MPMI Vol. 11, No. 11, 1998, pp. 1057–1068. Publication no. M-1998-0902-01R.

A Model for the Catabolism of Rhizopine in *Rhizobium leguminosarum* Involves a Ferredoxin Oxygenase Complex and the Inositol Degradative Pathway

Masoud Bahar,[1] John de Majnik,[1] Margaret Wexler,[1] Judith Fry,[2] Philip S. Poole,[2] and Peter J. Murphy[1]

[1]Department of Crop Protection, University of Adelaide, Waite Campus, Glen Osmond, South Australia 5064, Australia; [2]School of Animal and Microbial Sciences, University of Reading, Whiteknights, PO Box 228, Reading, U.K.

Accepted 15 July 1998.

Rhizopines are nodule-specific compounds that confer an intraspecies competitive nodulation advantage to strains that can catabolize them. The rhizopine (3-*O*-methyl-*scyllo*-inosamine, 3-*O*-MSI) catabolic *moc* gene cluster *mocCABRDE(F)* in *Rhizobium leguminosarum* bv. *viciae* strain 1a is located on the Sym plasmid. MocCABR are homologous to the *mocCABR* gene products from *Sinorhizobium meliloti.* MocD and MocE contain motifs corresponding to a TOL-like oxygenase and a [2Fe-2S] Rieske-like ferredoxin, respectively. The *mocF* gene encodes a ferredoxin reductase that would complete the oxygenase system, but is not essential for rhizopine catabolism. We propose a rhizopine catabolic model whereby MocB transports rhizopine into the cell and MocDE and MocF (or a similar protein elsewhere in the genome), under the regulation of MocR, act in concert to form a ferredoxin oxygenase system that demethylates 3-*O*-MSI to form *scyllo*-inosamine (SI). MocA, an NAD(H)-dependent dehydrogenase, and MocC continue the catabolic process. Compounds formed then enter the inositol catabolic pathway.

Additional keywords: nitrogen fixation, toluene.

Nodulation of legumes is dependent upon the correct combination of rhizobial and plant species. Competition between rhizobial strains within the same species also occurs for a particular host plant. This intraspecies competition depends upon factors such as the genes involved in nodulation, abiotic factors, and biotic factors. The biotic factors include bacteriocins and a variety of compounds synthesized during the symbiotic interaction (see review by Vlassak and Vanderleyden 1997). This latter group consists of compounds such as the betaine trigonelline (Boivin et al. 1990, 1991), homoserine (van Egeraat 1975), certain flavonoids (Hartwig et al. 1991), and rhizopines.

Rhizopines are novel inositol compounds that are synthesized in root nodules by bacteroids. These compounds are catabolized by the free-living rhizobia that induce the production of the rhizopine (Murphy and Saint 1992; Murphy et al. 1995). By far the most common rhizopine is 3-*O*-methyl-*scyllo*-inosamine (3-*O*-MSI) (Murphy et al. 1987). Strains of rhizobia capable of rhizopine synthesis and catabolism represent approximately 10% of *Sinorhizobium meliloti* (formerly *Rhizobium meliloti*) and *R. leguminosarum* bv. *viciae*. To date, rhizopines have not been found in other rhizobial species (Wexler et al. 1995). In the *S. meliloti* strains L5-30 and Rm220-3 the synthesis (*mos*) and catabolic (*moc*) genes have been isolated and shown to be closely linked on the Sym plasmid (Murphy et al. 1993; Saint et al. 1993). Furthermore, the regulation of the *mos* genes by the symbiotic regulatory gene *nifA* indicates that rhizopines are involved in the symbiotic interaction (Murphy et al. 1988; Saint et al. 1993). Although the *mos* genes do not enhance nodulation on their own, when in combination with the *moc* genes they confer a distinct advantage in competition for nodulation (Gordon et al. 1996).

The *mos* genes from the *S. meliloti* strains L5-30 and Rm220-3 have been characterized (Murphy et al. 1993; Rao et al. 1995) and more recently the *moc* genes from L5-30 were sequenced (Rossbach et al. 1994). Four genes, *mocCABR*, were shown to be required for 3-*O*-MSI catabolism. The *moc* genes were ascribed putative roles based on protein homologies deduced from the DNA sequence. It has been proposed that MocB functions as a transport protein, MocA as an inositol dehydrogenase, and MocR as a regulatory protein. MocC did not have any significant homologies to known proteins but was proposed to be involved in inositol catabolism (Rossbach et al. 1994).

In our efforts to understand the function of the different rhizopine genes, we have isolated and sequenced the *moc* gene cluster from a rhizopine-producing strain of *R. leguminosarum* bv. *viciae*. We report that for rhizopine catabolism in this species there is a requirement for genes homologous to *S. meliloti mocCABR* as well as genes constituting a ferredoxin-oxygenase catabolic system and genes involved in inositol catabolism.

Corresponding author: Peter J. Murphy
E-mail: pmurphy@waite.adelaide.edu.au

Current address of Masoud Bahar: Department of Crop Protection, Faculty of Agriculture, Isfahan University of Technology, Isfahan 84154, Iran.

Current address of Margaret Wexler: Nitrogen Fixation Laboratory, Norwich Research Park, Colney Lane, Norwich NR4 7UH, UK.

Nucleotide and/or amino acid sequence data are to be found at the GenBank data base as accession number AF076240.

RESULTS

Localization of the *moc* genes on the Sym plasmid.

Transfer of the Tn*5*-*mob* labeled Sym plasmid of *R. leguminosarum* bv. *viciae* 1a (LS2090) into a plasmidless *Agrobacterium* strain (K749) enabled the transconjugant to catabolize 3-*O*-MSI (data not shown). Sym plasmid localization of the *moc* genes was confirmed by probing plasmid gels (Eckhardt 1978) of the transconjugant with *nod* (pRT181) and *moc* (pPM1145) probes (data not shown).

Analysis of 1a *moc* clones.

The rhizopine catabolic (*moc*) region of *R. leguminosarum* bv. *viciae* strain 1a was isolated from a cosmid library by hybridization with the *S. meliloti* L5-30 *moc* genes. The plasmid pPM1291 (Fig. 1), which is a subclone from the library, contains four adjacent *Hin*dIII fragments of 9, 0.3, 0.9, and 3.8 kb. Through catabolism experiments, this plasmid was shown to contain the functional *moc* genes. To further delimit this region, subclones of the *moc* locus were prepared and transferred to the Moc⁻ strain, *R. leguminosarum* 8401(pRL1). Transconjugants were examined for their ability to degrade 3-*O*-MSI and its demethylated form, *S*I. Only pPM1291 could confer the ability to catabolize 3-*O*-MSI. None of the deleted or complementing clone combinations further delimited the 3-*O*-MSI catabolic region in pPM1291. However, a number of the constructs, either on their own or in combination, could catabolize *S*I (see Table 1, Figure 1). The minimal complementing clones needed to confer *S*I catabolism are pPM1305+pPM1296 (Fig. 1). These cover a 8-kb *Xho*I-*Hpa*I fragment in the center of pPM1291. Since pPM1235 catabolizes *S*I and is missing a 2.3-kb *Hin*dIII-*Hpa*I fragment from the right of pPM1296, an even smaller 5.7-kb *Xho*I-*Hin*dIII fragment would be sufficient for *S*I catabolism.

Defining the *moc* region.

To further define the *moc* genes, a combination of Tn*5*-B20 mutagenesis and DNA sequence analysis was used. The plasmid pPM1291 was mutagenized with the transposon Tn*5*-B20 (Simon et al. 1989), the position of the transposons mapped,

