BMC Microbiology

Research article

Characterization of the nodulation plasmid encoded chemoreceptor gene *mcpG* from *Rhizobium leguminosarum*

Christopher K Yost, Kirsten T Clark, Kate L Del Bel and Michael F Hynes*

Address: Department of Biological Sciences, University of Calgary, 2500 University Drive SW, Calgary, Alberta, Canada, T2N 1N4

Email: Christopher K Yost - ckjost@yahoo.com; Kirsten T Clark - suki6@sprint.ca; Kate L Del Bel - bkdel@ucalgary.ca;
Michael F Hynes* - hynes@ucalgary.ca

* Corresponding author

Abstract

Background: In general, chemotaxis in *Rhizobium* has not been well characterized. Methyl accepting chemotaxis proteins are sensory proteins important in chemotaxis of numerous bacteria, but their involvement in *Rhizobium* chemotaxis is unclear and merits further investigation.

Results: A putative methyl accepting chemotaxis protein gene (*mcpG*) of *Rhizobium leguminosarum* VF39SM was isolated and characterized. The gene was found to reside on the nodulation plasmid, pRleVF39d. The predicted *mcpG* ORF displayed motifs common to known methyl-accepting chemotaxis proteins, such as two transmembrane domains and high homology to the conserved methylation and signaling domains of well-characterized MCPs. Phenotypic analysis of *mcpG* mutants using swarm plates did not identify ligands for this putative receptor. Additionally, gene knockouts of *mcpG* did not affect a mutant strain's ability to compete for nodulation with the wild type. Notably, *mcpG* was found to be plasmid-encoded in all strains of *R. leguminosarum* and *R. etli* examined, though it was found on the nodulation plasmid only in a minority of strains.

Conclusions: Based on sequence homology *R. leguminosarum mcpG* gene codes for a methyl accepting chemotaxis protein. The gene is plasmid localized in numerous *Rhizobium* spp. Although localized to the sym plasmid of VF39SM *mcpG* does not appear to participate in early nodulation events. A ligand for McpG remains to be found. Apparent McpG orthologs appear in a diverse range of proteobacteria. Identification and characterization of *mcpG* adds to the family of *mcp* genes already identified in this organism.

Background

Using a process termed chemotaxis, motile bacteria are capable of sensing their external environment and responding appropriately by moving towards increasing concentrations of nutrients and away from increasing concentrations of toxic compounds. The chemotaxis signaling pathway has been well characterized due to extensive studies using the model organism *Escherichia coli*, as well as detailed molecular characterization in numerous taxonomically distinct bacteria [1,2]. Methyl accepting chemotaxis proteins (MCPs) play key roles in the chemotactic response of many bacteria. These sensor proteins have N-terminal domains that detect attractants and repellents. When a ligand binds to a MCP, information about the external environment is transmitted from the MCP to a two component signal transduction system (CheA/CheY) inside the cell. The effect is alteration of swimming behavior by changing the direction of flagellar rotation or, in some cases, the speed of flagellar rotation, resulting ultimately in movement towards higher
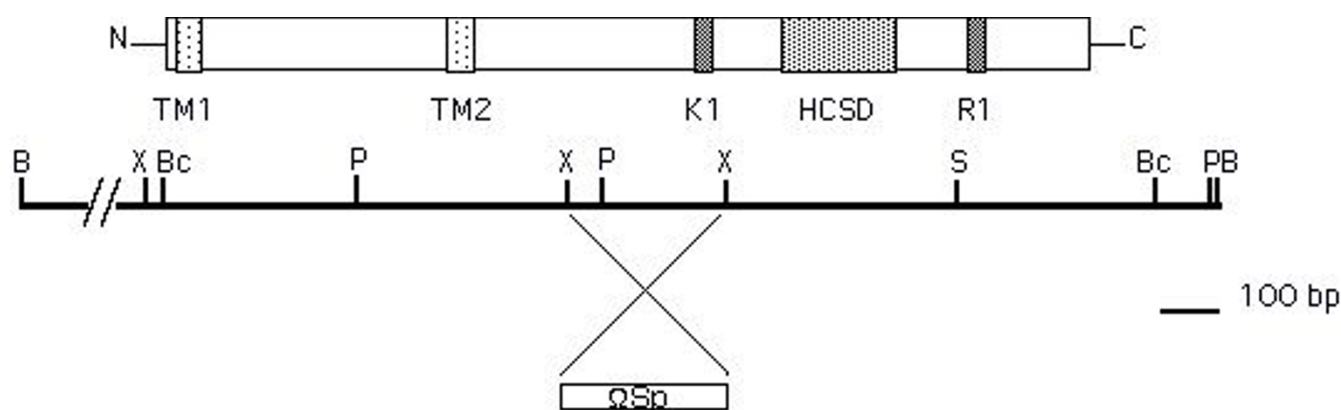
Figure 1
The restriction map of *mcpG* was drawn to scale relative to the translation product of a typical *mcp* ORF (TM, transmembrane domain; K1 and R1, methylation regions; HCSD, highly conserved signaling domain). The site of insertion of the Sp omega fragment and subsequent deletion in *mcpG* is also shown. Restriction enzymes are as follows: B, *Bam*HI; Bc, *Bcl*I; P, *Pst*I; S, *Sal*I; and X, *Xho*I.

gradients of attractant and away from high concentrations of repellents [see [1,3,4] for recent reviews].

MCPs contain highly conserved structural domains; typically there are two transmembrane domains, a highly conserved signaling domain involved with CheA interaction, and two domains for methylation of glutamate residues [5]. MCP homologues have been discovered in a wide range of bacteria [[6,7] for review]. However, the function of these proteins in chemotaxis in many cases is only predicted through homology to the archetypal MCPs from *E. coli* and consequently the ligands and the precise role of these putative MCPS in chemotactic signaling have yet to be determined. Despite detailed characterization of the signaling pathways used in bacterial chemotaxis, the ecological significance of chemotaxis in bacteria remains relatively uncharacterized.

