JOURNAL OF BACTERIOLOGY, Apr. 1999, p. 2244–2251 Vol. 181, No. 7
0021-9193/99/$04.00+0


# Molecular Analysis of a Novel Methanesulfonic Acid Monooxygenase from the Methylotroph *Methylosulfonomonas methylovora*

PAOLO DE MARCO,[1] PEDRO MORADAS-FERREIRA,[1] TIMOTHY P. HIGGINS,[2]† IAN McDONALD,[2] ELIZABETH M. KENNA,[2] AND J. COLIN MURRELL[2]*

*ICBAS and IBMC, University of Porto, 4150 Porto, Portugal,[1] and Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, United Kingdom[2]*

Received 19 October 1998/Accepted 26 January 1999

***Methylosulfonomonas methylovora*** **M2 is an unusual gram-negative methylotrophic bacterium that can grow on methanesulfonic acid (MSA) as the sole source of carbon and energy. Oxidation of MSA by this bacterium is carried out by a multicomponent MSA monooxygenase (MSAMO). Cloning and sequencing of a 7.5-kbp *Sph*I fragment of chromosomal DNA revealed four tightly linked genes encoding this novel monooxygenase. Analysis of the deduced MSAMO polypeptide sequences indicated that the enzyme contains a two-component hydroxylase of the mononuclear-iron-center type. The large subunit of the hydroxylase, MsmA (48 kDa), contains a typical Rieske-type [2Fe–2S] center with an unusual iron-binding motif and, together with the small subunit of the hydroxylase, MsmB (20 kDa), showed a high degree of identity with a number of dioxygenase enzymes. However, the other components of the MSAMO, MsmC, the ferredoxin component, and MsmD, the reductase, more closely resemble those found in other classes of oxygenases. MsmC has a high degree of identity to ferredoxins from toluene and methane monooxygenases, which are enzymes characterized by possessing hydroxylases containing μ-oxo bridge binuclear iron centers. MsmD is a reductase of 38 kDa with a typical chloroplast-like [2Fe–2S] center and conserved flavin adenine dinucleotide- and NAD-binding motifs and is similar to a number of mono- and dioxygenase reductase components. Preliminary analysis of the genes encoding MSAMO from a marine MSA-degrading bacterium, *Marinosulfonomonas methylotropha*, revealed the presence of *msm* genes highly related to those found in *Methylosulfonomonas*, suggesting that MSAMO is a novel type of oxygenase that may be conserved in all MSA-utilizing bacteria.**

Methanesulfonic acid (MSA) is a compound produced by natural processes. It results from the oxidation in the atmosphere of dimethylsulfide (DMS), which is produced by the decomposition of dimethylsulfoniopropionate, an algal osmolyte. Organic sulfur emissions to the atmosphere from the oceans are estimated to be on the order of $4 \times 10^{10}$ kg/year, up to 88% of which is DMS (30). It has been estimated that as much as 50% of the flux of the DMS oxidized in the atmosphere by the action of free radicals ends up in the form of MSA (2, 37). This suggests that approximately $10^{10}$ kg of S as MSA is produced on a global basis per year. MSA is a very hygroscopic compound which participates in the formation of clouds and then falls onto lands and oceans in wet and dry precipitations. MSA has been found to accumulate in the frozen layers of snow of Antarctica (46) and Greenland (58), but nowhere else in the environment can it be found in detectable levels. Since MSA is chemically very stable, the case for the existence of microbial MSA degradation is very strong. MSA ($CH_3SO_3H$) is the simplest of the sulfonates and is a substrate for the growth of certain methylotrophic microorganisms. The first bacterium isolated on MSA as the sole source of carbon and energy was the facultative methylotroph *Methylosulfonomonas methylovora* M2 (5, 28). It was isolated from garden soil after enrichment. Studies already published (23, 29) have described the physiology of MSA metabolism by strain M2. MSA is oxidized by strain M2 to formaldehyde, which can then be either assimilated through the serine cycle or fully oxidized (via formate) to yield $CO_2$ and $H_2O$ in order to produce reducing power and energy for the cell. Cell extracts of MSA-grown strain M2 showed an MSA-dependent, $Fe^{2+}$- and flavin adenine dinucleotide (FAD)-stimulated NADH-oxidase activity. This activity was shown to be inducible and was absent when *M. methylovora* M2 was grown on any other substrates, such as methanol or formate. It was shown that the inducible enzyme responsible for the oxidation of MSA was a multicomponent monooxygenase, MSA monooxygenase (MSAMO). $O_2$ consumption studies and cell extract assays demonstrated that MSAMO has a restricted substrate range that includes only the short-chain aliphatic sulfonates (methane- to butanesulfonate) and excludes all larger molecules, such as arylsulfonates and aromatic sulfonates. Specific activities of the enzyme in cell extracts with a range of substrates have been previously reported (23).

Another MSA-degrading methylotroph has been enriched and isolated from the marine environment (57). This bacterium, *Marinosulfonomonas methylotropha* PSCH4, was found to degrade MSA in a fashion identical to that described for the soil isolate strain M2, possessed an inducible multicomponent enzyme that resembles MSAMO, and also assimilated formaldehyde via the serine pathway. Despite all these similarities, the two strains originated from very different environments, and 16S ribosomal DNA analysis revealed that they belong to only distantly related genera, lying on separate branches within the α subgroup of the *Proteobacteria* (24).

The MSAMO appears from its substrate specificity to be a unique oxygenase with a rather narrow substrate range. However, the enzyme is rather unstable in cell extracts, and to date, it has not been possible to purify all the components of MSAMO to homogeneity. The work described here was initiated

* Corresponding author. Mailing address: Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom. Phone: 44 1203 523553. Fax: 44 1203 523568. E-mail: CM@dna.bio.warwick.ac.uk.

† Present address: School of Animal and Microbial Sciences, University of Reading, Reading, RG6 6AJ, United Kingdom.

in order to confirm at the molecular level that this is a unique oxygenase enzyme system which is present in all bacteria that can utilize MSA as the sole carbon and energy source.

## MATERIALS AND METHODS

**Materials.** Except where otherwise stated, all chemicals were of analytical grade and were supplied by Aldrich Chemical Co., Gillingham, Dorset, United Kingdom, or Sigma Chemical Co., Poole, Dorset, United Kingdom.

**Growth of the organism.** *M. methylovora* M2 (5) was cultivated and maintained on mineral salts medium Min E (29) containing 20 mM MSA.

**N-terminal sequencing.** Sequencing was carried out by the method already described in reference 22 with the modifications used by Matsudaira (35).

**DNA extraction.** Genomic DNA was extracted from cells of strain M2, grown as described in reference 29, by the lysozyme-Sarkosyl lysis-CsCl method, as described previously (43). Preparations of recombinant plasmid DNA from *Escherichia coli* were obtained by the method of Saunders and Burke (49).

