JOURNAL OF BACTERIOLOGY, Aug. 1993, p. 5205–5215
0021-9193/93/165205-11$02.00/0

Vol. 175, No. 16

Characterization of Genes for Synthesis and Catabolism of a New Rhizopine Induced in Nodules by *Rhizobium meliloti* Rm220-3: Extension of the Rhizopine Concept

CHRISTOPHER P. SAINT,[1]† MARGARET WEXLER,[1] PHILLIP J. MURPHY,[1] JACQUES TEMPÉ,[3] MAX E. TATE,[2] AND PETER J. MURPHY[1]*

Department of Crop Protection[1] and Department of Plant Science,[2] Waite Institute, University of Adelaide, Glen Osmond, South Australia 5064, Australia, and Institut des Science Vegetales, Centre National de la Recherche Scientifique, 91198 Gif Sur Yvette, France[3]

Received 29 March 1993/Accepted 10 June 1993

Rhizopines are selective growth substrates synthesized in nodules only by strains of rhizobia capable of their catabolism. We report the isolation and study of genes for the synthesis and catabolism of a new rhizopine, *scyllo*-inosamine (sIa), from alfalfa nodules induced by *Rhizobium meliloti* Rm220-3. This compound is similar in structure to the previously described rhizopine 3-*O*-methyl-*scyllo*-inosamine from *R. meliloti* L5-30 (P. J. Murphy, N. Heycke, Z. Banfalvi, M. E. Tate, F. J. de Bruijn, A. Kondorosi, J. Tempé, and J. Schell, Proc. Natl. Acad. Sci. USA 84:493–497, 1987). The synthesis (*mos*) and catabolism (*moc*) genes for the Rm220-3 rhizopine are closely linked and located on the *nod-nif* Sym plasmid. The *mos* genes are directly controlled by the NifA/NtrA regulatory system. A comparison of the sequence of the 5′ regions of the two *mos* loci shows very extensive conservation of sequence as well as strong homology to the *nifH* coding region. Restriction mapping and hybridization to DNA from the four open reading frames (ORFs) of the L5-30 *mos* locus indicate the absence of *mosA* and presence of the other three ORFs (ORF1 and *mosB* and -*C*) in Rm220-3. We suggest that the L5-30 *mosA* gene product is involved in the conversion of *scyllo*-inosamine to 3-*O*-methyl-*scyllo*-inosamine. Restriction fragment length polymorphism analysis of the *moc* regions of both strains shows that they are very similar. Regulation studies indicate that the *moc* region is not controlled by the common regulatory genes *nifA*, *ntrA*, and *ntrC*. We discuss the striking similarities in gene structure, location, and regulation between these two rhizopine loci in relation to the rhizopine concept.

Rhizobia form symbiotic associations with leguminous plants which result in the conversion of atmospheric nitrogen into a form which the plant can utilize. The rhizobia are thought to benefit from the interaction as they are provided with nutrients (derived from plant photosynthates) and a temporary shelter from the soil environment (the nodule). However, for much of the time rhizobia survive as saprophytic organisms in the soil or the rhizosphere in competition with other microorganisms. Many factors, such as the availability of nutrients from plant root exudates, determine which rhizobia survive and eventually predominate in the rhizosphere (5, 19).

A number of examples of plant-associated products increasing rhizobial growth rate have been described. These include compounds produced by host plants (for example, trigonelline [6, 10], which is catabolized by a variety of rhizobia) and certain flavonoids which increase rhizobial growth (26). Other compounds such as calystegins are produced by nonhost plants and are catabolized by only a limited number of rhizobia as well as a variety of other soil microorganisms (58, 59). A more selective compound found in the exudate of pea roots is L-homoserine, an amino acid catabolized by pea-nodulating *Rhizobium leguminosarum* bv. *viciae* strains but by few other rhizobia (62).

In a number of cases, the locations of the catabolic genes for these compounds have been investigated. Trigonelline (8, 10) and L-homoserine (21) catabolic genes are located on the Symbiotic (Sym) plasmid, which carries genes involved in nitrogen fixation (*nif*) and nodulation (*nod*), whereas calystegin catabolic genes in *Rhizobium meliloti* 41 are encoded on the cryptic plasmid pRme41*a* (59). The Sym plasmid location of catabolic genes points towards a symbiotic role for these compounds.

By far the most specific nutritional interaction of plants and rhizobia is the one described for the rhizopine produced in alfalfa (*Medicago sativa*) nodules induced by *R. meliloti* L5-30 (37, 57). This system involves a particular rhizobium directing the plant to produce a compound in nodules which can be utilized by the same rhizobium (but by few others) as a selective growth substrate. In this respect, this compound is analogous to *Agrobacterium* opines. A generic class of compounds called rhizopines (37)—nodule specific opine-like compounds—was coined to describe these substrates. Since then, a number of other *Rhizobium* strains which induce and catabolize rhizopines have been isolated (39). Furthermore, Scott et al. (51) isolated a compound, rhizolotine, from *Lotus* nodules which has many properties of a rhizopine.

Analyses of the synthesis and catabolism genes of the L5-30 rhizopine, 3-*O*-methyl-*scyllo*-inosamine (3-*O*-MSI), have strongly reinforced a symbiotic role for this compound. These synthesis (*mos*) and catabolism (*moc*) genes have been isolated and shown to be of bacterial origin, closely linked (4.5 kb apart), and located on the *nod-nif* Sym plasmid (37). This suggests not only that these genes have coevolved as a functional unit but also that they are important in symbiosis. Further support for a role in symbiosis was

* Corresponding author. Electronic mail address: pmurphy@waite.adelaide.edu.au.

† Present address: Department of Microbiology, Monash University, Clayton, Victoria 3168, Australia.

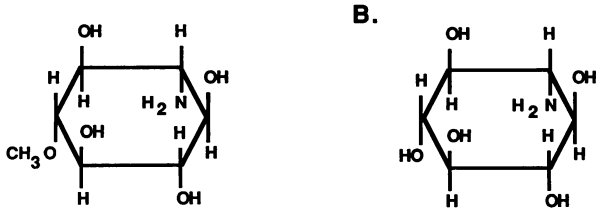

FIG. 1. Structures of 3-*O*-MSI (A) and *s*Ia (B).

gained by analyzing the regulation of the *mos* genes. This locus is directly controlled by a NifA/NtrA-regulated promoter, thus ensuring that the rhizopine is produced only when other symbiotic genes are functioning (38). Another feature which distinguishes rhizopines from the aforementioned compounds, which are plant secondary metabolites, is that they are synthesized by bacteroids within the nodule utilizing plant-derived products and catabolized as a nutrient source by the free-living bacteria.

From a screening of 20 *R. meliloti* strains, one (Rm220-3) was found to produce a compound in nodules which was revealed as a silver-staining spot of a mobility similar to that of 3-*O*-MSI upon high-voltage paper electrophoresis (HVPE) of extracts. Here, we describe identification of this compound as *scyllo*-inosamine (*s*Ia), which is a member of an inositol class of compounds closely related to 3-*O*-MSI (Fig. 1). We also demonstrate that *s*Ia is a rhizopine by virtue of the producing strain being able to catabolize this compound. The synthesis and catabolism genes have been isolated, their regulation has been investigated, and we compare our results with those for the previously described rhizopine genes from *R. meliloti* L5-30.

MATERIALS AND METHODS

Bacterial strains and plasmids. The *R. meliloti*, *Agrobacterium tumefaciens*, and *Escherichia coli* strains and the plasmids used and constructed during the course of this investigation are detailed in Table 1.

Media and culture conditions. *Rhizobium* and *Agrobacterium* strains were grown at 28°C in TY complex medium (4) or GTS minimal medium (30). *E. coli* was grown at 37°C in Luria-Bertani medium (36). All bacterial matings were performed at 28°C on TY medium as previously described (31).

Strain construction. Rm220-3 derivatives carrying Tn*5*-Mob were obtained from matings between *E. coli* S17-1(pSUP5011) and Rm220-3. Transconjugants were selected on GTS media containing 500 μg of kanamycin per ml. *E. coli* HB101(pJB3JI) was mass mated with Tn*5*-Mob-mutagenized Rm220-3 derivatives, and transconjugants were selected on GTS media containing 500 μg of kanamycin and 10 μg of tetracycline per ml. Transconjugants were mass mated with *A. tumefaciens* C58C1RS, and mobilization of Tn*5*-Mob-marked plasmids into C58C1RS by pJB3JI was selected for on TY media containing kanamycin and rifampin at 100 μg/ml each. Acquisition of an extra megaplasmid was identified by visualization on agarose gels following colony purification.

Clones were introduced into Rm1021 and its mutant derivatives for estimation of *s*Ia synthesis and catabolism by triparental mating (18) utilizing pRK2013 as the mobilizing plasmid.

