JOURNAL OF BACTERIOLOGY, June 1992, p. 4130–4139
0021-9193/92/124130-10$02.00/0

Vol. 174, No. 12

Nucleotide Sequence and Characterization of Four Additional Genes of the Hydrogenase Structural Operon from *Rhizobium leguminosarum* bv. viciae

ELENA HIDALGO, JOSÉ M. PALACIOS, JESÚS MURILLO,† AND TOMÁS RUIZ-ARGÜESO*

Departamento de Microbiología, Escuela Técnica Superior de Ingenieros Agrónomos, 28040-Madrid, Spain

Received 3 February 1992/Accepted 15 April 1992

The nucleotide sequence of a 2.5-kbp region following the hydrogenase structural genes (*hupSL*) in the H_2 uptake gene cluster from *Rhizobium leguminosarum* bv. viciae UPM791 was determined. Four closely linked genes encoding peptides of 27.9 (*hupC*), 22.1 (*hupD*), 19.0 (*hupE*), and 10.4 (*hupF*) kDa were identified immediately downstream of *hupL*. Proteins with comparable apparent molecular weights were detected by heterologous expression of these genes in *Escherichia coli*. The six genes, *hupS* to *hupF*, are arranged as an operon, and by mutant complementation analysis, it was shown that genes *hupSLCD* are cotranscribed. A transcription start site preceded by the −12 to −24 consensus sequence characteristic of NtrA-dependent promoters was identified upstream of *hupS*. On the basis of the lack of oxygen-dependent H_2 uptake activity of a *hupC*::Tn*5* mutant and on structural characteristics of the protein, we postulate that HupC is a *b*-type cytochrome involved in electron transfer from hydrogenase to oxygen. The product from *hupE*, which is needed for full hydrogenase activity, exhibited characteristics typical of a membrane protein. The features of HupC and HupE suggest that they form, together with the hydrogenase itself, a membrane-bound protein complex involved in hydrogen oxidation.

A few strains of *Rhizobium* and *Bradyrhizobium* spp. induce in legume nodules the synthesis of an H_2 uptake hydrogenase whose main function is assumed to be the oxidation of molecular H_2 produced by the nitrogenase system. The oxidation of H_2 results in an increase of the energy efficiency of symbiotic nitrogen fixation and provides mechanisms to prevent oxygen damage and hydrogen inhibition of the nitrogenase. At least in the soybean-*Bradyrhizobium japonicum* symbiosis, these benefits have been shown to stimulate plant yield (9).

The H_2 uptake hydrogenase from *B. japonicum* is a membrane-bound, nickel-containing enzyme composed of two subunits of about 35 and 65 kDa (9, 30). The large subunit from *B. japonicum* is immunologically cross-reactive with the hydrogenase from *Rhizobium leguminosarum* bv. viciae (22). During H_2 oxidation, the electrons from H_2 are injected into the electron transport chain apparently at the level of ubiquinone (UB) (7, 29) and transferred to O_2 either via *b*-, *c*-, and *aa*-type cytochromes or via cytochrome *o* (28, 30).

The genetic determinants for H_2 oxidation (*hup* genes) have been isolated from *R. leguminosarum* bv. viciae UPM791 (21). In this species, *hup* genes are clustered in a DNA fragment of about 15 kb of the symbiotic plasmid and organized in at least four transcriptional units designated regions *hup*I/II, *hup*III, *hup*IV, and *hup*V/VI (22). Region *hup*I/II contains the structural genes for the hydrogenase (*hupS* and *hupL*) which had previously been sequenced (14). The same genes have also been sequenced from *R. leguminosarum* bv. viciae B10 (39). By using *hup-lacZ* fusions, it was shown that region *hup*V/VI is induced in vegetative cells in response to low oxygen tensions in the culture medium (31). Regions *hup*I/II, *hup*III, and *hup*IV are expressed only in symbiosis with peas, what has made the genetic analysis of these regions difficult.

We present here the sequencing and characterization of four additional genes located downstream of the *R. leguminosarum* bv. viciae hydrogenase structural genes together with preliminary evidence for the function of one of the gene products (HupC) as a *b*-type cytochrome.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. Bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were routinely grown at 37°C in Luria-Bertani medium or in 2×YT medium for single-stranded DNA preparations (35). Conditions for routinely growing *Rhizobium* strains were as previously described (22).

Chemicals. Unless otherwise stated, all restriction endonucleases, DNA-modifying enzymes, and nucleotide triphosphates were purchased from Boehringer Mannheim GmbH (Mannheim, Germany) and were used according to the manufacturer's indications. Radioactive compounds were obtained from Amersham International plc (Amersham, United Kingdom). Medium constituents were from Oxoid Ltd. (Basingstoke, United Kingdom). Antibiotics used in growth media were from Boehringer and Serva Feinbiochemica (Heidelberg, Germany), except that the rifampin used for protein expression in *E. coli* was purchased from Sigma Chemical Co. (St. Louis, Mo.). All other chemicals were of reagent or electrophoresis grade.

Plant tests and hydrogenase activities. Hydrogenase activities of *Rhizobium* strains were tested in bacteroids from nodules induced in pea plants (*Pisum sativum* cv. Frisson). *Rhizobium* strains were used as inocula for surface-sterilized pea seedlings, and plants were grown under conditions

* Corresponding author.
† Present address: Department of Plant Pathology, University of California, Riverside, CA 92521-0122.

TABLE 1. Bacterial strains and plasmids used

Strain or plasmid	Characteristics[a]	Source or reference
R. leguminosarum		
UPM791	128C53; Str^r Hup$^+$ on peas	21
AL series	UPM derivatives with genomic Tn*5* insertions	22
E. coli BL21-DE3	*hsdS gal* (λ *cIts857 ind1 Sam7 nin5 lacUV5*-T7 gen*1*)	42
Plasmids		
pAL618	pLAFR1-derived cosmid containing *hup* region from *R. leguminosarum*	21
pAL series	pAL618::Tn*5*; Tc^r Km^r	22
pBluescript (pSK+, pSK−)	pUC19 derivative plasmids containing T7 promoter	41
pRLH series	pSK vectors carrying DNA fragments from pAL618 insert DNA	This work

[a] Abbreviations: Str, streptomycin; Tc, tetracycline; Km, kanamycin.

previously described (21). Bacteroids were prepared from nodules of 3-week-old plants by the method described by Leyva et al. (22), and bacteroid hydrogenase activity was determined by an amperometric method using O_2 or methylene blue (MB) as terminal electron acceptor (34). The protein contents of bacteroid suspensions were determined by the procedure of Lowry et al. (23) after alkaline digestion of cells at 90°C in 1 N NaOH for 10 min.

Recombinant DNA techniques. Plasmid DNA preparations, restriction enzyme digestions, agarose and polyacrylamide DNA electrophoresis, and DNA cloning were performed by standard procedures (35).

Sequence determination and analysis. A 6.7-kb *Kpn*I-*Hin*dIII DNA fragment from pAL618 (21) was subcloned in pBluescript vectors in both orientations. A series of overlapping nested deletion fragments covering both strands of the whole DNA region were generated by exonuclease III-nuclease S1 treatment (13). The deletion fragments were sequenced by the dideoxy chain termination method of Sanger et al. (36) using Sequenase version 2.0 from United States Biochemicals (Cleveland, Ohio) and the forward and reverse M13 primers from Pharmacia Biotechnology International (Uppsala, Sweden) with either single- or double-stranded DNA templates.