**DNA cloning.** Genomic DNA from strain M2 was digested by using restriction enzymes provided by BRL according to the manufacturer's instructions. *Sph*I-digested pUC19 DNA was treated with calf intestine alkaline phosphatase (Boehringer Mannheim) as recommended by the manufacturer. Size-fractionated genomic DNA from strain M2 was obtained by running *Sph*I-digested DNA in an agarose gel, cutting out the DNA fragments corresponding to the desired size range, and electroeluting the DNA from the agarose slice in dialysis tubing. The resulting DNA was purified by treatment with phenol-chloroform-isoamyl alcohol (25:24:1) and precipitation with ethanol. Ligation to pUC19 DNA, treated as described above, was performed by using BRL ligase according to the manufacturer's instructions. Competent cells of *E. coli* INV$\alpha$F′ (Invitrogen Corporation) were used as recipients for transformation with the ligation mixes. White Ap$^r$ colonies were picked and replicated onto nylon membranes (Hybond-N; Amersham). After the growth of colonies, cells were lysed in situ by the method of Grunstein and Hogness as described in reference 47. The DNA was fixed by exposure to UV light in a UV Stratalinker 2400 (Stratagene). The custom-made oligonucleotides 9R and 3F were used as probes in the following way: 100 pmol of each was labelled by using 1.85 MBq (50 $\mu$Ci) of [$\alpha$-$^{32}$P]ATP (specific activity, 259 TBq/mmol at a concentration of 0.74 MBq/$\mu$l; Amersham) and T4 kinase (BRL) according to the manufacturer's instructions. Filters containing 2,600 recombinant clones from the cloning experiment described above were prehybridized and hybridized in a Hybaid oven (Mini 10) under the conditions suggested in reference 47. Final washes were performed at increasing temperatures (58 to 65°C) in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) (47). Autoradiographs were obtained by exposing the radioactive filters to Fuji RX films for appropriate times (e.g., 8 to 24 h).

**DNA sequencing and sequence analysis.** Purified plasmid DNA was used as a template for sequencing. Synthetic custom-made oligonucleotides were used for primer walking. The *Taq* DyeDeoxy Terminator Cycle Sequencing Kit and a model 373A DNA Sequencing System gel apparatus (both from Applied Biosystems) were used according to the manufacturer's instructions. Analyses of DNA sequences were performed on a Sun Workstation running the Genetics Computer Group (GCG) Wisconsin Package, version 8.0.1-Unix (September 94). BLAST homology searches (1) were performed by using the Internet facility available at the National Center for Biotechnology Information at the National Institutes of Health, Bethesda, Md. (39a). Searches of the SwissProt database using the ScanProsite algorithm (4) were performed by Internet link to the ExPaSy Molecular Biology server of the Swiss Institute of Bioinformatics (53a).

**Protein sequence comparisons.** Protein sequences were aligned by using the PILEUP program (GCG), and the alignments were edited manually. Comparisons based on the alignments and the resulting phylogenetic trees were produced by using the programs SEQBOOT, PROTPARS, PROTDIST, NEIGHBOR, and CONSENSE contained in PHYLIP (Phylogenetic Inference Package), version 3.572c (18).

**Nucleotide sequence accession number.** The sequence of the entire 7.5-kbp *Sph*I DNA fragment of strain M2 was determined as described above and deposited in GenBank under accession no. AF091716.

## RESULTS

**Cloning of the *msm* genes from *M. methylovora* M2.** Cell extracts of *M. methylovora* M2 were examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) after growth on MSA or methanol. Five MSA-induced soluble polypeptides of approximately 50, 45, 35, 20, and 16 kDa were detected (Fig. 1). Ion-exchange chromatography of cell extracts from MSA-grown cells yielded three fractions, designated A, B, and C in order of elution, each of which was essential for reconstitution of activity in vitro (22). Fraction A, after further partial purification by gel filtration, yielded a major polypeptide of around 38 kDa which had spectral characteristics of an FAD-containing protein. This polypeptide, when analyzed by gel zymography, showed an FAD-dependent nitroblue tetrazolium-reducing activity, indicating that it was a reductase. However, this reductase component is very unstable, and it has not been possible to purify this component in active form. Fraction B was also subjected to purification by gel filtration. This yielded two major polypeptides of around 50 and 20 kDa. Preliminary evidence suggests that this component is likely to be the hydroxylase of MSAMO and contains Fe and a Rieske iron center (44a). Fraction C was further purified by gel filtration and MonoQ ion-exchange chromatography. The resulting sample was reddish and was shown to contain a polypeptide that had an apparent $M_r$ of 16,000 in SDS-PAGE. This was shown to be a ferredoxin (22). The 45-kDa MSA-induced soluble polypeptide was also purified to homogeneity by a combination of gel filtration and ion-exchange chromatography, although clearly it was not copurifying with the MSAMO enzyme.

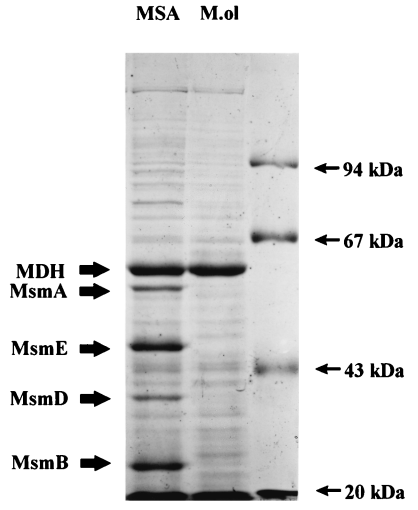


FIG. 1. SDS-PAGE profile of cell extracts of *M. methylovora* M2 grown on MSA or methanol (M.ol) (12). MDH is the large subunit of methanol dehydrogenase. MsmC at 16 kDa is running with the dye front (ca. 20 kDa). Polypeptide molecular mass markers are phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), and soybean trypsin inhibitor (20.1 kDa).

N-terminal amino acid sequences were obtained for fractions B (the 20-kDa polypeptide) and C and for the 45-kDa polypeptide. The degenerate oligonucleotide 3F [TTCGG(C/T)AAG CCGGG(C/T)GA(A/G)AAGGT(C/G)GACCT] was designed by back-translation of the N terminus of the 45-kDa polypeptide. Oligonucleotide 9R [ATGTCGTT(C/G)GC(C/G)AC (C/G)AGCTT] was designed from the N-terminal sequence of protein C. The codon usage table for *M. methylovora*, used to reduce the degeneracy of these oligonucleotides, had become available after the cloning of a near-complete *glnA* (glutamine synthetase) gene (13). Both oligonucleotides 3F and 9R were used as probes in cloning experiments.