***s*Ia synthesis and catabolism studies.** *s*Ia was chemically synthesized by the protocol of Anderson and Lardy (1), and its structure was confirmed by gas chromatography and mass spectrometry of the hexa-acetate. This was used as a standard in the identification of the *s*Ia produced by *R. meliloti* Rm220-3 and strains bearing cloned rhizopine synthesis genes. Synthetic *s*Ia was also used as a substrate in catabolism experiments. To test for *s*Ia synthesis, *R. meliloti* transconjugants were inoculated onto alfalfa plants grown on agar in glass tubes as previously described (30), and after 4 to 6 weeks resulting nodules were extracted for electrophoresis (37). To test for catabolism, *R. meliloti* transconjugants were inoculated into Bergersen minimal medium (3) supplemented with synthesized *s*Ia as a sole carbon source as previously described (37). HVPE and paper electrophoresis were performed by standard procedures (16, 23). Buffers used were formic-acetic acid, pH 1.7 (28.4 ml of 98% formic acid–59.2 ml of glacial acetic acid per liter), 0.05 M citric acid, pH 6.4 (10.5 g of citric acid per liter; pH was adjusted with NaOH), and 0.05 M $Na_2B_4O_7$, pH 9.4 (19.07 g of $Na_2B_4O_7 \cdot 10H_2O$ per liter). *s*Ia concentration in nodules was estimated by HVPE of extracts against chemically synthesized *s*Ia of known concentration.

Plasmid DNA extraction and manipulation. Plasmid DNA was purified by the alkali lysis method (48), followed by CsCl purification for which vector or cloned DNA was required in quantity. Detection of large plasmids was performed according to the method of Eckhardt (20). Total DNA was isolated as previously described (35), and construction of a gene bank in pVK102 was performed by standard procedures (48). Restriction endonuclease digestion and ligation with T4 DNA ligase was performed according to the manufacturer's instructions (Boehringer Mannheim). *E. coli* strains were transformed by standard procedures (14).

DNA-DNA hybridization. Plasmids and restriction fragments were separated by agarose gel electrophoresis and transferred by Southern blotting (54) to Hybond N+ (Amersham Ltd.). Restriction fragments used for the preparation of probes were excised from gels under long-wave UV light and purified by using the Geneclean protocol (Bio 101, La Jolla, Calif.). Radiolabelled probes were prepared by nick translation incorporating [α-^{32}P]dCTP by using a preparative kit (Bresatec, Ltd., Adelaide, Australia), followed by purification through Sephadex G-50. Hybridization and washing conditions were as described elsewhere (32). Fuji RX X-ray film at −80°C was used for autoradiography.

DNA sequencing. A 2.2-kb *Eco*RI fragment, derived originally from Rm220-3, was obtained from pPM1169 by *Eco*RI digestion followed by fragment isolation as described above. The purified fragment was ligated into *Eco*RI-digested M13mp18 (44), and recombinants were identified by standard procedures (48). DNA for sequencing was prepared by restriction digestion with *Xba*I and *Pst*I followed by exonuclease III-S1 nuclease digestion according to the method of Henikoff (27). The whole procedure was carried out with the Erase-a-base kit (Promega, Madison, Wis.). Dideoxy sequencing reactions (49) were performed with the Sequenase kit (U.S. Biochemicals, Cleveland, Ohio) using the universal primer supplied and ^{35}S-dATP (Bresatec, Ltd.). Both 6% linear and wedge polyacrylamide gels were used. DNA was sequenced on both strands and analyzed with the University of Wisconsin Genetics Computer Group sequence analysis series of programs (17).

Nucleotide accession number. The Rm220-3 *mos* locus nucleotide sequence determined in this study has been deposited in the GenBank data base under accession number L17073.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
R. meliloti		
Rm220-3	Mos$^+$ Moc$^+$ Str^r, produces *s*Ia and catabolizes *s*Ia and 3-*O*-M*S*I	Field isolate from Bielefeld, Germany, a gift from A. Pühler.
L5-30	Mos$^+$ Moc$^+$, produces and catabolizes 3-*O*-M*S*I	Isolate from Poland (33), a gift from J. Dénarié.
Rm1021	Mos$^-$ Moc$^-$	35
Rm1354	*nifA*::Tn*5*, Km^r derivative of Rm1021	56
Rm5002	*ntrC*::Tn*5*, Km^r derivative of Rm1021	55
Rm1491	*nifH*::Tn*5*, Km^r derivative of Rm1021	56
Rm1681	*ntrA*::Tn*5*, Km^r derivative of Rm1021	45
A. tumefaciens		
C58C1RS	Rif^r Str^r	13
E. coli		
HB101		11
S17-1	Chromosomally integrated RP4 derivative	53
Plasmids		
pRK2013	Km^r helper plasmid	22
pSUP5011	Tn*5*-Mob, Km^r	52
pJB3JI	Tra$^+$ Tc^r Km^s derivative of R68.45	12
pRmR2	Contains *nifH* and part of *nifD* of *R. meliloti* 102F34	47
pUC18	Cloning vector	63
pGEMEX-1	Expression cloning vector	Promega Corp.
pVK102	Tc^r Km^r cosmid cloning vector	29
pLAFR1	Tc^r cosmid cloning vector	24
pJRD184	Tc^r Ap^r	28
pJS201	Contains a 3.3-kb *nodABC* fragment from *R. meliloti* in pIN-II-A2	50
pPM1031	Mos$^-$ Moc$^+$, 15.1-kb fragment from L5-30 in pLAFR1, Tc^r	37
pPM1062	Mos$^+$ Moc$^-$, 3.4- and 6.9-kb *Eco*RI fragments of L5-30 in pLAFR1, Tc^r	37
pPM1146	2.5-kb *Pst*I fragment bearing L5-30 *mosB* and parts of *mosA* and *mosC* in pJRD184, Ap^s Tc^r	This study
pPM1153	Mos$^+$ Moc$^+$, 27-kb fragment of Rm220-3 in pVK102, Km^s Tc^r	This study
pPM1165	Mos$^-$ Moc$^-$, 3.5-kb *Hin*dIII fragment of pPM1153 in pVK102, Km^s Tc^r	This study
pPM1168	Mos$^-$ Moc$^-$, 3.5- and 1.0-kb *Hin*dIII fragments of pPM1153 in pVK102, Km^s Tc^r	This study
pPM1169	Mos$^+$ Moc$^-$, 3.5-, 1.0-, and 7.3-kb *Hin*dIII fragments of pPM1153 in pVK102, Km^s Tc^r	This study
pPM1171	Mos$^-$ Moc$^-$, pPM1153 with the 1.0-, 7.3-, and 7.1-kb *Hin*dIII fragments deleted, Km^s Tc^r	This study
pPM1175	Mos$^-$ Moc$^+$, pPM1153 with the 1.0- and 7.3-kb *Hin*dIII fragments deleted, Km^s Tc^r	This study
pPM1178	1-kb *Eco*RI-*Kpn*I *mos* fragment from pPM1062 in pUC18	This study
pPM1186	Mos$^-$ Moc$^+$, pPM1153 with the 3.5- and 7.1-kb *Eco*RI fragments deleted, Km^s Tc^r	This study
pPM1201	pGEMEX-1 containing a 965-bp *Kpn*I-*Xho*I fragment bearing the 3′ region of L5-30 ORF1	40
pPM1202	pGEMEX-1 containing a 763-bp *Bam*HI fragment internal to L5-30 *mosA*	40
pPM1203	pGEMEX-1 containing a 1,174-bp *Nsi*I-*Hin*dIII fragment bearing L5-30 *mosB*	40
pPM1204	pGEMEX-1 containing a 1,409-bp *Mlu*I-*Apa*I fragment bearing L5-30 *mosC*	40

RESULTS

Alfalfa nodules induced by *R. meliloti* Rm220-3 produce the rhizopine *s*Ia. An initial screening of 20 *R. meliloti* strains for the ability to induce rhizopines revealed that nodules induced by *R. meliloti* Rm220-3 produce a silver-staining compound, having a slightly greater mobility than 3-*O*-M*S*I from *R. meliloti* L5-30 nodules, when examined by HVPE in formic-acetic acid buffer, pH 1.7 (Fig. 2A, lanes 4 and 5, respectively). Further analysis of this compound revealed that it was a nonreducing polyol amine with charge characteristics and size similar to those of the reducing sugar glucosamine. The electrophoretic pH mobility profile indicated that no other ionizable groups were present. The electrophoretic borate-complexing behavior of this compound is characteristic of an equatorial inosamine which undergoes inversion upon heating to form a tridentate borate complex cation (23); the uninverted equatorial *s*Ia does not form a borate complex. Together, these data suggested that the compound was *s*Ia. *s*Ia was chemically synthesized and used in paper-chromatographic studies to verify the identity of the compound isolated from nodules. This compound comigrated with a chemically synthesized sample of *s*Ia to which nodule extract from strain Rm1021, a strain which does not produce *s*Ia (Fig. 2A, lane 2), was added (Fig. 2A, lane 3). After HVPE, the relative mobilities of *s*Ia, synthesized *s*Ia, and 3-*O*-M*S*I with respect to the orange G marker were −0.81, −0.81, and −0.87, respectively, and in 0.05 M citrate buffer (pH 6.4) they were −0.31, −0.31, and −0.34,

