SUPPORTING INFORMATION

The *Methanosarcina mazei MM2060* gene encodes a bifunctional kinase/decarboxylase enzyme involved in cobamide biosynthesis

Norbert K. Tavares[1], Carmen L. Zayas[2§], and Jorge C. Escalante-Semerena[1*]

[1]Department of Microbiology, University of Georgia, Athens, GA 30602, USA
[2]Department of Bacteriology, University of Wisconsin, Madison, 53706, USA

Running title: a new bifunctional archaeal L-Thr kinase / L-Thr-P decarboxylase

*To whom correspondence should be addressed: Jorge C. Escalante-Semerena, Department of Microbiology, University of Georgia, 212C Biological Sciences Building, 120 Cedar Street, Athens, GA, 30602, USA. Phone: 706-542-2651, Fax: 706-542-2815 email: jcescala@uga.edu, www.escalab.com

[§]Current address: Aurora Medical Center, Summit WI 53066

Keywords: Bacterial metabolism, kinase, decarboxylase, coenzyme biosynthesis, zinc finger, B_{12}

SUPPORTING FIGURES

Figure S1. Multiple sequence alignment of CobD from methanogenic archaea with CobD from S. enterica - Conserved residues are highlighted in red. Residues with similar properties are boxed in blue. The top row displays the secondary structure of the protein based on the crystal structure of CobD from *S. enterica.*

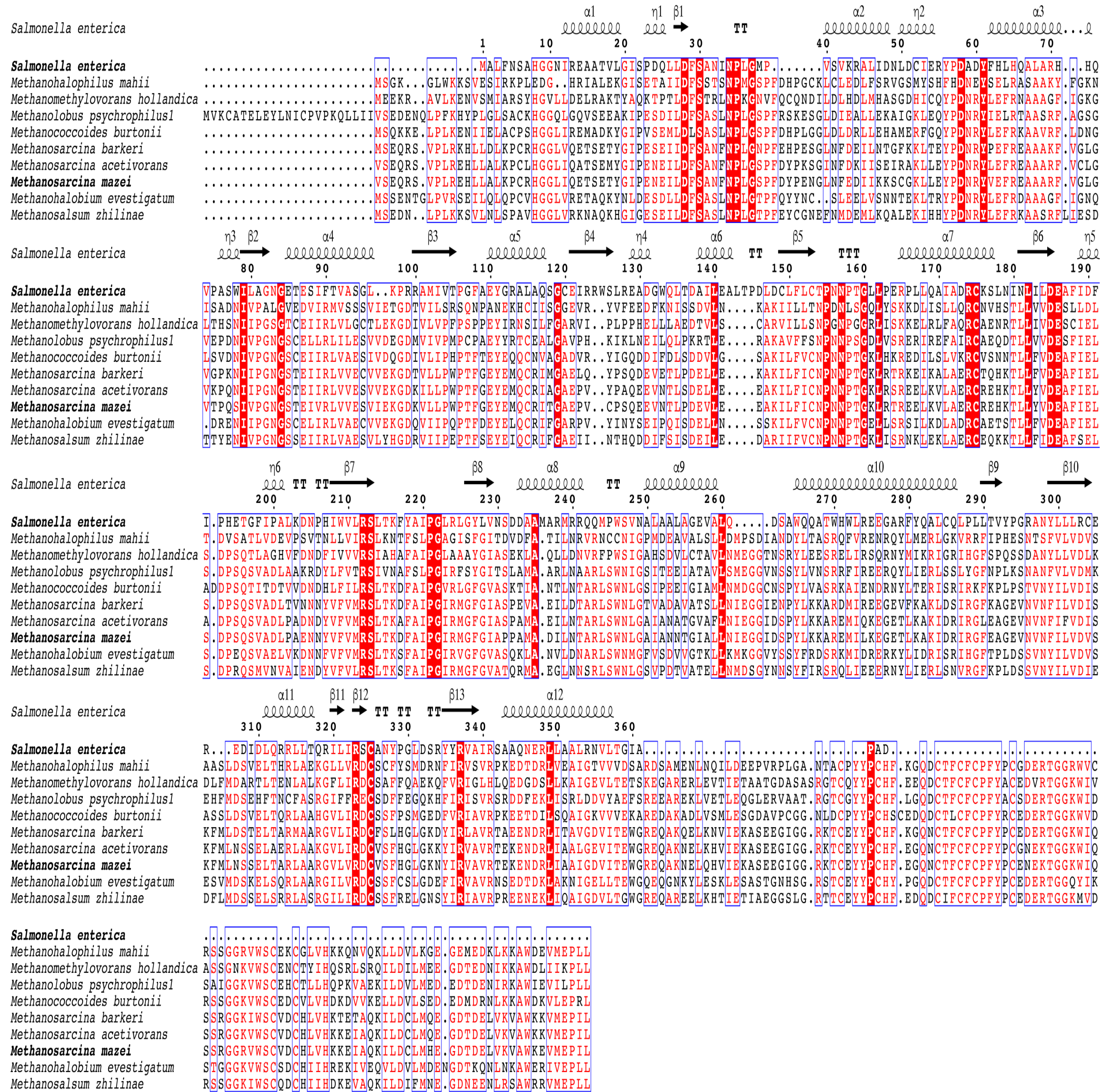
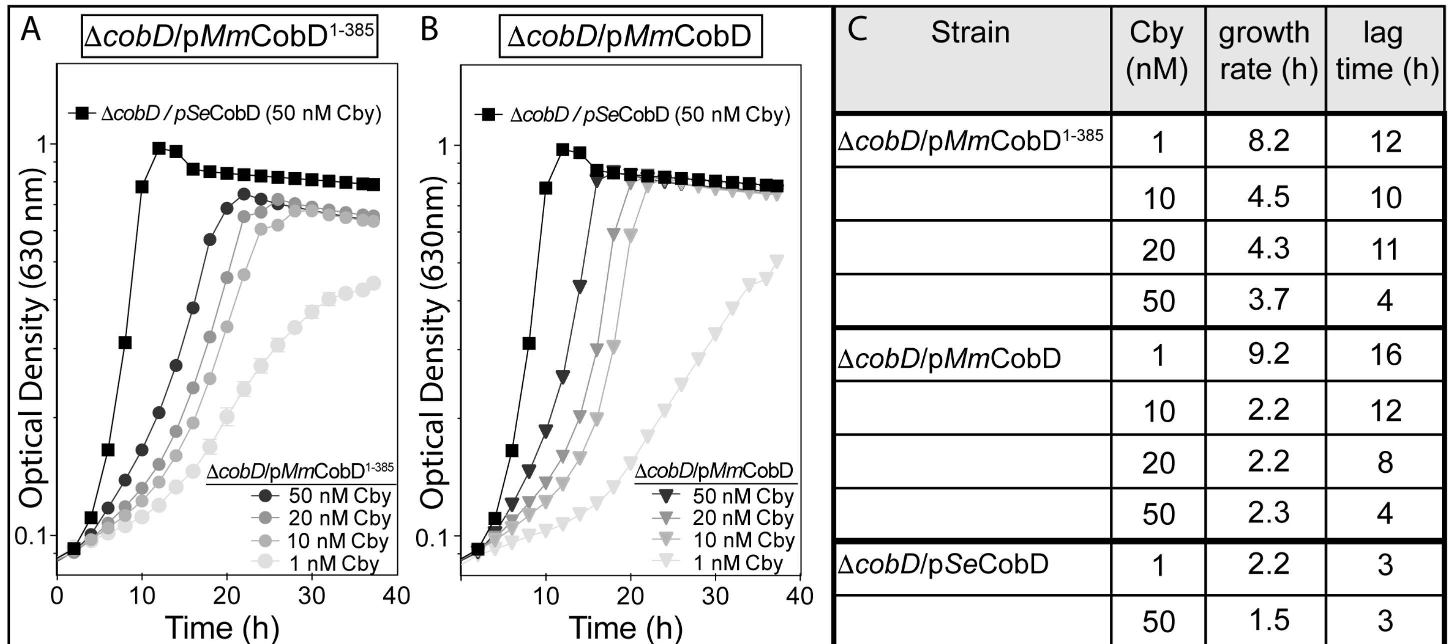


