JOURNAL OF BACTERIOLOGY, Apr. 1994, p. 1957–1968
0021-9193/94/$04.00+0

Vol. 176, No. 7

Genetics of the Serine Cycle in *Methylobacterium extorquens* AM1: Identification of *sgaA* and *mtdA* and Sequences of *sgaA*, *hprA*, and *mtdA*

LUDMILA V. CHISTOSERDOVA AND MARY E. LIDSTROM*

W. M. Keck Laboratories 138-78, California Institute of Technology, Pasadena, California 91125

Received 13 December 1993/Accepted 17 January 1994

In a previous paper, we reported identification of the 5′ part of *hprA* of *Methylobacterium extorquens* AM1, which encodes the serine cycle enzyme hydroxypyruvate reductase (L. V. Chistoserdova and M. E. Lidstrom, J. Bacteriol. 174:71–77, 1992). Here we present the complete sequence of *hprA* and partial sequence of genes adjacent to *hprA*. Upstream of *hprA*, the 3′ part of an open reading frame was discovered, separated from *hprA* by 263 bp. This open reading frame was identified as the gene encoding another serine cycle enzyme, serine glyoxylate aminotransferase (*sgaA*). Cells containing an insertion mutation into *sgaA* were unable to grow on C_1 compounds, demonstrating that the gene is required for C_1 metabolism. Sequencing downstream of *hprA* has revealed the presence of another open reading frame (*mtdA*), which is probably cotranscribed with *hprA*. This open reading frame was identified as the gene required for the synthesis of 5,10-methylenetetrahydrofolate dehydrogenase. Our data suggest that this enzyme plays an integral role in methylotrophic metabolism in *M. extorquens* AM1, either in formaldehyde oxidation or as part of the serine cycle.

Methylobacterium extorquens AM1 is a pink-pigmented serine cycle methylotroph able to grow on methanol and methylamine as well as on a variety of multicarbon substrates (39, 55). While much progress has been made recently in genetic studies of methanol and methylamine oxidation by *M. extorquens* AM1 and other methylotrophs (7, 8, 17, 26, 27), many questions concerning the genetics of the serine cycle remain unanswered. Up to now, three *M. extorquens* AM1 genes encoding serine cycle enzymes necessary for the synthesis of hydroxypyruvate reductase (HPR), phosphoenolpyruvate carboxylase (PEPC), and malyl coenzyme A (malyl-CoA) lyase have been identified. The gene for malyl-CoA lyase (*mclA*) was cloned by complementation of a malyl-CoA lyase mutant and located within a 1.6-kb region on the *M. extorquens* AM1 chromosome (15). A 4-kb fragment adjacent to the *mclA*-containing region was shown to be essential for acetyl-CoA-independent PEPC activity. This region was shown to be responsible for the synthesis of three polypeptides, one of which is the structural gene for the enzyme, while functions of the two others are as yet unknown (3). The structural gene for HPR (*hprA*) was cloned by using an oligonucleotide probe based on N-terminal amino acid sequence of purified HPR and was located about 3 kb upstream of the region responsible for PEPC activity (10). All of these genes are transcribed in the same direction (3), but nothing is known about their transcriptional regulation.

In another study, two large fragments of the *M. extorquens* AM1 chromosome that contain serine cycle genes were cloned (49). One of these complemented glycerate kinase mutants, and another complemented mutants with lesion(s) in the unknown acetyl-CoA oxidation pathway portion of the serine cycle (49). Neither of these fragments overlaps the *hprA-ppc-mclA*-containing region (10). The genes responsible for the synthesis of glycerate kinase and the enzymes participating in oxidation of acetyl-CoA to glyoxylate were not precisely mapped, and the number of genes involved in the latter pathway is not known. The gene for another serine cycle enzyme, serine hydroxymethyltransferase, was recently cloned and sequenced from an obligate methylotroph *Hyphomicrobium methylovorum* GM (36), but its location relative to other methylotrophy genes is unknown.

The goal of this study was to extend our knowledge of the *M. extorquens* AM1 chromosome region containing *hprA* and to determine whether it contains any other genes involved in C_1 metabolism.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown in LB medium in the presence of appropriate antibiotics as described by Maniatis et al. (31). *M. extorquens* AM1 was grown in the minimal medium described previously (15). Succinate (20 mM), methanol (100 mM), methylamine (20 mM), ethanol (40 mM), ethylamine (20 mM), or formate (40 mM) was used as the substrate. Methanol induction of mutants was carried out as described by Dunstan et al. (13). The following antibiotic concentrations were used for *M. extorquens* AM1: tetracycline, 10 μg/ml; kanamycin, 100 μg/ml; and rifamycin, 50 μg/ml. The growth responses of mutants were tested on plates containing the substrates listed above.

DNA-DNA hybridization. DNA-DNA hybridizations were carried out with dried agarose gels as described by Meinkoth and Wahl (34) at 68°C. For hybridizations, 6× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) was used, and 0.5× SSC was used for washes. Plasmid DNA used as a probe was labeled by the random primed method as described by the manufacturer (Boehringer GmbH, Mannheim, Germany).

DNA manipulations. Plasmid isolation, *E. coli* transformation, restriction enzyme digestion, ligation, blunting of ends with T4 DNA polymerase, and filling in ends with Klenow enzyme were carried out as described by Maniatis et al. (31). The chromosomal DNA of *M. extorquens* AM1 was isolated by the procedure of Saito and Miura (44).

* Corresponding author. Phone: (818) 395-2132. Fax: (818) 395-2940.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
E. coli		
DH5α	F⁻ *recA1* (*lacZYA-argF*) U169 *hsdR17 thi-1 gyrA66 supE44 endA1 relA1* Δ80*lac* Δ(*lacZ*) M15	Bethesda Research Laboratories, Inc.
S17-1	F⁻ *pro recA1* r⁻m⁺ RP4-2, integrated (Tc::Mu) (Km::Tn*7*[Sm^r Tp^r])	45
M. extorquens AM1		
AM1rif	Wild type, Rif^r	17
SGAT9	*sgaA*(*sgaA*::Km^r), Rif^r	This study
SGAT17	*sgaA*(*sgaA*::Km^r), Rif^r	This study
Sp2Tc−	*mtdA*(*mtdA*::Km^r), Rif^r	This study
Sp11Tc−	*mtdA*(*mtdA*::Km^r), Rif^r	This study
Sp1Tc$^+$	*mtdA*(*mtdA*::Km^r), Rif^r	This study
Sp10Tc+	*mtdA*(*mtdA*::Km^r), Rif^r	This study
HPR42	*hprA*(*hprA*::Km^r), Rif^r	10
20BL	*hpr-1*, Rif^r	20
Plasmids		
pAYC61	Ap^r, Tc^r, *mob*$^+$, Inc(ColE1)	8
pBE7.21	pBR322 (7.2-kb *Eco*RI), *mclA*	3
pLC3	pUC19 (3-kb *Sph*I), Ap^r, *sgaA*, *hprA*	10
pLC3.02	pUC19 (2-kb *Sph*I-*Eco*RI), Ap^r, *sgaA*	This study
pLC6.1	pRK310 (1-kb *Bam*HI-*Bgl*II), Ap^r, Tc^r, *mob*$^+$, IncP1	This study
pLC7.01	pAYC61 (2-kb *Sph*I-*Eco*RI), Ap^r, Tc^r, *sgaA*::Km^r, *mob*$^+$, Inc(ColE1)	This study
pLC7.24	pUC19 (3-kb *Xho*I), Ap^r, *mtdA*	This study
pLC13	pRK310 (3-kb *Sph*I), Ap^r, Tc^r, *sgaA*, *hprA*, *mob*$^+$, IncP1	10
pLC17.3Sp	pAYC61 (3-kb *Xho*I), Ap^r, Tc^r, *mtdA*::Km^r, *mob*$^+$, Inc(ColE1)	This study
pLC19.33	pUC19 (0.9-kb *Xho*I-*Bgl*II), Ap^r, Δ*mtdA*	This study
pLC19.34	pUC19 (0.9-kb *Xho*I-*Bgl*II), Ap^r, Δ*mtdA*	This study
pLC19.332	pUC19 (0.9-kb *Xho*I-*Bgl*II), Ap^r, *mtdA*::Km^r	This study
pLC21	pRK310 (3-kb *Sph*I), Ap^r, Tc^r, *sgaA*, *hprA*, *mob*$^+$, IncP1	10
pLC28.1b	pUC19 (2-kb *Sph*I-*Eco*RI), Ap^r, *sgaA*::Km^r	This study
pLC410a	pRK310 (3-kb *Xho*I), Ap^r, Tc^r, *mtdA*, *mob*$^+$, IncP1	This study
pLC410b	pRK310 (3-kb *Xho*I), Ap^r, Tc^r, *mtdA*, *mob*$^+$, IncP1	This study
pLC411b	pRK310 (2-kb *Sph*I-*Eco*RI), Ap^r, Tc^r, *sgaA*, *mob*$^+$, IncP1	This study
pRK2013	Km^r, *tra*$^+$, Inc(ColE1)	12
pRK310	Tc^r, *lacZ′ mob*$^+$, IncP1	12
pUC4K	Ap^r, Km^r	56
pUC19	Ap^r, *lacZ′*	Pharmacia

DNA sequencing. DNA sequencing was carried out with an Applied Biosystems automated sequencer by the University of California, Los Angeles, Sequencing Facility, or using Sequenase (U.S. Biochemical Corp., Cleveland, Ohio).

Computer analysis. Translation and analyses of DNA and DNA-derived polypeptide sequences were carried out by using PC/Gene (Genofit SA, Geneva, Switzerland) and DNA-Master (California Institute of Technology, Pasadena).