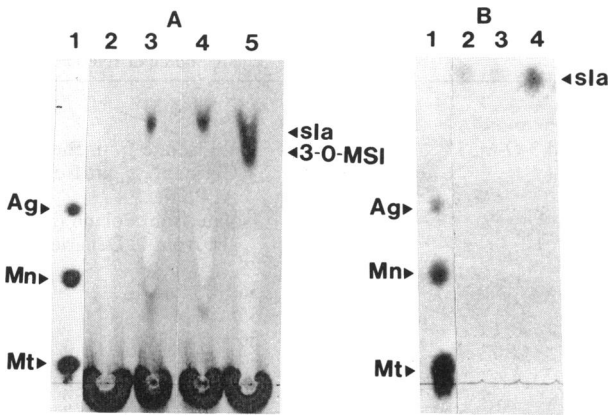

FIG. 2. Synthesis and catabolism of *s*Ia by Rm220-3. (A) Results of HVPE in formic-acetic acid buffer, pH 1.7, for nodule extracts of Rm1021 (lane 2), Rm1021 including chemically synthesized *s*Ia (lane 3), Rm220-3 (lane 4), and L5-30 (lane 5). (B) Results of HVPE of minimal medium supplemented with *s*Ia after a 5-day incubation with Rm220-3 (lane 2), L5-30 (lane 3), or Rm1021 (lane 4). Lanes 1, markers agropine (Ag), mannopine (Mn), and mannitol (Mt).

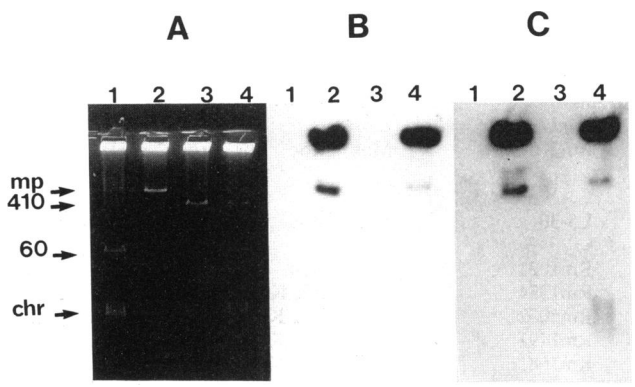

FIG. 3. Confirmation that *mos* and *moc* genes from Rm220-3 are on the *nod-nif* Sym plasmid. An Eckhardt plasmid gel (A) was transferred to membranes and probed with radiolabelled fragments from pPM1146 (bearing L5-30 *mosB* and parts of *mosA* and *mosC*) (B) and pJS201 (bearing *nodA*, *-B*, and *-C*) (C). Plasmid DNAs are as follows: lanes 1, HB101(pJB3JI) (60 kb); lanes 2, Rm220-3 (megaplasmid [mp]); lanes 3, C58C1RS (cryptic plasmid) (410 kb); lanes 4, C58C1RS transconjugant. chr, broken chromosomal DNA.

respectively (data not shown). Paper chromatography, in formic-acetic acid buffer, also revealed that *s*Ia from Rm220-3 and synthetic *s*Ia had identical mobilities. Two possible configurations for inosamine are possible, the *myo* and *scyllo* forms. HVPE in 0.05 M $K_2B_4O_7$ buffer (pH 9.2) gave mobility values of 0.45 for *myo*-inosamine and 0.01 for *s*Ia, as the uninverted *s*Ia does not form a borate complex and remains at the origin during electrophoresis (23). To further confirm its structure, the nodule compound was compared with an authentic sample (kindly provided by L. Anderson of the University of Wisconsin) and a synthetic sample of *s*Ia by gas chromatography-mass spectrometry of the hexa-acetate derivatives and found to be identical to *s*Ia.

We have estimated that Rm220-3 produces approximately 15 μg of *s*Ia per g of nodules (wet weight). Rm220-3 can also catabolize *s*Ia as a sole carbon and nitrogen source. After a 5-day incubation of Rm220-3 with synthetic *s*Ia as a substrate, most of the compound was utilized (Fig. 2B, lane 2). L5-30 could also utilize this substrate, whereas Rm1021 could not (Fig. 2B, lanes 3 and 4, respectively). As *s*Ia both is induced in the nodule by Rm220-3 and can be catabolized by this strain, this compound meets the definition of a rhizopine (39).

Isolation of the rhizopine synthesis and catabolism genes from Rm220-3. To isolate the rhizopine synthesis and catabolism genes from Rm220-3, it was assumed that because the rhizopines associated with Rm220-3 and L5-30 are structurally very similar there would be homology between the synthesis and catabolism genes of each strain. Accordingly, a clone bank of total DNA from Rm220-3 was prepared in the cosmid vector pVK102 and probed with a 2.5-kb *Pst*I fragment (pPM1146) from the L5-30 *mos* genes. In this manner, clone pPM1153 was isolated. This cosmid was mated into *R. meliloti* Rm1021 (which does not induce the production of the rhizopine or catabolize it) by triparental mating, and the transconjugant was tested for growth and catabolism of *s*Ia as a sole carbon source. Rm1021(pPM1153) was also inoculated onto alfalfa plants, and the nodules induced were analyzed for the production of *s*Ia. This strain could catabolize *s*Ia as a free-living bacterium as well as synthesize *s*Ia as an endosymbiont. pPM1153, therefore, contains functional *moc* and *mos* genes from Rm220-3.

***mos-moc* genes are on the *nod-nif* Sym plasmid.** When Southern blots of plasmid gels were probed with *nod* and *mos* probes, a plasmid band corresponding to the Rm220-3 megaplasmid hybridized with each probe (Fig. 3B and C, lanes 2). As many *R. meliloti* strains have two large megaplasmids (34) which often comigrate as a single band in agarose gels, it was necessary to show that the hybridization observed was that to the *nod-nif* Sym plasmid. Accordingly, transconjugants containing individually mobilized plasmids were constructed. These were prepared by introducing pSUP5011 bearing Tn*5*-Mob into Rm220-3 and using the helper plasmid pJB3JI to mobilize plasmids to *A. tumefaciens* C58C1RS. This recipient strain was chosen because it enabled easy visual resolution of the incoming plasmid from the resident plasmid and acted as a good recipient during conjugation experiments. Figure 3A, lane 4, shows DNA from a C58C1RS transconjugant containing a single megaplasmid species, a cryptic 410-kb plasmid, and the mobilizing plasmid pJB3JI. When plasmid DNA from this transconjugant was examined by DNA-DNA hybridization with a probe from the common *nod* region of pJS201 (Fig. 3B, lane 4) and a *mos* probe (pPM1146) (Fig. 3C, lane 4), the megaplasmid hybridized to both *nod* and *mos*, indicating that *mos* genes are situated on the *nod-nif* Sym plasmid of Rm220-3. As expected, since *moc* and *mos* genes are closely linked, a *moc* probe also hybridized to the Rm220-3 *nod-nif* Sym plasmid (data not shown). Evidence of Sym plasmid localization of these genes is further corroborated by physiological data, as the transconjugant C58C1RS(pSym) could catabolize *s*Ia (data not shown).

Analysis of *mos* genes from Rm220-3. pPM1153 was mapped with the restriction enzymes *Eco*RI and *Hin*dIII (Fig. 4). To localize the *moc* and *mos* functions on this plasmid, probes from the equivalent genes in L5-30 were prepared and hybridized to pPM1153.