Figure S2. The intracellular concentration of Cby is the limiting factor for the complementation of S. enterica *by* Mm*CobD* - Cobalamin-dependent growth was assessed in minimal medium supplemented with glycerol and Cby. Shown is the effect of increasing levels of Cby (1, 10, 20, and 50 nM) on the complementation of a *S. enterica cobD* strain with plasmids expressing (*A*) *Mm*CobD$^{1\text{-}385}$ (pMmCOBD7) or (*B*) *Mm*CobD (pMmCOBD18), relative to *Se*CobD (pCOBD6, black squares). Growth analysis was performed in duplicate independent experiments in three technical replicates with error bars indicating the standard error of the mean. Error bars are present but smaller than the symbols due to the low deviation between replicates in these representative growth curves. (*C*) Table of growth rates, lag times, and generation times as a function of Cby concentration.



Strain	Cby (nM)	growth rate (h)	lag time (h)
Δ*cobD*/p*Mm*CobD$^{1\text{-}385}$	1	8.2	12
	10	4.5	10
	20	4.3	11
	50	3.7	4
Δ*cobD*/p*Mm*CobD	1	9.2	16
	10	2.2	12
	20	2.2	8
	50	2.3	4
Δ*cobD*/p*Se*CobD	1	2.2	3
	50	1.5	3

Figure S3. Mm*CobD and* Se*CobD cannot decarboxylate L-Ser-P* - Representative ^{13}C-NMR spectra of duplicate independent experiments. Spectra were processed with MestReNova (Mestrelab Research). Reaction mixtures containing sodium phosphate buffer (5 mM, pH 8.5 at 25°C), $MgCl_2$ (5 mM), ATP (40 mM), L-Ser (40 mM), and protein (0.68 µM) were incubated for 1 h at 25ºC prior to the addition of D_2O (17% v/v). Each panel is labeled with the substrate and or protein reaction mixture. Each peak is labeled with the chemical shift value (ppm) and a color-coded letter corresponding to the carbon atom it represents, based on the corresponding chemical shifts for the standards in the no-enzyme control reactions. (*A*) Reaction and chemical structures of substrates with each carbon atom labeled with a color-coded letter. (*B*) *O*-Phospho-L-serine standard (L-Ser-P; red). (*C*) Ethanolamine phosphate (EA-P; boxed purple). (*D*) Equimolar mixture of L-Ser-P and EA-P standards. (*E*) Reaction containing L-Ser-P and *Se*CobD. (*F*) Reaction containing L-Ser-P and *Mm*CobD.

