JOURNAL OF BACTERIOLOGY, Oct. 1990, p. 5968–5979
0021-9193/90/105968-12$02.00/0

Vol. 172, No. 10

Nucleotide Sequence of a *Pseudomonas denitrificans* 5.4-Kilobase DNA Fragment Containing Five *cob* Genes and Identification of Structural Genes Encoding *S*-Adenosyl-L-Methionine: Uroporphyrinogen III Methyltransferase and Cobyrinic Acid *a*,*c*-Diamide Synthase

J. CROUZET,[1]* L. CAUCHOIS,[1] F. BLANCHE,[2] L. DEBUSSCHE,[2] D. THIBAUT,[2] M.-C. ROUYEZ,[1]†
S. RIGAULT,[1]‡ J.-F. MAYAUX,[1] AND B. CAMERON[1]

Laboratoire de Génétique, Institut des Biotechnologies,[1] and Département de Chimie Analytique,[2] Rhône-Poulenc Santé, Centre de Recherche de Vitry, 13 Quai Jules Guesde B.P. 14, 94 403 Vitry sur Seine Cédex, France

Received 6 March 1990/Accepted 11 July 1990

A 5.4-kilobase DNA fragment carrying *Pseudomonas denitrificans cob* genes has been sequenced. The nucleotide sequence and genetic analysis revealed that this fragment carries five different *cob* genes (*cobA* to *cobE*). Four of these genes present the characteristics of translationally coupled genes. *cobA* has been identified as the structural gene of *S*-adenosyl-L-methionine:uroporphyrinogen III methyltransferase (SUMT) because the encoded protein has the same NH_2 terminus and molecular weight as those determined for the purified SUMT. For the same reasons the *cobB* gene was shown to be the structural gene for cobyrinic acid *a*,*c*-diamide synthase. Genetic and biochemical data concerning *cobC* and *cobD* mutants suggest that the products of these genes are involved in the conversion of cobyric acid to cobinamide.

The cobalamin biosynthetic pathway probably involves 20 to 30 different enzymatic steps, consisting of (i) formation of uroporphyrinogen III (urogen III), which is the common intermediate for the synthesis of hemes, chlorophylls, cobalamins, F_{430}, and sirohemes; (ii) conversion of urogen III into cobyrinic acid, including the methylations at C-1, C-2, C-5, C-7, C-12, C-15, C-17, and C-20, the decarboxylation of the acetic side chain at C-12, the loss of C-20, and the introduction of cobalt; (iii) formation of cobinamide from cobyrinic acid by amidation of six of seven peripheral carboxylic groups, the seventh being amidated by (*R*)-1-amino-2-propanol; (iv) conversion of cobinamide into coenzyme B_{12} (for reviews on cobalamin synthesis, see references 3, 4, 18, 28, and 39). Only one enzymatic activity involved in the transformation of urogen III to coenzyme B_{12} has been purified (7), and no biosynthetic intermediate has been purified along the precorrin-3-to-cobyrinic-acid pathway. Cloned genes involved in cobalamin synthesis (*cob* genes) are valuable tools for the study of the coenzyme B_{12} biosynthesis at the biochemical and genetic levels. These genes should enable the identification of enzymatic activities and biosynthetic intermediates of the pathway and facilitate the understanding of the nature of biochemical and genetic regulation mechanisms operative in the *cob* regulon.

We have reported the cloning of at least 14 different genes, involved in cobalamin biosynthesis in *Pseudomonas denitrificans*, based on complementation data (9). Of these 14 genes, 12 are involved in the transformation of urogen III into cobinamide. The other two complement Cob mutants blocked in the conversion of cobinamide into cobalamin and are implicated in the last four steps of the cobalamin biosynthetic pathway (18). All 14 cloned genes are grouped on the *P. denitrificans* genome in four genomic regions, corresponding to complementation groups A, B, C, and D (9). In contrast, most of the *cob* genes in *Salmonella typhimurium* and *Bacillus megaterium* are clustered (23, 24, 45). We report the genetic analysis and nucleotide sequence of a 5.4-kilobase-pair (kb) fragment from complementation group C. Part of this fragment is carried by the previously described plasmid pXL190 (9). At least three genes were found on this fragment by complementation data (9). Plasmid pXL190 was shown to amplify *S*-adenosyl-L-methionine (SAM):urogen III methyltransferase (SUMT) when introduced in *P. denitrificans* (7). This key enzyme represents an important activity of the pathway since it is the first enzyme after the branch point of the cobalamin and heme pathways at urogen III. Previous biochemical studies (7) have shown that the enzyme is inhibited by its substrate at concentrations higher than 2 μM. In this study, the SUMT structural gene is identified on the 5.4-kb fragment together with four other *cob* genes. Although cobalamin is synthesized by a large number of bacteria, the cloning of *cob* genes has so far been reported for three organisms—*P. denitrificans* (9), *B. megaterium* (8), and *S. typhimurium* (1). This is the first report of a genetic analysis of *cob* genes at the molecular level.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are described in Table 1. For each of the plasmids constructed, a specific DNA fragment from the pXL157 insert (9) was cloned into either a derivative of RK2 (i.e., pRK290 to generate pXL723, or pRK404 to give pXL1630 and pXL1631) or a derivative of RSF1010 (i.e., pXL59 to lead to pXL302, pXL545, and pXL556, or pXL435 to generate pXL1397); for restriction sites, see Fig. 3. In addition, a spectinomycin resistance gene from pHP45Ω (37)

* Corresponding author.
† Present address: INSERM U. 152, Hôpital Cochin, 75674 Paris cédex 14, France.
‡ Present address: Résidence le Gilly, 73200 Gilly sur Isère, France.

TABLE 1. Bacterial strains and plasmids used

Bacterial strain, phage, or plasmid	Marker and replicon	Relevant properties	Reference or source
E. coli			
BS5548		F⁻ Δ*lacU169 rpsL thi cysG44 relA*	From P. Cossart and B. Gicquel-Sanzey, obtained by P1 transduction of the *cysG44* mutation from NF1400 (12) into MC4100 (10)
LE392		F⁻ *hsdR514 supE44 supF58 lacY1 galK2 galT22 metB1 trypR55* λ⁻	13
TG1		Δ(*lac-pro*) *thi supE hsdD5*/F′ *proAB lacI*^{q}*ZΔM15*	T. J. Gibson, Ph.D. thesis[a]
113-3 Cbl1		*metE* cannot convert cobinamide into cobalamin	9
MC1060		Δ(*lacIOPZYA*)*X74 galU galK strA2 hsdR*	11
P. denitrificans			
SC510		High cobalamin producer	9
SC510 Rif^r		Rif^r isolate (100 μg/ml) of SC510	7
SBL27 Rif^r		Rif^r isolate (100 μg/ml) of SBL27; SC510 derives from SBL27 by numerous mutagenesis	7
Plasmids			
pHP45Ω	Amp^r Sp^r ColE1		37
pRK2013	Km^r ColE1	Carries the *tra* genes of RK2	17
pRK2013::Tn*5*	Km^r ColE1	Carries the *tra* genes of RK2 and Tn*5*	9
pRK2013::Tn*5*Sp^r	Km^r ColE1	Carries the *tra* genes of RK2 and Tn*5*Sp^r	This study
pR751	Tp^r R751	Carries the *tra* genes of R751	25
pRK290	Tet^r RK2	Carries the Mob locus of RK2	15
pRK404	Tet^r RK2	Carries the Mob locus of RK2	14
pUC4KISS	Amp^r Km^r ColE1		Pharmacia, France S.A.
pXL59	Km^r RSF1010	Carries the Mob locus of RSF1010	9
pXL157	Km^r RSF1010	14-kb *P. denitrificans Sau*3AI fragment cloned into *Bam*HI site of pXL59	9
pXL190	Km^r RSF1010	4.2-kb *Cla*I-*Sau*3AI fragment from pXL157 cloned into *Bam*HI-*Cla*I sites of pXL59	9
pXL191	Km^r RSF1010	7.5-kb *Cla*I-*Sau*3AI fragment from pXL157 cloned into *Bam*HI-*Cla*I sites of pXL59	9
pXL435	Km^r RSF1010	Carries the Mob locus of RSF1010 multicloning site	9
Phage λ467		λ *b221 rex*::Tn*5 c*I857 *O*am29 *P*am80	13

[a] Thesis from University of Cambridge, Cambridge, England, 1984.

was cloned into the *Bam*HI site of pXL545 to give pXL545Ω. A kanamycin resistance gene from pUC4K-KISS (Pharmacia France S.A.) was inserted into the *Not*I restriction site (Fig. 1, position 747). The disrupted *Cla*I-*Rsa*I (*Cla*I and *Rsa*I at positions 0 and 1686, respectively) fragment was cloned into the multicloning site of pRK404 to generate pXL1630 or pXL1631 depending on the orientation of the kanamycin resistance gene transcription unit.

