

Marquette University

e-Publications@Marquette

---

Biological Sciences Faculty Research and  
Publications

Biological Sciences, Department of

---

1-2003

## BacS: An Abundant Bacteroid Protein in *Rhizobium etli* Whose Expression Ex Planta Requires *nifA*

Olivia Jeanne Jahn  
Marquette University

Guillermo Davila  
Centro de Investigación sobre Fijación de Nitrógeno

David Romero  
Centro de Investigación sobre Fijación de Nitrógeno

K. Dale Noel  
Marquette University, dale.noel@marquette.edu

Follow this and additional works at: [https://epublications.marquette.edu/bio\\_fac](https://epublications.marquette.edu/bio_fac)



Part of the [Biology Commons](#)

---

### Recommended Citation

Jahn, Olivia Jeanne; Davila, Guillermo; Romero, David; and Noel, K. Dale, "BacS: An Abundant Bacteroid Protein in *Rhizobium etli* Whose Expression Ex Planta Requires *nifA*" (2003). *Biological Sciences Faculty Research and Publications*. 63.

[https://epublications.marquette.edu/bio\\_fac/63](https://epublications.marquette.edu/bio_fac/63)

# BacS: An Abundant Bacteroid Protein in *Rhizobium etli* Whose Expression Ex Planta Requires *nifA*

Olivia J. Jahn,<sup>1</sup> Guillermo Davila,<sup>2</sup> David Romero,<sup>3</sup> and K. Dale Noel<sup>1</sup>

<sup>1</sup>Department of Biology, Marquette University, Milwaukee, WI 53233, U.S.A. <sup>2</sup>Programas de Evolución Molecular and

<sup>3</sup>Genética Molecular de Plásmidos Bacterianos, Centro de Investigación sobre Fijación de Nitrógeno, Cuernavaca, Morelos, Mexico

Submitted 14 February 2002. Accepted 20 September 2002.

***Rhizobium etli* CFN42 bacteroids from bean nodules possessed an abundant 16-kDa protein (BacS) that was found in the membrane pellet after cell disruption. This protein was not detected in bacteria cultured in tryptone-yeast extract. In minimal media, it was produced at low oxygen concentration but not in a mutant whose *nifA* was disrupted. N-terminal sequencing of the protein led to isolation of a *bacS* DNA fragment. DNA hybridization and nucleotide sequencing revealed three copies of the *bacS* gene, all residing on the main symbiotic plasmid of strain CFN42. A stretch of 304 nucleotides, exactly conserved upstream of all three *bacS* open reading frames, had very close matches with the NifA and sigma 54 consensus binding sequences. The only *bacS* homology in the genetic sequence databases was to three hypothetical proteins of unknown function, all from rhizobial species. Mutation and genetic complementation indicated that each of the *bacS* genes gives rise to a BacS polypeptide. Mutants disrupted or deleted in all three genes did not produce the BacS polypeptide but were Nod<sup>+</sup> and Fix<sup>+</sup> on *Phaseolus vulgaris*.**

Many leguminous plants obtain nitrogen via symbiosis in root and stem nodules with bacteria known as rhizobia, belonging to several branches of the  $\alpha$  and  $\beta$  subdivisions of the Proteobacteria (Moulin et al. 2001). In infected nodule cells, the bacteria are present as bacteroids, whose composition and structure differ from those of free-living bacteria (Franssen et al. 1992; Planque et al. 1979). Bacteroids fix nitrogen that is then assimilated by the host. The *nif* and *fix* genes involved in nitrogen fixation are induced as part of the conversion of the bacteria into bacteroids. The NifA protein activates the expression of *nifHDKE*, *fixABCX*, and *nifN* (genes for nitrogenase, electron transport to nitrogenase, and ferredoxin, respectively) in response to low oxygen concentration (Fischer 1994; Merrick 1992).

In *Rhizobium etli* CFN42, one of six plasmids carries most of the genes required to form nodules and fix nitrogen on bean roots (Brom et al. 1992; Cevallos et al. 1989; Palacios et al. 1985). This symbiotic (Sym) plasmid is 390 kb in length, with 10 families of internally reiterated DNA sequences of two to three elements each (Girard et al. 1991). In addition to well-established *nod*, *nif*, and *fix* genes, this and other symbiotic plasmids carry other genes that are expressed during symbiosis but

whose functions are unknown (Girard et al. 1996; Perret et al. 1999).

Work initiated to study bacteroid membrane proteins revealed an abundant bacteroid protein of *R. etli* CFN42, designated BacS. Its synthesis ex planta was dependent on microaerobic conditions and *nifA*. Three similar open reading frames (ORFs) on the Sym plasmid encoded BacS polypeptides. The possible role of this protein in symbiosis was tested by isolation of mutants in which all copies of *bacS* were deleted or disrupted.

## RESULTS

### Abundant bacteroid protein lacking in aerobic cultures.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) revealed an abundant, approximately 16-kDa protein in bacteroids of *R. etli* strain CE3 (Fig. 1). It was designated BacS (bacteroid protein of 16 kDa). This protein was not detected from aerobically grown CE3 bacteria by either Coomassie staining or immunoblot analysis, even when lanes were overloaded with the contents of aerobically grown cells (Fig. 1, lanes 5 and 6). After rupture of bacteroids by sonication, BacS was found in the cell fraction that pelleted after centrifugation at 15,000  $\times$  g for 60 min (Fig. 2). The protein was not released from the pelleted fraction with 2 M sodium chloride or 1 M magnesium chloride (not shown).

### Induction ex planta at low O<sub>2</sub> concentration, requiring *nifA*.

As noted above, strain CE3 grown under ambient atmospheric gas composition did not express detectable levels of BacS. However, strain CE3 grown with 1% O<sub>2</sub> (5% ambient air, 95% purified nitrogen) or no O<sub>2</sub> in the gas space above minimal liquid medium did express the protein, as evidenced by a band detected with antiserum against BacS migrating at the appropriate rate on SDS-PAGE (Fig. 3, lanes 4 to 7). Growth in nodule extract (Noel et al. 1988), root extract (Duelli 1999), limiting FeCl<sub>3</sub>, or in medium buffered at pH 5 or 7 did not result in detectable levels of the protein unless the cultures were also grown microaerobically (not shown). These additional conditions did not significantly affect the microaerobically induced level of expression, nor did substituting glucose for succinate as the major carbon source (not shown). CE3 grown microaerobically in tryptone-yeast (TY) extract rather than in minimal media did not induce detectable levels of the protein (not shown). Addition of nitrate to cultures also allowed expression of BacS under conditions of low oxygen concentration. The nitrate stimulated growth of CE3 when no oxygen was added to the headspace of the stoppered vessels after the nitrogen purge but led to apparently lower production

Corresponding author: K. D. Noel, Department of Biology, Marquette University, PO Box 1881, Milwaukee, WI 53201 U.S.A.; Telephone: 414-288-1475; Fax: 414-288-7357; E-mail: dale.noel@mu.edu

Nucleotide sequences data reported are available in the GenBank database under accession numbers AF481531 and AF481532.

of BacS relative to total proteins (Fig. 3, lanes 4 to 7). The level of expression of the protein ex planta was always much lower than in the bacteroids (Fig. 3, lane 8).

The *nifA* gene was required for production of BacS under these conditions. Under conditions that elicited synthesis of the protein in strain CE3 (Fig. 4, lane 1), strain CFNX247 (*nifA*<sup>-</sup>, Table 1) did not exhibit detectable levels (Fig. 4, lane 2).

### Cloning of the 5' portion of *bacS* and its use to reveal multiple *bacS* genes on pSym.

Twenty-nine amino acid (aa) residues were identified by N-terminal sequencing of BacS excised from an SDS-PAGE blot: T, K, I, S, E, R, S, F, E, T, I, E, N, P, G, S, X, D, L, K, V, P, D, Q, F, G, A, S, and R. The purity of the protein was such that, in each of the first five cycles of sequencing, the highest contaminating amino acid was present at 20% or less of the one indicated above. By use of degenerate oligonucleotide primers based on this sequence, a polymerase chain reaction (PCR) product whose DNA sequence matched the N-terminal protein sequence was cloned (plasmid pOJ1, Table 1).

