Biochem Biophys Res Commun 498(1):111-118; Puri RV, Singh N, Gupta RK, & Tyagi AK (2013) PLoS One 8(8):e71535
Endonuclease IV	Mycoplasma genitalium G37	AAC71456.1	Estevão S, van der Spek PE, van Rossum AMC, & Vink C (2014) Microbiology 160(Pt 6):1087-1100
Endonuclease IV	Mycoplasma pneumoniae M129	AAB96156.1	Estevão S, van der Spek PE, van Rossum AMC, & Vink C (2014) Microbiology 160(Pt 6):1087-1100
Endonuclease IV	Pyrobaculum aerophilum str. IM2	AAL64792.1	Sartori AA & Jiricny J (2003) J Biol Chem 278(27):24563-24576
Endonuclease IV	Pyrococcus furiosus DSM 3638	AAL80382.1	Kiyonari S, et al. (2009) Nucleic Acids Res 37(19):6439-6453
Endonuclease IV	Sulfolobus islandicus REY15A	ADX86710.1	Yan Z, Huang Q, Ni J, & Shen Y (2016) Extremophiles 20(5):785-793
Endonuclease IV	Thermococcus eurythermalis A501	AIU69743.1	Wang WW, et al. (2018) Int J Mol Sci 20(1)
Endonuclease IV	Thermotoga maritima MSB8	Q9WYJ7.1	Tomanicek SJ, Hughes RC, Ng JD, & Coates L (2010) Acta Crystallogr Sect F Struct Biol Cryst Commun 66(Pt 9):1003-1012; Haas BJ, Sandigursky M, Tainer JA, Franklin WA, & Cunningham RP (1999) J Bacteriol 181(9):2834-2839
Endonuclease IV	Thermus thermophilus HB8	BAD70657.1	Asano R, et al. (2011) Acta Crystallogr D Biol Crystallogr 67(Pt 3):149-155; Back JH, Chung JH, Park JH, & Han YS (2006) Biochem Biophys Res Commun 346(3):889-895
Endonuclease IV	Vibrio cholerae C6706	OFJ21245.1	Davies BW, et al. (2011) PLoS Pathog 7(2):e1001295
Fructoselysine 3-epimerase	Escherichia coli BL21(DE3)	ACT45027.1	Wiame E & Van Schaftingen E (2004) Biochem J 378(Pt 3):1047-1052
Hydroxypyruvate isomerase	Escherichia coli K-12	P30147.1	Ashiuchi M & Misono H (1999) Biochim Biophys Acta 1435(1-2):153-159
Hydroxypyruvate isomerase	Streptomyces coelicolor A3(2)	Q9Z596.1	Navone L, et al. (2015) Appl Environ Microbiol 81(19):6649-6659
2-keto-myo-inositol dehydratase	Bacillus subtilis subsp. subtilis str. 168	AIY95289.1	Yoshida KI, et al. (2004) Microbiology 150(Pt 3):571-580
2-keto-myo-inositol dehydratase	Caulobacter vibrioides CB15	AAK23281.1	Boutte CC, et al. (2008) PLoS Genet 4(12):e1000310
2-keto-myo-inositol dehydratase	Citrobacter sp. TBCP-5362	QCQ70428.1	Yuan C, Yang P, Wang J, & Jiang L (2019) Biochem Biophys Res Commun 517(3):427-432
2-keto-myo-inositol dehydratase	Clostridium perfringens str. 13	BAB79797.1	Kawsar HI, Ohtani K, Okumura K, Hayashi H, & Shimizu T (2004) FEMS Microbiol Lett 235(2):289-295
2-keto-myo-inositol dehydratase	Corynebacterium glutamicum ATCC 13032	AUH99791.1	Krings E, et al. (2006) J Bacteriol 188(23):8054-8061
2-keto-myo-inositol dehydratase	Geobacillus kaustophilus HTA426	BAD76175.1	Yoshida KI, et al. (2012) Microbiology 158(Pt 8):1942-1952
2-keto-myo-inositol dehydratase	Lactobacillus casei BL23	CAQ65355.1	Yebra MJ, et al. (2007) Appl Environ Microbiol 73(12):3850-3858
2-keto-myo-inositol dehydratase	Legionella pneumophila subsp. pneumophila str. Philadelphia 1	AAU27729.1	Manske C, Schell U, & Hilbi H (2016) Appl Environ Microbiol 82(16):5000-5014
2-keto-myo-inositol dehydratase	Mycoplasma hyopneumoniae 7448	AAZ53604.2	Galvao Ferrarini M, et al. (2018) Mol Microbiol 108(6):683-696