Media, bacteriological techniques, and chemicals. Bacteria were grown in LB (30), M9 (31), or PS4 (9) medium for routine culturing, complementation tests, and cobalamin production, respectively. The growth temperature was 37°C for *Escherichia coli* or 30°C for *P. denitrificans*, *Pseudomonas putida*, and *Agrobacterium tumefaciens*. The M9 medium was supplemented with L-cysteine (40 mg/liter). The antibiotic concentrations and culture conditions used for cobalamin synthesis in *P. denitrificans*, *P. putida*, and *A. tumefaciens* strains have been described previously (9). The cobalamin concentration was determined by the microbiological assay with the *E. coli* vitamin B_{12} auxotroph 113-3 Cbl1 as the indicator strain (9). To determine the concentration of the accumulated intermediates, *A. tumefaciens* or *P. putida* was cultured for 5 or 3 days, respectively, in a 250-ml Erlenmeyer flask containing 25 ml of PS4 medium, with a modified concentration of cobalt (5 μM for *A. tumefaciens* and 0.5 μM for *P. putida*), 2.5 μCi of ^{57}Co [specific activity, >4 mCi/μg of Co(II) chloride (Amersham France S.A.) in 0.1 M HCl] being added per flask. Assays for ^{57}Co-labeled corrinoids, including all corrinoids involved in the cobalamin pathway (i.e., intermediates between cobyrinic acid and coenzyme B_{12}) were monitored as described elsewhere (F. Blanche, D. Thibaut, M. Couder, and J. C. Muller, Anal. Chem., in press).

General methods. Recombinant DNA techniques used in this study include all the methods previously described (9). The procedures for mobilizing plasmid DNA from *E. coli* to other gram-negative bacteria and for complementing cobalamin-deficient mutants of *A. tumefaciens* and *P. putida* have already been described (9).

DNA sequencing. Overlapping DNA fragments of 2 to 3 kb, from the pXL191 insert, were generated by using appropriate restriction endonucleases and cloned in both orientation into M13 derivatives (M13mp18, M13mp19, M13tg130, or M13tg131 [27, 34, 43]). Progressive unidirectional deletions of each inserted DNA were created by using the procedure developed by Henikoff (21). Recombinants were transfected into strain TG1 to produce the desired single-stranded templates, which were purified as already described (*M13 Cloning and Sequencing Handbook*; Amersham France S.A.), except that after phenol extraction, DNA templates were extracted twice with ether before ethanol precipitation. Nucleotide sequences were determined unambiguously on both strands by the dideoxy-chain termination method of Sanger et al. (38) modified to use [α-^{35}S]dATP (>1,000

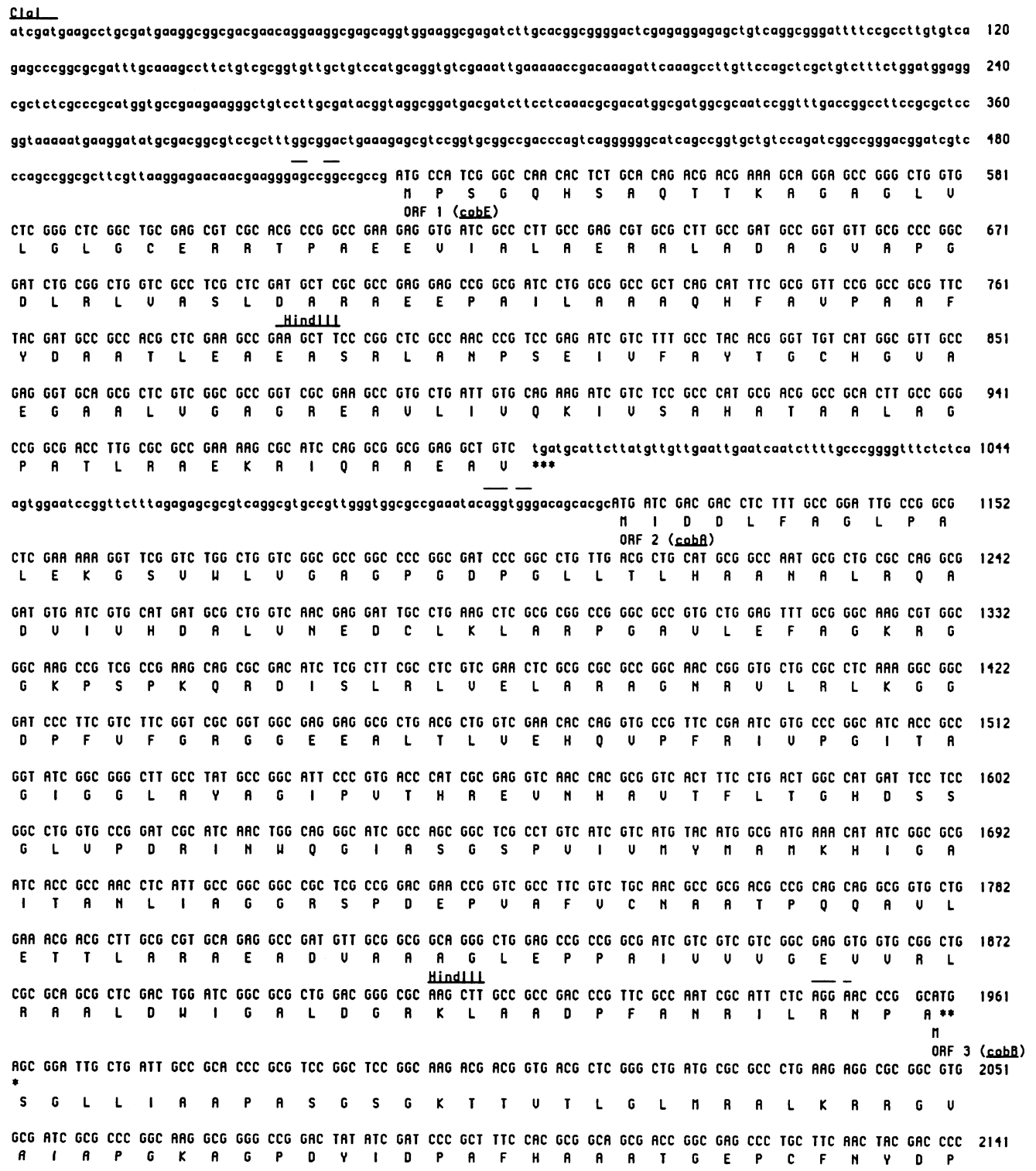

FIG. 1. Nucleotide sequence of the 5.4-kb fragment from complementation group C (9). The beginning of each ORF is indicated along with the predicted amino acid sequence of the gene. The positions of ORF1 to ORF5 potential ribosome-binding sites are indicated above the sequence by horizontal lines. Noncoding DNA is represented in lowercase letters.