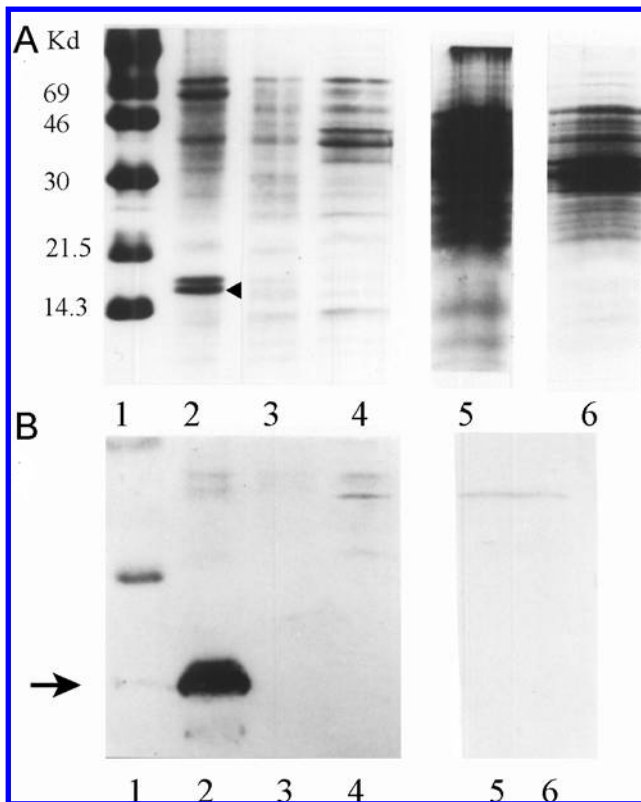
CE3 genomic DNA digested with *Bam*HI, *Eco*RI, *Eco*RV, and *Pst*I was analyzed by Southern hybridization using pOJ1 as probe (Fig. 5). With one exception, restriction digests with one or two of these enzymes resulted in two or three fragments hybridizing with the probe. Only the *Bam*HI and *Eco*RI double

digest gave a single hybridizing band (not shown). None of the restriction enzymes cut within the DNA sequence of the hybridization probe; therefore, the multiple hybridization bands indicated potentially three copies of the gene, with the *Bam*HI and *Eco*RI restriction sites conserved in all three genetic locations.

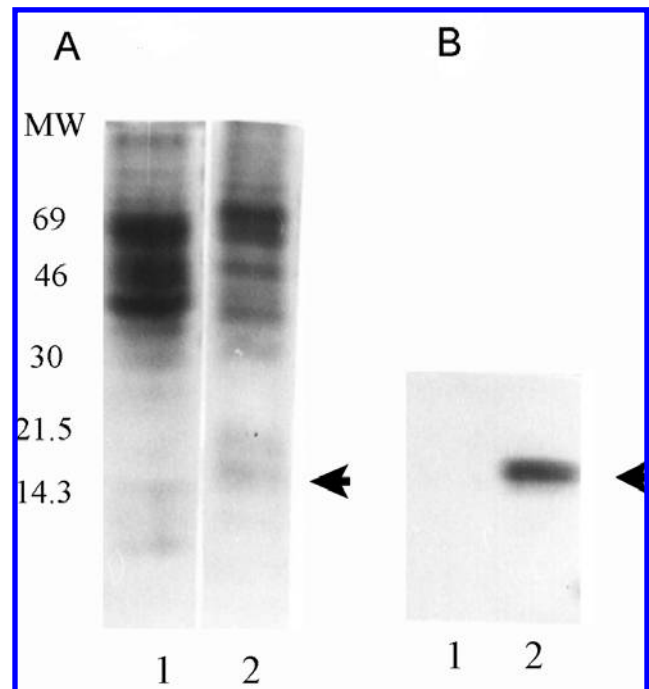
A strain cured of the pSym (CFNX89) did not produce BacS under inducing conditions ex planta (Fig. 4, lane 4) and lacked all of the above-described *bacS*-hybridizing restriction fragments (data not shown). The nucleotide sequence of the N-terminal region of BacS was compared with data obtained from total nucleotide sequencing of the CE3 symbiotic plasmid (V. González, P. Bustos, M. A. Ramírez-Romero, A. Medrano-Soto, H. Salgado, I. Hernández-González, J. C. Hernández-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodríguez, V. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Dávila, unpublished data). The comparison indicated high homology with three *Bam*HI fragments of the pSym (B03, B27, and B80 according to the nomenclature of Girard et al. 1991). The published sizes of these *Bam*HI fragments (Girard et al. 1991) matched the results of Southern hybridization with pOJ1 described above. Independently of the symbiotic plasmid genome sequencing project, cloned DNA fragments that included two *bacS* ORFs were sequenced.

### The *bacS* ORFs in pSym *Bam*HI fragments B03, B27, and B80.

The nucleotide sequences of the three *bacS* ORFs and 382 nucleotides (nt) upstream are shown in Figure 6. The three sequences are very highly conserved (Fig. 6), but diverge immediately upstream and downstream of the sequence shown. All three DNA sequences have the same potential NifA upstream activator sequences and a consensus  $\sigma^{54}$  promoter sequence (Fig. 6). The ORFs determined within the B03, B27, and B80 DNA (*bacS*<sub>03/27/80</sub>) would result in a 151-aa protein (Fig. 6) with



**Fig. 1.** BacS is an abundant 16-kDa bacteroid protein. Wild-type CE3 bacteroids or bacteria grown ex planta in rich medium (TY) or minimal medium (YGM) in ambient atmosphere were extracted with sodium dodecyl sulfate (SDS) and analyzed by SDS-polyacrylamide gel electrophoresis. **A**, Gels were stained with Coomassie blue. **B**, Immunoblots of identically loaded lanes were reacted with polyclonal antibodies to BacS. **A** and **B**, Molecular weight markers (lane 1), bacteroids (lane 2), bacteria grown in TY medium (lanes 3 and 5), or grown in YGM buffered at pH 7 (lanes 4 and 6). Lanes 5 and 6 were overloaded to emphasize the lack of immunologically detectable expression of *bacS*. The arrows indicate the location of BacS as determined by immunoblots of gels in which BacS was better resolved from the abundant protein migrating just above it (data not shown).



**Fig. 2.** BacS associates with the insoluble fraction of bacteroids. Envelope (lane 2) and soluble (lane 1) cell fractions of CE3 bacteroids (prepared as described in Methods) were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. **A**, Gel stained with Coomassie blue. **B**, Western blot of the lower portion of a gel run identically to that of **A** and reacted with polyclonal antisera to BacS. The arrows indicate the position of BacS.

a molecular mass of 16.6 kDa. There were no obvious subcellular targeting sequences for the protein upstream of the N terminus of the mature protein. Computer analysis (Brendel et al. 1992) also did not reveal appropriate sequences long enough to span the entire length of either membrane lipid bilayer.