Enzyme assays. Enzyme activities were determined in *M. extorquens* AM1 crude extracts prepared as described earlier (10). All measurements were done at room temperature in a total volume of 1 ml. The serine cycle enzyme activities were assayed as described earlier (10). Activity of 5,10-methylene tetrahydrofolate (THF) dehydrogenase (MTHFDH) was determined in continuous assay by recording the disappearance of NADP (45). Spectrophotometric methods (23, 61) were used for protein determination.

Matings. Triparental or biparental matings between *E. coli* and *M. extorquens* AM1 were performed overnight on nutrient agar. Cells were then washed with sterile medium and plated on selective medium at appropriate dilutions. In triparental matings, pRK2013 (12) was used as a helper plasmid. Rifamycin was used for *E. coli* counterselection.

Nucleotide sequence accession number. The sequence of 3,358 bp has been deposited with GenBank under accession number L27235.

RESULTS

Sequencing of an *hprA*-containing fragment. A 3-kb *Sph*I fragment of the *M. extorquens* AM1 chromosome cloned in pUC19 (pLC3) was described earlier (10). The 5′ part of *hprA*, coding for the serine cycle HPR, was identified in the middle of this fragment, and the fragment was shown to be able to complement an *hprA* mutant, indicating that the whole gene was present. Cloning of an overlapping region of the *M. extorquens* AM1 chromosome (pBE7.21), containing the gene for PEPC synthesis and a few other genes possibly involved in formaldehyde assimilation, was also described earlier (3). An *Xho*I deletion of 3 kb from this DNA fragment was cloned in this study, adjacent to *hprA* (pLC7.24; Fig. 1). A series of subclones was generated from pLC3 and pLC7.24 in pUC19, and the resulting plasmids used for sequencing. A 3.3-kb fragment containing *hprA* and adjacent DNA regions was sequenced on both strands (Fig. 2). The sequence analysis revealed an open reading frame of 945 bases, including the stop codon, corresponding to *hprA*, which encodes a polypeptide with a calculated molecular mass of 35 kDa. The deduced amino acid sequence of the N-terminal part of the polypeptide was in complete agreement with that of purified HPR, which is known to be 37 kDa, based on sodium dodecyl sulfate-gel electrophoresis (9). Upstream of *hprA*, another partial open reading frame of 1,137 bases was identified, including the stop codon, separated from *hprA* by 266 bp, coding for a polypep-

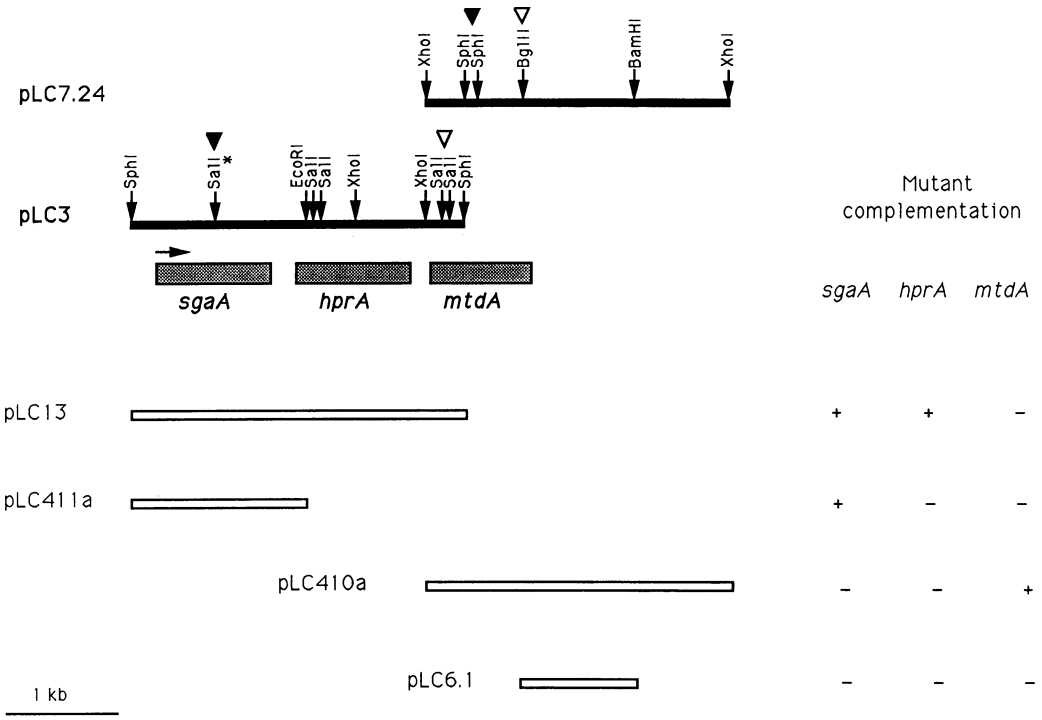

FIG. 1. Physical map of the *M. extorquens* AM1 chromosomal region containing *sgaA*, *hprA*, and *mtdA* and the results of mutant complementation with plasmids containing the indicated subfragments. Transcription is from left to right. *, *Sal*I restriction site shown only for pLC3. Triangles indicate sites where insertion mutations were obtained in the course of this study. A mutation in *hprA* is described in reference 10. Open triangles indicate sites, mutants in which are not investigated in detail.

tide of calculated molecular mass of 42 kDa (Fig. 2). Downstream of *hprA*, a third open reading frame of 867 bases was present, separated from *hprA* by 143 bp, coding for a polypeptide of calculated molecular mass of 32 kDa, and preceded by a putative ribosome binding site, GAGG. No known promoter sequences were found in front of *hprA* or the downstream open reading frame, although a few copies of the characteristic motif GAAA were present upstream of both *hprA* and the downstream open reading frame. This sequence has been implicated in the −10 sequence for the *moxF* promoter of *M. extorquens* AM1 (27).

An inverted repeat sequence was found in the sequenced region, separating the upstream open reading frame and *hprA* (calculated ΔG [25.5°C] of −19 kcal [ca. −58.5 kJ]), having the potential to form a stable stem-loop-type structure (underlined in Fig. 2), which could serve as a terminator for the upstream open reading frame.

Sequence comparisons to the HPR amino acid sequence. The amino acid sequence derived from *hprA* was compared with the only known sequence for this enzyme, cucumber HPR (16). The latter enzyme is a part of the glycolate pathway involved in photorespiration (54), playing a role similar to that of *M. extorquens* AM1 HPR in the serine cycle. Alignment of the two sequences is shown in Fig. 3. The similarity is very low for the N-terminal parts of the proteins but stronger for the C termini, although one gap of 24 amino acid residues and a few smaller gaps had to be introduced for optimal alignment. Considerable similarity for HPR from *M. extorquens* AM1 was found also with the plasmid-encoded VanH protein from *Enterococcus faecium* BM4147 (whose homology with 2-hydroxycarboxylic acid dehydrogenases was reported earlier [4]), formate dehydrogenase from *Pseudomonas* sp. strain 101 (41), phosphoglycerate dehydrogenase from *E. coli* (53), D-2-hydroxyisocaproate dehydrogenase from *Lactobacillus casei* (25), and D-lactate dehydrogenase enzymes from *Lactobacillus plantarum* (50) and *Lactobacillus delbrueckii* subsp. *bulgaricus* (6). With all these enzymes (data not shown), as with cucumber HPR, the enzyme from *M. extorquens* AM1 showed low similarity in the N terminus and stronger similarity in the C terminus. Overall identity between HPR from *M. extorquens* AM1 and the proteins mentioned above varied between 19 and 27%. The complete fingerprint characteristic for the pyridine nucleotide cofactor-binding site was found, highly conserved in other NAD^+-linked dehydrogenases (62) (Fig. 3).

Identification of the upstream open reading frame. The amino acid sequence deduced from the upstream open reading frame was compared against the protein data bank, and similarity was found with human peroxisomal alanine:glyoxylate aminotransferase (42, 51), mitochondrial serine:pyruvate aminotransferase from rat liver (38), and the small subunit of soluble hydrogenase from *Anabaena cylindrica* (14) and *Synechococcus* sp. strain PCC 6716 (56), all apparently belonging to the same family of enzymes. Alignment of the sequence deduced from this open reading frame with the sequences of human alanine:glyoxylate aminotransferase (22% identity) and the small subunit of soluble hydrogenase from *A. cylindrica* (29% identity) is shown in Fig. 4, with amino acid residues conserved for all three enzymes indicated. Similarity with aminotransferases made the upstream open reading frame a candidate for the gene responsible for a serine cycle enzyme, serine glyoxylate aminotransferase (SGAT), and we tentatively designated it *sgaA*.