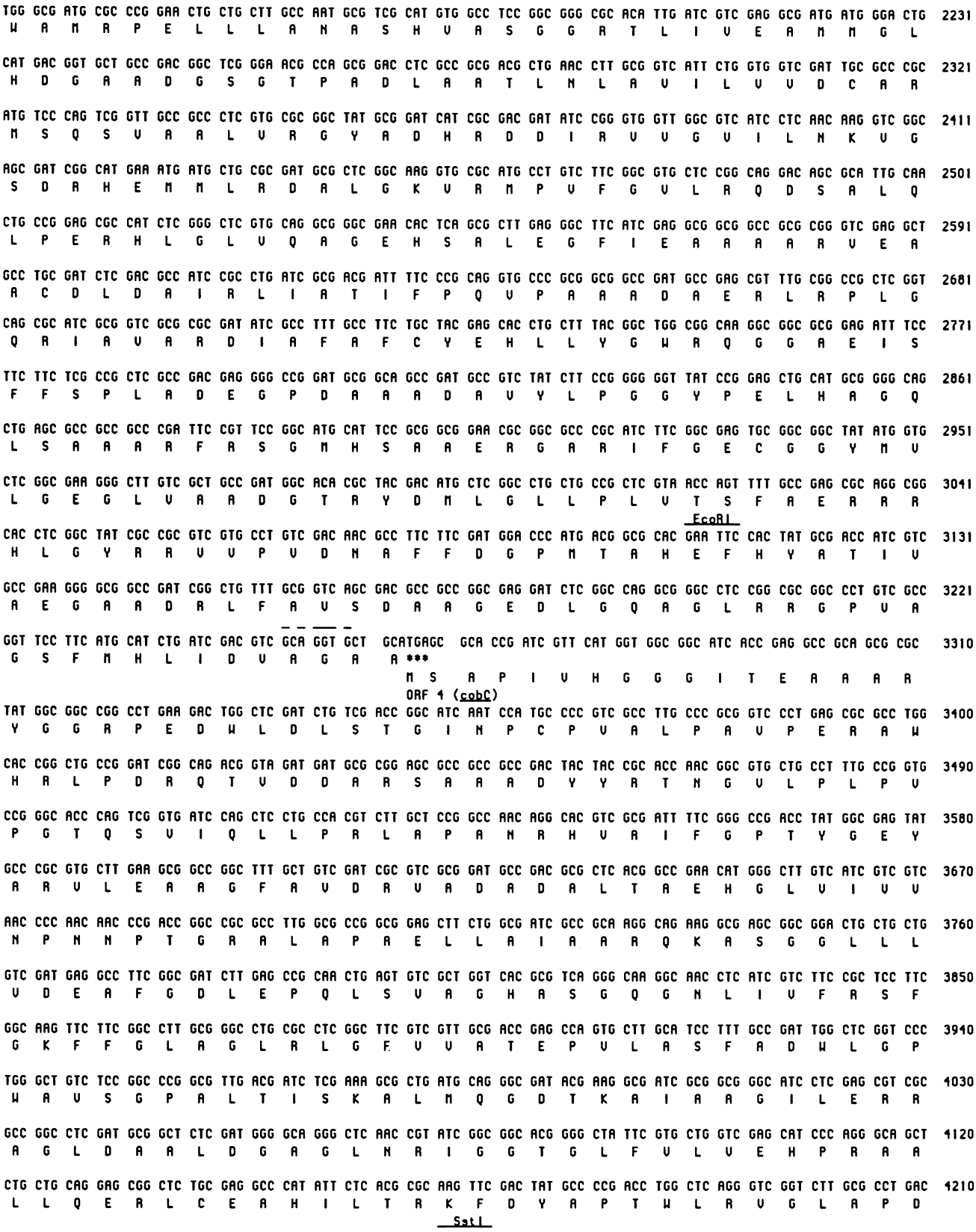

FIG. 1—*Continued.*

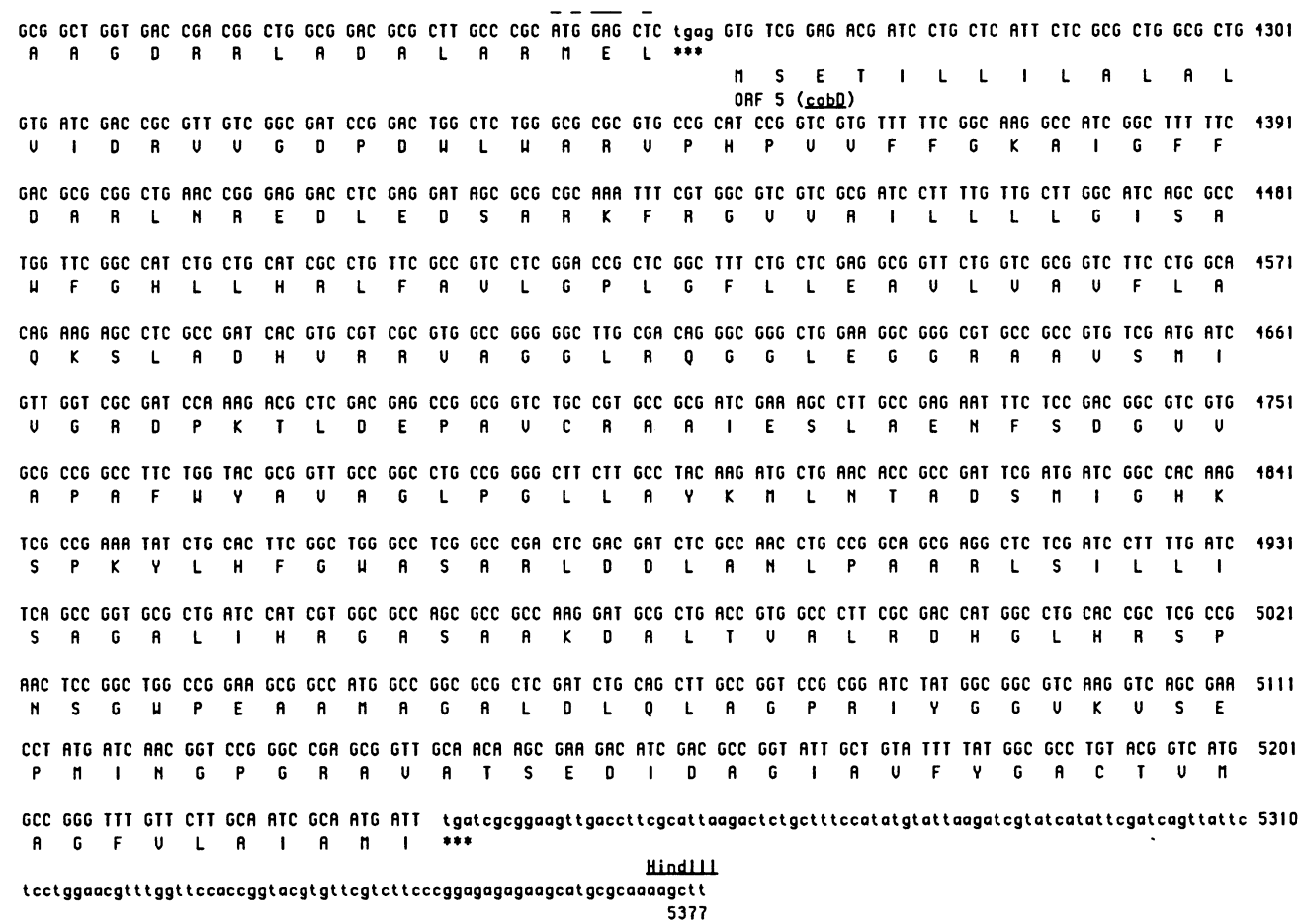

FIG. 1—*Continued.*

Ci/ml; Amersham France S.A.), and 7-deaza-dGTP, since this nucleotide helps to resolve compression areas which occur in sequencing DNA with a high G+C content (32). Reactions were analyzed on 0.4-mm 8% acrylamide sequencing gels, which were exposed to X-Omat AR film (Eastman Kodak Co., Rochester, N.Y.). DNA and protein sequences were examined by using Centre de Traitement Interuniversitaire d'Informatique à Orientation Biomédicale (CITI2) sequence analysis programs.

Mutagenesis. Three different mutagenesis strategies were used to identify correctly the *cob* genes on the sequenced DNA fragment. First, chromosomal random Tn*5* insertions were obtained on SBL27 Rif^r (7). Tn*5*Sp^r was provided by pRK2013::Tn*5*Sp^r. This plasmid was constructed by inserting a spectinomycin resistance gene from pHP45Ω into the *Bam*HI site of Tn*5* in pRK2013::Tn*5* (6, 9). Biparental mating (14) was carried out by mixing exponentially grown cultures of SBL27 Rif^r and MC1060 (pRK2013::Tn*5*Sp^r). Rif^r Sp^r conjugants were obtained at a low frequency of 10^{-8} clone per recipient cell after 5 days of incubation at 30°C. Plasmid DNAs from 12 clones were prepared, but no Km^r transformant was found after transformation into *E. coli*; genomic DNA was also prepared, digested with *Eco*RI, blotted after electrophoresis, and hybridized with pRK2013::Tn*5*Sp^r, labeled by nick translation with [α-^{32}P]dCTP (3,000 Ci/mmol; Amersham France S.A.). Under these conditions only one fragment hybridized with the probe (if the plasmid was still present, two *Eco*RI fragments of 6 and 40 kb would hybridize). Therefore, the Rif^r Sp^r conjugants have lost the introduced plasmid and integrated the Tn*5*Sp^r transposon into their genome. A total of 10,000 Rif^r Sp^r clones were screened for their cobalamin production as previously done for the *A. tumefaciens* Cob mutants (9), and 30 *cob*::Tn*5*Sp^r mutants were identified as producing less than 100-fold less cobalamin than SBL27 Rif^r.

In a second procedure, Tn*5* was inserted into the 2.3-kb *Eco*RI-*Hin*dIII fragment cloned in pXL723. Tn*5* mutagenesis was performed as described by de Bruijn and Lupski (13), using λ 467 to infect LE392(pXL723). After plasmid DNA purification, seven Tn*5* insertions were identified on the 2.3-kb *Eco*RI-*Hin*dIII insert of pXL723.

