Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Cloning, purification and preliminary crystallographic analysis of cobalamin methyltransferases from *Rhodobacter capsulatus*

Arefeh Seyedarabi,[a] Thomas Hutchison,[a] Teng Teng To,[a] Evelyne Deery,[b] Amanda Brindley,[b] Martin J. Warren[b] and Richard W. Pickersgill[a]*

[a]School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, England, and [b]Centre for Molecular Processing, School of Biosciences, University of Kent, Giles Lane, Canterbury, Kent CT2 7NJ, England

Correspondence e-mail: r.w.pickersgill@qmul.ac.uk

Received 11 August 2010
Accepted 21 October 2010

Of the 30 biosynthetic steps necessary for the production of cobalamin (vitamin B_{12}), eight involve the addition of *S*-adenosylmethionine-derived methyl groups to the tetrapyrrole framework. These eight methyl additions are catalysed by six canonical methyltransferase domains and one noncanonical methyltransferase domain. Recombinant forms of four methyltransferases from *Rhodobacter capsulatus*, CobJ, CobM, CobF and CobL, and of the C-terminal noncanonical domain of CobL (CobL-C) have been crystallized, some in more than one crystal form. Most of the crystals diffracted to beyond 2.5 Å resolution and all are suitable for structure determination. Crystals of CobM and CobJ, which are involved in ring contraction, and of CobL, which is involved in two methylations and decarboxylation, are reported for the first time.

1. Introduction

There are seven methyltransferases that add eight *S*-adenosylmethionine-derived methyl groups to the tetrapyrrole framework and display a high level of regiospecificity and temporal specificity. The discrepancy between the number of methyltransferases and the number of methylations arises from the fact that the first methyltransferase in the sequence, CobA, carries out two methylations. The remaining enzymes catalyse a single methylation reaction. To complicate things further, there are two similar though distinct routes for vitamin B_{12} biosynthesis, which differ in a requirement for molecular oxygen and the timing of cobalt insertion. These two routes are referred to as the aerobic and anaerobic pathways. However, both pathways require the same temporal methylation pattern although the intermediates differ, mainly owing to the presence or absence of cobalt. Table 1 lists the methyltransferases associated with both the late cobalt insertion (aerobic) and the early cobalt insertion pathways (anaerobic) of B_{12} biosynthesis. Within this pathway, six of the seven methyltransferases display sequence similarity to other class III methyltransferases, indicating that these enzymes have evolved from a common ancestor. Indeed, most of the methyltransferases of the aerobic route have a homologue in the anaerobic pathway, indicating that the aerobic and anaerobic routes split after the primordial B_{12} pathway had evolved (Holliday *et al.*, 2007). The structure of the canonical cobalamin methyltransferase class III domain, CbiF, has been solved (Schubert *et al.*, 1998), as has the structure of the noncanonical fold CbiT (Keller *et al.*, 2002). The sequence identities of the three canonical methyltransferases crystallized here are, in order of decreasing similarity, 20.0% (CobJ/CobM), 16.0% (CobJ/CobF) and 12.5% (CobF/CobM). The canonical (N-terminal) domain of CobL has 21.5, 18.9 and 13.6% identity to CobJ, CobM and CobF, respectively. The noncanonical (C-terminal) domain of CobL has 41% identity to CbiT. Our current understanding of the evolution of complex metabolic pathways is based upon the idea that pathways evolved from a collection of broad-specificity enzymes, which sub-

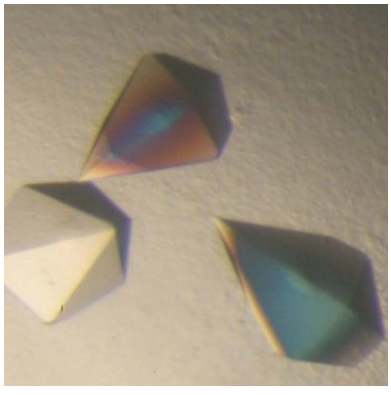


Table 1
Methyltransferases of cobalamin biosynthesis.

