**Rhizobium leguminosarum** contains a group of genes that appear to code for methyl-accepting chemotaxis proteins

Christopher K. Yost, Patrice Rochepeau and Michael F. Hynes

**INTRODUCTION**

The study of bacterial chemotaxis has provided substantial information about the molecular mechanisms by which bacteria process and react to environmental information. Most of the knowledge of chemotaxis comes from research using enteric organisms such as *Escherichia coli* and *Salmonella typhimurium*. In these bacteria, many of the chemotaxis genes are contained within a single operon (*cheA, cheY, cheW, cheR, cheB*) (for reviews see Stock et al., 1991; Manson, 1992). Proteins CheA and CheY belong to a two-component regulatory system that modulates direction of flagellar rotation and thus determines the behaviour of the bacterium (reviewed by Matsumura et al., 1990; Parkinson, 1993). Methyl-accepting chemotaxis proteins (MCPs) may also play a role in this signalling pathway by serving as environmental sensors. They are transmembrane proteins with the carboxyl-terminus located in the cytoplasm and the amino-terminus exposed to the external environment (Boyd et al., 1983; Hazelbauer et al., 1990). The amino-terminal domain detects attractants and repellents, and the carboxyl-terminal domain relays this information to CheA, allowing the

**Methyl-accepting chemotaxis proteins (MCPs)** play important roles in the chemotactic response of many bacteria. Oligonucleotide primers designed to amplify the conserved signalling domain of MCPs by PCR were used to identify potential MCP-encoding genes in *Rhizobium leguminosarum*. Using a PCR-derived probe created from these primers a genomic library of *R. leguminosarum* VF39SM was screened; at least five putative MCP-encoding genes (termed *mcpB* to *mcpF*) were identified and isolated from the library. One of these putative genes (*mcpC*) is located on one of the indigenous plasmids of VF39SM. Fifteen different cosmids showing homology to an *mcpD* probe were also isolated from a genomic library. The complete DNA sequences of *mcpB*, *mcpC* and *mcpD* were obtained. All three genes code for proteins with characteristics typical of MCPs. However, the protein encoded by *mcpB* has a relatively large periplasmic domain compared to that in other MCPs. Partial DNA sequences of *mcpE* and *mcpF* had strong similarity to sequences from the methylation domains of known MCPs. Mutants defective in *mcpB*, *mcpC* or *mcpD* or *mcpE* were created using insertional mutagenesis strategies. Mutation of *mcpB* resulted in impairment of chemotaxis to a wide range of carbon sources on swarm plates; phenotypes for the other three mutants have yet to be elucidated. The *mcpB*, *mcpC* and *mcpD* mutants were tested for loss of nodulation competitiveness. When co-inoculated with the wild-type, the *mcpB* and *mcpC* mutants formed fewer nodules than the wild-type, whereas the *mcpD* mutant was just as competitive as the wild-type. The results overall suggest that *R. leguminosarum* possesses *mcp*-like genes, and that at least some of these play a role in early steps in the plant–microbe interaction.

**Keywords**: rhizobia, chemotaxis, sensory proteins

---

**Abbreviations**: DIG, digoxigenin; MCP, methyl-accepting chemotaxis protein.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are: *mcpB*, AF022807; *mcpC*, AF036168; *mcpD*, U81828.
bacterium to swim towards attractants and away from repellents (Hazelbauer et al., 1990). Sensory adaptation occurs through methylation and demethylation of glutamate residues contained in the carboxyl-terminal of the protein (Hazelbauer, 1988). Generally, the binding of attractants leads to increased methylation of the MCP. Conservation of MCP amino acid sequences among bacteria occurs predominately in the carboxyl-terminal region, where the MCP interacts with the signalling components CheW and CheA and where methylation occurs. A surge of interest in the chemotactic behaviour of other bacterial species has demonstrated that the chemotactic signalling pathway is highly conserved. Elements of the signalling pathway including cheA, cheY, cheW, cheR, cheB and MCP homologues have been found in a wide range of bacteria, such as Bacillus subtilis (Hanlon & Ordal, 1994; Zuberi et al., 1990), Caulobacter crescentus (Alley et al., 1992), Myxococcus xanthus (McCleary et al., 1990; McCleary & Zusman, 1990), and the archaeon Halobacterium halobium (Alam et al., 1989).