A total of 2,600 recombinant clones containing *Sph*I DNA fragments of strain M2 were screened, and one that gave a positive signal when challenged with either probe 3F or probe 9R was selected and analyzed. It contained a recombinant plasmid, designated pDM5, consisting of pUC19 with an *Sph*I insert of around 7.5 kbp (Fig. 2). A preliminary restriction map was obtained, and subclones containing a variety of DNA restriction fragments were produced. These subclones were used to prepare DNA for double-stranded sequencing. The sequence of the entire 7.5-kbp *Sph*I DNA fragment was determined as described in Materials and Methods.

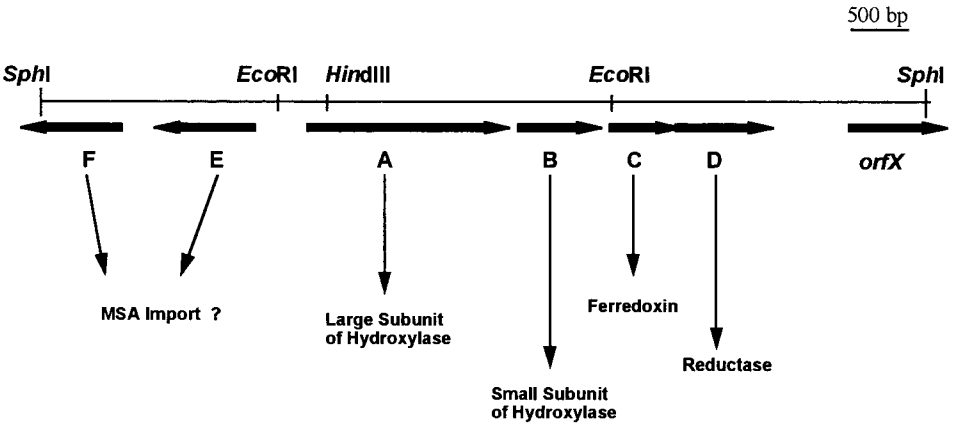


FIG. 2. Restriction map of the insert of plasmid pDM5. The length of the whole *Sph*I fragment is 7,509 bp.

**DNA sequence analysis of the 7.5-kbp *Sph*I DNA fragment of strain M2.** The sequence was analyzed by using the BLAST algorithm (1). Seven open reading frames (ORFs), two of which are incomplete, were identified within this sequence. Four ORFs, designated *msmABCD*, showed high degrees of identity to known mono- and dioxygenase enzyme components. Two other ORFs (not described in this work), *msmE* and *msmF* (incomplete), showed similarity to components of bacterial ABC membrane transport systems. One incomplete ORF, designated *orfX* (extending from nucleotide [nt] 6907 to the end of the GenBank sequence), has significant similarity to the sequence of bacterial transcriptional positive regulators of the LysR family. The G+C content of this *Sph*I fragment of the chromosome of strain M2 is 64.8 mol%, which is in good agreement with the experimental datum (61 mol%) obtained with total DNA from this strain (29). The N-terminal sequences of predicted proteins MsmB, MsmC, and MsmE (the last is not reported in this paper) are in agreement with the corresponding sequences obtained by Edman degradation of the MSA-specific polypeptides of 20, 16, and 45 kDa, respectively.

**Analysis of the genes *msmABCD*.** The genes *msmA*, *msmB*, *msmC*, and *msmD* are transcribed in the same orientation, whereas *msmE* and *msmF* are transcribed in the opposite direction (Fig. 2). The intergenic region between *msmA* and *msmB* covers 135 nt, and that between *msmB* and *msmC* is 107 nt. The 3′ end of *msmC* overlaps with the beginning of *msmD* by 23 nt. Putative ribosome binding sites (51) were identified at locations 5 to 8 nt upstream of each of genes *msmABCD* (Table 1). Downstream of the 3′ end of *msmD* (nt 5909 to 5939 and nt 5984 to 6018 of the GenBank sequence) two inverted repeats, which are probably involved in the termination of translation, are present (Table 1). Genes *msmABCD* are probably transcribed into a single mRNA and, as such, constitute an operon for the coordinated expression of MSAMO.

The analysis of these four ORFs reveals a codon usage which is very similar to that found in *glnA* of *M. methylovora* and to that found in the soluble methane monooxygenase genes of *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* Bath (39). Excluding start and stop codons, more than 87% of all codons in genes *msmABCD* end with a G or a C. In each group of codons encoding the same amino acid, the bias for a C or a G in the third position ranges between 62 and 100%. In all four cases, the stop codon used is UGA.

**Analysis of gene *msmA*.** *msmA* codes for a polypeptide of 414 amino acids (aa) which shows significant identity to the iron-sulfur-containing large ($\alpha$) subunits of the hydroxylase components of several known mono- and dioxygenases. The predicted protein has a pI of 6.73 and an $M_r$ of 48,473. Alignments of the sequence of the predicted MsmA polypeptide with similar hydroxylase ($\alpha$) subunits revealed that, beside some regions and residues of clear conservation, MsmA also has some quite novel characteristics.

A sequence starting at residue 85 clearly resembles those found to ligate Rieske-type [2Fe–2S] centers in other proteins involved in the transfer of electrons (CXH–$X_{16-17}$–CXXH) (34). However, in this case the sequence intervening between the two cysteine-histidine groups is unusually long (26 residues) (Fig. 3a). A search of the SwissProt database using the ScanProsite algorithm (4) indicated that very few known Rieske-type proteins have an intervening sequence longer than 18 aa. Three of the examples found are nitrite reductases of fungi, which have 22-aa spacers. Two other examples are bacterial iron-sulfur subunits of cytochrome *c* reductases (spacers of 21 aa). The other bacterial proteins found in this search, with spacer sequences that in some cases reach a length of 26 aa (as in MsmA), are all hypothetical translations of DNA sequences with no known function.

Preliminary cloning and sequencing data from the marine MSA-degrading bacterium *M. methylotropha* PSCH4 revealed the presence of *msm* genes highly related to those found in the soil isolate strain M2 (50). In particular, the sequence of the large subunit of the hydroxylase has a Rieske-like motif with a 26-residue spacer region that is a perfect match with the corresponding sequence of MsmA.

Other motifs and single residues are conserved throughout

TABLE 1. 5′ Ends of genes *msmA*, *msmB*, *msmC*, and *msmD*

| Gene[a] | nt position[b] | Sequence[c] |
|---|---|---|
| *msmA* | 2392–2411 | GAGCAGAGGAAGCGCCG**ATG** |
| *msmB* | 3772–3795 | CGCGAGGAAACGAGAACCGCA**ATG** |
| *msmC* | 4424–4443 | ACACAAAAGGCGCCCTG**ATG** |
| *msmD* | 4773–4798 | CGAAGAAGAAGACGACATGGACG**ATG** |
| InvRep1 | 5909–5939 | TGCGGCCTCAGGCGAACTGCCTGAGGCCGAG |
| InvRep2 | 5984–6018 | CCACGCGCCCGAGCGCCTCGCGCGACAGCGCATGG |

[a] InvRep, inverted repeat.
[b] Numbering in the GenBank sequence.
[c] Boldfaced letters, start codons. Underlining denotes putative ribosome binding sites in *msmABCD* and complementary sequences in inverted repeats.