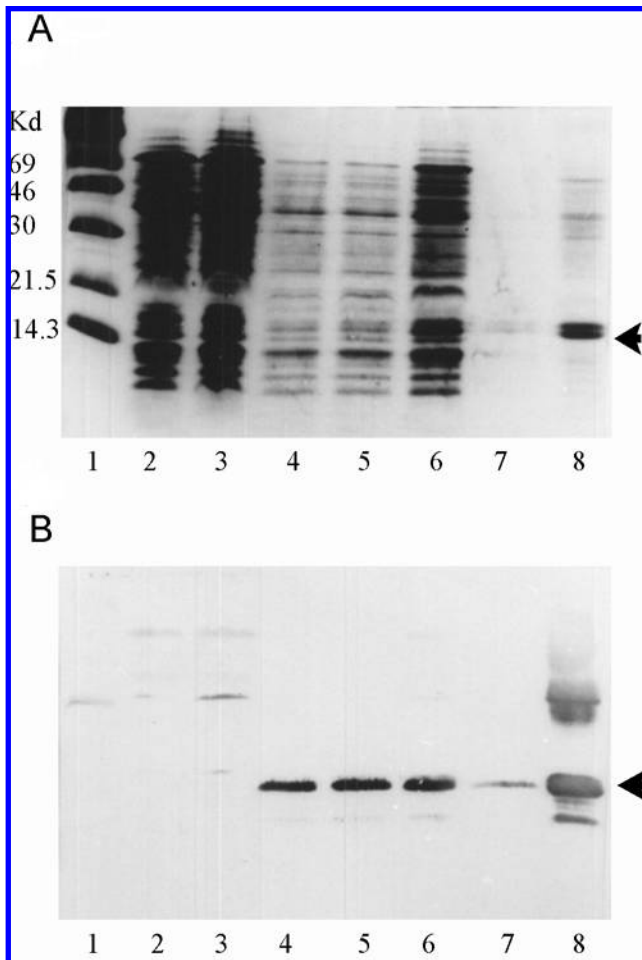
A search of the data bases showed significant homology with hypothetical proteins of unknown function (Fig. 7) encoded by three ORFs: Y4wP from the symbiotic plasmid of *Sinorhizobium fredii* NGR234 (Freiberg et al. 1997), Np107027.1 from *Mesorhizobium loti* (Kaneko et al. 2000), and SMc00777 AA from *S. meliloti* (Galibert et al. 2001). The hypothetical protein encoded by the *y4wP* ORF has a length of 135 aa residues and an estimated molecular mass of 14.9 kDa. The Y4wP protein had 50 identical residues and 87 similar residues with BacS<sub>03</sub> (Fig. 7). The hypothetical protein from *M. loti* has 152 aa residues, similar to the 151 residues in BacS. The homology spans almost the entire length of both proteins with 47% identity and 67% similarity (Fig. 7). The hypothetical protein from *S. meliloti* has 148 aa residues. The homology spans residues 3 to 146 of SMc00777AA, with 35% identity and 53% similarity (Fig. 7). The predicted amino acid

sequences of BacS and these three homologs are aligned in Figure 7.

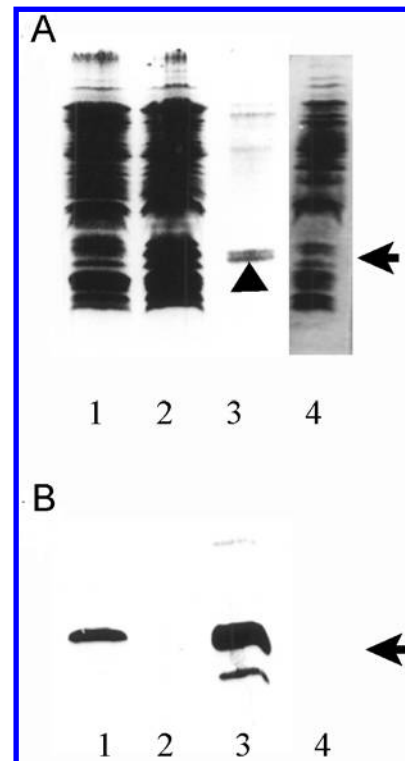
In some experiments, the BacS polyclonal antibodies appeared to react weakly with *Phaseolus vulgaris* leghemoglobin on Western blots of the soluble nodule plant proteins (data not shown). Prompted by this observation, leghemoglobin amino acid sequences were compared with the sequences of BacS and its homologs in the other rhizobial species (Fig. 7). Interestingly, without gapping, there was a conservation in the spacing of five (PvLb) or six (PsLb) of the first seven residues that are conserved in all the rhizobial BacS homologs (Fig. 7). One of these conserved residues was a histidine (residue 87 in BacS) (Fig. 7) which, in leghemoglobin, is one of two conserved histidines crucial in coordinating the iron of the heme and directing its binding to oxygen. At the position of the other leghemoglobin histidine, BacS also has a histidine (residue 118 in BacS) (Fig. 7). At this position, the three BacS homologs have a glutamine.

#### Phenotypes of BacS<sup>-</sup> mutants.

*R. etli* CFNX250 (D. Romero, A. Corvera, A. Geniaux, S. Brom, E. Valencia-Morales, C. Rodríguez, and B. Valderrama, unpublished) is Nod<sup>+</sup>Fix<sup>+</sup> but has a pSym deletion (resulting in pGM1, Table 1) that eliminates two-thirds of the pSym of strains CFN42 and CE3, including two of the *bacS* ORFs. The remaining ORF (in *Bam*HI fragment B03) allows the production of BacS in normal amounts (Fig. 8, lane 1). This copy of *bacS* was mutated in strain CFNX250 by insertion of DNA that confers tetracycline resistance to generate *bacS* isolates. These isolates did not produce detectable levels of BacS ex



**Fig. 3.** BacS is produced ex planta under microaerobic conditions. Wild-type CE3 grown under aerobic, 1% O<sub>2</sub>, or anaerobic conditions with or without nitrate was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. **A**, Gel stained with Coomassie blue. **B**, Immunoblot of identically loaded lanes was reacted with polyclonal antibodies to BacS. **A** and **B**, Bacteria were grown in YGM medium (buffered at pH 7) with nitrate (lanes 2, 4, and 6) or without nitrate (lanes 3, 5, and 7). The gas above the liquid phases of the cultures was ambient air (lanes 2 and 3), nitrogen supplemented with 5% air (lanes 4 and 5), or only nitrogen (lanes 6 and 7). Arrows indicate the position of BacS. Lane 1, molecular weight markers; lane 8, samples of a bacteroid extract.



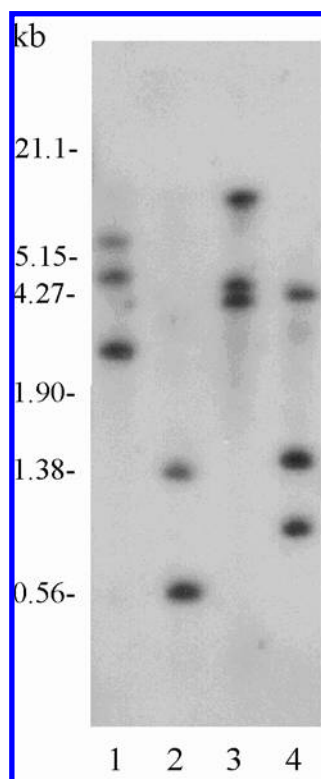
**Fig. 4.** *nifA* and pSym are required for production of BacS. Strains CE3 (wild type), CFNX247 (*nifA*<sup>-</sup>), and CFNX89 (pSym<sup>-</sup>) were grown under 1% oxygen conditions and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. **A**, Portions of two gels stained with Coomassie blue. **B**, Portions of immunoblots of lanes that had been loaded identically to those in **A** were reacted with polyclonal antibodies to BacS. **A** and **B**, CE3 (lane 1), CFNX247 (lane 2), or CFNX89 (lane 4) cells were grown at 1% O<sub>2</sub> in YSM buffered at pH 7. CE3 bacteroids were processed for lane 3. An arrow indicates the position of BacS.

planta or in bacteroids (e.g., strain CE461, Fig. 8). However, all mutants nodulated and fixed nitrogen on *P. vulgaris*, in a manner that was not observed to differ from wild-type strains CFNX250 or CE3.

When wild-type *bacS*<sub>27</sub> (on plasmid pOJB27F) or *bacS*<sub>80</sub> (on plasmid pOJB80F) were transferred into these *bacS* mutants, the transconjugant strains produced BacS ex planta under inducing conditions (not shown). These complementation results indicated that all three of the *bacS* genes can be expressed under these conditions.

