Biochimica et Biophysica Acta 1840 (2014) 464–475

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Dissecting cobamide diversity through structural and functional analyses of the base-activating CobT enzyme of *Salmonella enterica*

Chi Ho Chan [a,1,2], Sean A. Newmister [b,1,3], Keenan Talyor [b], Kathy R. Claas [a], Ivan Rayment [b], Jorge C. Escalante-Semerena [c,*]

[a] Department of Bacteriology, University of Wisconsin, Madison, WI, USA
[b] Department of Biochemistry, University of Wisconsin, Madison, WI, USA
[c] Department of Microbiology, University of Georgia, Athens, GA, USA

ARTICLE INFO

Article history:
Received 13 June 2013
Received in revised form 20 August 2013
Accepted 30 September 2013
Available online 10 October 2013

Keywords:
B_{12} biosynthesis
Lower base activation
Directed enzyme evolution
Cobamide diversity

ABSTRACT

Background: Cobamide diversity arises from the nature of the nucleotide base. Nicotinate mononucleotide (NaMN):base phosphoribosyltransferases (CobT) synthesize α-linked riboside monophosphates from diverse nucleotide base substrates (e.g., benzimidazoles, purines, phenolics) that are incorporated into cobamides.
Methods: Structural investigations of two members of the CobT family of enzymes in complex with various substrate bases as well as in vivo and vitro activity analyses of enzyme variants were performed to elucidate the roles of key amino acid residues important for substrate recognition.
Results: Results of in vitro and in vivo studies of active-site variants of the *Salmonella enterica* CobT (*Se*CobT) enzyme suggest that a catalytic base may not be required for catalysis. This idea is supported by the analyses of crystal structures that show that two glutamate residues function primarily to maintain an active conformation of the enzyme. In light of these findings, we propose that proper positioning of the substrates in the active site triggers the attack at the C1 ribose of NaMN.
Conclusion: Whether or not a catalytic base is needed for function is discussed within the framework of the in vitro analysis of the enzyme activity. Additionally, structure-guided site-directed mutagenesis of *Se*CobT broadened its substrate specificity to include phenolic bases, revealing likely evolutionary changes needed to increase cobamide diversity, and further supporting the proposed mechanism for the phosphoribosylation of phenolic substrates.
General Significance: Results of this study uncover key residues in the CobT enzyme that contribute to the diversity of cobamides in nature.

1. Introduction

A cobamide (Cba) is a cobalt-containing cyclic tetrapyrrole involved in enzyme catalyzed carbon skeleton rearrangements, methyl-group transfers and reductive dehalogenation [1,2]. Cobamides are related to heme, chlorophyll and coenzyme F_{430} in structure though they play different roles in nature [3]. A diagram of cobamides is shown in Fig. 1. In cobamides, nitrogen atoms of the tetrapyrrole rings equatorially coordinate a cobalt ion. In their coenzyme form, cobamides have a 5′-deoxyadenosyl (Ado) group as the axial *Coβ* ligand that participates in radical chemistry reactions such as those catalyzed by methylmalonyl-CoA mutase, ethanolamine ammonia-lyase and diol dehydratase [4–6]. Cobamides also serve as a transient methyl group carrier in methyltransferase reactions [7].

Vitamin B_{12} (cyanocobalamin, CNCbl) is a cobamide that is an essential nutrient to humans. In CNCbl, the axial *Coα* ligand is 5,6-dimethylbenzimidazole (DMB) and the axial *Coβ* ligand is a cyano (CN) group. Cobamides differ from each other by the nature of their nucleotide base, which ranges from DMB, to purines, to phenolics [8].

The biosynthesis of cobamides is complex and is only performed by some archaea and bacteria [9]. Adenosylcobamide (AdoCba) biosynthesis has been studied extensively in several bacteria and some archaea [10]. Given the complexity of the molecule, it is not surprising that a large number of gene products are required for its assembly. For example, in *Salmonella enterica*, the AdoCba biosynthetic pathway involves at least 25 gene products. Some microorganisms synthesize a range of different cobamides. For example, *S. enterica* synthesizes three, one that contains DMB as the nucleotide base, and two others that contain adenine and 2-methyladenine in lieu of DMB [11,12]. When adenine is the nucleotide base, this cobamide is known as pseudo-B_{12}. The reasons that drive

* Corresponding author at: Department of Microbiology, University of Georgia, 527 Biological Sciences Building, 120 Cedar Street, Athens, GA 30602, USA. Tel.: +1 706 542 2651; fax: +1 706 542 2815.
E-mail address: jcescala@uga.edu (J.C. Escalante-Semerena).
[1] These authors contributed equally to this work.
[2] Present address: BioTechnology Institute, University of Minnesota, St. Paul, MN 55108.
[3] Present address: Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109.

0304-4165/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagen.2013.09.038

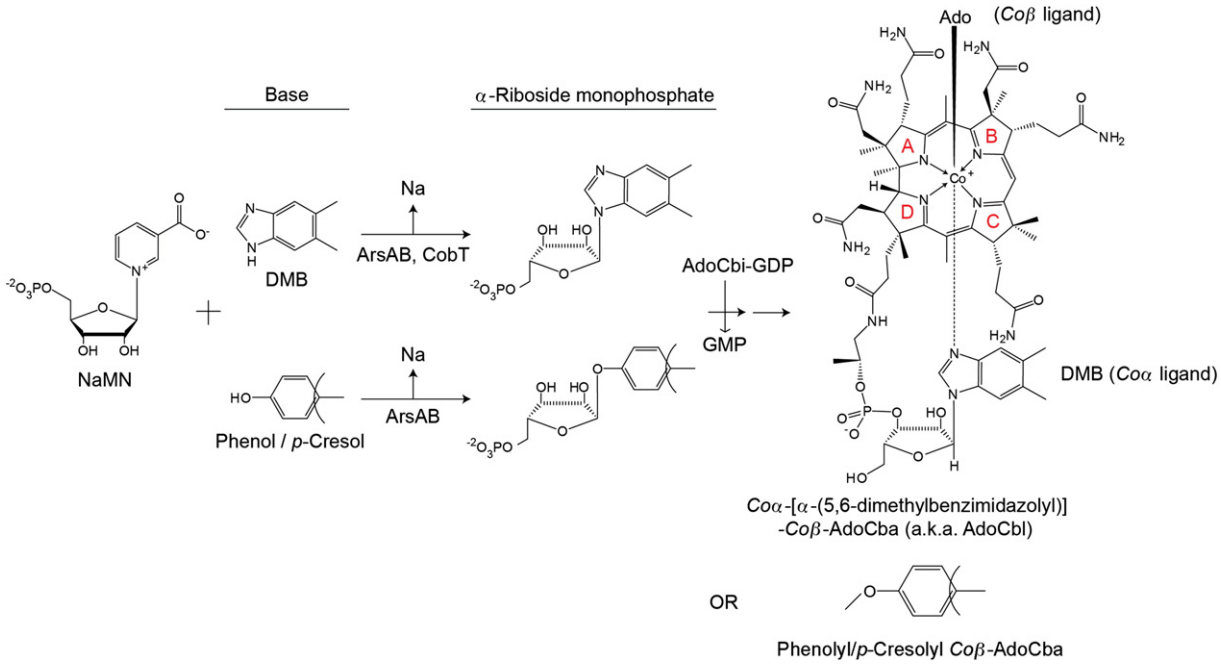

Fig. 1. Schematic of the nicotinate mononucleotide (NaMN):base phosphoribosyltransferase reaction (left). The α-riboside monophosphate formed is combined with precursor adenosylcobinamide-GDP (AdoCbi-GDP) to produce adenosyl-cobamide (right). Phenolic bases do not form an α-coordination bond with the cobalt ion.

the need to synthesize different cobamides are unclear, but such diversity may reflect physiological needs under different conditions.

In *S. enterica*, the nicotinate mononucleotide (NaMN):base phosphoribosyltransferase (*Se*CobT) enzyme catalyzes the synthesis of α-riboside monophosphates (α-ribotides) that are subsequently utilized in AdoCba biosynthesis (Fig. 1) [13]. In vitro, *Se*CobT was shown to phosphoribosylate a wide range of benzimidazole and purine analogs that are not produced by *S. enterica*, showing that this enzyme has broad substrate specificity [13,14]. Products formed with benzimidazole and purine analogs feature an α-*N*-glycosidic bond between N1 of imidazole and C1 of ribose. In *S. enterica*, the CobB sirtuin deacetylase compensates for the lack of CobT during the assembly of the nucleotide loop in Δ*cobT* strains [15]. Thus, to eliminate background activity, all strains used in this work carried null alleles of *cobT* and *cobB*.

Previous structural studies of *Se*CobT revealed how the active site accommodates diverse aromatic compounds, and a general-base catalysis mechanism through a conserved glutamate residue was proposed [16]. Remarkably, products were captured in the structure when *Se*CobT was co-crystallized with aromatic bases and NaMN with little movement in the active site [14,17]. Interestingly, *Se*CobT cannot catalyze the phosphoribosylation of phenolic bases.

In contrast, the acetogenic anaerobe *Sporomusa ovata* [18] synthesizes cobamides with phenolic bases [19]. Incorporation of phenolic compounds into cobamides is catalyzed by the *Se*CobT homolog ArsAB in *S. ovata* [20]. *So*ArsAB phosphoribosylates phenol and *p*-cresol yielding phenolic α-ribotides that have a unique α-*O*-glycosidic bond between the oxygen in phenols and the C1 of ribose (Fig. 1). Like *Se*CobT, *So*ArsAB also incorporates DMB and purine analogs into α-ribotides. Unlike *Se*CobT, *So*ArsAB is a heterodimeric enzyme, where both subunits are *Se*CobT orthologs. Differences in the structures of *So*ArsAB and *Se*CobT have been proposed to account for the phenolic activity in *So*ArsAB [21].

The work reported here expands our understanding of the active site of the CobT family of enzymes. We demonstrate that two glutamate residues (Glu174 and Glu317) in the active site that are important for activity are also critical to the positioning of the base substrate. Whether either one of these residues plays a role as a catalytic base remains an open question. Crystal structures of variant proteins in which either glutamate was changed have been determined to a 1.9 Å resolution and provided insight into the roles of these residues in the activity of this enzyme. To gain a greater understanding of the residues involved in substrate recognition and activity, *Se*CobT was engineered to synthesize α-*p*-cresolyl-riboside monophosphate using structure-guided analyses of the *So*ArsAB active site. Structures of this engineered variant in complex with *p*-cresol, *p*-cresol and NaMN at a 2.0 Å resolution illustrate some of the structural changes *Se*CobT needs to undergo to be able to use phenolic substrates.