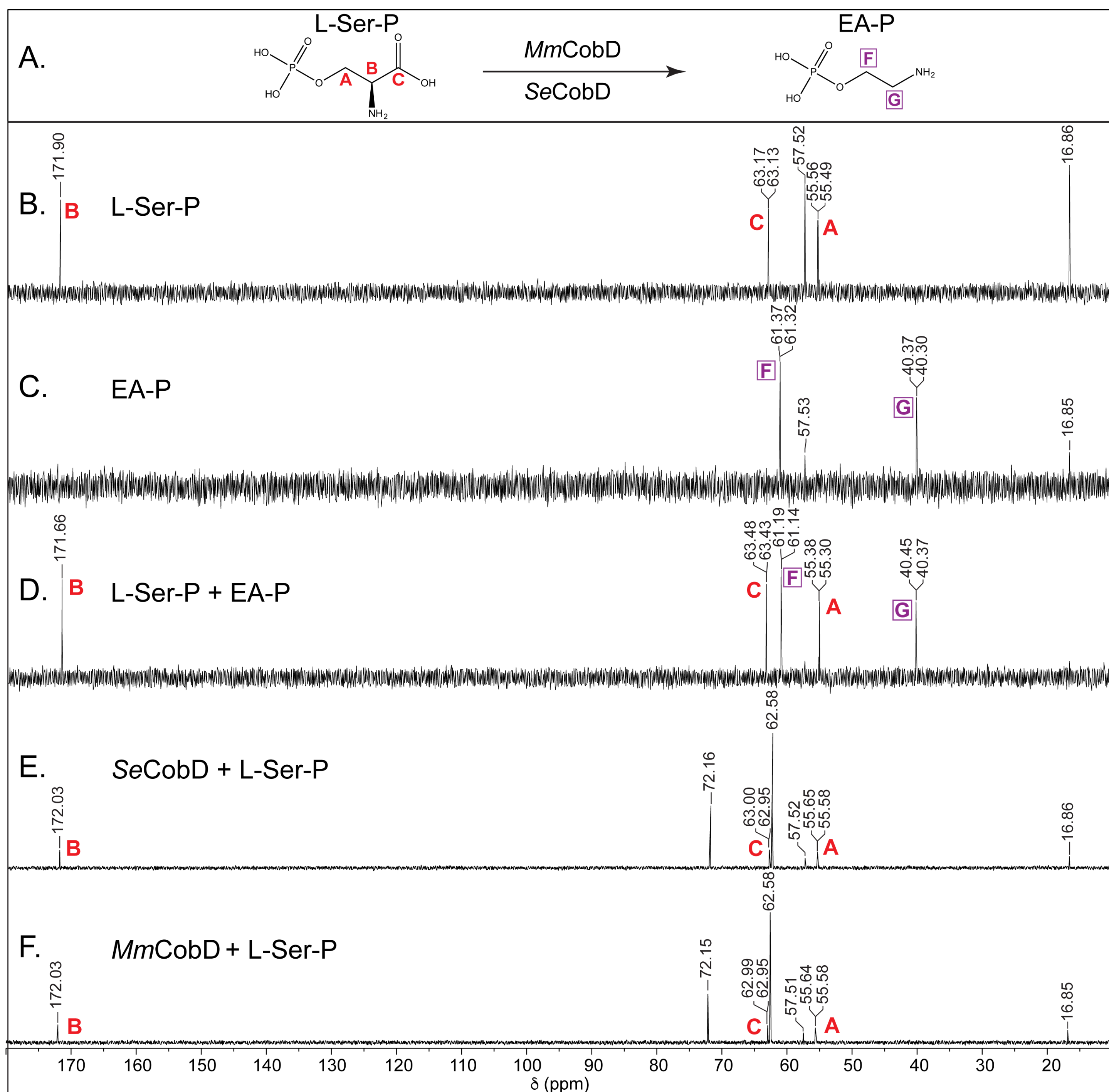


Figure S4. Mm*CobD cannot use L-Ser as a substrate* - Representative ^{13}C-NMR spectra of duplicate independent experiments of *Mm*CobD kinase reaction. Spectra were processed with MestReNova (Mestrelab Research). Reaction mixtures containing sodium phosphate buffer (5 mM, pH 8.5 at 25°C), $MgCl_2$ (5 mM), ATP (40 mM), L-Ser (40 mM), and protein (0.68 μM) were incubated for 1 h at 25ºC prior to the addition of D_2O (17% v/v). Each panel is labeled with the substrate and or protein reaction mixture. Each peak is labeled with the chemical shift value (ppm) and a color-coded letter corresponding to the carbon atom it represents, based on the corresponding chemical shifts for the standards in the no-enzyme control reactions. (*A*) Reaction and chemical structures of substrates with each carbon atom labeled with a color-coded letter. (*B*) L-serine standard (L-Ser; green). (*C*) *O*-Phospho-L-serine standard (L-Ser-P; red). (*D*) Ethanolamine phosphate (EA-P; boxed purple). (*E*) ATP standard (blue). (*F*) Equimolar mixture of ATP, L-Ser, L-Ser-P and EA-P. (*G*) Reaction containing ATP, L-Ser, and *Mm*CobD.