For the precise locations of Tn*5* insertions in pAL plasmids, restriction fragments containing the Tn*5* kanamicin resistance gene together with the Tn*5* left arm and adjacent DNA region were subcloned in pBluescript vectors. These constructs were sequenced by using a 17-mer synthetic oligonucleotide (5′-ATGAACGTTACCATGTT-3′) complementary to the Tn*5* inverted repeat as primer.

Nucleotide and derived amino acid sequences were analyzed by using the University of Wisconsin Genetics Computer Group Sequence Analysis Software Package (6), DNA Strider (C. Marck, Departemente de Biologie, Centre d'Etudes Nucleares de Saclay, Gif-sur-Ivette, France), and Brujene (J. A. Vara, Centro de Biología Molecular, Universidad Autónoma de Madrid, Spain) computer programs.

Bacteroid RNA isolation and location of 5′ end of mRNA. RNA isolation from bacteroids was carried out as previously described (31).

The location of the mRNA 5′ end was determined by primer extension using a 17-mer synthetic oligonucleotide (5′-CCCTGGCGGCGAATGAC-3′) complementary to the mRNA corresponding to amino acids 10 to 15 from HupS. The synthetic oligonucleotide was labeled with γ-^{32}P by using polynucleotide kinase. About 5 ng of labeled oligonucleotide was mixed with 10 μg of bacterial total RNA in 15 μl of hybridization solution (10 mM Tris-HCl [pH 8.3], 50 mM KCl). The hybridization mixture was maintained for 3 min at 90°C and allowed to cool slowly to 35°C. Extension was accomplished by adding 4.5 μl of a mixture of the four triphosphate deoxynucleotides (4 mM each), 2.5 μl of TR buffer (250 mM Tris-HCl [pH 8.3], 100 mM MgCl, 50 mM dithiothreitol), 2 μl of sodium pyrophosphate, and 25 U of avian myeloblastosis virus reverse transcriptase (purchased from Pharmacia) and incubating the mixture for 90 min at 42°C. After the extension, the reaction was stopped by the addition of 2 μl of 250 mM EDTA, and the mixture was treated with 10 U of bovine pancreatic RNase (30 min at 37°C) to remove single-stranded RNA. Extended chains were precipitated with ethanol, washed with 80% ethanol, and loaded onto a sequencing gel.

Expression and labeling of proteins in *E. coli*. Proteins encoded in the *hup*I/II region were expressed in *E. coli* cells by using the bacteriophage T7 RNA polymerase-promoter system described by Tabor and Richardson (44). *E. coli* BL21-DE3, harboring gene 1 from T7 RNA polymerase under the control of a *lacUV5* promoter, was used as expression host. After adequate DNA restriction fragments were subcloned in pBluescript vectors, constructs were transformed into BL21-DE3. Conditions for isopropyl-β-D-thiogalactopyranoside (IPTG) induction of T7 RNA polymerase-dependent expression and labeling with L-[^{35}S]methionine were as described by Studier and Moffatt (42). After being labeled, cells were spun down, and crude cell extracts were prepared as previously described (22). Crude extracts were resolved in sodium dodecyl sulfate (SDS)-polyacrylamide discontinuous gels (19). Gels were stained, dried, and exposed to Kodak X-OMAT X-ray film.

Nucleotide accession number. The sequence data presented in this paper can be obtained from the EMBL data library under accession number X52974.

RESULTS

Determination and analysis of the nucleotide sequence of region *hup*I/II. The detailed restriction map of the 6.7-kb *Kpn*I-*Hin*dIII DNA fragment containing the region *hup*I/II is presented in Fig. 1B along with the general organization of the *hup* cluster contained in the 20-kb DNA insert of cosmid pAL618 (22). Different restriction fragments from this 6.7-kb DNA segment were subcloned in pBluescript vectors, and series of deletion derivatives of the subclones were generated. These deletion derivatives were used to determine the nucleotide sequence of a fragment of 5,596 bp (Fig. 2). The analysis of the sequence revealed the presence of six complete open reading frames (ORFs), designated ORF11 to ORF16, all oriented in the same direction (from left to right in Fig. 1B). The probability that these ORFs encode polypeptides was examined with the Brujene computer program employing a *R. leguminosarum* codon usage table (48). According to this analysis, all six ORFs had a high probability of coding for a protein.

The DNA sequences and amino acid translations of ORF11 and ORF12 were previously published (14). On the basis of their high homology with the genes for the *B.*

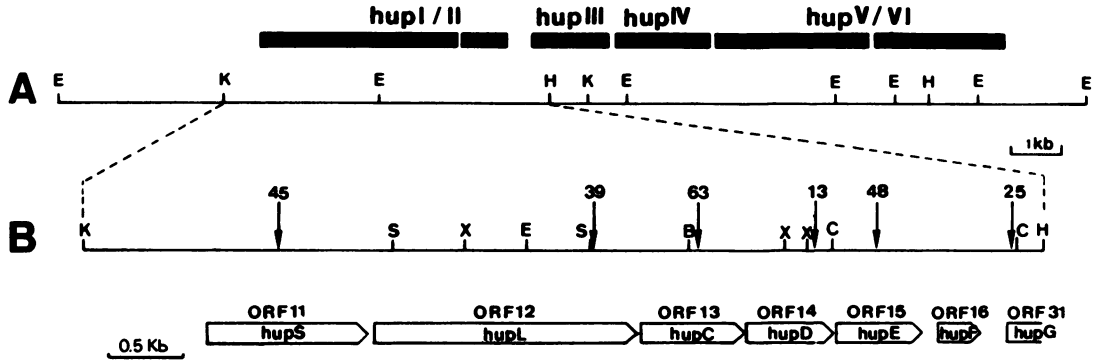

FIG. 1. Physical and genetic maps of region *hup*I/II. (A) Restriction map of cosmid pAL618 DNA insert and positions of putative *hup* transcriptional units (full bars) defined by Tn*5* mutagenesis and complementation analysis (22). (B) Detailed restriction and genetic maps of the 6.7-kb *Kpn*I-*Hin*dIII fragment containing the *hup*I/II region. Locations and directions of ORFs are indicated by open horizontal arrows. Positions of Tn*5* insertions in genomic AL mutants were determined by sequencing and are indicated by vertical arrows below the numbers of the corresponding mutants. Restriction enzymes: B, *Bgl*I; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; S, *Sal*I; X, *Xho*I.

japonicum hydrogenase subunits (37), ORF11 and ORF12 were identified as the genes for the small and large subunits of the *R. leguminosarum* bv. viciae hydrogenase and were designated *hupS* (ORF11) and *hupL* (ORF12). The percentages of amino acid identity of the predicted proteins from *R. leguminosarum* bv. viciae UPM791 *hupS-hupL* with the hydrogenase small and large subunits predicted from the corresponding genes in *B. japonicum* (37), *Rhodobacter capsulatus* (20), *Azotobacter vinelandii* (26), *Azotobacter chroococcum* (10), and *E. coli* (27) were 89.3 and 89.4, 76.2 and 74.6, 79.8 and 68.6, 75.8 and 64.6, and 68.0 and 63.5, respectively.