Chemotaxis has long been suggested to play a role in the early events of nodulation between rhizobia and host legumes [8–16]. Many legume root exudate compounds are chemoattractants for rhizobia. Both *Rhizobium leguminosarum* and *Sinorhizobium meliloti* exhibit chemotaxis towards the flavonoid compounds that induce symbiotic nodulation (*nod*) genes [13,15]. During infection the *Rhizobium* site of entry is commonly the tip of a developing root hair [17]. The legume plant may enhance infection by directing the rhizobia to the proper site of infection through secretion of chemoattractants. Chemotaxis towards plant root exudates could amplify *nod* gene induction by stimu-

lating movement towards increasing inducer concentrations present at root surfaces [15].

In many species of rhizobia, symbiotic nitrogen fixation genes are localized on large plasmids, commonly referred to as sym plasmids (pSyms) or nodulation plasmids. In addition to these plasmids, most rhizobia harbor multiple large plasmids, some of which are still cryptic, and others whose function has only recently been elucidated [18–25]. It has been shown that these cryptic plasmids may play a role in rhizosphere competitive fitness [[19,20,24,25], Hynes unpublished]. Genes with homology to MCP-encoding genes have been previously reported on plasmids in strains from *Rhizobium leguminosarum* [16,26], *Rhizobium* sp. NGR234 [27] and *Sinorhizobium meliloti* [28]. In one instance, mutation of a plasmid residing *mcp* gene, *mcpC* from *R. leguminosarum* biovar *vicae* VF39SM, resulted in loss of ability to compete against wild type in the formation of nodules on Trapper peas [16]. The presence and significance of plasmid encoded *mcp* genes in rhizobia remains relatively unstudied. Studying sym plasmid localized *mcp* genes is of particular interest as co-localization with nodulation genes may suggest a role for nodulation related chemotaxis. This work characterizes *mcpG*, a sym plasmid encoded *mcp* gene, isolated from *R. leguminosarum* biovar *vicae* VF39SM.

Results
mcpG Cloning and Gene Characterization
mcpG was identified in a R. leguminosarum VF39SM cosmid library using a DNA probe derived from the highly

A

```
McPG 374:DLAKRTEQQAASVEETAAALEEITTTYADSSNKAQEAGQLVRKTKDSAERSGSYVRDAYD
TSR  290:DLSSRTEQQAASLEETAASMEQLTATVKQNAENARQASHLALSASETAQRGGKVVDNVVQ
              *             *          *

McPG 434:AMGKIESSATEIGSIIGVIDEIAFQTNLLALNAGVEAARAGEAGKGFAVVAQEVRELAQR
TSR  350:TMRDISTSSQKIADIISVIDGIAFQTNILALNAAVEAARAGEQGRGFAVVAGEVRNLAQR

McPG 494:SAKAAKEIKELINASNGHYKSGVALVGETGKALKEIAEQVQQVDGNVGAIVGASQEQATG
TSR  410:SAQAAREIKSLIEDSVGKVDVGSTLVESAGETMAEIVSAVTRVTDIDGEIASASDEQSRG

McPG 554:LKEINTAVNRMDQGTQQNAAIVEEATAA
TSR  470:IDQVGLAVAEMDRVTQQNAALVEESAAA
                                *
```

B

```
McPG    1:RTIVGPLAKLQLSLKAHADGNLDGDRRLEARGDEIGKLARAVAGLRDAISAKA.
HAMP  218:RRLLRPLRRLAEAAQRIADGDLTVRLPVDGR.DEIGELARAFNEHADRLEETIA
```

Figure 2
A: ClustalW alignment of McpG with Tsr. The sequence alignment shown spans the C-terminal domains of each protein. Tsr residues 297, 304, 311, 493 are known sites for methylation by CheR [60,61] B: ClustalW alignment of McpG with the consensus sequence for the HAMP domain. The consensus sequence was obtained from the SMART collection of conserved domains http://smart.embl-heidelberg.de/. The program Macboxshade (MD Baron, Institute for Animal Health, Surrey, UK) was used to visualize the alignments. Yellow shaded residues are identical while similar residues are shaded green.

conserved signaling domain of R. leguminosarum VF39SM mcpD [16]. The mcpG gene was subcloned from a cosmid as a BamHI fragment. Figure 1 provides a restriction map of mcpG. The entire mcpG coding region plus adjacent DNA regions was sequenced using a combination of subclones and use of appropriate primers.

The *mcpG* gene sequence has been deposited in GenBank under the accession number AF141674. The open reading frame of *mcpG* is predicted to be 654 aa, with a molecular mass of 68.48 kDa. Transmembrane domains are predicted in McpG using two separate transmembrane helix prediction programs. TopPred II [29] predicts two transmembrane domains (TMD), the first spanning aa residues 12 to 32 and the second from aa 196 to 216. TM-HMM [30] also predicts two transmembrane regions similar in location to TopPred II; first TMD from aa 7 to 29 and the second from aa 194 to 216. These predictions suggest the N-terminal domain of McpG is located in the periplasmic space and likely functions as a sensor domain. The software predictions are in good agreement with the secondary structure of a typical MCP protein, whereby two transmembrane domains flank the N-terminal domain of the protein (Figure 1).