## DISCUSSION

BacS was highly expressed in CE3 bacteroids but was not detectably expressed in the free-living form unless the bacteria were cultivated under microaerobic conditions. Production of BacS ex planta under low oxygen does not occur in a *nifA* mutant strain. Upstream of all three reiterated *bacS* genes, three matches with the consensus upstream activator sequence (UAS) for NifA (Ditta 1989) are present (Fig 6, underlined sequences beginning at nucleotides 9, 43, and 72). A sequence that corresponds with the consensus  $\sigma^{54}$ -dependent promoter occurs approximately 100 nt downstream of the last UAS (Fig. 6, nucleotide 183). Although these observations provide strong evidence of regulation via NifA and microaerobic conditions in minimal media, no expression was detected in microaerobic TY cultures and the level of expression in bacteroids was much higher than in bacteria under any condition tested ex planta. Therefore, there must be conditions in addition to microaerobiosis that affect the expression of this protein. There may be another regulator in addition to NifA, or NifA activity may be limited under the ex planta conditions tested.



**Fig. 5.** CE3 genomic DNA contains multiple copies of *bacS*. CE3 genomic DNA was digested with restriction enzymes *Bam*HI (lane 1), *Eco*RI (lane 2), *Eco*RV (lane 3), and *Pst*I (lane 4). After electrophoresis, DNA fragments were blotted and probed with the <sup>32</sup>P-labeled pOJ1. The migration of standard DNA is indicated by the positions labeled at the left of the blot.

The BacS protein was found among the nonsoluble proteins when centrifuged at 15,000 to 150,000 × *g*. After sonication, much of it appears to be lost with the broken cell debris. However, its structure does not have features consistent with integral inner or outer membrane proteins and, in some experiments, it was found also in the soluble fraction. BacS from bacteroids is stained by silver after meta-periodate treatment, which may indicate that it is a glycoprotein. However, it does not migrate at a higher molecular weight in SDS-PAGE than expected, as is common for most glycoproteins.

The three ORFs for *bacS* are located in and near the *nod* and *nif* region of pSym. Also, all three are located within the type IV amplicon, one of the defined regions of DNA that is prone to the formation of frequent tandem amplification (Romero et al. 1995). All the known amplicons in *R. etli* are concentrated within the region of pSym in which most of the symbiotic genes are located (Romero et al. 1995). Further, a 135-kb region sufficient to reestablish nodulation and nitrogen fixation in otherwise *pSym*<sup>-</sup> *R. etli* (D. Romero, unpublished) contains one of the *bacS* genes.

Homologs to *bacS* are present in all rhizobial genomes sequenced thus far. The *bacS* homolog ORF, *y4wP*, in *S. fredii* NGR234 is directly upstream of *nifHDK*, downstream of *nifS* and *nifW*, and has a putative NifA-dependent upstream activator and a putative  $\sigma^{54}$ -dependent promoter. In *M. loti*, the homologous ORF is just outside the symbiosis island on the chromosome. Therefore, in both of these species, as in *R. etli*, the *bacS* homolog is found near symbiotically significant genes. In both *M. loti* and *S. meliloti*, a His-GTG-tRNA occurs directly downstream of the ORFs. As more sequence data becomes available, the relative location of the *bacS* genes may provide more clues to the BacS function.

One possible clue to the function of BacS is that its sequence aligns in a certain way with leghemoglobins (Fig. 7). Conservation of the primary structure in globins across phylogeny is low (Moens et al. 1996). The sequences from diverse organisms can be aligned only by the relative positioning of very few residues, particularly the two conserved histidine–glutamine residues. Computer analysis of the secondary structure of BacS indicates that these two residues are located in two different helices which could correspond to the two helices where the histidine–glutamine residues from the leghemoglobin model are located. Conservation and alignment of the first six of seven identical residues between the BacS rhizobial homologs and leghemoglobin, including the exact alignment of the histidine–glutamine residues that are crucial in hemoglobin function, along with the possible similarities in secondary structure, indicate that BacS may be related to hemoglobins.

In a paper by Perret and associates (1999), *y4wP* transcription was not detected under conditions of induction with daidzein or in bacteroids from *Vigna unguiculata* determinate and *Cajanus cajan* indeterminate nodules. However, they did not test bacteroids from bean nodules or bacteria grown ex planta under conditions of low oxygen, the conditions in which expression of *bacS* was detected in *R. etli*.

*nifA*-dependent expression ex planta, the location of the three *bacS* ORFs in and near the *nod* and *nif* region on pSym, and the bacteroids' high level of *bacS* expression all suggest an importance of this protein to *R. etli* during symbiosis. Therefore, the lack of an Nif<sup>-</sup> or Fix<sup>-</sup> phenotype in BacS<sup>-</sup> mutants is surprising. It may be that absence of BacS can be compensated for by another protein, similar to the complementing proteins NodO and NodE in *R. leguminosarum* bv. *viciae*, in which either protein is sufficient to promote normal infection thread growth despite having disparate biochemical activities (Walker and Downie 2000). Alternatively, BacS may be required under a set of conditions or a different host not yet tested.



## MATERIALS AND METHODS

### Bacterial strains and growth media.

*R. etli* strains ultimately were derived from bean nodule isolate *R. etli* CFN42 (Quinto et al. 1982), with strain CE3 (Table

1) being the wild-type control. Liquid cultures were grown at 30°C in TY extract + CaCl<sub>2</sub> (Noel et al. 1984), minimal medium YGM (Tao et al. 1992), or minimal medium YSM which was identical to YGM except that 0.4% succinate was substituted for the glucose. Solidified medium contained 1.5% Bacto

1	AATTCCTT <u>GTCCCCAGCGCAACCGT</u>
27	CGACATTTCTCGCAAAT <u>GTTCGGGCATCCGCCACTTTTGT</u> CAGGTT <u>GTTCGTGATTGGAACGTTTT</u>
92	GTTTCGCCGGTCTCTCCGTCAAATACACATCAACCTGCCCTCAACCCGCCATTGATAAAAGAAAA
157	TCATCTTCTTGGCGCGTTCGGAGCAGTT <u>GGCATGGTGTTCCT</u> GATCCCTGGTTACAAGCGGCAGA
222	AAGCAGTCTGCATGATGCCGGTTCATGAAACGAACGGAACGGAGGCCACCTGACCGTTCGTCATC
287	<u>GGATCCGCGTTCTACCAAGCGGCCATGCATCCGCCGGCGGAAGGGTCAATCCCGATGTCACAGCT</u>
	A <sub>80</sub> TAG <sub>80</sub> G <sub>80</sub>
352	CGCAACCAACTCAAAGAGAAGGAAAATGCA <b>ATG</b> ACCAAGATTTCCGAAAGATC <b>CTTT</b> GAAACGAT
1	M T K I S E R <b>S</b> F E T I
	<b>S</b> <sub>80</sub>
417	CGAAAACCCGGGGTTCATCG <b>TC</b> GCTCAAGGTGCCAGATCAGTTCGGCGCATC <b>TT</b> GTTGAAAAGGGGA
13	E N P G S S <b>S</b> L K V P D Q F G A <b>S</b> V E K G
	<b>P</b> <sub>80</sub> <b>S</b> <sub>80</sub>
482	ACAAG <b>A</b> AGGTGACAGAGGCTTTTTTGA <b>AACT</b> TGCGTCCGGCGCTGAAGCGACACAGAAAATGCTG
34	N K <b>K</b> V T E A F L K <b>L</b> A S G A E A T Q K M L
	<b>Q</b> <sub>80</sub> <b>F</b> <sub>27/80</sub>
547	CCTCCGATCCTCGAAACGACAAGTCTATTCCGGCAACGAATTGTGGTGGAAAGACGATCGCTGCACT
56	P P I L E T T S L F G N E L W W K T I A A L
612	GCAG <b>G</b> CCGACGCCGAGGCC <b>AG</b> CTTCTCACATTTGCAAGCTTTGCTGGGCGCG <b>AAATTC</b> GCCGTTCGC
78	<b>Q</b> A D A E A <b>S</b> F S H L Q A L L G A N S P S
	<b>Q</b> <sub>80</sub> <b>G</b> <sub>80</sub>
677	AGATCCTCGAACAGCAGTCGACCTTTTTCCGCAAGCG <b>CG</b> TTGAGACAAGTTTGCAGCAC <b>G</b> CCAAG
99	Q I L E Q Q S T F F R K <b>R</b> V E T S L Q H <b>A</b> K
	<b>R</b> <sub>27</sub> <b>T</b> <sub>27</sub>
742	GAAGTCCGGG <b>T</b> GCTCTCAAGCAGGGCGGTGGAGGAAATCTCAAAGCCGGTCAAGGATGCTTTTGA
121	E V R <b>V</b> L S S R A V E E I S K P V K D A F D
	<b>E</b> <sub>80</sub>
807	CAAGGT <b>G</b> CTGAC <b>G</b> AC <b>CT</b> CAAG <b>G</b> CGAC <b>G</b> TAA <b>AACTG</b> AGGAGACCCTAGAAATACCGCGCGGTGAT
143	K <b>V</b> L T <b>D</b> L K A T
	<b>V</b> <sub>80</sub> <b>N</b> <sub>80</sub> <b>P</b> <sub>80</sub> <b>K</b> <sub>80</sub>