It has frequently been suggested that chemotaxis and motility must play an important role in the interaction of root-nodule-forming bacteria (collectively known as rhizobia) with their legume hosts (Currier & Strobel, 1977; Ames & Bergman, 1981; Gulash et al., 1984; Caetano-Anollés et al., 1988, 1992; Munoz Aguilar et al., 1988; Bauer & Caetano-Anollés, 1990; Dharmatilake & Bauer, 1992; reviewed by Bauer, 1991, and Vande Broek & Vanderleyden, 1995). To confirm this hypothesis more information is needed about the chemotactic signalling pathways in rhizobia. Quite recently, Greck et al. (1995) have shown that Sinorhizobium meliloti contains a chemotaxis operon homologous to the E. coli chemotaxis operon. Genes identified using sequence homology to E. coli genes include cheA, cheR, cheB, two cheY genes and two ORFs. The products of the two ORFs show amino acid similarity to the signalling domain of MCPs, but their predicted amino acid sequence lacks the other functional domains, such as transmembrane regions and the two methylation domains, characteristic of MCPs. The roles of the two ORFs in chemotaxis remain unclear. Previously, conflicting reports of in vivo methylation studies questioned the existence of MCP-encoding genes in Sinorhizobium and Rhizobium. Robinson & Bauer (1993) observed no increase or decrease in protein methylation after Sinorhizobium meliloti cells were incubated with attractants such as L-amino acids or D-mannitol. However, Armitage et al. (1988) demonstrated an increase in protein methylation after Rhizobium leguminosarum cells were incubated with the chemoattractant L-serine. The latter experiment was in agreement with a study conducted by Morgan et al. (1993). Probing an S. meliloti Western blot with an antibody directed against an E. coli MCP (Trg) gave three hybridizing bands: a doublet between 65 and 70 kDa, and a band at 80 kDa. This provided further evidence that MCPs might exist in the rhizobia. When this research was initiated, the existence of MCP-encoding genes in Rhizobium had not been clearly established. More recently, genes with homology to MCPs have been found on plasmids in a strain of R. leguminosarum (Brito et al., 1996) and Rhizobium sp. NGR234 (Freiberg et al., 1997); however, no role in chemotaxis has been shown. This paper documents research conducted to identify and isolate MCP-encoding genes from R. leguminosarum bv. viciae strain VF39SM.

**METHODS**

**Bacterial strains, plasmids, and media.** The bacterial strains and plasmids used are listed in Table 1. R. leguminosarum strains were grown on TY medium (Beringer, 1974) at 30 °C. E. coli strains were grown on LB medium (Sambrook et al., 1989) at 37 °C. Swarm medium was composed of 0:01% yeast extract, 1 mM MgSO4 and 0:3% (w/v) agar. The chemotactic response of R. leguminosarum to specific carbon sources was assayed using swarm plates containing Vincent’s minimal medium (VMM; Vincent, 1970) with 0:15% agarose and the potential chemoattractant as the sole carbon source (1 mM final concentration). Carbon sources were purchased from Sigma-Aldrich. When necessary, Rhizobium strains were cultured in media containing antibiotics at the following concentrations: neomycin, 100 pg ml⁻¹; spectinomycin, 500 pg ml⁻¹; tetracycline, 5 µg ml⁻¹; streptomycin, 500 µg ml⁻¹; and gentamicin, 30 µg ml⁻¹. Antibiotic concentrations used when cultivating E. coli strains were as follows: kanamycin, 50 µg ml⁻¹; tetracycline, 10 µg ml⁻¹; streptomycin, 500 µg ml⁻¹; gentamicin, 15 µg ml⁻¹; and spectinomycin 100 µg ml⁻¹.

**Recombinant DNA techniques.** All restriction endonucleases and modifying enzymes were purchased from Gibco-BRL, and used according to the manufacturer’s specifications. Southern blots were hybridized with non-radioactive DIG-labelled DNA probes using reagents and protocols obtained from Boehringer Mannheim. Hybridizations were performed overnight at 67 °C followed by washes of 2 x SSC, 0:1% SDS at room temperature, and 0:1 x SSC, 0:1% SDS at 67 °C. All washes were done in duplicate. Hybridization signals were detected via chemiluminescence using CSPD substrate and a protocol supplied by Boehringer Mannheim.

**Visualization of Rhizobium plasmids.** When necessary the plasmids of VF39SM were visualized on agarose gels using a modified Eckhardt technique (Eckhardt, 1978) described by Hynes et al. (1985), as modified by Hynes & McGregor (1990).

**DNA sequencing and analysis.** DNA sequencing was performed using an Applied Biosystems automated sequencer operated by University Core DNA Services (University of Calgary). ExoIII deletions were used to obtain the complete DNA sequences of mcpB, mcpC and mcpD. Enzymes and protocols for the ExoIII deletions were obtained from Pharmacia. Partial DNA sequences of mcpE and mcpF were obtained from subclones in pBSIISK(+) using T7 and T3 primers. Sequence homology searches were performed using the BLAST sequence alignment program (Altschul et al., 1990).