2. Materials and methods

2.1. Growth media, conditions and analysis

Cultures were grown at 37 °C in a no-carbon essential (NCE) minimal medium [22] supplemented with glycerol (22 mM) as the carbon source. $MgSO_4$ (1 mM), trace minerals [23,24], ampicillin (100 μg/mL), kanamycin (50 μg/mL) and chloramphenicol (12.5 μg/mL) were used where indicated. For all growth experiments, strains were grown in triplicate in 96-well microtiter dishes; a sample (2 μL) of overnight cultures grown in lysogenic broth (LB) [25,26] was used to inoculate 198 μL of fresh minimal medium. Growth was analyzed using a computer-controlled BioTek ELx808 Ultra microplate reader (BioTek Instruments, Inc.). Cell density measurements at 650 nm were acquired every 1800 s for 36 h; plates were shaken for 1745 s between readings. Data were analyzed using the GraphPad Prism v4 software package (GraphPad Software).

2.2. Strain and plasmid constructions

All strains used in this study were derivatives of *S. enterica* sv Typhimurium strain LT2 or *Escherichia coli* strain DH5α. All *S. enterica* strains are derived from strain TR6583. Strain genotypes are described in Table 5. Plasmids were introduced by electroporation into *S. enterica* and by chemically competent transformation by heat shock into *E. coli*. All primers used in this work are listed in Table 5.

Table 1
Data collection and refinement statistics.

complex	317A	317A	174A	174A	YMM	YMM
		α-RP + NA	DMB	adenine	*p*-cresol	*p*-cresol + NaMN
pdb id	4KQH	4KQI	4KQG	4KQF	4KQK	4KQJ
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_1$	$P2_12_12$
Wavelength	1.54	0.979	1.54	1.54	0.979	1.54
Source	Cu K_α	19BM	Cu K_α	Cu K_α	19 BM	Cu K_α
Resolution range[a]	50–1.9 (2.0–1.9)	50–1.4 (1.42–1.4)	50–1.9 (2.0–1.9)	50–1.9 (2.0–1.9)	50–1.47 (1.5–1.47)	50–1.9 (2.0–1.9)
Reflections: measured	288139	989867	209381	303189	513574	257886
Reflections: unique	24438	59605	24307	24215	97861	22398
Redundancy	11.8 (8.4)	8.7 (5.5)	8.6 (5.3)	12.5 (9.0)	3.2 (2.7)	11.4 (6.1)
% Complete	99.9 (99.3)	100 (99.8)	99.9 (99.2)	99.8 (98.8)	95.2 (88.9)	99.0 (93.5)
Average I/σ	14.5 (6.3)	78.0 (9.8)	9.9 (3.8)	16.4 (7.3)	26.6 (4.6)	14.2 (5.1)
R_{merge} (%)[b]	11.3 (26.5)	4.6 (17.9)	15.5 (35.2)	10.0 (23.5)	3.1 (16.1)	12.7 (29.6)
R_{cryst}[c]	16.9	17.3	17.0	15.3	16.8	18.2
R_{free}	22.4	19.5	22.5	22.0	19.5	22.2
Protein atoms	2441	2451	2509	2499	5077	2472
Ligand atoms	19	37	20	24	44	34
Waters	260	252	296	285	411	175
Avg. B (Å^2)	12.6	17.9	8.4	12.5	14.8	8.9
Ramachandran (%)	––	––	––	––	––	––
Most favored	98.8	98.3	97.7	98.6	98.2	97.7
Allowed	0.9	1.4	1.7	1.1	1.5	1.7
Disallowed	0.3	0.3	0.6	0.3	0.3	0.3
rms deviations	––	––	––	––	––	––
Bond lengths (Å)	0.019	0.019	0.019	0.020	0.013	0.018
Bond angles (deg)	1.99	1.99	1.97	1.89	1.67	2.07

R_{work} refers to the R_{factor} for the data utilized in the refinement and R_{free} refers to the R_{factor} for 5% of the data that were excluded from the refinement.

[a] Data in parentheses represent highest resolution shell.

[b] $R_{merge} = (\sum |I_{(hkl)} - I| \times 100.) / \sum |I_{(hkl)}|$, where the average intensity I is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given observation.

[c] $R_{factor} = \sum | F_{obs} - F_{calc} | \times 100. / \sum | F_{obs} |$.

2.2.1. Plasmid pCOBT15

This plasmid encoded *Se*CobT^{E317A}, and was generated with two sets of overlapping primers (cobT13 and cobT14) coding for the amino acid change using plasmid pCOBT10 DNA as template [27]. Briefly, primers cobT13 and cobT14 encoding the E317A changes were used to amplify the *cobT* gene on pCOBT10. The amplification product was transformed in to *E. coli* DH5α, and its sequenced verified.

2.2.2. Plasmid pCOBT19

This plasmid was a derivative of plasmid pT7-5 [28], and encoded variant *Se*CobT^{E317A}. The *cobT* allele encoding *Se*CobT^{E317A} was subcloned from pCOBT15 using BamHI and HindIII restriction enzymes.

2.2.3. Plasmids pCOBT42 and pCOBT57

These plasmids encoded *Se*CobT^{E174A} and CobT^{E174A,E317A} variants, respectively, and were constructed following the QuikChange™ site-directed mutagenesis protocol (Stratagene) using plasmids pJO27 and pCOBT19 DNA as templates, respectively.

2.2.4. Plasmid pCOBT141

This plasmid encoded *Se*CobT^{WT} fused to an *N*-terminal hexahistidine (His_6) tag. Plasmid pCOBT141 was constructed using polymerase incomplete primer extension (PIPE)-based cloning [29].

Table 2
Specific activities of *Se*CobT variants on DMB and adenine.

Variant	Specific Activity (μmol min^{-1} mg^{-1})	
	DMB	Adenine
*Se*CobT^{WT}	30 ± 2	1.6 ± 0.6
*Se*CobT^{E174A}	0.5 ± 0.4	0.03 ± 0.01
*Se*CobT^{E317A}	1.5 ± 0.4	0.07 ± 0.01
*Se*CobT^{E174A,E317A}	2e-3 ± 1e-3	1e-3 ± 2e-4

Conditions used for the assay are described under the Materials and methods section.

2.2.5. Plasmid pCOBT184

This plasmid encoded the *Se*CobT^{S80Y,Q88M,L175M} variant, and was constructed following the QuikChange™ site-directed mutagenesis protocol. Each substitution was introduced one at a time into plasmid pCOBT141. *So*ArsAB variants were constructed following the QuikChange™ site-directed mutagenesis protocol (Stratagene) using plasmid pARSAB4 DNA as template [20]. Primers were purchased from IDT http://www.idtdna.com/ and designed using the PrimerX website http://www.bioinformatics.org/primerx/.

2.3. In vivo assessments of SeCobT and SoArsAB functions by cobalamin-dependent methionine biosynthesis

S. enterica strains JE2607 (*cobB cobT recA*) and JE2501(*cobB cobT*) were used to assess the function of *Se*CobT and *So*ArsAB variants, respectively, under conditions that demanded Cbl-dependent methionine synthesis. All strains lacked the *cobB* gene to ensure that base incorporation was due to plasmid-encoded *Se*CobT variants. Generation times were calculated using the equation $g = n / t$, where g is the generation time, n is the number of generations, and t is the time it took to reach the final cell population. To calculate the number of generations (n), we used the equation $N_t = N_0 2^n$, where N_t and N_0 are the final and initial optical densities (OD) of the culture, respectively.

2.4. Purification of SeCobT variants

*Se*CobT variant proteins encoded in pCOBT19, pCOBT42 and pCOBT57 were overproduced in strain JE2017 (*cobA*/pGP1-2 T7 *rpo*$^+$ (T7 RNAP) *kan*$^+$) and purified as described [13]. The *Se*CobT variant encoded in pCOBT184 was overproduced in strain JE13607, an *E. coli* BL21(λDE3) derivative lacking the *cobT* gene [20]. This variant was overproduced and purified using a procedure similar to the one described for the purification of *So*ArsAB variants [21]. *Se*CobT variants concentrated to 6 mg/mL were dialyzed in Tris–HCl buffer (20 mM, pH 7.5) containing NaCl (100 mM) before storing them at −80 °C [17] until used.

Table 3
Doubling time (h) of a *S. enterica cobT cobB* strain harboring plasmids encoding *Se*CobT variants.

Variant	Cbi	Cbi + DMB	Cbi + Adenine	Cbi + Cbl
Vector only	NG	NG (600 μM)	NG (1 mM)	2 ± 0.1
*Se*CobT^{WT}	2 ± 0.1	2 ± 0.1 (10 nM)	2 ± 0.1 (10 nM)	2 ± 0.1
*Se*CobT^{E174A}	NG	19 ± 5 (1 μM) 2 ± 0.4 (10 μM)	NG (1 mM)	2 ± 0.3
*Se*CobT^{E317A}	NG	25 ± 3 (10 nM) 4 ± 0.1 (50 nM)	35 ± 4 (0.5 mM) 3 ± 0.5 (1 mM)	2 ± 0.1
*Se*CobT^{E174A,E317A}	NG	NG (0.6 mM)	NG (1 mM)	2 ± 0.1

Cells were grown in NCE minimal medium with glycerol as the sole source of carbon and energy. Indicated in parentheses is the concentration of base substrate present in the reaction mixture. Up to 1 mM (for adenine) and 0.6 mM (for DMB) were used to restore growth to WT levels as indicated. Cbi, [$(CN)_2$Cbi]; Cbl, (CNCbl); NG, no growth.

2.5. Reaction with p-cresol as substrate

Phosphoribosylation reactions using *Se*CobT^{WT}, *Se*CobT^{S80Y,Q88M,L175M} or *So*ArsAB with NaMN and *p*-cresol were performed at 37 °C for 6 h. Each reaction was performed in 600 μL of glycine buffer (100 mM, pH 10) containing the enzymes (60 μg each) NaMN and *p*-cresol (3 mM each). The reaction mixture was applied onto a C18 Sep-Pak column equilibrated and washed with water and eluted with 100 μL of methanol. Ten-microliter samples were analyzed on a Shimadzu HPL chromatograph equipped with a Kinetex 2.6 μm C18 100 Å (100 × 4.60 mm) column (Phenomenex) as described [21].