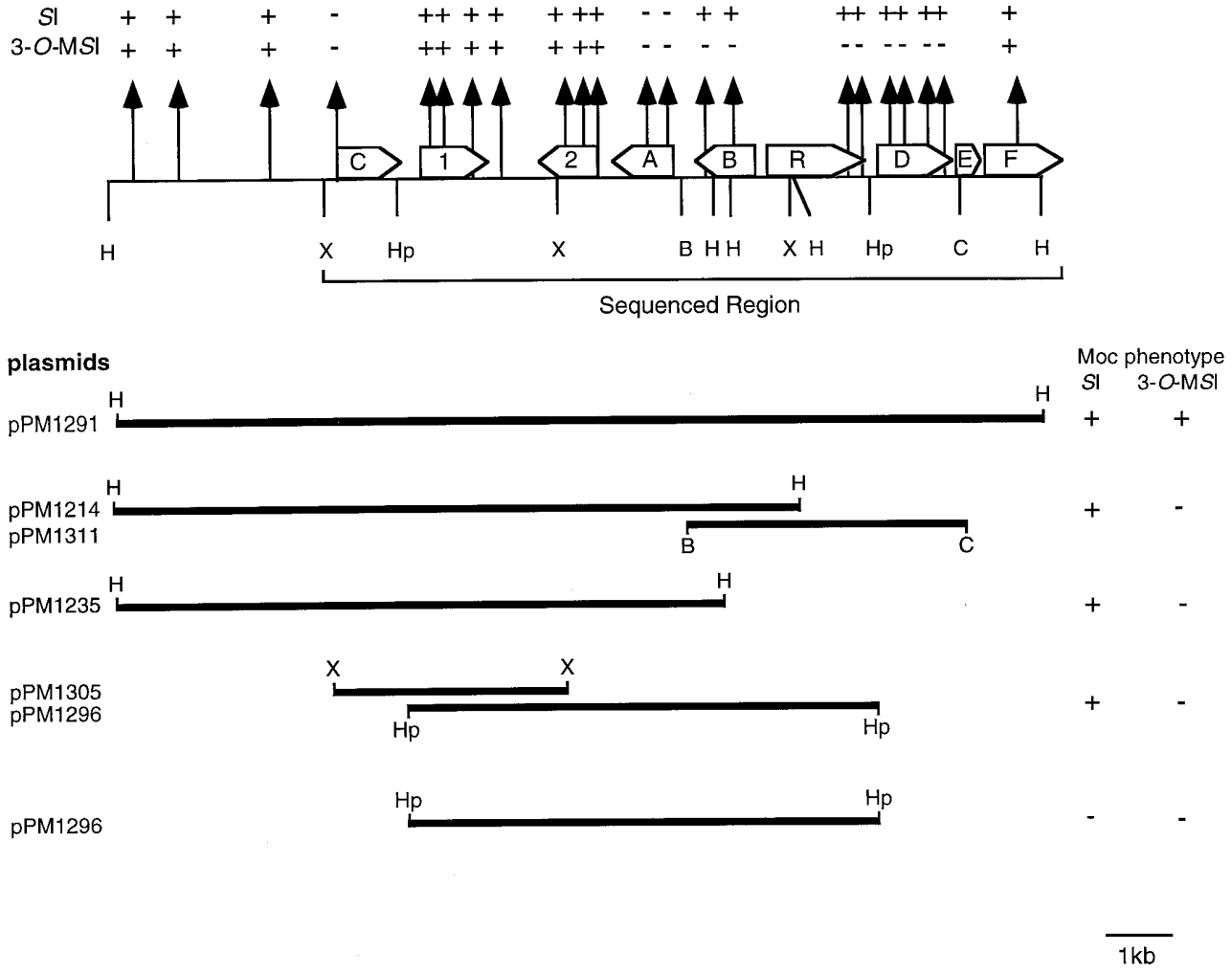

Fig. 1. Structure of rhizopine *moc* cluster in *Rhizobium leguminosarum* bv. *viciae* 1a. Deduced open reading frames shown by large open arrows. Narrow vertical arrows indicate location of Tn*5*-B20 inserts, above which catabolic phenotype for 3-*O*-MSI and *S*I is indicated by "+" (catabolism) and "–" (non-catabolism). Deletion and complementing clones are shown below map. Catabolic phenotypes (+, –) for 3-*O*-MSI and *S*I are shown on right. Region sequenced is indicated by a thin line. Only restriction enzyme sites discussed in the text are shown. B, *Bam* HI; C, *Cla*I; H, *Hin*dIII; Hp, *Hpa*I, X, *Xho*I.

the mutated plasmids transferred into 8401(pRL1), and 3-*O*-MSI and *S*I catabolic phenotypes determined. The results of this mutagenesis are shown in Figure 1. Based on the Tn*5*-B20 and deletion studies, a fragment of 10,804 nucleic acids was sequenced (Fig. 1). Within this region, nine putative open reading frames (ORFs) were detected by the programs MAP and CODONPREFERENCE with a *Rhizobium leguminosarum* data base of codon usage. Based on homologies with the previously identified *moc* ORFs in *S. meliloti*, six of the nine ORFs were named *mocC*, ORF1, ORF2, *mocA*, *mocB*, and *mocR*. The remaining three ORFs were named *mocD*, *mocE*, and *mocF*. The size and position of these ORFs in the DNA sequence is shown in Table 2.

Genes required for 3-*O*-MSI catabolism.

Tn*5*-B20 mutagenesis showed that if the *mocCABR* and *mocD* genes are disrupted, 3-*O*-MSI is not catabolized. If ORF1 and ORF2 are disrupted, 3-*O*-MSI is still catabolized, showing that these ORFs are not essential. In addition, complementing clones pPM1214 and pPM1311, which incorporate *mocCABRD*, but not *mocE* or *mocF*, did not catabolize 3-*O*-MSI. The plasmid pPM1291, which lacks the C-terminal end of *mocF*, could confer the ability to catabolize 3-*O*-MSI, and a Tn*5*-B20 insertion in the central region of *mocF* did not affect 3-*O*-MSI catabolism. These results indicate that, in addition to *mocCABRD, mocE* but not *mocF* is essential for 3-*O*-MSI catabolism.

Genes required for *S*I catabolism.

Tn*5*-B20 mutagenesis indicated that *mocA* and *mocC* are the only genes in the *moc* cluster required for *S*I metabolism.

Table 2. Deduced open reading frames (ORFs) from DNA sequence

ORF	DNA sequence coordinates	Shine-Dalgarno sequence/ intervening bases/ start codon/ number of residues/ stop codon	Calculated molecular weight
MocC	215 to 1102	GAAGGAAA-6-ATG-295-TGA	32,357
ORF1	1438 to 2442	GAGGAAG-2-ATG-334-TAG	36,136
ORF2	3995 to 3102	GGGAGAGA-8-ATG-297-TAG	32,831
MocA	5091 to 4171	AGG-8-ATG-306-TGA	32,728
MocB	6311 to 5382	GGAG-7-GTG-309-TGA	32,808
MocR	6457 to 7923	-ATG-488-TAA	54,292
MocD	8086 to 9171	GGAGG-6-ATG-361-TGA	41,164
MocE	9217 to 9534	GGAG-5-ATG-105-TAA	11,568
MocF	9539 to 10804	-ATG-421-TAA	44,779

Table 1 Bacterial strains and plasmids used in this study

Strains and plasmids	Description	Source or reference
Rhizobium leguminosarum bv. *viciae*		
1a	Wild type strain, Mos^+, Moc^+	Jensen 1987
LS2090	1a(pSym::Tn*5*-*mob*), Moc^+, Mos^+, Sr^r, Km^r, Tc^r	L. Scøt, WPBS, Aberystwyth, UK
3841	Wild-type strain, Str^r derivative of strain 300 bv. *viciae*	Johnston and Beringer 1975
8401(pRL1)	Mos^-, Moc^-, Rif^r	Downie et al. 1983
RU360	Strain 3841::Tn*5* defective for growth on *myo*-inositol	Poole et al. 1994
RU361	Strain 3841::Tn*5* defective for growth on *myo*-inositol	Poole et al. 1994
Sinorhizobium meliloti		
Rm1021	Mos^-, Moc^-	Meade et al. 1982
Azorhizobium radiobacter		
K749	C58, pTi^-, pTi^-, Rif^r, Sr^r	Donner et al. 1993
Escherichia coli		
DH5α	Cloning strain	Sambrook et al. 1989
JM109	Cloning strain	Yanisch-Perron et al. 1985
HB101	Cloning strain	Boyer and Roulland-Dussoix 1969
MC1061	Cloning strain	Casadaban and Cohen 1980
Plasmids		
pGEM-5Zf(+)	Amp^r, cloning and sequencing vector	Promega (Madison, WI)
pGEM-7ZF(+)	Amp^r, cloning and sequencing vector	Promega
pJB3JI	Tra^+, Tc^r, Km^s derivative of R68.45	Brewin et al. 1980
pKT231	Km^r, broad host range Inc Q cosmid	Bagdasarian et al. 1981
pRK2013	Km^r, helper plasmid	Figurski and Helinski 1979
pUCD2608	Km^r, broad host range Inc W cosmid	P. Rogowoski, CNRS, Lyon, France
pVK102	Tc^r, broad host range Inc P-1 cosmid	Knauf and Nester,1982
pPM1145	5.4-kb *Eco*RI fragment from pPM1031, in pJRD184, Tc^r, contains *S. meliloti moc* genes	see Murphy et al. 1987
pPM1214	0.3, 0.9, 9 kb *Hin*dIII fragments from pPM1291, in pVK102, Tc^r	This study
pPM1235	9 kb *Hin*dIII fragment from pPM1291 in pVK102, Tc^r	This study
pPM1291	0.3, 0.9, 3.6, 9.0 kb *Hin*dIII fragments, functional 1a *moc* region, in pVK102, Tc^r	This study
pPM1296	7.2 kb *Hpa*I fragment from pPM1291, in pKT231, Km^r	This study
pPM1305	3.5 kb *Xho*I fragment from pPM1291, in pVK102, Tc^r	This study
pPM1309	5.9 kb *Xho*I-*Hin*dIII fragment from pPM1291, in pGEM-7Zf(+), Amp^r	This study
pPM1311	3.7 kb *Bam*HI-*Cla*I fragment from pPM1291, in pUCD2608, Km^r	This study
pPM1314	0.3, 0.9, 3.6 kb *Hin*dIII fragments from pPM1291, in pGEM-7Zf(+), Amp^r	This study
pPM1319	3.3 kb *Xho*I fragments from pPM1291, in pGEM-7Zf(+), Amp^r	This study
pPM1340	3.5 kb *Xho*I fragment from pPM1291, in pGEM-7Zf(+), Amp^r	This study
pPM1343	0.3, 0.9, 3.6 kb *Hin*dIII fragments from pPM1291, in pGEM-7Zf(+), Amp^r, opposite direction to pPM1314	This study
pRT181	2.1 kb *Eco*RI-*Bam*HI *nodBAD* fragment from *R. leguminosarum* bv. *trifolii*, in pKT230	B. Rolfe, ANU, Canberra, Australia
Tn*5*-B20	Transposon with *lac* gene	Simon et al. 1989

This is supported by results obtained with the use of combinations of complementing subclones (Fig. 1).

DNA sequence analysis of the *moc* cluster: MocC.

A search of sequence data bases indicated that MocC from *R. leguminosarum* strain 1a has extensive homology (84% identity and 90% similarity) to MocC of *S. meliloti* L5-30 (Rossbach et al. 1994). The *R. leguminosarum* protein is 30 amino acids shorter than the protein in *S. meliloti*, lacking residues at both the C-terminal and N-terminal ends. In addition, there is considerable homology (39% identity and 62% similarity) to a partial ORF, *E83A*, that is part of the *myo*-inositol catabolism *iol* operon in *Bacillus subtilis* (Yoshida et al. 1994). The specific role of *E83A* has not been characterized in the *myo*-inositol catabolism pathway. However, *E83A* is located 6 kb upstream of a gene that encodes an inositol dehydrogenase, *E83G* (called *idh*) that has homology to *mocA*. No other proteins homologous to MocC were found in data bases. It is currently not possible to conclusively predict how *mocC* is involved in rhizopine catabolism in strain 1a.

ORF1 and ORF2.

Tn*5*-B20 insertions in the intergenic region between *mocC* and *mocA* did not affect the catabolism of *S*I or 3-*O*-M*S*I, indicating that this region is not important for rhizopine catabolism. There are two ORFs in this region, which are transcribed convergently. The first is designated ORF1 and is transcribed in the same direction as *mocC*. It has 81% identity and 90% similarity at the amino acid level with ORF334 from *S. meliloti* L5-30. The second ORF is termed ORF2 and is transcribed in the same direction as *mocA*. The deduced amino acid sequence of ORF2 has 79% identity and 86% similarity with ORF293 from L5-30 (Rossbach et al. 1994).