ORF13, ORF14, ORF15, and ORF16 are located downstream of *hupL*, and their nucleotide and derived amino acid sequences are shown in Fig. 2. The likely translational initiation codon for all four ORFs is ATG, which is preceded in every case by a purine-rich motif characteristic of potential ribosome-binding sites. All the ORFs except ORF13 terminate with TGA; ORF13 terminates with TAG. With the exception of ORF16, which is separated by 96 bp from the termination codon of ORF15, the other ORFs of the gene cluster, including *hupS* and *hupL*, are closely linked: ORF13 is contiguous with ORF14, and the intergenic spaces between *hupS* and *hupL*, *hupL* and ORF13, and ORF14 and ORF15 are 23, 13, and 11 bp, respectively. No promoter-related signals were observed in the intergenic region between ORF15 and ORF16. This arrangement suggests that the six ORFs constitute a single operon corresponding to the previously defined *hup*I/II region (22). Following the alphabetical gene designation adopted in the *E. coli* hydrogenase operons (25, 27), we named the corresponding genes *hupC* (ORF13) to *hupF* (ORF16). In the DNA located downstream of *hupF* in cosmid pAL618, a second gene cluster containing 11 genes has been identified (data to be presented elsewhere). The first of these genes, *hupG*, corresponds to ORF31, and a partial sequence of it is presented in Fig. 2. *hupG* apparently belongs to transcriptional unit *hup*III, previously defined by complementation analysis (22), since the Tn*5* insertion of Hup$^-$ mutant AL25, ascribed to region *hup*III, has been located in ORF31 (Fig. 2). The 166-bp DNA region between *hupF* and *hupG* has a higher content of A+T bases (57.6%) than the average of the overall sequence (38.7%), suggesting that it may be separating two operons. However, no significant inverted repeats characteristic of rho-independent transcriptional terminators (5) were found in this region.

Expression of *hup* gene products in *E. coli* cells. Polypeptides encoded by the *hup*I/II DNA region were identified by using an in vivo gene expression procedure based on the bacteriophage T7 RNA polymerase-promoter system (42). DNA restriction fragments containing all or part of the ORFs from the *hup*I/II and *hup*III regions (Fig. 3A) were subcloned in pBluescript vectors, and the resulting plasmids (pRLH) were expressed in *E. coli* BL21-DE3.

Seven protein bands with apparent molecular weights of approximately 65, 39, 26, 25, 17, 16, and 13 kDa were specifically labeled in strain BL21-DE3(pRLH382) (Fig. 3B, lane 5). Based on the specifically labeled protein bands obtained from strains BL21-DE3 carrying plasmids pRLH20, pRLH53, pRLH21, and pRLH82 (Fig. 3B, lanes 2, 3, 4, and 6, respectively), these polypeptides were identified as the gene products from *hupL*, *hupS*, *hupD*, *hupC*, *hupG*, *hupE*, and *hupF*, respectively. The strongly labeled protein band of about 43 kDa present in strain BL21-DE3(pRLH21) corresponds to a truncated protein from partial translation of *hupL*. In some cases (Fig. 3A, lanes 5 and 6), the gene product from *hupS*, the hydrogenase small subunit, is only weakly labeled. This could be because of a higher sensitivity of this protein toward proteolytic digestion, as has been reported for the *B. japonicum* hydrogenase small subunit (43). The sizes of the polypeptides translated from *hupS* to *hupF* agree quite well with those of the proteins predicted from the nucleotide sequence (Table 2).

Mutant characterization and complementation analysis. Several Hup$^-$ Tn*5* mutants of *R. leguminosarum* bv. viciae UPM791 were previously generated by site-directed mutagenesis of DNA associated with region *hup*I/II (22). Taking advantage of the nucleotide sequence availability, these mutants have now been precisely mapped by sequencing the DNA contiguous to the genomic Tn*5* insertion with a primer complementary to a 17-bp sequence of the transposon. Tn*5* insertions in mutants AL45, AL39, AL63, AL13, and AL48 were located within *hupS*, *hupL*, *hupC*, *hupD*, and *hupE*, respectively (Fig. 1B and 2). No Tn*5* insertions in *hupF* were available.

The effect of the Tn*5* insertions on the Hup phenotype was evaluated in symbiosis with peas (Table 3). Bacteroids from mutants carrying Tn*5* insertions in the hydrogenase struc-

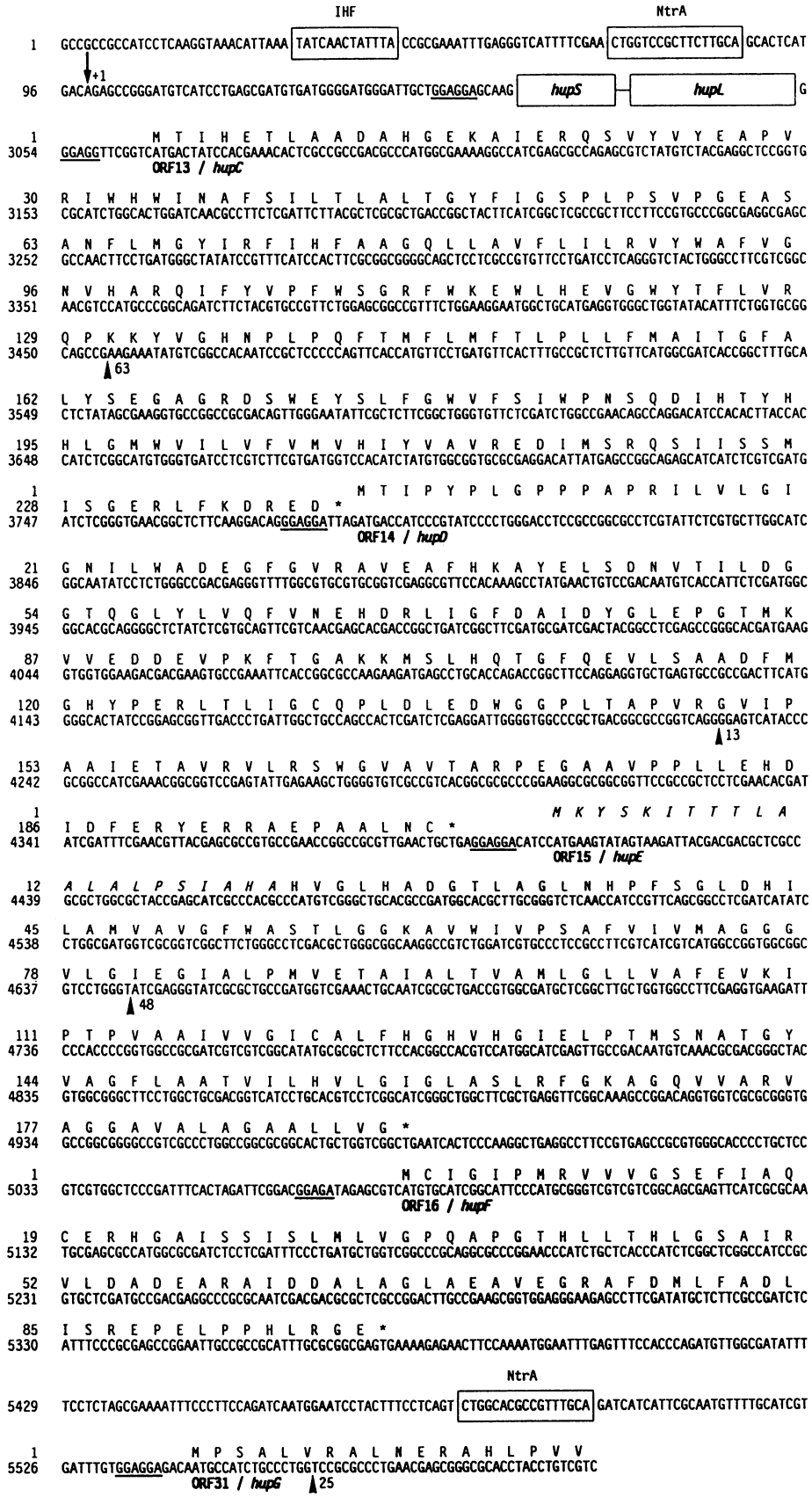

FIG. 2. Nucleotide sequence of *hupC* to *hupF* genes and flanking regions, including previously published *hupS* upstream region (14). Predicted amino acid sequences from ORFs are given above the nucleotide sequence. Potential ribosome-binding sites are underlined. The transcription initiation (+1) site is shown by a vertical arrow. The putative NtrA- and IHF-binding sites are boxed and indicated. Positions of Tn*5* insertions in the AL mutants are marked by arrowheads with corresponding designation numbers.