The enzymes are given in order according to their position in the biosynthetic pathway (with the exception of CobA, which is common to both pathways). The methyltransferase of the early cobalt insertion (aerobic) pathway is given followed by the equivalent enzyme of the late cobalt insertion (anaerobic) pathway. Noncanononical refers to a structure which does not belong to the cobalamin-biosynthetic methyltransferase fold. CobL-N and CobL-C refer to the N- and C-terminal domains of CobL, respectively.

Methylates	Enzyme	Pathway	Catalyses the transformation	Reference to structure
C2 and C7	CobA	Both	Uroporphyrinogen III to precorrin-2	Vévodová *et al.* (2004), PDB code 1s4d
C20	CobI	Aerobic	Precorrin-2 to precorrin-3A	None
C20	CbiL	Anaerobic	Cobalt factor II to cobalt factor III	Frank *et al.* (2007), Wada *et al.* (2007), PDB codes 2qbu, 2eok, 2eon
C17	CobJ	Aerobic	Precorrin-3B to precorrin-4	This work
C17	CbiH	Anaerobic	Cobalt factor III to cobalt factor IV	None
C11	CobM	Aerobic	Precorrin-4 to precorrin-5	This work
C11	CbiF	Anaerobic	Cobalt factor IV to cobalt factor V	Schubert *et al.* (1998), PDB code 2cbf
C1	CobF	Aerobic	Precorrin-5 to precorrin-6X	This work; structural genomics project, PDB code 2npn
C1	CbiD	Anaerobic	Cobalt factor V to cobalt factor VI	Noncanonical fold; structural genomics project, PDB code 1sr8
C5	CobL-N	Aerobic	Precorrin-6Y to precorrin-7	This work
C5	CbiE	Anaerobic	Cobalt factor VI to cobalt factor VII	Structural genomics project, PDB code 2bb3
C15	CobL-C	Aerobic	Precorrin-7 to hydrogenobyrinic acid	This work; structural genomics project, PDB code 3hm2
C15	CbiT	Anaerobic	Cobalt factor VII to hydrogenobyrinoic acid	Keller *et al.* (2002), PDB codes 1f38, 1kxz, 1l3b, 1l3c, 1l3i; structural genomics project, PDB code 2yxd

sequently refined their binding pattern to become particular for one substrate. This theory is known as the patchwork evolution model and infers that the recruitment of single enzymes from different pathways is the driving force for pathway evolution, in which enzymes evolve to catalyse different reactions on the same structural scaffold. This explains the appearance of enzyme superfamilies, in which enzymes retain a common structural or mechanistic strategy for catalysis within a shared protein framework.

The alternative theory is the retrograde model of pathway evolution, in which the pathway evolves backwards from a key metabolite. Moreover, it was suggested that the enzymes of a pathway were descended from a common ancestral gene by a process of gene duplication followed by functional differentiation, driven by substrate depletion. In this model, the enzymes in a pathway would have overlapping substrate specificities with an inherent molecular memory of the original key metabolite. Such a theory was used to explain the existence of operons and end-product inhibition of the first enzyme of the pathway and could be used to explain the prevalence of class III methyltransferases in the cobalamin-biosynthetic pathway. Solving these structures in complex with substrates or products will help to unravel the determinants of substrate specificity and regiospecificity and the chemistry at work in the ring contraction, deacylation and decarboxylation reactions, and will illuminate the evolutionary relationships between these enzymes.

2. Materials and methods

2.1. Protein expression and purification

The DNA fragments corresponding to amino-acid residues 1–245, 1–255, 1–261 and 1–396 encoded by the *cobJ*, *cobF*, *cobM* and *cobL* genes, respectively, were amplified from *Rhodobacter capsulatus* genomic DNA by polymerase chain reaction (PCR) using the FastStart High Fidelity PCR system (Roche). The PCR products were initially cloned into modified pET3a (Novagen), which contained one *Spe*I site 5′ to *Bam*HI, by *Nde*I and *Spe*I digests (Table 2). The same restriction enzymes were used to subclone the products into pET14b (Novagen), which had also been modified by adding an *Spe*I site 5′ to *Bam*HI. The DNA fragment corresponding to amino acids 214–396 encoded by the C-terminal domain of the *cobL* gene was amplified by PCR using the *pET3a::cobL(1–396)* construct as template and cloned directly into the modified pET14b. The pET14b vector added an N-terminal His_6 tag and a thrombin cleavage site to all five enzymes. Recombinant clones were selected for by ampicillin resistance. The