2.6. Synthesis and purification of nicotinic acid mononucleotide (NaMN)

At the time this work was performed, NaMN was not commercially available hence it was synthesized in house. The enzymatic synthesis of NaMN was modified from a published method [30]. Briefly, a 3-mL reaction mixture containing sodium phosphate buffer, pH 7.0 (250 μmol), 5-phosphorylribose 1-pyrophosphate (PRPP, 5 μmol), quinolinic acid (QA, 5 μmol), magnesium chloride (100 μmol), and phosphoribosylpyrophosphate (PRPP)-quinolinate phosphoribosyltransferase (NadC, 5 mg) was incubated at 37 °C. NadC was purified using the ASKA library (A Complete Set of *E. coli* genes K-12 ORF Archive) [31]. To circumvent substrate inhibition, additional 5 μmol of PRPP and QA were added to the reaction mixture at 1.5 h intervals. A total of 20 μmol each of PRPP and QA was added, and the reaction mixture was incubated for a total of 6 h. Ten 3-mL NadC reactions were performed, all of which were pooled and desalted over a 225-mL Sephadex G-10 column (GE Healthcare) at room temperature. Deionized water was further stripped of contaminants using a Millipore Milli-Q water purification system. Water treated this way was used to equilibrate the column and to resolve components of the NadC reaction mixture at a rate of 1 mL/min using a GE Healthcare P1 peristaltic pump. Fractions were collected and absorbance was measured at 265 nm to identify fractions containing NaMN. The latter were lyophilized and solids were re-suspended in water for purification using DEAE chromatography as described elsewhere [32]. Three 5-mL GE Healthcare HiTrap DEAE FF (FastFlow) pre-packed columns were equilibrated with water at a rate of 2 mL/min using an ÄKTA explorer FPLC system (GE Healthcare). Desalted products of the NadC reaction were loaded onto the column, which was washed with 30 mL of water. NaMN was eluted off the column with NaCl (30 mM). To elute the remaining reaction components, a linear gradient of 300 mM NaCl was used. NaMN elution was monitored at 260 nm. Purified NaMN was desalted as outlined above. Conductivity (YSI Scientific Conductance Meter Model 35) was measured to determine whether or not NaMN was salt-free. A standard curve of NaCl concentrations was used to determine that <1 mM NaCl remained in the solution. The NaMN concentration was determined using the molar extinction coefficient at 265 nm of 3300 M^{-1}. NaMN was lyophilized and the solid was divided into 2-mL serum vials containing approximately 5 mg, then flushed with nitrogen before sealing and storing at −20 °C.

Table 4
Doubling times (h) of a *S. enterica cobT cobB* strain harboring plasmids encoding *So*ArsAB variants.

Variant	Cbi	Cbi + DMB	Cbi + Cbl
Vector only	NG	NG	2.4 ± 0.01
*So*ArsAB^{WT}	3.4 ± 0.1	2.7 ± 0.01	2.5 ± 0.01
*So*ArsA^{E176A}ArsB	NG	2.7 ± 0.01	2.8 ± 0.02
*So*ArsA^{E319A}ArsB	NG	2.6 ± 0.02	2.3 ± 0.01
*So*ArsA^{E176A,E319A}ArsB	NG	NG	3 ± 0.01
*So*ArsAArsB^{E171A,E303A}	4.2 ± 0.04	4.0 ± 0.03	3.2 ± 0.02

Cells were grown in NCE minimal medium with glycerol as the sole source of carbon and energy. When added to the culture medium Cbi, [$(CN)_2$Cbi], was at 15 nM; Cbl, [CNCbl], was at 15 nM; and DMB was at 150 μM; NG, no growth.

2.7. In vitro phosphoribosyltransferase activity assay

Conditions for the phosphoribosyltransferase activity assay when NaMN and DMB are substrates are described elsewhere [33]. These conditions were also used when measuring activity with adenine. [^{14}C,C-8] Adenine was purchased from Moravek Biochemicals, Inc., and [^{14}C,C-2] DMB was synthesized as described [33]. *Se*CobT proteins were diluted in glycine–NaOH buffer (50 mM, pH 10.0) containing bovine serum albumin (0.05% wt/v) and sodium azide (4.6 mM). KCl (0.5 M final) was added to stop the reactions. Specific activities were determined for the formation of α-DMB-riboside monophosphate (α-RP, when DMB was used as substrate), and of α-adenine-riboside monophosphate (α-AMN, when adenine was used as substrate). The rate of α-RP synthesis was performed using 50 ng of *Se*CobT^{E174A} and *Se*CobT^{E317A} variants, while the rate of α-AMN synthesis required 1.25 μg of *Se*CobT^{E174A} and *Se*CobT^{E317A} variants, respectively. Control experiments employing *Se*CobT^{WT}, required 2.5 ng of enzyme for the synthesis of α-RP, and 25 ng of enzyme for the synthesis of α-AMN. Substrate concentrations were used at saturating levels for each protein and initial rates were measured. Activities for α-RP and α-AMN formations were assayed twice over a period of 0.75 and 1.5 h, respectively. The specific activity of *Se*CobT^{E174A,E317A} was determined for α-RP and α-AMN formation using 1 μg and 17 μg of protein, respectively, over a period of 2 and 4 h, respectively.

2.8. Mass spectrometry

The molecular mass of α-*p*-cresolyl-riboside monophosphate was determined by electrospray ionization (ESI) mass spectrometry on an Agilent LC/MSD TOF spectrometer as described elsewhere [20].

2.9. Crystallization of SeCobT^{E174A} in complex with adenine and DMB

All protein and substrate complexes were screened at room temperature by hanging drop vapor diffusion using a 144-condition sparse matrix screen developed in the Rayment Laboratory. Single, diffraction quality crystals of *Se*CobT^{174A} in complex with adenine were grown by mixing equal volumes of 6 mg/mL *Se*CobT^{174A}, adenine (1 mM), Tris-HCl buffer (20 mM, pH 7.5), NaCl (100 mM), and dimethylsulfoxide (DMSO, 1%, v/v) with a reservoir solution composed of 2-(*N*-morpholino)ethanesulfonic acid (MES-NaOH) buffer (100 mM, pH 6.0) containing $(NH_4)_2SO_4$ (1.05 M), and glycerol (5%, v/v). Hanging drops were nucleated after 4 h with a fine cat whisker. The crystals grew to a size of 0.4 × 0.4 × 0.3 mm within 4 days and were transferred to a synthetic mother liquor composed of MES-NaOH buffer (100 mM, pH 6.0), $(NH_4)_2SO_4$ (1 M), adenine (1 mM), DMSO (1%, v/v), and glycerol (5%, v/v). The crystals were transferred into a cryoprotectant solution with the same composition except for the presence of glycerol (17%, v/v). The crystals were flash cooled in liquid N_2.

Single, diffraction quality crystals of *Se*CobT^{174A} in complex with DMB were grown by mixing equal volumes of 6 mg/mL *Se*CobT^{E174A},

Table 5
List of strains, plasmids and primers used in this study.

Strains	Genotype	Reference
S. enterica strains		
JE2017	*cobA367*::Tn*10*d(*tet*$^+$)/pGP1-2 T7 *rpo*$^+$ (T7 RNAP) *kan*$^+$	Lab collection
TR6583	*metE205 ara-9*	K. Sanderson via J. Roth
Derivatives of TR6583		
JE2501	*cobB1176*::Tn*10*d(*tet*$^+$) *cobT109*::MudJ	[46]
JE2607	*cobB1176*::Tn*10*d(*tet*$^+$) *cobT109*::MudJ *recA1*	[15]
E. coli strains		
JE13607	BL21(λDE3) *cobT762*::*kan*$^+$	[20]
Plasmid	**Genotype and Description**	
pTEV5	Overexpression vector that fuses the *N*-terminus of the protein of interest to a H_6 tag, which can be removed by rTEV protease, *bla*$^+$	[47]
pBAD33	Complementation vector P_{araBAD} *cat*$^+$	[48]
pT7-5	Cloning vector, *bla*$^+$	[49]
pJO27	*S.e. cobT*$^+$ encodes *Se*CobT^{WT} in pT7-5, *bla*$^+$	[50]
pCOBT10	*S.e. cobT*$^+$ encodes *Se*CobT^{WT} in pSU18, *cat*$^+$	[33]
pCOBT15	*S.e. cobT1475* (encodes *Se*CobT^{E317A}) in pSU18, *cat*$^+$	
pCOBT19	*S.e. cobT1475* (encodes *Se*CobT^{E317A}) in pT7-5, *bla*$^+$	
pCOBT42	*S.e. cobT1476* (encodes *Se*CobT^{E174A}) in pT7-5, *bla*$^+$	
pCOBT57	*S.e. cobT1477* (encodes SeCobT$^{E174A,\ E317A}$) in pT7-5, *bla*$^+$	
pCOBT141	*S.e. cobT*$^+$ encodes *Se*CobT^{WT} in pTEV5, *bla*$^+$	
pCOBT184	*S.e. cobT1478* (encodes *Se*CobT$^{S80Y,\ Q88M,\ L175M}$) in pTEV5, *bla*$^+$	
pARSAB4	*S.o.* arsAB$^+$ encodes *So*ArsAB^{WT} in pBAD33, *cat*$^+$	[20]
pARSAB8	*S.o. arsAB* (ArsA^{E176A}ArsB) in pBAD33, *cat*$^+$	
pARSAB9	*S.o. arsAB* (ArsA^{E319A}ArsB) in pBAD33, *cat*$^+$	
pARSAB16	*S.o. arsAB* (ArsA^{E176A,E319A}ArsB) in pBAD33, *cat*$^+$	
pARSAB18	*S.o. arsAB* (ArsA–ArsB^{E171A,E303A}) in pBAD33, *cat*$^+$	
Primer Name	**Primer Sequence**	
cobT13 5′	TAT TTG GCG TAG GGG AAC	
cobT13 3′	CGC TTC CTG CAC CTA ACC	
cobT14 5′	TGC AGG AAG CGG CGC GGC	
cobT14 3′	GTT TTC CCA GTC ACG AC	
pTEV5-PIPE-5′	GCC CTG AAA ATA CAG GTT TTC ACT AGT TG	
pTEV5-PIPE-3′	CCA TGG AAT TCT CGA GCT CCC G	
CobT PIPE pTEV5 5′	AAC CTG TAT TTT CAG GGC ATG CAG ACA CTA CAC GCT TTA CTC CGT G	
CobT PIPE pTEV5 3′	AGC TCG AGA ATT CCA TGG TTA TGT TGC GTT TGC GTT CCC CTC	
CobT E174A 5′	AGG GGC GCT GGG AAT GGC GAA C	
CobT E174A 3′	GCG CCC CTA CGC CAA ATA AGG T	
CobT S80Y 5′	GAA GGC GTA GCG GTT TAT CCC AAA ATC GTG ACG	
CobT S80Y 3′	CGT CAC GAT TTT GGG ATA AAC CGC TAC GCC TTC	
CobT Q88M 5′	CAA AAT CGT GAC GGC GAT TAT GGC GGC GAA TAT GAC GCG	
CobT Q88M 3′	CGC GTC ATA TTC GCC GCC ATA ATC GCC GTC ACG ATT TTG	
CobT L175M 5′	CTT ATT TGG CGT AGG GGA GAT GGG AAT GGC GAA CAC TAC	
CobT L175M 3′	GTA GTG TTC GCC ATT CCC ATC TCC CCT ACG CCA AAT AAG	
ArsA E176A 5′	CTG TTT TTG TCT GGG AGC GAT GGG CAT TGG TAA TAC	
ArsA E176A 3′	GTA TTA CCA ATG CCC ATC GCT CCC AGA CAA AAA CAG	
ArsA E319A 5′	AGT CCG GCT GGG AGC GGG TAT CGG GGC TTC	
ArsA E319A 3′	GAA GCC CCG ATA CCC GCT CCC AGC CGG ACT	
ArsB E171A 5′	GTG ATC GGC TTG GGC GCG ATG GGG CTG GGC GGT TT	
ArsB E171A 3′	AAA CCG CCC AGC CCC ATC GCG CCC AAG CCG ATC AC	
ArsB E303A 5′	CTT AAA ATG AAC CTG GGA GCG GGG ACA GGT GCA GCA CTC	
ArsB E303A 3′	GAG TGC TGC ACC TGT CCC CGC TCC CAG GTT CAT TTT AAG	