In a third method, a kanamycin resistance (Km^r) cartridge (from plasmid pUC4KISS) was introduced by marker exchange mutagenesis into *P. denitrificans* chromosome at the *Not*I restriction site (position 747) on the 5.4-kb fragment. Plasmids pXL1630 and pXL1631, which contain the Km^r cartridge on the disrupted *Cla*I-*Rsa*I insert, were each introduced into SC510 Rif^r by bacterial mating. The exoconjugants obtained were subjected to another biparental mating with MC1060(pR751), and the presence of pR751 and of the Km^r cartridge was selected. This selection favored the loss of the pXL1630 (or pXL1631) replicon and the marker exchange, because pXL1630, pXL1631, and pR751 belong to the same incompatibility group, IncP. It was verified by

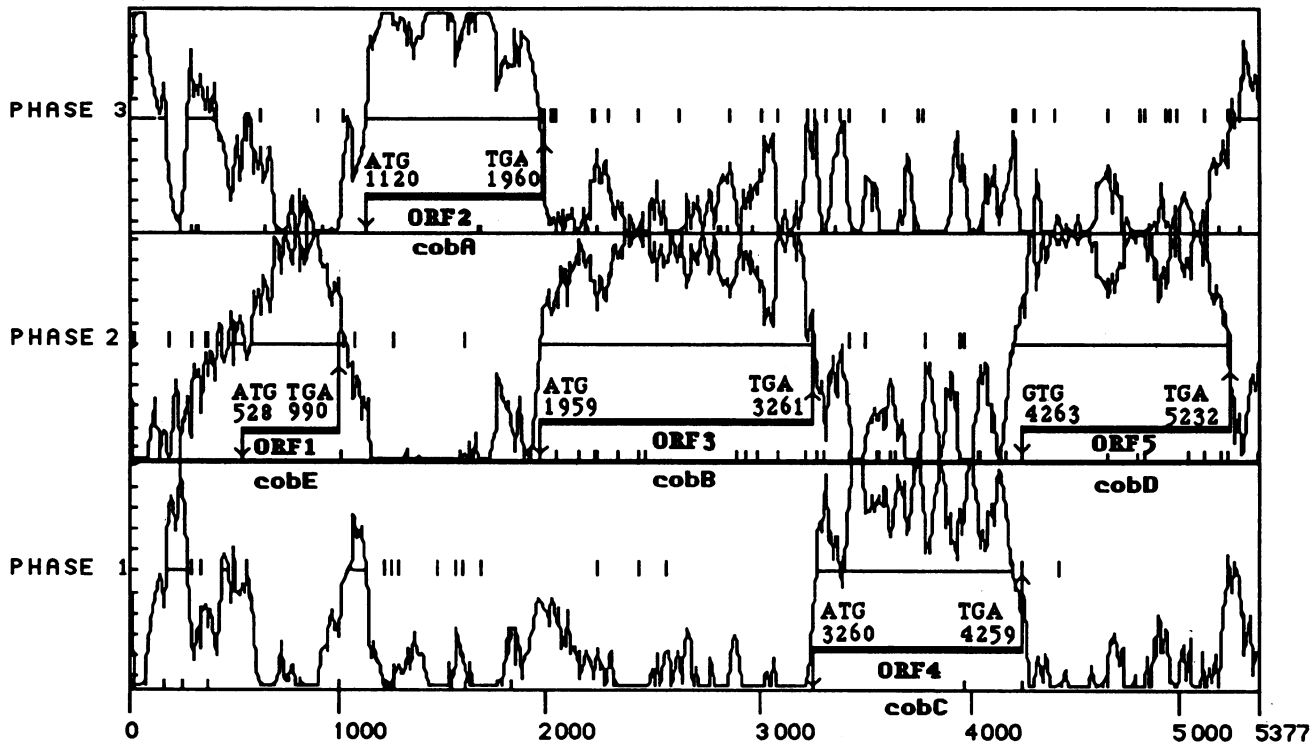

FIG. 2. Codon preference plot of the 5.4-kb DNA sequence as deduced from the codon usage method described in the text. The window size is 25 codons. This analysis is shown for the three frames of the strand going 5′→3′ from *Cla*I to *Hin*dIII sites. Vertical dashes across the half points represent stop codons, and those at the bottom of each box represent AUG codons. Horizontal lines across the half points correspond to 50% coding probability and allow us to identify the ORFs, whose respective positions of initiation and stop codons are indicated.

Southern blotting that the two strains generated were derived from SC510 Rif^r and contained a chromosomal kanamycin resistance gene at the *Not*I site.

Nucleotide sequence accession number. The sequence of the 5.4-kb fragment has been assigned the GenBank accession number M32223.

RESULTS

DNA sequence of the 5.4-kb fragment from complementation group C. A 5.4-kb fragment from pXL157 was sequenced. This fragment is bounded by a *Cla*I site at its left end and a *Hin*dIII site at its right end and has two internal *Hin*dIII sites. It was chosen because the amplification of one of its subfragments in *P. denitrificans* SC510 Rif^r increases SUMT activity by a factor of 50 (7). Complementation analysis of the 5.4-kb fragment showed that it carried at least three *cob* genes (9).

The sequence shown in Fig. 1 was analyzed by the program of Staden and McLachlan (42, 43), using codon preference to identify the coding sequences. A codon preference table was established from the codon usage in known sequenced *Pseudomonas* genes obtained from GenBank release 59. By this method, the identified open reading frames (ORFs) are those for which the codon usage is the closest to the one defined for other *Pseudomonas* genes. It was assumed that codon usage in *P. denitrificans* would not be substantially different from that in other *Pseudomonas* species. The principal assumptions made for this analysis were that all the genes should have similar codon preferences and that these preferences should be sufficiently strong to be used as a means of discrimination. The result of this analysis is shown in Fig. 2. Five ORFs (named ORF1 to ORF5) characterized by a high coding probability can be identified on the same strand (going 5′→3′ from the *Cla*I to *Hin*dIII sites) (Table 2); however, no convincing coding sequence could be detected on the other strand (data not shown). All five ORFs are preceded by potential ribosome-binding sequences having some homology to the consensus ribosome-binding site of *E. coli* (TAAGGAGGTG. . .5 to 9 base pairs . . .ATG) (40), as shown by the overlined bases in Fig. 1. Either the ribosome-binding site or the initiation codon overlaps with the termination codon of the preceding ORF in four of five ORFs (ORF2 to ORF5). It suggests that these ORFs may be translationally coupled (33). ORF1 and ORF2 are separated by an intergenic region of 130 base pairs. No sequences exhibiting the characteristics of a ρ-independent terminator (36) could be detected in the intergenic region between ORF1 and ORF2 or 3′ to ORF5. A GUG codon has been considered as the possible initiation codon of

TABLE 2. ORFs of the 5.3-kb fragment from complementation group C

ORF	Position of:		Mol wt of encoded polypeptide
	First codon	Last codon	
ORF1	528	989	15,500
ORF2	1120	1959	29,200
ORF3	1959	3260	45,600
ORF4	3260	4258	35,000
ORF5	4263	5231	34,100

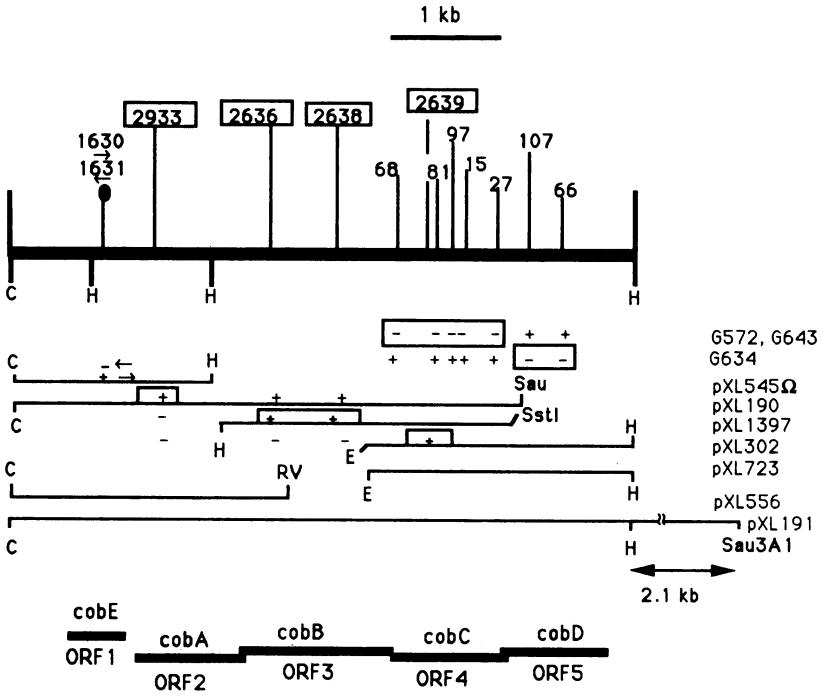

FIG. 3. Genetic analysis of the 5.4-kb fragment from complementation group C. Insertions 1630 and 1631, which correspond to Km^r cassette insertions into the chromosome of strain SC510 Rif^r, are shown with arrows beneath them. These arrows indicate the polarity of the transcription of the Km^r gene. The chromosomal Tn*5*Sp^r insertions in strain SBL27 Rif^r, whose number are boxed, are shown along with the Tn*5* insertions obtained on plasmid pXL723, numbered 15, 27, 66, 68, 81, 97, and 107. For the Tn*5* insertions, a minus or plus sign indicates that insertion does or does not, respectively, inactivate the complementation of mutants G572, G643, or G634. The inserts of plasmids studied for complementation of mutants are shown. Part of the pXL191 insert consisting of a 2.1-kb fragment is not represented to scale. Signs above each insert indicate that the plasmid does (+) or does not (−) complement Cob mutants aligned with the signs. The position of each ORF is indicated.