The C-terminal domain of McpG contains strong sequence homology to the methylation, and signaling domains of well characterized MCP proteins. Figure 2 shows an alignment of the C-terminal domains of McpG and the *E. coli* MCP Tsr. Results from a BLASTP alignment search of the GenBank database show that *mcpG* ORF exhibits the greatest similarity to putative MCPs from *Agrobacterium tumefaciens* C58 (McpU, Genbank accession AAK87918) and *S. meliloti* (McpY, Genbank accession AAG37852). However, a BLASTP alignment search using only the sequence flanked by the transmembrane domains (aa residues 34–190), presumed to be the sensor domain of McpG, yields different results. Figure 3 shows the alignment of the sensor region to ORFs of probable MCPs submitted to the Genbank database from genome sequencing projects (*Pseudomonas syringae*, http://genome.ornl.gov/microbial/psyr/; *Rhodospirillum rubrum*, http://genome.ornl.gov/microbial/rrub/; *Xanthamonas campestris*, http://cancer.lbi.ic.unicamp.br/xanthomonas/). A MCP sensor region detects attractants and repellents, directly or indirectly; therefore genes that are potential orthologs are those with significant homologies in the N-terminal sensor regions. To date, BLASTP searches of the sensor regions from the previously characterized *R. legu-*

Figure 3
ClustalW alignment of the sensor region of McpG with the sensor regions of putative MCP proteins from *Xanthamonas campestris*, *Rhodospirillum rubrum*, and *Pseudomonas syringae*. The sensor region of each respective MCP is defined as the N-terminal area flanked by the two predicted transmembrane domains. The Genbank Accession numbers for the complete protein sequences used in the alignment are as follows: *Xanthamonas campestris*, NP_637412; *Rhodospirillum rubrum*, ZP_00014650; *Pseudomonas syringae*, ZP_00124464.

minosarum VF39SM MCPs [16] have not detected orthologs in any bacteria, including members of the *Rhizobiaceae* family (data not shown).

The genome sequencing of *R. leguminosarum* bv. *viciae* 3841 is ongoing at the Welcome Trust Sanger Institute http://www.sanger.ac.uk/Projects/R_leguminosarum. Searching the database for sequence homology to *mcpG* yields a single gene that is 99% identical to the *mcpG* sequence. In strain 3841 a gene precedes the *mcpG* homologue with homology to a family of monooxygenases. The highest homology (80% identical) is to an alkanesulfonate monooxygenase from *A. tumefaciens* C58 (Genbank accession AAL44243). Downstream of the *mcpG* homologue is a gene coding for a monocarboxylate permease (Genbank accession AJ421944) that is required for optimal growth with alanine as a sole carbon and nitrogen source [31]. Our sequencing results on upstream and downstream regions of *mcpG* from VF39 show that the gene organization in this strain is the same as in 3841.

In addition to the conserved methylation and signaling domains found in McpG (figure 2a) another conserved domain likely exists. Searching the NCBI Conserved Domain Database http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml revealed that the *mcpG* ORF has significant homology to the HAMP motif (figure 2b). The HAMP domain was suggested as a conserved amino acid motif found in histidine kinases, adenylyl cyclases, methyl-accepting proteins and phosphatases by Aravind and Ponting [32]. McpG has a single putative HAMP motif,

spanning from amino acid residues 218–267. This places the domain immediately downstream of the 2nd transmembrane domain (196–216 aa). This close proximity of the HAMP domain to a transmembrane domain is typical of characterized HAMP containing proteins [32]. It is suggested that the HAMP domain in MCPs has a role in inhibiting autophosphorylation of CheA within the MCP-CheW-CheA multimolecular complex [32].

The C-terminal sequence of McpG (GDWEEF) may be significant. This sequence is highly similar to the C-terminus of McpD [16] and variations on it, with the last four amino acids highly conserved, are also found in many of the predicted MCP proteins detected in the genomic sequences of *S. meliloti* [28] and *A. tumefaciens* [33]. A function for this motif has yet to be determined.

Plasmid localization and distribution of *mcpG* in *Rhizobium spp*

A DNA probe from the ca. 500 bp *Pst*I fragment of *mcpG* was used to search for *mcpG* homologues in other *Rhizobium* spp. The probe originates from the 5' region of the gene, thereby lacking the highly conserved region of *mcp* genes, and is therefore unique to *mcpG* (Fig. 1); the fact that this probe did not hybridize to other *mcp* genes was confirmed by Southern blots to total DNA of VF39, which gave only one band as expected, in spite of the presence of at least 17 *mcp* gene homologues in this strain [16]. As evidenced by lanes A-C in Figure 4, *mcpG* resides on pRleVF39d of VF39SM. Additionally, *mcpG* homologues are found on plasmids from a number of *R. leguminosarum*
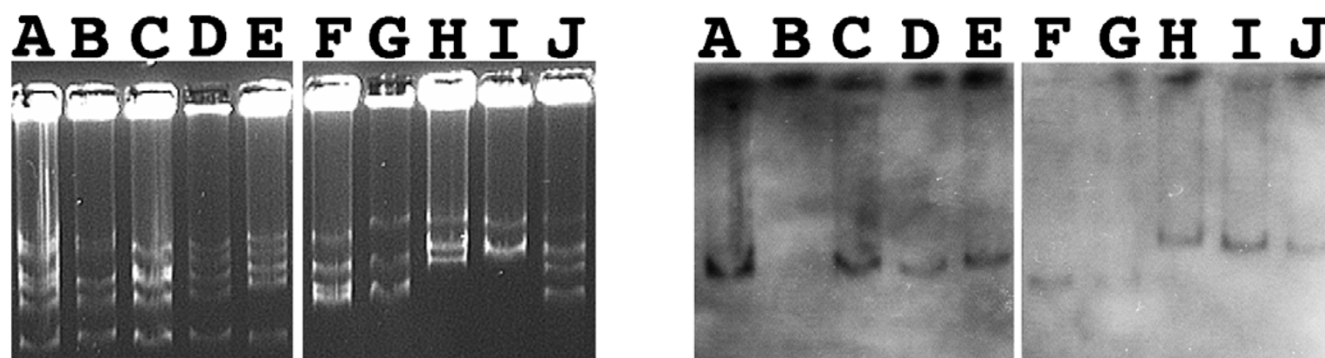
Figure 4
Panel 1 is a digital image of an Eckhardt gel showing plasmid profiles from various *R. leguminosarum* strains. The contents of each lane are as follows: A, VF39SM; B, LRS39401(=VF39SM cured of pSym); C, LRS39601; D, 3841; E, T3CB; F, W14-2; G, 162Y10; H, 8002; I, 8401; and J, GF160. Panel 2 is a Southern blot, of the Eckhardt gel shown in panel 1, probed with the *Pst*I fragment from *mcpG* (Fig. 1). The lane order is identical to that of the Eckhardt gel from panel 1.