DMSO (1 mM), Tris–HCl buffer (20 mM, pH 7.5), NaCl (100 mM), and ethanol (0.85%, v/v) with a reservoir solution composed of MES-NaOH buffer (100 mM, pH 6.0), methyl ether polyethylene glycol 5000 (MEPEG 5K, 10%, w/v), $(NH_4)_2SO_4$ (200 mM), and glycerol (5%, v/v). Hanging drops were nucleated after 4 h with a fine cat whisker. Crystals grew to a size of 0.3 × 0.3 × 0.2 mm within 4 days and were transferred to a synthetic mother liquor composed of MES-NaOH buffer (95 mM, pH 6.0), MEPEG 5K (9.5%, w/v), $(NH_4)_2SO_4$ (190 mM) NaCl (95 mM), DMB (1 mM), ethanol (0.85%, v/v), and glycerol (4.75%, v/v). The crystals were transferred to a cryoprotectant solution with the same composition except for the presence of MEPEG 5K (11.5%, w/v) and glycerol (17%, v/v) in a stepwise fashion. The crystals were flash frozen in liquid N_2.

2.10. Crystallization of SeCobT^{E317A} in the substrate free state and in complex with reaction products α-DMB riboside monophosphate and nicotinate

Single, diffraction-quality crystals of *Se*CobT^{E317A} were grown by mixing equal volumes of 6 mg/mL *Se*CobT^{E317A}, Tris–HCl buffer (20 mM, pH 7.5), and NaCl (100 mM) with a reservoir solution composed of MES-NaOH buffer (100 mM, pH 6.0), polyethylene glycol 1500 (PEG 1.5K, 20%, w/v), Li_2SO_4 (50 mM) and ethylene glycol (2.5%, v/v). Crystals were also grown in the presence of either adenine (1 mM) or DMB (1 mM), however no electron density was observed for either substrate. The mixture was centrifuged at 16,000 ×*g* for 10 min to remove nuclei. Four-microliter droplets of the cleared mixture were nucleated with a fine cat whisker and suspended over a reservoir solution. Crystals

grew to a size of 0.4 × 0.4 × 0.3 mm within 7 days and were transferred to a synthetic mother liquor composed of MES-NaOH buffer (95 mM, pH 6.0), PEG 1.5K (19%, w/v), Li_2SO_4 (48 mM), NaCl (95 mM), adenine (1 mM), DMSO (1%, v/v), and ethylene glycol (2.4% v/v). The crystals were transferred to a cryoprotectant solution with the same composition except for the presence of PEG 1.5K (20%, w/v) and ethylene glycol (16.2%, v/v). Crystals with reaction products α-DMB riboside monophosphate and nicotinate were obtained by transferring the crystals into the same synthetic mother liquor as above with the addition of 1 mM DMB and 10 mM NaMN. The crystals were soaked for 12 h and transferred to the same cryoprotectant solution as above and flash cooled in liquid N_2.

2.11. Crystallization of *Se*CobT^{S80Y,Q88M,L175M} in the presence of p-cresol

Single, diffraction quality crystals of *Se*CobT^{S80Y,Q88M,L175M} were grown by mixing equal volumes of 6 mg/mL *Se*CobT^{S80Y,Q88M,L175M}, Tris–HCl buffer (20 mM, pH 7.5), and NaCl (100 mM) with a reservoir solution composed of *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES–NaOH) buffer (100 mM, pH 9.0), MEPEG 5K (24%, w/v), $MgSO_4$ (50 mM), *p*-cresol (5 mM), DMSO (1%, v/v), and ethylene glycol (5%, v/v). Four-microliter droplets were nucleated with a fine cat whisker and suspended over a reservoir solution. Crystals grew to a size of 0.3 × 0.3 × 0.2 mm within 7 days and were transferred to a synthetic mother liquor composed of CHES–NaOH buffer (100 mM, pH 9.0), MEPEG 5K (24%, w/v), $MgSO_4$ (50 mM) NaCl (100 mM), *p*-cresol (5 mM), DMSO (1%, v/v), and ethylene glycol (5%, v/v). The crystals were transferred to a cryoprotectant solution with the same composition except for the presence of 17% ethylene glycol. The crystals were flash cooled in liquid N_2. Crystals with *p*-cresol and NaMN were obtained by transferring the crystals to the above cryoprotectant solution with the addition of 10 mM NaMN. Crystals that were frozen after 5 min of incubation contained both reaction products, whereas crystals that were frozen after 10 days of incubation contained only *p*-cresol.

2.12. Data collection and refinement

All except one X-ray diffraction data set were collected at 100 K with a Bruker AXS Platinum 135 CCD detector equipped with Montel optics and controlled by the Proteum software suite (Brucker AXS Inc.) (Table 1). These data sets were integrated with SAINT version 7.06A software and internally scaled with SADABS version 2005/1. The X-ray data for the *Se*CobT^{E317A} mutant protein in complex with α-DMB riboside monophosphate and nicotinate were collected at 100 K on the Structural Biology Center beamline 19BM at the Advanced Photon Source in Argonne, IL. Although the crystals diffracted beyond the 1.4 Å resolution at that X-ray source, a cutoff at that resolution was chosen to optimize the quality of both the high and low resolution data with two data collection scans. These diffraction data were integrated and scaled with HKL3000 [34]. The variant structures were determined by molecular replacement using PHASER [35] in which the substrate-free *Se*CobT (PDB id: 1L4B) was used as a search model. Model refinement was performed by alternate cycles of manual building with Coot and restrained refinement with Refmac5 [36]. Data collection and refinement statistics are presented in Table 1.

3. Results and discussion

3.1. Glutamate 174 and 317 are essential for optimal activity

It was proposed in the initial characterization of *Se*CobT that the γ-carboxyl group of Glu317 in the active site served as a general base to catalyze the phosphoribosyl-transfer [16,17]. To test this idea we replaced of Glu317 with alanine, a substitution that decreased the activity of the enzyme for DMB and adenine by 20–60 fold in vitro (Table 2). The same substitution was made in the adjacent glutamate residue (174) in the active site, and similar effects on activity were measured. Remarkably, the double substitution variant (*Se*CobT^{E317A,E174A}) retained activity in the nmol/min/mg range (a decrease of 15,000 fold from *Se*CobT^{WT}) showing that under the assay conditions used, a protein-derived catalytic base was not required for activity.

AdoCbl synthesis was restored in a *S. enterica cobT cobB* strain harboring a plasmid encoding the *Se*CobT^{E174A} variant when DMB (10 μM) was provided in the medium. Under such conditions the mutant strain grew at a rate similar to that of the wild-type strain. In contrast, restoration of AdoCbl synthesis in the *S. enterica cobT cobB* strain harboring a plasmid encoding the *Se*CobT^{E317A} variant required 200-fold less exogenous DMB (50 nM) to reach wild-type growth rates (Table 3). Both *Se*CobT^{E174A} and *Se*CobT^{E317A} variants had similar activities when adenine was the substrate, but only the *Se*CobT^{E317A} variant supported pseudo-B_{12} synthesis in vivo when supplemented with 1 mM adenine. In contrast, the *Se*CobT^{E174A} variant expressed from a plasmid failed to support pseudo-B_{12} synthesis even when the concentration of adenine in the medium was 1 mM (Table 3). Growth was not supported by the double variant *Se*CobT^{E174A,E317A} under any condition tested. The limited impact of changing Glu174 or Glu317 raises two questions: i) is the role of any of these residues catalytic? and ii) do these residues play an equivalent role in the heterodimer *Se*CobT homolog *So*ArsAB, which is able to catalyze the synthesis of phenolic riboside monophosphates?

3.2. Equivalent glutamate residues in SoArsA are critical for activity of the heterodimeric SoArsAB

Variant *So*ArsAB proteins were generated with alanine substitutions in the equivalent glutamate residues to investigate whether or not these residues were needed for activity of this *Se*CobT homolog. Wild-type *So*ArsAB catalyzes the synthesis of α-phenolyl, α-*p*-cresolyl and α-benzimidazolyl riboside monophosphates [20]. Both *So*ArsA and *So*ArsB have glutamate residues equivalent to Glu174 and Glu317 in *Se*CobT [21]. *So*ArsAB variants with single Glu-to-Ala changes in the *So*ArsA subunit were active in vivo when DMB was provided in the medium, a result that was similar to the one obtained with *Se*CobT variants (Table 4). Notably, *So*ArsAB lost all of its in vivo activity when both glutamate residues in *So*ArsA were changed to alanine suggesting that the active site within the *So*ArsB subunit does not contribute to the activity despite possessing equivalent glutamate residues. Substitutions in *So*ArsB did not affect activity in vivo (Table 4), a result that was expected given the absence of any base substrate binding in the crystal structure of *So*ArsAB [21]. The role of the *So*ArsB subunit remains unclear.

3.3. Conservation of the acidic residues in SeCobT orthologues

Alignment of *Se*CobT orthologs (PFAM 02277; EC 2.4.2.21) encoded in representative genomes of bacteria and archaea revealed that acidic residues are conserved in the active site of this family of enzymes (Fig. 2A). Structural alignments of *Se*CobT and two archaeal CobT orthologs whose structures have been determined, showed a valine residue replacing the equivalent Glu317 in *Se*CobT (Fig. 2B). Together with the mutational analyses of *Se*CobT and *So*ArsAB, this alignment suggested that an acidic residue at a position equivalent to Glu174 in *Se*CobT alone might be sufficient for biological activity in archaeal CobT orthologs.