**Identification of MCP-encoding genes in R. leguminosarum VF39SM.** Primers designed from the DNA sequence of dcrA, an MCP-encoding gene from Desulfovibrio vulgaris (Dolla et al., 1992), were used to amplify VF39SM genomic DNA by
### Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristic(s)*</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5a</td>
<td><em>endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1 Δ(argFlacZYA) U169 + 80lacZ ΔM15</em></td>
<td>Gibco-BRL</td>
</tr>
<tr>
<td>S17-1</td>
<td>Sp'; RP4 tra region, mobilizer strain</td>
<td>Simon <em>et al.</em> (1983)</td>
</tr>
<tr>
<td><strong>R. leguminosarum</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VF395M</td>
<td><em>biovar vicieae, St'</em></td>
<td>Priefer (1989)</td>
</tr>
<tr>
<td>LRS93901</td>
<td>VF395M cured of pRleVF39b</td>
<td>Hynes &amp; McGregor (1990)</td>
</tr>
<tr>
<td>LRS93901</td>
<td>VF395M cured of pRleVF39c</td>
<td>Hynes &amp; McGregor (1990)</td>
</tr>
<tr>
<td>LRS93901</td>
<td>VF395M cured of pRleVF39d</td>
<td>Hynes &amp; McGregor (1990)</td>
</tr>
<tr>
<td>LRS93901</td>
<td>VF395M cured of pRleVF39e</td>
<td>Hynes &amp; McGregor (1990)</td>
</tr>
<tr>
<td>LRS93901</td>
<td>VF395M cured of pRleVF39f</td>
<td>Hynes &amp; McGregor (1990)</td>
</tr>
<tr>
<td>VF-MCPB(\textsuperscript{-})</td>
<td>VF395M, <em>mcpB</em>: :ΩTc; Tc(\textsuperscript{c})</td>
<td>This work</td>
</tr>
<tr>
<td>VF-MCPD(\textsuperscript{-})</td>
<td>VF395M, <em>mcpC</em>: :ΩSp; Sp(\textsuperscript{c})</td>
<td>This work</td>
</tr>
<tr>
<td>VF-MCPD(\textsuperscript{-})</td>
<td>VF395M, <em>mcpD</em>: :ΩNm; Nm(\textsuperscript{c})</td>
<td>This work</td>
</tr>
<tr>
<td>VF-MCPD(\textsuperscript{-})</td>
<td>VF395M, <em>mcpE</em>: :ΩSp; Sp(\textsuperscript{c})</td>
<td>This work</td>
</tr>
<tr>
<td>VF-MCPD(\textsuperscript{-}/D(\textsuperscript{-})</td>
<td>VF395M, <em>mcpC</em>: :ΩSp, <em>mcpD</em>: :ΩNm; Sp(\textsuperscript{c})</td>
<td>This work</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pJQ200mp18</td>
<td>Suicide vector with sacB system; Gm(\textsuperscript{c})</td>
<td>Quandt &amp; Hynes (1993)</td>
</tr>
<tr>
<td>pJQ200SK</td>
<td>Suicide vector with sacB system; Gm(\textsuperscript{c})</td>
<td>Quandt &amp; Hynes (1993)</td>
</tr>
<tr>
<td>pRK7813</td>
<td>Broad-host-range cloning vector; Tc(\textsuperscript{c})</td>
<td>Jones &amp; Gutterson (1987)</td>
</tr>
<tr>
<td>pBSISK(\textsuperscript{(+)})</td>
<td>Cloning vector; Ap(\textsuperscript{c})</td>
<td>Stratagene</td>
</tr>
<tr>
<td>MCPB.B</td>
<td>8 kb VF395M <em>BamHI</em> fragment cloned in pBSISK(\textsuperscript{(+)}); contains a fragment of <em>mcpB</em> gene</td>
<td>This work</td>
</tr>
<tr>
<td>VGL-747</td>
<td>VF395M cosmid clone in pRK7813; cosmid contains the entire <em>mcpB</em> gene</td>
<td>This work</td>
</tr>
<tr>
<td>MCPB.E</td>
<td>1.8 kb VF395M <em>EcoRI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPC.B</td>
<td>6.7 kb VF395M <em>BamHI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPC.E</td>
<td>3 kb VF395M <em>EcoRI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPCD.B</td>
<td>7 kb VF395M <em>BamHI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPCD.P</td>
<td>2.5 kb VF395M <em>PstI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPE.B</td>
<td>5 kb VF395M <em>BamHI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPE.C</td>
<td>1.7 kb VF395M <em>ClaI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPE.F</td>
<td>6.7 kb VF395M <em>BamHI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPE.PL</td>
<td>1.2 kb VF395M <em>PstI</em> fragment cloned in pBSISK(\textsuperscript{(+)})</td>
<td>This work</td>
</tr>
<tr>
<td>MCPB(\textsuperscript{ΩTc})</td>
<td>ΩTc cassette cloned in internal NotI site of MCPB.E; contained in pJQ200mp18</td>
<td>This work</td>
</tr>
<tr>
<td>MCPC(\textsuperscript{ΩSp})</td>
<td>ΩSp cassette cloned in internal XhoI site of MCPC.E; contained in pJQ200SK</td>
<td>This work</td>
</tr>
<tr>
<td>MCD(\textsuperscript{ΩNm})</td>
<td>ΩNm cassette cloned in internal HindIII site of MCPD.P; contained in pJQ200SK</td>
<td>This work</td>
</tr>
<tr>
<td>MCE(\textsuperscript{ΩSp})</td>
<td>ΩSp cassette cloned in internal SmaI site of MCPE.C; contained in pJQ200SK</td>
<td>This work</td>
</tr>
</tbody>
</table>