A probe prepared from a 2.5-kb *Pst*I DNA fragment, which is internal to the L5-30 *mos* region and which contains *mosB* and parts of *mosA* and *mosC* (pPM1146), hybridized to 3.5- and 1.0-kb *Hin*dIII fragments of pPM1153 (Fig. 4). pPM1168, a subclone of pPM1153 which contains these fragments, when transferred to rhizobia was not sufficient to

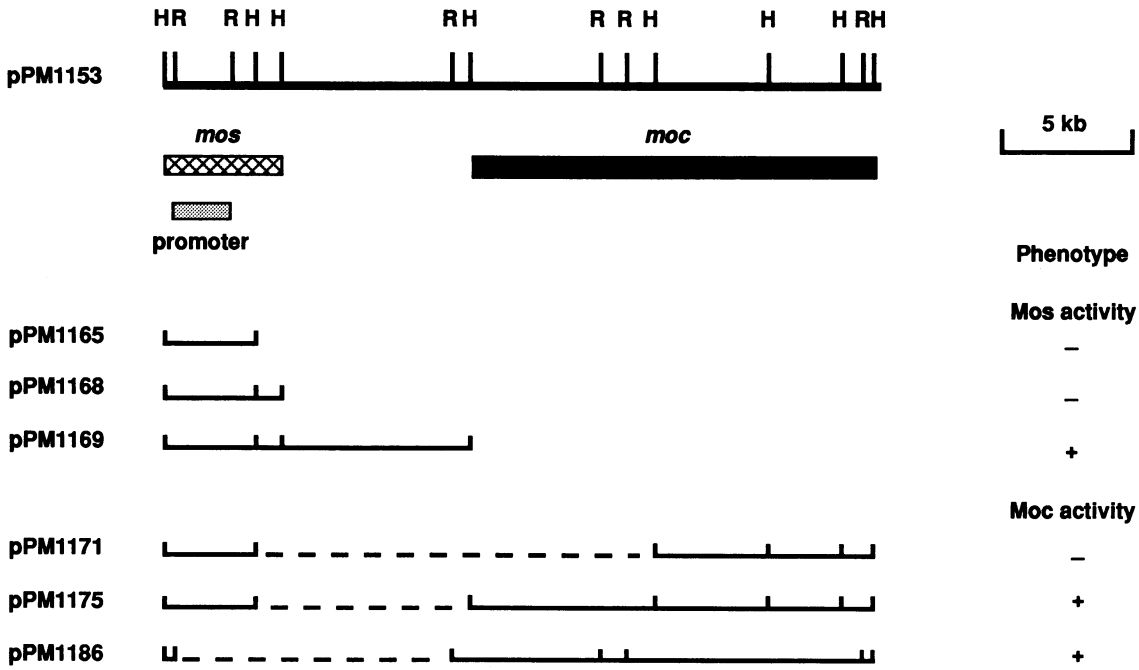

FIG. 4. *mos* and *moc* clones of Rm220-3. Regions of hybridization with the original cosmid clone pPM1153 using a *Pst*I radiolabelled fragment from pPM1146 (containing L5-30 *mosB* and parts of *mosA* and *mosC*) (crosshatched bar), radiolabelled pPM1031 (containing *moc* genes from L5-30) (closed bar), and a radiolabelled *Eco*RI-*Kpn*I fragment from pPM1178 (containing the promoter region of the *mos* locus from L5-30) (stippled bar) are shown. The various subclones constructed from pPM1153 are also shown along with the respective Moc and Mos phenotypes they bestow when introduced into Rm1021. Abbreviations: H, *Hin*dIII; R, *Eco*RI.

confer the ability to produce the rhizopine in nodules. However, pPM1169, which also contains an adjacent 7.3-kb *Hin*dIII fragment, is sufficient for rhizopine production in situ (Fig. 4). These fragments total 11.8 kb, but it is likely, by analogy with the L5-30 *mos* region, which is 4.8 kb in size, that only a portion of the nonhybridizing 7.3-kb *Hin*dIII fragment is required to express the Rm220-3 Mos phenotype.

The L5-30 *mos* genes were found to be regulated by the *nifA* gene (38). To determine whether Rm220-3 *mos* genes are similarly regulated, pPM1169, bearing the complete suite of *mos* genes from Rm220-3, was introduced into a variety of Rm1021 regulatory mutants. Transconjugants were used to induce nodules on alfalfa, and these were subsequently extracted and examined for the presence of *s*Ia. This compound was not produced when pPM1169 was present in a NifA$^-$ or NtrA$^-$ background, but normal production was obtained in NtrC$^-$ and NifH$^-$ mutants bearing this plasmid (Table 2). This suggests that NifA and the common bacterial regulatory sigma factor NtrA are involved in *mos* regulation in Rm220-3. Either there could be a requirement for NifA/NtrA-regulated functions for *s*Ia production or *mos* genes could be directly regulated, as is the case for these genes in strain L5-30 (38). *s*Ia production by a NifH$^-$ mutant, which is Fix$^-$, suggests that this control is direct and that nitrogen fixation per se is not required for *mos* gene function.

TABLE 2. Expression of Mos and Moc in regulatory mutants of *R. meliloti*[a]

Strain	Relevant phenotype	Mos
Rm1354	NifA$^-$	−
Rm5002	NtrC$^-$	+
Rm1491	NifH$^-$	+
Rm1681	NtrA$^-$	−
Rm1021	WT[b]	+

[a] Plasmids pPM1169 (for the Mos phenotype) and pPM1153 (for the Moc phenotype) were mated into the above strains, and phenotypes were determined as described in Materials and Methods. All strains expressed Moc.
[b] WT, wild type.

NifA regulation of L5-30 *mos* genes occurs directly via the *mos* promoter, which is similar to the *nifHDK* operon promoter (38). In addition, proximal to the *mos* promoter there is a 57-bp region which is highly homologous to the 5′ end of the *nifH* coding region. Initially, we investigated by hybridization whether Rm220-3 might also contain a similar promoter and 5′ region. A 1.0-kb *Eco*RI-*Kpn*I fragment from pPM1178 which contains the NifA-regulated promoter from the L5-30 *mos* locus (38) was labelled and probed against *Eco*RI-digested pPM1153 DNA. A 2.2-kb fragment of DNA hybridized (Fig. 4). Similarly, a 190-bp *Alu*I fragment (derived from pRmR2) which contains the 5′ portion of the *nifH* coding region also hybridized to this fragment (data not presented). Together, these data indicate that the structures of the promoter and proximal region of Rm220-3 are similar to those of the equivalent regions in strain L5-30.

To further investigate the structure of the 5′ region of the Rm220-3 *mos* locus, the 2.2-kb *Eco*RI fragment from Rm220-3 was subcloned into M13mp18, nested deletions were prepared by using exonuclease III, and these were sequenced by the dideoxy method. Figure 5A shows the sequence from this fragment of the Rm220-3 *mos* locus compared with those of the 5′ regions of the L5-30 *mos* locus and the *nifH* locus of *R. meliloti* Rm102F34. A total of 77.4% of the bases are identical among all three loci, and there is a 97.5% conservation of sequence between Rm220-3 and L5-30. Rm220-3 has the two consensus sequences required for