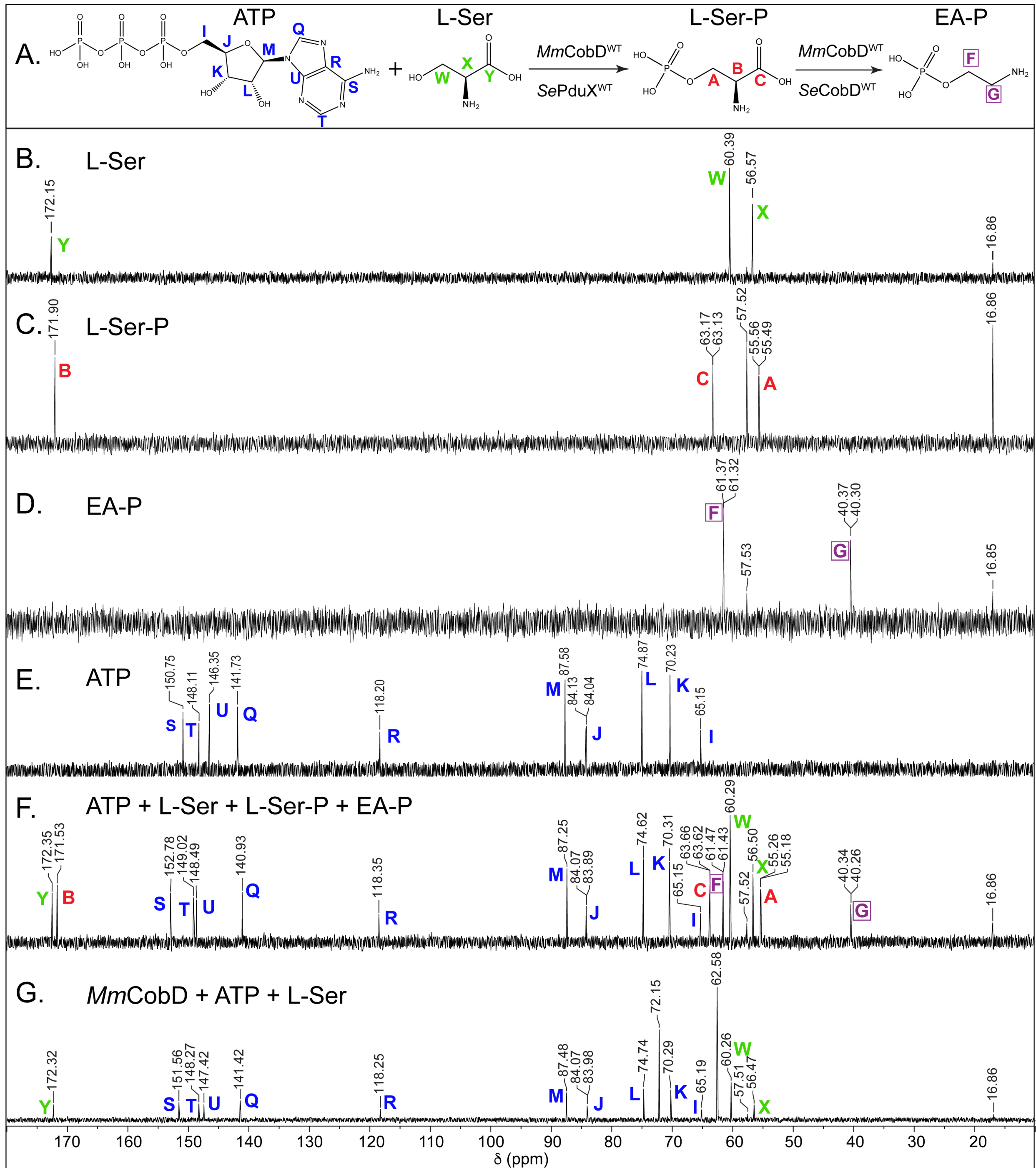


Figure S5. Optimization of Mm*CobD reaction conditions* - ATPase activity assay measured with ADP-Glo™ Assay Kit (Promega). Values were compared to a standard curve for ATP luminescence *vs* conversion and converted into units of ATP produced (mM) per μM of protein, with the standard error of the mean of triplicate reactions represented by the error bars. Unless otherwise indicated, reaction mixture contained HEPES buffer (50 mM, pH 8.5 at 25°C), $MgCl_2$ (1 mM), ATP (0.1 mM), L-Thr (0.1 mM), protein (100 nM), incubated at 25°C for 1 h. (*A*) *Mm*CobD activity as a function of pH. (*B*) *Mm*CobD activity in the presence of salts (100 mM) or other reductants; TCEP (1 mM), sodium dithionite (1 mM). (*C*) *Mm*CobD activity as a function of added divalent metals (1 mM). (*D*) Enzymatic activity of truncations and *Mm*CobD in the presence of ATP (10 mM) and L-threonine (L-Thr, 50 mM) or (*E*) L-serine (L-Ser, 50 mM). *Se*CobD and *Se*PduX proteins were used as negative and positive controls, respectively. (*F*) *MmCobD* activity as a function of L-threonine (L-Thr) concentration (1-100 μM). (*G*) *Mm*CobD activity as a function of ATP concentration (1 - 100 μM) with L-Thr (50 mM). (*H*) *Mm*CobD (6 μM) activity in the presence of ATPase/kinase inhibitors ADP (200 mM), AMP (200 mM), sodium pyrophosphate (PPi, 10 mM), sodium triphosphate (PPPi, 10 mM), adenosine 5′-[γ-thio]triphosphate (ADP-γ-S, 0.1 mM), sodium *ortho*-vanadate (Na_3VO_4, 1 mM), and sodium beryllium fluoride (BeF_2, 2 mM) in HEPES buffer (50 mM, pH 7.5 at 25°C). (*I*) *Mm*CobD and *Mm*CobD$^{1-385}$ activity in the presence ADP-γ-S, (0.1 mM). *Se*EtuP acetate kinase was used as a positive control; protein concentrations (0.1 mM).