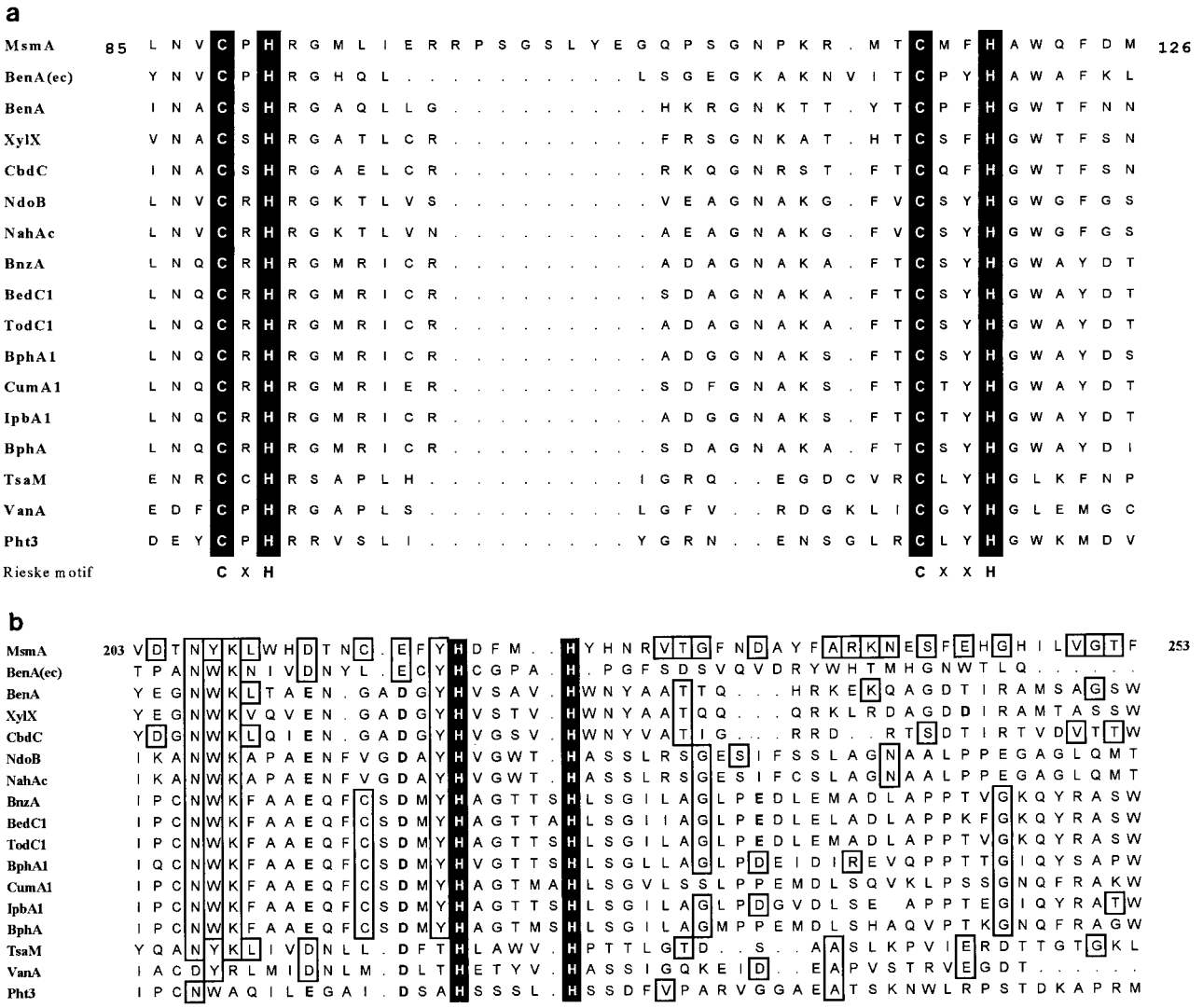


FIG. 3. Alignment of large (α) hydroxylase components in the regions binding the Rieske-type [2Fe–2S] center (a) and in the region purported to bind the mononuclear Fe center (b). MsmA, MSAMO (this work); NdoB, naphthalene dioxygenase (32); NahAc, naphthalene dioxygenase (52); BnzA, benzene 1,2-dioxygenase (26); BedC1, benzene dioxygenase (55); TodC1, toluene dioxygenase (62); BphA1, biphenyl dioxygenase (3); CumA1, cumene dioxygenase (accession no. D37828) (43a); IpbA1, isopropylbenzene dioxygenase (31); BphA, biphenyl dioxygenase (54); BenA, benzoate dioxygenase (40); XylX, toluate dioxygenase (21); CbdC, 2-halobenzoate 1,2-dioxygenase (19); BenA(ec), *E. coli* homologue of benzoate 1,2-dioxygenase (6); TsaM, toluenesulfonate methyl-monooxygenase (27); VanA, vanillate-*O*-demethylase (8); Pht3, phthalate dioxygenase (41). The numbering refers to the amino acid sequence of MsmA.

the rest of the sequence of MsmA, as can be seen in the alignment in Fig. 3b. Particularly evident is the conservation of those residues (histidines and aspartates/glutamates) that are likely to be involved in ligating a mononuclear iron center where $O_2$ is reduced and activated.

Other regions of MsmA show a clear divergence with the other known hydroxylase α-subunits. Since the large hydroxylase subunits are believed to harbor the site for substrate recognition (44), it is not surprising that MsmA has sections of sequence quite divergent from the sequences of the other homologues; MSA is a small, charged sulfur compound, very different from the aromatic molecules that serve as the substrates of most of the other known enzymes.

Phylogenetic analysis was carried out on an alignment of 51 hydroxylase large subunits. Maximum parsimony and distance matrix (Dayhoff-PAM) followed by neighbor-joining methods were used, excluding all gap-containing columns of the alignment. MsmA was found to be only loosely associated with an *E. coli* homologue of benzoate 1,2-dioxygenase (Fig. 3) (6), with low bootstrap scores. The topology of the tree as a whole showed a high dependence on the algorithm chosen for the analysis. The inclusion of the MSAMO-specific spacer region of the Rieske motif of MsmA in the alignment did not give clearer results in the phylogenetic analysis. To demonstrate the relatedness of MsmA to its homologues, we report in Table 2 the identity and similarity values of five of the highest-scoring hits in the BLAST search.