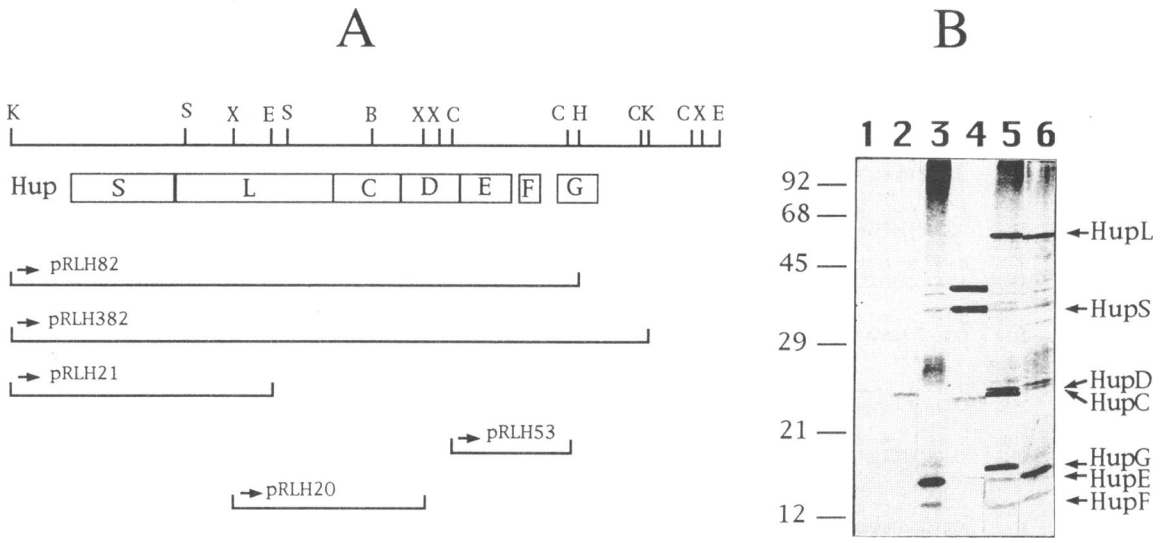

FIG. 3. Identification of proteins encoded in the *hup*I/II region by its expression in a T7 RNA polymerase-promoter system. (A) The extent of each restriction fragments carried by the pSK derivative plasmids (pRLH) is aligned with the location of the *hup* genes identified by sequencing in the *hup*I/II region and part of the *hup*III region. Arrows indicate the direction of transcription of T7 promoter from the vector. Abbreviations of restriction enzymes are as in Fig. 1. (B) Autoradiogram of SDS-polyacrylamide (15%) gel containing [^{35}S]methionine-labeled proteins from *E. coli* BL21-DE3 bearing plasmids indicated in panel A or pSK as follows: pSK (lane 1), pRLH20 (lane 2), pRLH53 (lane 3), pRLH21 (lane 4), pRLH382 (lane 5), and pRLH82 (lane 6). The positions and sizes (in kilodaltons) of the proteins used as molecular weight standards are shown at the left. From top to bottom, they are phosphorylase B, bovine serum albumin, egg albumin, carboanhydrase, soybean trypsin inhibitor, and cytochrome *c*. Arrows at the right indicate positions of the bands corresponding to the *hup* gene products in lane 5 (plasmid pRLH382).

tural genes (AL45 and AL39) or in *hupD* (AL13) were unable to oxidize H_2 with either O_2 or MB as terminal electron acceptor. Mutant AL48, bearing a Tn*5* insertion in *hupE*, exhibited a 50% reduction of these activities compared with the wild-type strain, suggesting that the gene products from *hupE* and/or *hupF* are not essential for hydrogenase activity in pea bacteroids but are required to attain full levels of activity. This implies that, besides the hydrogenase structural genes, the *hupD* gene product is absolutely required for hydrogenase activity in *R. leguminosarum* bv. viciae cells. The Tn*5* insertion in *hupC* (mutant AL63) completely suppressed the bacteroid hydrogenase activity with oxygen as electron acceptor, but a low H_2 uptake activity was still detected with MB. This MB-dependent hydrogenase activity is increased in bacteroids from the merodiploid strain AL63(pAL63) carrying multiple copies of pAL618 with a Tn*5* insertion in *hupC*.

Complementation analysis was carried out by examining

TABLE 2. Characteristics of predicted gene products from region *hup*I/II

ORF (gene)	Position on sequence	No. of amino acid residues	Mol wt (10^3)		pI	% Homology to *E. coli* hydrogenase genes[a]
			From sequence	From gel		
ORF11 (*hupS*)	156–1238	360	39.153	40	7.25	68 (*hyaA*)
ORF12 (*hupL*)	1262–3052	596	66.115	65	6.51	63 (*hyaB*)
ORF13 (*hupC*)	3066–3785	239	27.873	25	7.47	50 (*hyaC*)
ORF14 (*hupD*)	3786–4394	202	22.050	26	4.55	43 (*hyaD*)
ORF15 (*hupE*)	4406–4981	191	19.025	16	8.48	
ORF16 (*hupF*)	5078–5374	99	10.417	13	5.58	32 (*hypC*)

[a] Expressed as percent amino acid identity with the corresponding gene product. *E. coli* hydrogenase genes are from *hya* (27) and *hyp* (25) operons.

TABLE 3. Symbiotic hydrogenase activities of *R. leguminosarum* UPM791 derivative strains carrying Tn*5* insertions in genomic or genomic and cosmidic copies of *hup* genes

Strain	Location of Tn*5* insertion[a]	Hydrogenase activity[b]	
		O_2	MB
UPM791		1,120	1,270
UPM791(pAL618)		2,085	2,300
AL45	*hupS*	0	0
AL45(pAL45)	*hupS* (*hupS*)	0	0
AL39	*hupL*	0	0
AL39(pAL39)	*hupL* (*hupL*)	0	0
AL39(pAL45)	*hupL* (*hupS*)	0	0
AL39(pAL63)	*hupL* (*hupC*)	0	60
AL39(pAL13)	*hupL* (*hupD*)	170	180
AL63	*hupC*	0	14
AL63(pAL63)	*hupC* (*hupC*)	0	53
AL63(pAL39)	*hupC* (*hupL*)	0	49
AL63(pAL13)	*hupC* (*hupD*)	184	210
AL13	*hupD*	0	0
AL13(pAL13)	*hupD* (*hupD*)	0	0
AL13(pAL39)	*hupD* (*hupL*)	765	820
AL13(pAL63)	*hupD* (*hupC*)	940	980
AL48	*hupE*	650	680

[a] Locations of the Tn*5* insertions in pAL618 derivative cosmids (pAL) are in parentheses.