Construction of an insertion mutation in *sgaA*. To prove the identity of the *sgaA* gene and its significance for C_1 metabo-

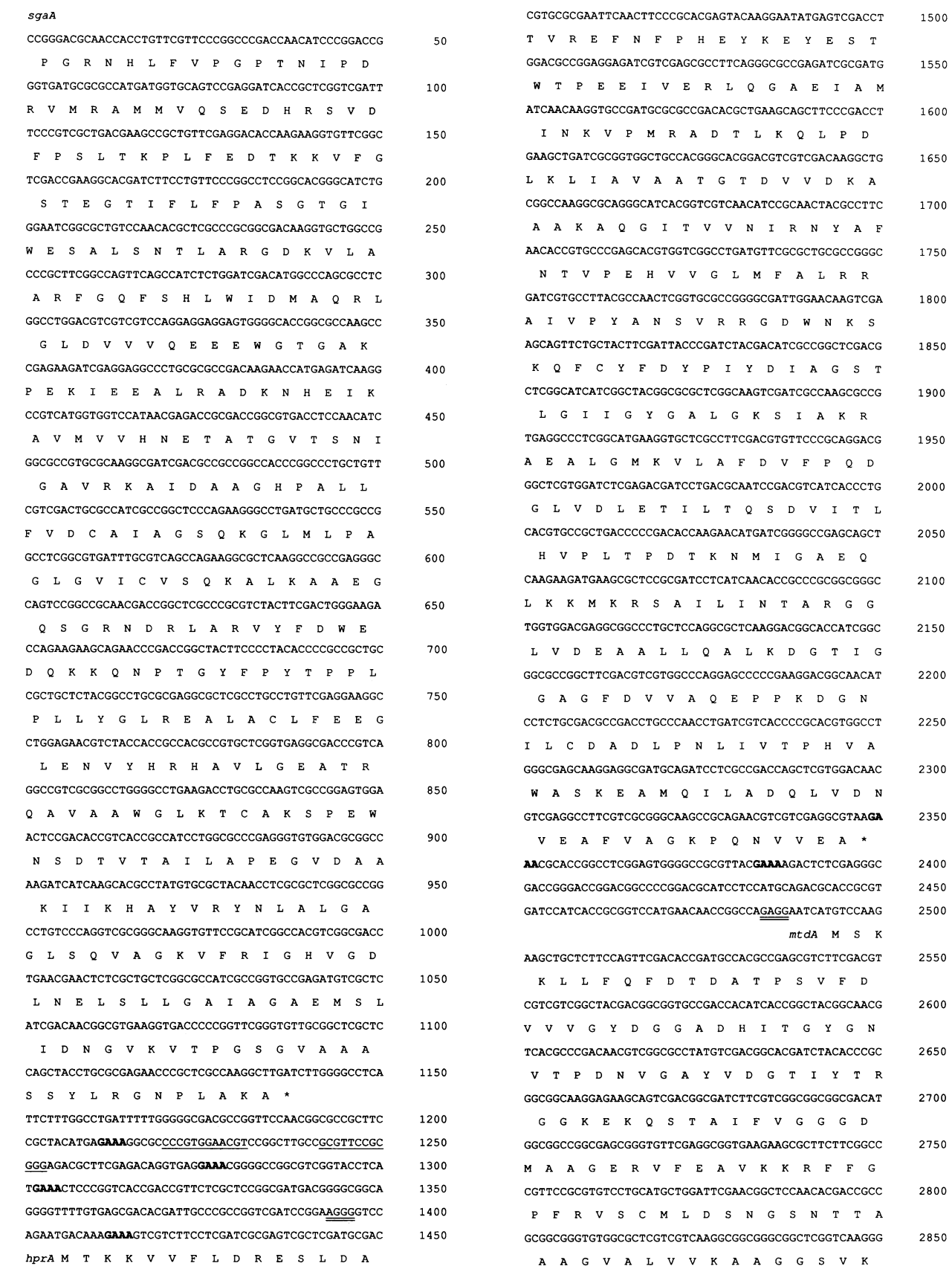

FIG. 2. Nucleotide sequence of a 3,358-bp *M. extorquens* AM1 chromosome region containing *sgaA*, *hprA*, and *mtdA*. Amino acid sequences are deduced from coding regions. Inverted repeat sequences with possible terminator function are underlined, putative Shine-Dalgarno sequences are double underlined, the motif GAAA is shown in bold, and asterisks indicate stop codons.

```
CAAGAAGGCCGTCGTGCTCGCGGGCACCGGCCCGGTCGGCATGCGCTCGG    2900
G K K A V V L A G T G P V G M R S
CGGCGCTGCTCGCCGGCGAGGGCGCCGAGGTCGTGCTGTGCGGGCGCAAG    2950
 A A L L A G E G A E V V L C G R K
CTCGACAAGGCGCAGGCCGCGGCCGATTCCGTGAACAAGCGCTTCAAGGT    3000
  L D K A Q A A A D S V N K R F K
GAACGTCACCGCGGCCGAAACCGCGGACGACGCTTCGCGCGCCGAGGCCG    3050
V N V T A A E T A D D A S R A E A
TGAAGGGCGCCCATTTCGTCTTCACCGCCGGTGCGATCGGCCTTGAACTG    3100
 V K G A H F V F T A G A I G L E L
CTGCCGCAGGCAGCCTGGCAGAACGAGAGTTCGATCGAGATCGTGGCCGA    3150
  L P Q A A W Q N E S S I E I V A
CTACAACGCCCAGCCGCCGCTCGGCATCGGCGGGATCGATGCGACCGACA    3200
D Y N A Q P P L G I G G I D A T D
AAGGCAAGGAATACGGCGGAAAGCGCGCCTTCGGTGCGCTCGGCATCGGC    3250
 K G K E Y G G K R A F G A L G I G
GGCTTGAAGCTCAAGCTGCACCGCGCCTGCATCGCCAAGCTGTTCGAGTC    3300
  G L K L K L H R A C I A K L F E
GAGCGAAGGCGTCTTCGACGCCGAGGAGATCTACAAGCTGGCCAAGGAAA    3350
S S E G V F D A E E I Y K L A K E
TGGCCTGA
 M A *
```

FIG. 2—*Continued.*

```
                                MTKKVVFLDRESLDATVREFNFPHEYKEY     29
MAKPVQIEVWNPNGKYRVVSTKPMPGTRWINLLIEQDCRVEICTEKKTIL     50

ESTWTPEEIVERLQGAEIAMINKVPMRADTLKQLPDLKLIAVAATGTDVV     79
SVEDILALIGDKCDGVIGQLTEDWGEVLFSALSRAGGKAFSNMAVGYNNV    100

DKAAAKAQGITVVNIRNYAFNTVPEHVVGLMFALRRAIVPYANSVRRGDW    129
DVNAANKYGVAVGNTPGVLTETTAELAASLSLAAARRIVEADEFMRAG--    148

                                ▲■■GGG ■■   ■■●
NKSKQFCYFDYPIYDIAGSTLGIIGYGALGKSIA-KRAEALGMKVLAFDV    178
-RYDGWLPNLFVGNLLKGQTVGVIGAGRIGSAYARMMVEGFKMNLIYFDL    197

FPQDGLVDLETILTQS----------------------DVITLHVPLT    204
YQSTRLEKFVTAYGEFLKANGEAPVTWRRASSMDEVLREADVISLHPVLD    247

PDTKNMIGAEQLKKMKRSAILINTARGGLVDEAALLQALKDGTIGGAGFD    254
KTTFHLVNKESLKAMKKDAILINCSRGPVIDEAALVDHLRDNPMFRVGLD    297

VVAQEPPKDGNILCDADLPNLIVTPHVAWASK---EAMQILADQLVDN-V    300
VFEDEPYMKPGLA-DMKNA-IIV-PHIASASKWTREGMATLAALNVLGKI    344

EAF-VAGKPQNVVEA                                       314
KGYPVWSDP-NRVEPFLDENVSPPAASPSIVNAKALGNA               383
```

FIG. 3. Alignment of the derived amino acid sequences for HPRs from *M. extorquens* AM1 (upper line) and cucumber (15). Identical amino acid residue matches between two sequences are shown by double dots; matches between two similar amino acid residues are shown by single dots. The sequences believed to be involved in binding of the ADP moiety of NAD^+ are underlined. The reported amino acid sequence fingerprint for this binding site (59) is shown above the sequences. A triangle represents a basic or hydrophilic residue, a square represents a small and hydrophobic residue, a circle represents an acidic residue, and G represents glycine.

lism, an insertion mutation in the gene was obtained. The strategy for introduction of a specific mutation into *sgaA* was based on the exchange of DNA via homologous recombination in vivo between the wild-type gene and a gene that had been inactivated by an insertion mutation in vitro (43).

The 2-kb *Sph*I-*Eco*RI fragment from pLC3 containing the whole *sgaA* gene, was cloned (pLC3.02), and a 1.4-kb DNA fragment from pUC4K containing a kanamycin resistance (Km^r) gene was inserted into the *Sal*I site located approximately in the middle of the gene (Fig. 1), so that the Km^r gene was transcribed in the same direction as *sgaA*. The plasmid carrying the interrupted *sgaA* (pLC28.1b) was ligated with the suicide vector pAYC61 (7), and the resulting plasmid (pLC7.01) was transformed into *E. coli* S17-1. The resulting *E. coli* strain was used as a donor in biparental matings with *M. extorquens* AM1, and the progeny was selected on succinate plates containing kanamycin. Twenty-four percent of the Km^r colonies were also tetracycline sensitive (Tc^s), which should be the result of a double-recombination event leading to a complete gene replacement. To confirm the presence of the in vitro-mutated *sgaA* at the proper location in the chromosome of *M. extorquens* AM1, chromosomal DNA of the parental and the mutant strains was digested with *Sph*I and used for DNA-DNA hybridization analysis. pLC3 DNA, used as a probe, hybridized with the expected fragments, 3 kb for the wild-type chromosome and 4.4 kb for the mutant strains (data not shown).

Analysis of *sgaA* insertion mutants. Growth responses and the presence of SGAT activity were determined for six separate insertion mutant isolates. All of the mutants lost their ability to grow on C_1 compounds (methanol, methylamine, and formate) and grew normally on C_2 compounds (ethanol and ethylamine), proving the specific involvement of SGAT in the serine cycle for formaldehyde assimilation. Activity of SGAT was absent in the mutants. Two mutants, SGAT9 and SGAT17, were used for further enzymological analysis. The levels of a few key enzymes of the serine cycle were measured in the mutants (Table 2) grown on succinate and induced with methanol (13). Wild-type *M. extorquens* AM1 grown on succinate, methanol, or induced with methanol was used for comparison. For the two latter growth conditions, activities of serine cycle enzymes were similar, so only the values obtained for methanol-grown cells are shown in Table 2. Except for SGAT, all enzymes, including HPR, were present in the mutants at levels similar to the wild-type levels. These data suggest that the insertions in SGAT9 and SGAT17 did not interfere with transcription or regulation of *hprA*.