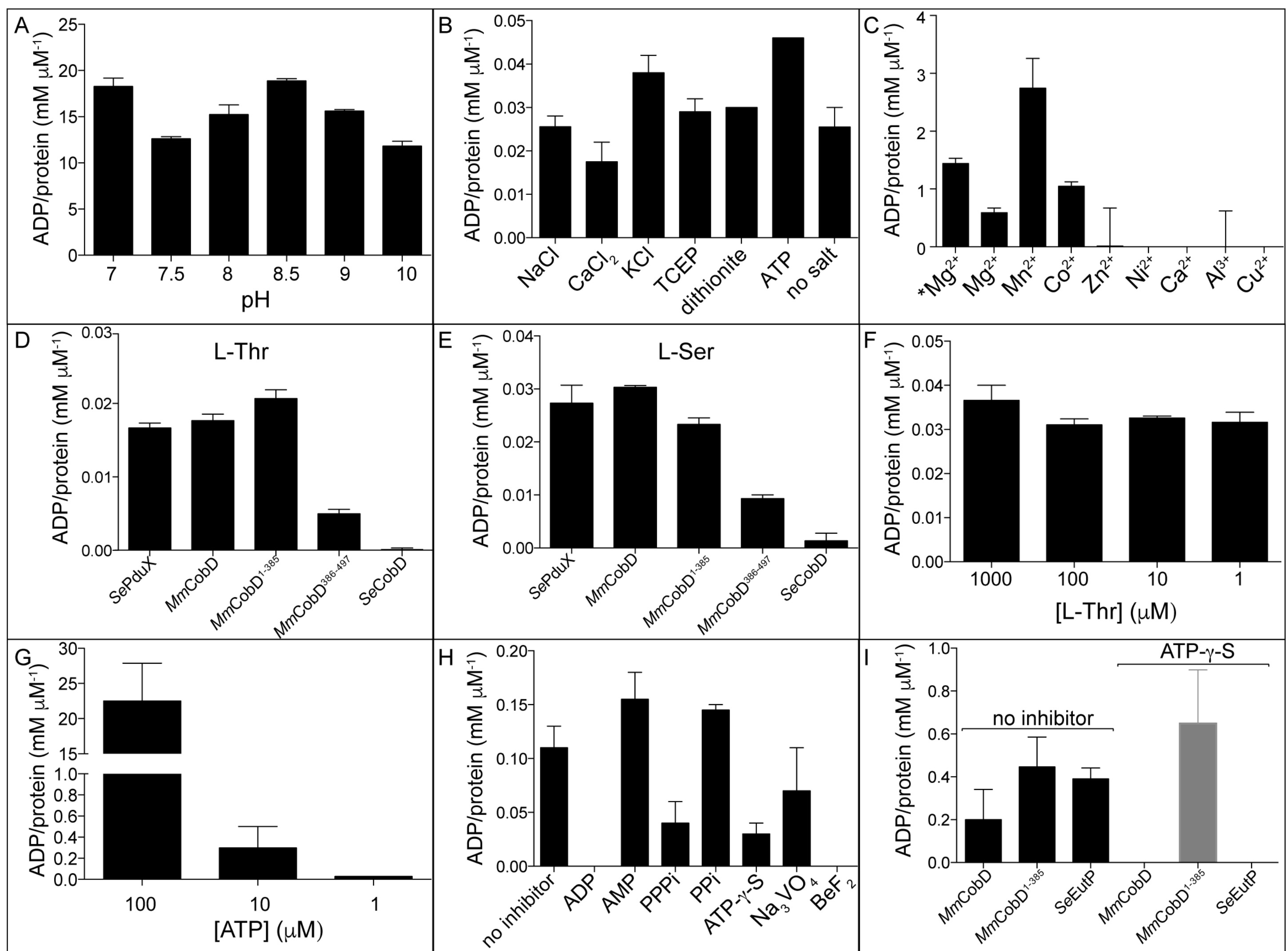


Figure S6. Gel permeation analysis for the determination of the oligomeric state of Mm*CobD proteins* - Gel filtration analysis of *Mm*CobD and truncated variants. Samples were applied to a HiPrep 26/60 Sephacryl S-100 High-resolution gel filtration column using isocratic elution with sodium phosphate (50 mM, pH 7.4 at 4°C) containing NaCl (150 mM). Calibration was performed with Bio-Rad gel filtration standards supplemented with BSA and DNaseI with linear regression to generate the standard curve. *Mm*CobD and *Mm*CobD$^{1\text{-}385}$ correspond to the approximate molecular mass of a dimer and *Mm*CobD$^{386\text{-}497}$ corresponds to the molecular mass of a monomer.