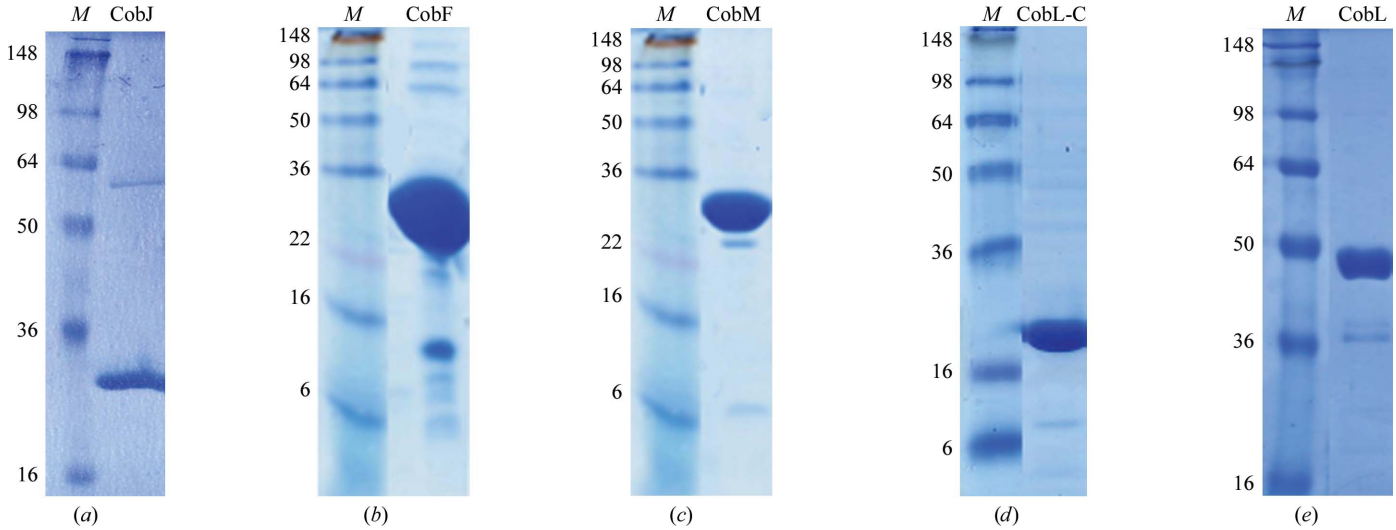
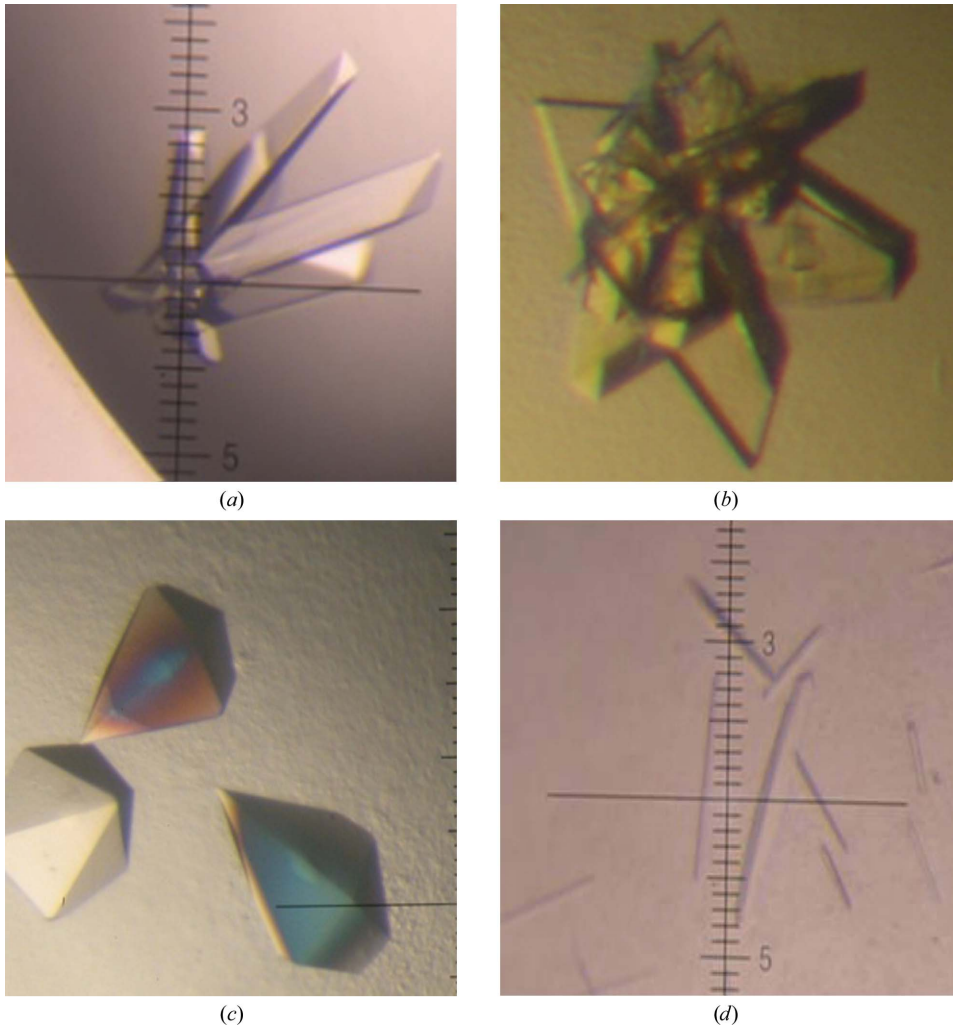