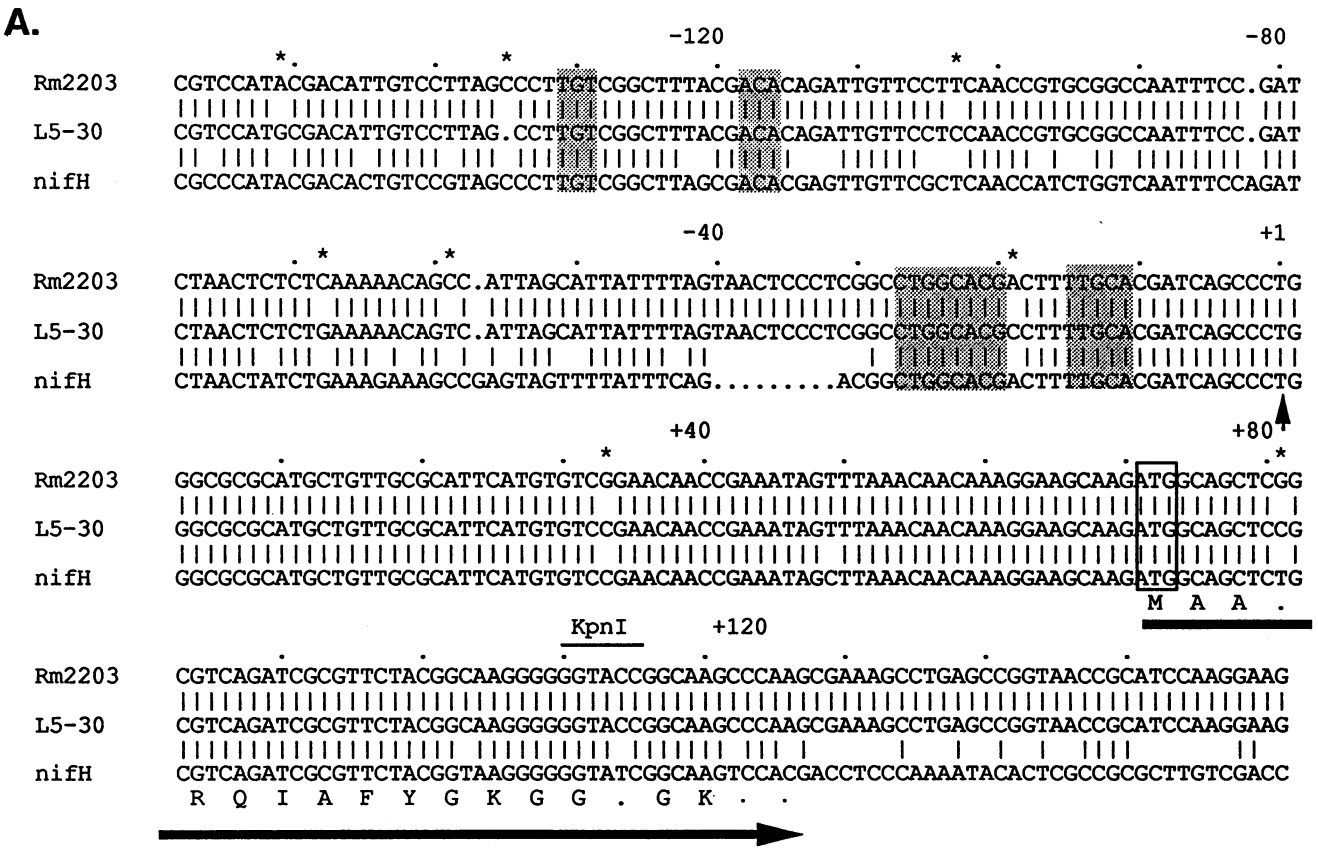
A.
-120
-80
Rm2203 CGTCCATACGACATTGTCCTTAGCCCTTGTCGGCTTTACGACACAGATTGTTCCTTCAACCGTGCGGCCAATTTCC.GAT
L5-30 CGTCCATGCGACATTGTCCTTAG.CCTTGTCGGCTTTACGACACAGATTGTTCCTCCAACCGTGCGGCCAATTTCC.GAT
nifH CGCCCATACGACACTGTCCGTAGCCCTTGTCGGCTTAGCGACACGAGTTGTTCGCTCAACCATCTGGTCAATTTCCAGAT
-40
+1
Rm2203 CTAACTCTCTCAAAAACAGCC.ATTAGCATTATTTTAGTAACTCCCTCGGCCTGGCACGACTTTTGCACGATCAGCCCTG
L5-30 CTAACTCTCTGAAAAACAGTC.ATTAGCATTATTTTAGTAACTCCCTCGGCCTGGCACGCCTTTTGCACGATCAGCCCTG
nifH CTAACTATCTGAAAGAAAGCCGAGTAGTTTTATTTCAG.....ACGGCTGGCACGACTTTTGCACGATCAGCCCTG
+40
+80
Rm2203 GGCGCGCATGCTGTTGCGCATTCATGTGTCGGAACAACCGAAATAGTTTAAACAACAAAGGAAGCAAGATGGCAGCTCGG
L5-30 GGCGCGCATGCTGTTGCGCATTCATGTGTCCGAACAACCGAAATAGTTTAAACAACAAAGGAAGCAAGATGGCAGCTCCG
nifH GGCGCGCATGCTGTTGCGCATTCATGTGTCCGAACAACCGAAATAGCTTAAACAACAAAGGAAGCAAGATGGCAGCTCTG
M A A .
KpnI
+120
Rm2203 CGTCAGATCGCGTTCTACGGCAAGGGGGGTACCGGCAAGCCCAAGCGAAAGCCTGAGCCGGTAACCGCATCCAAGGAAG
L5-30 CGTCAGATCGCGTTCTACGGCAAGGGGGGTACCGGCAAGCCCAAGCGAAAGCCTGAGCCGGTAACCGCATCCAAGGAAG
nifH CGTCAGATCGCGTTCTACGGTAAGGGGGGTATCGGCAAGTCCACGACCTCCCAAAATACACTCGCCGCGCTTGTCGACC
R Q I A F Y G K G G . G K . .
B.
100 bp
K
R K C B X P B S N R B H M P H A P R
ORF1 mosA mosB mosC
R. meliloti L5-30 (pPM1062)
pPM1201
pPM1202
pPM1203
pPM1204
H R K S N R B H M P H A P H
R. meliloti 220-3 (pPM1169)
ORF1 mosB mosC
1 kb

FIG. 5. (A) DNA sequence of a promoter and 5′ region of ORF1 and location of a deletion in the *mos* locus of Rm220-3. The sequence of the promoter and 5′ region of ORF1 from Rm220-3 is also presented and contrasted with the known sequences for a comparable region of *mos* from L5-30 and the promoter and 5′ region of the *R. meliloti nifH* gene. Differences between the Rm220-3 and L5-30 sequences are indicated (stars). The consensus sequences for NifA and NtrA regulation are shaded, and the ATG start sites are boxed. The 57-bp *nifH* homologous region present in L5-30 and Rm220-3 and the start of transcription for the *nifH* locus are indicated (horizontal and vertical arrows, respectively). (B) The region of the Rm220-3 and L5-30 DNA sequences in panel A is shown at the top. A comparison of the restriction map for the *mos* region of pPM1062 (from strain L5-30) and the corresponding region of pPM1169 (from strain Rm220-3) is also shown. The ORFs for L5-30 *mos* have been determined by sequencing (40), and the corresponding regions for Rm220-3 *mos* have been determined by hybridization and restriction studies. The relative positions of *mos* ORFs (arrows), the size and extent of deletion in pPM1169 (triangle), the relative positions of the probe fragments from pPM1201 to -1204 used to confirm the presence of ORF1, *mosB*, and *mosC* and absence of *mosA* from Rm220-3 (closed bars), and the region of Rm220-3 DNA sequenced (hatched box) are indicated. Abbreviations: B, *Bam*HI; H, *Hin*dIII; C, *Cla*I; K, *Kpn*I; R, *Eco*RI; P, *Pst*I; S, *Sac*I; M, *Mlu*I; N, *Nsu*I; A, *Apa*I; X, *Xho*I.

NifA/NtrA regulation (25) at approximately −120 bp (TGT-N_{10}-ACA) and at approximately −20 bp (CTGGCACG-N_4-TTGCA), respectively. Thus, Rm220-3 *mos* genes are directly controlled by a NifA-regulated promoter. The leader sequence of Rm220-3 is highly homologous to the analogous region from L5-30, differing by only 1 bp. Furthermore, the region of Rm220-3 open reading frame 1 (ORF1) which has been sequenced is almost identical to the L5-30 ORF1. Within the first 57 bp of ORF1, the region which is also homologous to the *nifH* gene coding region, there is a 1-bp difference between the *mos* regions of Rm220-3 and L5-30, and this is at a point where the L5-30 sequence diverges from the *nifH* sequence. After the first 57 bp, the Rm220-3 and L5-30 sequences remain identical but diverge from the *nifH* sequence.

Since the promoter and ORF1 of L5-30 show remarkable similarity to equivalent regions in Rm220-3, we investigated the downstream region of the Rm220-3 *mos* locus by restriction analysis and hybridization studies. Comparison of restriction fragments present in pPM1169, containing Rm220-3 *mos* genes, and pPM1062, which contains the L5-30 *mos* locus, show many similarities (Fig. 5B). These data are also consistent with there being a 1.1-kb deletion in Rm220-3. The L5-30 *mos* locus consists of four ORFs (termed ORF1 and *mosA*, *-B*, and *-C*) arranged in an operon structure (Fig. 5B) (40). To determine which region of the Rm220-3 *mos* locus was encompassed in the deletion, we hybridized fragments from the four separate L5-30 ORFs (pPM1201 to pPM1204 [Fig. 5B]) to the Rm220-3 *mos*-containing plasmid pPM1169. We confirmed the absence of homology to *mosA* and presence of homology to ORF1, *mosB*, and *mosC* in Rm220-3. DNA hybridization results leading to this conclusion are shown in Fig. 6. Here, a probe (pPM1201) prepared from the 3′ region of the L5-30 *mos* ORF1 was hybridized to the *mos* regions of Rm220-3 and L5-30. This probe was homologous to a 2.3-kb *Kpn*I-*Eco*RI fragment from L5-30 (Fig. 5B and 6, lane 2) as predicted. The same probe hybridized to a 1.2-kb *Kpn*I-*Eco*RI fragment from Rm220-3 (Fig. 6, lane 1). The difference between the sizes of these two fragments can be explained by a deletion of 1.1 kb in Rm220-3. Probing with a fragment (from pPM1202) covering an internal region of L5-30 *mosA* resulted in hybridization of a 0.8-kb *Bam*HI fragment from L5-30, but no equivalent band was present in Rm220-3 (Fig. 6, lanes 4 and 3, respectively). The deletion in Rm220-3 can therefore be explained by the absence of a region corresponding to *mosA* from strain L5-30. Downstream of this deletion, the restriction maps of Rm220-3 and L5-30 shown in Fig. 5B are identical. Confirmation that Rm220-3 has homology to L5-30 *mosB* and *mosC* was obtained, as common hybridizing 2.0-kb *Sac*I-*Pst*I bands (Fig. 6, lanes 5 and 6) and common 1.0-kb *Hin*dIII fragments (Fig. 6, lanes 7 and 8) were found when the DNAs from the *mos* regions were hybridized with probes to *mosB* (pPM1203) and *mosC* (pPM1204), respectively. The 3′ end of the *mosC* probe also hybridizes with a 7.3-kb *Hin*dIII fragment which is adjacent to the Rm220-3 *mos* region (Fig. 4 and 6, lane 7). In the case of the cloned *mos* region of L5-30, the same region is attached to the vector and yields a large (>20 kb) additional hybridizing fragment (Fig. 7, lane 8).