**Fig. 6.** Nucleotide and protein sequences are exactly conserved in all three *bacS* loci in *Rhizobium etli* CE3, except where noted. The amino acid sequence shown below the open reading frame (ORF) nucleotide sequence is that of the predicted BacS<sub>03</sub> protein. The probable translation initiation codon (nucleotide 382) and two nonsense codons (nucleotides 835 and 841) that terminate the ORF are shown in bold. A potential σ<sup>54</sup> promoter sequence (nucleotide 183) and three NifA upstream activator sequences (nucleotides 9, 43, and 72) are underlined. The double underlines indicate the highly conserved regions in the promoter and upstream activator consensus sequences. The *Bam*HI (nucleotide 287), *Sma*I (nucleotide 423), and *Eco*RI (nucleotide 663) restriction sites also are underlined. The amino acid residues in bold indicate differences between *bacS*<sub>03</sub>, *bacS*<sub>27</sub>, and *bacS*<sub>80</sub>. The nucleotide that differs in each case also is shown in bold. (In some cases, the changed nucleotide sequence results in the same amino acid.) The upstream regulatory region is almost completely conserved among all three ORFs, the only difference being five additional nucleotides present between the *Bam*HI restriction site and the initiation codon of *bacS*<sub>80</sub> (an A between nucleotides 304 and 305, a TAG between nucleotides 315 and 316, and a G between nucleotides 328 and 329). The three sequences diverge after the TAA nonsense codon.

agar. Antibiotics were added in the following concentrations: gentamicin at 30 µg/ml, kanamycin at 30 µg/ml, nalidixic acid at 20 µg/ml, spectinomycin at 100 µg/ml, streptomycin at 200 µg/ml, and tetracycline at 5 or 1 µg/ml.

*Escherichia coli* strains were grown at 37°C in Luria-Bertani (LB) medium (Sambrook et al. 1989) in liquid shake cultures or on 1.5% Bacto agar. Antibiotics were added in the following concentrations: 30 µg of gentamicin per ml, 100 µg of spectinomycin per ml, and 15 µg of tetracycline per ml.

#### Plant tests.

*P. vulgaris* cv. Midnight (bean) seeds were surface sterilized with commercial bleach, germinated for 2 days, and transferred to growth pouches (Mega International, Minneapolis, MN) containing nitrogen-free RBN plant medium (Noel et al. 1984; Wacek and Brill 1976). The seedlings were inoculated with bacteria grown in TY broth, and Fix activity was determined by acetylene reduction of intact roots in stoppered serum vials (Noel et al. 1982, 1984).

#### Protein analysis by SDS PAGE.

Bacterial cells and isolated bacteroids were prepared in SDS buffer. The extracts were separated by SDS-PAGE (Laemmli 1970) with 12 or 15% acrylamide in the separating gel. The gel was stained with Coomassie blue (Dzandu et al. 1984) or electroblotted onto nitrocellulose. The blot was incubated with rabbit polyclonal antiserum as described by Carlson and associates (1987). Bound antibodies were detected with goat alkaline-phosphatase-conjugated anti-rabbit immunoglobulin G (Sigma-Aldrich, St. Louis) that was developed with 5-bromo-4-chloro-3-indolyl phosphate (Blake et al. 1984). Molecular weight markers (RPN756 from Amersham) included myosin, 200 kb; phosphorylase b, 92.5 kb; bovine serum albumin, 69 kb; ovalbumin, 46 kb; carbonic anhydrase, 30 kb; trypsin inhibitor, 21.5 kb; and lysozyme, 14.3 kb.

#### Antibody production.

Subcutaneous injections into New Zealand white rabbits were administered using a total of 0.4 to 0.5 cc of a suspension of

<b>BacS</b>	1	<b>MTKISERSFETIENP</b> ---	<b>GSSSLKVPDQFGASV</b>	<b>E</b>	<b>KGNKKVTEAFLKLAGAEATQKML</b>
<b>Y4wP</b>	1	<b>MSRTAEKPFETVENSIVGOSTAINRGEQLTHAI</b>	<b>E</b>	-----	<b>LQELA</b>
<b>MLhp</b>	9	<b>TIENVEFP</b> ---	<b>SFDASKATDQIRAF</b>	<b>E</b>	<b>KGVEQSKEAYAKLKTGAETQKAL</b>
<b>SMhp</b>	3	<b>TRKTEDAFSLSFDP</b> -----	<b>TKFADSFREF</b>	<b>E</b>	<b>KGAAQSKEAYAKMKTATEEAAKTV</b>
<b>PvLb</b>	1		<b>MGAFT</b>	<b>E</b>	<b>KQEALVNSSWEAFKGNIPQYSVVF</b>
<b>PsLb</b>	1		<b>MG-FT</b>	<b>E</b>	<b>KQEALVNSSWELFKQNPSGYSVLF</b>
					*
<b>BacS</b>	56	<b>PPILETTSLFGNE</b>	<b>L WWKTIA A LQADAEASFS</b>	<b>H LQALL G AN S P S QILEQ</b>	
<b>Y4wP</b>	40	<b>GSVFKSTMMVGFE</b>	<b>L SLNTFA A VQANTFANLH</b>	<b>H LSEFL G AK S P S QVFDL</b>	
<b>MLhp</b>	67	<b>ESTYETAKTVSSD</b>	<b>L SLKTIA A LRANAEGFS</b>	<b>H LEALI G AK S L S EVVEL</b>	
<b>SMhp</b>	53	<b>EATVESAQSGTVE</b>	<b>L GLKAID A LRTNAESSLS</b>	<b>H MEALL G VR S L S ELVEL</b>	
<b>PvLb</b>	31	<b>YTSILEKAPAAKN</b>	<b>L FSFLAN G VDPTNPKLTA</b>	<b>H AESLF G LV R D S AAQLR</b>	
<b>PsLb</b>	30	<b>YNIILKKAPAAKG</b>	<b>L FSFLKD A GVEDSPKLQA</b>	<b>H AEQVF G LV R D S AAQLR</b>	
					*
<b>BacS</b>	104	<b>Q ST F F RK RVETSLOHA</b>	<b>K EVRVLSSRAVEEIS</b>	<b>KP V K DAFD K</b>	
<b>Y4wP</b>	88	<b>Q CS F L RK RIDMGVEQV</b>	<b>K ESQALSTKALIDMS</b>	<b>KP I K EFFE K</b>	
<b>MLhp</b>	105	<b>Q TA F L RK RVELTVEQA</b>	<b>K DFQAVASKAEEVS</b>	<b>KP I K TAFE K</b>	
<b>SMhp</b>	101	<b>Q TA F I RK QAElaveQA</b>	<b>K TMQEATRkVAENVA</b>	<b>KP G K DAAE K</b>	
<b>PvLb</b>	79	<b>A NG A V VA DAALGSIHS</b>	<b>Q KGVNDSQFLVVKEA</b>	<b>LL K T LKAE V</b>	
<b>PsLb</b>	78	<b>T KG E V VL NATLGAIHV</b>	<b>Q RGVTDPHFVVVKEA</b>	<b>LL Q T IKKA S</b>	
<b>BacS</b>	144	<b>VLTDL K AT</b>	151		
<b>Y4wP</b>	128	<b>AIADL K TA</b>	135		
<b>MLhp</b>	145	<b>AIKD I K AA</b>	152		
<b>SMhp</b>	141	<b>AMSSL K KG</b>	148		
<b>PvLb</b>	119	<b>GDKWT D ELSTALELAYDEFAAGIKKAYA</b>	146		
<b>PsLb</b>	118	<b>GNNWS E ELNTAWEVAYDGLATAIKKAMT</b>	145		