ORF5, since the next AUG codon is found nearly 400 base pairs downstream. The coding probability sharply rises to a high value just after the end of ORF4. The mean G+C content of this 5.4-kb fragment is 65.7%.

Genetic analysis of the 5.4-kb fragment. The 5.4-kb fragment was subjected to a genetic analysis to determine whether the five ORFs were *cob* genes. pXL723 was constructed by subcloning the 2.3-kb *Eco*RI-*Hin*dIII fragment into pRK290. Seven Tn*5* insertions were selected (named 15, 27, 66, 68, 81, 97, and 107) and mapped by standard procedures (Fig. 3). The mutated plasmids were then transferred into three Cob mutants, G572 (a *P. putida* Cob mutant), and G634 and G643 (*A. tumefaciens* Cob mutants), as previously described (9). Strains G572 and G643 were not complemented by insertions 15, 27, 68, 81, or 97, and G634 was not complemented by insertions 66 or 107 (Fig. 3). Two classes of insertions were identified, those that prevent the complementation of G572 and G643 and those that prevent the complementation of G634. For each class, the insertions were mapped in the same ORF previously identified by the nucleotide sequence. It was therefore concluded that ORF4 and ORF5 are two *cob* genes, named *cobC* and *cobD*, respectively.

A Tn*5*Sp^r was used for making transposon mutagenesis into *P. denitrificans* SBL27 Rif^r. SC510, derived by several mutagenesis steps from SBL27, produces more cobalamin. Of the Cob mutants obtained, four insertions, 2636, 2638, 2639, and 2933, were identified and precisely mapped by Southern blot analysis on the 5.4-kb fragment (Fig. 3). Insertion 2639 was found in ORF4 and complemented by pXL302, which contains *cobC* and *cobD*. Insertions 2636 and 2638, mapped in ORF3, were not complemented by pXL302 but were complemented by pXL1397, which contains ORF3 and ORF4. Therefore, it was concluded that ORF3 is a *cob* gene, *cobB*. Insertion 2933, mapped in ORF2, was complemented by plasmid pXL190 but not by plasmid pXL1397. Mutants for which Tn*5*Sp^r has been mapped into ORF3 and ORF4 were complemented by plasmid pXL1397. It was therefore concluded that insertion 2933 inactivates *cobA*, corresponding to ORF2.

The Km^r cassette from plasmid pUC4K was introduced into ORF1 on the genome of SC510 Rif^r, through marker exchange mutagenesis. The transcription of the Km^r gene from transposon Tn*903* is known to continue downstream of the gene, allowing the expression of genes at the 3′ end (2). Insertions 1630 and 1631 differ in the orientation of transcription of the Km^r gene (Fig. 3). The transcription of the resistance gene in insertion 1630 has the same polarity as ORF1 and all the *cob* genes. This insertion would therefore inactivate the expression of ORF1 and not disrupt totally the transcription of ORF2 to ORF5 if they are part of the same operon as ORF1. It follows that insertion 1631 would inactivate the expression of ORF1 and probably also the expression of ORF2 to ORF5. Both insertions 1630 and 1631 resulted in Cob$^-$ phenotypes with reduced cobalamin productions in PS4 medium of 0.7 and 0.1% of the SC510 Rif^r level, respectively. Plasmid pXL545Ω, which carries a 0.8-kb *Cla*I-*Hin*dIII fragment containing only a complete ORF1 (Fig. 3), nearly complemented insertion 1630, as strain SC510 Rif^r::1630 pXL545Ω produced 35% of the cobalamin level of SC510 Rif^r pXL545Ω. Plasmid pXL545Ω did not complement insertion 1631 (strain SC510 Rif^r::1631

pXL545Ω produced only 1% of the cobalamin produced by SC510 Rif^r). Insertion in ORF1 leads to a Cob$^-$ phenotype and can be nearly complemented in *trans* by a fragment carrying ORF1, only if the insertion does not exert a strong polar effect on downstream transcription through ORF2 to ORF5. These results indicate that ORF1 is most probably the *cobE* gene. Moreover, it is probable that all these *cob* genes are part of the same operon or at least that expression of *cobE* is linked to the expression of some downstream genes. Another possibility would be that the *cobA* promoter is disrupted by the *Not*I insertion.

Codon usage of the five *cob* genes is reported in Table 3. *cobA* to *cobD* show very similar codon usage with few observable differences. One difference is that ACU represents 20% of the utilized *cobA* threonine codons, whereas it is not used in the others. A second difference is that the CAA codon for glutamine is not used in *cobA* and *cobD* but is used in the other three genes. Other minor differences are not likely to be significant. *cobE* codons, however, seem more different than those used in the other four genes. For instance, some codons present in the other genes, such as AAU, AGC, and GAC, are totally absent in *cobE*. On the other hand, the UCU codon is present only in *cobE*. Such a difference is not explained and is less important than that between highly and poorly expressed genes in *E. coli* (20). The codon usage in genes *cobA* to *cobD* is similar to the codon usage in *Pseudomonas aeruginosa* (44), with some exceptions: (i) the most frequently used threonine codons are ACC and ACG, unlike ACC in *P. aeruginosa*; (ii) the two tyrosine codons are almost equally utilized in *P. denitrificans*, in contrast to *P. aeruginosa*, where UAC is used most frequently; and (iii) the histidine codon most often used is CAU in *P. denitrificans* and CAC in *P. aeruginosa*. The observed differences concern few codons, and only minor differences were found with the codon usage of the *Pseudomonas* genes (data not shown). This explains why it was a good choice to search for coding sequences in the fragment by using the program of Staden and McLachlan (41, 42) and a codon usage table obtained from the known *Pseudomonas* genes.

***cobA* is the structural gene encoding SUMT activity.** Since plasmid pXL190 (the insert of which is part of the 5.4-kb fragment) was shown to lead to a 50-fold increase in SUMT activity in strain SC510 Rif^r (7), it was likely that one of the identified *cob* gene was the structural gene for SUMT. NH_2-terminal sequencing of the purified SUMT was performed, and the sequence of the first 10 amino acids was found to be Met-Ile-Asp-Asp-Leu-Phe-Ala-Gly-Leu-Pro. This sequence is identical to the first 10 amino acids of the protein encoded by the *cobA* gene. *P. denitrificans* SUMT was shown to be a homodimeric enzyme of M_r 30,000 ± 1,000 (7). The *cobA* gene is predicted to encode a protein of M_r 29,200, in good agreement with the biochemical estimate. The amino acid composition of *P. denitrificans* SUMT was determined (7) and agreed well with the composition of the CobA protein (data not shown). The NH_2-terminal methionine was not removed by methionine aminopeptidase. This is in agreement with the rules of methionine excision found for *E. coli* proteins (5, 22), as the penultimate amino acid of the coding sequence is isoleucine. All of these results demonstrate that *cobA* is the structural gene for SUMT.

No significant homology with *P. denitrificans* SUMT was found in protein data bases. However, the protein encoded by the *E. coli cysG* (T. Peakman, J. Crouzet, J.-F. Mayaux, S. Busby, S. Mohan, R. Nicholson and J. Cole, Eur. J. Biochem., in press) shows striking homology to CobA (Fig.