The DNA sequence separating ORF1 and ORF2, like those found in the intergenic regions of the other genes, was divergent from that of L5-30. Similar to ORF334 and ORF293, no conserved motif was found within the sequences of ORF1 and ORF2 to predict their possible function. ORF1 showed homology (37% identity, 57% similarity over 240 amino acids) to RdmF from *Streptomyces purpurascens*, whose function is unknown (Niemi and Mäntsälä 1995). The glucose-fructose oxidoreductase of *Zymomonas mobilis* (Kanagasundaram and Scopes 1992), MocA from L5-30 (Rossbach et al. 1994), and strain 1a are also homologous to ORF1 in parts of the sequence. ORF2 was found to be homologous (43% identity and 55% similarity) over a 104 amino acid region of CelC in *Bacillus stearothermophilus* (Lai and Ingram 1993). This gene encodes a cleavage enzyme for cellobiose phosphate in the cellobiose phosphotransferase system. Although these genes appear not to have a role in rhizopine catabolism, their conservation in both *S. meliloti* and *R. leguminosarum* suggests they may have some other role.

MocA.

MocA is located approximately 3 kb downstream of MocC and exhibits strong homology (80% identity and 88% similarity) to the MocA protein in L5-30 (Rossbach et al. 1994). MocA homology is also found with the gene product StrI, a glyceraldehyde-3-phosphate dehydrogenase involved in the formation of the inositol-derived compound streptidine from *Streptomyces griseus* (33% identity and 55% similarity) (Mansouri and Piepersberg 1991), and the *idh* gene product that is an inositol dehydrogenase from *B. subtilis* (24% identity and 47% similarity) (Fujita et al. 1991; Yoshida et al. 1994). Both of these gene products contain a conserved motif Gly-X-Gly-X_2-Gly in their N-terminal end, which is characteristic for NADH-binding βαβ-fold proteins. This evidence suggests that, as for MocA in L5-30, the *mocA* gene product in strain 1a plays a functional role in the catalysis of an NADH-dependent dehydrogenase reaction involved in rhizopine catabolism.

MocB.

MocB is separated from MocA by 290 nucleotides, a distance very similar to that observed between MocB and MocA in L5-30 (Rossbach et al. 1994). The deduced amino acid sequence of MocB from strain 1a shows striking similarity (89% identity and 96% similarity) to MocB from *S. meliloti* strain L5-30. With a GTG start site rather than an ATG further down stream, the deduced amino acid sequence showed an amino-terminal sequence characteristic of a signal peptide sequence and contained a cleavage site at the 20 to 21 amino acid position (Pugsley 1993; Sjöström et al. 1987). Overall, the signal peptide of strain 1a MocB is closely related to that in strain L5-30. Both the 1a and L5-30 MocB proteins have homology to D-galactose binding protein (MglB; Hogg et al. 1991), D-ribose binding protein (Groarke et al. 1983), and D-xylose binding protein (Sumiya and Henderson 1989; Sofia et al. 1994) as determined by the program BLAST.

MocR.

The sixth ORF encodes a 488 amino acid protein corresponding to the 493 amino acids that constitute MocR in L5-30. There is 86% identity and 93% similarity between MocR in strain L5-30 (Rossbach et al. 1994) and strain 1a. The *mocR* gene in both strains 1a and L5-30 is located immediately downstream of *mocB*, with transcription of *mocR* occurring in the direction opposite to that of *mocB*. Identical to MocR in L5-30, alanine, glycine, and valine residues occupy positions 42, 46, and 52, respectively in MocR of strain 1a. This putative helix-turn-helix motif is conserved in the amino-terminal end of DNA-binding proteins (Pabo and Sauer 1984), indicating that MocR may function as a DNA-binding protein. The search for proteins homologous to MocR showed significant homology (36 to 38% identity and 56 to 62% similarity) between the first 80 amino acid residues of YrdX in *Rhodobacter sphaeroides* (Neidle and Kaplan 1992), PtsJ in *Salmonella typhimurium* (Titgemeyer et al. 1995), FarR in *Escherichia coli* (Buck and Guest 1989), and LldR in *E. coli* (Dong et al. 1993). These proteins have been determined to be transcriptional regulatory proteins belonging to the GntR class of regulators (Haydon and Guest 1991). However, the remainder of MocR was less homologous (20 to 25% identities) or completely divergent from these proteins. This central and C-terminal portion of MocR from both strains 1a and L5-30 has similarity to unspecified aminotransferases in *Sulfolobus sulfataricus* (Sensen et al. 1996) and tyrosine aminotransferase in *E. coli* and *B. subtilis* (Cosmina et al. 1993), as well as to many other aminotransferases involving the cofactor pyridoxal phosphate. From these homologies the role of MocR is not totally clear, but it has been suggested to be a regulatory protein (Rossbach et al. 1994).

MocD.

The *mocD* gene is downstream of *mocR* and consists of 1,083 nucleotides. Tn*5*-B20 insertion within the gene or deletion of the *mocD* region results in failure to catabolize 3-*O*-M*S*I, although the mutants or deleted subclones are still able to catabolize *S*I. MocD has regions of notable homology to ω-3 fatty acid desaturases of plants (Yadav et al. 1993) and cyanobacteria (Wada et al. 1990). The high degree of homology was mainly associated with three highly conserved histidine boxes (Fig. 2A). The first box, containing His-Glu-Cys-Gly-His, was found in residues 96 to 100. The second box, consisting of His-Ala-Arg-His-His, was detected 31 amino acids distant from the first box, and the third box, His-Val-Glu-His-His, was located 145 amino acids after the second histidine box. The same motifs ($HX_{(3\ or\ 4)}H$, $HX_{(2\ or\ 3)}HH$, and $HX_{(2\ or\ 3)}HH$) have been reported in membrane desaturases of eukaryotes and cyanobacteria as well as from the bacterial membrane enzymes alkane hydroxylase (Kok et al. 1989a, 1989b) and xylene monooxygenase (Suzuki et al. 1991). Analysis of the hydropathy profile of MocD by the method of Kyte and Doolittle (1982) revealed that the three conserved histidine regions are present in hydrophilic domains (Fig. 2B).

MocE.

The *mocE* gene is in the immediate vicinity of *mocD* and is a short ORF consisting of 315 nucleotides. Deletion analysis showed that this region is essential for 3-*O*-M*S*I catabolism. Figure 3 shows the alignment of MocE with ferredoxin components from various species of bacteria. NdoA, a ferredoxin in the naphthalene catabolism system in *Pseudomonas putida* (Kurkela et al. 1988), contains 104 amino acids with 40% identity and 60% similarity to MocE. BnzC, a 106 amino acid ferredoxin component involved in the degradation pathway of benzene (Irie et al. 1987) has 37% identity and 55% similarity to MocE. MocE also has 37% identity and 55% similarity to a 106 amino acid ferredoxin component involved in the toluene catabolism system in *Pseudomonas putida* (TodB) (Zylstra and Gibson 1989). Among these genes, the Cys-X-His and Cys-X2-His motifs, which are characteristic of Rieske-like ferredoxin components involved in hydrocarbon catabolism (Tanaka et al. 1974; Morrice et al. 1988; Neidle et al. 1991), are conserved.

MocF.

Five base pairs downstream of *mocE*, and transcribed in the same direction as *mocE*, is *mocF*. A mutant with a Tn*5*-B20

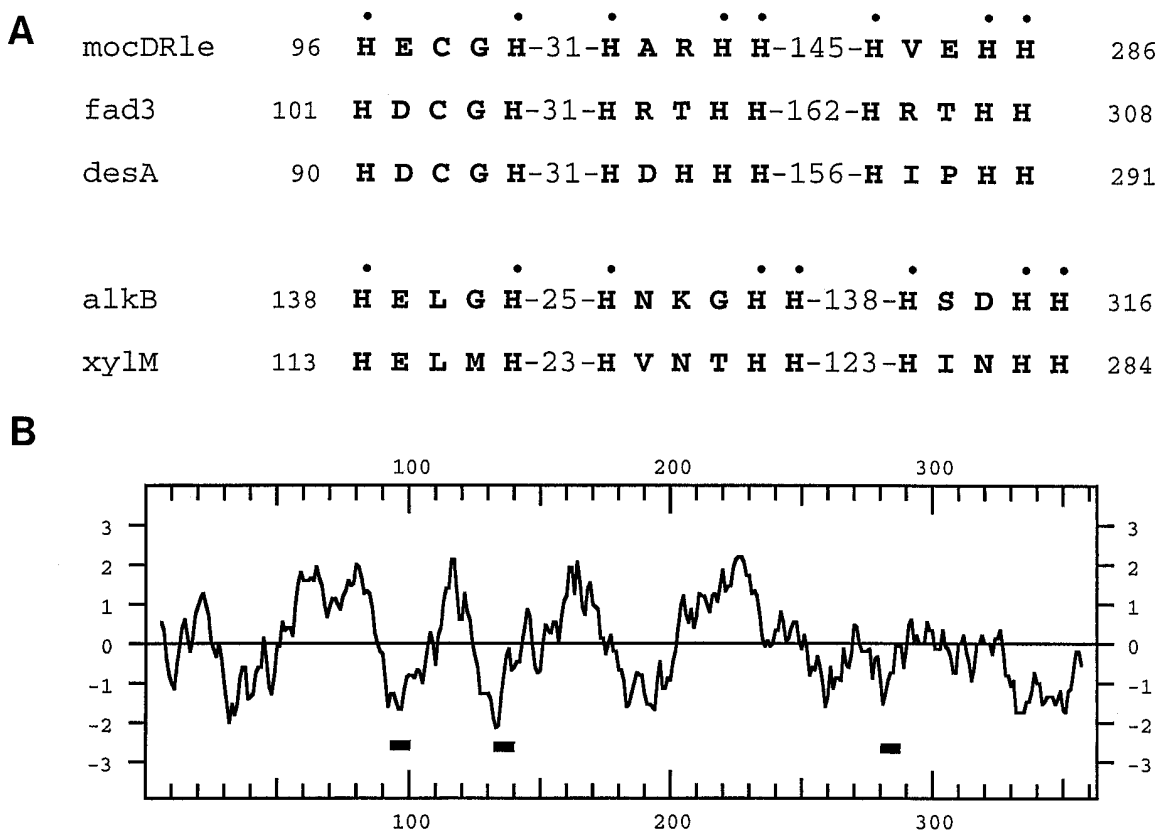

Fig. 2. A, Conservation of histidine boxes that act as metal-chelating ligands contributing to binding of oxygen in reaction center in the following: mocDRle, MocD from *Rhizobium leguminosarum* bv. *viciae* strain 1a; fad3, ω-3 fatty acid desaturase from *Arabidopsis thaliana;* desA, fatty acid desaturase from *Synechocystis* spp.; alkB, alkane hydroxylase from *Pseudomonas olevorans;* and xylM, xylene monooxygenase from *Pseudomonas putida*. Black dots show conserved histidine residues. Numbers at ends of sequence show positions of start and finish of the boxes. Numbers within the sequence show number of amino acids between histidine boxes. **B,** Hydrophobicity plot of MocD from *R. leguminosarum* bv. *viciae* strain 1a. Hydropathy indices calculated by the method of Kyte and Doolittle (1982). Conserved histidine regions indicated by horizontal bars. Vertical scale, negative values are hydrophilic regions. Horizontal scale, amino acid position.