3.4. Crystal structures of SeCobT^{E174A} and SeCobT^{E317A} variants reveal the complexity of the role of Glu174 and Glu317

To better understand the molecular effects of substitutions at Glu174 and 317 in *Se*CobT, both variants were subjected to extensive crystallization studies. Structures of substrate-free *Se*CobT^{E317A}, *Se*CobT^{E317A} in complex with DMB riboside monophosphate (α-RP) and nicotinate, *Se*CobT^{E174A} in complex with adenine, and *Se*CobT^{E174A} in complex with DMB were determined to a 2.0 Å resolution or better. Interestingly,

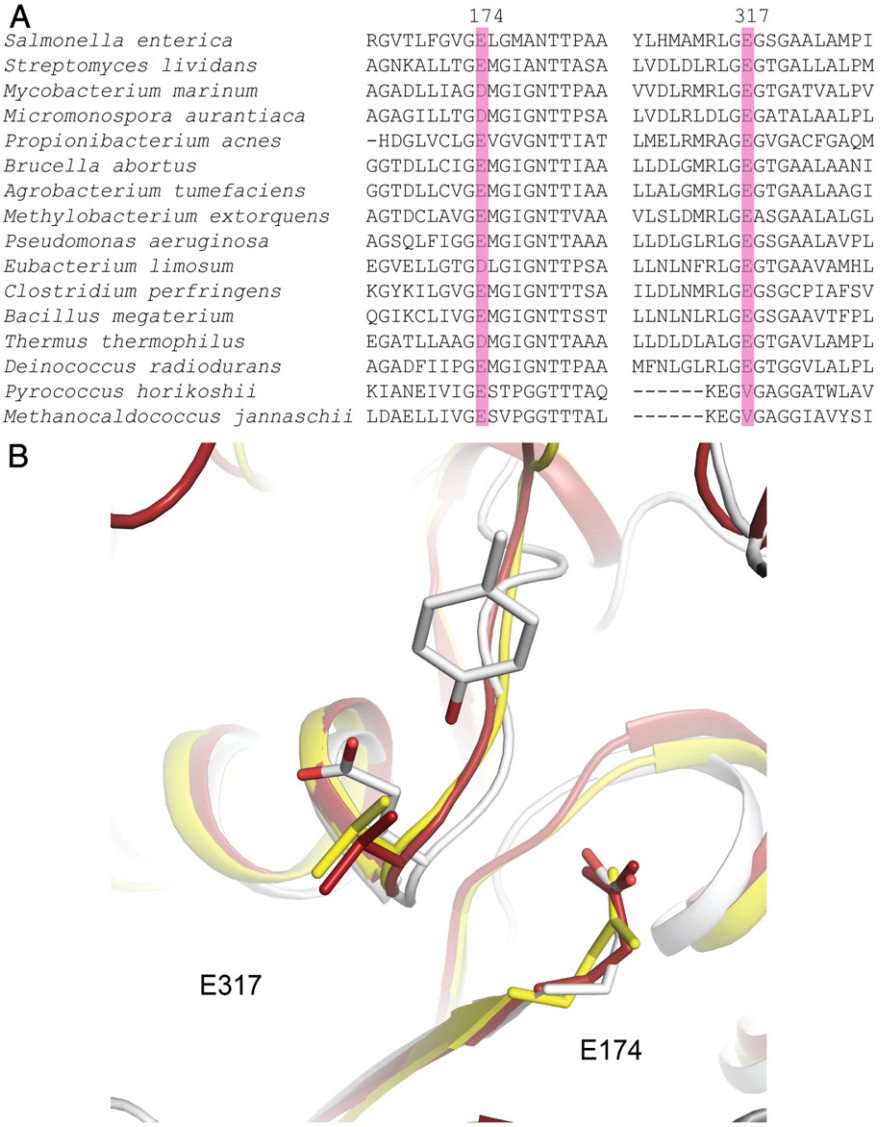

Fig. 2. A. Sequence alignments (ClustalW) of the Glu174 and Glu317 regions in CobT orthologs. Alignments of these regions in archaeal sequences were aided by structural comparisons. In archaea, the *C*-terminal region containing the acidic residue is not conserved. B. Structural alignments of *P. horikoshii* (Yellow, 3U4G) and *M. jannaschii* (Red, 3L0Z) CobT with *S. enterica* CobT complexed with *p*-cresol (White, 1JHU). The structures of the archaeal proteins were determined in the absence of any ligand and it is unknown whether they exhibit the same enzymatic function as *Se*CobT.

crystals of the *Se*CobT^{E317A} variant grown in the presence of DMB or adenine did not exhibit these substrates in their active sites. As with *Se*CobT^{WT}, the variants crystallized with a monomer in the asymmetric unit, while the biologically active dimer is generated by a crystallographic two-fold axis. This results in two identical active sites per dimer, where each is formed by components of both protomers. Each subunit contains a large and small domain. The large domain exhibits a six-stranded parallel beta sheet with connecting helices characteristic of the Rossmann fold, while the small domain is built from helices contributed by both the *N*- and *C*-termini of the polypeptide chain to form a three-helix bundle (Fig. 3A).

In the *Se*CobT^{E174A} variant, DMB and adenine adopt the same orientation as observed in *Se*CobT^{WT} (data not shown for *Se*CobT^{E174A} in complex with adenine) [14]. Interestingly, the change of Glu174 to alanine introduced conformational flexibility at Gly316–Glu317 (Fig. 3B). Two conformations were observed in the crystal structure. In one conformation the residues were in the same position as in *Se*CobT^{WT}, and in the other, the carboxylate moiety on the side chain of Glu317 was shifted away from its native orientation thereby breaking a close hydrogen bond with the substrate DMB. This shift was facilitated by the rotation of the carbonyl oxygen of Gly316 nearly 180° toward the position that was formerly occupied by the side chain of Glu174 in the *Se*CobT^{WT} structure. Although Glu174 was further away from the base substrate at a distance improbable for direct proton abstraction, structural analysis of this variant showed that this residue plays a role in stabilizing Glu317 in the active conformation.

As noted above, the *Se*CobT^{E317A} variant crystallized devoid of substrates, despite the addition of either DMB or adenine during crystallization, and was thus considered a substrate-free form of the enzyme (Fig. 4). However, the mutation induced substantial structural changes in the active site pocket relative to that of substrate-free *Se*CobT^{WT}. Substitution of alanine for Glu317 caused a backbone shift in the loop that underlies the active site (Met313 to Ala317). This shift likely resulted from the loss of a single hydrogen bond between Oε2 of Glu317 and the side chain of Gln88 (Fig. 4). The disruption of this hydrogen bond caused a collapse of the active site in which several residues (Gln88, Leu315, and Arg314) shifted inwards to fill the substrate binding pocket, thereby precluding substrate binding. The replacement of Glu317 to alanine released the disfavored torsional angles in the 317 position likely triggering the collapse. There was also a substantial change in

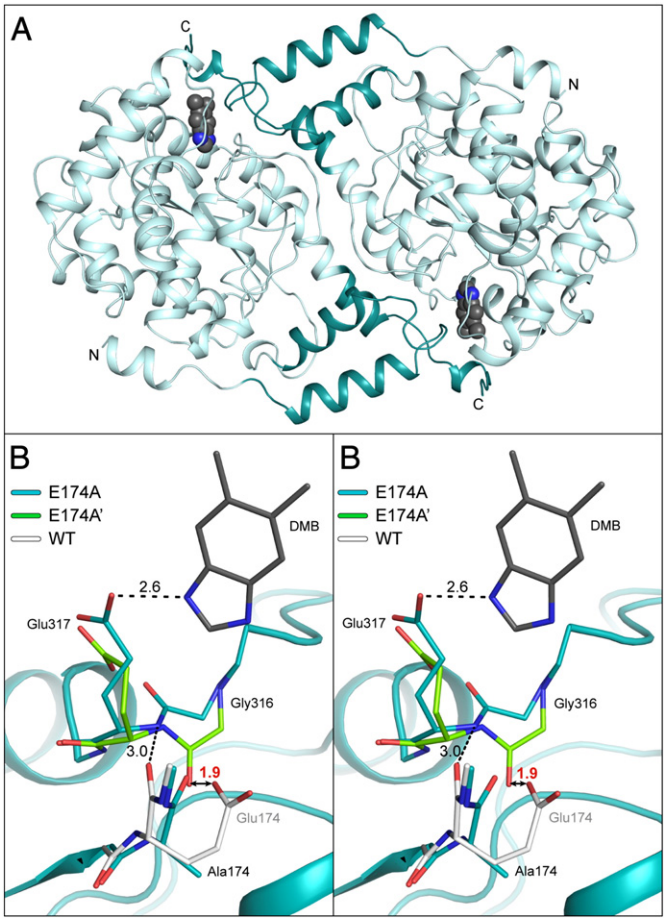

Fig. 3. A. Cartoon representation of *Se*CobT^{E174A} dimer. The large domains are shown in pale cyan, the small domains are shown in teal. DMB is shown in space fill as dark gray. B. Stereo view of the active site for *Se*CobT^{E174A} (teal). This view reveals that the backbone for residues 316 and 317 in *Se*CobT^{E174A} adopts multiple conformations (teal and green). One of these (teal) is essentially identical to the wild type protein where in the wild-type protein this is maintained by a hydrogen bond between Glu174 and the backbone amide hydrogen on Glu317 (white; from coordinates 1D0S). The alternate conformation (green) is formed by a rotation of the peptide bond at Gly316, which breaks the interaction between Oε2 of Glu317 with the DMB substrate. The interaction between Gly316 and Glu174 (black arrow) in the wild-type protein prevents the backbone from adopting the alternate conformation.

the water structure. Additionally, residues Gly202–Leu205, which comprised much of the NaMN binding region, were disordered in the *Se*CobT^{E317A} substrate-free form (Fig. 4). This was likely due to the movement of Arg314. In addition, residue Glu174 rotated away from the substrate-binding site. Finally, the component of the active site contributed by the adjacent subunit (Leu30–Pro32) adopted a new conformation in which the loop encroached further into the adjacent active site (Fig. 4), again precluding substrate binding. These changes in the active site architecture explained the decrease in activity for both adenine and DMB, and suggested that Glu317 played an important role in maintaining the integrity of the active site. The structure of the substrate-free form of *Se*CobT^{E317A} most likely represented an inactive conformation trapped by the crystal lattice, since the variant did exhibit biological activity in vitro and in vivo (Tables 2, 3). However, when crystals of *Se*CobT^{E317A} grown in the presence of DMB were soaked with NaMN the reaction occurred in the crystal resulting in a complex with α-RP and nicotinate. In this case, the overall orientation of the residues in the active site was restored to that of *Se*CobT^{WT} suggesting that both substrates contributed to the active conformation of this variant (Fig. 4). Interestingly, the occupancy for both reaction products was low despite the overall high quality of the crystallographic data.