Figure 1
SDS–PAGE analysis of the purified *R. capsulatus* methyltransferases (*a*) CobJ, (*b*) CobF, (*c*) CobM, (*d*) CobL-C and (*e*) CobL as visualized by Coomassie Blue staining. The molecular masses of the marker proteins in the lanes marked *M* are given in kDa.

expression constructs were transformed into competent *Escherichia coli* BL21 Star (DE3) pLysS cells. A small 10 ml overnight culture grown at 310 K in Luria–Bertani broth containing 100 µg ml^{-1} ampicillin and 34 µg ml^{-1} chloramphenicol was used to inoculate a 1 l culture. Cells were grown to an OD_{600} of 0.6. To express the recombinant proteins, the cells were induced by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.5 m*M*. After overnight induction at 291 or 303 K, the cells were harvested by centrifugation at 4950*g* for 20 min at 277 K. Subsequently, the cell pellets were resuspended in 25 m*M* Tris–HCl buffer pH 8.0 containing 10 m*M* imidazole and 400–500 m*M* NaCl (buffer *A*) and frozen at 193 K. For purification of the overproduced proteins, the frozen cell pellets were thawed and sonicated, followed by centrifugation at 26 900*g* for 30 min at 277 K to remove cell debris. A 5 ml HisTrap nickel immobilized-metal affinity chromatography column (GE Healthcare) equilibrated with buffer *A* was attached to an ÄKTA FPLC system (GE Healthcare). The soluble fraction was applied onto the column and the column was washed with 40 ml buffer *A* before the target protein was eluted with a linear gradient of imidazole (from 10 to 500 m*M* in 20 column volumes) using buffer *B* consisting of 25 m*M* Tris–HCl buffer pH 8.0, 500 m*M* imidazole and 400–500 m*M* NaCl. The fractions containing the His_6-tagged Cobs (CobJ, CobF, CobM, CobL and CobL-C) were identified by SDS–PAGE analysis. The His_6-tagged Cobs were concentrated using 10 kDa molecular-weight cutoff centrifugal concentrators (Vivaspin, Sartorius) and applied onto a Superdex 200 size-exclusion column (GE Healthcare). The CobL-C protein was eluted in 25 m*M* 2-(*N*-morpholino)ethanesulfonic acid pH 6.0 and 250 m*M* NaCl, while all of the other Cobs were eluted in 50 m*M* Tris–HCl pH 8.0 and 100 m*M* NaCl. The protein fractions from the Superdex column were further analysed by SDS–PAGE analysis (Fig. 1) and concentrated. CobJ protein was concentrated to 10–20 mg ml^{-1}, CobF to 25–50 mg ml^{-1}, CobM to 20–40 mg ml^{-1}, CobL to 8–10 mg ml^{-1} and CobL-C to 2–