TABLE 3. Codon usage analysis of the *cobA* to *cobE* genes

Amino acid	Codon	% Codon usage in:				
		cobA	*cobB*	*cobC*	*cobD*	*cobE*
F	UUU	25.0	17.6	18.2	42.9	33.3
F	UUC	75.0	82.4	81.8	57.1	66.7
L	UUA	0.0	0.0	0.0	0.0	0.0
L	UUG	6.5	8.9	9.3	8.9	6.7
L	CUU	12.9	13.3	20.9	20.0	20.0
L	CUC	29.0	37.8	32.6	28.9	40.0
L	CUA	0.0	0.0	2.3	0.0	0.0
L	CUG	51.6	40.0	34.9	42.2	33.3
I	AUU	18.8	23.5	15.4	16.7	16.7
I	AUC	81.3	76.5	84.6	83.3	83.3
I	AUA	0.0	0.0	0.0	0.0	0.0
M	AUG	100.0	100.0	100.0	100.0	100.0
V	GUU	3.7	6.5	8.3	21.4	25.0
V	GUC	55.6	51.6	62.5	42.9	41.7
V	GUA	0.0	3.2	4.2	3.6	0.0
V	GUG	40.7	38.7	25.0	32.1	33.3
S	UCU	0.0	0.0	0.0	0.0	16.7
S	UCC	25.0	42.1	27.3	11.8	50.0
S	UCA	0.0	5.3	9.1	5.9	0.0
S	UCG	62.5	21.1	27.3	41.2	33.3
P	CCU	5.3	14.3	17.4	6.7	0.0
P	CCC	26.3	33.3	21.7	0.0	14.3
P	CCA	0.0	4.8	13.0	6.7	14.3
P	CCG	68.4	47.6	47.8	86.7	71.4
T	ACU	20.0	0.0	0.0	0.0	0.0
T	ACC	30.0	25.0	57.1	33.3	14.3
T	ACA	0.0	16.7	0.0	16.7	0.0
T	ACG	50.0	58.3	42.9	50.0	85.7
A	GCU	0.0	6.6	12.5	2.0	9.8
A	GCC	41.5	40.8	33.9	52.9	53.7
A	GCA	9.8	7.9	10.7	9.8	9.8
A	GCG	48.8	44.7	42.9	35.3	26.8
Y	UAU	50.0	63.6	66.7	60.0	0.0
Y	UAC	50.0	36.4	33.3	40.0	100.0
*[a]	UAA	0.0	0.0	0.0	0.0	0.0
*	UAG	0.0	0.0	0.0	0.0	0.0
H	CAU	71.4	57.1	57.1	55.6	75.0
H	CAC	28.6	42.9	42.9	44.4	25.0
Q	CAA	0.0	22.2	25.0	0.0	20.0
Q	CAG	100.0	77.8	75.0	100.0	80.0
N	AAU	22.2	20.0	12.5	16.7	0.0
N	AAC	77.8	80.0	87.5	83.3	100.0
K	AAA	37.5	0.0	20.0	22.2	33.3
K	AAG	62.5	100.0	80.0	77.8	66.7
D	GAU	53.3	53.6	63.2	44.4	100.0
D	GAC	46.7	46.4	36.8	55.6	0.0
E	GAA	38.5	33.3	20.0	45.5	38.5
E	GAG	61.5	66.7	80.0	54.5	61.5
C	UGU	0.0	0.0	0.0	50.0	50.0
C	UGC	100.0	100.0	100.0	50.0	50.0
*	UGA	0.0	0.0	0.0	0.0	0.0
W	UGG	0.0	100.0	100.0	100.0	0.0
R	CGU	10.5	5.9	12.0	25.0	22.2
R	CGC	63.2	58.8	44.0	40.0	55.6
R	CGA	5.3	2.9	4.0	15.0	0.0
R	CGG	15.8	26.5	24.0	15.0	22.2
S	AGU	0.0	5.3	9.1	0.0	0.0
S	AGC	12.5	26.3	27.3	41.2	0.0
R	AGA	0.0	0.0	0.0	0.0	0.0
R	AGG	5.3	5.9	16.0	5.0	0.0
G	GGU	12.5	10.4	14.3	15.2	30.8
G	GGC	75.0	62.5	65.7	66.7	38.5
G	GGA	3.1	8.3	2.9	3.0	7.7
G	GGG	9.4	8.8	17.1	15.2	23.1

[a] *, Translation termination codon.

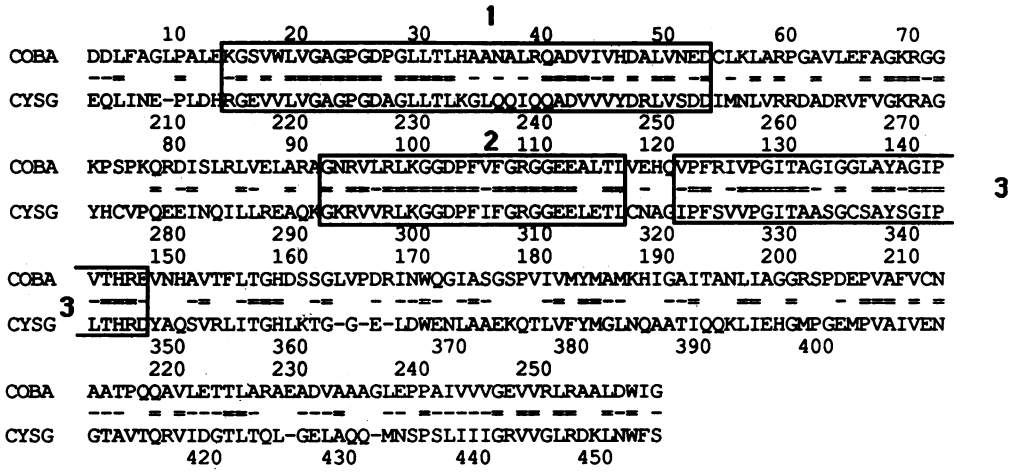

FIG. 4. Alignment of amino acid sequences of *P. denitrificans* SUMT and *E. coli* CysG. The comparison goes from positions 3 to 259 for CobA and from 204 to 454 for CysG. Similarities are indicated below the sequences as follows: =, same amino acid; –, amino acids belonging to the same group (hydroxyl/small aliphatic: A, G, S, T; acid and acid amide: N, D, E, Q; basic: H, R, K; aliphatic: M, I, L, V; or aromatic: F, Y, W). The proteins present more than 41% of strict homology in the alignment indicated. Dashes within the sequence represent gaps in the alignment.

4). More than 41% of strict homology is observed between the two proteins. This homology occurs on most of the SUMT sequence (from residues 3 to 259) and corresponds to the carboxy terminus of the *E. coli* protein (from residues 204 to 454). Interestingly, there are three domains of higher homology between CobA and CysG (Fig. 4). It is probable that these domains are involved in the active site of the enzyme. The *S. typhimurium cysG* gene is believed to catalyze at least the same reaction as SUMT (23), since *cysG* mutants are blocked in both cobalamin and siroheme synthesis. These pathways are supposed to share the conversion of urogen III into precorrin-2, which is catalyzed by SUMT (7). Thus the homology detected between *P. denitrificans* SUMT and *E. coli* CysG might be specific for SUMT activity. It is noteworthy that *cobA* is clustered with *cob* genes in *P. denitrificans*, whereas *cysG* is not found near the *cobI* to *cobIII* region in *S. typhimurium* (23).

Genetic studies of *cobA*. Plasmids pXL545 and pXL556 (Fig. 3) were transformed into the *E. coli cysG* strain B5548. There was no complementation for the *cysG* mutation on minimal M9 medium, whereas plasmid pHM1 (29), which carries the *E. coli cysG* gene, did complement the mutation. These results could be interpreted by either an insufficient level of expression of *cobA* for complementation or the inability of *cobA* to complement the *E. coli cysG* mutant. The most likely explanation for the latter would be that the CysG protein carries two domains: the amino-terminal domain, which catalyzes the conversion of precorrin-2 into siroheme (consisting of an oxydation and Fe^{2+} chelation), and the CobA homologous carboxy-terminal domain, responsible for SUMT activity. We have reported the isolation of *P. denitrificans* Cob mutants with reduced SUMT activity (i.e., less than 2% of the parent level) which continue to grow on minimal medium without cysteine (7). This suggests that CobA does not support the transformation of precorrin-2 into siroheme.

Identification of the structural gene encoding cobyrinic acid *a,c*-diamide synthase. The purification of *P. denitrificans* cobyrinic acid *a,c*-diamide synthase, a homodimer of M_r 2 × 45,000, will be described elsewhere (L. Debussche, D. Thibaut, B. Cameron, J. Crouzet, and F. Blanche, submitted for publication). Since the substrate for this reaction can be either cobyrinic acid or hydrogenobrynic acid, the enzyme must be active on both the corrin and the descobalto corrin nuclei. This activity is responsible for the amidation of carboxylic groups at positions *a* and *c*. NH_2 groups are provided by glutamine, and one molecule of ATP is hydrolyzed for each amidation. The scheme of the reaction, using cobyrinic acid as the substrate, is shown in Fig. 5. The amplified enzymatic activity has been purified from strain SC510 Rif^r pXL191 (Debussche et al., submitted). Plasmid pXL191 carries the 5.4-kb fragment (Fig. 3). The sequence of the NH_2-terminal first 15 amino acids of the purified cobyrinic acid *a,c*-diamide synthase has been determined and matches the amino-terminal sequence of CobB (Fig. 1), except that methionine, encoded by the initiation codon, has been removed. The first amino acid of this NH_2-terminal sequence is a serine, which is among the amino acids proposed to favor methionine excision (5, 22). The *cobB* gene is predicted to encode a protein of molecular weight 45,600, in good agreement with the size of the purified monomer. It is concluded that *cobB* is the structural gene for cobyrinic acid *a,c*-diamide synthase.