[b] Bacteroid hydrogenase activities are expressed as nanomoles of H_2 taken up per hour per milligram of protein with oxygen or MB as terminal electron acceptor. Zeros indicate no activity detected.

the symbiotic Hup phenotype of merodiploid strains carrying Tn*5* insertions in the genomic copies of *hup* genes with pAL618 derivative cosmids bearing Tn*5* insertions in different *hup* genes (Table 3). From the O_2-dependent hydrogenase activities of the merodiploid strains, it is inferred that *hupS*, *hupL*, and *hupC* genes belong to a single operon transcribed from upstream of *hupS*. This operon seems also to contain *hupD*, since mutants AL63 (*hupC*::Tn*5*) and AL39 (*hupL*::Tn*5*) were not complemented (only low levels of hydrogenase activity were detected) by cosmid pAL13 (*hupD*::Tn*5*). However, the full complementation observed in strains AL13(pAL39) and AL13(pAL63) suggests that transcription of *hupD* is not dependent on the *hupSLC* promoter only. A possible explanation for these results is that the *hupD* gene is transcribed from the *hupSLC* promoter and from a secondary promoter located upstream of *hupD*. This secondary, weaker promoter would allow a low level of transcription of this gene, and in merodiploid strains such as AL13(pAL39) and AL13(pAL63) carrying multiple copies of *hupD*, its activity would result in higher levels of *hupD* gene product and, consequently, higher levels of hydrogenase activity.

The apparent expression of the *hupD* gene in strains in which the transcription of *hupC* is prevented by the Tn*5* insertions, such as in AL39(pAL63), AL63, AL63(pAL63), and AL63(pAL39), accounts for the MB-dependent hydrogenase activities observed in these mutants. Since none of them exhibited any detectable capacity to oxidize H_2 with O_2 as terminal electron acceptor, a role of *hupC* in electron transfer from hydrogenase to O_2 is predicted. In addition to this role, *hupC* may also participate in the formation or stabilization of the hydrogenase complex. This assessment is based on the low levels of MB-dependent hydrogenase activity obtained when all copies of the *hupC* gene in merodiploid strains are affected by the Tn*5* mutations compared with the much higher values of activity observed in strains, such as AL13(pAL39) and AL13(pAL63), carrying at least one unaffected copy of *hupC*.

Identification of *hupS* promoter. The 5′ end of the *hupS* RNA transcript was determined by primer extension using an oligonucleotide complementary to the coding strand of *hupS* gene as a primer and total RNA from bacteroids produced by *R. leguminosarum* bv. viciae UPM791. The base A located at position 99 (56 bases upstream of *hupS*) was identified as the transcriptional initiation site (Fig. 4).

The analysis of the nucleotide sequence upstream of the transcriptional start site of the *hupS* gene identified the sequence 5′-CTGGTCCGCTTCTTGCA-3′ between positions 71 and 87. This sequence is remarkably similar to the NtrA-binding consensus motif (5′-CTGGYRYRN$_4$TTGCA-3′) between positions −12 and −24 in *nif* promoters (12). The GC bases at positions 85 and 86 are located 13 nucleotides upstream of the transcriptional start site, which agrees quite well with the position of the GC conserved couple of bases of the NtrA-binding site and strongly suggests that transcription of this *hup* operon is dependent on a σ^{54} (NtrA) factor. The NtrA-binding motif was also present 44 bp upstream of ORF31, suggesting that the operon starting with the *hupG* gene contains an NtrA-dependent promoter. However, the transcriptional start site for this putative operon has not been identified yet.

Upstream of the NtrA-binding site, a sequence (5′-TATCAACTATTTA-3′) showing good homology with the integration host factor (IHF)-binding motif 5′-WATCAA nnnnTTR-3′ (4) was identified between nucleotides 31 and 43 (Fig. 2). In *nif* promoters, the IHF protein bends the DNA to mediate the contact of upstream binding activator proteins with the RNA polymerase-σ^{54} complex (15). The putative IHF sequence in the *hupS* promoter is preceded by an AT-rich stretch of nucleotides. This trait is also present in *nif* promoters.

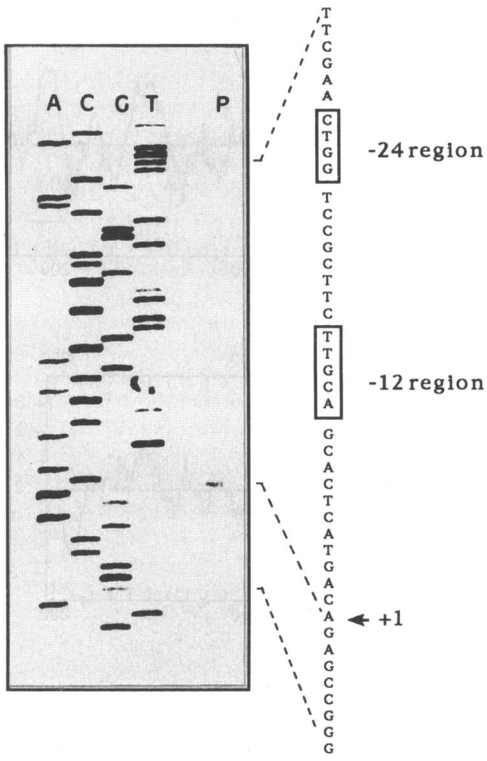

FIG. 4. Determination of transcriptional start point of *hupS* by primer extension. Lanes A, C, G, and T show the sequence of plasmid pRLH21 obtained with the primer 5′-CCCTGGCGGC GAATGAC-3′ (complementary to bases 183 to 201 in Fig. 2); lane P shows the extension product from the same primer hybridized with total RNA from *R. leguminosarum* UPM791 pea bacteroids.

To investigate the presence of a potential promoter immediately upstream of *hupD*, a primer extension experiment was carried out with the oligonucleotide 5′-GCGCCGGCG GAGGTCCCA-3′ complementary to the coding strand (bases 20 to 38) of *hupD* and total RNA from UPM791 bacteroids. In this experiment, no extension product indicative of a possible transcription initiation site was detected (data not shown).

Features of the potential gene products. The properties of the predicted gene products from *hupC* to *hupF* are listed in Table 2.

(i) **HupC.** The protein predicted from ORF13 (HupC) is very hydrophobic (51% nonpolar amino acids) and contains a high percentage of aromatic amino acid residues (17.8%) and no cysteines. The hydropathy profile of HupC shows four highly hydrophobic regions (I to IV) that could correspond to potential transmembrane segments (Fig. 5). When the hydrophobicity index for each amino acid was calculated as an average of 19 neighboring residues, a value higher than 1.6 was obtained for several amino acids within regions II, III, and IV. The critical value of 1.6 has been used to predict membrane-spanning domains in a large number of membrane proteins (18).