**Analysis of gene *msmB*.** The predicted amino acid sequence of MsmB is a 181-aa polypeptide of 20,478 Da with a pI of 5.58. Its sequence shows, at least in its N-terminal part, similarity with a few small (β) subunits of terminal hydroxylase components of oxygenases (identity and similarity values are given in Table 3). The lack of similarity found in the C-terminal region led us to check the other two forward reading frames to make sure that a sequencing error was not the cause of such a result. Neither of the two alternative predicted C termini showed

TABLE 2. Identity and similarity values obtained by BLAST search using the amino acid sequence of MsmA

| Large (α) subunit of hydroxylase[a] | Identity (%) | Similarity (%) | Size of region compared (aa) |
|---|---|---|---|
| IpbA1 | 32 | 51 | 198 |
| BphA1 | 32 | 50 | 198 |
| BnzA | 31 | 50 | 198 |
| CumA1 | 30 | 49 | 198 |
| BenA | 26 | 43 | 324 |

[a] See the legend to Fig. 3 for references for each sequence.

similarity with known oxygenase β-subunits. The solidity of the sequence of our predicted MsmB was also confirmed by "third-position GC bias," "codon preference," and "rare codon frequency" analyses. All these methods confirmed that the reading frame chosen was continuous and coherent and clearly scored better than the other two. It has been noted before that β-subunits are less conserved than their α counterparts (40). Recent work (44) suggested that these small hydroxylase subunits do not participate in the constitution of the active site of the oxygenases but rather provide an external structure that holds the α-subunits in place. This kind of function may mean that the constraints on the protein sequence are less strong in these subunits than in the rest of the enzyme.

**Analysis of gene *msmC*.** We have previously described the biochemical and molecular characteristics of MsmC (22), but for the sake of completeness its properties are briefly described here. MsmC has the characteristics of a ferredoxin. It is a small acidic protein of 122 aa with an $M_r$ of 13,748 and a pI of 3.9. Its sequence shows a canonical motif of Rieske-type [2Fe–2S] center-binding proteins (CXH–$X_{17}$–CXXH) and is similar to those of other known bacterial ferredoxins. Phylogenetic comparisons and treeing using various algorithms showed that MsmC is most closely related to TmoC and TbuB, ferredoxins of the toluene-3-monooxygenase of *Pseudomonas pickettii* (9) and the toluene-4-monooxygenase of *Pseudomonas mendocina* (60), respectively. These two enzymes belong, together with soluble methane monooxygenase (sMMO), to a separate group of oxygenases characterized by hydroxylases containing μ-oxo bridge binuclear iron centers.

**Analysis of gene *msmD*.** *msmD* codes for a 366-aa polypeptide with significant identity to reductase components of known oxygenases. The predicted polypeptide, MsmD, has an $M_r$ of 38,852 and a pI of 6.51. It is similar to reductases that have chloroplast-like [2Fe–2S] centers, and a motif very similar to that found in this kind of protein (CXXXXCXXC–$X_{29}$–C) (34) can be found in MsmD at residue 56. In MsmD, the region between the last two C residues of this motif is 31 rather than 29 aa long, but this kind of variation is also found in many other known plant-like [2Fe–2S]-binding motifs. Other conserved residues can be found along the sequence of MsmD, particularly in those regions that Byrne et al. (9) and Neidle et al. (40) propose as FAD- and NAD-binding motifs (see the alignment in Fig. 4 and the identity and similarity values in Table 4).

## DISCUSSION

The genes coding for MSAMO of *M. methylovora* M2 have been identified and characterized. The four polypeptides comprising MSAMO are almost certainly the products of the coordinated expression of an operon (*msmABCD*), as found for many other bacterial oxygenase systems.

Aromatic dioxygenases have already been classified on the basis of their quaternary structure (hydroxylase in the form of a homo- or heteropolymeric protein; the presence or absence of a free ferredoxin), the type (plant-like or Rieske-type) of [2Fe–2S] center(s) present in the electron transfer chain components, and the cofactor used by the reductase (flavin mononucleotide or FAD) (34). This type of classification, in our view, can be extended to include many nonaromatic oxygenases and, indeed, most of the known monooxygenases. Probably one more class should be added to include a few more such enzymes in this classification scheme. Based on the characteristics of their hydroxylases, enzymes like sMMO (10, 53), toluene-4-monooxygenase (60), toluene-3-monooxygenase (9), and phenol hydroxylase (42) could be placed into a new class (IV). These enzymes have similar hydroxylases which bind and activate the $O_2$ molecule at a binuclear iron center (where the two Fe atoms may be linked by a μ-oxo bridge). The other type of known bacterial oxygenases have hydroxylases with a mononuclear iron center. The hydroxylase of MSAMO is, by sequence similarity, of the mononuclear-iron-center type. The electron transfer chain elements of MSAMO (the reductase MsmD and the ferredoxin MsmC), however, resemble more closely those found in oxygenases of class IV. A similar situation is also found in the *benABC* (40), *carAB* (48), and *xylXYZ* (21) gene clusters.

MSAMO is just another example in support of the theory (20, 34) that these oxygenase enzyme complexes have evolved not only by means of mutation but also by exchange and reshuffling of homologous elements to yield a variety of combinations with the ability to degrade a wide variety of compounds.

The discovery of aromatic-degrading oxygenases in gram-positive bacteria (3, 33) showing high degrees of relatedness to their counterparts from gram-negative bacteria demonstrates that lateral gene transfer, even between organisms only distantly related, may be another important factor in the evolution of these enzymes.

Bacteria that degrade MSA are probably ubiquitous. Strain M2 was isolated from a British soil sample; one different strain was enriched from a North Sea water sample; more recently, two further strains belonging to the genera *Methylobacterium* and *Hyphomicrobium*, respectively, were isolated from a Portuguese soil sample (14). Preliminary data from the cloning and sequencing of the MSAMO genes of the marine strain (*M. methylotropha* PSCH4) show a very high level of conservation (identity levels between 60.6 and 83.8%) with the sequence obtained from strain M2 (50). Preliminary DNA hybridization experiments with the Portuguese soil isolates show that these too have genes with a high degree of identity to the *msm* operon of strain M2 (14). These data make a very strong case for the hypothesis that a conserved MSAMO enzyme is present in a variety of natural bacterial strains. If this is the

TABLE 3. Identity and similarity values obtained by BLAST search using the sequence of MsmB

| Small (β) subunit of hydroxylase[a] | Identity (%) | Similarity (%) | Size of region compared (aa) |
|---|---|---|---|
| TftB | 29 | 48 | 79 |
| XylY | 26 | 50 | 73 |
| IpbAb | 35 | 53 | 47 |
| CumA2 | 41 | 51 | 38 |
| CmtAc | 23 | 44 | 114 |

[a] TftB, trichlorophenoxyacetic acid oxygenase (11); XylY, toluate dioxygenase (21); IpbAb, isopropylbenzene dioxygenase (16); CumA2, cumene dioxygenase (accession no. D37828) (43a); CmtAc, *p*-cumate dioxygenase (15).