Fig. 7. Alignment of the amino acid sequences of BacS, Y4wP, MLhp, SMhp, and leghemoglobin from *Phaseolus vulgaris* and *Pisum sativum*. The sequences are aligned first to compare BacS and the three rhizobial hypothetical proteins (Y4wP, MLhp, and SMhp). Identical residues between BacS<sub>03</sub> and these hypothetical rhizobial proteins are in bold and similar residues are underlined. Residues that are identical for all four polypeptides are separated by spaces. Y4wP is from the symbiotic plasmid of *Sinorhizobium fredii* NGR234, MLhp (Np107027.1) from *Mesorhizobium loti*, and SMhp (SMc00777AA) from *Sinorhizobium meliloti*. Below these three BacS homologs, two leghemoglobin sequences are presented without any gapping, except for one which was necessary to maintain alignment between the leghemoglobins. The amino acid sequences for the leghemoglobins from *Phaseolus vulgaris* (PvLb) and *Pisum sativum* (PsLb) have residues in bold where they are identical at the same position with residues that are conserved in BacS and all three hypothetical rhizobial proteins. Asterisks indicate the positions of the well-known distal and proximal heme-coordinating histidines in the leghemoglobins.

BacS protein antigen, emulsified from the BacS portion of a nitrocellulose blot of an SDS-PAGE gel, in phosphate-buffered saline, pH 7.2 and Freund's adjuvant (Sigma-Aldrich) (Harlow and Lane 1988). Boosters were given intramuscularly or intravenously once every 2 weeks for a total of 4 to 8 weeks. To increase the monospecificity, acetone powders of CE3 cell cultures grown aerobically were added to the undiluted antisera at a 1% (wt/vol) concentration and incubated for 30 min at 4°C (Harlow & Lane, 1988). The resulting suspension was centrifuged at 10,000 × g for 10 min and the supernatant used as the source of antibodies.

### Isolation of bacteroids and separation into cell envelope and soluble portions.

Bacteroids and the soluble nodule plant proteins were harvested from crushed nodules as described by Tao and associates (1992). Isolated bacteroids were resuspended in 50 mM Tris-HCl, sonicated on ice with nine 20-sec bursts, and centrifuged at 1,000 × g for 20 min to pellet the unbroken cells. The supernatant was removed and centrifuged at 15,000 × g for 60 min. The resulting soluble and insoluble (cell envelope) fractions were analyzed by SDS PAGE and Western blot.

### Expression of the protein ex planta.

Anaerobic and microaerobic cultures were grown in tubes sealed with serum stoppers, in which the headspace was purged of air with nitrogen after the inoculum was added. Sodium nitrate (10 mM final) was added in some cases. In cultures described as having 1% O<sub>2</sub>, a volume of air was reintroduced into the sealed vessels to give a concentration of 1% O<sub>2</sub> in the headspace of the vessel. Aerobic cultures were grown with loose caps and no purging with nitrogen. All cultures were grown at 30°C on an orbital shaker at 200 rpm.

### Partial purification and N-terminal sequencing of the BacS protein.

The nonsoluble fraction of bacteroids was isolated by sonication and differential centrifugation as described above, except that final centrifugation to pellet the membranes was at 13,000 × g for 90 min. After SDS-PAGE in 15% (wt/vol) acrylamide, the contents of the gel were electroblotted to polyvinylidene difluoride membranes, stained with Amido black (Harlow and Lane 1988). The portion of the membrane containing the BacS protein was removed and N-terminal sequencing was performed by Liane Mende-Mueller at the Medical college of Wisconsin on a Beckman LF 3000G Peptide Sequencer.

### Southern analysis.

Plasmid and genomic DNA were isolated by use of cetyltrimethylammonium bromide (Ausubel et al. 1994; Del Sal et al. 1989). DNA digested with restriction enzymes was fractionated by electrophoresis on 0.7% (wt/vol) agarose and electroblotted to Nytran-Plus membrane (Schleicher and Schuell). Prehybridization, hybridization, and washes were performed at 65°C. After prehybridization in buffer (containing 0.25 M Na<sub>2</sub>HPO<sub>4</sub>, pH 7.2, 5% SDS, and salmon sperm DNA at 10 µg/ml) overnight, the membrane was incubated with the denatured plasmid DNA (labeled with <sup>32</sup>P-dATP by random priming [Pharmacia]) in fresh buffer for 5 h. The membrane then was rinsed twice in wash A (20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.2, and 5% SDS) and twice in wash B (20 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.2, and 1% SDS) for 15 min each, sealed in plastic, and subjected to autoradiography.

### Cloning of the PCR product and a BamHI-EcoRI fragment overlapping a portion of bacS.

Using *R. meliloti* codon preferences from J. M. Cherry (*personal communication*), degenerate oligonucleotide sequences

**Table 1.** Bacterial strains and plasmids

Strain	Relevant properties <sup>a</sup>	Reference or source
<i>Rhizobium etli</i>		
CE3	Str <sup>r</sup> derivative of nodule isolate CFN42	Noel et al. 1984
CFNX89	CFN42 lacking the pSym	Brom et al. 1992
CFNX247	CFN42 <i>nifA</i> ::Sp/Sm	Girard et al. 1996
CFNX250	CE3 derivative carrying truncated pSym plasmid pGM1	D. Romero, unpublished data
CE459 <sup>b</sup>	CFNX250::pOJB3QT, <i>bacS</i> ::Tc and <i>bacS</i> <sup>+</sup>	This study
CE461 <sup>b</sup>	CE459, <i>bacS</i> ::Tc	This study
<i>Escherichia coli</i>		
DH5α	Cloning host strain	Gibco-B
MT616	MT607 + pRK600 helper plasmid	Finan et al. 1986
JM109	Cloning host strain	Yanisch et al. 1985
INV•F'	Cloning host strain	Invitrogen
Plasmids		
pBKS <sup>-</sup>	Cloning vector pBluescript KS <sup>-</sup>	Stratagene
pBG27	<i>R. etli</i> CFN42 pSym BamHI fragment B27 in vector pBR328	Girard et al. 1991
pGM1	Modified p42d(pSym), <i>nod</i> <sup>+</sup> <i>fix</i> <sup>+</sup> contained within <i>R. etli</i> CFNX250	D. Romero, unpublished data
pHP45Ω-Tc	Donor of Tc interposon	Fellay et al. 1987
pJQ200	Suicide vector Gm <sup>r</sup> , <i>sacB</i> , for gene replacement	Quandt and Hynes 1993
pCR2.1	TA vector for cloning PCR products	Invitrogen
pRK2013	RK2 derivative, Km <sup>r</sup> Tra <sup>+</sup>	Ditta et al. 1980
pFAJ1700	Tc <sup>r</sup> RK2-derived cloning vector	Dombrecht et al. 2001
pOJ1	pBKS <sup>-</sup> carrying 75-bp <i>bacS</i> PCR product corresponding to N-terminal 25 amino acids of BacS	This study
pOJ2	pBKS <sup>-</sup> carrying <i>Bam</i> HI- <i>Eco</i> RI fragment that includes the 5' portion of <i>bacS</i> <sub>27</sub>	This study
pOJB3	pCR2.1 carrying 2.6-kb PCR product from B02-B03 region of pGM1	This study
pOJB3Q	pJQ200 carrying 2.6-kb PCR product from B02-B03 region of pGM1	This study
pOJB3QT <sup>b</sup>	pOJB3Q::Tc	This study
pTA27	pCR2.1 carrying 2.6-kb PCR product from B27-B28 region of pSym	This study
pTA80	pCR2.1 carrying 2.6-kb PCR product from B79-B80 region of pSym	This study
pOJB27F	pFAJ1700 carrying 2.6-kb PCR product excised from pTA27	This study
pOJB80F	pFAJ1700 carrying 2.6-kb PCR product excised from pTA80	This study

<sup>a</sup> PCR = polymerase chain reaction.