species representing all three biovars of this species. Strains 3841 and T3CB, which are both derived from the same parent strain, carry the *mcpG* gene on the nodulation plasmid pRL10JI. The remaining strains represented in Figure 4 harbor *mcpG* on cryptic plasmids. Based on the published descriptions of these strains and our unpublished results, the pSym in each of these strains is: the largest plasmid in strain 14–2 (lane F), the smallest plasmid in strain 8002 (lane H) and the second smallest plasmid in strain GF160 (lane J). Strain 8401 (lane I) is 8002 cured of its pSym, and the identity of the pSym in strain 162Y10 is not known, though it is not one of the two smallest comigrating bands, one of which is hybridizing to our probe. *mcpG* is also found on plasmids in *Rhizobium etli* strains: Brazil5, F8, and CFN42 (data not shown). In strain CFN42, the plasmid carrying the *mcpG* homologue is pRetCFN42c, which is not the pSym [19].

Insertional mutagenesis of mcpG
Mutagenesis of *mcpG* was accomplished by disrupting the ORF through insertion of a Sp[r] cassette within the coding region of *mcpG*. Figure 1 illustrates the location of the gene disruption relative to the predicted ORF. The mutated gene was introduced into VF39SM via double recombination using the strategy of Quandt and Hynes [34]. Southern blots of genomic DNA from putative *mcpG* mutants confirmed the replacement of the wild type gene with the insertionally disrupted mutant gene (data not shown).

The *mcpG* mutant was screened for altered chemotactic response to a number of carbon sources. Of particular interest were carbon sources whose catabolic genes reside on pRleVF39d, the plasmid that carries *mcpG*. These carbon

sources are: adonitol, alanine, hydroxy-L-proline, and trigonelline [[24,35], Hynes unpublished]. On swarm plates, the *mcpG* mutant maintained the same chemotactic response as wild type to these particular carbon sources (data not shown). Additionally, the chemotactic behavior of the *mcpG* mutant to a wide range of other carbon sources (various sugars and amino acids) was identical to the wild type (data not shown). The chemotactic phenotype of the *mcpG* mutants towards alanine, lactate, and pyruvate, which are known substrates of the monocarboxylate transporter which is encoded by the *mct* gene located adjacent to *mcpG* [31] was also unaltered.

The effect of a *mcpG* mutation on the ability of VF39SM to nodulate competitively was tested using a nodulation competition assay. Figure 5 shows the results of the competition experiment. Two independently isolated *mcpG* mutants were unaffected in their ability to compete for nodulation against the wild type strain VF39SM.

Discussion
The sequence homology of the *mcpG* ORF to known MCP proteins and its secondary structure strongly suggest that *mcpG* codes for a methyl accepting chemotaxis protein. Notably, sequence data from regions surrounding *mcpG* suggest that, as is the case in *R. leguminosarum* 3841, putative oxygenase genes flank *mcpG* and a monocarboxylate transport gene is in close proximity [31]. However alanine, lactate and pyruvate chemotaxis was unaffected in swarm plate assays with a *mcpG* mutant strain. In fact, swarm plate analysis with numerous carbon sources did not reveal a mutant phenotype. Given the large number of compounds that are chemoattractants for rhizobia [36] and the metabolic diversity of *Rhizobium* [37,38] identify-
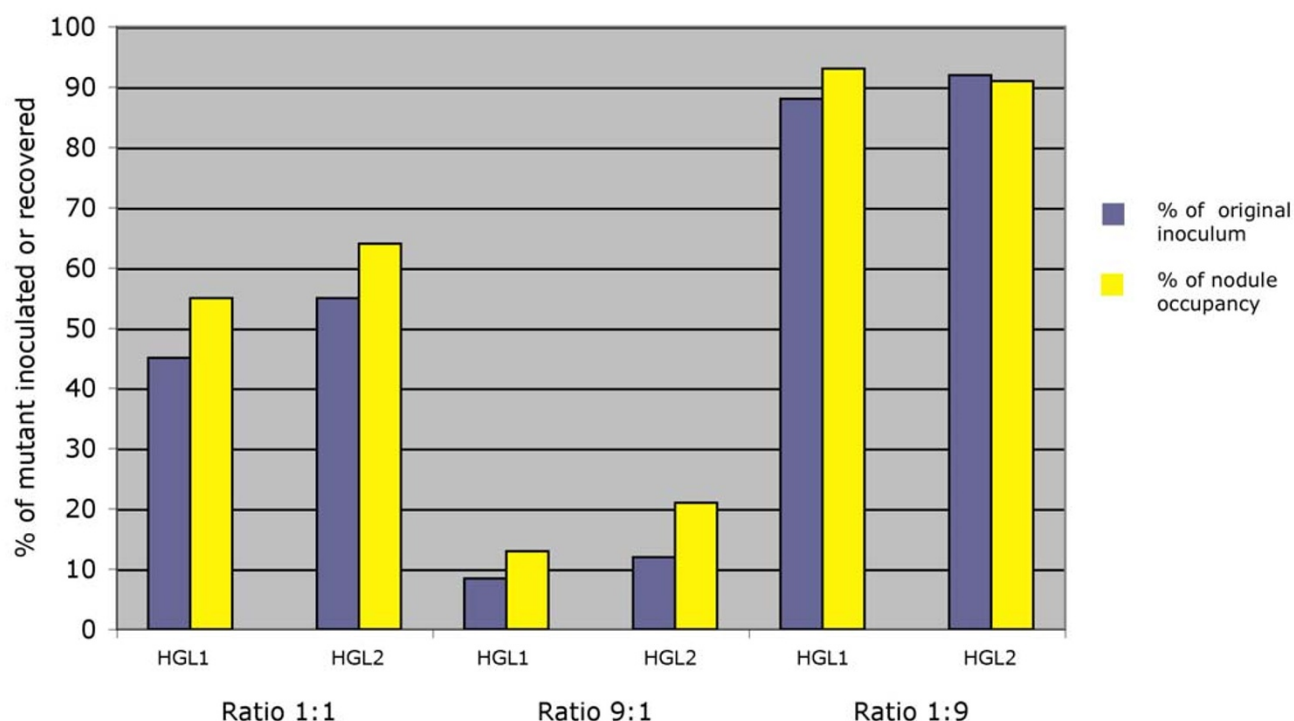
Figure 5
Results of nodulation competition assay between wild type VF39SM and the two *mcpG*- strains, HLG1 and HLG2. The ratios are expressed as VF39SM:HLG1 or HLG2. The graph indicates that the % of nodules occupied by each mutant strain was in accordance to the initial inoculation used. The Chi-square test was used to confirm there was no statistically significant difference between the initial inoculum % and the recovery % from nodules.