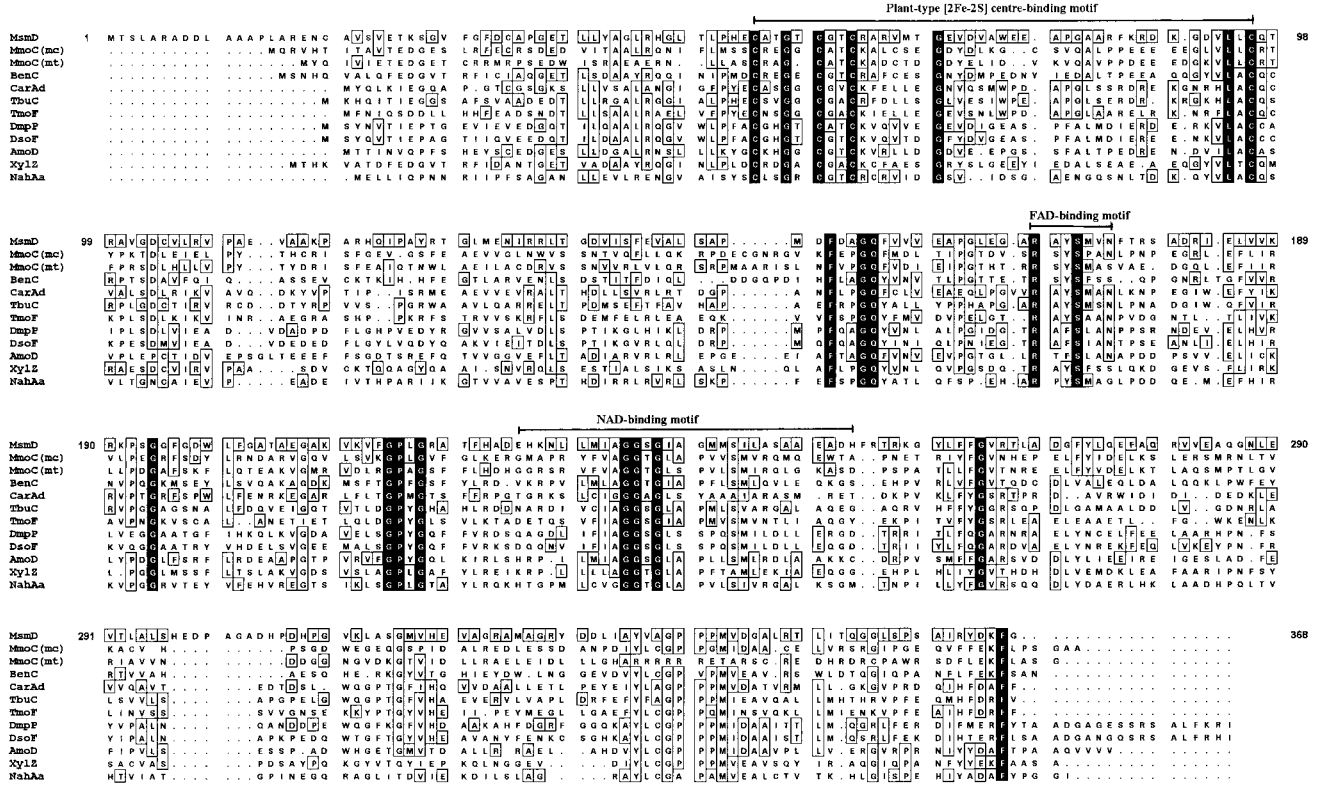


FIG. 4. Alignment of the reductase of MSAMO, MsmD, with its homologues. MmoC(mc), sMMO of *M. capsulatus* Bath (53); MmoC(mt), sMMO of *M. trichosporium* OB3b (10); BenC, benzoate dioxygenase (40); CarAd, carbazole dioxygenase (48); TbuC, toluene-3-monooxygenase (9); TmoF, toluene-4-monooxygenase (56); DmpP, phenol hydroxylase (42); DsoF, DMS monooxygenase (25); AmoD, alkene monooxygenase (45); XylZ, xylene monooxygenase (21); NahAa, naphthalene dioxygenase (52). The numbering refers to the amino acid sequence of MsmD.

case, then the distribution and activity of MSA-degrading bacteria can be studied directly in environmental samples by applying techniques of molecular biology in order to avoid the laborious and biased step of enrichment in the laboratory. Studies of this kind have now been performed by using a few other highly conserved enzymes such as methane monooxygenase (36, 38), polychlorinated biphenyl oxygenase (17), mercury reductase (7), and nitrogenase (61). These studies have shown that molecular biology can be used to obtain information about the diversity and distribution of organisms that are present in the environment but do not withstand the artificial culture conditions of the lab. Similar techniques can now be carried out by using the MSAMO genes, which offer a distinct advantage with these types of methylotrophs, which are very difficult to cultivate in the laboratory (57).

MSAMO, although possessing only monooxygenase activity (22, 23), appears to be a hybrid enzyme complex. It comprises a two-component hydroxylase (MsmA and MsmB), which is most similar in amino acid sequence to dioxygenases containing Rieske iron centers, whereas the electron transfer chain elements (MsmC and MsmD) are structurally more similar to those found in a number of well-documented monooxygenase enzymes. However, the distinction between monooxygenases and dioxygenases is not absolute (discussed in references 20 and 34). For example, both toluene and naphthalene dioxygenases catalyze the oxidation of indene and indan to 1-indenol and 1-indanol, respectively (58). Another example is the dihydroxylation of the vinylic side chain of 4-methoxy styrene by the rather nonspecific oxygenase 4-methyl benzoate monooxygenase (putidamonooxin), with both atoms being derived from molecular oxygen (56). Spectral characteristics indicate the presence of iron and a Rieske-type iron center in crude hydroxylase preparations (44a). Unfortunately, the hydroxylase and reductase components of MSAMO are very unstable in cell extracts and to date have not been purified to homogeneity in active form, and therefore it is not possible to make crystals of MSAMO and determine its X-ray crystal structure. Once this is achieved, it will be possible to determine the role that the unique sequences centered around the Rieske center of MsmA may play in the catalysis of sulfonated alkanes.

TABLE 4. Identity and similarity values obtained by BLAST search using the sequence of MsmD

| Reductase[a] | Identity (%) | Similarity (%) | Size of region compared (aa) |
|---|---|---|---|
| CarAd | 31 | 46 | 330 |
| TbuC | 29 | 44 | 329 |
| DmpP | 29 | 43 | 327 |
| DsoF | 27 | 44 | 326 |
| XylZ | 26 | 43 | 343 |

[a] See the legend to Fig. 4 for references.

## ACKNOWLEDGMENTS

We thank the NERC (grant GR3/8242), the EC MAST program (grant CEC 910604), and the Junta Nacional para a Investigação Científica e Tecnológica, Portuguese Ministry of Science and Technology (grant Praxis XXI/BPD/9977/96), for financial support.