2-keto-myo-inositol dehydratase	Propionibacterium freudenreichii subsp. shermanii CIRM-BIA1	CBL57393.1	Loux V, et al. (2015) BMC Genomics 16:296
2-keto-myo-inositol dehydratase	Rhizobium leguminosarum bv. viciae 3841	CAK06988.1	Fry J, Wood M, & Poole PS (2001) Mol Plant Microbe Interact 14(8):1016-1025
2-keto-myo-inositol dehydratase	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	AAL23244.1	Kröger C & Fuchs TM (2009) J Bacteriol 191(2):545-554
2-keto-myo-inositol dehydratase	Sinorhizobium meliloti 1021	CAC41789.1	Kohler PR, Zheng JY, Schoffers E, & Rossbach S (2010) Appl Environ Microbiol 76(24):7972-7980
2-keto-myo-inositol dehydratase	Streptomyces coelicolor A3(2)	CAB88959.1	Yu L, Gao W, Li S, Pan Y, & Liu G (2016) Microbiology 162(3):537-551
2-keto-myo-inositol isomerase	Bacillus subtilis subsp. subtilis str. 168	QJF42344.1	Zhang RG, et al. (2002) Proteins 48(2):423-426; Yoshida K, et al. (2006) Appl Environ Microbiol 72(2):1310-1315
Ketose 3-epimerase	Agrobacterium fabrum str. C58	AAK88700.1	Kim K, Kim HJ, Oh DK, Cha SS, & Rhee S (2006) J Mol Biol 361(5):920-931; Kim HJ, Hyun EK, Kim YS, Lee YJ, & Oh DK (2006) Appl Environ Microbiol 72(2):981-985
Ketose 3-epimerase	Arthrobacter globiformis M30	BAW27657.1	Yoshida H, et al. (2018) Acta Crystallogr F Struct Biol Commun 74(Pt 10):669-676; Yoshihara A, et al. (2017) J Biosci Bioeng 123(2):170-176
Ketose 3-epimerase	[Clostridium] bolteae ATCC BAA-613	EDP19602.1	Jia M, et al. (2014) Appl Microbiol Biotechnol 98(2):717-725
Ketose 3-epimerase	[Clostridium] scindens ATCC 35704	EDS06411.1	Zhang W, et al. (2013) PLoS One 8(4):e62987
Ketose 3-epimerase	Clostridium sp. BNL1100	AEY67409.1	Mu W, et al. (2013) Biotechnol Lett 35(9):1481-1486
Ketose 3-epimerase	Desmospora sp. 8437	EGK07060.1	Zhang W, et al. (2013) J Agric Food Chem 61(47):11468-11476
Ketose 3-epimerase	Dorea sp. CAG:317	CDD07088.1	Zhang W, et al. (2015) J Mol Catal B Enzym 120:68-74
Ketose 3-epimerase	Flavonifractor plautii ATCC 29863	EHM40452.1	Park CS, et al. (2016) PLoS One 11(7):e0160044
Ketose 3-epimerase	Mesorhizobium japonicum MAFF 303099	BAB50456.1	Uechi K, Sakuraba H, Yoshihara A, Morimoto K, & Takata G (2013) Acta Crystallogr D Biol Crystallogr 69(Pt 12):2330-2339; Uechi K, Takata G, Fukai Y, Yoshihara A, & Morimoto K (2013) Biosci Biotechnol Biochem 77(3):511-515
Ketose 3-epimerase	Pseudomonas cichorii ST-24	BAA24429.1	Yoshida H, et al. (2007) J Mol Biol 374(2):443-453; Itoh H, et al. (1994) Biosci Biotechnol Biochem 58:2168–2171
Ketose 3-epimerase	Rhodobacter sphaeroides SK-011	ACO59490.1	Zhang L, Mu W, Jiang B, & Zhang T (2009) Biotechnol Lett 31(6):857-862
Ketose 3-epimerase	Ruminiclostridium cellulolyticum H10	ACL75304.1	Chan HC, et al. (2012) Protein Cell 3(2):123-131; Mu W, et al. (2011) J Agric Food Chem 59(14):7785-7792
Ketose 3-epimerase	Ruminococcus sp. 5_1_39BFAA	EES75522.1	Zhu Y, et al. (2012) Biotechnol Lett 34(10):1901-1906
Ketose 3-epimerase	Sinorhizobium sp. RAC02	AOF93213.1	Zhu Z, et al. (2019) RSC Adv 9:2919-2927
Ketose 3-epimerase	Treponema primitia ZAS-1	WP_010256447.1	Zhang W, Zhang T, Jiang B, & Mu W (2016) J Sci Food Agric 96(1):49-56