*Antibiotics: Ap, ampicillin; Gm, gentamicin; Nm, neomycin; Sp, spectinomycin; St, streptomycin; Tc, tetracycline.*
PCR. The primer sequences used were 5'-CATGCTCTTC-
TCCCGAGCTTTGC-3' and 5'-GAATCGCCGACAGAC-
CAACC-3', and PCR cycling conditions were 1 min at 94 °C,
30 s at 60 °C and 1 min at 72 °C for 30 cycles. Using the
protocol outlined by Boehringer Mannheim, the amplified
product was PCR labelled with DIG. The DIG-labelled
product was gel-purified using band interception as described
by Sambrook et al. (1989) and was subsequently used as a
probe to identify MCP-encoding genes from a size selected
partial genomic library of VF39SM.

The partial genomic library was constructed by digesting
VF39SM genomic DNA with BamHI to completion. The
digest was electrophoresed in 0.8% agarose gel and the smear
of fragments ranging in size from 5 to 12 kb was isolated using
Prep-a-Gene (Bio-Rad). The fragments were ligated to
BamHI-digested pBSIIKS(+) (+). The ligation products were
used to transform E. coli DH5a.

A putative MCP-encoding gene was isolated from this first
round of screening, and named mcpB. Its identity was verified
by subcloning and sequencing the region that hybridized to the
PCR probe. The subcloned DNA was labelled by the random
primer method (Boehringer Mannheim) and used to reprobe
the partial genomic library. This second round of screening
discovered four previously undetected putative MCP-
encoding genes, termed mcpC through mcpF. Each fragment
was subcloned and sequenced to identify regions of homology
to MCP-encoding genes. New primers C1 (5'-AGGCCGAC-
CGAGCAGAGCCG-3') and C2 (5'-CTTGATTTCTTGG-
CCGC-3') were designed from the DNA sequence of
VF39SM genomic DNA. PCR cycling parameters were as
described above. Subsequently, this probe was used to probe a
complete genomic library of VF39SM.

Preparation of a VF39SM genomic library. Total genomic
DNA from VF39SM was partially digested with SspI/AI to
give a maximum number of fragments in the 30–40 kb size
range. The fragments were ligated to the dephosphorylated
cosmid pRK7813. The ligation mixture was packaged in vitro
using the Boehringer Mannheim packaging kit according to
the manufacturer’s recommendation. Titration of the genomic
library yielded approximately 2500 c.f.u. Cosmid DNA was
isolated as alkaline lysis minipreps from 100 independent
tetracycline-resistant clones and verified on 0.8
agarose gels. Cosmid DNA from each individually purified clone was then
macerate was plated on the appropriate selective media to
form the nodule. For each competition experiment 100
nodules were crushed and plated.

RESULTS

Cloning of putative mcp genes

PCR reactions using VF39SM genomic DNA and the
dcrA primers (Dolla et al., 1992) amplified a single
fragment approximately 350 bp in size. This fragment
was used to probe a genomic blot of VF39SM digested
with a variety of restriction endonucleases. Digestion
of VF39SM genomic DNA with BamHI gave at least seven
hybridizing bands ranging in size from 5 to 12 kb (data
not shown). Consequently, to enrich for fragments of
this size a size-fractionated library of VF39SM was
created. Four hundred colonies containing inserts, based
on blue/white selection, were included in the library.
Five putative mcp clones were subsequently isolated
from this library. Fig. 1 provides restriction maps of
mcpB through mcpF. Fifteen additional clones showing
homology to an mcpD probe (all unique based on the
size of the hybridizing fragment) were identified from a
cosmid library of VF39SM. No further characterization

Insertional mutagenesis of VF39SM mcp genes. mcp genes
were mutated by inserting antibiotic-resistance cassettes into
the ORFs of the cloned mcp genes. The antibiotic-resistance
cassettes used were those developed by Prentki & Krisch
in VF39SM was performed using a suicide vector and a positive
selection strategy (Quandt & Hynes, 1993) for double
homologous recombinational events. Southern blotting (data
not shown) was used to verify the replacement of the wild-type
gene with the mutant gene.