Biochemical studies on *cobC* and *cobD* mutants. Of the five genes, *cobA* and *cobB* have been determined by genetic analysis and nucleotide sequencing to encode identified enzymatic activities. The enzymatic activities of the gene products of *cobC*, *cobD*, and *cobE* remain to be elucidated. One approach to localizing these activities along the cobalamin pathway has been to study the intracellular accumulation of biosynthetic intermediates in mutants complemented by *cobC* and *cobD*. Accumulation of intermediates in a mutant is most probably caused by blocking at a step which utilizes the accumulated intermediates as precursors or substrates. Since four of the *cob* genes, defined by the nucleotide sequence and the genetic study, seem to be translationally coupled, it is likely that a transposon insertion in the upstream region (for instance, a *cobA* insertion) would have a polar effect on the expression of the following genes. Therefore, we chose to study mutants obtained after treatment with *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine. *A. tumefaciens* Cob mutants G634 and G643 and a *P. putida* Cob G572 mutant were studied for their intracellular content of corrinoids. Since these mutants were *A. tumefaciens* and

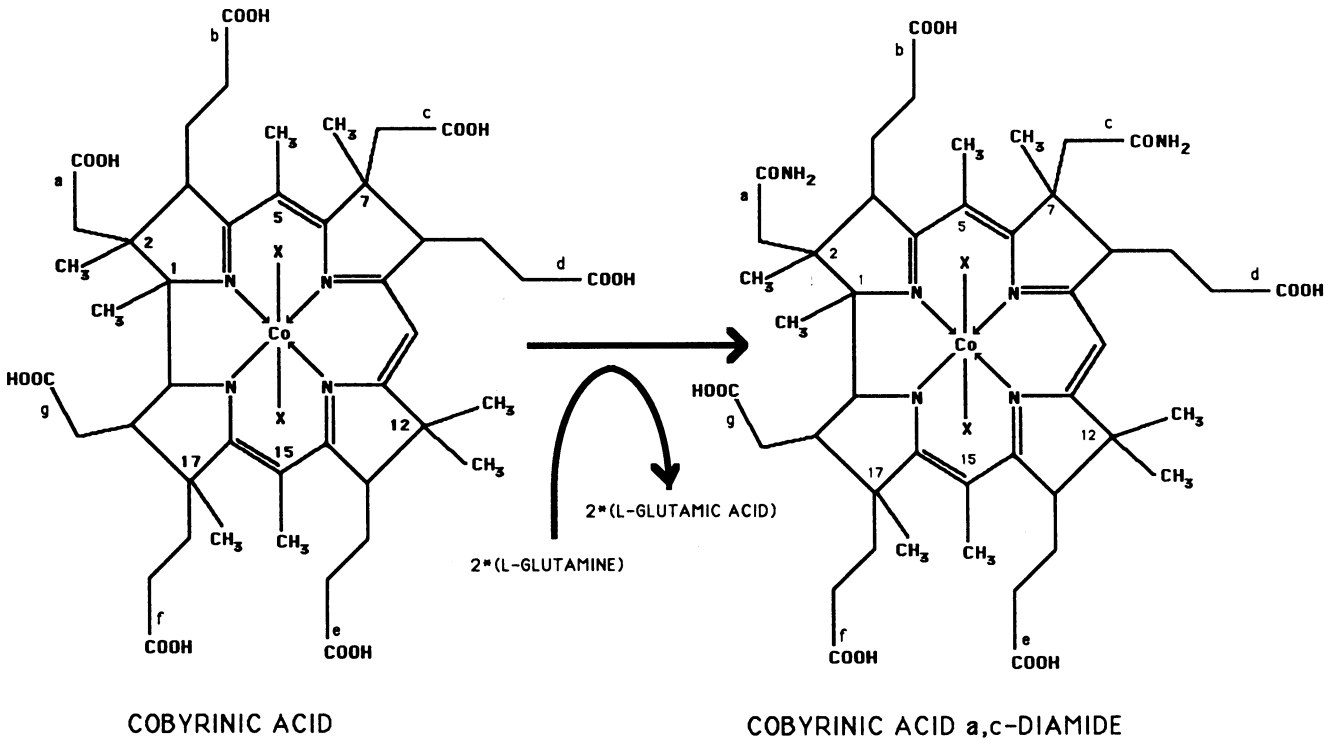

FIG. 5. Reaction catalyzed by cobyrinic acid *a,c*-diamide synthase.

P. putida strains complemented by *P. denitrificans cob* genes, it is likely that they share the same pathway for cobalamin synthesis. This would imply that the Cob mutants are blocked in the step catalyzed by the product of the complementing gene. Studies of these mutants should give indications about the activity of the protein encoded by the complementing gene. Strain G643 carries a mutation in *cobC* as described above. This partial mutant accumulates cobyric acid and cobyrinic acid pentaamide (Table 4) and synthesizes 14% of the coenzyme B_{12} produced by the parent strain. Strain G572 (mutated in *cobC*) also accumulates cobyric acid. These mutants are blocked before cobinamide (7) and accumulate the intermediate that precedes cobinamide. It is therefore likely that the CobC protein is involved in the transformation of cobyric acid into cobinamide. Strain G634, an *A. tumefaciens* Cob mutant, is blocked before cobinamide and is complemented by *cobD*. This mutant accumulates cobyric acid and must also be blocked after cobyric acid. It is improbable that the results obtained with mutants G572 and G643 can be explained by a polar effect on the *cobD* gene, as Tn*5* insertions in *cobC* do not inactivate the complementation of a *cobD* mutant by a low-copy-number plasmid (an RK2 derivative). The *cobD* gene, like *cobC*, codes for a protein that is involved in the transformation of cobyric acid into cobinamide. At least two polypeptides are involved in this reaction, in which the carboxylic group at position *f* of the corrin nucleus is amidated with (*R*)-1-amino-2-propanol. This suggests that the enzyme responsible for this activity might be a heterodimer. The phenotype of one Cob mutant (isolated from among 10,000 SBL27 Rif^r::Tn*5*Sp^r isolates) was suppressed by the addition of 40 mM (*R*)-1-amino-2-propanol to PS4 medium. This mutant is believed to be blocked in the synthesis of (*R*)-1-amino-2-propanol. The *Eco*RI genomic fragment, into which the Tn*5*Sp^r has inserted, was cloned and used as a probe against the previously described plasmids (pXL151 to pXL154, pXL156 to pXL161 and pXL519), which carry all the *P. denitrificans* cloned *cob* genes (9). No hybridization was observed, suggesting that the gene in which Tn*5*Sp^r has been inserted cannot be found among the previously cloned *cob* genes. It can be deduced that CobC and CobD are implicated in the transformation of cobyric acid into cobinamide at the level of the amidation of the *f* carboxylic group, and not in the synthesis of the aminopropanol moiety. We assume that the adenosylation of the corrinoids occurs before the cobyric acid since, so far, all the corrinoid intermediates isolated from *P. denitrificans* cultures have been found in their 5′-deoxy-5′-adenosyl forms (F. Blanche, unpublished results).

Mutant SC510 Rif^r::1630 was not studied for intermediate accumulation since the insertion might show polar effects on the expression of the downstream genes, even if such a mutant is partially complemented by pXL545Ω (up to 35% of the parent strain). Polar effects would probably lead to the

TABLE 4. Intracellular corrinoid accumulation into strains of *A. tumefaciens* C58-C9 Rif^r Nal^r, *P. putida* KT2440 Rif^r Nal^r, and their Cob mutants[a]

Mutant	Complemented by *cob* gene	% Corrinoid accumulation
C58-C9 Rif^r Nal^r		CNB_{12}, 100
G643	*cobC*	Cobyrinic acid pentaamide, 1.5; cobyric acid, 11; CNB_{12}, 14
G634	*cobD*	Cobyric acid, 15; CNB_{12}, 4.1
KT2440 Rif^r Nal^r		CNB_{12}, 100
G572	*cobC*	Cobyric acid, 5.2

[a] For each parent strain the level of the accumulated intermediate is standardized to 100% of the cyanocobalamin synthetized by the parent strain. The cyanoforms of the intermediates are assayed, CNB_{12}:cyanocobalamin.

accumulation of intermediates not directly due to *cobE* inactivation, which would lead to a false interpretation.

DISCUSSION

This study allowed us to identify five *P. denitrificans cob* genes, *cobA*, *cobB*, *cobC*, *cobD*, and *cobE*. Four of these genes show characteristics of translationally coupled genes; however, an effective translational coupling remains to be established. For instance, a promoter could be present in an ORF and allow expression of the next ORF. Such an organization of genes implicated in the same pathway has already been reported. For example, four genes in *E. coli*, *bioB*, *bioF*, *bioC*, and *bioD*, are involved in biotin synthesis and present the same arrangement (35). The genes for purine synthesis in *Bacillus subtilis* are clustered and organized into three groups of overlapping genes belonging to the same operon (16).

For the CobA to CobE proteins, no sequence homology was found with the proteins in GenBank by using the program of Kanehisa (26). The function of the protein was investigated by biochemical analysis. The enzymatic activity of the gene product, or the likely step in the cobalamin pathway at which the encoded protein may act, has been determined for all genes except *cobE*. *cobA* is the structural gene for SUMT, which is a key enzyme in coenzyme B_{12} synthesis, since it is the first enzyme after the branch point from heme synthesis. The CobA protein shows extensive homology with the *E. coli* CysG protein. Three domains are very well conserved between these two proteins. This might indicate portions of the enzymes which are important for the methyltransferase activity. Enzymatic studies should confirm whether CobA, like CysG, can catalyze the transformation of precorrin-2 into sirohemes, assuming that this reaction is enzymatically catalyzed in the cell. Metabolic pathways are regulated mostly by the activity of the first enzyme of the pathway. This activity can be modulated by a low level of expression or by a feedback regulation mechanism that represses enzymatic activity. SUMT and δ-aminolevulinic acid synthase (the first enzyme of the common pathway for heme and coenzyme B_{12} synthesis) are good candidates for such regulations of the cobalamin pathway. In certain organisms, δ-aminolevulinic acid synthetase is inhibited by heme accumulation (19). SUMT activity should determine the metabolic flux entering the cobalamin pathway from the total flux upstream of urogen III. It is already known that *P. denitrificans* SUMT activity is inhibited in vitro by urogen III at concentrations above 2 μM (7).