## REFERENCES


1. **Altshul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Andreae, M. O.** 1986. The ocean as a source of atmospheric sulfur compounds, p. 331–362. *In* P. Buat-Menard (ed.), The role of air-sea exchange in geochemical cycling. Reidel, New York, N.Y.
3. **Asturias, J. A., E. Diaz, and K. N. Timmis.** 1995. The evolutionary relationship of biphenyl dioxygenase from Gram-positive *Rhodococcus globerulus* P6 to multicomponent dioxygenases from Gram-negative bacteria. Gene **156:** 11–18.
4. **Bairoch, A., and R. Apweiler.** 1998. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. Nucleic Acids Res. **26:** 38–42.
5. **Baker, S. C., D. P. Kelly, and J. C. Murrell.** 1991. Microbial degradation of methanesulphonic acid: a missing link in the biogeochemical sulphur cycle. Nature **350:**627–628.
6. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1474.
7. **Bruce, K. D., A. M. Osborn, A. J. Pearson, P. Strike, and D. A. Ritchie.** 1995. Genetic diversity within *mer* genes directly amplified from communities of noncultivated soil and sediment bacteria. Mol. Ecol. **4:**605–612.
8. **Brunel, F., and J. Davison.** 1988. Cloning and sequencing of *Pseudomonas* genes encoding vanillate demethylase. J. Bacteriol. **170:**4924–4930.
9. **Byrne, A. M., J. J. Kukor, and R. H. Olsen.** 1995. Sequence analysis of the gene cluster encoding toluene-3-monooxygenase from *Pseudomonas pickettii* PK01. Gene **154:**65–70.
10. **Cardy, D. L. N., V. Laidler, G. P. C. Salmond, and J. C. Murrell.** 1991. Molecular analysis of the methane monooxygenase (MMO) gene cluster of *Methylosinus trichosporium* OB3b. Mol. Microbiol. **5:**335–342.
11. **Danganan, C. E., R. W. Ye, D. L. Daubaras, L. Xun, and A. M. Chakrabarty.** 1994. Nucleotide sequence and functional analysis of the genes encoding 2,4,5-trichlorophenoxyacetic acid oxygenase in *Pseudomonas cepacia* AC1100. Appl. Environ. Microbiol. **60:**4100–4106.
12. **Davey, M.** 1996. Ph.D. thesis. University of Warwick, Coventry, United Kingdom.
13. **De Marco, P.** 1996. Ph.D. thesis. University of Warwick, Coventry, United Kingdom.
14. **De Marco, P., A. A. Bordalo e Sá, P. Moradas-Ferreira, and J. C. Murrell.** Unpublished data.
15. **Eaton, R. W.** 1996. *p*-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the *cmt* operon. J. Bacteriol. **178:**1351–1362.
16. **Eaton, R. W., and K. N. Timmis.** 1986. Characterization of a plasmid-specified pathway for catabolism of isopropylbenzene in *Pseudomonas putida* RE204. J. Bacteriol. **168:**123–131.
17. **Erb, R. W., and I. Wagner-Döbler.** 1993. Detection of polychlorinated byphenyl degradation genes in polluted sediments by direct DNA extraction and polymerase chain reaction. Appl. Environ. Microbiol. **59:**4065–4073.
18. **Felsenstein, J.** 1995. Phylogeny inference package, version 3.57c. University of Washington, Seattle.
19. **Haak, B., S. Fetzner, and F. Lingens.** 1995. Cloning, nucleotide sequence, and expression of the plasmid-encoded genes for the two-component 2-halobenzoate 1,2-dioxygenase from *Pseudomonas cepacia* 2CBS. J. Bacteriol. **177:**667–675.
20. **Harayama, S., M. Kok, and E. L. Neidle.** 1992. Functional and evolutionary relationships among diverse oxygenases. Annu. Rev. Microbiol. **46:**565–601.
21. **Harayama, S., M. Rekik, A. Bairoch, E. L. Neidle, and L. N. Ornston.** 1991. Potential DNA slippage structures acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal *benABC* and *Pseudomonas putida* TOL pWW0 plasmid *xylXYZ*, genes encoding benzoate dioxygenases. J. Bacteriol. **173:**7540–7548.
22. **Higgins, T. P., P. De Marco, and J. C. Murrell.** 1997. Purification and molecular characterization of the electron transfer protein of methanesulfonic acid monooxygenase. J. Bacteriol. **179:**1974–1979.
23. **Higgins, T. P., M. Davey, J. Trickett, D. P. Kelly, and J. C. Murrell.** 1996. Metabolism of methanesulfonic acid involves a multicomponent monooxygenase system. Microbiology **124:**251–260.
24. **Holmes, A. J., D. P. Kelly, S. C. Baker, A. S. Thompson, P. De Marco, E. M. Kenna, and J. C. Murrell.** 1997. *Methylosulfonomonas methylovora* gen. nov., sp. nov., and *Marinosulfonomonas methylotropha* gen. nov., sp. nov.: novel methylotrophs able to grow on methanesulfonic acid. Arch. Microbiol. **167:** 46–53.
25. **Horinouchi, M., K. Kasuga, H. Nojiri, H. Yamane, and T. Omori.** 1997. Cloning and characterization of genes encoding an enzyme which oxidizes dimethyl sulfide in *Acinetobacter* sp. strain 20B. FEMS Microbiol. Lett. **155:** 99–105.
26. **Irie, S., S. Doi, T. Yorifuji, M. Takagi, and K. Yano.** 1987. Nucleotide sequencing and characterization of the genes encoding benzene oxidation enzymes of *Pseudomonas putida*. J. Bacteriol. **169:**5174–5179.
27. **Junker, F., R. Kiewitz, and A. M. Cook.** 1997. Characterization of the *p*-toluenesulfonate operon *tsaMBCD* and *tsaR* in *Comamonas testosteroni* T-2. J. Bacteriol. **179:**919–927.
28. **Kelly, D. P., and S. C. Baker.** 1990. The organosulphur cycle: aerobic and anaerobic processes leading to turnover of organic C1-sulphur compounds. FEMS Microbiol. Rev. **87:**241–246.
29. **Kelly, D. P., S. C. Baker, J. Trickett, M. Davey, and J. C. Murrell.** 1994. Methanesulphonate utilization by a novel methylotrophic bacterium involves an unusual monooxygenase. Microbiology **140:**1419–1426.
30. **Kelly, D. P., and N. A. Smith.** 1990. Organic sulfur compounds in the environment. Biogeochemistry, microbiology and ecological aspects. Adv. Microb. Ecol. **11:**345–385.
31. **Kesseler, M., E. R. Dabbs, B. Averhoff, and G. Gottschalk.** 1996. Studies on the isopropylbenzene 2,3-dioxygenase and the 3-iso-propylcatechol 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. Microbiology **142:**3241–3251.
32. **Kurkela, S., H. Lehvaslaiho, E. T. Palva, and T. H. Teeri.** 1988. Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB9816. Gene **73:**355–362.
33. **Masai, E., A. Yamada, J. M. Healy, T. Hatta, K. Kimbara, M. Fukuda, and K. Yano.** 1995. Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **61:**2079–2085.
34. **Mason, J. R., and R. Cammack.** 1992. The electron transport proteins of hydroxylating bacterial dioxygenases. Annu. Rev. Microbiol. **46:**277–305.
35. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **21:**10035–10038.
36. **McDonald, I. R., E. M. Kenna, and J. C. Murrell.** 1995. Detection of methanotrophic bacteria in environmental samples with the PCR. Appl. Environ. Microbiol. **61:**116–121.
37. **Mihalopoulos, N., B. C. Nguyen, C. Boissard, J. M. Campin, J. P. Putaud, S. Belviso, I. Barnes, and K. H. Becker.** 1992. Field studies of dimethylsulfide oxidation in the boundary layer: variations of dimethylsulfide, methanesulfonic acid, sulfur dioxide, non-sea-salt sulfate and Aitken nuclei at a coastal site. J. Atm. Chem. **14:**459–477.
38. **Murrell, J. C., V. McGowan, and D. L. N. Cardy.** 1993. Detection of methylotrophic bacteria in natural samples by molecular probing techniques. Chemosphere **26:**1–11.
39. **Murrell, J. C.** 1992. Genetics and molecular biology of methanotrophs. FEMS Microbiol. Rev. **88:**233–248.