insertion located in the middle of *mocF* was still able to catabolize both *S*I and 3-*O*-MSI, (Fig. 1) indicating that this ORF is not crucial for rhizopine catabolism in strain 1a. Comparison of the deduced amino acid sequence of *mocF* with sequences in the data bases showed considerable homology to a number of ferredoxin-like reductases (Fig. 4). These reductases have similar modes of action. For example, rhodocoxin reductase (ThcD), from *Rhodococcus* sp. strain NI86/21 (Nagy et al. 1995), is involved in the degradation of a thiocarbomate herbicide requiring the participation of a flavoprotein.

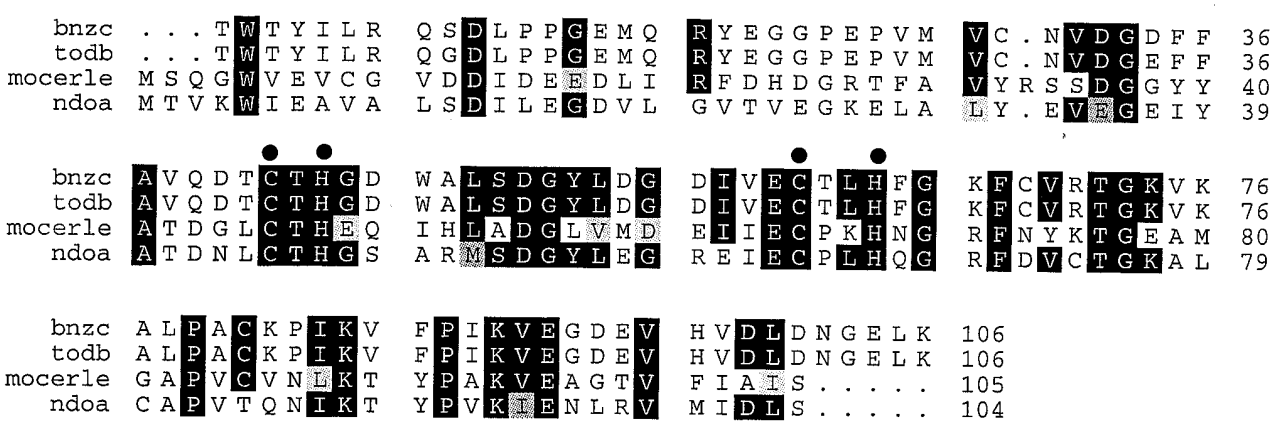

Fig. 3. Alignment of deduced amino acid sequence of the following: mocerle, *mocE* from *Rhizobium leguminosarum* bv. *viciae* strain 1a; ndoa, naphthalene 1, 2-dioxygenase system ferredoxin component (NdoA) from *Pseudomonas putida*; todb, toluene 1, 2-dioxygenase system ferredoxin component (TodB) from *P. putida;* and bnzc, a ferredoxin component involved in benzene degradation. Residues identical for at least three sequences indicated by black boxes; similar residues shown by gray boxes. Highly conserved cysteine (C) and histidine (H) residues indicated by black dots.

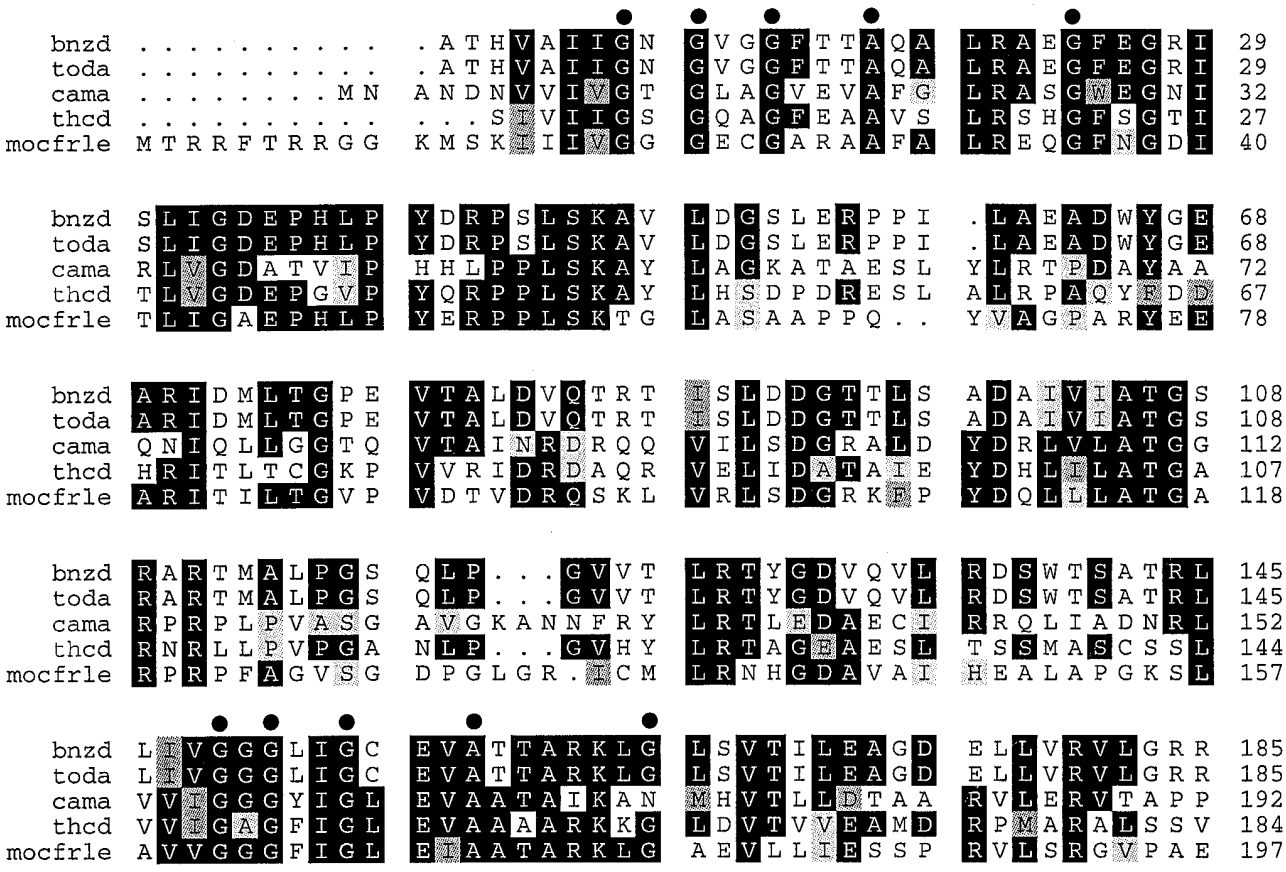

Fig. 4. Alignment of deduced amino acid sequence of the following: mocfrle, MocF from *Rhizobium leguminosarum* bv. *viciae* strain 1a; thcd, rhodocoxin reductase (ThcD) from *Rhodococcus* spp.; cama, putidaredoxin reductase component involved in camphor degradation (CamA) from *Pseudomonas putida;* toda, ferredoxin reductase component involved in toluene degradation (TodA) from *P. putida*; bnzd, ferredoxin reductase component involved in benzene degradation (BnzD) from *P. putida*. Residues identical for at least three sequences indicated by black boxes; similar residues shown by gray boxes. Consensus sequence (Gly-X-Gly-X_2-Gly-X_3-Ala-X_6-Gly) at two different sites shown by black dots. Hhomology extends 240 amino acids farther than shown.

The rhodocoxin reductase and an iron-sulfur protein mediate the transfer of electrons from NADH to P450 to activate an oxygenase reaction. Similarly, CamA from *P. putida* (Peterson et al. 1990) is involved in the degradation of camphor in a manner similar to that of ThcD, but instead utilizes putidaredoxin and putidaredoxin reductase. MocF also has considerable homology with TodA (Zylstra and Gibson 1989) and BnzD (Irie et al. 1987) from *P. putida*, which are involved in toluene and benzene degradation, respectively. Both of these are ferredoxin reductase components that interact with ferredoxin in ring-hydroxylating dioxygenase reactions.

The consensus sequence (Gly-X-Gly-X2-Gly-X3-Ala-X6-Gly) characteristic of ferredoxin-type NAD-dependent enzymes was found at two different sites in *mocF*, at residue positions 19 to 35 and 161 to 177. Two hydrophobic regions rich in leucine, alanine, and valine are located before and after the consensus sequences that start with a charged residue (Fig. 4). These features are also present in the enzymes mentioned above (Hanukoglu and Gutfinger 1989). The presence of two NAD-binding motifs in *mocF* together with considerable homology to ferredoxin reductases suggests that this ORF may encode for ferredoxin reductase in strain 1a. The close proximity of *mocF* to the putative ferredoxin encoding gene, *mocE*, also supports this hypothesis.