Proteins encoded by genes *cobB*, *cobC*, and *cobD* are involved in late reactions in the pathway. CobB catalyzes the *a*,*c*-amidation of cobyrinic acid as well as of hydrogenobyrinic acid. CobC and CobD are probably implicated in the transformation of cobyric acid into cobinamide. Enzymatic studies should confirm whether CobC and CobD are involved in the amidation of the *f* carboxylic group of cobyric acid, as suggested by this work. CobA to CobD are involved in different steps dispersed along the biosynthetic pathway. Why are these genes so closely linked in the same operon? This question remains to be answered. The clustering of *cob* genes might reflect the necessity of these genes to be expressed at the same level. It is difficult to imagine that enzymes in this pathway are not regulated in the same way. They should share the same kind of regulation, and only the expression level should vary from operon to operon. It would be interesting to test this hypothesis by studying the various expression levels of these genes.

ACKNOWLEDGMENTS

We thank J. Lunel, P. E. Bost, and J.-C. Brunie for their support during this work. We express our gratitude to S. Lévy-Schil for critical reading of the manuscript and helpful discussions. We also thank M. Couder and L. Maton for excellent technical assistance, T. Ciora for part of the sequence of the *cobA* gene, and the Centre de Traitement Interuniversitaire d'Informatique à Orientation Biomédicale for nucleic and protein sequences analysis programs. B. Gicquel-Sanzey and P. Cossart, F. de Bruijn, G. Ditta, and H. Krish are acknowledged for their gifts of strains, plasmids, or phages.

LITERATURE CITED

1. **Anderson, D. I., and J. R. Roth.** 1989. Mutations affecting regulation of cobinamide biosynthesis in *Salmonella typhimurium*. J. Bacteriol. **171**:6726–6733.
2. **Bagdasarian, M., R. Lurz, B. Rückert, F. C. Franklin, M. M. Bagdasarian, J. Frey, and K. Timmis.** 1981. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host vector system for gene cloning in Pseudomonas. Gene **16**:237–247.
3. **Battersby, A. R., C. J. R. Fookes, G. W. J. Matcham, and E. MacDonald.** 1980. Biosynthesis of the pigments of life: formation of the macrocycle. Nature (London) **285**:17–21.
4. **Battersby, A. R., and E. MacDonald.** 1982. Biosynthesis of the corrin macrocycle. p. 107–144. *In* D. Dolphin (ed.), B12, vol. 1. John Wiley & Sons, Inc., New York.
5. **Ben-Bassat, A., K. Bauer, C. Sy, K. Mijamo, A. Boosman, and S. Chang.** 1987. Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. J. Bacteriol. **169**:751–757.
6. **Berg, D. E., and C. M. Berg.** 1983. The prokaryotic transposable element Tn*5*. Bio/Technology **1**:417–435.
7. **Blanche, F., L. Debussche, D. Thibaut, J. Crouzet, and B. Cameron.** 1989. Purification and characterization of *S*-adenosyl-L-methionine:uroporphyrinogen III methyltransferase from *Pseudomonas denitrificans*. J. Bacteriol. **171**:4222–4231.
8. **Brey, R. N., C. D. B. Banner, and J. B. Wolf.** 1986. Cloning of multiple genes involved with cobalamin (vitamin B_{12}) biosynthesis in *Bacillus megaterium*. J. Bacteriol. **167**:623–630.
9. **Cameron, B., K. Briggs, S. Pridmore, G. Brefort, and J. Crouzet.** 1989. Cloning and analysis of genes involved in coenzyme B_{12} biosynthesis in *Pseudomonas denitrificans*. J. Bacteriol. **171**:547–557.
10. **Casadaban, M. J., and S. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA **76**:4530–4533.
11. **Casadaban, M. J., A. Martinez-Arias, S. T. Shapira, and J. Chou.** 1983. β-Galactosidase gene fusion for analysing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–308.
12. **Cossart, P., and B. Gicquel-Sanzey.** 1982. Cloning and sequence of the *crp* gene of *E. coli*. Nucleic Acids Res. **10**:1363–1378.
13. **De Bruijn, F. J., and J. R. Lupski.** 1984. The use of transposon Tn*5* mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids—a review. Gene **27**:131–149.
14. **Ditta, G., T. Schmidhauser, E. Yakobson, P. Lu, X.-W. Liang, D. R. Finlay, D. Guiney, and D. R. Helinski.** 1985. Plasmids related to the broad host range vector pRK290, useful for gene cloning and for monitoring gene expression. Plasmid **13**:149–154.
15. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium melitoti*. Proc. Natl. Acad. Sci. USA **77**:7347–7351.
16. **Ebbole, D. J., and H. Zalkin.** 1987. Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for *de novo* purine nucleotide synthesis. J. Biol. Chem. **262**:8274–8287.
17. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA

76:1648–1652.
18. **Friedmann, H. C.** 1975. Biosynthesis of corrinoids, p. 75–110. *In* B. M. Babior (ed.), Cobalamin. John Wiley & Sons, Inc., New York.
19. **Granick, S., and S. I. Beale.** 1978. Hemes, chlorophylls and related compounds: biosynthesis and metabolic regulation. Adv. Enzymol. Relat. Areas Mol. Biol. **46**:33–203.
20. **Grosjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199–209.
21. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
22. **Hirel, P.-H., J.-M. Schmitter, P. Dessen, and S. Blanquet.** 1989. Extent of N-terminal methionine excision within *E. coli* proteins is governed by the side chain of the penultimate aminoacids. Proc. Natl. Acad. Sci. USA **86**:8247–8251.
23. **Jeter, R. M., B. M. Olivera, and J. R. Roth.** 1984. *Salmonella typhimurium* synthesizes cobalamin (vitamin B_{12}) de novo under anaerobic growth conditions. J. Bacteriol. **159**:206–213.
24. **Jeter, R. M., and J. R. Roth.** 1987. Cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium*. J. Bacteriol. **169**:3189–3198.
25. **Jobanputra, R. S., and N. Datta.** 1974. Trimethoprim R factors in enterobacteria from clinical specimens. J. Med. Microbiol. **7**:169–177.
26. **Kanehisa, M.** 1984. Use of statistical criteria for screening potential homologies in nucleic acids sequences. Nucleic Acids Res. **12**:203–215.
27. **Kieny, M. P., R. Lathe, and J. P. Lecocq.** 1983. New versatile cloning vectors based on bacteriophage M13. Gene **26**:91–99.
28. **Leeper, F. J.** 1989. The biosynthesis of porphyrins, chlorophylls and vitamin B_{12}. Nat. Prod. Rep. **6**:171–203.
29. **Macdonald, H., and J. Cole.** 1985. Molecular cloning and functional analysis of *cysG* and *nirB* genes of *E. coli* K12, two closely-linked genes required for NADH-dependent reductase activity. Mol. Gen. Genet. **200**:328–334.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Mizusawa, S., S. Nishimura, and F. Sela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
33. **Normark, S., S. Bergtröm, T. Edlund, T. Grundström, B. Jaurin, F. Lindberg, and O. Olsson.** 1983. Overlapping genes. Annu. Rev. Genet. **17**:499–525.
34. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
35. **Otsuka, A. J., M. R. Buoncristiani, P. K. Howard, J. Flamm, C. Johnson, R. Yamamoto, K. Uchida, C. Cook, J. Ruppert, and J. Matsuzaki.** 1988. The *Escherichia coli* biotin biosynthetic enzyme sequences predicted from the nucleotide sequence of the *bio* operon. J. Biol. Chem. **263**:19577–19585.
36. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
37. **Prentki, P., and H. M. Krisch.** 1984. *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene **29**:303–313.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Scott, A. I., N. E. Mackenzie, P. J. Santander, P. E. Fagerness, G. Muller, E. Schneider, R. Seldmeier, and G. Worner.** 1984. Biosynthesis of vitamin B_{12}—timing of the methylation steps between uro'gen III and cobyrinic acid. Bioorg. Chem. **12**:356–352.
40. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
41. **Staden, R.** 1984. Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. Nucleic Acids Res. **12**:551–567.
42. **Staden, R., and A. D. McLachlan.** 1982. Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res. **10**:141–156.
43. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
44. **West, S. E. H., and B. Iglewsi.** 1988. Codon usage in *Pseudomonas aeruginosa*. Nucleic Acids Res. **16**:9323–9335.
45. **Wolf, J. B., and R. N. Brey.** 1986. Isolation and genetic characterization of *Bacillus megaterium* cobalamin biosynthesis-deficient mutants. J. Bacteriol. **166**:51–58.