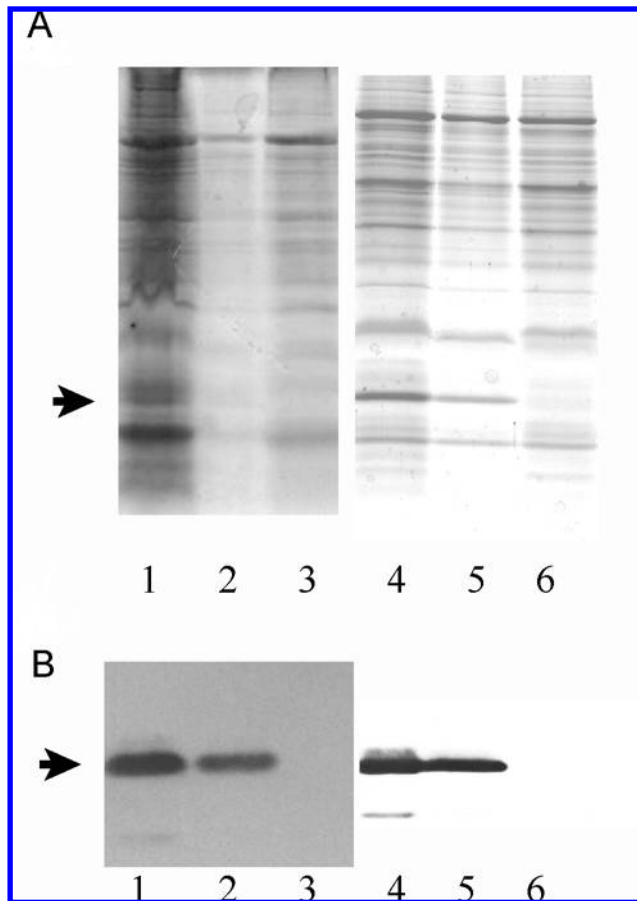
<sup>b</sup> Strains and plasmids with a tetracycline resistance cassette (Tc) inserted in the *Sma*I restriction site of the *bacS* gene.



were designed for use as primers in PCR amplification (Ericom Power Block II) of the portion of the structural gene encoding the N-terminus of the protein. The sequences were TAGGATCC-AAG-ATC-TC(C/G)-GA(A/G)-CG(G/T/C)-TC(C/G)-TT and CCGAATTC-CC-GAA-CTG-(A/G)TC-(T/G/C)GG-(G/C)AC-(C/T)TT, including (underlined) *Bam*HI or *Eco*RI restriction sites (synthesized by Operon). Genomic DNA from *R. etli* CE3 was used as the template. The PCR product was isolated from an agarose gel and ligated into pBKS- after digestion of the insert and vector DNA with *Bam*HI and *Eco*RI to obtain pOJ1 (Table 1). The *Bam*HI-*Eco*RI fragment from pBG27 (Table 1) was inserted similarly into pBKS- to obtain pOJ2 (Table 1). These recombinant plasmids were cloned in *E. coli* strain DH5 $\alpha$  by transformation of CaCl<sub>2</sub>-treated cells (Sambrook et al. 1989).

### Sequence analysis.

A Sequenase (United States Biochemical) or Fidelity kit (Oncor Inc.) was used to sequence the *Bam*HI-*Eco*RI fragment (Fig. 6) of plasmid pOJ2. The *bacS*<sub>03</sub> ORF within the *Bam*HI fragment of the pSym (Fig. 6) was PCR-amplified using *R. etli* CFNX250 as template. This PCR product and the *bacS*<sub>27</sub> ORF sequence in template pBG27 using the same four primers were sequenced at the DNA Sequencing Facility at the University of Wisconsin-Milwaukee (Perkin-Elmer Applied Biosystem Automated).



**Fig. 8.** Nod<sup>+</sup>Fix<sup>+</sup> mutant strain CE461 does not express *bacS* ex planta or in the nodule. **A**, Gels stained with Coomassie blue. **B**, Immunoblots of identically loaded lanes were reacted with polyclonal antibodies to BacS. **A** and **B**, CFNX250 (*bacS*<sup>+</sup>) (lanes 1 and 4), CE459 (*bacS*<sup>+</sup>/*bacS*<sup>-</sup>) (lanes 2 and 5), and CE461 (*bacS*<sup>-</sup>) (lanes 3 and 6), grown under anaerobic conditions ex planta (lanes 1, 2, and 3) or bacteroids harvested from *Phaseolus vulgaris* 19 days post inoculation (lanes 4, 5, and 6).

### Construction of mutants.

A 2.6-kb DNA fragment that stretches from 1 kb 5' of *bacS*<sub>03</sub> to 1 kb 3' of this ORF within the symbiotic plasmid was PCR-amplified from CFNX250 DNA by using primers AGTGGGCCCGTTCGACGCCTTCGCCGATTGA and CA-TTCTAGATGTGGCAGCCGACGATACCAGCAT (Operon), which were based on the total pSym sequence (VV. González, P. Bustos, M. A. Ramírez-Romero, A. Medrano-Soto, H. Salgado, I. Hernández-González, J. C. Hernández-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodríguez, M. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Dávila, unpublished data). The PCR product was cloned into vector pCR2.1 (Invitrogen) to generate plasmid pOJB3, which was digested with *Apa*I and *Xba*I (underlined in the primer sequences above), and the 2.6-kb PCR fragment was cloned into vector pJQ200 to obtain a plasmid pOJB3Q. The Tc<sup>r</sup> cassette released from pHP45-Tc by *Sma*I digestion was ligated to pOJB3Q after it was linearized at the unique *Sma*I site within the *bacS* open reading frame (Fig. 6) to obtain pOJB3QT. The Gm<sup>r</sup> Tc<sup>r</sup> *E. coli* JM109 transformant carrying this plasmid was mated with *E. coli* MT616 and *R. etli* CFNX250. One Nal<sup>r</sup>, Tc<sup>r</sup>, Gm<sup>r</sup>, Str<sup>r</sup> isolate, called CE459, was cultivated on TY containing 10% sucrose and tetracycline at 1 μg/ml for selection of isolates in which a second crossover eliminated the pJQ200 portion of pOJB3QT and the wild-type *bacS* locus, resulting in strains carrying only the disrupted *bacS* locus. Three isolates, of which strain CE461 is representative, were screened by PCR and Southern analysis for the presence of the insertional mutation.

### Complementation analysis.

A 2.6-kb fragment of DNA from either the B27 and B28 or B79 and B80 region of the symbiotic plasmid was PCR amplified by using primers based on pSym sequence information (V. González, P. Bustos, M. A. Ramírez-Romero, A. Medrano-Soto, H. Salgado, I. Hernández-González, J. C. Hernández-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodríguez, M. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Dávila, unpublished data) and genomic DNA from CE3 was used as template. The primers used for the B27 and B28 region were TACCGTTGACGACACATTGG and GAAGACAATAACCGGACGCT. Those used for the B79 and B80 region were GCTGACCGCCGATTGAGAT and CATGCGCTTGCCCTCCTTACG. The PCR product was ligated into pCR2.1 to generate plasmids pTA27 and pTA80.