Expression of *sgaA* in *M. extorquens* AM1. In a previous study, the 3-kb *Sph*I fragment carrying both *sgaA* and *hprA* was used for complementation of an *hprA* insertion mutant. The fragment was able to restore a C_1-positive phenotype when cloned in either orientation with respect to the *lac* promoter (10). Here we used this fragment (plasmids pLC13 and pLC21), and also a smaller fragment containing only *sgaA* (*Sph*I-*Eco*RI; pLC411b), to complement *sgaA* insertion mutants (Fig. 1). *E. coli* DH5α strains carrying the corresponding plasmids were used as donors in triparental matings with *sgaA* mutants of *M. extorquens* AM1. Tc^r colonies were selected on succinate plates. Transconjugants which had received plasmids containing *sgaA* downstream of the *lac* promoter of pRK310 (pLC13 and pLC411b) regained the ability to grow on methanol and methylamine. Transconjugants carrying a plasmid with the opposite orientation of *sgaA* with respect to the *lac* promoter (pLC21) were not able to grow on C_1 compounds. Activities of SGAT and HPR were measured in transconju-

```
AGAT        SHKLLVTPPKALLKPLSIPNQLLLGPGPSNLPPRIMAAGG    40
SGAT                      PGRNHLFVPGPTNIPDRVMRAMM    23
SHG                     MDDKLMLMIPGPTPVPEAALLALA    24

AGAT    LQMIGSMSKDMYQIMDEIKEGIQYVFQTRNPLTLVISGSGHC    82
SGAT    VQSEDHRSVDFPSLTKPLFEDTKKVFGSTEGTIFLFPASGTG    65
SHG     KHPIGHRTSEF-SNMMGEVTQNLKWLHQTESDVLMLNVSGTG    65

AGAT    ALEAALVNVLEPGDSFLVGANGIWGQRAVDIGERIGARVHPM   124
SGAT    IWESALSNTLARGDKVLAARFGQFSHLWIDMAQRLGLDVVVQ   107
SHG     AVEAGMINFLSPGDRILVGSNGKFGERWVEVGQAFGLNVEAI   107

AGAT    TKDPGGHYTLQEVEEGL---AQHKPVLLFLTHGESSTGVLQP   163
SGAT    EEEWGTGAKPEKIEEALRADKNHEIKAVMVVHNETATGVTSN   149
SHG     TAEWGQPLDPDKFAQKLQADTNKEIKAVIITHSETSTGVIND   149

AGAT    L-DGFGELCHRYKCLLLVDSVASLGGTPLYMDRQGIDILYSG   204
SGAT    I-----------------GAVRKAIDAAGHPALLFVDCAIAG   174
SHG     LVAINSHVKEHGQALIIVDAVTSLGAYNVPVDALGLDVVASG   191

AGAT    SQKALNAPPGTSLISFSDKAKKKMYSRKTKPFSFYLDIKWLA   246
SGAT    SQKGLMLPAGLGVICVSQLALKAAEGQSGRNDRLARVYFDWE   216
SHG     SQKGYMIPPGLGFVSVSPKAWEAY-----KTAKLPKYYLDLG   228

AGAT    NFWGCDDQPRMYHHTIPVISLYSLRESLALIAEQGLENSWRQ   288
SGAT    -DQKKQNPTGYFPYTPPLPLLYGLREALACLFEEGLENVYHR   257
SHG     -KYRKATAKNTTPFTPPVNLMVALHTTLGMMKKEGLESIFTR   269

AGAT    HREAAAYLHGRLQALGLQLFVKDPALRLPTVTTVAVPAGYDW   330
SGAT    HAVLGEATRQAVAAWGLKTCAKSPEWNSDTVTAILAPEGVDA   299
SHG     HERQKNATRAAMKALNLPLFAAD-ECASPAITAVATP-GMEA   309

AGAT    RDIVSYVIDHFDIEIMGGLGPSTGKVLRIGLLGCNATRENVD   372
SGAT    AKIIKHAYVRYNLALGAGLSQVAGKVFRIGHVGDLNELSLLG   341
SHG     DKIRSLMKKRFDIALAGGQDHLSNKIFRVGHLGFVSDRDILS   351

AGAT    RVTEALRAALQHCPKKKL                          390
SGAT    AIAGAEMSLIDNGVKVTPGSGVAAASSYLRENPLAKA       378
SHG     CIASLEVVLLELGHENFNSGAGVAAAARVFSN            383
```

FIG. 4. Alignment of the derived amino acid sequences for the mature form of human alanine:glyoxylate aminotransferase (AGAT) (48), SGAT from *M. extorquens* AM1, and the small subunit of soluble hydrogenase from *A. cylindrica* (SHG) (13). The amino acid residues conserved for all three enzymes are indicated by asterisks above the sequences.

gants carrying pLC13, pLC21, or pLC411b grown on succinate or induced with methanol (Table 3). SGAT was present at high levels in constructs having *sgaA* under the *lac* promoter in cells grown on succinate or methanol. For cells containing the construct with *sgaA* in the opposite orientation, SGAT was present at low levels in both succinate-grown and methanol-induced cells, suggesting that a promoter for *sgaA* may be present on the cloned DNA fragment, but some regulatory elements necessary for its proper methanol induction are absent. HPR activity was normal in SGAT9 containing pLC411b, as expected. In SGAT9 containing pLC13, an additional constitutive expression of the cloned *hprA* apparently occurs, presumably as a result of the *lac* promoter, while in SGAT9 containing pLC21, only a small increase over the chromosomal activity is observed.

Sequence comparisons with the downstream open reading frame. The deduced amino acid sequence for the downstream open reading frame was compared with sequences in the protein data bank, but no sequences with high similarity were identified. However, some very low similarity was found to eukaryotic MTHFDH (5, 40, 46). Since MTHFDH is known to be present in serine cycle methylotrophs (24), the downstream open reading frame appeared to be a candidate for the gene encoding this enzyme. The amino acid sequence of this open reading frame has been compared with a few MTHFDH sequences available, including murine mitochondrial NAD-dependent dehydrogenase-cyclohydrolase (5), human NAD-dependent bifunctional enzyme (40), and the corresponding dehydrogenase-cyclohydrolase domains of the trifunctional human enzyme (MTHFDH-methenyl THF cyclohydrolase-formyl THF synthase [21]) and the cytosolic and mitochondrial forms of the yeast enzyme (46, 48), and a region of 21 amino acid residues has been found in which 7 residues are invariant (Fig. 5). Nothing is known about the function of these residues. The size of the product of the downstream open reading frame (32 kDa) matches sizes of NAD-dependent MTHFDH from *Clostridium formicoaceticum* (30 kDa [28]), bifunctional en-

TABLE 2. Activities of some enzymes participating in C_1 assimilation measured in mutant and wild-type *M. extorquens* AM1

Strain	Phenotype	Growth conditions[a]	Enzyme activity (nmol/min/mg of protein)[b]						
			STHM	SGAT	HPR	MCL	PEPC	GK	MTHFDH
AM1rif (wild type)	C_1^+	Succ	20	18	330	—	—	—	30
		MeOH	41	201	1,257	30	93	154	89
SGAT9	C_1^-	Succ	—	0	320	—	—	—	25
		MeOH	28	0	995	28	—	—	74
SGAT17	C_1^-	Succ	—	0	300	—	—	—	28
		MeOH	35	0	1,100	20	—	—	72
Sp2Tc−	C_1^-	Succ	4	—	725	—	—	—	8
		MeOH	10	293	2,778	20	66	180	10
Sp11Tc−	C_1^-	Succ	3	—	709	—	—	—	9
		MeOH	11	317	2,986	15	39	108	10
Sp1Tc+	C_1^+	Succ	6	—	462	—	—	—	30
		MeOH	38	177	1,374	22	44	120	75
Sp10Tc+	C_1^+	Succ	10	—	420	—	—	—	44
		MeOH	35	160	1,335	18	41	108	85
HPR42[c]	C_1^-	Succ	—	—	2	—	—	—	10
		MeOH[d]	—	—	28	—	—	—	30
20BL[c]	C_1^-	Succ	—	—	3	—	—	—	28
		MeOH[d]	—	—	29	—	—	—	74

[a] Succ, cells were grown on succinate; MeOH, cells were grown on methanol except as indicated otherwise.
[b] Data are averages of two to four independent measurements. Values agreed ±15%. —, activity was not measured. STHM, serine transhydroxymethylase; MCL, malate thiokinase-malyl-CoA lyase; GK, glycerate kinase.
[c] Serine cycle enzyme activities for this mutant are discussed in reference 10.
[d] Cells were grown on succinate, washed, and incubated with methanol for at least 24 h to allow for induction.
[e] Enzyme activities for this mutant are discussed in reference 20.

TABLE 3. Expression of *sgaA* in *M. extorquens* AM1

Strain	Phenotype	Growth conditions[a]	Enzyme activity (nmol/min/mg of protein)[b]	
			SGAT	HPR
SGAT9(pLC411b)	C_1^+	Succ	105	320
		MeOH	95	1,100
SGAT9(pLC13)	C_1^+	Succ	165	1,765
		MeOH	150	2,590
SGAT9(pLC21)	C_1^-	Succ	6	867
		MeOH[c]	7	1,250

[a] Succ, cells were grown on succinate; MeOH, cells were grown on methanol unless indicated otherwise.

[b] Results are averages of two to four independent measurements. Values agreed ±15%.