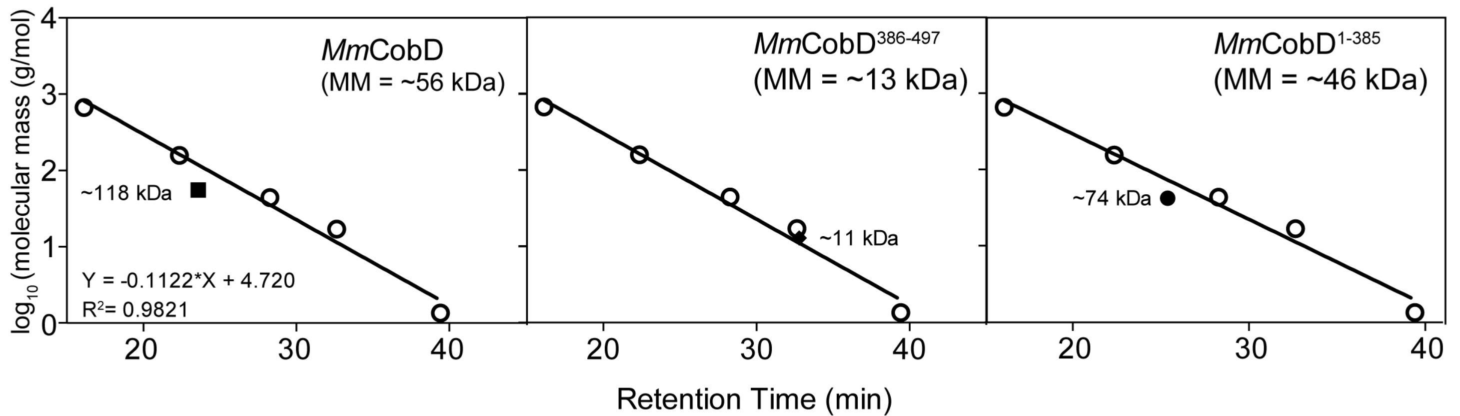


Figure S7. Purification of Mm*CobD proteins* - SDS-PAGE gel of proteins purified normoxically. Lane (*1*) *Mm*CobD, (*2*) *Mm*CobD$^{1\text{-}385}$, (*3*) *Mm*CobD$^{386\text{-}497}$, (*4*) *Se*CobD. Molecular mass markers are shown on the left most lane. Red letters identify (a) high and (b) low molecular weight proteins analyzed by MALDI TOF (see Fig. S8).