ing a ligand for McpG could be difficult. Swarm plate analysis of mutant strains created from using previously isolated VF39SM *mcp* genes failed to reveal mutant phenotypes except in the case of a mutation in *mcpB* which resulted in impairment of a chemotactic response to a variety of carbon sources on swarm plates [16]. Metabolism of the test compound is a prerequisite for the swarm assay, so it is possible that a mutant phenotype is hidden in swarm plate analysis by an overriding redox chemotactic response. Redox chemotaxis has been observed in members of the alpha-proteobacteria family; particularly in *Azospirillum brasilense* where energy taxis is the dominant chemotactic behavior [39], but also in *Rhodobacter sphaeroides* [40]. The capability of redox chemotaxis in *R. leguminosarum* VF39SM is being investigated and an alternative chemotaxis assay that does not require metabolism of the chemoattractant and therefore eliminates overriding effects of any redox chemotaxis is currently being optimized.

Finding putative orthologs of *mcpG* in a diverse range of proteobacteria suggests that the Mcp senses a com-

pound(s) found in the habitats of all these bacteria. All of these organisms have been isolated from soil environments and both *X. campestris* and *P. syringae* are known to colonize the rhizosphere of plant species. Interestingly, no *mcpG* orthologs are found in the genome sequences of *Agrobacterium tumefaciens*, *M. loti*, or *S. meliloti*. Based on the annotated database of *X. campestris*, a family of 19 *mcp* genes may exist; the genome sequences of *P. syringae*, and *R. rubrum*, to date, are drafts and therefore it is premature to conclude the size of the *mcp* gene families in these species. The large *mcp* gene family of *X. campestris* is very similar to that reported for *R. leguminosarum* VF39SM [16].

The results from the nodulation competition assay suggest *mcpG* does not play a role in the early events of plant infection. Previous work with *mcp* genes from VF39SM has shown that at least 2 *mcps* (*mcpB* and *mcpC*) do contribute to the strain's ability to compete for infection sites on pea plants [16]. Although *mcpG* is not needed during early infection events it may offer a competitive advantage in the rhizosphere.

Table 1: Bacterial Strains and Plasmids

Bacterial Strain or Plasmid	Relevant Characteristics	Source or Reference
E. coli		
DH5α	*EndA1, hsdR17, supE44, thi-1, recA1, gyrA96, relA1, • (argF-lacZYA)*, U169, φ 80d*lacZ* ΔM15	Life Technologies
S17-1	Sp^r. RP4 *tra* region, mobilizer strain	[51]
Rhizobium leguminosarum		
VF39SM	biovar *viciae*, Sm^r	[52]
HLG1	VF39SM, *mcpG*:: ΩSp; Sm^r, Sp^r	This work
HLG2	VF39SM, *mcpG*:: ΩSp; Sm^r, Sp^r	This work
LRS39401	VF39SM cured of pRleVF39d	[48]
LRS39601	VF39SM cured of pRleVF39f	[48]
T3CB	biovar *viciae*, 1062 derivative,	[53]
W14-2	biovar *trifolii*, Sm^r	[18]
8002	biovar *phaseoli*	[54]
8401	8002 cured of pSym plasmid	[54]
GF160	biovar *viciae*, Sm^r,	[55]
3841	biovar *viciae*, JB300 derivative	[56]
162Y10	biovar *trifolii*	[57]
Plasmids		
pBSIISK(+)	Cloning vector, Ap^r	Stratagene
pCYΩSp	Sp resistance gene from pHP45Ω [58] flanked by MCS of p1918	[35]
p1918	vector with a MCS composed of inverted repeats of the pUC19 MCS, Ap^r	[59]
pJQ200mp18	delivery vector for gene replacement, Gm^r	[34]
pKC42	6 kb *Bam*HI VF39SM genomic fragment containing *mcpG* cloned into pBSIISK(+)	This work
pCYG1	*Bcl*I fragment with *mcpG* ORF cloned into pJQ200mp18	This work
pCYG2	Derivative of pCYG1 with ΩSp cassette cloned into *mcpG* gene	This work

Plasmids in *Rhizobium* sp. carry important genetic determinants, such as genes required for symbiotic nitrogen fixation, bacteriocin production, and catabolism of various carbon sources [18,21–25,41,42] The plasmid profiles of rhizobia vary greatly, both in size and number of plasmids. The origin and evolution of these plasmids is relatively unknown. Interestingly, in all strains tested, only *R. leguminosarum* VF39SM and *R. leguminosarum* 3841 had *mcpG* on the sym plasmid; in the remaining strains *mcpG* was found on non-sym plasmids. Genes for adonitol metabolism have been localized to plasmids in each of the strains used in Figure 4[24]. Notably, the adonitol catabolism genes appear to localize to the same plasmid as *mcpG* in all the tested strains. The localization of *mcpG* to different plasmids in a number of *Rhizobium* spp. may prove useful in supplementing ongoing studies regarding plasmid origin and evolution in *Rhizobium*.