[c] Cells were grown on succinate, washed, and incubated with methanol to allow for induction.

zymes (MTHFDH-methenyl THF cyclohydrolase) isolated from *Clostridium thermoaceticum* (30 kDa [28]) and mammals (34 kDa, 35), and the amino-terminal domain (33 kDa) of the trifunctional enzyme from mammalian cells (52, 58). For convenience, we will further refer to this open reading frame as *mtdA*. Alignment of the amino acid sequence deduced from *mtdA* with MTHFDH from the mouse is shown in Fig. 5, and residues invariant for all enzymes whose sequences are available are indicated. Although the similarity is very low (less than 15% identity), the two sequences can be aligned without introducing a single gap.

Construction of an insertion mutant in *mtdA*. The *Pst*I site of the pUC19 linker and the *Bgl*II site of the insert of pLC7.24 were used to clone a 0.9-kb *Xho*I-*Bgl*II DNA fragment containing most of *mtdA* into pUC19 (pLC19.34). The *Sph*I site was then removed from the pUC19 linker by cutting pLC19.34 with *Hin*dIII and *Pst*I, followed by treatment of the ends with Klenow enzyme and ligation of the blunt ends. The resulting plasmid, pLC19.33, had two closely located *Sph*I sites approximately in the middle of the insert and in the beginning of *mtdA* (Fig. 1). These sites were treated with T4 DNA polymerase and used for insertion of a *Hinc*II fragment from pUC4K carrying the Km^r gene. In the resulting plasmid (pLC19.332), the Km^r gene was transcribed in the same direction as *mtdA*, while *mtdA* had a deletion of 123 bp and was truncated by 27 bp at the 3′ end. This plasmid was ligated into the *Kpn*I site of pAYC61, and the resulting plasmid (pLC17.3Sp) was transformed into *E. coli* S17-1, which served as a donor strain in biparental matings with *M. extorquens* AM1. Km^r transconjugants were routinely selected on succinate minimal plates. In this case, however, no transconjugants (of 150 colonies checked) were isolated that were Km^r and Tc^s, the criterion for double recombination. All recombinants were Tc^r, suggesting that they contained the suicide vector as a result of a recombinational event. Our inability to obtain null mutants or double recombinants in *mtdA* could indicate that the gene is essential for growth not only on C_1 compounds but also on multicarbon compounds. This hypothesis is consistent with the idea of *mtdA* being the gene for MTHFDH, since this enzyme participates in basic cell metabolic pathways, including purine biosynthesis (30).

Analysis of *mtdA* insertion mutants. The transconjugants resistant to both Km and Tc fell into two phenotypic classes with regard to their ability to grow on C_1 compounds. About 80% of them were methanol and methylamine negative; the rest were able to grow on C_1 compounds. DNA-DNA hybrid-

```
                                                MSKKLLFQFNT    11
NEAVVISGRKLAQQIKQEVQQEVEEWVASGNKRPHLSVILVGDNPASHSY    50

DATPSVFDVVVGYDGGADHITGYGNVTPDNVGAYVDGTIYTRGGKEKQST    61
VLNKTRAAAEVGINSETIVKPASVSEEELLNSIRKLNNDEDVDGLLVQLP   100

AIFVGGGDMAAGERVFEAVKKRFFGPFRVSCMLDSNGSNTTAAAGVALVV   111
LPEHIDERKVCNAVSPDKDVDGFHVINVGRMSLDQYSMLPATPWGVWEII   150

KAAGGSVKGKKAVVLAGTGPVGMRSAALLAGEGAEVVLCGRKLDKAQAAA   161
KRTGIPTLGKNVVVAGRSKNVGMPIAMLLHTDGAHERPGGDATVTISHRY   200

DSVNKRFKVNVTAAETADDASRAEAVKGAHFVFTAGAIGLELLPQAAWQN   211
TPKEQLKKHTILADIVISAAGIPNLITADMIKEGAAVIDVGINRVQDPVT   250

ESSIEIVADYNAQPPLGIGGIDATDKGLEYGGKRAFGALGIGGLKLKLHR   261
AKPKLVGDVDFEGVKKAGYITPVPGGVGPMTVAMLMKNTIIAAKKVLRP   300

ACIAKLFESSEGVFDAEEIYKLAKEMA                          288
EELEVFKSKQRGVATN                                     316
```

FIG. 5. Alignment of the deduced amino acid sequences for MTH FDH from *M. extorquens* AM1 (upper line) and the mature form of bifunctional enzyme (MTHFDH-methenyl THF cyclohydrolase) from mouse (5). Identical amino acid residue matches between the two sequences are shown by double dots; matches between two similar amino acid residues are shown by a single dot. Amino acid residues conserved in all the sequences available for MTHFDH enzymes (5, 20, 38, 44) are indicated by asterisks above the sequences.

ization analysis was used to determine the difference between these two classes at the insertion site. *Xho*I chromosomal digests of the parental strain and representatives of both classes of recombinants were hybridized with pLC7.24 DNA, the Km^r gene cartridge, or the suicide vector (data not shown). Representatives of the first class, not able to grow on C_1 compounds, Sp2Tc− and Sp11Tc−, had the gene interrupted in its native location. The 3-kb band seen for the wild-type chromosome was absent, replaced with a heavier band of about 12 kb, which is the correct size to be the result of recombination with pLC17.3Sp at a site 5′ to the Km^r gene insertion site. Probing with the Km^r fragment and vector DNA identified the expected fragments for this recombinational event. Representatives of the second class of mutants, able to grow on C_1 compounds, Sp1Tc− and Sp10Tc+, apparently carried an intact copy of *mtdA* in its natural site (3-kb band) and another copy of the gene as a part of the integrated suicide vector. This conclusion was supported by probing with the Km^r fragment and the vector (data not shown). When the *Bgl*II site or two closely located *Sal*I sites (positions 2627 and 2668 in Fig. 2; also shown in Fig. 1) were used for insertion of the Km^r gene, double-crossover recombinants were also not possible to isolate. In these cases, as with the *Sph*I insertions, all recombinants obtained were Km^r and Tc^r, and all had the suicide vector integrated into the chromosome (data not shown).

To further clarify the phenotype of these insertion mutants, the activity of MTHFDH and also activities of key enzymes of the serine cycle were measured. In the C_1-positive mutants Sp1Tc$^+$ and Sp10Tc$^+$, all enzymes were present at approximately the wild-type level, and a normal pattern of methanol induction was observed (Table 2), confirming the presence of a wild-type *mtdA*. In the C_1-negative mutants, Sp2Tc− and Sp11Tc−, MTHFDH was about threefold lower than the wild-type level in succinate-grown cells, and it was at a similar

TABLE 4. Expression of *mtdA* in *M. extorquens* AM1

Strain	Phenotype	Growth conditions[a]	Activity of MTHFDH (nmol/min/mg of protein)[b]
Sp2Tc-(pLC410a)	C_1^+	Succ	270
		MeOH	280
Sp11Tc-(pLC410a)	C_1^+	Succ	275
		MeOH	280
HPR42(pLC410a)	C_1^-	Succ	260
		MeOH[c]	265
HPR42(pLC410b)	C_1^-	Succ	10
		MeOH[c]	30
HPR42(pLC6.1)	C_1^-	Succ	11
		MeOH[c]	32
HPR42(pLC13)	C_1^+	Succ	10
		MeOH	33
HPR42(pLC21)	C_1^+	Succ	9
		MeOH	29

[a] See footnote *a* of Table 3.
[b] See footnote *b* of Table 3.
[c] See footnote *c* of Table 3.

low level in methanol-induced cells. Other enzyme levels in these mutants were in the range observed for other serine cycle mutants, except HPR and SGAT activities, which were increased. The reason for this is unknown. In the case of HPR, it could be due to altered transcripts produced as a result of the insertion. An increase in SGAT activity could be due to an increase in HPR activity, since HPR is used as an internal coupling enzyme in the SGAT assay (10).

Previously we described an insertion mutant in *hprA*, HPR42, unable to grow on C_1 compounds (10), but MTHFDH was not measured in this mutant. We have now measured MTHFDH activity in HPR42 for comparison with the *mtdA* mutants. It is seen from Table 2 that MTHFDH activity in HPR42 is lowered about threefold compared with the wild type, but the methanol induction pattern (about threefold) is present.

In another mutant, 20BL, selected after chemical mutagenesis (20) and defective in *hprA* and some other, yet unknown function, MTHFDH was present at the wild-type level (Table 2).

Expression of *mtdA* in *M. extorquens* AM1. A 3-kb *Xho*I fragment containing the entire *mtdA* and an adjacent downstream fragment was ligated into pRK310, resulting in plasmids pLC410a (in which *mtdA* is under the control of the *lac* promoter) and pLC410b (in which *mtdA* is in the opposite orientation with respect to the *lac* promoter). Both plasmids were transformed into *E. coli* DH5α, which was used as a donor in triparental matings with Sp2Tc− and Sp11Tc−. A 1-kb *Bgl*II-*Bam*HI fragment containing the 3′ portion of *mtdA* and an adjacent downstream fragment, whose sequence is not yet available, was also used for complementation. This fragment was under the control of the *lac* promoter (pLC6.1; Fig. 1) and served as a negative control for *mtdA* complementation. The leftmost portion of the *Xho*I fragment in pLC410a and pLC410b contains a gene for malate thiokinase (11) and cannot be responsible for complementation of *mtdA* mutations.