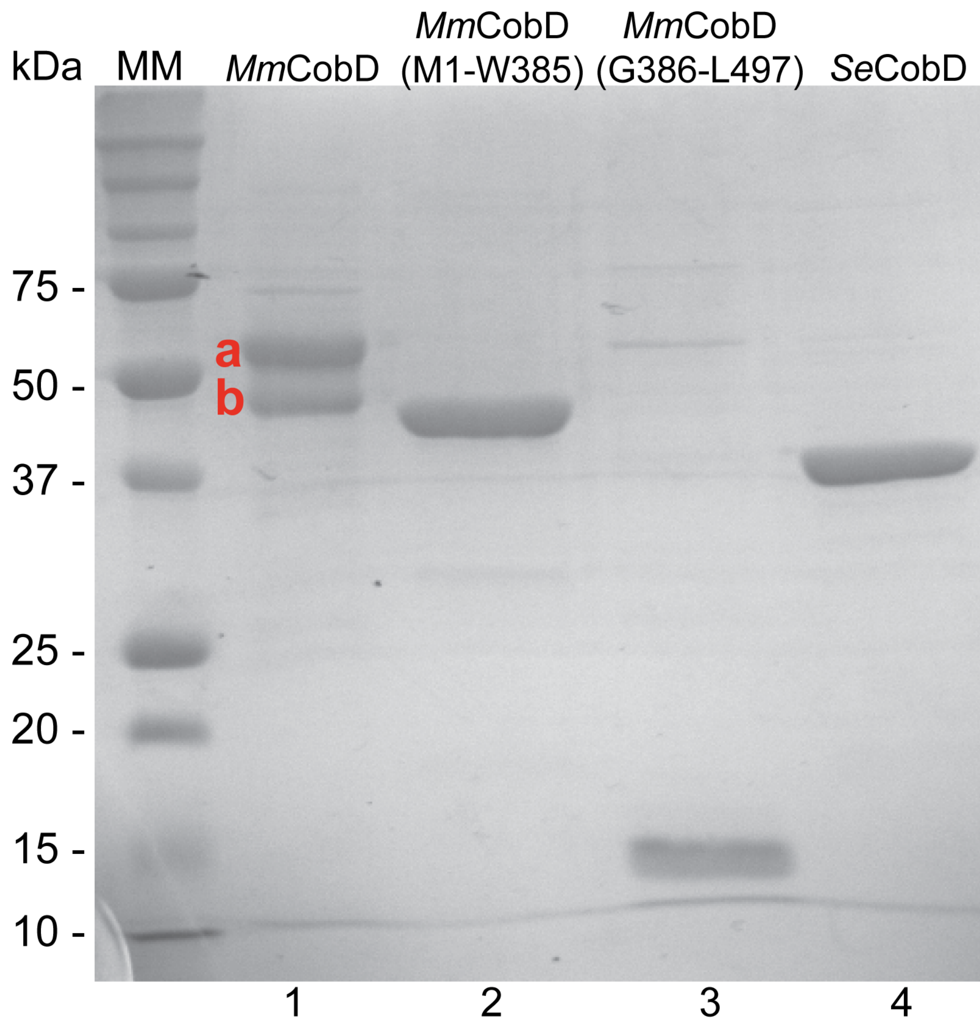
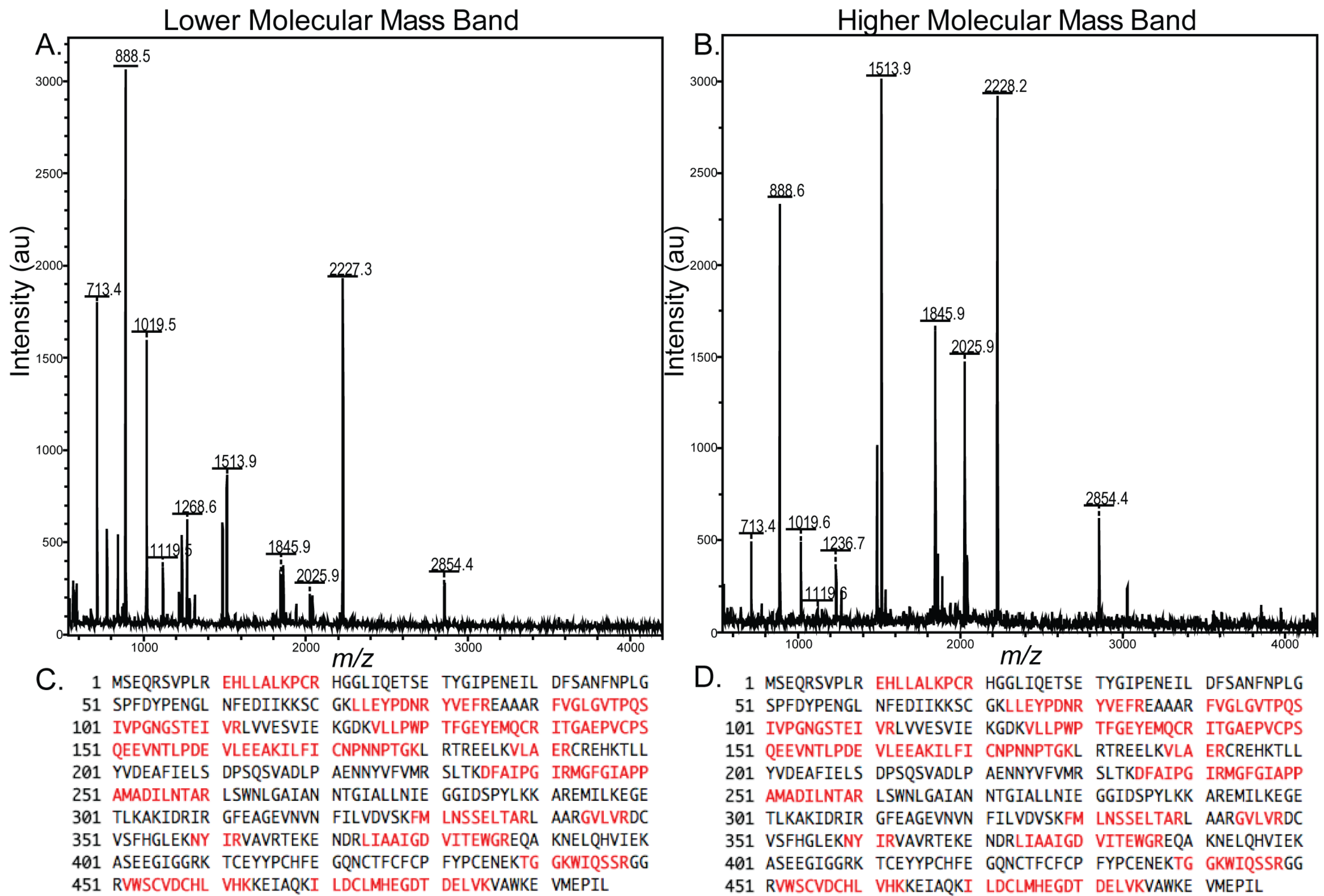


Figure S8. MALDI TOF analysis of MmCobD proteins. Panels A and B show the mass spectra of bands labeled **a** and **b** in lane 1 of Fig. S7. Panels C and D show the coverage of the analysis, which identified these proteins as *Mm*CobD.



SUPPORTING TABLES

Table S1. Strains and plasmids list - *S. enterica* strains are derivatives of sv. Typhimurium strain LT2. Strains and plasmids were constructed during the course of this work unless stated otherwise.