Nodulation competition experiments. Peas (Pisum sativum
cv. Trapper) were surface-sterilized by washing in 50% bleach
for 5 min, then with 70% ethanol, followed by three rinses in
sterile distilled water. The seeds were germinated on water
agar plates for 3 d and then transferred to modified magenta
jars which were modelled to resemble Leonard jars (Vincent,
1970). The peas were grown in a vermiculite substrate.

After transfer to the magenta jars the germinating pea seeds
were inoculated with the appropriate VF39SM strains. For
competition experiments, the wild-type was co-inoculated
with the mcp mutant in a 1:1 ratio. The ratios were confirmed
by performing viable plate counts on the inoculum. The
inoculated peas were then grown for 5 weeks, after which the
nodules were harvested, surface-sterilized and crushed. The
macerate was plated on the appropriate selective media to
determine whether the wild-type or the mcp mutant strain had
formed the nodule. For each competition experiment 100
nodules were crushed and plated.

![Image of restriction maps of the original BamHI fragments cloned from the size-fractionated genomic library of VF39SM.](image)

Fig. 1. Restriction maps of the original BamHI fragments cloned from the size-fractionated genomic library of VF39SM (see text for details). B, BamHI; E, EcoRI; H, HindIII; K, KpnI; N, NotI; P, PstI; S, SalI; Sa, SalII; Sm, Smal; X, Xhol. The locations and omega cassettes used for insertional mutagenesis of mcpB, mcpC, mcpD and mcpE are indicated by the triangles above each restriction map.
Rhizobium leguminosarum chemotaxis chemoreceptors

**McpC**

Identities = 153/383 (39%), Positives = 205/383 (53%)

McpC:

<table>
<thead>
<tr>
<th>Identity</th>
<th>Alignment Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA-1A-2</td>
<td>716 aa</td>
</tr>
</tbody>
</table>

McpA:

<table>
<thead>
<tr>
<th>Identity</th>
<th>Alignment Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA-1A-2</td>
<td>624 aa</td>
</tr>
</tbody>
</table>

**McpE**

<table>
<thead>
<tr>
<th>Identity</th>
<th>Alignment Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA-1A-2</td>
<td>261 aa</td>
</tr>
</tbody>
</table>

**McpF**

<table>
<thead>
<tr>
<th>Identity</th>
<th>Alignment Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIA-1A-2</td>
<td>282 aa</td>
</tr>
</tbody>
</table>

![Fig. 2](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAIgAAAABwCAIAAADG1rBAAAAGXRFWHRTb2Z0d2FyZQBBZG9iZSBJbWFnZVJlYWR5ccllPAAAA5pVFRFZnJhZ2ulPSJXZ3JhZ2VZPSJhbGlnaW9uZSJGPD94gAABCcAAAOAIAAAA8AAAABhJRU5ErkJggg==)

Fig. 2. Predicted amino acid sequences of McpC, McpE and McpF derived from complete or partial DNA sequencing of mcpC, mcpE and mcpF were submitted to a BLASTX search (Altschul et al., 1990). The C. crescentus McpA sequence is shown below; the numbers flanking the McpA sequence represent the amino acid residue numbers of McpA. Plus signs between the sequences indicate similar residues. The methylation regions are represented by bold characters. The bold asterisks indicate sites of potential methylation in McpA (Alley et al., 1992). The McpC sequence contains both the methylation domains and the highly conserved signalling region of McpA, while the sequences from McpE and McpF contain the first methylation domain. The locations of the omega antibiotic-resistance cassettes used in insertional mutagenesis in relation to the amino acid sequence of each putative MCP are also indicated.

**Sequencing of the mcp clones**

The BamHI fragments originally cloned from the VF39SM partial genomic library, labelled MCB.B through MCPF.B, were subcloned to isolate smaller restriction fragments that hybridized to the dcrA primer derived PCR probe. The subclones were then sequenced from vector T7 or T3 primers, and deduced amino acid sequences were analysed by BLASTX for homology with protein databases. High similarities were reported to sequences of known MCPs. The highest alignment scores occurred with MCP sequences from R. leguminosarum (Brito et al., 1996), Rhodobacter capsulatus (Michotey et al., 1996) and Caulobacter crescentus (Alley et al., 1992). Sequence alignments of McpC, McpE and McpF with McpA of C. crescentus are shown in Fig. 2. The nucleotide sequence of mcpB has been deposited in GenBank (accession number AF036168) indicates that it encodes a protein containing two predicted transmembrane regions in the amino-terminal domain, and has all the characteristics common to known MCPs: a periplasmic amino-terminus, putative methylation domains, and a conserved signalling domain.