Table 2
List of primers used for PCR amplification of the methyltransferases.

Gene	Primers	Restriction sites
cobJ	Forward: CAT**CATATG**AGCGGTTGGGTCACG	*Nde*I
	Reverse: CAT**ACTAGT**GGCTTGCCGATCACGGC	*Spe*I
cobF	Forward: CAT**CATATG**ATCGAGCTTTCCCTG	*Nde*I
	Reverse: CAT**ACTAGT**CACGCCGCGGGCACATTG	*Spe*I
cobM	Forward: CAT**CATATG**ACGGTGCATTTCATCG	*Nde*I
	Reverse: CAT**ACTAGT**CATTCGCTGCCCTCCGG	*Spe*I
cobL	Forward: CTA**CATATG**TCTGATCCGTGGTTG	*Nde*I
	Reverse: CAT**ACTAGT**CATCTTTGCCCGCTCC	*Spe*I
cobL-C	Forward: CAT**CATATG**GGCACGGGGCTTTCGCAG	*Nde*I
	Reverse: as for *cobL* above	



(*a*) (*b*) (*c*) (*d*)

Figure 2
Crystals of *R. capsulatus* methyltransferases: (*a*) CobJ (type I), (*b*) CobF (type I), (*c*) CobM (type I) and (*d*) CobL. This is the first report of CobJ, CobM and CobL crystals; CobF crystals have been reported previously from a different species.

Table 3
Crystallographic data statistics.

Values in parentheses are for the highest resolution shell. All crystals contained an N-terminally hexahistidine-tagged methyltransferase, with the exception of the CobM (I) crystal which had the tag removed as described in §2.