Since both the Mtd mutants and the conjugative plasmid carried Tc resistance, transconjugants could only be selected by complementation on methanol minimal plates. Only pLC410a, not pLC410b or pLC6.1, was able to complement the mutants. Activity of MTHFDH was measured in transconjugants (Table 4) and found to be about 25-fold increased compared with the mutant levels when pLC410a was present. HPR42 does not carry Tc resistance, and so in this case it was possible to obtain transconjugants containing all plasmids. The same pattern of *mtdA* expression was observed for HPR42 carrying pLC410a, although HPR42 could not be complemented for growth on methanol or methylamine by this DNA fragment. MTHFDH was at the mutant level in HPR42 when pLC410b or pLC6.1 was present (Table 4).

We have shown here that the *hprA* insertion mutant HPR42 expresses a low level of MTHFDH. Earlier we had shown that this mutant can be complemented by the DNA fragment upstream of *mtdA* and including only the 5′ end of *mtdA* (3-kb *Sph*I fragment; pLC13 and pLC21; Fig. 1). The transconjugants grew on C_1 compounds with practically the same rate as the wild type (10). It was of interest to measure MTHFDH in these constructs, to determine whether this low-level expression could support growth on C_1 compounds. MTHFDH in HPR42 carrying pLC13 or pLC21 was as low as in the mutant alone and was induced about threefold on methanol. Therefore, it seems that MTHFDH at an activity level found in Sp2Tc− and Sp11Tc− (10 nmol/min/mg of protein) is not sufficient for growth on methanol, but at 30 nmol/min/mg of protein normal growth on methanol occurs.

Since MTHFDH has been proposed to directly participate in C_1 assimilation during growth on formate (24), both types of *mtdA* mutants and also transconjugants containing *mtdA* in *trans* were tested for growth on formate minimal plates. Although all of the strains grew very slowly, it was possible to differentiate between C_1-negative *mtdA* mutants, which were the slowest, and C_1-positive mutants, and transconjugants carrying pLC411a, which grew with the wild-type rate.

DISCUSSION

In our recent work, we have identified *hprA*, coding for the serine cycle HPR (10). The gene was localized in close proximity to two other serine cycle genes responsible for the synthesis of PEPC and malyl-CoA lyase (3, 10). These data suggested the possibility that serine cycle genes, like genes involved in the oxidation of methanol (27) and methylamine (7, 8, 26), could be clustered on the chromosome of *M. extorquens* AM1. To determine whether there were other genes in this chromosomal region involved in the serine cycle, *hprA* and the chromosomal regions adjacent to it have been sequenced and analyzed. A new gene involved in the serine cycle, *sgaA*, encoding SGAT was identified upstream of *hprA*. An insertion mutation into the gene has shown that the enzyme is specific for C_1 metabolism and is not required for growth on multicarbon compounds. SGAT was sequenced for the first time for a serine cycle methylotroph, and it showed similarity to the small subunit of soluble hydrogenases as well as to a family of serine pyruvate aminotransferases. The nature of similarity of one of the two subunits of soluble hydrogenase with some aminotransferases is not known, and the possible evolutionary relationship of these enzymes is unclear. Soluble hydrogenase does not show any similarity with its functional counterpart, membrane-bound hydrogenase (14). However, within the group of aminotransferases many sequences are available, and similarity has been found for primary structures of tyrosine, histidinol-phosphate, and aspartate aminotransferases (33). For all of these enzymes, a few invariant residues that play an important role in structure and function of aminotransferases were found (1, 33, 60). However, enzymes belonging to the serine:pyruvate aminotransferase family, including SGAT from *M. extorquens* AM1, do not show any similarity with these enzymes in their primary structure, and

most of the invariant residues were not found in corresponding locations of these sequences. Recently, Oda and coauthors (38) compared the secondary structure of mitochondrial serine:pyruvate aminotransferase with those of rat tyrosine, ornithine, and aspartate aminotransferases and found significant similarity between the secondary structures of the enzymes. The secondary structures were tentatively separated into three segments, A, B, and C, with segment B consisting of a cluster of β structure, and segments A and C consisting of alternating α helix and β structure. Since lysine 250 of mitochondrial aspartate aminotransferase was known to be the cofactor (pyridoxal phosphate)-binding lysine (22) and was located in the B segment, Oda and coauthors proposed that lysine 207, located in the B segment of rat serine:pyruvate aminotransferase, could be the cofactor-binding residue (38). Lys-207 corresponds to a conserved Lys residue for the three amino acid sequences aligned in Fig. 4, which would make Lys-209 of human alanine:glyoxylate aminotransferase and Lys-177 of *M. extorquens* AM1 SGAT potential candidates for pyridoxal phosphate-binding residues. The Lys residue is also conserved in the sequence of the small subunit of soluble hydrogenase, although a requirement of pyridoxal phosphate for this enzyme activity was not reported (14). *sgaA* seems to be transcribed from its own promoter, but the promoter region and elements involved in its regulation by single carbon compounds remain to be identified.

The gene responsible for the synthesis of the serine cycle HPR, *hprA*, was also sequenced for the first time for a methylotroph. Another HPR, fulfilling a similar function in the glycolate pathway in photorespiration, has been sequenced from cucumber cotyledons (16). The two enzymes show rather low similarity, especially in the N-terminal regions, and a few gaps are needed to align the enzymes. Both HPR enzymes belong to a newly identified family of D-isomer-stereospecific 2-hydroxyacid dehydrogenases (50), and a characteristic fingerprint is present responsible for the β-α-β fold common for NAD-binding domain of the NAD-dependent dehydrogenases (62).

A region of 263 bp separates *sgaA* and *hprA*; this region apparently contains a promoter for *hprA*, since the gene is expressed in *M. extorquens* AM1 when no *lac* promoter is present in front of it (10), and a DNA fragment upstream of *hprA* shows promoter activity when cloned into a promoter probe vector (11). Four copies of a motif GAAA are present upstream of *hprA*. This motif was shown to be a part of the −10 region of the putative promoter for *moxF* in *M. extorquens* AM1 and *Methylobacterium organophilum* XX (27, 29). It is also found in front of the *moxF* and *moxZ* genes in *Paracoccus denitrificans* (18, 19), but it is not known whether it serves as a promoter sequence in this organism. This GAAA motif is also present in front of the *moxF* gene in methane utilizers *Methylomonas albus* BG8 and *Methylomonas* sp. strain A4, but in these bacteria, this motif apparently is not involved in promoter structures (11, 59). A few copies of this motif were also found upstream of the *moxW* gene in *M. organophilum* XX (63) and upstream of *pqqD* and *pqqC* genes of *M. extorquens* AM1 (37), and in this case they also do not participate in promoter structures. Mutational analysis is needed to determine the role of this motif in transcriptional regulation of methylotrophic genes.

A gene *mtdA*, showing some similarity with eukaryotic MTHFDH, was discovered downstream of *hprA*. It is interesting that neither sequence of the MTHFDH proteins in Fig. 6 revealed the presence of the fingerprint for a pyridine nucleotide cofactor-binding site, characteristic for other dehydrogenases (62) and present in the HPR sequence (this work).

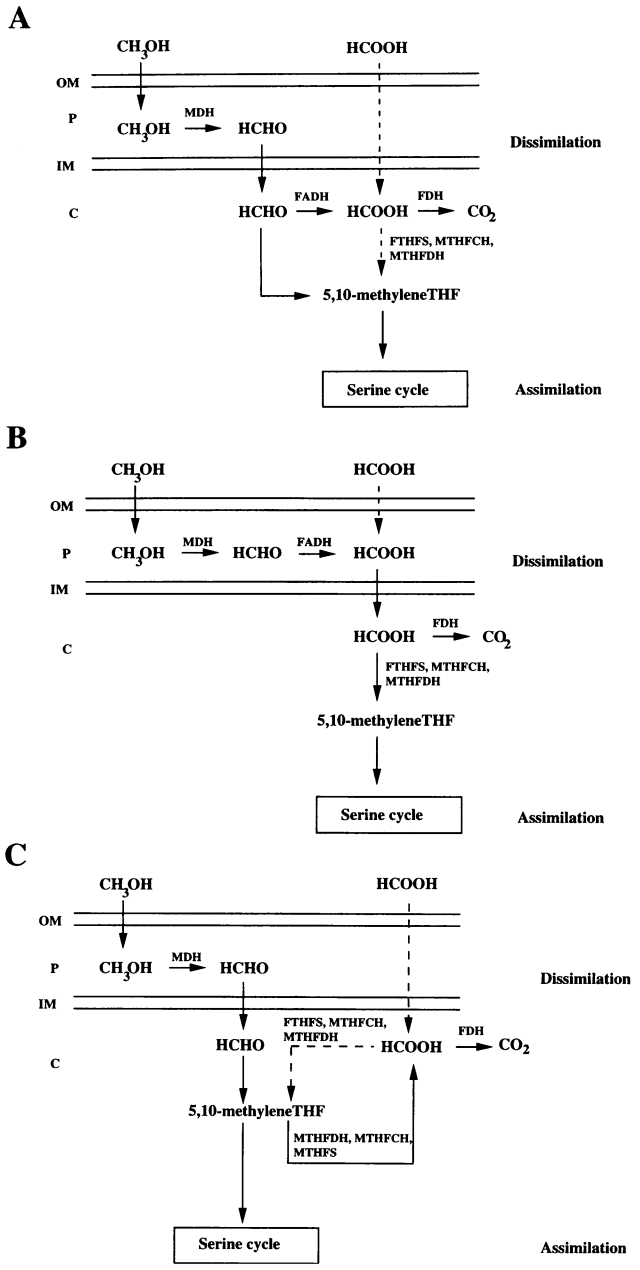

FIG. 6. Possible role of MTHFDH and other enzymes of the MTHF pathway in growth of *M. extorquens* AM1 on methanol (solid lines) and proposed role of the pathway in growth on formate (broken lines). (A) Currently held scheme, assuming that the MTHF pathway is not required for growth on methanol. (B) Scheme that would require the MTHF pathway for growth on methanol involving oxidation of formaldehyde in the periplasm and reduction of formate to 5,10-methylene THF by the MTHF pathway. (C) Scheme that would require the MTHF pathway for growth on methanol not involving formaldehyde dehydrogenase. The MTHF pathway works in the opposite direction, oxidizing 5,10-methylene THF to formate. OM, outer membrane; P, periplasm; IM, inner membrane; C, cytoplasm; MDH, methanol dehydrogenase; FADH, formaldehyde dehydrogenase; FDH, formate dehydrogenase; FTHFS, formyl THF synthase; MTHFCH, methenyl THF cyclohydrolase.