The involvement of inositol catabolism genes in rhizopine degradation.

The similarity in structure between 3-*O*-M*S*I and inositol suggests inositol catabolic pathways may also be involved in rhizopine catabolism. Accordingly, the *moc* plasmid pPM1291 was mated into the control strain *R. leguminosarum* bv. *viciae* 3841 and into its Tn*5*-*lac* derived mutants (RU360 and RU361), which are unable to catabolize inositol (Poole et al. 1994). After 5 days of incubation with *S*I or 3-*O*-M*S*I catabolic media it was shown that only strain 3841(pPM1291) could catabolize *S*I and 3-*O*-M*S*I. Strains RU360 and RU361 with or without the *moc* region were not able to catabolize the rhizopines (data not shown). These results demonstrate that inositol catabolic genes, as well as the *moc* region, are involved in rhizopine degradation.

In order to determine precisely which genes in the *myo*-inositol degradation pathway are disrupted in RU360 and RU361, Tn*5*-*lacZ* and the surrounding region were cloned from both strains and the regions adjacent to the transposon sequence. Sequence analysis indicated that RU361 is mutated in a gene homologous (70% identity over 217 residues) to malonate semialdehyde oxidative decarboxylase from *Streptomyces coelicolor*, which is involved in the final step of the proposed pathway of *myo*-inositol degradation (Poole et al. 1994). The mutation in RU361 inactivates a gene homologous (39% identity over 81 residues) to the *iolD* gene of *B. subtilis*, which codes for acetolactate synthase. This gene is found in the *myo*-inositol degradation operon of *B. subtilis* (Yoshida et al. 1997).

DISCUSSION

The *mocCABR* genes required for the catabolism of the rhizopine 3-*O*-M*S*I in both *R. leguminosarum* bv. *viciae* and *S. meliloti* are highly conserved, whereas the intergenic regions are far less conserved. The spacing and direction of translation of these genes in the two species are also similar. There is also conservation between ORF1 and ORF2 and the equivalent ORFs in L5-30, even though these ORFs are not required for rhizopine catabolism. The assignment of functions for *mocCABR* was based on homologies consistent with those found by Rossbach et al. (1994). The proposed roles are as follows: *mocA* as a NAD(H)-dependent inositol dehydrogenase; *mocB* as a transport protein; and *mocR* with a regulatory role. The function of *mocC* cannot be conclusively determined, although it is interesting that *mocC* has homology with a gene of unknown function that is linked to inositol catabolic genes in *Bacillus subtilis* (Yoshida et al. 1994).

For the degradation of *S*I, the demethylated form of 3-*O*-M*S*I, only *mocC* and *mocA* are required. However, for 3-*O*-M*S*I catabolism, the *mocCABRDE* genes are all necessary, suggesting that the catabolism of 3-*O*-M*S*I proceeds via *S*I. Since *S*I and 3-*O*-M*S*I are not catabolized in inositol-catabolic mutants of *R. leguminosarum*, other genes, besides those in the *moc* cluster, are required for rhizopine catabolism.

The *moc* oxygenase system.

The oxidation reactions of aromatic hydrocarbons (such as toluene, benzene, and xylene) are similar in that the initial reaction is carried out by the association of three protein components: an oxygenase, a ferredoxin, and a ferredoxin reductase. Basically, two electrons from NAD(P)H are transferred to a ferredoxin reductase. The reduced flavoprotein transfers electrons to a [2Fe-2S] ferredoxin, and finally the electrons are transferred from ferredoxin to the oxygenase component (Harayama and Kok 1992). MocDE and F show highly conserved motifs for this three-component system.

MocD: a putative oxygenase.

Analysis of Tn*5*-B20 mutants and deletion studies within the *moc* region of strain 1a indicate that the *mocD* gene is involved in the catabolism of 3-*O*-M*S*I but not *S*I. MocD has homology to fatty acid desaturase enzymes from plants and cyanobacter, and to monooxygenases involved in the catabolism of hydrocarbons. The homologous regions are largely confined to regions around characterized histidine boxes ($HX_{(3\text{ or }4)}H$, $HX_{(2\text{ or }3)}HH$ and $HX_{(2\text{ or }3)}HH$) that are present in hydrophilic regions. The histidine residues are proposed to act as ligands for the iron atoms embodied in these enzymes, which contribute to the binding of oxygen in the reaction center (Shanklin et al. 1994).

The spacing of the histidine residues more closely corresponds with the fatty acid desaturases that are involved in the introduction of the second and third double bonds into fatty acids (Sato et al. 1986; Yadav et al. 1993), yet it is difficult to conceive how this reaction could be involved in rhizopine catabolism. The more likely role would be similar to that reported for alkane hydroxylase (AlkB) (Kok et al. 1989b) and xylene monooxygenase (Suzuki et al. 1991) enzymes. Both of these enzymes are monooxygenases that are involved in the breakdown of aromatic hydrocarbons, albeit in different reactions. AlkB from *Pseudomonas oleovorans* introduces molecular oxygen into the terminal carbon atom of aliphatic hydrocarbons, which opens up the aromatic ring in the degradation of n-alkanes (Ueda et al. 1972; Kok et al. 1989b). The xylene monooxygenase (XO) that is encoded by the closely linked genes *xylA* and *xylM* of *P. putida* is a TOL

plasmid (pWW0) hydroxylase that incorporates one atom of oxygen into the methyl side chain of toluene or xylene (Suzuki et al. 1991). This oxidizes it to carboxylic acid, which can be further catabolized via the Krebs cycle (Worsey and Williams 1975).

The oxidation of a methyl side group in the xylene monooxygenase system resembles the manner in which 3-*O*-MSI is converted into *S*I. It is therefore proposed that MocD functions as an oxygenase for the conversion of 3-*O*-MSI to *S*I, which can then be further catabolized by other genes in the *moc* cluster.

MocE: a putative ferredoxin.

The ferredoxin component usually consists of a small protein of around 100 amino acids with conserved Cys-X-His and Cys-X2-His regions (Neidle et al. 1991). This motif is present in MocE, which is immediately downstream of *mocD*. Similar to the ferredoxin components for oxygenase systems involved in the catabolism of naphthalene, benzene, and toluene (NdoA, BnzC, and TodB1, respectively) MocE contains a threonine (T) residue in the first conserved cysteine-histidine box (Erickson and Mondello 1992). It is therefore reasonable to assume that MocE encodes a ferredoxin-like component that provides the required electrons for the oxidation of the methyl side chain in 3-*O*-MSI.

Ferredoxins are low molecular weight, non-haem iron proteins acting as electron donors in a range of electron transport chains. They possess characteristic motifs of cysteine residues that act as ligands for iron-sulfur clusters. The positioning of the cysteine residues is important in defining the structure and properties of the iron-sulfur chromaphore (Stout 1982). In plant, algal, and halobacterial ferredoxins there is a nearly invariant structure of C-X_4-C-X_2-C-X_{29}-C (see Meyer et al. 1986). Typical bacterial (clostridial)-type ferredoxins contain two cysteine motifs with the sequence of C-X_2-C-X_2-C-X_3-C (George et al. 1985). Distinct from this, *nif*-associated ferredoxins have the spacing of C-X_2-C-X_8-C-X_3-C in the second motif. The FdxN gene in *S. meliloti* (Klipp et al. 1989; Masepohl et al. 1992), which is essential for symbiotic nitrogen fixation, and its homolog in *R. leguminosarum* (Grönger et al. 1988) also have this *nif*-like ferredoxin structure. In iron-sulfur proteins such as those found in *Neurospora crassa* (Harnisch et al. 1985), four cysteines are arranged in two pairs, C-X_4-C and C-X-C. Distinct from these, *P. putida* contains a new class of [2Fe-2S] Rieske-like ferredoxins involved in the catabolism of benzene (Morrice et al. 1988; Tan et al. 1993) with a conserved sequence, C-X-H-$X_{15\text{-}17}$-C-X_2-H, that is also present in MocE. The type of ferredoxin component involved in the catabolism of rhizopines is more similar to those involved in hydrocarbon catabolism than to those involved in electron transfer in *nif* reactions. A Rieske oxygenase-like component has also been shown to be involved in the catabolism of stachydrine in *S. meliloti* (M. Burnet and D. Tepfer, *personal communication*).

Moc F: a putative ferredoxin-like reductase.

The *mocF* gene is likely to be cotranscribed with *mocE,* since it is 4 bp downstream and does not possess a recognizable Shine-Dalgarno sequence. The sequence shows similarity to a ferredoxin reductase, the third component of oxygenase systems. Two copies of the consensus sequence (G-X-G-X_2-G-X_3-A-X_6-G) that is characteristic of NAD-binding ferredoxin reductases (Hanukoglu and Gutfinger 1989) are present in MocF. These regions are proposed to be the site of interaction of both the flavin moiety and the NADH electron donor (Erickson and Mondello 1992). The function of all these enzymes is to transfer electrons from NADH to the ferredoxin, which in turn transfers them to the oxygenase.

The conserved motif and the close proximity of the oxygenase and ferredoxin components are compelling evidence that MocF provides the ferredoxin reductase to complete the requirement for an oxygenase system. However, MocF has been shown not to be essential for 3-*O*-MSI catabolism. It is possible that a ferredoxin reductase located elsewhere in the chromosome may substitute for *mocF.*

The catabolic model for 3-*O*-MSI.