Enzyme (crystal type)	CobJ (I)	CobJ (II)	CobJ (III)	CobF (I)	CobF (II)	CobF (III)	CobM (I)	CobM (II)	CobL-C (I)	CobL-C (II)	CobL
Synchrotron	DLS IO3	ESRF ID14-1	ESRF ID14-1	SRS PX10.1	ESRF ID14-1	ESRF ID29	SRS PX10.1	ESRF ID29	ESRF ID29	DLS IO2	ESRF ID14-4
Wavelength (Å)	0.981	0.933	0.933	1.117	0.933	1.011	1.117	1.011	1.011	0.980	0.939
Detector	ADSC Q315R	ADSC Q210	ADSC Q210	MAR Mosaic	ADSC Q210	ADSC Q315R	MAR Mosaic	ADSC Q315R	ADSC Q315R	ADSC Q315	ADSC Q315R
Space group	$P2_12_12_1$	$P2_1$	$C222_1$	$P2_12_12_1$	$P2_1$	$C222_1$	$P6_5$	$P2_122_1$	$C2$	$P2_12_12_1$	$I2_12_12_1$
Unit-cell parameters											
a (Å)	85.64	68.31	49.73	61.75	54.15	61.38	84.52	51.05	114.35	62.67	103.99
b (Å)	110.99	69.43	112.85	89.42	84.78	89.28	84.52	54.50	44.80	86.39	112.33
c (Å)	116.19	72.65	91.25	107.14	58.61	107.99	144.64	175.85	84.11	149.96	373.69
β (°)		118.16			98.65				119.75		
Unique reflections	54689 (7910)	41959 (6172)	7019 (1007)	95539 (13791)	41043 (5940)	22829 (3264)	39472 (5720)	17755 (2525)	10402 (1514)	21312 (3045)	36000 (5188)
Multiplicity	3.9 (4.0)	3.6 (3.7)	6.7 (6.9)	6.9 (6.4)	3.6 (3.5)	4.5 (4.3)	9.8 (8.5)	7.0 (7.2)	3.6 (3.7)	5.5 (5.6)	14.4 (14.8)
Resolution (Å)	58.56–2.22 (2.34–2.22)	34.72–1.97 (2.08–1.97)	48.09–2.75 (2.90–2.75)	26.78–1.50 (1.58–1.50)	33.23–1.90 (2.00–1.90)	53.99–1.93 (2.03–1.93)	42.26–2.00 (2.11–2.00)	87.93–2.50 (2.64–2.50)	72.93–2.70 (2.85–2.70)	74.98–2.78 (2.93–2.78)	107.57–3.20 (3.37–3.20)
Completeness (%)	99.4 (99.9)	98.8 (100.0)	99.7 (100.0)	99.9 (100.0)	99.5 (98.9)	99.9 (99.7)	100.0 (100.0)	100.0 (100.0)	99.9 (100.0)	99.9 (100.0)	98.7 (98.6)
$R_{\rm merge}$† (%)	8.1 (52.5)	19.5 (79.2)	10.3 (63.1)	6.4 (41.3)	8.1 (40.7)	11.6 (27.5)	6.8 (62.0)	14.2 (30.5)	6.0 (21.4)	13.2 (63.2)	17.3 (65.2)
$R_{\rm p.i.m.}$‡ (%)	4.6 (30.9)	12.4 (47.9)	4.3 (25.6)	2.6 (17.6)	4.9 (25.0)	6.3 (15.1)	2.3 (22.3)	5.8 (12.2)	3.7 (13.1)	6.3 (30.0)	4.6 (17.4)
Mean $I/\sigma(I)$	12.3 (2.6)	7.3 (2.7)	12.2 (2.7)	15.2 (3.7)	13.1 (3.1)	9.0 (4.6)	19.2 (3.3)	9.9 (5.6)	14.6 (5.3)	9.4 (3.0)	17.3 (5.4)
Wilson B factor (Å^2)	38.3	16.9	66.5	16.3	18.7	20.3	33.2	35.1	56.5	63.4	66.8
Mol. per ASU§	4	2	1	2	2	1	2	2	2	4	4
$V_{\rm M}$¶ (Å^3 Da^{-1})	2.51	2.76	2.34	2.39	2.15	2.39	2.51	1.93	2.04	2.21	2.92

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for each unique Bragg reflection with indices hkl and $\langle I(hkl)\rangle$ is the average intensity for multiple measurements of this reflection. ‡ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where N is the multiplicity and the other variables are as defined for $R_{\rm merge}$ (Weiss, 2001). § Number of molecules in the asymmetric unit. ¶ Matthews coefficient (Matthews, 1968).

6 mg ml^{-1}. The protein concentrations were calculated using the extinction coefficients for each of the proteins as calculated from the sequence with the help of the *ProtParam* program.

His$_6$-tagged CobM and CobL proteins at 7 and 10 mg ml^{-1}, respectively, were treated with thrombin (Novagen) to cleave off the His$_6$ tag. 1 U of thrombin was sufficient to cleave 1 mg protein in thrombin cleavage buffer consisting of 20 m*M* Tris–HCl pH 8.4, 150 m*M* NaCl and 2.5 m*M* $CaCl_2$ over the course of 18–20 h at room temperature. A final concentration of 1 m*M* phenylmethanesulfonylfluoride was added to stop the thrombin activity after the incubation period. Thrombin was removed by size-exclusion chromatography (Superdex 200). The crystals described below are all from His-tagged protein (with the additional N-terminal sequence MGSSHHHHHHSSGLVPRGSHM), with the exception of the CobM (type I) crystals, which were from cleaved protein with the additional N-terminal sequence GSHM.