HupC has significant homology with proteins predicted

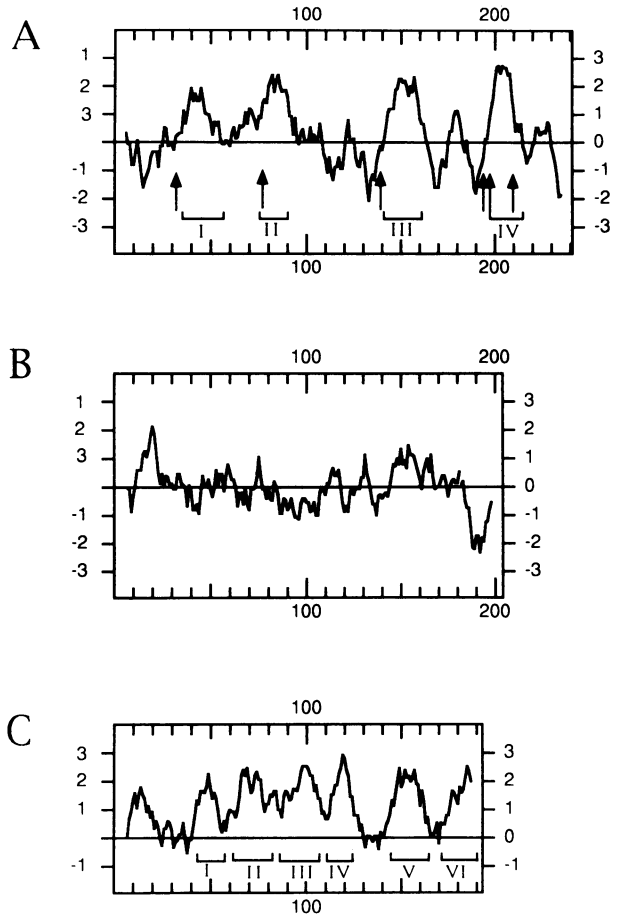

FIG. 5. Hydropathy profiles of HupC (A), HupD (B), and HupE (C). The Kyte-Doolittle algorithm (18) was used, and the hydropathy of an average of 11 amino acid residues was plotted as a function of amino acid number. Horizontal segments in panels A and C indicate predicted membrane-spanning domains and are designated by roman numbers. Vertical arrows in panel A show the positions of histidines conserved in HupC and in HupC-homologous proteins from other bacteria.

from ORFs that follow the hydrogenase structural genes in other bacteria (Fig. 6). The overall amino acid identity was 58.7% with ORF3 from *A. vinelandii* (26), 48.8% with HyaC from *E. coli* (27), 56.6% with ORFX from *R. capsulatus* (32), and 66.6% with 111 amino acids translated from the nucleotide sequence located immediately downstream of *hupL* in *B. japonicum* (37). Besides, the hydropathy profile of HupC has a striking similarity to that published for *E. coli* HyaC (27). Interestingly, six histidines (positions 31, 74, 136, 191, 194, and 208 in HupC) are conserved in the four proteins (Fig. 6). Three of these histidines are located at the beginnings of the potential membrane-spanning regions I, II, and IV. Histidines within membrane-spanning domains have been described as heme ligands in cytochromes of the *b* type (11).

(ii) HupD. The potential protein from ORF14 has a predicted size of 22.1 kDa and is mostly hydrophilic except for an 11-amino-acid stretch in the N terminus that exhibits an average hydrophobic index higher than 2 according to the Kyte-Doolittle scale (18). This hydrophobic domain may serve to anchor the protein to the membrane. HupD has homology to HyaD from the *E. coli* hydrogenase 1 operon (46% identity) (27) and to a sequence of 35 amino acid residues translated from DNA located immediately downstream of the structural genes for the periplasmic hydrogenase of *Desulfovibrio fructosovorans* (48% identity) (33) (Fig. 7). No significant similarity to any other protein included in the SwissProt data base was identified.

(iii) HupE. ORF15 encodes a potential protein of 191 amino acids with a predicted size of 19 kDa. The first 21 amino acid residues have the structure of a leader peptide sequence (47), with two lysines at positions 2 and 5, a hydrophobic stretch, and a putative cleavage site after the alanine at position 21. A striking feature of HupE protein is its extreme hydrophobicity (72% nonpolar amino acid residues), with six potential membrane-spanning segments (I to VI) (Fig. 5).

HupE showed no significant homology to any of the gene products from the *E. coli hya* operon. A search for similarity with proteins listed in the SwissProt data base revealed that HupE displayed a limited degree of similarity to chain 4 of NADH-UB oxidoreductases from maize (38) and tobacco (40) chloroplasts. These proteins contain 515 and 509 amino acids, respectively, and multiple membrane-spanning do-

```
Rl HupC     1                                    MTIHETLAADAHGEKAIERQSVYVYEAPVRIWHWINAFSILTLALTGYFIGSPLPSVP.GEASANFLMGYIRFIHFAAG
Av ORF3     1                                    *AL*KS*ETGDGQ**VRKQTA*******L*L***VT*L**VV*GV******A***TM*.***MD*Y*******A*****
Ec HyaC     1                                          MQQ*SDNVV*H**F*********LTVLCMAV*MV******K*****S.***TYL*Y*****L***S**
Rc ORFX     1  MKGVSDERINAPVRGPDEIFEASRLTGDATREDL*SIRR*T**********V***V**LA*TI*VV*****A*****MQI***TDQ*V******A*****
Bj          1                MMDAVAPASDARPDLASRDAS**R*VG*PT**********C**V*****IV*MV***L**T***T*A.****D**V******A*****

Rl HupC    79  QLLAVFLILRVYWAFVGNVHARQIFYVPFWSGRFWKE.WLHEVGWYTFLVRQPKKYVGHNPLPQFTMFLMFTLPLLFMAITGFALYSEGAGRDSWEYSLF
Av ORF3    79  YV**IGFLG*********H***EL*L**VHRKAW***L*.***R**L**EKT****I*****G*LA**CF*VVGAV**SV******A**L*****ADR**
Ec HyaC    68  MVFT*V*LM*I*******RYS*EL*I**V*RKSW*QGV*Y.*IR**L**AKR*SADI****IA*AA**GY*LMS.V**I*********HSQYAIFAPFRY
Rc ORFX   101  VVMS*AFFG*I*******R**W*M**I*IFNK*Y***FVF.*LR**F**EEE****I*****AHAA**TFI**GIT**M***W***A****QGGVTD***
Bj         88  *V****FLT*IL******H*S***

Rl HupC   178  GWVFSIWP...NSQDIHTYHHLGMWVILVFVMVHIYVAVREDIMSRQSIISSMISGERLFK.DRED
Av ORF3   178  ***IPLFG...Q***V**W******YLV******V*L*******V*******T*VG*W*M**D**P*
Ec HyaC   166  .V*EFFYWTGG**M***SW*R****L*GA**IG*V*M*L******DDTV**T*VN*Y*SH*FGKISNKERS
Rc ORFX   200  ***LGYVQ...***RL**L******A*VI*AII***A*****V*****MV*T*********D**IE
```

FIG. 6. Comparison of the deduced amino acid sequences from *R. leguminosarum hupC* with sequences of homologous proteins deduced from genes involved in hydrogen metabolism in other bacteria. Asterisks indicate conserved amino acid residues (relative to the *R. leguminosarum* bv. viciae sequence). Vertical arrowheads indicate the positions of conserved histidines in *R. leguminosarum* bv. viciae HupC and homologous proteins from other bacteria. Abbreviations: Av, *A. vinelandii* (26); Ec, *E. coli* (27); Rc, *R. capsulatus* (32); Bj, *B. japonicum* (amino acid sequence was translated from DNA located downstream of *hupL* [37]).