The mcpD gene sequence has been deposited in GenBank (accession number U81828). Two potential translational start sites exist for the mcpD ORF. The first is the ATG codon at nt 40, while the second is an ATG start codon at nt 67. Neither is preceded by a strong ribosome-binding site. The gene is, however, expressed in VF39SM as indicated by a promoterless lacZ fusion to mcpD (data not shown). The predicted protein based on the first start codon has 624 amino acids and a molecular mass of 767 kDa. Based on a Kyte–Doolittle plot, two transmembrane domains are present. These encompass amino acid residues 18–37 and 298–317. The area flanked by the transmembrane domains is, therefore, 261 aa in size and is presumably located in the periplasm. The complete DNA sequence of mcpC (GenBank accession number AF036168) indicates that it encodes a protein containing two predicted transmembrane regions in the amino-terminal domain, and has all the characteristics common to known MCPs: a periplasmic amino-terminus, putative methylation domains, and a conserved signalling domain.

of any of these clones is available to verify that they code for MCPs.

Sequencing of the mcp clones

The BamHI fragments originally cloned from the VF39SM partial genomic library, labelled MCB.B through MCPF.B, were subcloned to isolate smaller restriction fragments that hybridized to the dcrA primer derived PCR probe. The subclones were then sequenced from vector T7 or T3 primers, and deduced amino acid sequences were analysed by BLASTX for homology with protein databases. High similarities were reported to sequences of known MCPs. The highest alignment scores occurred with MCP sequences from R. leguminosarum (Brito et al., 1996), Rhodobacter capsulatus (Michotey et al., 1996) and Caulobacter crescentus (Alley et al., 1992). Sequence alignments of McpC, McpE and McpF with McpA of C. crescentus are shown in Fig. 2. The nucleotide sequence of mcpB has been deposited in GenBank (accession number AF036168). DNA sequence data were obtained using the clone MCB.B and from a genomic cosmid clone carrying the entire mcpB gene. The predicted MCB protein is 716 aa in length with a molecular mass of 767 kDa. Based on a Kyte–Doolittle plot, two transmembrane domains are present. These encompass amino acid residues 18–37 and 298–317. The area flanked by the transmembrane domains is, therefore, 261 aa in size and is presumably located in the periplasm. The complete DNA sequence of mcpC (GenBank accession number AF036168) indicates that it encodes a protein containing two predicted transmembrane regions in the amino-terminal domain, and has all the characteristics common to known MCPs: a periplasmic amino-terminus, putative methylation domains, and a conserved signalling domain.

The mcpD gene sequence has been deposited in GenBank (accession number U81828). Two potential translational start sites exist for the mcpD ORF. The first is the ATG codon at nt 40, while the second is an ATG start codon at nt 67. Neither is preceded by a strong ribosome-binding site. The gene is, however, expressed in VF39SM as indicated by a promoterless lacZ fusion to mcpD (data not shown). The predicted protein based on the first start codon has 624 amino acids and a molecular mass of 767 kDa. A Kyte–Doolittle plot predicts two transmembrane regions, the first from residues 1 to 26 or 8 to 26, depending on the transcriptional initiation site. The second trans-
Fig. 3. Continued on facing page.
membrane domain extends from residue 179 to 201. Based on these observations, McpD should contain a periplasmic domain of approximately 147 aa.

The sequence alignment in Fig. 3 shows the strong similarity of the McpB and McpD protein sequences to the C. crescentus McpA sequence (Alley et al., 1992), the R. leguminosarum McpA (Brito et al., 1996), and the E. coli Tsr sequence (Boyd et al., 1983). The strongest similarities occur within the methylation and signalling domains of the MCPs.

**Mutagenesis of the mcp genes**

mcpB, mcpC, mcpD and mcpE were mutated via insertional mutagenesis, using antibiotic-resistance cassettes as the selectable markers (Prentki & Krisch, 1984). Restriction endonuclease sites used for insertions of antibiotic resistance cassettes were ascertained from the sequence data generated in this study. Figs 1 and 2 indicate the location of gene disruption sites in each mcp clone. Insertion of the QTc fragment into the NotI site of mcpB disrupted the gene within the putative second methylation site. The QSp fragment was inserted just downstream of the highly conserved domain of mcpC using a XhoI site. In mcpD the QKm cassette inserted at the HindIII site ahead of the first methylation region. With mcpE, insertional mutagenesis also occurred using a SmaI site located slightly upstream of the first methylation region. In addition to the single gene knockouts created, an mcpC, mcpD double mutant was obtained. The mutant was created by mating MCDQNm into VF-MCPC−, and then selecting for double recombinants.