2.2. Crystallization

Crystallization was performed at 292 K using the hanging-drop vapour-diffusion method in 24-well plates. The methyl donor *S*-adenosylmethionine (SAM; Sigma) was used to aid crystallization of the proteins. Two crystallization screens exploiting sparse-matrix sampling (Jancarik & Kim, 1991), Crystal Screen and Crystal Screen 2 (Hampton Research), with a total of 98 different conditions were used to screen for crystallization conditions. The SAM concentration was varied in the range 0–10 m*M*. 1–2 µl drops consisting of a 1:1 ratio of protein and reservoir solution were equilibrated against 500 µl of each reservoir solution. CobJ, CobM and CobL crystals are reported here for the first time, whereas CobF and CobL-C have previously been crystallized from different species and reported by structural genomics projects (Table 1).

His$_6$-tagged CobJ crystals grew in 2–4 d from Crystal Screen condition No. 46 [0.2 *M* calcium acetate hydrate, 0.1 *M* sodium cacodylate trihydrate pH 6.5 and 18% polyethylene glycol (PEG) 8000]. Three crystal forms (types I–III; Table 3) were characterized; slightly varying calcium acetate and PEG concentrations may have contributed to these different forms. His$_6$-tagged CobF crystals (types I–III; Table 3) were obtained within 2–5 d in Crystal Screen condition No. 37 (0.1 *M* sodium acetate trihydrate pH 4.6 and 8% PEG 4000). His$_6$-tagged CobM crystals (type II) were obtained within 2–5 d in Crystal Screen condition No. 17 (0.2 *M* lithium sulfate monohydrate, 0.1 *M* Tris pH 8.5 and 30% PEG 4000) and also in Crystal Screen 2 condition No. 38 (0.1 *M* HEPES pH 7.5 and 20% PEG 10 000), while crystals of His$_6$-cleaved CobM (type I) were obtained within 5–7 d in Crystal Screen condition No. 41 (0.1 *M* HEPES pH 7.5, 10% 2-propanol and 20% PEG 4000). His$_6$-tagged CobL-C crystals were obtained within 1–2 d in numerous conditions including Crystal Screen condition No. 9 (0.2 *M* ammonium acetate, 0.1 *M* trisodium citrate dihydrate pH 5.6 and 30% PEG 4000; type II in Table 3) and Crystal Screen condition No. 40 (0.1 *M* trisodium citrate dihydrate pH 5.6, 20% 2-propanol and 30% PEG 4000; type I in Table 3). The other conditions in which CobL-C crystals were obtained were Crystal Screen conditions 6, 7, 17 and 18. His$_6$-tagged CobL crystals grew after 1–2 weeks from Crystal Screen condition No. 3 (0.4 *M* ammonium phosphate) and Crystal Screen condition No. 14 (0.2 *M* calcium chloride, 0.1 *M* Na HEPES pH 7.5 and 28% PEG 400). Fig. 2 shows crystals of CobJ, CobF, CobM and CobL; in all cases, the best crystals grew in the presence of 1–5 m*M* SAM. Crystals grown in the absence of SAM diffracted less well than those grown in its presence.

2.3. Cryoprotection and data collection

Cryoprotectants consisted of the reservoir condition with a slightly increased precipitant concentration (typically increased by 5–10% of the original concentration depending upon the age of the crystallization trial) augmented with 20% glycerol for His$_6$-tagged CobJ (types I, II and III); 25% 2-methyl-1,3-propanediol (MPD) for His$_6$-tagged CobF (types I, II and III); 15% glycerol for His$_6$-tagged CobM (type II); 25% MPD for His$_6$-cleaved CobM (type I); 5 and 15%

Table 4
Molecular-replacement results.

These results and inspection of the resultant σ_A-weighted $2F_{obs} - F_{calc}$ Fourier syntheses suggest that the structures of CobF, CobM and CobL-C can readily be solved by molecular replacement. In each case the methyltransferase dimer was used as the search model.