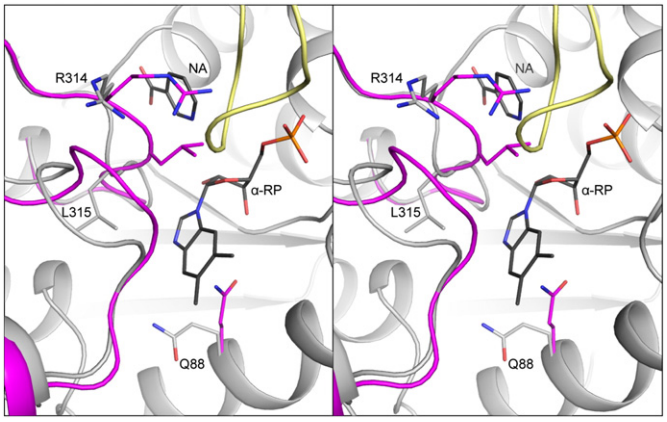

Fig. 4. Stereo view of the active site of the substrate-free form of CobT^{E317A} (magenta) aligned with CobT^{E317A} in complex with reaction products α-DMB-riboside monophosphate (α-RP) and nicotinate (NA) (light gray). Several conformational rearrangements take place in the absence of substrate (magenta): the peptide backbone near Leu315 is rearranged, the side chains of Glu88 and Arg314 shift, and the loop that folds over the second substrate binding site in product complex (yellow; Val198–Pro207) is disordered. The active site residues are restored to active conformation (light gray) in the presence of reaction products.

3.5. Are the roles of Glu174 and Glu317 strictly structural?

Together with the reduced in vitro activities of the *Se*CobT^{E174A} and *Se*CobT^{E317A}, these results appear to support the idea that glutamates 174 and 317 are important for the maintenance of the active site architecture, but are not critical for enzymatic activity. This idea is consistent with the expectation that removal of a catalytic base from *Se*CobT would result in an ~10^5 loss in activity based on the precedence set for active site residues derived from study of the serine proteases [37]. Notably, we measured only a modest decrease (~20–60 fold) in the activity of the *Se*CobT^{E174A} and *Se*CobT^{E317A} variants, and the *Se*CobT$^{E174A\ E317A}$ variant retained activity at the nmol/min/mg level. However, the conclusion that Glu174 and Glu317 are not catalytic residues deserves further discussion for two reasons. Firstly, it should be noted that the in vitro activity assay for *Se*CobT is performed at the non-physiological pH of 10, because at neutral pH, the activity of *Se*CobT is undetectable [13]. Thus it is plausible that at high pH a catalytic base is not required to remove a proton from the substrate, but might be more important under physiological conditions. Also, if the activity of the *Se*CobT^{E174A,E317A} variant at pH 10 is four orders of magnitude lower than the activity of *Se*CobT^{WT} (Table 2), its activity at physiological pH is likely to be $\lesssim 10^{-5}$ relative to *Se*CobT^{WT}. Secondly, such a precipitous loss of activity is consistent with the fact that, in vivo, not even the elevated level of the *Se*CobT^{E174A,E317A} variant encoded on a plasmid (at 15-fold greater level than a chromosomally encoded gene) can restore AdoCbl synthesis in a *S. enterica cobT* strain. The caveats imposed by the pH conditions under which the activity of the enzyme is measured (pH 10), make it difficult to ascertain whether or not Glu174 or Glu317 acts as the base that triggers the attack on NaMN, as proposed earlier [16]. Studies of a *Se*CobT ortholog that retains full activity at pH 7 would allow us to determine the precise role of residues equivalent to Glu174 and Glu317 in *Se*CobT.

3.6. Changes in the primary sequence of SeCobT allow the resulting variant to phosphoribosylate phenolic substrates

A striking difference between *Se*CobT and *So*ArsAB is the ability of the latter to form an α-*O*-glycosidic bond between the hydroxyl of the phenolic moiety and the C1 of ribose. Structural determination of *So*ArsAB in complex with *p*-cresol revealed that the overall structures of these two enzymes were remarkably similar [21]. Examination of the structures, coupled with a ClustalW sequence alignment of active site residues of *So*ArsA, *So*ArsB and *Se*CobT suggested how *So*ArsAB might catalyze the reaction with phenolic compounds [21]. This led to

a list of potential residues that could be targeted for site-directed mutagenesis to test the hypothesis that the specificity of *Se*CobT might be expanded to include phenolic compounds (Fig. 5).

Subsequent mutagenesis of the *Se*CobT active site identified a variant with three amino acid substitutions S80Y, Q88M and L175M that did phosphoribosylate *p*-cresol. This was the minimum set of changes in *Se*CobT needed to convert *p*-cresol into a product. These substitutions mimicked those of the *So*ArsA active site. HPLC was used to resolve the product formed with *p*-cresol and NaMN using *So*ArsAB and *Se*CobT^{S80Y,Q88M,L175M} compared to *Se*CobT (Fig. 6A). The product was confirmed by mass spectrometry to match the predicted mass of α-*p*-cresolyl-riboside monophosphate (Fig. 6B).

In the lattice of *Se*CobT crystals, NaMN was hydrolyzed to nicotinate when *p*-cresol was also present in the active site, but product formation, i.e., α-*p*-cresolyl-riboside monophosphate was not observed. In contrast, product formation did occur in the crystal lattice with other benzimidazole and purine bases [14]. Structural analyses of *So*ArsAB, a heterodimeric ortholog of *Se*CobT from *S. ovata* that phosphoribosylates phenolics, showed that the hydroxyl group in *p*-cresol was rotated away from Glu319 (Glu317 equivalent in *Se*CobT) [21], and thereby was brought into close proximity to the C1 of ribose when NaMN was modeled into the active site.

3.7. Structural analyses of the SeCobT^{S80Y,Q88M,L175M} variant provide support for the role of water in the formation of phenolic riboside monophosphates

To further characterize the changes in *Se*CobT that led to the phenolic activity, crystal structures of the *Se*CobT^{S80Y,Q88M,L175M} variant were determined with *p*-cresol and *p*-cresol and NaMN. The overall structure of the *Se*CobT^{S80Y,Q88M,L175M} variant in complex with *p*-cresol was the same as *Se*CobT (Fig. 7A). The orientation of *p*-cresol in the variant did not change in comparison to *Se*CobT. However, when crystals grown in the presence of *p*-cresol were briefly soaked in a solution containing NaMN, both substrates were well defined in the *Se*CobT^{S80Y,Q88M,L175M} active site, but the reaction did not occur (Fig. 7). These results differed from those obtained with *Se*CobT^{WT} crystals in which NaMN was hydrolyzed when incubated with *p*-cresol. Importantly, *p*-cresol bound to the variant in crystals soaked with NaMN for long time periods adopted a rotated conformation similar to that in *So*ArsAB in one of the active sites in the dimer (Fig. 5A). An ordered water molecule between the hydroxyl

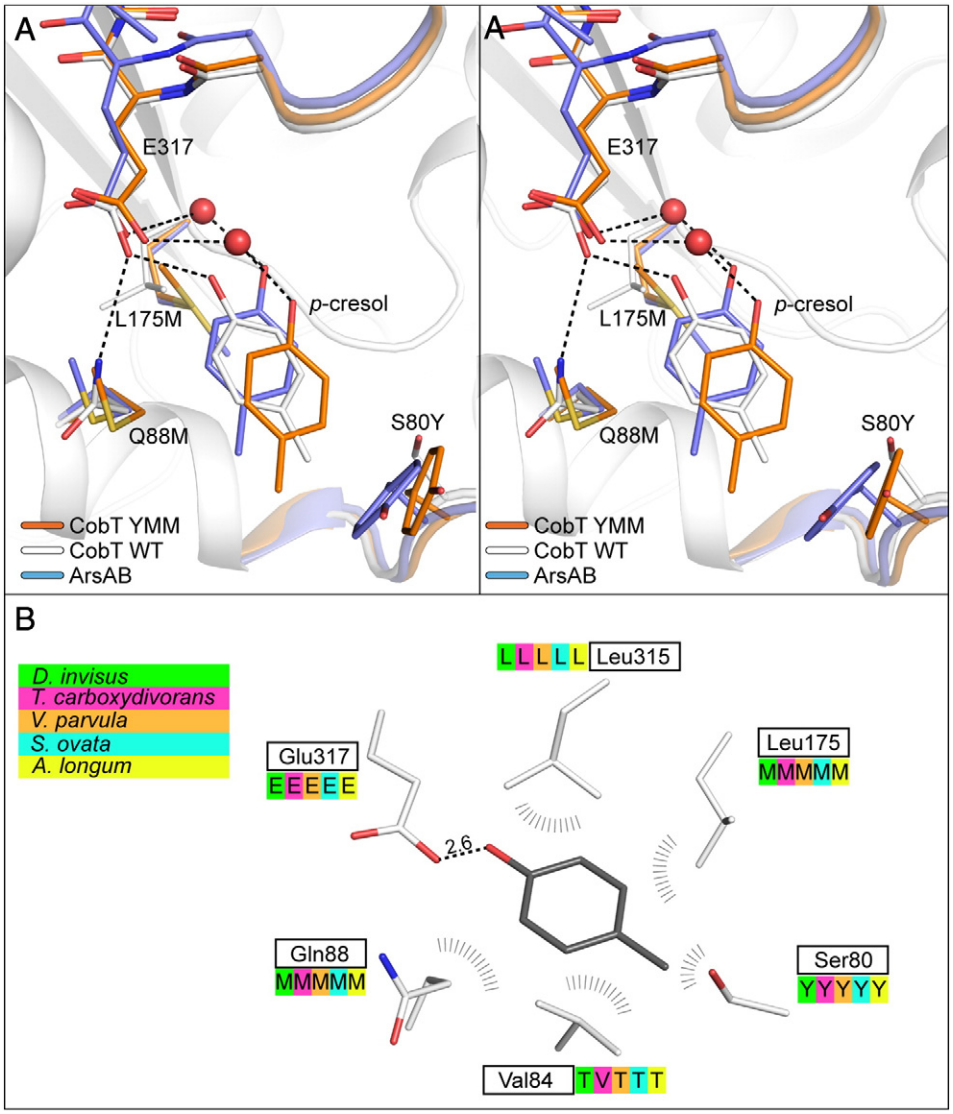

Fig. 5. A. Structure alignments of *Se*CobT^{S80Y,Q88M,L175M} (CobT YMM, orange) and *So*ArsAB (blue) with *Se*CobT (white). Both *Se*CobT^{S80Y,Q88M,L175M} and *Se*CobT^{WT} form a close hydrogen bond with *p*-cresol, while in *So*ArsAB this bond is mediated by a water molecule. B. Sequence alignments (ClustalW) of several ArsA orthologues in *Dialister invisus*, *Thermsinus carboxydivorans*, *Veillonella parvula*, *Sporomusa ovata*, and *Acetonema longum* superimposed on a contact diagram for *p*-cresol. This alignment shows the conservation in residues surrounding the substrate base.