Two classes of insertion mutants in *mtdA* were obtained, and both were the result of insertion of the plasmid by single-crossover recombination. In one, recombination occurred 3′ to the Km^r gene, which generates an intact copy of *mtdA* downstream of the vector, separated from its native promoter (10, 11). These recombinants were unable to utilize C_1 compounds and had low, noninducible levels of MTHFDH, presumably due to promotion from vector sequences. In the second class, an intact *mtdA* was present upstream of an interrupted *mtdA*, a class that must have occurred via some type of duplication or rearrangement event. This class had wild-type levels of MTHFDH and grew normally on C_1 compounds. Two other expected classes of recombinants, one in which the crossover occurred 5′ to the Km^r resistance gene and another in which double-crossover recombination occurred, were not obtained. In both cases, no intact *mtdA* would have been generated. The lack of these two classes of insertion mutants suggests that *mtdA* is required for growth on both succinate and C_1 compounds. The ability of the class of recombinants expressing low levels of MTHFDH to grow on succinate but not on C_1 compounds further suggests that MTHFDH is required at higher levels for methylotrophic growth than for growth on succinate. This hypothesis is supported by the observation that a DNA fragment containing *mtdA* was able to complement the C_1-negative mutants, with concomitant restoration of high MTHFDH activity. These data and the similarity to eukaryotic MTHFDHs allowed us to identify *mtdA* as the gene for MTHFDH.

It was surprising to find *mtdA* closely linked to serine cycle genes and required for growth on methanol and methylamine. MTHFDH is not generally thought to be a part of the serine cycle. It has been assumed that in serine cycle methylotrophs, periplasmically generated formaldehyde is transported to the cytoplasm, where it either is oxidized to formate for dissimilatory purposes or reacts nonenzymatically with THF to form 5,10-methylene THF for entry into the serine cycle for assimilatory purposes (Fig. 6A) (2). MTHFDH is thought to be required for growth on formate, in which case the formate is reduced to 5,10-methylene THF via a pathway (the MTHF pathway) involving MTHFDH, methenyl THF cyclohydrolase, and formyl THF synthase (Fig. 6A) (2).

Our results suggest two possibilities for involvement of MTHFDH in methylotrophic metabolism. First, it may be that formaldehyde must be oxidized to formate in the periplasm and then formate is transported to the cytoplasm and reduced to methylene THF via the MTHF pathway in order to enter the serine cycle (Fig. 6B). Alternatively, formaldehyde could be transported to the cytoplasm as in Fig. 6A, and all of it converted to 5,10-methylene THF. Part of it could be oxidized to formate via the MTHF pathway and part could enter the serine cycle (Fig. 6C). This alternative mechanism of formaldehyde oxidation in serine methylotrophs was proposed earlier by Marison and Attwood (32) on the basis of enzyme activities but never confirmed by mutant analysis. MTHFDHs, along with methenyl THF cyclohydrolase and formyl THF synthase, are known to be present in *M. extorquens* AM1 and show induced levels in methanol-grown cells (24, 32). The induction by C_1 substrates and the requirement for higher levels of MTHFDH activity during growth on C_1 compounds as opposed to growth on succinate suggest a central role for MTHFDH in methylotrophic metabolism. It is not yet clear which scheme (Fig. 6B or C) is correct, but further studies are under way to discriminate between these two possibilities.

ACKNOWLEDGMENT

This work was supported by NIH grant GM36294.

REFERENCES

1. **Alfano, J. R., and M. Kahn.** 1993. Isolation and characterization of a gene coding for a novel aspartate aminotransferase from *Rhizobium meliloti*. J. Bacteriol. **175:**4186–4196.
2. **Anthony, C.** 1982. The biochemistry of methylotrophs. Academic Press, Inc., London.
3. **Arps, P. J., G. F. Fulton, E. C. Minnich, and M. E. Lidstrom.** 1993. Genetics of serine pathway enzymes in *Methylobacterium extorquens* AM1: phosphoenolpyruvate carboxylase and malyl coenzyme A lyase. J. Bacteriol. **175:**3776–3783.
4. **Arthur, M., C. Molians, S. Dutka-Malen, and P. Courvalin.** 1991. Structural relationship between the vancomycin resistance protein VanH and 2-hydroxycarboxylic acid dehydrogenases. Gene **103:** 133–134.
5. **Belanger, C., and R. E. MacKenzie.** 1989. Isolation and characterization of cDNA clones encoding the murine NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase. J. Biol. Chem. **264:**4837–4843.
6. **Bernard, N., T. Ferain, D. Garmyn, P. Hols, and J. Delocur.** 1991. Cloning of the D-lactate dehydrogenase gene from *Lactobacillus delbruekii* subsp. *bulgaricus* by complementation in *Escherichia coli*. FEBS Lett. **290:**61–64.
7. **Chistoserdov, A. Y., L. V. Chistoserdova, W. McIntire, and M. E. Lidstrom.** The genetic organization of the *mau* gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of *mau* mutants. Submitted for publication.
8. **Chistoserdov, A. Y., Y. D. Tsygankov, and M. E. Lidstrom.** 1991. Genetic organization of methylamine utilization genes from *Methylobacterium extorquens* AM1. J. Bacteriol. **173:**5901–5908.
9. **Chistoserdova, L. V., and M. E. Lidstrom.** 1991. Purification and characterization of hydroxypyruvate reductase from the facultative methylotroph *Methylobacterium extorquens* AM1. J. Bacteriol. **173:**7228–7232.
10. **Chistoserdova, L. V., and M. E. Lidstrom.** 1992. Cloning, mutagenesis, and physiological effect of a hydroxypyruvate reductase gene from *Methylobacterium extorquens* AM1. J. Bacteriol. **174:**71–77.
11. **Chistoserdova, L. V., and M. E. Lidstrom.** Unpublished results.
12. **Ditta, G., T. Schmidhauser, F. Yakobson, P. Lu, X. Liang, D. Finlay, D. Guiney, and D. Helinski.** 1985. Plasmids related to the broad host range vector, pRK290, useful for gene cloning and monitoring gene expression. Plasmid **13:**149–153.
13. **Dunstan, P. M., C. Anthony, and W. T. Drabble.** 1972. Microbial metabolism of C1 and C2 compounds. The involvement of glycollate in the metabolism of methanol and acetate by *Pseudomonas* AM1. Biochem. J. **128:**99–106.
14. **Ewart, G. D., K. C. Reed, and G. D. Smith.** 1990. Soluble hydrogenase of *Anabaena cylindrica*. Cloning and sequencing of a potential gene encoding the tritium exchange subunit. Eur. J. Biochem. **187:**215–223.
15. **Fulton, G. L., D. N. Nunn, and M. E. Lidstrom.** 1984. Molecular cloning of a malyl CoA lyase gene from *Pseudomonas* AM1, a facultative methylotroph. J. Bacteriol. **160:**718–723.
16. **Greenler, J. M., J. S. Sloan, B. W. Schwartz, and Wayne M. Becker.** 1989. Isolation, characterization and sequence analysis of a full-length cDNA clone encoding NADH-dependent hydroxypyruvate reductase from cucumber. Plant Mol. Biol. **13:**139–150.
17. **Harms, N.** 1993. Genetics of methanol oxidation in *Paracoccus denitrificans*, p. 235–244. *In* J. C. Murrell and D. P. Kelly (ed.), Microbial growth on C-1 compounds. Intercept Limited, Andover, England.
18. **Harms, N., G. E. De Vries, K. Maurer, J. Hoogendijk, and A. H. Stouthamer.** 1987. Isolation and nucleotide sequence of the methanol dehydrogenase structural gene from *Paracoccus denitrificans*. J. Bacteriol. **169:**3969–3975.
19. **Harms, N., W. N. M. Reijnders, H. Anazawa, C. J. N. M. van der Palen, R. J. M. van Spanning, L. F. Ottmann, and A. H. Stouthamer.** 1993. Identification of a two-component regulatory system controlling methanol dehydrogenase synthesis in *Paracoccus denitrificans*. Mol. Microbiol. **8:**457–470.
20. **Heptinstall, J., and J. R. Quayle.** 1970. Pathways leading to and