Phenotypic analysis of the mutants was performed using VMM swarm plates to assess chemotactic behaviour towards various carbon sources (adonitol, arabinose, arginine, aspartate, erythritol, galactose, glutamate, glycerol, histidine, malate, manitol, melibiose, ornithine, raffinose, ribose, rhamnose, serine, sorbitol and trigonelline). To date, phenotypes for mutations in mcpC, mcpD and mcpE have not been identified. However, a mutation in mcpB inhibited chemotaxis to all carbon sources tested. This phenotype does not conform to the E. coli model whereby a MCP is responsive to a specific set of ligands. Fig. 4 provides examples of the swarming patterns of the mutants. VF-MCPB− is impaired in its ability to swarm away from the inoculation point. Cells accumulate in a localized area, resulting in a denser circle of growth relative to the wild-type and the other mutant strains. This is particularly evident in plate (a). In addition, the swarm diameter of VF-MCPB− is smaller relative to the other strains. This behaviour is most pronounced in the raffinose swarm plate.

Possible effects on the ability of the mcp mutants to interact with legume host plants were assayed by nodulation competition tests. The principle of this type of study is that if a plant is co-inoculated with two strains of rhizobia, formation of more nodules by one strain indicates greater 'competitiveness'; this may be due to a variety of factors, including faster growth in the

---

**Fig. 3.** Amino acid sequence alignment of McpB and McpD with R. leguminosarum McpA, C. crescentus McpA, and E. coli Tsr. The shaded rectangular boxes denote the approximate location of the transmembrane domains for each MCP. The boxed residues indicate identical amino acids among the five chemotaxis proteins. McpB and McpD have high homology to the methylation domains and signalling domain of both of the McpA proteins and of Tsr. The sequence alignment was performed using the CLUSTAL V program, and displayed using the SeqVu program (Garvan Institute).
Fig. 4. Swarm plates showing the chemotactic behaviour of the various VF39SM MCP mutants. Strains are labelled as follows: A, VF-MCPB--; B, VF-MCPC--; C, VF-MCPD--; D, VF-MCPE--; E, VF39SM. (a) Yeast extract swarm plate. (b) VMM raffinose swarm plate. (c) VMM adonitol swarm plate. The swarm plates were prepared as described in Methods.

Fig. 5. Nodulation competition experiments: inoculation and recovery ratios of mcp mutants versus the wild-type VF39SM. Trapper peas were co-inoculated with VF39SM and an individual mcp mutant strain as described in Methods. The mcp mutants co-inoculated were as follows: set 1, VF-MCPB--; set 2, VF-MCPC--; set 3, VF-MCPD--; set 4, VF-MCPC'/D-. After 5 weeks incubation, 100 nodules were crushed to determine which strain formed each nodule. Since bacteria in each nodule are usually of clonal origin a selectable marker can be used to determine the number of nodules formed by a strain. Ratios are expressed as the mcp mutant/VF39SM. An asterisk indicates that the recovery ratio differed significantly from the initial inoculation ratio. Statistical significance was confirmed using the chi-squared test.

DISCUSSION

Our results provide evidence for the existence of a family of mcp-like genes in R. leguminosarum VF39SM. Positive hybridization to a probe containing DNA corresponding to the conserved signalling domain of MCPs, and sequence similarities between MCPs and the deduced amino acid sequences for McpB through McpF, suggest that these genes do code for MCP-like proteins.
In addition, the identification of 15 further clones showing homology to the highly conserved domain of mcp genes strongly suggests that, as in Desulfovibrio (Deckers & Voordouw, 1994) and Halobacterium (Zhang et al., 1996), Rhizobium spp. may contain numerous genes encoding proteins showing similarity to MCPs. Further evidence for this comes from recent publications showing that genes homologous to the mcp genes we have identified are present on the nodulation plasmids of another strain of R. leguminosarum (Brito et al., 1996) and of Rhizobium sp. NGR234 (Freiberg et al., 1997); also, genes encoding proteins with amino acid sequence similarity to the signalling domain of MCPs in S. meliloti have been described (Greck et al., 1995). None of these genes appears to be closely related to any of the genes from strain VF39SM for which we have complete or partial sequence data. Further investigation will be required before a role (if any) for these genes in chemotaxis can be defined. We do believe that the majority of the clones we have identified from VF39SM code for proteins with homology to MCPs for two reasons: (1) Western blots using antibodies against the MCP highly conserved domain reveal the presence of at least 10 proteins that react with the antibody (data not shown); and (2) most clones hybridized at least as strongly to several mcp-derived probes as did the five genes for which we have DNA sequence data. Since five clones chosen more or less at random from those hybridizing with mcp probes code for MCP-like proteins, it is highly probable that some of the other clones do so as well. However, it is also possible that some of the clones carrying DNA homologous to mcp genes are pseudogenes, or code for proteins like the tcpI and acfB gene products in Vibrio cholerae (Harkey et al., 1994; Everiss et al., 1994), which show similarity to the MCP signalling domain, but apparently do not function as chemoreceptors, although they have important roles in other facets of signal transduction.