Analysis of *moc* genes from Rm220-3. The genes for the catabolism of the L5-30 rhizopine 3-*O*-MSI have been cloned into pLAFR1 to give pPM1031 (37). This plasmid was used to probe pPM1153, and four contiguous *Hin*dIII fragments of 7.1, 4.4, 2.8, and 1.2 kb hybridized (Fig. 4). Partial digestion of pPM1153 with *Hin*dIII followed by religation and transformation into *E. coli* HB101 produced clones pPM1171 and pPM1175. These clones were introduced into Rm1021, and the resulting strains were tested for catabolism of *s*Ia. Rm1021(pPM1171) does not catabolize *s*Ia, whereas Rm1021(pPM1175) does, indicating that the 7.1-kb fragment is required for catabolic activity. A similar experiment involving partial digestion with *Eco*RI yielded pPM1186, and catabolic studies with Rm1021(pPM1186) indicate that this strain also catabolizes *s*Ia.

To determine similarities between the *moc*-containing regions of L5-30 and Rm220-3, total DNA from these strains

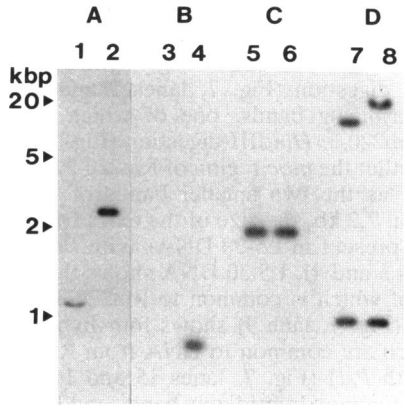

FIG. 6. Hybridization analysis of the Rm220-3 *mos* locus with probes from individual ORFs from the L5-30 *mos* locus. An autoradiogram of a Southern blot of restricted pPM1169 (from Rm220-3) and pPM1062 (from L5-30) probed against radiolabelled restriction fragments from pPM1201 (ORF1) (A), pPM1202 (*mosA*) (B), pPM1203 (*mosB*) (C), and pPM1204 (*mosC*) (D) is shown. Lanes: 1, pPM1169 (*Kpn*I-*Eco*RI); 2, pPM1062 (*Kpn*I-*Eco*RI); 3, pPM1169 (*Bam*HI-*Eco*RI); 4, pPM1062 (*Bam*HI-*Eco*RI); 5, pPM1169 (*Sac*I-*Pst*I); 6, pPM1062 (*Sac*I-*Pst*I); 7, pPM1169 (*Hin*dIII); 8, pPM1062 (*Hin*dIII).

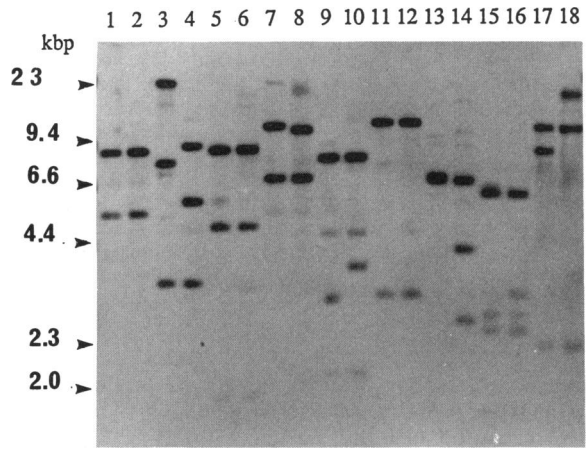

FIG. 7. Restriction fragment length polymorphism analysis of the rhizopine-catabolic (*moc*) regions from Rm220-3 and L5-30. An autoradiogram of the Southern blot of restricted total DNA from Rm220-3 and L5-30 probed against radiolabelled pPM1031 bearing the *moc* region from L5-30 is shown. Lanes: 1, L5-30 (*Eco*RI); 2, Rm220-3 (*Eco*RI); 3, L5-30 (*Bam*HI); 4, Rm220-3 (*Bam*HI); 5, L5-30 (*Sma*I); 6, Rm220-3 (*Sma*I); 7, L5-30 (*Xho*I); 8, Rm220-3 (*Xho*I); 9, L5-30 (*Cla*I); 10, Rm220-3 (*Cla*I); 11, L5-30 (*Sal*I); 12, Rm220-3 (*Sal*I); 13, L5-30 (*Hin*dIII); 14, Rm220-3 (*Hin*dIII); 15, L5-30 (*Pst*I); 16, Rm220-3 (*Pst*I); 17, L5-30 (*Nde*I); 18, Rm220-3 (*Nde*I).

was extracted and digested with a variety of restriction endonucleases and the fragments were separated by agarose gel electrophoresis, transferred to a membrane, and probed with radiolabelled pPM1031, which contains the L5-30 *moc* locus. The resulting autoradiogram is shown in Fig. 7. Lane 1 shows L5-30 total DNA restricted with *Eco*RI and the characteristic 8.7- and 5.4-kb fragments known to make up the *moc* insert. Equivalent fragments are also present in lane 2, which contains Rm220-3 total DNA also digested with *Eco*RI. A 1.0-kb fragment characteristic of the L5-30 *moc* region was also found in both L5-30 and Rm220-3, but these cannot be seen in Fig. 7, as they are below the 1.5-kb cutoff point used to calculate the DNA sequence divergence. In the cases of *Sma*I (Fig. 7, lanes 5 and 6) and *Sal*I (Fig. 7, lanes 11 and 12), all digests contain strongly hybridizing equivalent bands. *Xho*I digestion (Fig. 7, lanes 7 and 8) shows two strongly hybridizing bands, one of which is common to L5-30 and Rm220-3. *Hin*dIII digestion (Fig. 7, lanes 13 and 14) indicates that the *moc* region of Rm220-3 bears one extra *Hin*dIII site, as the two smaller bands of 4.4 and 2.8 kb together equal 7.2 kb, the size of the extra fragment running as a doublet present in L5-30 DNA. With *Bam*HI digestion (Fig. 7, lanes 3 and 4), L5-30 DNA shows three hybridizing bands, one of which is common to Rm220-3. *Cla*I-digested L5-30 DNA (Fig. 7, lane 9) shows four hybridizing bands, three of which are common to DNA from Rm220-3 (Fig. 7, lane 10). With *Pst*I (Fig. 7, lanes 15 and 16), L5-30 DNA shows three common hybridizing bands and Rm220-3 has an additional hybridizing band. *Nde*I-digested L5-30 and Rm220-3 DNAs (Fig. 7, lanes 17 and 18) contain two common hybridizing bands, and each has a unique hybridizing band. The percent DNA sequence divergence, estimated by the method described by Nei and Li (42, 43), is 1.5%. This figure is based on restriction fragments between 1.5 and 25 kb in size. These experiments delineated the *moc* locus to a region of approximately 15 kb, which is of a size similar to that of the 15.1-kb region required for 3-*O*-M*S*I catabolism in L5-30. We conclude that the *moc* loci from strains L5-30 and Rm220-3 are very similar by restriction fragment length polymorphism analysis.

We have undertaken a preliminary study of the regulation of *moc* genes. pPM1153 containing the complete suite of *moc* genes was introduced into the NifA$^-$, NtrC$^-$, NifH$^-$, and NtrA$^-$ Rm1021 strains, and catabolism studies were performed. However, *moc* genes were found to be fully active in all these mutants (Table 2).

DISCUSSION

We have isolated and studied genes for the synthesis and catabolism of the rhizopine *s*Ia induced in alfalfa nodules by *R. meliloti* Rm220-3. This rhizopine is structurally closely related to the rhizopine, 3-*O*-M*S*I, induced by *R. meliloti* L5-30, with both compounds classed as substituted inositols.

The isolation of these rhizopine genes was aided by their strong hybridization to the equivalent genes of L5-30. Rm220-3 rhizopine genes were originally cloned from total DNA as a 27-kb fragment in pPM1153. Subsequent subcloning, hybridization, and phenotypic studies have delineated the *mos* locus to a 12-kb region and *moc* genes to a contiguous 15-kb region within the 27-kb fragment. Hybridization with the L5-30 *mos* gene probes further localized the *mos* genes of Rm220-3 to a 4-kb region (Fig. 5B and 6). Therefore, the *mos* and *moc* loci are in close juxtaposition, being separated by approximately 7 kb. In addition, it was demonstrated that rhizopine genes from Rm220-3 are located on the *nod-nif* Sym plasmid. This was shown by plasmid mobilization using Tn*5*-Mob in conjunction with hybridization studies with *nod* and *mos* genes and expression of the Moc phenotype. The close linkage of the synthesis and catabolism genes and their Sym plasmid location parallel those of the rhizopine genes from L5-30.