from serine during growth of *Pseudomonas* AM1 on C1 compounds or succinate. Biochem. J. **117**:563–572.
21. **Hum, D. W., A. W. Bell, R. Rosen, and R. E. MacKenzie.** 1988. Primary structure of a human trifunctional enzyme. Isolation and characterization of a cDNA encoding methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase-formyltetrahydrofolate synthetase. J. Biol. Chem. **263**:15946–15950.
22. **Huynh, Q. K., R. Sakakibara, T. Watanabe, and H. Wada.** 1980. Primary structure of mitochondrial glutamic oxaloacetic transaminase from rat liver: comparison with that of the pig heart isozyme. Biochem. Biophys. Res. Commun. **97**:474–479.
23. **Kalb, V. F., and R. W. Bernlohr.** 1977. A new spectrophotometric assay for protein in cell extracts. Anal. Biochem. **82**:362–371.
24. **Large, P. J., and J. R. Quayle.** 1963. Microbial growth on C1 compounds. 5. Enzyme activities in extracts of *Pseudomonas* AM1. Biochem. J. **87**:386–396.
25. **Lerch, H.-P., H. Blocker, H. Kallwass, J. Hoppe, H. Tsai, and J. Collins.** 1989. Cloning, sequencing and expression in *Escherichia coli* of the D-2-hydroxyisocaproate dehydrogenase gene of *Lactobacillus casei*. Gene **78**:47–57.
26. **Lidstrom, M. E., and A. Y. Chistoserdov.** 1993. Molecular biology and genetics of methylamine dehydrogenases, p. 381–400. *In* J. C. Murrell and D. P. Kelly (ed.), Microbial growth on C-1 compounds. Intercept Limited, Andover, England.
27. **Lidstrom, M. E., and Y. D. Tsygankov.** 1991. Molecular genetics of methylotrophic bacteria, p. 173-3-4. *In* I. Goldberg and J. S. Rokem (ed.), Biology of methylotrophs. Butterworth-Heinemann, New York.
28. **Ljungdahl, L. G., W. E. O'Brien, M. R. Moore, and M.-T. Liu.** 1980. Methylenetetrahydrofolate dehydrogenase from *Clostridium formicoaceticum* and methylenetetrahydrofolate dehydrogenase, methenyltetrahydrofolate cyclohydrolase (combined) from *Clostridium thermoaceticum*. Methods Enzymol. **66**:599–609.
29. **Machlin, S. M., and R. S. Hanson.** 1988. Nucleotide sequence and transcriptional start site of the *Methylobacterium organophilum* XX methanol dehydrogenase structural gene. J. Bacteriol. **170**:4739–4747.
30. **MacKenzie, R. E.** 1984. Biogenesis and interconversion of substituted tetrahydrofolates, p. 255–306. *In* R. L. Blackely and S. J. Benkovic (ed.), Folates and pterins. John Wiley & Sons, New York.
31. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Marison, I. W., and M. M. Attwood.** 1982. A possible mechanism for the oxidation of formaldehyde to formate. J. Gen. Microbiol. **128**:1441–1446.
33. **Mehta, P. K., T. I. Hale, and P. Christen.** 1989. Evolutionary relationships among aminotransferases. Tyrosine aminotransferase, histidinol-phosphate aminotransferase, and aspartate aminotransferase are homologous proteins. Eur. J. Biochem. **186**:249–253.
34. **Meinkoth, J., and G. Wahl.** 1984. Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. **138**:267–284.
35. **Mejia, N. R., E. M. Rios-Orlandi, and R. E. MacKenzie.** 1986. NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase from ascites tumor cells. Purification and properties. J. Biol. Chem. **261**:9509–9513.
36. **Miyata, A., T. Yoshida, K. Yamaguchi, C. Yokoyama, T. Tanabe, H. Ton, T. Mitsunaga, and Y. Izumi.** 1993. Molecular cloning and expression of the gene for serine hydroxymethyltransferase from an obligate methylotroph *Hyphomicrobium methylovorum* GM2. Eur. J. Biochem. **212**:745–750.
37. **Morris, C. J., F. Biville, E. Turlin, E. Lee, K. Ellermann, W.-H. Fan, R. Ramamoorthi, A. L. Springer, and M. E. Lidstrom.** 1994. Isolation, phenotypic characterization, and complementation analysis of mutants in *Methylobacterium extorquens* AM1 unable to synthesize pyrroloquinoline quinone and sequences of *pqqD*, *pqqG*, and *pqqC*. J. Bacteriol. **176**:1746–1755.
38. **Oda, T., H. Miyajima, Y. Suzuki, and A. Ichiyama.** 1987. Nucleotide sequence of the cDNA encoding the precursor for mitochondrial serine pyruvate aminotransferase of rat liver. Eur. J. Biochem. **168**:537–542.
39. **Peel, D., and J. R. Quayle.** 1961. Microbial growth on C-1 compounds: isolation and characterization of *Pseudomonas* AM1. Biochem. J. **81**:465–469.
40. **Peri, K. G., C. Belanger, and R. Mackenzie.** 1989. Nucleotide sequence of the human NAD-dependent methylene tetrahydrofolate dehydrogenase-cyclohydrolase. Nucleic Acids Res. **17**:8853.
41. **Popov, V. O., I. A. Shumilin, T. B. Ustinnikova, V. S. Lamzin, and T. S. Egorov.** 1990. NAD-dependent formate dehydrogenase of the methylotrophic bacterium *Pseudomonas sp.* 101. Bioorg. Khim. **16**:324–335. (In Russian.)
42. **Purdue, P. E., Y. Takada, and C. J. Danpure.** 1990. Identification of mutations associated with peroxisome-to-mitochondrion mistargeting of alanine/glyoxylate aminotransferase in primary hyperoxaluria type 1. J. Cell Biol. **111**:2341–2351.
43. **Ruvkun, G. B., and J. R. Ausubel.** 1981. A general method for site-directed mutagenesis in procaryotes. Nature (London) **289**:85–88.
44. **Saito, H., and K.-I. Miura.** 1963. Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim. Biophys. Acta **72**:619–629.
45. **Scrimgeour, K. G., and F. M. Huennekens.** 1963. N5, N10-methylenetetrahydrofolic dehydrogenase. Methods Enzymol. **6**:368–372.
46. **Shannon, K. W., and J. C. Rabinowitz.** 1988. Isolation and characterization of the *Saccharomyces cerevisiae* MIS1 gene encoding mitochondrial C1-tetrahydrofolate synthase. J. Biol. Chem. **263**:7717–7725.
47. **Simon, R., U. Priefer, and A. Puhler.** 1983. Vector plasmids for in vivo and in vitro manipulations of Gram-negative bacteria, p. 98–106. *In* A. Puhler (ed.), Molecular genetics of the bacteria-plant interactions. Springer-Verlag KG, Berlin.
48. **Staben, C., and J. C. Rabinowitz.** 1986. Nucleotide sequence of the *Saccharomyces cerevisiae* ADE3 gene encoding C1-tetrahydrofolate synthase. J. Biol. Chem. **261**:4629–4637.
49. **Stone, S., and P. M. Goodwin.** 1989. Characterization and complementation of mutants of *Methylobacterium* AM1 which are defective in C-1 assimilation. J. Gen. Microbiol. **135**:227–235.
50. **Taguchi, H., and O. Takahisa.** 1991. D-Lactate dehydrogenase is a member of the D-isomer-specific 2-hydroxyacid dehydrogenase family. Cloning, sequencing, and expression in *Escherichia coli* of the D-lactate dehydrogenase gene of *Lactobacillus plantarum*. J. Biol. Chem. **266**:12588–12594.
51. **Takada, Y., N. Kaneko, H. Ezumi, P. E. Purdue, and C. J. Danpure.** 1990. Human peroxisomal L-alanine:glyoxylate aminotransferase. Evolutionary loss of a mitochondrial targeting signal by point mutation of the initiation codon. Biochem. J. **268**:517–520.
52. **Tan, L. U. L., and R. E. MacKenzie.** 1979. Methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase-formyltetrahydrofolate synthetase from porcine liver. Location of the activities in two domains of the multifunctional polypeptide. Can. J. Biochem. **57**:806–812.
53. **Tobey, K. L., and G. A. Grant.** 1986. The nucleotide sequence of the *serA* gene of *Escherichia coli* and the amino acid sequence of the encoded protein, D-3-phosphoglycerate dehydrogenase. J. Biol. Chem. **261**:12179–12183.
54. **Tolbert, N. E.** 1981. Metabolic pathways in peroxysomes and glyoxysomes. Annu. Rev. Biochem. **50**:133–157.
55. **Tsuji, K., H. C. Tsien, R. S. Hanson, S. R. Depalma, R. Scholtz, and S. Laroche.** 1990. 16S ribosomal RNA sequence analysis for determination of phylogenetic relationship among methylotrophs. J. Gen. Microbiol. **136**:1–10.
56. **van der Oost, J., H. S. van Walraven, J. Bogerd, A. B. Smit, G. D. Ewart, and G. D. Smith.** 1989. Nucleotide sequence of the gene proposed to encode the small subunit of the soluble hydrogenase of the thermophilic unicellular cyanobacterium *Synechococcus* PCC6716. Nucleic Acids Res. **17**:10098.
57. **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.
58. **Villar, E., B. Schuster, D. Peterson, and V. Schirch.** 1985. C1-tetrahydrofolate synthase from rabbit liver. Structural and kinetic properties of the enzyme and its two domains. J. Biol. Chem.

260:2245–2252.

59. **Waechter-Brulla, D., A. A. DiSpirito, L. V. Chistoserdova, and M. E. Lidstrom.** 1993. Methanol oxidation genes in the marine methanotroph *Methylomonas* sp. strain A4.
60. **Watson, R. J., and V. K. Rastogi.** 1993. Cloning and nucleotide sequencing of *Rhizobium meliloti* aminotransferase genes: an aspartate aminotransferase required for symbiotic nitrogen fixation is atypical. J. Bacteriol. **175:**1919–1928.
61. **Whitaker, J. R., and P. E. Granum.** 1980. An absolute method for protein determination based on difference in absorbance at 235 and 280 nm. Anal. Biochem. **109:**156–159.
62. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1986. Prediction of the occurrence of the ADP-binding β-α-β fold in proteins, using amino acid sequence fingerprint. J. Mol. Biol. **187:**101–107.
63. **Xu, H. H., M. Viebahn, and R. S. Hanson.** 1993. Identification of methanol-regulated promoter sequences from the facultative methylotrophic bacterium *Methylobacterium organophilum* XX. J. Gen. Microbiol. **139:**743–752.