Xylose Isomerase	Actinoplanes missouriensis 431	P12851.3	Jenkins J, et al. (1992) Biochemistry 31(24):5449-5458; van Tilbeurgh H, et al. (1992) Biochemistry 31(24):5467-5471
Xylose Isomerase	Actinoplanes sp. ATCC 31351	AAA92578.1	Saari GC, Kumar AA, Kawasaki GH, Insley MY, & O'Hara PJ (1987) J Bacteriol 169(2):612-618
Xylose Isomerase	Arthrobacter sp. NRRL B3728	P12070.3	Collyer CA, Henrick K, & Blow DM (1990) J Mol Biol 212(1):211-235; Rangarajan M & Hartley BS (1992) Biochem J 283 (Pt 1):223-233
Xylose Isomerase	Bacillus subtilis subsp. spizizenii ATCC 6633	EFG92849.1	Amore R, Wilhelm M, & Hollenberg CP (1989) Appl Microbiol Biotechnol 30:351-357
Xylose Isomerase	Bacteroides stercoris HJ-15	AEK21499.1	Ha SJ, Kim SR, Choi JH, Park MS, & Jin YS (2011) Appl Microbiol Biotechnol 92(1):77-84
Xylose Isomerase	Burkholderia cenocepacia J2315	B4ENA5.1	Vieira IPV, et al. (2019) AMB Express 9(1):73
Xylose Isomerase	Escherichia coli K-12	AMH36887.1	Batt CA, Jamieson AC, & Vandeyar MA (1990) Proc Natl Acad Sci U S A 87(2):618-622

Xylose Isomerase	Lachnoclostridium phytofermentans ISDg	ABX41597.1	Seike T, et al. (2019) Biotechnol Biofuels 12:139
Xylose Isomerase	Lactobacillus reuteri KLR1002	OTA44610.1	Staudigl P, Haltrich D, & Peterbauer CK (2014) J Agric Food Chem 62(7):1617-1624
Xylose Isomerase	Lactococcus lactis subsp. lactis IO-1	BAL51426.1	Park JH & Batt CA (2004) Appl Environ Microbiol 70(7):4318-4325
Xylose Isomerase	Pectobacterium atrosepticum SCRI1043	CAG73017.1	Sapunova LI, Lobanok AG, Kazakevich IO, Shliakhotko EA, & Evtushenkov AN (2006) Appl Biochem Microbiol 42(3):246-251
Xylose Isomerase	Streptomyces rubiginosus	AAA26838.1	Whitlow M, et al. (1991) Proteins 9(3):153-173; Waltman MJ, Yang ZK, Langan P, Graham DE, & Kovalevsky A (2014) Protein Eng Des Sel 27(2):59-64
Xylose Isomerase	Streptomyces violaceoruber S21	ARF65364.1	Callens M, Kersters-Hilderson H, Van Opstal O, & De Bruyne CK (1996) Enzyme Microb Technol 8:696-700
Xylose Isomerase	Streptomyces violaceusniger	AAA26839.1	Lavie A, Allen KN, Petsko GA, & Ringe D (1994) Biochemistry 33:5469-5480; Suekane M, Tamura M, & Tomimura C (1978) Agric Biol Chem 42(5):909-917
Xylose Isomerase	Thermoanaerobacter ethanolicus JW 200	ACU01780.1	Fan L, Zhang Y, Qu W, Wang J, & Shao W (2011) Biotechnol Lett 33(3):593-598
Xylose Isomerase	Thermoanaerobacterium saccharolyticum PB8	AIE39924.1	Lee YE, Ramesh MV, & Zeikus JG (1993) J Gen Microbiol 139 Pt 6:1227-1234
Xylose Isomerase	Thermotoga neapolitana 5068	AAB06798.1	Vieille C, Hess JM, Kelly RM, & Zeikus JG (1995) Appl Environ Microbiol 61(5):1867-1875
Xylose Isomerase	Thermus thermophilus HB8	BAL42599.1	Lehmacher A & Bisswanger H (1990) Biol Chem Hoppe Seyler 371(6):527-536
Xylose Isomerase	Vibrio sp. XY-214	BAI23199.1	Umemoto Y, Shibata T, & Araki T (2012) Mar Biotechnol (NY) 14(1):10-20
L-xylulose-5-P 3-epimerase	Escherichia coli O157:H7 str. EDL933	AAG59393.1	Shi R, et al. (2008) J Bacteriol 190(24):8137-8144; Yew WS & Gerlt JA (2002) J Bacteriol 184(1):302-306
L-xylulose-5-P 3-epimerase	Klebsiella pneumoniae ATCC 13882	ABF60040.1	Campos E, et al. (2008) J Bacteriol 190(20):6615-6624