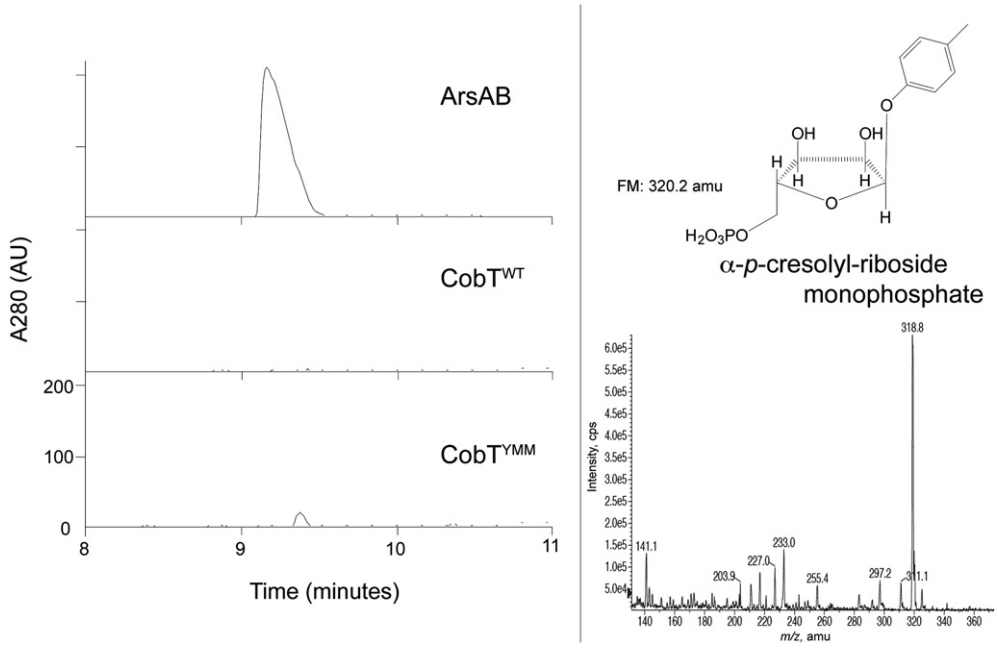

Fig. 6. A. HPLC separation of product formation in 6 h from enzymes ArsAB, CobT^{WT}, and CobT^{YMM} (CobT^{S80Y,Q88M,L175M}) using *p*-cresol and NaMN. B. Mass analysis (negative ion mode) of the material under the peak eluted from the *Se*CobT^{S80Y,Q88M,L175M} reaction (319 amu) was consistent with the predicted mass of α-*p*-cresolyl-riboside monophosphate (320 amu).

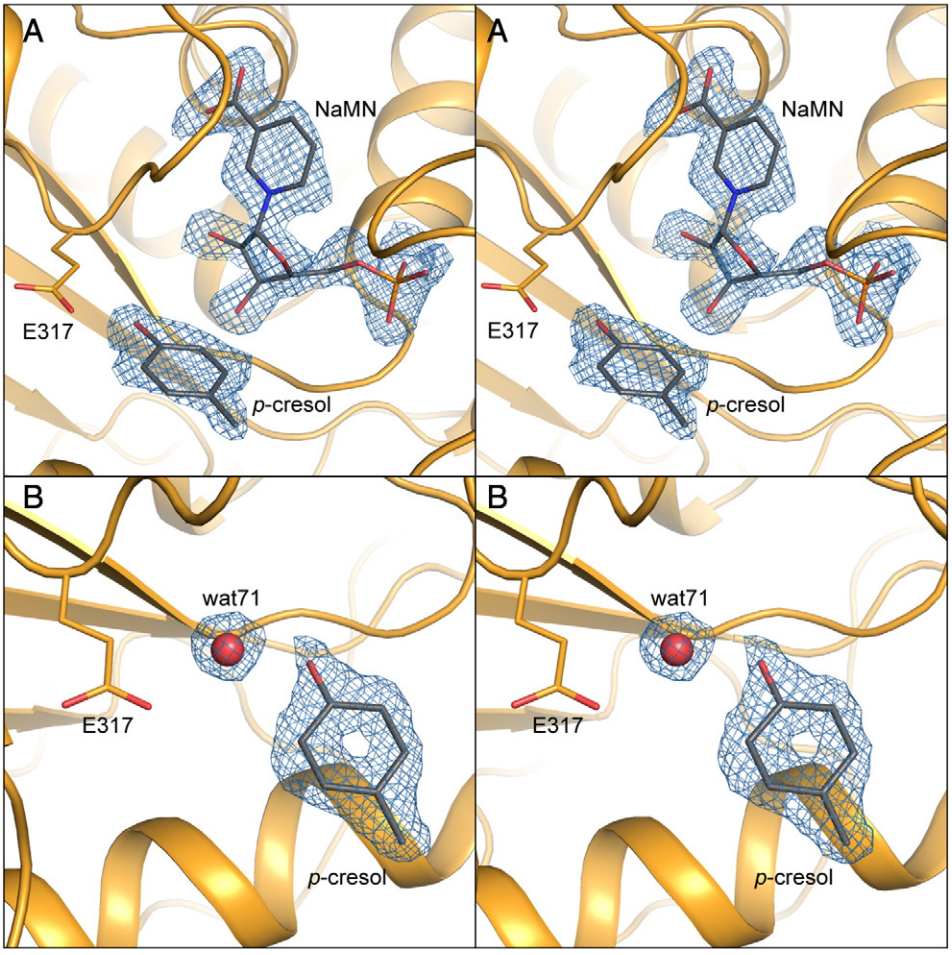

Fig. 7. Stereo view of the electron density for both NaMN and *p*-cresol in *Se*CobT^{S80Y,Q88M,L175M} (A) and *p*-cresol and *Se*CobT^{S80Y,Q88M,L175M} (B). Electron density (1.8 σ) was calculated from coefficients of the form $F_o - F_c$ where *p*-cresol and NaMN were omitted from phase calculation and refinement.

group of *p*-cresol and Glu317 was observed in the *Se*CobT^{S80Y,Q88M,L175M} variant. Interestingly, neither NaMN nor any reaction products were observed in this structure. Notably, an ordered water molecule was proposed to facilitate the proton abstraction that triggered the α-*O*-glycosidic bond formation in *So*ArsAB [21]. These results support the proposed mechanism of phenolic phosphoribosylation through an active site water molecule [21].

Although these results indicate that *Se*CobT^{WT} can be changed into an enzyme capable of phenolic phosphoribosylation, the CobT^{S80Y,Q88M,L175M} variant synthesized far less α-*p*-cresolyl riboside monophosphate than the *So*ArsAB enzyme toward phenolic substrates. This is not surprising given the numerous small differences that likely contribute to the optimization of *So*ArsAB activity. The observed differences could be due to the constraints placed on Glu317 by the hydroxyl side chain of Ser319, which forms a close hydrogen bond (2.6 Å) with the carbonyl oxygen of Gly316 in *Se*CobT^{S80Y,Q88M,L175M}. Ser319 is replaced with an isoleucine in *So*ArsA, which was shown to be important for activity in *So*ArsAB [21]. However, alignments do not show strong conservation at this position (Fig. 2), indicating that individual substitutions need to be considered in the context of the surrounding structure.

4. Conclusion

The lack of substrate specificity in the CobT family of enzymes plays a critical part in the diversity of the nucleoside base found in cobamides. The structural studies of *Se*CobT variants and comparisons of *Se*CobT orthologues presented here reveal that the conserved glutamate residues play a role in orienting the substrates and stabilizing the active site. It is less clear, however, whether the alluded glutamate residues may also function as catalytic bases under physiological conditions. The structural studies reported here suggest some of the changes that are necessary to allow *Se*CobT to catalyze the phosphoribosylation of phenolic compounds. Together with in vivo and in vitro analyses of *Se*CobT variants, our results of the crystallographic analyses advance our understanding of the mechanism of the NaMN:base phosphoribosyltransferase family of enzymes.

The structural and functional studies in *Se*CobT enzymes described here and elsewhere provide insights into why these enzymes facilitate the diversity in cobamides. Cobamide dependent enzymes likely dictate the chemical nature of the base. Some cobamide dependent enzymes displace the cobamide base, but provide a histidinyl residue as an alternative ligand. Other enzymes rely on the existing base of the cobamide [38]. The coordination of the cobalt atom by the α-ligand to modify its reduction potential and charge is an important aspect in AdoCba biosynthesis and reactions catalyzed through this cofactor [39–41].

In nature, organisms have evolved at least one way to remodel the α-ligand to fit the need of their set of cobamide dependent enzymes through an amidohydrolase known as CbiZ [42–45]. A complete analysis of the source of the base(s), AdoCba biosynthetic and remodeling machineries, and cobamide dependent enzymes is required to better understand the forces driving cobamide diversity in nature.

5. Accession numbers

X-ray coordinates for *Se*CobT^{E317A} in the substrate free form and in complex with α-DMB riboside monophosphate and nicotinate, *Se*CobT^{E174A} in complex with DMB and adenine, *Se*CobT^{S80Y,Q88M,L175M} in complex with *p*-cresol and in complex with *p*-cresol and nicotinate mononucleotide have been deposited in the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, N. J. with Protein Data Bank entries 4KQH, 4KQI, 4KQG, 4KQF, 4KQK, and 4KQJ respectively.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgments

This work was supported in part by the NIH grants GM083987 and GM086351 to I.R. and R37 GM40313 to J.C.E.-S. Use of the SBC 19BM beamline at the Argonne National Laboratory Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract No. W-31-109-ENG-38.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2013.09.038.