The following model of a *moc* oxygenase system is proposed for 3-*O*-MSI catabolism steps in *R. leguminosarum* bv. *viciae* strain 1a. Most likely MocB is a rhizopine-binding protein interacting with chromosomal gene products to import 3-*O*-MSI into the bacterial cytoplasm. Inside the cell, catabolism of this compound may be carried out through the oxidation of the methyl side chain by sequential activities of MocF (or a similarly encoded protein elsewhere in the genome), MocE, and MocD. In this redox process, electrons from NADH would be transferred to the *mocF* product (ferredoxin reductase) that passes electrons to MocE as a Rieske-type, iron-sulfur designated ferredoxin. Electrons from ferredoxin would then be transferred to MocD, which would catalyze the oxidation of 3-*O*-MSI to *S*I or a closely related compound. Although the regulatory mechanism is not clear, it appears that this oxidation process may be performed under the regulation of MocR in *R. leguminosarum* bv. *viciae* strain 1a, since this protein is required for 3-*O*-MSI catabolism but not *S*I catabolism. In this regard, it is interesting that MocR shows significant homology to ORFT2 in *Rhodobacter sphaeroides* (Neidle and Kaplan 1992), which is assumed to be a transcriptional regulator located adjacent to *rdxA*, a *fixG* homolog, and which encodes a bacterial-type ferredoxin. *S*I or a related compound formed in the first step of degradation may then be catabolized by MocA, which is an NAD(H)-dependent dehydrogenase, and by MocC. Although the enzymatic role of MocC is unknown, it is proposed that this gene product, in association with inositol-catabolic genes that are located elsewhere in the chromosome, completes the steps of rhizopine catabolism.

The model raises the question as to why 3-*O*-MSI catabolism requires MocB, the putative transport protein, whereas *S*I catabolism does not. MocB shares homology with sugar-binding periplasmic proteins, which possess signal peptides and serve as recognition units for membrane-bound proteins involved in the active transport of compounds into cells (Ames 1986). The specificity is considered to reside in the C-terminal end of the protein (Müller-Hill 1983) and closely related compounds often require different binding proteins (Deshusses and Reber 1977). It is therefore conceivable that *S*I can be transported into the cells via a transporter, possibly involved in the *myo*-inositol system, while MocB transports 3-*O*-MSI.

The complete catabolic process is not yet known. It is possible the oxygenase step in the conversion of 3-*O*-MSI to *S*I

could involve the transient formation of the formaldehyde hemiketal of *S*I in a manner similar to the conversion of the methyl group to an alcohol group of toluene controlled by XylMA (Assinder and Williams 1990).

These studies indicate an absolute requirement for the inositol catabolism pathway in the breakdown of 3-*O*-M*S*I. One of the mutations used in this study is in malonate semialdehyde oxidative decarboxylase, which converts malonic semialdehyde co-enzyme A and NAD^+ to acetyl-CoA, CO_2, and $NADH+H^+$ in the final step in the proposed pathway of *myo*-inositol (Anderson and Magasanik 1971). The other inositol pathway mutant inhibiting rhizopine catabolism disrupts acetolactate synthesis, suggesting acetolactate may need to be synthesized from pyruvate for degradation of *myo*-inositol.

The model proposed outlines at least six genes in the *moc* cluster; however, these genes do not function alone, as other genes associated with transport and inositol catabolism must be involved in rhizopine catabolism. The process of utilizing pre-existing pathways appears to be characteristic of both the rhizopine *moc* and *mos* genes.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media.

Bacterial strains and plasmids are described in Table 1. *Rhizobium* strains were grown at 28°C on TY medium (Beringer 1974). *E. coli* strains were grown at 37°C on Luria-Bertani or Terrific Broth media (Sambrook et al. 1989). When required, antibiotics were added to the medium at the following concentration (in μg/ml): rifampicin 100, kanamycin 200, streptomycin 200, and tetracycline 10 for *Rhizobium;* and kanamycin 25, tetracycline 10, and ampicillin 100 for *E. coli.* For solid media, 1.5% agar was added.

Microbial techniques.

To generate *moc* mutants, plasmid pPM1291 was mutagenized with the Tn*5*-B20 transposon (Simon et al. 1989). For this purpose, CsCl-purified DNA of pPM1291 was transformed into *E. coli* A118 bearing the chromosomally inserted Tn*5*-B20 (Surin et al. 1990) and transformants selected on kanamycin and tetracycline. The mutated plasmids were first transferred to *S. meliloti* Rm1021 by triparental mating with the helper plasmid pRK2013, with selection on GTS medium containing kanamycin and tetracycline. The position of transposons was mapped by restriction enzyme digestion. For rhizopine catabolism tests plasmids were isolated, transformed into *E. coli,* and mobilized into a noncatabolizing strain, *R. leguminosarum* 8401(pRL1).

The Tn*5*-*mob* labeled Sym plasmid was transferred to *A. radiobacter* K749 with the helper plasmid pJB3JI.

General DNA techniques.

DNA manipulations were as described by Sambrook et al. (1989) and the Promega Protocols and Applications Guide (Promega, Madison, WI). DNA fragments were isolated and purified from agarose gels with the Geneclean kit (BIO 101, La Jolla, CA) or with homemade Sephadex columns (Wang and Rossman 1994). pGEM DNA containing inserts for sequencing was CsCl purified and acid phenol extracted (Promega) to remove nicked and linear DNA. This DNA was treated with exonuclease III and size selected on agarose gels to obtain overlapping regions for sequencing. A modified mini alkaline lysis/PEG precipitation procedure (Applied Biosystems, Foster City, CA) was used to isolate template DNA for sequencing. For hybridization studies, DNA was transferred by Southern blotting onto Boehringer-Mannheim positively charged membranes (Boehringer-Mannheim, Mannheim, Germany) and fixed by baking for 30 min at 120°C. Digoxigenin (DIG)-labeled probes were prepared by random primed incorporation of DIG-labeled dUTP with a preparative kit (Boehringer-Mannheim). Membranes were washed twice in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.1% sodium dodecyl sulfate (SDS) at room temperature for 5 min, then twice in 0.1× SSC, 0.1% SDS at 65°C for 20 min. DIG-labeled DNA was detected with the chemiluminescent substrate AMPPD (3-[2′-Spiroadamantane]-4-methoxy-4-[3″-phosphoryloxy]-phenyl-1, 2-dioetane), according to the manufacturer's instructions (Boehringer-Mannheim), using autoradiography.

DNA sequence determination.

The *moc* region of *R. leguminosarum* bv. *viciae* strain 1a was sequenced in both directions by preparing subclones derived from plasmid pPM1291 (Table 1). Cloned DNA from plasmids pPM1309 and pPM1314 was used for sequencing in one direction, and pPM1319, pPM1340, and pPM1343 for sequencing in the opposite direction. These plasmids were subjected to controlled exonuclease III digestion (Erase-a-base system, Promega) to generate nested deletions, which were then sequenced by the dideoxy method (Sanger et al. 1977) with the M13 forward or reverse sequencing primer. Sequence reactions were performed with the PRISM Ready Reaction Dye Primer Cycle Sequencing Kit (Applied Biosystems) and were analyzed by an Applied Biosystems 373A automated DNA sequencer as described by the manufacturer. To fill in sequencing gaps, 20-bp oligonucleotides from the previously determined sequence were synthesized and used as primers for DyeDeoxy terminator sequencing reactions (Applied Biosystems) as described by the manufacturer. The oligonucleotides were purified in two steps by ion-exchange on a MonoQ column (Amersham-Pharmacia Biotech, Uppsala, Sweden) with a gradient elution at pH 13 and 1:1 ethanol:acetone.

For identification of the genes mutated in RU360 and RU361, the region surrounding the transposon was cloned and subclones sequenced with the universal, reverse, or primer p113 (aggtcacatggaagtcagatc), which binds to the end of IS50.

DNA sequence analysis.

Sequencing data were aligned and analyzed with the Sequence Editor v1.03 (Applied Biosystems) and DNA Strider personal computer programs. Similarities between pairs of nucleic acid and amino acid sequences were carried out with the GCG (Genetics Computer Group, Madison, WI) version 7.0 suite of programs (Devereux et al. 1984).ORFs were determined by the CODONPREFERENCE and MAP programs. DNA and protein data base searches were performed with the BLAST and BESTFIT programs (Altschul et al. 1990). Sequences were aligned with the program PILEUP.

Nucleotide accession number.

The GenBank nucleotide accession number for the *R. leguminosarum* bv. *viciae* strain 1a *mocCABRDEF* gene cluster is AF076240.

Rhizopine catabolism tests.

To examine rhizobial strains for catabolism of *S*I and 3-*O*-M*S*I, the strains were grown in TY medium with aeration at 28°C. After 36 h the absorbence (580 nm) of each culture was read and 500 µl of the culture was harvested by centrifugation. The pelleted cells were washed with 0.85% NaCl and resuspended in Bergersen's minimal medium (Bergersen 1961). The optical density (OD) adjusted to 0.5 and 200 µl of the prepared cells was then mixed with 5 µl of *S*I or 3-*O*-M*S*I (as a sole carbon source) in sterile polystyrene tubes (Elkay Products, Shrewsbury, MA) and the cultures were shaken at 28°C for 3 days. All treatments were performed at least twice and a bacterial strain able to catabolize rhizopines (Moc^+) and a strain unable to catabolize rhizopine (Moc^-) were used as controls. Standards of 3-*O*-M*S*I and *S*I were applied to the chromatogram. Rhizopine catabolic activity of strains was determined by high voltage paper electrophoresis (HVPE) of the incubation medium; disappearance of the rhizopine indicated catabolism had occurred. For this purpose, 4-µl aliquots of the test samples were briefly centrifuged in a bench top centrifuge and the supernatant applied to the center of Whatman 1 Chr chromatography paper. After the aliquots were dried with cool air, the loaded paper was saturated with formic/acetic acid buffer (28.4 ml of 98% formic acid, 59.2 ml of glacial acetic acid in 1 liter of water, pH 1.75) and electrophoresed for 20 to 45 min at 3,000 V in formic/acetic acid buffer with distilled perchloroethylene as the inert buffer. The paper was dried and stained with 0.4% $AgNO_3$ in acetone. After air drying, the paper was submersed in 2% NaOH in ethanol and then gently steamed until brown spots corresponding to rhizopine standards appeared. The paper was dipped in thiosulfate fixative (sodium thiosulfate 100 g and sodium metabisulfate 15 g in 1 liter of water) and washed in gently flowing tap water for 1 to 2 h.

*S*I was synthesized by the protocol of Anderson and Lardy (1950) as described previously (Saint et al. 1993).

3-*O*-M*S*I was purified from nodules as described (Murphy et al. 1987) except that nodule extracts were prepared in H_2O instead of HCl. Nonspecific substrates were removed by biological enrichment and cation exchange chromatography, with the concentrated rhizopine being filter sterilized. The amount of rhizopine added to the catabolic tests was such that the final concentration in a 4-µl sample gave an easily visible spot after HVPE and $AgNO_3$ staining. This represents a final concentration of approximately 200 µg/ml.