The ecological significance of plasmid encoded *mcp* genes remains relatively uncharacterized. VF39SM contains at least 4 plasmid encoded *mcp* genes [[16,35], this study]. Additionally the pSym-encoded *mcp* gene isolated in a *R. leguminosarum* bv. *vicae* strain by Brito et al. [26] is different from these 4 genes. Notably, searching the DNA sequence of *Sinorhizobium meliloti* 1021 http:// sequence.toulouse.inra.fr/meliloti.html [28] for homology to *mcp* genes reveals only one plasmid encoded *mcp* gene. This gene resides on pSymA while no ORFs with homology to MCPs are found on pSymB. Both pSymA and pSymB are extremely large plasmids, and, similarly to *R. leguminosarum* plasmids, contain genes for catabolism of many substrates, yet they do not harbor *mcp* genes to the same extent. This contrast suggests that the ecological role of plasmid encoded *mcp* genes in *R. leguminosarum* may be distinctive and significant.

Conclusions

Based on sequence homology the *mcpG* ORF codes for a methyl accepting chemotaxis protein, and appears to be homologous to an MCP found in a diverse group of proteobacteria. The ligand for *mcpG* remains to be discovered and, in terms of early nodulation events, *mcpG* does not contribute to nodulation efficacy. *mcpG* is found widely distributed amongst *Rhizobium* spp., but is located on different types of plasmids in different strains. Isolation and characterization of this gene adds to the family of previously described *mcp* genes in *R. leguminosarum* [16,26,35].

Methods

Bacterial strains, plasmids and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 1. R. leguminosarum strains were routinely cultured on TY medium [43] at 30°C while E. coli strains were cultured on LB medium [44] at 37°C. Chemotactic response of R. leguminosarum to specific carbon compounds was assayed using swarm media containing Vincent's minimal medium [45] 0.15% agarose and a sole carbon source as the potential chemoattractant [16]. The concentration of the carbon source was 1 mM. Carbon sources were purchased from Sigma-Aldrich (ON, Canada). When necessary, Rhizobium strains were cultured in media containing antibiotics at the following concentrations: spectinomycin, 500 μg/ml; streptomycin, 500 μg/ml. E. coli was cultured with the following antibiotic concentrations when required: ampicillin, 100 μg/ml; and spectinomycin, 100 μg/ml. Antibiotics were obtained from Sigma-Aldrich (ON, Canada)

DNA manipulation and sequencing

Restriction enzymes and modifying enzymes were purchased from Life Technologies (ON, Canada) and used according to the manufacturer's instructions.

Plasmid profiles of *Rhizobium* strains were visualized on agarose gels using a modified Eckhardt procedure [46] described by Hynes *et al.* [47] and modified by Hynes and McGregor [48]. Probe labeling, southern blots, and detection procedures were performed using the non-radioactive DIG labeling and detection system as specified per the manufacturer's instructions (Roche Biochemicals, Laval, PQ, Canada)

DNA sequencing of the *mcpG* gene was accomplished using both subcloning and primer walking approaches. Contigs were assembled using DNASIS (Hitachi Software Engineering, CA, USA). Primers for sequencing were designed using the Oligo software application (National Biociences Inc., MN, USA) and synthesized by Operon (CA, USA). Sequencing was performed by the UC DNA Services sequencing facilities (University of Calgary, AB, Canada). Template DNA was prepared according to the facility's specifications. Sequence characterization was performed using DNASIS (Hitachi Software Engineering, CA, USA). Sequence alignments were performed using the BLAST [49] and clustalW [50] programs. To predict membrane spanning regions in the *mcpG* ORF the programs TopPred II and TMHMM were used [30,31].

Mutagenesis of *mcpG* was accomplished using an insertional mutagenesis strategy. A 2.1 Kb *Bcl*I fragment containing the *mcpG* ORF was cloned into the *Bam*HI site of pJQ200mp18. A spectinomycin resistance gene cassette was excised from p1918::Sp using *Sal*I. The cassette was

inserted into the *mcpG* gene through an *Xho*I digestion of pJQ200mp18::*mcpG*. The resultant vector had a 341 bp deletion and a Sp^r cassette inserted within the *mcpG* ORF. The disrupted gene was introduced into VF39SM via double recombination using a protocol described by Quandt and Hynes [34]. Correct gene replacement was confirmed by Southern hybridization.

Nodulation competition experiments

Trapper peas (*Pisum sativum*) were surface sterilized by washing the seeds in 50% Sodium Hypochlorite for 5 minutes, followed by a second wash in 70% ethanol. Following the washes the seeds were rinsed 3 times in sterile distilled water. The seeds were germinated by placing them on water agar plates (12.5 g agar for 1 litre of distilled water) and incubating them at room temperature in the dark for 3 days. Seedlings were transferred to modified magenta jars that were designed to resemble Leonard Jars [24,45]. The peas were grown in a vermiculite substrate.

Once the peas were transferred to the magenta jars the seedlings were co-inoculated with VF39SM and a *mcpG* mutant strain in approximate 9:1, 1:1, and 1:9 ratios. The exact ratios were confirmed by performing viable plate counts on the inoculant cultures. The inoculated peas were then grown for 4 weeks, after which the nodules were harvested, and surface sterilized by washing them in a 20% solution of bleach for 5 min, followed by a 5 min wash in 70% ethanol. The nodules were then rinsed twice in sterile distilled water. Surface sterilized nodules were placed individually in microfuge tubes containing 50 μl of sterile distilled H_2O and crushed using inoculating sticks. 5 μl of the macerate was spotted in duplicate onto TY plates containing Sm, and TY plates containing Sm and Sp to distinguish which strain had formed the nodule, the wild-type or the *mcpG* mutant strain. For each competition experiment set >50 nodules were sampled.