References

[1] E.N. Marsh, Coenzyme B12 (cobalamin)-dependent enzymes, Essays Biochem. 34 (1999) 139–154.
[2] B. Kräutler, W. Fieber, S. Osterman, M. Fasching, K.-H. Ongania, K. Gruber, C. Kratky, C. Mikl, A. Siebert, G. Diekert, The cofactor of tetrachloroethene reductive dehalogenase of *Dehalospirillum multivorans* is Norpseudo-B12, a new type of natural corrinoid, Helv. Chim. Acta 86 (2003) 3698–3716.
[3] A.R. Battersby, Tetrapyrroles: the pigments of life, Nat. Prod. Rep. 17 (2000) 507–526.
[4] R. Banerjee, Radical carbon skeleton rearrangements: catalysis by coenzyme B12-dependent mutases, Chem. Rev. 103 (2003) 2083–2094.
[5] V. Bandarian, G.H. Reed, Ethanolamine ammonia-lyase, in: R. Banerjee (Ed.), Chemistry and Biochemistry of B12, John Wiley & Sons, Inc., New York, 1999, pp. 811–833.
[6] T. Toraya, Enzymatic radical catalysis: coenzyme B12-dependent diol dehydratase, Chem. Rec. 2 (2002) 352–366.
[7] R.G. Matthews, M. Koutmos, S. Datta, Cobalamin-dependent and cobamide-dependent methyltransferases, Curr. Opin. Struct. Biol. 18 (2008) 658–666.
[8] P. Renz, Biosynthesis of the 5,6-dimethylbenzimidazole moiety of cobalamin and of other bases found in natural corrinoids, in: R. Banerjee (Ed.), Chemistry and Biochemistry of B12, John Wiley & Sons, Inc., New York, 1999, pp. 557–575.
[9] M.J. Warren, E. Raux, H.L. Schubert, J.C. Escalante-Semerena, The biosynthesis of adenosylcobalamin (vitamin B12), Nat. Prod. Rep. 19 (2002) 390–412.
[10] J.C. Escalante-Semerena, M.J. Warren, Biosynthesis and use of cobalamin (B_{12}), in: A. Böck, R. Curtiss III, J.B. Kaper, P.D. Karp, F.C. Neidhardt, T. Nyström, J.M. Slauch, C.L. Squires (Eds.), *EcoSal — Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, ASM Press, Washington, D. C., 2008
[11] M.G. Johnson, J.C. Escalante-Semerena, Identification of 5,6-dimethylbenzimidazole as the *Coα* ligand of the cobamide synthesized by *Salmonella typhimurium*. Nutritional characterization of mutants defective in biosynthesis of the imidazole ring, J. Biol. Chem. 267 (1992) 13302–13305.
[12] B. Keck, P. Renz, *Salmonella typhimurium* forms adenylcobamide and 2-methyladenylcobamide, but no detectable cobalamin during strictly anaerobic growth, Arch. Microbiol. 173 (2000) 76–77.
[13] J.R. Trzebiatowski, J.C. Escalante-Semerena, Purification and characterization of CobT, the nicotinate-mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase enzyme from *Salmonella typhimurium* LT2, J. Biol. Chem. 272 (1997) 17662–17667.
[14] C.G. Cheong, J.C. Escalante-Semerena, I. Rayment, Structural investigation of the biosynthesis of alternative lower ligands for cobamides by nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase from *Salmonella enterica*, J. Biol. Chem. 276 (2001) 37612–37620.
[15] A.W. Tsang, J.C. Escalante-Semerena, CobB, a new member of the SIR2 family of eucaryotic regulatory proteins, is required to compensate for the lack of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase activity in *cobT* mutants during cobalamin biosynthesis in *Salmonella typhimurium* LT2, J. Biol. Chem. 273 (1998) 31788–31794.
[16] C.G. Cheong, J.C. Escalante-Semerena, I. Rayment, Capture of a labile substrate by expulsion of water molecules from the active site of nicotinate mononucleotide: 5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella enterica*, J. Biol. Chem. 277 (2002) 41120–41127.
[17] C.G. Cheong, J.C. Escalante-Semerena, I. Rayment, The three-dimensional structures of nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) from *Salmonella typhimurium* complexed with 5,6-dimethybenzimidazole and its reaction products determined to 1.9 Å resolution, Biochemistry 38 (1999) 16125–16135.
[18] B. Möller, R. Ossmer, B.H. Howard, G. Gottschalk, H. Hippe, *Sporomusa*, a new genus of Gram-negative anaerobic-bacteria including *Sporomusa-Sphaeroides* spec-nov and *Sporomusa-Ovata* spec-nov, Arch. Microbiol. 139 (1984) 388–396.
[19] E. Stupperich, H.J. Eisinger, Biosynthesis of *para*-cresolyl cobamide in *Sporomusa ovata*, Arch. Microbiol. 151 (1989) 372–377.
[20] C.H. Chan, J.C. Escalante-Semerena, ArsAB, a novel enzyme from *Sporomusa ovata* activates phenolic bases for adenosylcobamide biosynthesis, Mol. Microbiol. 81 (2011) 952–967.
[21] S.A. Newmister, C.H. Chan, J.C. Escalante-Semerena, I. Rayment, Structural insights into the function of the nicotinate mononucleotide:phenol/p-cresol phosphoribosyltransferase (ArsAB) enzyme from *Sporomusa ovata*, Biochemistry 51 (2012) 8571–8582.
[22] D. Berkowitz, J.M. Hushon, H.J. Whitfield Jr., J. Roth, B.N. Ames, Procedure for identifying nonsense mutations, J. Bacteriol. 96 (1968) 215–220.

[23] W.E. Balch, R.S. Wolfe, New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere, Appl. Environ. Microbiol. 32 (1976) 781–791.
[24] R. Atlas, Handbook of Media for Environmental Microbiology, CRC Press, Boca Raton, 1995, p. 6.
[25] G. Bertani, Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*, J. Bacteriol. 62 (1951) 293–300.
[26] G. Bertani, Lysogeny at mid-twentieth century: P1, P2, and other experimental systems, J. Bacteriol. 186 (2004) 595–600.
[27] B. Cormack, Directed Mutagenesis Using the Polymerase Chain Reaction, Greene Publishing Associates & Wiley Interscience, New York, 1997.
[28] A.H. Rosenberg, B.N. Lade, D.S. Chui, S.W. Lin, J.J. Dunn, F.W. Studier, Vectors for selective expression of cloned DNAs by T7 RNA polymerase, Gene 56 (1987) 125–135.
[29] H.E. Klock, E.J. Koesema, M.W. Knuth, S.A. Lesley, Combining the polymerase incomplete primer extension method for cloning and mutagenesis with micro-screening to accelerate structural genomics efforts, Proteins Struct. Funct. Bioinf. 71 (2008) 982–994.
[30] A.J. Andreoli, M. Ikeda, Y. Nishizuka, O. Hayaishi, Quinolinic acid: a precursor to nicotinamide adenine dinucleotide in *Escherichia coli*, Biochem. Biophys. Res. Commun. 12 (1963) 92–97.
[31] M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, Complete set of ORF clones of *Escherichia coli* ASKA library (A complete set of *E. coli* K-12 ORF archive): unique resources for biological research, DNA Res. 12 (2005) 291–299.
[32] F.M. Dickinson, P.C. Engel, The preparation of pure salt-free nicotinamide coenzymes, Anal. Biochem. 82 (1977) 523–531.
[33] K.R. Claas, J.R. Parrish, L.A. Maggio-Hall, J.C. Escalante-Semerena, Functional analysis of the nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) enzyme, involved in the late steps of coenzyme B12 biosynthesis in *Salmonella enterica*, J. Bacteriol. 192 (2010) 145–154.
[34] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.
[35] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.
[36] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 3553–3567.
[37] P. Carter, J.A. Wells, Dissecting the catalytic triad of a serine protease, Nature 332 (1988) 564–568.
[38] R. Banerjee, S.W. Ragsdale, The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes, Annu. Rev. Biochem. 72 (2003) 209–247.
[39] T.A. Stich, M. Yamanishi, R. Banerjee, T.C. Brunold, Spectroscopic evidence for the formation of a four-coordinate Co(2+)cobalamin species upon binding to the human ATP:Cobalamin adenosyltransferase, J. Am. Chem. Soc. 127 (2005) 7660–7661.
[40] T.A. Stich, N.R. Buan, J.C. Escalante-Semerena, T.C. Brunold, Spectroscopic and computational studies of the ATP:Corrinoid adenosyltransferase (CobA) from *Salmonella enterica*: insights into the mechanism of adenosylcobalamin biosynthesis, J. Am. Chem. Soc. 127 (2005) 8710–8719.
[41] P.E. Mera, M. St Maurice, I. Rayment, J.C. Escalante-Semerena, Residue Phe112 of the human-type corrinoid adenosyltransferase (PduO) enzyme of *Lactobacillus reuteri* is critical to the formation of the four-coordinate Co(II) corrinoid substrate and to the activity of the enzyme, Biochemistry 48 (2009) 3138–3145.
[42] J.D. Woodson, J.C. Escalante-Semerena, CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B_{12} precursor cobinamide in archaea, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 3591–3596.
[43] J.D. Woodson, J.C. Escalante-Semerena, The *cbiS* gene of the archaeon *Methanopyrus kandleri* AV19 encodes a bifunctional enzyme with adenosylcobinamide amidohydrolase and alpha-ribazole-phosphate phosphatase activities, J. Bacteriol. 188 (2006) 4227–4235.
[44] J.D. Woodson, C.L. Zayas, J.C. Escalante-Semerena, A new pathway for salvaging the coenzyme B_{12} precursor cobinamide in archaea requires cobinamide-phosphate synthase (CbiB) enzyme activity, J. Bacteriol. 185 (2003) 7193–7201.
[45] M.J. Gray, J.C. Escalante-Semerena, The cobinamide amidohydrolase (cobyric acid-forming) CbiZ enzyme: a critical activity of the cobamide remodelling system of *Rhodobacter sphaeroides*, Mol. Microbiol. 74 (2009) 1198–1210.
[46] J.R. Trzebiatowski, G.A. O'Toole, J.C. Escalante-Semerena, The *cobT* gene of *Salmonella typhimurium* encodes the NaMN: 5,6-dimethylbenzimidazole phosphoribosyltransferase responsible for the synthesis of N^1-(5-phospho-alpha-D-ribosyl)-5,6-dimethylbenzimidazole, an intermediate in the synthesis of the nucleotide loop of cobalamin, J. Bacteriol. 176 (1994) 3568–3575.
[47] C.J. Rocco, K.L. Dennison, V.A. Klenchin, I. Rayment, J.C. Escalante-Semerena, Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, Plasmid 59 (2008) 231–237.
[48] L.M. Guzman, D. Belin, M.J. Carson, J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, J. Bacteriol. 177 (1995) 4121–4130.
[49] S. Tabor, Expression using the T7 RNA polymerase/promoter system, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), Current Protocols in Molecular Biology, vol. 2, Wiley Interscience, New York, 1990, (16.12.11.-16.12.11).
[50] G.A. O'Toole, M.R. Rondon, J.C. Escalante-Semerena, Analysis of mutants of defective in the synthesis of the nucleotide loop of cobalamin, J. Bacteriol. 175 (1993) 3317–3326.