39a. **National Center for Biotechnology Information Website.** September 1998, copyright date. [Online.] http://www.ncbi.nlm.nih.gov. [1 October 1998, last date accessed.]

40. **Neidle, E. L., C. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama.** 1991. Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. J. Bacteriol. **173:**5385–5395.
41. **Nomura, Y., M. Nakagawa, N. Ogawa, S. Harashima, and Y. Oshima.** 1992. Genes in PHT plasmid encoding the initial degradation pathway of phthalate in *Pseudomonas putida*. J. Ferment. Bioeng. **74:**333–344.
42. **Nordlund, I., J. Powlowski, and V. Shingler.** 1990. Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J. Bacteriol. **172:**6826–6833.
43. **Oakley, C. J., and J. C. Murrell.** 1988. *nifH* genes in the obligate methane oxidising bacteria. FEMS Microbiol. Lett. **49:**53–57.

43a. **Omori, T.** Unpublished data.

44. **Parales, R. E., M. D. Emig, N. A. Lynch, and D. T. Gibson.** 1998. Substrate specificities of hybrid naphthalene and 2,4-dinitrotoluene dioxygenase enzyme systems. J. Bacteriol. **180:**2337–2344.

44a. **Reichenbecher, W., and J. C. Murrell.** Unpublished data.

45. **Saeki, H., and K. Furuhashi.** 1994. Cloning and characterization of a *Nocardia corallina* B-276 gene cluster encoding alkene monooxygenase. J. Ferment. Bioeng. **78:**339–406.
46. **Saigne, C., and M. Legrand.** 1987. Measurements of methanesuphonic acid in Antarctic ice. Nature **330:**240–242.
47. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. **Sato, S. I., N. Ouchiyama, T. Kimura, H. Nojiri, H. Yamane, and T. Omori.** 1997. Cloning of genes involved in carbazole degradation of *Pseudomonas* sp. strain CA10: nucleotide sequences of genes and characterization of meta-cleavage enzymes and hydrolase. J. Bacteriol. **179:**4841–4849.
49. **Saunders, S. E., and J. F. Burke.** 1990. Rapid isolation of miniprep DNA for double strand sequencing. Nucleic Acids Res. **18:**4948.
50. **Scanlan, J., N. Baxter, and J. C. Murrell.** Unpublished data.
51. **Shine, J., and L. Dalgarno.** 1975. Determination of cistron specificity in bacterial ribosomes. Nature **254:**34–38.
52. **Simon, M. J., T. D. Osslund, R. Saunders, B. D. Ensley, S. Suggs, A. Harcourt, W. C. Suen, D. L. Cruden, D. T. Gibson, and G. J. Zylstra.** 1993. Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas*

*putida* strains G7 and NCIB 9816-4. Gene **127:**31–37.
53. **Stainthorpe, A. C., V. Lees, G. P. C. Salmond, H. Dalton, and J. C. Murrell.** 1990. The methane monooxygenase gene cluster of *Methylococcus capsulatus* (Bath). Gene **91:**27–34.
53a. **Swiss Institute of Bioinformatics, ExPaSy Molecular Biology Server.** September 1998, posting date. [Online.] http://expasy.hcuge.ch. [1 October 1998, last date accessed.]
54. **Taira, K., J. Hirose, S. Hayashida, and K. Furukawa.** 1992. Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. J. Biol. Chem. **267:**4844–4853.
55. **Tan, H. M., H. Y. Tang, C. L. Joannou, N. H. Abdel-Wahab, and J. R. Mason.** 1993. The *Pseudomonas putida* ML2 plasmid-encoded genes for benzene dioxygenase are unusual in codon usage and low in G+C content. Gene **130:** 33–39.
56. **Thompson, A. S., N. J. P. Owens, and J. C. Murrell.** 1995. Isolation and characterization of methanesulfonic acid-degrading bacteria from the marine environment. Appl. Environ. Microbiol. **61:**2388–2393.
57. **Wackett, L. P., L. D. Kwart, and D. T. Gibson.** 1988. Benzylic monooxygenation catalysed by toluene dioxygenase from *Pseudomonas putida*. Biochemistry **27:**1360–1367.
58. **Wende, P., F.-H. Bernhardt, and K. Pfleger.** 1989. Substrate-modulated reactions of putidamonooxin. The nature of the active oxygen species formed and its reaction mechanism. Eur. J. Biochem. **181:**189–197.
59. **Whung, P.-Y., E. S. Saltzman, M. J. Spencer, P. A. Mayewski, and N. Gundestrup.** 1994. Two-hundred-year record of biogenic sulfur in a south Greenland ice core (20D). J. Geophys. Res. **99:**1147–1156.
60. **Yen, K. M., and M. R. Karl.** 1992. Identification of a new gene, *tmoF*, in the *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. J. Bacteriol. **174:**7253–7261.
61. **Zehr, J. P., M. Mellon, S. Braun, W. Litaker, T. Steppe, and H. W. Paerl.** 1995. Diversity of heterotrophic nitrogen fixation genes in a marine cyanobacterial mat. Appl. Environ. Microbiol. **61:**2527–2532.
62. **Zylstra, G. J., and D. T. Gibson.** 1989. Toluene degradation by *Pseudomonas putida* F1. J. Biol. Chem. **264:**14940–14946.