List of abbreviations

aa, amino acid; Ap, ampicillin; Kb, kilobase pairs; ORF, open reading frame; Sm, streptomycin; Gm, gentamicin; Sp, spectinomycin

Authors' Contributions

CKY drafted the manuscript, designed and executed the strategy for making the *mcpG* mutants, did the work on mutant chemotactic phenotypes, completed the sequencing, conducted Southern hybridizations and participated in the nodulation competition experiments. KTC subcloned and sequenced *mcpG*. KLDB conducted the nodulation competition experiments. MFH was involved in subcloning experiments, sequence assembly and bioinformatics, and performed the Eckhardt gels, as well as some of the Southern Blots. As principal investigator in the lab,

MFH also provided intellectual guidance for the entire the study, and was involved in writing the paper.

All authors read and approved the final manuscript.

Acknowledgements

References
1. Armitage JP **Bacterial tactic response.** *Adv Microb Physiol* 1999, **41:**229-289
2. Alexandre G and Zhulin IB **More than one way to sense chemicals.** *J Bacteriol* 2001, **183:**4681-4686
3. Falke JJ, Bass RB, Butler SL, Chervitz SA and Danielson MA **The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases and adaptation enzymes.** *Annu Rev Cell Dev Biol* 1997, **13:**457-512
4. Manson MD, Armitage JP, Hoch JA and Macnab RM **Bacterial locomotion and signal transduction.** *J Bacteriol* 1998, **180:**1009-1022
5. Boyd A, Kendall K and Simon MI **Structure of the serine chemoreceptor in *Escherichia coli*.** *Nature* 1983, **301:**623-626
6. Alam M and Hazelbauer GL **Structural features of methyl-accepting taxis proteins conserved between archaebacteria and eubacteria revealed by antigenic cross-reaction.** *J Bacteriol* 1991, **173:**5837-5842
7. Zhulin IB **The superfamily of chemotaxis transducers: from physiology to genomics and back.** *Adv Microb Physiol* 2001, **45:**157-198
8. Currier WW and Strobel GA **Chemotaxis of *Rhizobium* spp. to a glycoprotein produced by birdsfoot trefoil roots.** *Science* 1977, **196:**434-435
9. Ames P and Bergman K **Competitive advantage provided by bacterial motility in the formation of nodules by *Rhizobium meliloti*.** *J Bacteriol* 1981, **148:**728-729
10. Gulash M, Ames P, Larosiliere RC and Bergman K **Rhizobia are attracted to localized sites on legume roots.** *Appl Environ Microbiol* 1984, **48:**149-152
11. Caetano-Anollés G, Wall LG, Micheli ATD, Macchi EM, Bauer WD and Favelukes G **Role of motility and chemotaxis in efficiency of nodulation by *Rhizobium meliloti*.** *Plant Physiol* 1988, **86:**1228-1235
12. Caetano-Anollés G, Wrobel-Boerner E and Bauer WD **Growth and movement of spot inoculated *Rhizobium meliloti* on the root surface of alfalfa.** *Plant Physiol* 1992, **98:**1181-1189
13. Munoz Aguilar JM, Ashby AM, Richards AJM, Loake GJ, Watson MD and Shaw CH **Chemotaxis of *Rhizobium leguminosarum* biovar *phaseoli* towards flavonoid inducers of symbiotic nodulation genes.** *J Gen Microbiol* 1988, **134:**2741-2746
14. Bauer WD and Caetano-Anollés G **Chemotaxis, induced gene expression and competitiveness in the rhizosphere.** *Plant Soil* 1990, **129:**45-52
15. Dharmatilake AJ and Bauer WD **Chemotaxis of *Rhizobium meliloti* towards nodulation gene-inducing compounds from alfalfa roots.** *Appl Environ Microbiol* 1992, **58:**1153-1158
16. Yost CK, Rochepeau P and Hynes MF ***Rhizobium leguminosarum* contains a group of genes that appear to code for methyl-accepting chemotaxis proteins.** *Microbiology* 1998, **144:**1945-1956
17. Van Rhijn P and Vanderleyden J **The *Rhizobium*-plant symbiosis.** *Microbiol Rev* 1995, **59:**124-142
18. Baldani JL, Weaver RW, Hynes MF and Eardly BD **Utilization of carbon substrates, electrophorectic enzyme patterns and symbiotic performance of plasmid cured rhizobia.** *Appl Environ Microbiol* 1992, **58:**2308-2314
19. Brom S, García-de los Santos A, Stepkowsky T, Flores M, Dávila G, Romero D and Palacios R **Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance.** *J Bacteriol* 1992, **174:**5183-5189
20. Moënne-Loccoz and Weaver RW **Plasmids influence growth of rhizobia in the rhizosphere of clover.** *Soil Biol Biochem* 1995, **27:**1001-1004
21. García-de los Santos A, Brom S and Romero D ***Rhizobium* plasmids in bacteria-legume interactions.** *World J Microbiol & Biotech* 1996, **12:**119-125
22. Mercado-Blanco J and Toro N **Plasmids in Rhizobia: the role of nonsymbiotic plasmids.** *Mol Plant-Microbe Interact* 1996, **9:**535-545
23. Rochepeau P, Selinger LB and Hynes MF **Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM.** *Mol Gen Genet* 1997, **256:**387-396
24. Oresnik IJ, Pacarynuk LA, O'Brien SAP, Yost CK and Hynes MF **Plasmid-encoded catabolic loci in *Rhizobium leguminosarum* bv. *trifolii*. Evidence for a plant-inducible rhamnose locus involved in competition for nodulation.** *Mol Plant Microbe Interact* 1998, **11:**1175-1185
25. Brom S, García-de los Santos A, Cervantes L, Palacios R and Romero D **In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons.** *Plasmid* 2000, **44:**34-43
26. Brito B, Palacios J-M, Ruiz-Argüeso T and Imperial J **Identification of a gene for a chemoreceptor of the methyl-accepting type in the symbiotic plasmid of *Rhizobium leguminosarum* bv. *viciae* UPM791.** *Biochim Biophys Acta* 1996, **1308:**7-11
27. Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A and Perret X **Molecular basis of symbiosis between *Rhizobium* and legumes.** *Nature* 1997, **387:**394-401
28. Galibert F, Finan TM, Long SR, Pühler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, becker A, Boistard P, Bothe G, Bountry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetell D, Purnell B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh KC and Batut J **The composite genome of the legume symbiont *Sinorhizobium meliloti*.** *Science* 2001, **293:**668-72
29. Carlos MG and von Heigne G **TopPred II: an improved software for membrane protein structure predictions.** *Comput Appli Biosci* 1994, **10:**685-686
30. Krogh A, Larsson B, von Heijne G and Sonnhammer EL **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305:**567-580
31. Hosie AHF, Allaway D and Poole PS **A monocarboxylate permease of *Rhizobium leguminosarum* is the first member of a new subfamily of transporters.** *J Bacteriol* 2002, **184:**5436-5448
32. Aravind L and Ponting CP **The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signaling proteins.** *Fems Microbiol Lett* 1999, **176:**111-116
33. Wood DW, Setubal JC, Kaul R, Monks D, Chen I, Wood GE, Chen Y, Woo L, Kitajima JP, Okura VK, Almeida NF Jr, Zhou Y, Bovee D Sr, Chapman P, Clendenning J, Deatherage G, Gillet W, Grant C, Guenthner D, Kutyavin T, Levy R, Li M, McClelland E, Palmieri A, Raymond C, Rouse G, Saenphimmachak C, Wu Z, Gordon D, Eisen JA, Paulsen I, Karp P, Romero P, Zhang S, Yoo H, Tao Y, Biddle P, Jung M, Krespan W, Perry M, Gordon-Kamm B, Liao L, Kim S, Hendrick C, Zhao Z, Dolan M, Tingey SV, Tomb J, Gordon MP, Olson MV and Nester EW **The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58.** *Science* 2001, **294:**2317-2323
34. Quandt J and Hynes MF **Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria.** *Gene* 1993, **127:**15-21
35. Yost CK **Characterization of *Rhizobium leguminosarum* genes homologous to chemotaxis chemoreceptors.** *Ph.D.Thesis, University of Calgary, Calgary, Alberta* 1998
36. Parke D, Rivelli M and Ornston LN **Chemotaxis to aromatic and hydroaromatic acids: comparisons of *Bradyrhizobium japonicum and Rhizobium trifolii*.** *J Bacteriol* 1985, **163:**417-422
37. Parke D and Ornston LN **Nutritional diversity of *Rhizobiaceae* revealed by auxanography.** *J Gen Microbiol* 1984, **130:**1743-1750
38. Stowers MD **Carbon metabolism in *Rhizobium* species.** *Annu Rev Microbiol* 1985, **39:**89-108
39. Alexandre G, Greer SE and Zhulin IB **Energy taxis is the dominant behavior in *Azospirillum brasilense*.** *J Bacteriol* 2000, **182:**6042-6048