Strains	Relevant genotype	Reference / Source
Salmonella enterica		
JE7088	Δ*metE*2702 *ara-9*	Laboratory collection
Derivatives of JE7088		
JE21644	/ pBAD24 *bla*$^+$	
JE21557	/ pTEV5 *bla*$^+$	
JE12656	*pduX516*	
JE16777	*pduX516 cobD1371*	
JE2216	*cobD1302*::Tn*10*d(*cat*$^+$)	Laboratory collection
JE21654	*cobD1302*::Tn*10*d(*cat*$^+$) / pBAD24 *bla*$^+$	
JE21653	*cobD1302*::Tn*10*d(*cat*$^+$) / pTEV5 *bla*$^+$	
JE6158	*cobD1302*::Tn*10*d(*cat*$^+$) / pCOBD6 *bla*$^+$	Laboratory collection
JE21656	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD7 *bla*$^+$	
JE21781	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD9 *bla*$^+$	
JE21657	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD13 *bla*$^+$	
JE21646	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD17 *bla*$^+$	
JE21659	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD18 *bla*$^+$	
JE21782	*cobD1302*::Tn*10*d(*cat*$^+$) / pMmCOBD19 *bla*$^+$	
JE18765	*pduX516* / pMmCOBD7 *bla*$^+$	
JE21619	*pduX516* / pMmCOBD9 *bla*$^+$	
JE18970	*pduX516* / pMmCOBD13 *bla*$^+$	
JE19204	*pduX516* / pMmCOBD17 *bla*$^+$	
JE21336	*pduX516* / pMmCOBD18 *bla*$^+$	
JE19205	*pduX516* / pMmCOBD19 *bla*$^+$	
JE16786	*pduX516 cobD1371* / pBAD24 *bla*$^+$	
JE19200	*pduX516 cobD1371* / pMmCOBD17 *bla*$^+$	
Escherichia coli		
C43 (λDE3)	F$^-$ *ompT gal hsdS*$_B$ (rB$^-$mB) [*dcm*] [*Ion*]	[1]
DH5α	F$^-$/ *endA1 hsdR17*(rk$^-$, mk$^+$) *glnV44 thi-1 recA1 gyrA96* (Nx^R) *relA1* U169 *deoR* (Φ80d *lacZ* Δ*M15* Δ(*lacZYA-argF*) *phoA supE44*	[2]
Plasmids		
pCOBD6	*S. enterica cobD*$^+$ in pT7-7 *bla*$^+$	Laboratory collection
pMmCOBD7	*M. mazei cobD19* (encoding *Mm*CobD$^{1\text{-}385}$) in pBAD24 *bla*$^+$	
pMmCOBD9	*M. mazei cobD19* (encoding *Mm*CobD$^{1\text{-}385}$) in pTEV5 *bla*$^+$	
pMmCOBD13	*M. mazei cobD20* (encoding *Mm*CobD$^{386\text{-}497}$) in pBAD24 *bla*$^+$	
pMmCOBD17	*M. mazei cobD*$^+$ in pBAD24 *bla*$^+$	
pMmCOBD18	*M. mazei cobD*$^+$ in pTEV5 *bla*$^+$	
pMmCOBD19	*M. mazei cobD20 cobD*$^{386\text{-}497}$ in pTEV5 *bla*$^+$	
pPDU15	*S. enterica pduX*$^+$ in pBAD30 *bla*$^+$	Laboratory collection
pPDU23	*S. enterica pduX*$^+$ in pTEV5 *bla*$^+$	Laboratory collection
pBAD24	cloning/complementation vector *bla*$^+$	[3]
pBAD30	cloning/complementation vector *bla*$^+$	[3]
pTEV5	cloning/overexpression vector, *N*-terminal rTEV cleavable His$_6$ tag *bla*$^+$	[4]

Table S2. Primers used in this study - Restriction sites are indicated in boldface type.

Name	Sequence
Deletion primers	
cobD_wan5'_b	CAGTTACTGGATTTTAGCGCAAACATTAATCCGCTGGGTATGGTGTAGGCTGGAGCTGCTTC
cobD_wan3'_b	ACTACATCAATCAGCAGGGGCTATACCGGTAAGCACATTCATATGAATATCCTCCTTAG
pduX_wan5'_b	ATGCGCGCACACTATTCGTACCTGAAAGGTGATAATGTGGTGTAGGCTGGAGCTGCTTC
pduX_wan3'_b	GCCAGTGACCATCTTGAGTAAATGTTGTTTTGGCCAGTGCATATGAATATCCTCCTTAG
Cloning primers	
pduX_KpnI5'	TAGG**GGTACC**ATGCGCGCACACTATTCGTA
pduX_Hind3'	TAGG**AAGCTT**TCACTGCAGTTTGACCCCG
CobD-Nhe5'	TAGG**GCTAGC**ATGGCGTTATTCAACAGCG
CobD-SacI3'	TAGG**GAGCT**CTCAATCAGCAGGGGCTATACC
MmCobD_Nhe5'	TAGG**GCTAGC**ATGTCAGAGCAAAGAAGTGTACCT
MmCobD_Sma3'	TAGG**CCCGGG**TCATAAGATAGGCTCCATCACTTCTTTC
MmCobD_short_Kpn5'	TAGG**GGTACC**ATGTCAGAGCAAAGAAGTGTACCT
MmCobD_short_Nhe5'	TAGG**GCTAGC**ATGTCAGAGCAAAGAAGTGTACCT
MmCobD_short_Hind3'	TAGG**AAGCTT**TTATTACCATTCGGTGATAACGTCC
MmCobDzf_NheI5'	TAGG**GCTAGC**ATGGGCAGAGAACAGGCAAA
MmCobDzf_SacI3'	TAGG**GAGCTC**TCATAAGATAGGCTCCATCACTTC
Sequencing primers	
pduX_flank5'	GCGTAATGCGACATTTATCCA
pduX_flank3'	TGAGGCGATTCAGGGTATCAT
cobD_flank5'	CGACATTGGCCTCGGTTT
cobD_flank3'	GAAACGCCCTGGCTTAAT

REFERENCES

[1] Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, *J. Mol. Biol. 260*, 289-298.

[2] Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., De Cruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z., and Graham, M. W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants, *Nucl. Acids Res. 17*, 3469-3478.

[3] Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, *J. Bacteriol. 177*, 4121-4130.

[4] Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, *Plasmid 59*, 231-237.