ACKNOWLEDGMENTS

We wish to thank Nicki Featherstone for expert technical help, Leif Scøt and Barry Rolfe for strains, and Chikezie Owuama for reading the manuscript. Allan Downie is thanked for strain A118 and for pointing out the GTG start for MocB. This work was supported by a grant from the Australian Research Council (A09601063). M. B. acknowledges the support of an Iranian Government Postgraduate Scholarship.

LITERATURE CITED

Altschul, S .F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Ames, G. F.-L. 1986. Bacterial periplasmic transport systems: Structure, mechanism and evolution. Annu. Rev. Biochem. 55:397-425.

Anderson, L., and Lardy, H. A. 1950. Stereochemical studies in the aminodesoxyinositol series *meso*-inosamine-2 and *scyllo*-inosamine. J. Am. Chem. Soc. 72:3141-3147.

Anderson, W. A., and Magasanik, B. 1971. The pathway of *myo*-inositol degradation in *Aerobacter aerogenes* conversion of 2-deoxy-5-keto-D-gluconic acid to glycolytic intermediates. J. Biol. Chem. 246:5662-5675.

Assinder, S. J., and Williams, P. A. 1990. The TOL plasmids: Determinants of the catabolism of toluene and the xylenes. Adv. Microb. Physiol. 31:1-69.

Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J., and Timmis, K. N. 1981. Specific-purpose plasmid cloning vectors II. Broad host range high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene 16: 237-247.

Bergersen, F. J. 1961. The growth of *Rhizobium* in synthetic media. Aust. J. Biol. Sci. 14:349-360.

Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum.* J. Gen. Microbiol. 84:188-198.

Boivin C., Barran, L. R., Malpica, C. A., and Rosenberg, C. 1991. Genetic analysis of a region of the *Rhizobium meliloti* pSym plasmid specifying catabolism of trigonelline, a secondary metabolite present in legumes. J. Bacteriol. 173:2809-2817.

Boivin, C., Camut, S., Malpica, C. A., Truchet, G., and Rosenberg, C. 1990. *Rhizobium meliloti* genes encoding catabolism of trigonelline are induced under symbiotic conditions. Plant Cell 2:1157-1170.

Boyer, H. W., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41:459-472.

Brewin, N. J., Beringer, J. E., and Johnston, A. W. B. 1980. Plasmid-mediated transfer of host-range specificity between two strains of *Rhizobium leguminosarum*. J. Gen. Microbiol. 120:413-420.

Buck, D., and Guest, J. R. 1989. Overexpression and site-directed mutagenesis of the succinyl-CoA synthetase of *Escherichia coli* and nucleotide sequence of a gene (g30) that is adjacent to the *suc* operon. Biochem. J. 260:737-747.

Casadaban, M. J., and Cohen, S. N. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli.* J. Mol. Biol. 138:179-207.

Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. 1993. Sequence and analysis of the genomic locus responsible for surfactin synthesis in *Bacillus subtilis*. Mol. Microbiol. 8:821-831.

Deshusses, J., and Reber, G. 1977. *myo*-inositol transport in *Klebsiella aerogenes*: *scyllo*-inositol, a non-metabolizable substrate for the study of the *myo*-inositol transport system. Eur. J. Biochem. 72:87-91.

Devereux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-395.

Dong, J. M., Taylor, J. S., Latour, D. J., Iuchi, S., and Lin, E. C. C. 1993. Three overlapping *lct* genes involved in L-lactate utilization by *Escherichia coli*. J. Bacteriol. 175:6671-6678.

Donner, S. C., Jones, D. A., McClure, N. C., Rosewarne, G. M., Tate, M. E., Kerr, A., Fajardo, N. N., and Clare, B. G. 1993. Agrocin 434, a new plasmid encoded agrocin from the biocontrol *Agrobacterium* strains K84 and K1026, which inhibits biovar 2 agrobacteria. Physiol. Mol. Plant Pathol. 42:185-194.

Downie, J. A., Ma, Q.-S., Knight, C. D., Hombrecher, G., and Johnston, A. W. B. 1983. Cloning of the symbiotic region of *Rhizobium leguminosarum:* The nodulation genes are between the nitrogenase genes and a *nifA*-like gene. EMBO J. 2:947-952.

Eckhardt, T. 1978. A rapid method for the identification of deoxyribonucleic acid in bacteria. Plasmid 1:584-588.

Erickson, B. D., and Mondello, F. J. 1992. Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated-biphenyl-degrading enzyme in *Pseudomonas* strain LB400. J. Bacteriol. 174:2903-2912.

Figurski, D. H., and Helinski, D. R. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function *in trans*. Proc. Natl. Acad. Sci. USA 76:1648-1652.

Fujita, Y., Shindo, K., Miwa, Y., and Yoshida, K. 1991. *Bacillus subtilis* inositol-dehydrogenase-encoding gene (*idh*); sequence and expression in *Escherichia coli*. Gene 108:121-125.

George, D. G., Hunt, L. T., Yeh, L.-S. L., and Barker, W. C. 1985. New perspectives on bacterial ferredoxin evolution. J. Mol. Evol. 22:20-31.

Gordon, D. M., Ryder, M. H., Heinrich, K., and Murphy, P. J. 1996. An

experimental test of the rhizopine concept in *Rhizobium meliloti*. Appl. Environ. Microbiol. 62:3991-3996.

Groarke, J. M., Mahoney, W. C., Hope, J. N., Furlong, C. E., Robb, F. T., Zalkin, H., and Hermodson, M. A. 1983. The amino acid sequence of D-ribose-binding protein from *Escherichia coli* K12. J. Biol. Chem. 258:12952-12956.

Grönger, P., Priefer, U. B., Klipp, W., and Pühler, A. 1988. The *Rhizobium leguminosarum nifA, nifB* gene region can be used for complementation of *Rhizobium meliloti* mutants. Page 367 in: Nitrogen Fixation: Hundred Years After. H. Bothe, F. M. de Bruijn, and W. E. Newton, eds. Gustav Fischer, Stuttgart, Germany.

Hanukoglu, I., and Gutfinger, T. 1989. cDNA sequence of adrendoxin reductase. Identification of NADP-binding sites in oxidoreductases. Eur. J. Biochem. 180:479-484.

Harayama, S., and Kok, M. 1992. Functional and evolutionary relationships among diverse oxygenases. Annu. Rev. Microbiol. 46:565-601.

Harnisch, U., Weiss, H., and Sebald, W. 1985. The primary structure of the iron-sulfur subunit of ubiquinol-cytochrome *c* reductase from *Neurospora*, determined by cDNA and gene sequencing. Eur. J. Biochem. 149:95-99.

Hartwig U. A., Joseph, C. M., and Phillips, D. A. 1991. Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti.* Plant Physiol. 95:797-803.

Haydon, D. J., and Guest, J. R. 1991. A new family of bacterial regulatory proteins. FEMS Microbiol. Lett. 79:291-296.

Hogg, R. W., Voelker, C., and Carlowitz, I. V. 1991. Nucleotide sequence and analysis of the *mgl* operon of *Escherichia coli* K12. Mol. Gen. Genet. 229:453-459.

Irie, S., Doi, S., Yorifuji, T., Takagi, M., and Yano, K. 1987. Nucleotide sequencing and characterization of the genes encoding benzene oxidation enzymes of *Pseudomonas putida*. J. Bacteriol. 169:5174-5179.

Jensen, E. S. 1987. Inoculation of pea by application of *Rhizobium* in the planting furrow. Plant Soil 97:63-70.

Johnston, A. W. B., and Beringer, J. E. 1975. Identification of the *Rhizobium* strains in pea root nodules using genetic markers. J. Gen. Microbiol. 87:343-350.

Kanagasundaram, V., and Scopes, P. K. 1992. Cloning, sequence analysis, and expression of the structural gene encoding glucose-fructose oxidoreductase from *Zymomonas mobilis*. J. Bacteriol. 174:1439-1447.

Klipp, W., Reiländer, H., Schlüter, A., Reiner, K., and Pühler, A. 1989. The *Rhizobium meliloti fdxN* gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with *nifA* and *nifB*. Mol. Gen. Genet. 216:293-302.

Knauf, V. C., and Nester, E. W. 1982. Wide host range cloning vector: A cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid 8:45-54.

Kok, M., Oldenhuis, R., van der Linden, M. P. G., Meulenberg, C. H. C., Kingma, J., and Witholt, B. 1989a. The *Pseudomonas oleovorans alkBAC* operon encodes two structurally related rubredoxins and an aldehyde dehydrogenase. J. Biol. Chem. 264:5442-5451.

Kok, M., Oldenhuis, R., van der Linden, M. P. G., Raatjes, P., Kingma, J., van Leyveld, P. H., and Witholt, B. 1989b. The *Pseudomonas oleovorans* alkane hydroxylase gene. J. Biol. Chem. 264:5435-5441.

Kurkela, S., Lehväslaiho, H., Palva, E. T., and Teeri, T. H. 1988. Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB9816. Gene 73:355-362.

Kyte, J., and Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Lai, X., and Ingram, L. O. 1993. Cloning and sequencing of a cellobiose phosphotransferase system operon from *Bacillus stearothermophilus* XL-65-6 and functional expression in *Escherichia coli.* J. Bacteriol. 175:6441-6450.

Mansouri, K., and Piepersberg, W. 1991. Genetics of streptomycin production in *Streptomyces griseus:* Nucleotide sequence of five genes, *strFGHIK*, including a phosphate gene. Mol. Gen. Genet. 228:459-469.

Masepohl, B., Kutsche., M., Riedel, K.-U., Schmehl, M., Klipp, W., and Pühler, A. 1992. Functional analysis of the cysteine motifs in the ferredoxin-like protein FdxN of *Rhizobium meliloti* involved in symbiotic nitrogen fixation. Mol. Gen. Genet. 233:33-41.

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. 149:114-122.

Meyer, J., Bruschi, M. H., Bonicel, J. J., and Bovier-Lapierre, G. E. 1986. Amino acid sequence of [2Fe-2S] ferredoxin from *Clostridium pasteurianum*. Biochemistry 25:6054-6061.