To construct hybrid plasmids for complementation of the *bacS*<sup>-</sup> mutants, pTA27, pTA80, and pFAJ1700 were digested with *Kpn*I and *Xba*I. After ligation of pTA27 or pTA80 with pFAJ1700, two plasmids, pOJB27F and pOJB80F, were isolated from *E. coli* JM109 transformant colonies having resistance to tetracycline at 15 μg/ml. Triparental mating (with *E. coli* carrying the mobilization plasmid, pRK2013) (Ditta et al. 1980) was used to move pOJB27F or pOJB80F from JM109 to the *bacS*<sup>-</sup> mutants. The isolates were cultivated on TY containing streptomycin at 200 μg/ml, nalidixic acid at 30 μg/ml, and tetracycline at 5 μg/ml for selection of isolates having either pOJB27F or pOJB80F.

### ACKNOWLEDGMENTS

This work was supported by grants R15 GM57775-01 from NIH and DE-FG02-98ER20307 from DOE. We thank S. Brom, M. Cevallos, and L. Girard from Centro de Investigacion sobre Fijacion de Nitrogeno in Cuernavaca for helpful discussions, pBG27, and *R. etli* CFNX89 and CFNX247; and L. Mende-Mueller at the Medical College of Wisconsin for the N-terminal sequencing of BacS, and we acknowledge the use of the DNA Sequencing Facility at the University of Wisconsin-Milwaukee.

## LITERATURE CITED

- Ausubel, M., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. A., and Struhl, K. 1994. Current Protocols in Molecular Biology. New York: Greene Publishing Associates, Inc., Brooklyn, NY, U.S.A.
- Blake, M. S., Johnston, K. H., Russel-Jones, G. J., and Gotschlich 1984. A rapid and sensitive method for detection of alkaline phosphatase conjugated anti-antibody on western blots. *Anal. Biochem.* 136:175-179.
- Brendel, V., Bucher, P., Nourbakhsh, I., Blaisdell, B. E., and Karlin, S. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. U.S.A.* 89:2002-2006.
- Brom, S., Garcia de los Santos, A., Stepkowski, T., Flores, M., Davila, G., Romero, D., and Palacios, R. 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J. Bacteriol.* 174:5183-5189.
- Carlson, R. W., Kalembasa, S., Turowski, D., Pachori, P., and Noel, K. D. 1987. Characterization of the lipopolysaccharide from a *Rhizobium phaseoli* mutant that is defective in infection thread development. *J. Bacteriol.* 169:4923-4928.
- Cevallos, M. A., Vazquez, M., Davalos, A., Espin, G., Sepulveda, J., and Quinto, C. 1989. Characterization of *Rhizobium phaseoli* Sym plasmid regions involved in nodule morphogenesis and host-range specificity. *Mol. Microbiol.* 3:879-889.
- Del Sal, G., Manfioletti, G., and Schneider, C. 1989. The CTAB-DNA precipitation method: A common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. *BioTechniques* 7:514-520.
- Ditta, G. 1989. Regulation of *nif* genes in *Rhizobium*. *Mol. Plant-Microbe Interact.* 11-24.
- Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* 77:7347-7351.
- Dombrecht, B., Vanderleyden, J., and Michiels, J. 2001. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. *Mol. Plant-Microbe Interact.* 14:426-430.
- Duelli, D., 1999. Ph.D. thesis. Marquette University, Milwaukee, WI, U.S.A.
- Dzandu, J. K., Deh, M. E., Barratt, D. L., and Wise, G. E. 1984. Detection of erythrocyte membrane proteins, sialoglycoproteins, and lipids in the same polyacrylamide gel using a double-staining technique. *Proc. Natl. Acad. Sci. U.S.A.* 81:1733-1737.
- Fellay, R., Frey, J., and Krisch, H. 1987. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. *Gene* 52:147-154.
- Finan, T., Kunkel, B., De Vos, G., and Signer, E. 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* 167:66-72.
- Fischer, H. 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* 58:352-386.
- Franssen, H. J., Vijn, I., Yang, W. C., and Bisseling, T. 1992. Developmental aspects of the *Rhizobium*-legume symbiosis. *Plant Mol. Biol.* 19:89.
- Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A., and Perret, X. 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* 387:394-401.
- Gailbert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Huber, F., Barnett, M. J., Becker, A., Boistard, P., Both, G., Boutry, M., Bowser, L., Buhmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F., Weidner, S., Wells, D. H., Wong, K., Yeh, K., Batut, J. 2001 The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293:668-672.
- Girard, M. L., Flores, M., Brom, S., Romero, D., Palacios, R., and Davila, G. 1991. Structural complexity of the symbiotic plasmid of *Rhizobium leguminosarum* biovar *phaseoli*. *J. Bacteriol.* 173:2411-2419.
- Girard, M. L., Valderrama, B., Palacios, R., Romero, D., and Davila, G. 1996. Transcriptional activity of the symbiotic plasmid of *Rhizobium etli* is affected by different environmental conditions. *Microbiology (UK)* 142:2847-2856.
- Harlow, E., and Lane, D. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.
- Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S. 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* 7(6):331-338.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the bacteriophage T4. *Nature* 227:680-685.
- Merrick, M. J. 1992. Regulation of nitrogen fixation genes in free-living and symbiotic bacteria. *Biol. Nitrogen Fixation* 835-876.
- Moen, L., Vanfleteren, J., Van de Peer, Y., Peeters, K., Kapp, O., Czeluzniak, J., Goodman, M., Blaxter, M., and Vinogradov, S. 1996. Globins in nonvertebrate species: dispersal by horizontal gene transfer and evolution of the structure-function relationships. *Mol. Biol. Evol.* 13:324-333.
- Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. 2001. Nodulation of legumes by members of the B-subclass of Proteobacteria. *Nature* 411:948-950.
- Noel, K. D., Carneol, M., and Brill, W. J. 1982. Nodule protein synthesis and nitrogenase activity of soybeans exposed to fixed nitrogen. *Plant Physiol.* 70:1236-1241.
- Noel, K. D., Diebold, R. J., Cava, J. R., and Brink, B. A. 1988. Rhizobial purine and pyrimidine auxotrophs: Nutrient supplementation, genetic analysis, and the symbiotic requirement for *de novo* purine biosynthesis. *Arch. Microbiol.* 149:499-506.
- Noel, K. D., Sanchez, A., Fernandez, L., Leemans, J., and Cevallos, M. A. 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J. Bacteriol.* 158:148-155.
- Palacios, R., Flores, M., Martinez, E., and Quinto, C. 1985. Nitrogen fixation genes and DNA reiteration. Pages 173-179 in: Nitrogen Fixation Research Progress. C. Veerger and W. E. Newton, eds. Martinus Nijhoff Publishers, The Hague, The Netherlands.
- Perret, X., Freiberg, C., Rosenthal, A., Broughton, W. J., and Fellay, R. 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. *Mol. Microbiol.* 32:415-425.
- Planque, K., van Nierop, J. J., and Burgers, A. 1979. The lipopolysaccharide of free-living and bacteroid forms of *Rhizobium leguminosarum*. *J. Gen. Microbiol.* 110:151-159.
- Quandt, J., and Hynes, M. F., 1993. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* 127:15-21.
- Quinto, C., de la Vega, H., Flores, M., Fernandez, L., Ballado, T., Soberon, G., and Palacios, R. 1982. Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. *Nature (Lond.)* 299:724-726.
- Romero, D., Martinez-Salazar, J., Girard, L., Brom, S., Davila, G., Palacios, R., Flores, M., and Rodriguez, C. 1995. Discrete amplifiable regions (amplicons) in the symbiotic plasmid of *Rhizobium etli* CFN42. *J. Bacteriol.* 173:973-980.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.
- Tao H., Brewin, N. J., and Noel, K. D. 1992. *Rhizobium leguminosarum* CFN42 lipopolysaccharide antigenic changes induced by environmental conditions. *J. Bacteriol.* 174:2222-2229.
- Wacek, T. S., and Brill, W. J. 1976. Simple, rapid assay for screening nitrogen-fixing ability in soybean. *Crop Sci.* 16:519-522.
- Walker, S. A., and Downie, J. A. 2000. Entry of *Rhizobium leguminosarum* bv. *viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. *Mol. Plant-Microbe Interact.* 13:754-762.
- Yanisch-Perron, C., Viera, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and puc19 vectors. *Gene* 33:103-119.