We have demonstrated that *mos* genes of Rm220-3 are NifA/NtrA regulated. When a plasmid (pPM1169) containing a complete suite of genes required for *s*Ia synthesis was introduced into NifA$^-$ or NtrA$^-$ *R. meliloti* strains, this plasmid did not bestow the ability to produce *s*Ia in nodules. The DNA sequence of the *mos* promoter and part of L5-30 ORF1 (38) shows remarkable homology to the NifA-regulated promoter of the *nifHDK* operon, which encodes the nitrogenase complex (15, 47), and the first 57 bp of the *nifH* gene (7). Sequencing a comparable region of Rm220-3 has revealed very extensive homology (97.5%) between the equivalent regions of these two *mos* loci (Fig. 5A). The two consensus sequences for regulation by NifA and NtrA, which act in concert to control many symbiotic genes (25), are conserved in the *mos* 5′ region of Rm220-3. This, together with the results of the regulatory studies, indicates that, as with L5-30, the *mos* genes are directly regulated by a symbiotic promoter.

Recently (40), we have shown that ORF1 from L5-30 does not produce a protein in nodules, and a frameshift mutation indicates that it is not required for rhizopine production. These results are consistent with the L5-30 *mos* locus (and also, because of its similar structure, presumably the Rm220-3 *mos* locus) having acquired a duplicated copy of *nifH* and its regulatory region, resulting in symbiotic regulation of this replicon. Reasoning along the lines that the *mos* locus evolved by insertion into a duplicated copy of a *nifH* gene, we have accordingly tried to detect further remnants of such a gene. However, no further *nifH* homology with Rm220-3 and L5-30 clones bearing regions up to 40 kb downstream of ORF1 could be detected (data not shown).

This suggests that if *mos*, or indeed *mos* and *moc* genes, were inserted into a complete *nifH* gene subsequent evolution removed the 3′ region of this gene. A more likely explanation is that this locus results from rearranged fragments of symbiotic genes. The accompanying article (40) describing the mosaic structure of the L5-30 *mos* locus, a region of rearranged symbiotic genes, bears this out.

The similarity of the L5-30 and Rm220-3 rhizopines and the similarity between the respective *mos* loci suggest that biosynthetic steps for the synthesis of *s*Ia and 3-*O*-M*S*I are similar. The *mos* locus of L5-30 has recently (40) been shown to consist of four ORFs (ORF1 and *mosA*, *-B*, and *-C*) arranged in an operon structure. Probes were prepared from each ORF and hybridized to pPM1169 containing the Rm220-3 *mos* region. Homology to ORF1, *mosB*, and *mosC*, but not to *mosA*, was detected. This, together with restriction mapping and 5′ sequence data, indicates that the two *mos* loci are very similar but differ by a 1.1-kb fragment encompassing *mosA* which is absent from Rm220-3. Since the structures of *s*Ia and 3-*O*-M*S*I are very similar, differing only in a methyl group, it is likely that, in L5-30, 3-*O*-M*S*I synthesis occurs via *s*Ia and that L5-30 *mosA* is involved in this methylation step, with the preceding steps being common to both strains.

Rm220-3 can catabolize both rhizopines, 3-*O*-M*S*I and *s*Ia. The catabolism genes of Rm220-3 have been located on a 15-kb fragment whose size is similar to the 15.1 kb required for *moc* activity in L5-30. To determine the degree of similarity between these two different loci, total DNA from both strains was digested with a number of different restriction enzymes and probed with plasmid pPM1031, which contains the L5-30 *moc* genes. Many common hybridizing bands were identified (Fig. 7), and the percent DNA sequence divergence was estimated to be 1.5%. This compares favorably with the results of a larger study of restriction fragment length polymorphisms using DNA from 85 *R. leguminosarum* isolates probed with DNA from the *lac* operon, which found that the average DNA sequence divergence for this region ranged between 1.4 and 15.8% with an average of 5.7% (64). When DNA from the Sym region was used to probe DNA from the same 85 isolates, an even greater degree of polymorphism was seen, with the average DNA sequence divergence for this region being 11.4%. Hence, a value of 1.2% DNA sequence divergence for the *moc* loci from L5-30 and Rm220-3 suggests that these sequences are very similar.

We investigated the regulation of the *moc* genes from Rm220-3 and found that neither NifA, NtrC, nor NtrA (RpoN GlnF) played a part in their regulation (Table 2). The *nifA* regulon controls symbiotic functions, not processes in free-living bacteria (25); therefore, its involvement was not expected. Regulation by NtrC or NtrA was thought more likely, since NtrC is known to be involved in nitrogen metabolism by free-living bacteria not only in rhizobia but in several species of *Enterobacteriaceae* and NtrA is involved in the regulation of a wide variety of catabolic processes in several genera of gram-negative bacteria (25, 60). Similarly, Boivin and coworkers (9) have recently reported that *R. meliloti* genes for trigonelline catabolism are not controlled by any of the common general or symbiotic regulatory genes. Genes for trigonelline catabolism in *R. meliloti* are induced at all stages of the rhizobium-legume association (10). Trigonelline is an abundant legume secondary metabolite unlike *s*Ia, which is produced during symbiosis in limiting amounts by *mos*-endowed bacteroids. Therefore, it would seem prudent for rhizopine-catabolic genes to be nonactive in bacteroids, so as to allow significant feeding of free-living cells.

So far, only three rhizobial isolates we have examined can synthesize and catabolize rhizopines. These three strains all produce inositol-based rhizopines; however, there may be other compounds fulfilling the role of rhizopines which have so far gone undetected, as our definition of rhizopines is functional rather than chemical (39). There may be a wide range of different rhizopines, each with a limited number of strains capable of catabolizing it. Nevertheless, it may well be that not all strains produce rhizopines, and it may be a particular refinement which some rhizobia have developed to survive in the rhizosphere. This is different from what is found with opines, as all agrobacteria are known to induce their production and they fall into a few catabolic classes which can be utilized by a range of *Agrobacterium* strains as well as by other bacteria (2, 41, 46, 61).

Studies on the inositol class of rhizopines from *R. meliloti* Rm220-3 and L5-30 have indicated common function at both the physiological and the genetic levels. Physiologically, a specific rhizobium is able to induce the synthesis of, and catabolize, a selective growth substrate. Genetically, synthesis and catabolism of the substrate are controlled by closely linked genes on the Sym plasmid, and synthesis is NifA regulated. Whether rhizopines analyzed subsequently can be likewise defined awaits further research.

ACKNOWLEDGMENTS

We thank J. Dénarié, A. Pühler, M. John, J. Schmidt, C. Ronson, and W. Grzemski for strains and plasmids. A. Dunbar and J. Groom are thanked for photographic work, and E. Cabot is thanked for help with typing the manuscript.

This work was supported by the Australian Wool Corporation grants UAD16 and UAD57 and by the Australian Research Council.

REFERENCES

1. **Anderson, L., and H. A. Lardy.** 1950. Stereochemical studies in the aminodesoxyinositol series *meso*-inosamine-2 and *scyllo*-inosamine. J. Am. Chem. Soc. **72**:3141–3147.
2. **Beauchamp, C. J., J. W. Kloepper, R. Lifshitz, P. Dion, and H. Antoun.** 1991. Frequent occurrence of the ability to utilize octopine in rhizobacteria. Can. J. Microbiol. **37**:158–164.
3. **Bergersen, F. J.** 1961. The growth of *Rhizobium* in synthetic media. Aust. J. Biol. Sci. **14**:349–360.
4. **Beringer, J. E.** 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. **84**:188–198.
5. **Beringer, J. E., N. Brewin, A. W. B. Johnston, H. M. Schulman, and D. A. Hopwood.** 1979. The *Rhizobium*—legume symbiosis. Proc. R. Soc. Lond. Ser. B. **204**:219–233.
6. **Bernard, T., J. A. Pocard, B. Perroud, and D. Le Rudulier.** 1986. Variation in the response of salt-stressed *Rhizobium* strains to betaines. Arch. Microbiol. **143**:359–364.
7. **Better, M., B. Lewis, D. Corbin, G. Ditta, and D. R. Helinski.** 1983. Structural relationships among *Rhizobium meliloti* symbiotic promoters. Cell **35**:479–485.
8. **Boivin, C., L. R. Barran, C. A. Malpica, and C. Rosenberg.** 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.
9. **Boivin, C., S. Camut, C. A. Malpica, G. Truchet, and C. Rosenberg.** 1990. *Rhizobium meliloti* genes encoding catabolism of trigonelline are induced under symbiotic conditions. Plant Cell **2**:1157–1170.
10. **Boivin, C., C. Malpica, C. Rosenberg, J. Dénarié, A. Goldmann, V. Fleury, M. Maille, B. Message, and D. Tepfer.** 1989. Metabolic signals in the rhizosphere: catabolism of calystegins and trigonelline by *Rhizobium meliloti*, p. 401–407. *In* B. Lugtenberg