All the putative mcp genes for which DNA sequencing (complete or partial) was carried out had high homology to the methylation domains characteristic of MCPs. The amino acid sequences deduced from the DNA sequences of mcpB, mcpC and mcpD show all the characteristic features of MCPs, such as transmembrane domains, methylation sites, and the signalling domain. The genes also code for proteins of a size typical for MCPs. Of the other putative MCP-encoding genes so far identified in the rhizobia, those from S. meliloti (Greck et al., 1995) appear unusual in that the methylation sites and transmembrane regions are missing, whereas those from NGR234 (Freiberg et al., 1997) and R. leguminosarum bv. viciae UPM791 (Brito et al., 1996) appear to code for fairly typical MCPs.

Our results suggest that some of the mcp genes we have identified play a role in chemotaxis. While mutations in mcpC, mcpD and mcpE conferred no alteration in chemotactic properties towards a variety of sugars, amino acids and other carbon sources, mutations in mcpB appeared to create a general chemotactic block. The reason for this remains unclear, but mcpB mutations resemble mcpA mutants in Rhodobacter sphaeroides (Ward et al., 1995), which lost chemotaxis to a wide range of carbon sources, rather than a discrete set. In addition, this phenotype was apparent only under aerobic conditions. This suggests that the role of certain MCPs in chemotaxis of members of the alpha sub-group of the proteobacteria may deviate significantly from the E. coli paradigm.

Further evidence that several of the mcp genes we have identified have a biological function comes from our nodulation competition experiments. It has long been speculated that chemotaxis may play a role in competition between Rhizobium strains in the legume rhizosphere, and indeed some studies (Ames & Bergman, 1981; Caetano-Anollés et al., 1988; Bauer & Caetano-Anollés, 1990) have shown that non-motile and some non-chemotactic strains cannot compete success-

**Fig. 6.** (a) Typical Eckhardt gel showing VF39SM and the plasmid-cured derivative strains prepared by Hynes & McGregor (1990). The strain in each lane is as follows: A, VF39SM; B, LRS39201 (cured of pRleVF39b); C, LRS39301 (cured of pRleVF39c); D, LRS39401 (cured of pRleVF39d); E, LRS39501 (cured of pRleVF39e); F, LRS39601 (cured of pRleVF39f). (b) Southern blot of an Eckhardt gel with the same strain order as that in (a) probed with the 1.5 kb EcoRI fragment of MCPC.B. A band is present in all lanes except that (F) containing LRS39601. The letters A-F indicate the positions of the six plasmids of strain VF39 (pRleVF39a, pRleVF39b, etc.) as described by Hynes et al. (1988).
fully with their wild-type counterparts. Our results establish not only that chemotaxis per se is important in competition for nodulation, but that subtle, and as yet uncharacterized, effects on chemotactic behaviour may also affect competitiveness. Not only was the mcpB (chemotaxis minus) mutant greatly reduced in its ability to compete, but the mutations in mcpC, which caused no apparent chemotactic phenotype in laboratory assays, also had drastic effects on competition for nodule occupancy. This suggests that these genes may code for proteins that sense highly specific compounds in the plant environment, and that chemotaxis towards these compounds may be an important aspect of signal exchange between plant and bacterial partners in the symbiosis.

Identifying ligands for a large family of putative chemoreceptors could prove to be a daunting task, given the great metabolic diversity of Rhizobium species (Stowers, 1985; Parke & Ornston, 1984), and the wide diversity of compounds that can serve as chemooattractants for rhizobia (Parke et al., 1985). Adding to the difficulty, some of the potential chemooattractants may be compounds from root exudates that have not yet been identified. The finding that certain putative mcp genes are located on plasmids, both in strain VF39SM, as is the case for mcpC and at least three other cosmid clones, and in other rhizobia (Brito et al., 1996; Freiberg et al., 1997) may be of some use in determining functions. It is not unlikely that mcp genes located on plasmids code for proteins that sense attractants catabolized by enzymes encoded by those plasmids. Many catabolic genes have been mapped to plasmids in rhizobia. Further work on each of the clones we have obtained will determine the precise role the putative mcp genes play in chemotaxis and other cellular functions, and should establish conclusively whether a large family of such genes exists in rhizobia.

ACKNOWLEDGEMENTS

We thank Melanie Berndt and Christa Martin tole for technical assistance, and Alberto Mendoza for many useful discussions. The financial support of a Natural Sciences and Engineering Council research grant to M.F.H. is gratefully acknowledged.

REFERENCES


Received 22 December 1997; revised 6 March 1998; accepted 16 March 1998.