Methyltransferase	PDB code of search model†	Sequence identity (%)	Probable no. of dimers per asymmetric unit	*MOLREP* score
CobJ	1s4d	24	2	0.296‡
CobF	2npn	40	1	0.455
CobM	2cbf	37	1	0.544
CobL-C	3hm2	41	1	0.314
CobL	3hm2	21	2	0.417§

† The original citations for 1s4d and 2cbf are Vévodová *et al.* (2004) and Schubert *et al.* (1998), respectively; the other two structures, 2npn and 3hm2, are unpublished results from structural genomics projects. ‡ The CobJ score is low, suggesting that experimental phases will be needed; this was confirmed by inspection of the electron-density map. § Poor discrimination between the first solution and the next best solution (score 0.415) suggests that this is not a solution despite the high score. Type I crystals of each of the methyltransferases were used for these calculations (see Table 3).

glycerol, respectively, for His_6-tagged CobL-C types I and II; and 30% PEG 400 for His_6-tagged CobL. Diffraction data from 11 crystals were collected at 100 K and processed with *MOSFLM* (Leslie, 2006). *POINTLESS* was used to aid identification of the space group for each data set before scaling using *SCALA* (Evans, 2006). Where there was ambiguity, the space group given in Table 3 is that with the highest score in molecular replacement calculated using *MOLREP* (Vagin & Teplyakov, 2010; Table 4).

3. Results and discussion

Four methyltransferases from *R. capsulatus*, CobJ, CobF, CobM and CobL, were successfully cloned, overproduced and crystallized (Figs. 1 and 2). The C-terminal domain of CobL (CobL-C; a non-canonical methyltransferase domain) was also successfully cloned, overproduced and crystallized. In each case, the authenticity of the cloned DNA was confirmed by sequencing. All five constructs expressed large quantities of protein and well diffracting crystals were grown of all of them, with the possible exception of CobL, which gave reasonable diffraction (diffraction limit of 3.2 Å). CobL has two domains, a canonical methyltransferase domain and a noncanonical methyltransferase domain (see Table 1), and there may be some flexibility between the two domains that might contribute to the lower diffraction limit for these crystals. The inclusion of SAM was crucially important in obtaining well diffracting crystals of these SAM-binding methyltransferases. Crystals were mostly grown with the hexahistidine tag intact, although removal of the tag from CobM did give slightly better diffracting crystals (CobM type I). The proteins had good stability, with the exception of the C-terminal domain of CobL (CobL-C), which was stabilized by the addition of sodium chloride. The quality of the diffraction data is presented in Table 3. It is immediately clear from molecular-replacement calculations that CobF, CobM and CobL-C can be readily solved using existing methyltransferase structures and refinement and validation of these structures is under way (Table 4). Experimental phases are being determined for CobJ and CobL using SeMet-labelled proteins.

We acknowledge support from the Biotechnology and Biology Research Council (BBSRC; grant No. BB/E00213) and the use of the Diamond Light Source (DLS; Oxford), ESRF (Grenoble) and SRS (Daresbury). We thank Dr Nora Cronin for help arranging data collection.

References

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Frank, S., Deery, E., Brindley, A. A., Leech, H. K., Lawrence, A., Heathcote, P., Schubert, H. L., Brocklehurst, K., Rigby, S. E. J., Warren, M. J. & Pickersgill, R. W. (2007). *J. Biol. Chem.* **282**, 23957–23969.
Holliday, G. L., Thornton, J. M., Marquet, A., Smith, A. G., Rebeille, F., Mendel, R., Schubert, H. L., Lawrence, A. D. & Warren, M. J. (2007). *Nat. Prod. Rep.* **24**, 972–987.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Keller, J. P., Smith, P. M., Benach, J., Christendat, D., deTitta, G. T. & Hunt, J. F. (2002). *Structure*, **10**, 1475–1487.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C. & Warren, M. J. (1998). *Nature Struct. Biol.* **5**, 585–592.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vévodová, J., Graham, R. M., Raux, E., Schubert, H. L., Roper, D. I., Brindley, A. A., Scott, A. I., Roessner, C. A., Stamford, N. P. J., Stroupe, M. E., Getzoff, E. D., Warren, M. J. & Wilson, K. S. (2004). *J. Mol. Biol.* **344**, 419–433.
Wada, K., Harada, J., Yaeda, Y., Tamiaki, H., Oh-oka, H. & Fukuyama, K. (2007). *FEBS J.* **274**, 563–573.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.