40. Gauden DE and Armitage JP **Electron transport-dependent taxis in *Rhodobacter sphaeroides*.** *J Bacteriol* 1995, **177:**5853-5859

41. Oresnik IJ, Twelker S and Hynes MF **Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins.** *Appl Environ Microbiol* 1999, **65:**2833-2840

42. Venter AP, Twelker S, Oresnik IJ and Hynes MF **Analysis of the genetic region encoding a novel rhizobiocin from *Rhizobium leguminosarum* bv. *viciae* strain 306.** *Can J Microbiol* 2001, **47:**495-502

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

44. Sambrook J, Fitsch EF and Maniatis T **Molecular Cloning: A Laboratory Manual.** *Cold Spring Harbor, Cold Spring Harbor Press* 1989,

45. Vincent JM **A Manual for the Practical Study of Root-nodule Bacteria (IBP handbook no. 15).** *Oxford, Blackwell Scientific Publications* 1970,

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

47. Hynes MF, Simon R and Pühler A **The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate pATC58.** *Plasmid* 1985, **13:**99-105

48. Hynes MF and McGregor NF **Two plasmids other than the nodulation plasimd are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*.** *Mol Microbiol* 1990, **4:**567-574

49. Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W and Lipman D **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25:**3389-3402

50. Thompson JD, Higgins DG and Gibson TJ **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22:**4673-4680

51. Simon R, Priefer U and Pühler A **A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria.** *Bio/Technology* 1983, **1:**784-791

52. Priefer UB **Genes involved in lipopolysaccharide production and symbiosis are clustered on the chromosome of *Rhizobium leguminosarum* biovar *viciae* VF39.** *J Bacteriol* 1989, **171:**6161-6168

53. Hynes MF, Quandt J, O'Connell MP and Pühler A **Direct selection for curing and deletion of *Rhizobium* plasmids using transposons carrying the *Bacillus subtilis sacB* gene.** *Gene* 1989, **78:**111-120

54. Lamb JW, Hombrecher G and Johnston AWB **Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli*.** *Mol Gen Genet* 1982, **186:**449-452

55. Finan TM, Wood JM and Jordan DC **Succinate transport in *Rhizobium leguminosarum*.** *J Bacteriol* 1981, **148:**193-202

56. Poole PS, Schofield NA, Reid CJ, Drew EM and Walshaw DL **Identification of chromosomal genes located downstream of *dctD* that affect the requirement for calcium and the lipopolysaccharide layer of *Rhizobium leguminosarum*.** *Microbiol* 1994, **140:**2797-2809

57. Moënne-Loccoz Y, Sen D, Krause ES and Weaver RW **Plasmid profiles of rhizobia used in inoculants and isolated from clover fields.** *Agron J* 1994, **86:**117-121

58. Fellay R, Frey J and Krisch H **Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria.** *Gene* 1987, **52:**147-154

59. Schweizer HP **Two plasmids, X1918 and Z1918 for easy recovery of the *xylE* and *lacZ* reporter genes.** *Gene* 1918, **134:**89-91

60. Rice MS and Dahlquist FW **Sites of deamidation and methylation in Tsr, a bacterial chemotaxis sensory transducer.** *J Biol Chem* 1991, **266:**9746-9753

61. Kim KK, Yokota H and Kim S-H **Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor.** *Nature* 1999, **400:**787-792