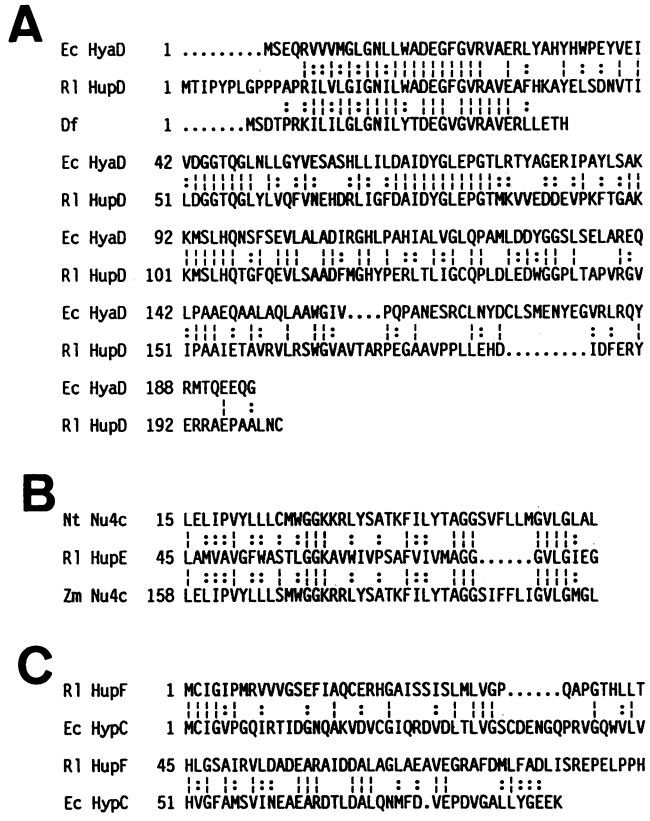

FIG. 7. Homology of *R. leguminosarum* HupD (A) and HupF (C) to proteins involved in hydrogen metabolism in other bacteria and homology of a fragment of HupE (B) to a fragment of chain 4 from chloroplast NADH-UB oxidoreductases (Nu4c) from plants. Symbols: ¦, identical residues; :, conservative change of residue. Abbreviations: Ec, *E. coli* (HyaD [27] or HypC [24]); Df, *D. fructosovorans* (amino acid sequence was translated from DNA downstream from hydrogenase structural gene *hupL* [33]); Nt, *Nicotiana tabacum* (40); Zm, *Zea mays* (38).

mains. The homology to HupE extends over a stretch of 250 residues in the central part of the protein. The segment exhibiting the highest homology is shown in Fig. 7B.

(iv) HupF. The last ORF in region *hup*I/II codes for a small protein (99 amino acid residues, 10.4 kDa) with no significant similarities to proteins listed in the SwissProt data base. However, HupF exhibited homology to the predicted protein from *hypC* gene of *E. coli* (33% identity) (Fig. 7C). The gene products from the *hyp* operon are required for the synthesis of the three hydrogenase isoenzymes of *E. coli*, but no specific functions have been assigned to them (25).

DISCUSSION

The region *hup*I/II of the H_2 uptake gene cluster cloned in cosmid pAL618 from *R. leguminosarum* bv. viciae UPM791 contains the hydrogenase structural genes (22). These genes were previously sequenced and designated *hupS* and *hupL* (14). They are contiguous and oriented in the same direction, with the gene for the small subunit (*hupS*) preceding the gene for the large subunit (*hupL*). In this study, we have determined the nucleotide sequence of a DNA fragment of 2,554 bp located immediately downstream of *hupL* and described the gene organization of the complete *hup*I/II DNA region. The analysis of the sequence revealed the presence of four additional ORFs (ORF13 to ORF16) that follow *hupL* and have the same orientation as the structural genes. When expressed in *E. coli* cells, these ORFs encoded polypeptides of molecular sizes which agree quite well with those of the proteins predicted from the nucleotide sequences. Several arguments support the idea that the gene products from these ORFs play a role in H_2 uptake by pea bacteroids; consequently, the corresponding genes have been designated *hupC* to *hupF*. First, Tn*5* insertions within ORF13, ORF14, and ORF15 impair the symbiotic Hup phenotype. The effect is not due to the polarity of the Tn*5* mutation, since the phenotype alteration was different depending on the location of the Tn*5* insertion (Table 3). Second, analysis of the primary structures of the gene products from ORF13 to ORF16 showed that at least three of these products have significant sequence similarity to genes belonging to operons involved in hydrogenase activity in other bacteria (Fig. 6 and 7).

Tn*5* complementation analysis clearly showed that the structural genes *hupS* and *hupL*, together with *hupC* and *hupD* are transcribed from a promoter located upstream of *hupS*. Similarly, the ORF3 gene product from the *R. capsulatus hup* cluster has recently been shown to belong to the same operon as the hydrogenase structural genes (3). Although there is no direct evidence that *hupE* and *hupF* belong to the same operon as *hupSLCD*, the facts that *hupS* to *hupF* are closely linked and oriented in the same direction and that no promoterlike sequences were identified in the 5′ flanking regions of *hupD* to *hupF* favor the idea of a single operon (*hupSLCDEF*).

The transcription initiation site of the putative *hup SLCDEF* operon was located in the 5′ noncoding region of *hupS* (Fig. 4). The analysis of the sequence upstream of this transcription initiation site revealed the presence of NtrA (σ^{54} factor)- and IHF-binding consensus motives also found in *nif* and *fix* genes (12). An IHF-like protein from *R. capsulatus* has recently been shown to specifically bind the promoter of the hydrogenase structural genes in this bacterium (45). The existence of NtrA and IHF motives is coherent with the observation that the *hup*I/II region is activated only in symbiosis with peas (31) and suggests a coregulation of *nif* and *hup* genes in *R. leguminosarum* bv. viciae. However, no obvious NifA-binding consensus sequences were found upstream of the NtrA site. Sequences showing homology to the NtrA-dependent promoters between positions −24 and −12 have also been found in the promoters of the *hyc* operon coding for the hydrogenase 3 components of *E. coli* (24) and the *hoxFUYH* operon coding for the NADH-reducing hydrogenase of *Alcaligenes eutrophus* (46).

The allocation of precise functions in hydrogenase synthesis or activity to any of the gene products from *hupC* to *hupF* is difficult from the available data. However, three lines of evidence support the idea that HupC is a *b*-type cytochrome. First, the fact that mutant AL63 carrying a Tn*5* insertion within *hupC* shows some H_2 uptake activity with MB but not with O_2 as electron acceptor suggests that HupC is involved in electron transfer to O_2 (Table 3). A mutation in ORF3 (*hupM*) from *R. capsulatus* also resulted in the inability of the cells to transfer electrons to O_2 (3). Second, the hydropathy profile of the amino acid sequence deduced from *hupC* predicted a secondary structure characteristic of an integral membrane protein with four potential membrane-spanning domains. The presence of transmembrane domains is a common characteristic of *b*-type cytochomes (49). In these

integral membrane proteins, several histidine residues located within membrane-spanning regions are important because they serve as heme ligands (11). A comparison of HupC with *b*-type cytochromes included in the SwissProt data base revealed that HupC had little overall homology (less than 25% identity; data not shown) with *b*-type cytochromes. However, a detailed analysis of the sequences around the histidines located in the membrane-spanning regions in HupC and homologous proteins revealed that a stretch of 10 amino acids preceding the histidine at position 74 is highly similar to the conserved motif [(D, E, N, or Q)xxxG(F, Y, or W)RxxH] included in the PROSITE data base (1) as specific for the cytochrome *b* component of both the respiratory chain complex III (b-c_1 complex) and the photosynthetic b_6-f complex. The histidine of this motif has been identified as the first ligand of one of the two hemes present in b-b_6 cytochromes (49). The cytochrome b-b_6 signature is located at the beginning of membrane-spanning segment II in both HupC and b-b_6 cytochromes. Third, it has recently been demonstrated that the gene product from *Wolinella succinogenes hydC*, a gene that follows the hydrogenase structural genes in this bacterium, is a cytochrome *b* that reacts with quinones and anchors the enzyme to the membrane (17). This protein has 26% amino acid identity with *R. leguminosarum* bv. viciae HupC, and both proteins present a very similar hydropathy profile (16).