(ed.), Molecular signals in the microbe-plant symbiotic and pathogenic systems. Springer-Verlag KG, Berlin.
11. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
12. **Brewin, N. J., J. E. Beringer, and A. W. B. Johnston.** 1980. Plasmid-mediated transfer of host-range specificity between two strains of *Rhizobium leguminosarum*. J. Gen. Microbiol. **120**: 413–420.
13. **Clare, B. G., A. Kerr, and D. A. Jones.** 1990. Characteristics of the nopaline catabolic plasmid in *Agrobacterium* strains K84 and K1026 used for biological control of crown gall disease. Plasmid **23**:126–137.
14. **Cohen, S. N., A. C. Y. Chang, and C. L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
15. **Corbin, D., L. Barran, and G. Ditta.** 1983. Organization and expression of *Rhizobium meliloti* nitrogen fixation genes. Proc. Natl. Acad. Sci. USA **80**:3005–3009.
16. **Dahl, G. A., P. Guyon, A. Petit, and J. Tempé.** 1983. Silver nitrate-positive opines in crown gall tumors. Plant Sci. Lett. **32**:193–203.
17. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
18. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range DNA cloning system for gram negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77**:7347–7351.
19. **Dowling, D. N., and W. J. Broughton.** 1986. Competition for nodulation of legumes. Annu. Rev. Microbiol. **40**:131–157.
20. **Eckhardt, T.** 1978. A rapid method for the identification of deoxyribonucleic acid in bacteria. Plasmid **1**:584–588.
21. **Economou, A., F. K. L. Hawkins, and A. W. B. Johnston.** 1988. pRLIJI specifies the catabolism of L-homoserine and contains a gene, *rhi*, whose transcription is reduced in the presence of *nod* gene inducer molecules, p. 462. *In* H. Bothe, F. J. de Bruijn, and W. E. Newton (ed.), Nitrogen fixation: hundred years after. Gustav Fischer, Stuttgart, Germany.
22. **Figurski, D. H., and D. R. Helinski.** 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.
23. **Frahn, J. L., and J. A. Mills.** 1974. Formation of borate complexes of inosamines (aminodeoxyinositols) and their separation by paper electrophoresis. Aust. J. Chem. **27**:853–864.
24. **Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel.** 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. Gene **18**:289–296.
25. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**: 567–591.
26. **Hartwig, U. A., C. M. Joseph, and D. A. Phillips.** 1991. Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*. Plant Physiol. **95**:797–803.
27. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
28. **Heusterspreute, M., V. H. Thi, S. Emery, S. Tournis-Gamble, N. Kennedy, and J. Davison.** 1985. Vectors with restriction site banks. IV. pJRD184, a 3793-bp plasmid vector having 43 unique cloning sites. Gene **39**:299–304.
29. **Knauf, V. C., and E. W. Nester.** 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. Plasmid **8**:45–54.
30. **Kondorosi, E., Z. Banfalvi, and A. Kondorosi.** 1984. Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: identification of nodulation genes. Mol. Gen. Genet. **193**:445–452.
31. **Kondorosi, A., G. B. Kiss, T. Forrai, E. Vincze, and Z. Banfalvi.** 1977. Circular linkage map of *Rhizobium meliloti* chromosome. Nature (London) **268**:525–527.
32. **Kondorosi, A., E. Kondorosi, C. [illegible]churst, W. J. Broughton, and Z. Banfalvi.** 1982. Mobilization of a *Rhizobium meliloti* megaplasmid carrying nodulation and nitrogen fixation into other rhizobia and *Agrobacterium*. Mol. Gen. Genet. **188**:433–439.
33. **Kowalski, M.** 1967. Transduction in *Rhizobium meliloti*. Acta Microbiol. Pol. **16**:7–12.
34. **Long, S. R.** 1989. *Rhizobium* genetics. Annu. Rev. Genet. **23**:483–506.
35. **Meade, H. M., S. R. Long, G. B. Ruvkin, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. **149**:114–122.
36. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Murphy, P. J., N. Heycke, Z. Banfalvi, M. E. Tate, F. J. de Bruijn, A. Kondorosi, J. Tempé, and J. Schell.** 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.
38. **Murphy, P. J., N. Heycke, S. P. Trenz, P. Ratet, F. J. de Bruijn, and J. Schell.** 1988. Synthesis of an opine-like compound—a rhizopine—in alfalfa nodules is symbiotically regulated. Proc. Natl. Acad. Sci. USA **85**:9133–9137.
39. **Murphy, P. J., and C. P. Saint.** 1991. Rhizopines in the legume-*Rhizobium* symbiosis, p. 378–390. *In* D. P. S. Verma (ed.), Molecular signals in plant-microbe communication. CRC Press, Inc., Boca Raton, Fla.
40. **Murphy, P. J., S. P. Trenz, W. Grzemski, F. J. de Bruijn, and J. Schell.** 1993. The *Rhizobium meliloti* rhizopine *mos* locus is a mosaic structure facilitating its symbiotic regulation. J. Bacteriol. **175**:5193–5204.
41. **Nautiyal, C. S., P. Dion, and W. S. Chilton.** 1991. Mannopine and mannopinic acid as substrates for *Arthrobacter* sp. strain MBA209 and *Pseudomonas putida* NA513. J. Bacteriol. **173**: 2833–2841.
42. **Nei, M.** 1987. Molecular evolutionary genetics. Columbia University Press, New York.
43. **Nei, M., and W.-H. Li.** 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76**:5269–5273.
44. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxyribonucleotide-directed mutagenesis. Gene **26**:101–106.
45. **Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel.** 1987. *R. meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J. Bacteriol. **169**:2424–2431.
46. **Rossignol, G., and P. Dion.** 1987. Octopine, nopaline and octopinic acid utilization in *Pseudomonas*. Can. J. Microbiol. **31**:68–74.
47. **Ruvkun, G. B., V. Sundaresan, and F. M. Ausubel.** 1982. Directed transposon Tn*5* mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. Cell **29**:551–559.
48. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
49. **Sanger, F., S. Nicklen, and R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
50. **Schmidt, J., M. John, E. Kondorosi, A. Kondorosi, U. Wieneke, J. Schröder, and J. Schell.** 1984. Mapping of the protein-coding regions of *Rhizobium meliloti* common nodulation genes. EMBO J. **3**:1705–1711.
51. **Scott, D. B., R. Wilson, G. J. Shaw, A. Petit, and J. Tempe.** 1987. Biosynthesis and degradation of nodule-specific *Rhizobium loti* compounds in *Lotus* nodules. J. Bacteriol. **169**:278–282.
52. **Simon, R.** 1984. High frequency mobilization of gram-negative bacterial replicons by the *in vitro* constructed Tn*5*-Mob transposon. Mol. Gen. Genet. **196**:413–420.
53. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range

mobilization system for *in vivo* genetic engineering. Transposon mutagenesis in Gram negative bacteria. Bio/Technology **1:**784–791.

54. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.
55. **Szeto, W. W., B. T. Nixon, C. W. Ronson, and F. M. Ausubel.** 1987. Identification and characterization of the *Rhizobium meliloti ntrC* gene: *R. meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. J. Bacteriol. **169:**1423–1432.
56. **Szeto, W. W., J. L. Zimmerman, V. Sundaresan, and F. M. Ausubel.** 1984. A *Rhizobium meliloti* symbiotic regulatory gene. Cell **36:**1035–1043.
57. **Tempé, J., and A. Petit.** 1983. La piste des opines, p. 14–32. *In* A. Pühler (ed.), Molecular genetics of the bacteria plant interaction. Springer-Verlag KG, Berlin.
58. **Tepfer, D., A. Goldmann, V. Fleury, M. Maille, B. Message, N. Pamboukdjian, C. Bovin, J. Dénarié, C. Rosenberg, J. Y. Lallemand, C. Descoins, I. Charpin, and N. Amarger.** 1988. Calystegins, nutritional mediators in plant-microbe interactions, p. 139–144. *In* R. Palacios and D. P. S. Verma (ed.), Molecular genetics of plant-microbe interactions. American Phytopathological Society Press, St. Paul, Minn.
59. **Tepfer, D., A. Goldmann, N. Pamboukdjian, M. Maille, A. Lepingle, D. Chevalier, J. Dénarié, and C. Rosenberg.** 1988. A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudates of *Calystegia sepium*. J. Bacteriol. **170:**153–161.
60. **Thöny, B., and H. Hennecke.** 1989. The −24/−12 promoter comes of age. FEMS Microbiol. Rev. **63:**341–358.
61. **Tremblay, G., R. Gagliardo, W. S. Chilton, and P. Dion.** 1987. Diversity among opine utilizing bacteria: identification of coryneform isolates. Appl. Environ. Microbiol. **53:**1519–1524.
62. **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.
63. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
64. **Young, J. P. W., and M. Wexler.** 1988. Sym plasmid and chromosomal genotypes are correlated in field populations of *Rhizobium leguminosarum*. J. Gen. Microbiol. **134:**2731–2739.