Morrice, N., Geary, P., Cammack, R., Harris, A., Beg, F., and Aitken, A. 1988. Primary structure of protein B from *Pseudomonas putida*, member of a new class of 2Fe-2S ferredoxins. FEBS Lett. 231:336-340.

Müller-Hill, B. 1983. Sequence homology between Lac and Gal repressors and three sugar-binding periplasmic proteins. Nature 302:163-164.

Murphy, P. J., Heycke, N., Banfalvi, Z., Tate, M. E., de Bruijn, F., Kondorosi, A., Tempé, J., and Schell, J. 1987. Genes for the catabolism and synthesis of an opine-like compound in *Rhizobium meliloti* are closely linked and on the *sym* plasmid. Proc. Natl. Acad. Sci. USA 84:493-497.

Murphy, P. J., Heycke, N., Trenz, S., Ratet, P., de Bruijn, F. J., and Schell, J. 1988. Synthesis of an opine-like compound - a rhizopine - in alfalfa nodules is symbiotically regulated. Proc. Natl. Acad. Sci. USA 85:9133-9137.

Murphy, P. J., and Saint, C. P. 1992. Rhizopines in the legume-*Rhizobium* symbiosis. Pages 377-390 in: Molecular Signals in Plant-Microbe Communication. D. P. S. Verma, ed. CRC Press, Boca Raton, FL.

Murphy, P. J., Trenz, S. P., Grzemski, W., de Bruijn, F. J., and Schell, J. 1993. The *Rhizobium meliloti* rhizopine *mos* locus is a mosaic structure facilitating its symbiotic regulation. J. Bacteriol. 175:5193-5204.

Murphy P. J., Wexler, M., Grzemski, W., Rao, J. P., and Gordon, D. 1995. Rhizopines - their role in symbiosis and competition. Soil Biol. Biochem. 27:525-529.

Nagy, I., Schoofs, G., Compernolle, F., Proost, P., Vanderleyden, J., and De Mot, R. 1995. Degradation of the thiocarbomate herbicide EPTC (*S*-ethyl dipropylcarbamothioate) and biosafening by *Rhodococcus* sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase. J. Bacteriol. 177:676-687.

Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M., and Harayama, S. 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.

Neidle, E. L., and Kaplan, S. 1992. *Rhodobacter sphaeroides rdxA*, a homolog of *Rhizobium meliloti fixG*, encodes a membrane protein which may bind cytoplasmic [4Fe-4S] clusters. J. Bacteriol. 174: 6444-6454.

Niemi, J., and Mäntsälä, P. 1995. Nucleotide sequences and expression of genes from *Streptomyces pupurascens* that cause the production of new anthracyclines in *Streptomyces galilaeus*. J. Bacteriol. 177:2942-2945.

Pabo, C. O., and Sauer, R. T. 1984. Protein DNA recognition. Annu. Rev. Biochem. 53:293-321.

Peterson, J. A., Lorence, M. C., and Amarneh, B. 1990. Putidaredoxin reductase and putidaredoxin. Cloning, sequence determination, and heterologous expression of the proteins. J. Biol. Chem. 265:6066-6073.

Poole, S., P., Blyth, A., Reid, C. J., and Walters, K. 1994. *myo*-inositol catabolism and catabolite regulation in *Rhizobium leguminosarum* bv. *viciae*. Microbiology 140:2787-2795.

Pugsley, A. P. 1993. The complete general secretory pathway in Gram-negative bacteria. Microbiol. Rev. 57:50-108.

Rao, J. P., Grzemski, W., and Murphy, P. J. 1995. *Rhizobium meliloti* lacking *mosA* synthesizes the rhizopine *scyllo*-inosamine in place of 3-*O*-methyl-*scyllo*-inosamine. Microbiology 141:1683-1690.

Rossbach, S., Kulpa, D. A., Rossbach, U., and de Bruijn, F. J. 1994. Molecular and genetic characterization of the rhizopine catabolism (*mocABRC*) genes of *Rhizobium meliloti* L5-30. Mol. Gen. Genet. 245:11-24.

Saint, C. P., Wexler, M., Murphy, P. J., Tempé, J., Tate, M. E., and Murphy, P. J. 1993. Characterization of genes for synthesis and catabolism of a new rhizopine induced in nodules by *Rhizobium meliloti* Rm220-3: Extension of the rhizopine concept. J. Bacteriol. 175:2505-2515.

Sambrook, J., Fritsch, E. F., and Maniatis, T. A. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., and Coulson, R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

Sato, N., Seyama, Y., and Mutata, N. 1986. Lipid-linked desaturation of palmitic acid in monogalactosyl diacyl glycerol in the bluegreen alga (cyanobacterium) *Anabaena variabilis* studied *in vivo*. Plant Cell Physiol. 27:819-835.

Sensen, C. W., Klenk, H.-P., Singh, R. K., Allard, G., Chan, C. C.-Y., Liu, Q. Y., Penny, S. L., Young, F., Eschenk, M. E., Gassterland, T., Doolittle, W. F., Ragan, M. A., and Charlebois, R. L. 1996. Organization characteristics and information content of an archaeal genome: 156kb of sequence from *Sulfolobus solfataricus* P2. Mol. Microbiol. 22:175-191.

Shanklin, J., Whittle, E., and Fox, G. 1994. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-coA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry 33:12787-12794.

Simon, R., Quandt, J., and Klipp, W. 1989. New derivatives of transposon Tn*5* suitable for mobilization of replicons, generation of operon fusions and induction of genes in Gram-negative bacteria. Gene 80: 161-169.

Sjöström, M., Wold, S., Wieslander, Å., and Rilfors, L. 1987. Signal peptide amino acid sequences in *Escherichia coli* contain information related to final protein localization. A multivariate data analysis. EMBO J. 6:823-831.

Sofia, H. J., Burland, V., Daniels, D. L., Plunkett, G., III, and Blattner, F. R. 1994. Analysis of the *Escherichia coli* genome V. DNA sequence of the region from 76.0 to 81.5 minutes. Nucleic Acids Res. 22:2576-2586.

Stout, C. D. 1982. Iron-sulfur protein crystallography. Pages 97-146 in: Iron-Sulfur Proteins. T. G. Spiro, ed. Wiley-Interscience, New York.

Sumiya, M., and Henderson, P. J. F. 1989. The D-xylose binding protein of *Escherichia coli*. Biochem. Soc. Trans. 17:553-554.

Surin, B. P., Watson, J. M., Hamilton, W. D. O., Economou, A., and Downie, J. A. 1990. Molecular characterization of the nodulation gene, *nodT*, from two biovars of *Rhizobium leguminosarum*. Mol. Microbiol. 4:245-252.

Suzuki, M., Hayakawa, T., Shaw, J. P., Rekik, M., and Harayama, S. 1991. Primary structure of xylene monooxygenase: Similarities to and differences from the alkane hydroxylation system. J. Bacteriol. 173: 1690-1695.

Tan, H. M., Tang, H. Y., Joannou, C. L., Abdel-Wahab, N. H., and Mason, J. R. 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.

Tanaka, M., Haniu, M., and Yasunobu, T. 1974. The amino acid sequence of putidaredoxin, an iron-sulfur protein from *Pseudomonas putida*. J. Biol. Chem. 249:3689-3701.

Titgemeyer, F., Reizer, J., Reizer, A., Tang, J., Parr, T. R., Jr., and Saier, M. H., Jr. 1995. Nucleotide sequence of the region between *crr* and *cysM* in *Salmonella typhimurium:* Five novel ORFs including one encoding a putative transcriptional regulator of the phosphotransferase system. DNA Seq. 5:145-152.

Ueda, T., Lode, E. T., and Coon, M. J. 1972. Enzymatic ω-oxidation VI. Isolation of homogeneous reduced diphosphopyridine nucleotide-rubredoxin reductase. J. Biol. Chem. 247:2109-2116.

van Egeraat, A. W. S. M. 1975. The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings. Plant Soil 42:381-386.

Vlassak, K. M., and Vanderleyden, J. 1997. Factors influencing nodule occupancy by inoculant rhizobia. Crit. Rev. Plant Sci. 16:163-229.

Wada, H., Gombos, Z., and Murata, N. 1990. Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. Nature 347: 200-203.

Wang, Z., and Rossman, T. G. 1994. Isolation of DNA fragments from agarose gel by centrifugation. Nucleic Acids Res. 22:2862-2863.

Wexler, M., Gordon, D., and Murphy, P. J. 1995. The distribution of inositol rhizopine genes in *Rhizobium* populations. Soil Biol. Biochem. 27:531-537.

Worsey, M., and Williams, P. A. 1975. Metabolism of toluene and xylenes by *Pseudomonas putida* (*arvilla*) mt-2: Evidence for a new function of the TOL plasmid. J. Bacteriol. 124:7-13.

Yadav, N. S., Wierzbicki, A., Aegerter, M., Caster, C. S., Pérez-Grau, L., Kinney, A. J., Hitz, W. D., Booth, R., Jr., Schweiger, B., Stecca, K. L., Allen, S. M., Blackwell, M., Reiter, R. S., Carlson, T. J., Russel, S. H., Feldmann, K. A., Pierce, J., and Browse, J. 1993. Cloning of higher plant ω-3 fatty acid desaturases. Plant Physiol. 103:467-476.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33:103-119.

Yoshida, K.-I., Aoyama, D., Ishio, I., Shibayama, T., and Fujita, Y. 1997. Organization and transcription of the *myo*-inositol operon, *iol*, of *Bacillus subtilis*. J. Bacteriol. 179:4591-4598.

Yoshida, K., Sano, H., Miwa, Y., Ogasawara, N., and Fujita, Y. 1994. Cloning and nucleotide sequencing of a 15kb region of the *Bacillus subtilis* genome containing the *iol* operon. Microbiology 140:2289-2298.

Zylstra, G. J., and Gibson, D. T. 1989. Toluene degradation by *Pseudomonas putida* F1. Nucleotide sequence of the *todC1C2BADE* genes and their expression in *Escherichia coli*. J. Biol. Chem. 264:14940-14946.