The pathway of electron transfer from hydrogen to O_2 in vegetative and symbiotic cells of *B. japonicum* has been examined in detail (9, 30). The electrons from H_2 seem to be injected into the electron transport chain at the level of UB (7, 29). However, the primary electron acceptor from hydrogenase has not been clearly identified. From difference spectrum analysis, a *b*-type cytochrome (component b_{559}) was proposed to be specifically involved in H_2 oxidation by *B. japonicum* vegetative and symbiotic cells and to function as an electron carrier between hydrogenase and UB (7, 8). Later on, the existence of a low-potential cytochrome solely reducible by H_2 has been questioned, and instead, a cytochrome b_{558} (cytochrome *b'* or cytochrome *o*) located after the UB was shown to be not specifically reduced by H_2 in *B. japonicum* cells (28). These findings are consistent with the idea of HupC being a cytochrome *b* involved in H_2 oxidation in *R. leguminosarum*. The location of the *hupC* gene in the same operon as the hydrogenase structural genes strongly suggests that HupC is specifically required for the physiological oxidation of H_2 and points toward it being the primary electron acceptor of the *R. leguminosarum* bv. viciae hydrogenase. However, the possibility that HupC can also receive electrons from substrates other than H_2 cannot be discarded. The precise location and function of HupC should await the construction and physiological characterization of nonpolar HupC$^-$ mutants.

With the available data, we can only speculate about the possible roles of the *hupD*, *hupE*, and *hupF* gene products. HupD is a hydrophilic protein that seems to be essential for hydrogenase activity in pea bacteroids. It may be involved in processing or assembling the hydrogenase components. HupE showed a limited degree of similarity to chain 4 of the NADH-UB oxidoreductase from chloroplasts (Fig. 7B). The presence of components related to the NADH-UB oxidoreductase complex in the *R. leguminosarum* bv. viciae hydrogenase system should not be surprising, since it has been shown in *B. japonicum* that UB is the respiratory chain component which accepts the electrons from H_2 (7, 29). Moreover, recent work pointed out that several subunits of the NADH-UB oxidoreductase complex have significant sequence similarities to subunits of the hydrogenase 3 (*hyc* operon) from *E. coli* (2).

It is possible that the hydrogenase enzyme and the products from the *hupCDE* genes form a membrane protein complex (H_2-UB oxidoreductase complex) that reduces UB to ubiquinol with H_2. The function of this complex would be equivalent to that of the specific NADH- or flavin-linked dehydrogenases that also reduce UB with hydrogen-containing substrates such as NADH or succinate.

ACKNOWLEDGMENTS

This work has been financed by grants from Comisión Interministerial de Ciencia y Tecnología (projects PBT87-0029 and BIO090-0206). We are grateful to Centro Nacional de Biotecnología for providing us with computer assistance and access to sequence data bases.

REFERENCES

1. **Bairoch, A.** 1991. PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acids Res. **19:**2241–2245.
2. **Böhm, R., M. Sauter, and A. Böck.** 1990. Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. Mol. Microbiol. **4:**231–243.
3. **Cauvin, B., A. Colbeau, and P. M. Vignais.** 1991. The hydrogenase structural operon in *Rhodobacter capsulatus* contains a third gene, *hup*M, necessary for the formation of a physiologically competent hydrogenase. Mol. Microbiol. **5:**2519–2527.
4. **Craig, N. L., and H. A. Nash.** 1984. *E. coli* integration host factor binds to specific sites in DNA. Cell **39:**707–716.
5. **d'Aubenton Carafa, Y., E. Brody, and C. Thermes.** 1990. Prediction of Rho-independent *Escherichia coli* transcription terminators. J. Mol. Biol. **216:**835–858.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
7. **Eisbrenner, G., and H. J. Evans.** 1982. Carriers in the electron transport from molecular hydrogen to oxygen in *Rhizobium japonicum* bacteroids. J. Bacteriol. **149:**1005–1012.
8. **Eisbrenner, G., and H. J. Evans.** 1982. Spectral evidence for a component involved in hydrogen metabolism of soybean nodule bacteroids. Plant Physiol. **70:**1667–1672.
9. **Evans, H. J., A. R. Harker, H. Papen, S. A. Russell, F. J. Hanus, and M. Zuber.** 1987. Physiology, biochemistry and genetics of the uptake hydrogenase in rhizobia. Annu. Rev. Microbiol. **41:**335–361.
10. **Ford, C., N. Garg, R. P. Garag, K. H. Tibelius, M. G. Yates, D. J. Arp, and L. C. Seefeldt.** 1990. The identification, characterization, sequencing and mutagenesis of the genes (*hup*SL) encoding the small and large subunits of the H_2-uptake hydrogenase of *Azotobacter chroococcum*. Mol. Microbiol. **4:**999–1008.
11. **Friden, H., and L. Hedrstedt.** 1990. Role of His residues in *Bacillus subtilis* cytochrome b_{558} for haem binding and assembly of succinate:quinone oxidoreductase (complex II). Mol. Microbiol. **4:**1045–1056.
12. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20:**567–592.
13. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155:**156–165.
14. **Hidalgo, E., A. Leyva, and T. Ruiz-Argüeso.** 1990. Nucleotide sequence of the hydrogenase structural genes from *Rhizobium leguminosarum*. Plant Mol. Biol. **15:**367–370.
15. **Hoover, T. R., E. Santero, S. Porter, and S. Kustu.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. Cell **63:**11–22.
16. **Kröger, A. (Institut für Mikrobiologie, J. W. Goethe-Universität, Frankfurt am Main, Germany).** 1991. Personal communication.
17. **Kröger, A., F. Droβ, and F. Theis.** 1991. The Ni/Fe-hydrogenase of *Wolinella succinogenes*, p. 201. Abstr. Third Int. Conf.

Mol. Biol. Hydrogenases.
18. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
20. **Leclerc, M., A. Colbeau, B. Cauvin, and P. M. Vignais.** 1988. Cloning and sequencing of the genes encoding the large and the small subunits of the H_2 uptake hydrogenase (*hup*) of *Rhodobacter capsulatus*. Mol. Gen. Genet. **214:**97–107.
21. **Leyva, A., J. M. Palacios, T. Mozo, and T. Ruiz-Argüeso.** 1987. Cloning and characterization of hydrogen uptake genes from *Rhizobium leguminosarum*. J. Bacteriol. **169:**4929–4934.
22. **Leyva, A., J. M. Palacios, J. Murillo, and T. Ruiz-Argüeso.** 1990. Genetic organization of the hydrogen uptake (*hup*) cluster from *Rhizobium leguminosarum*. J. Bacteriol. **172:**1647–1655.
23. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193:**265–275.
24. **Lutz, S., R. Böhm, A. Beier, and A. Böck.** 1990. Characterization of divergent NtrA-dependent promoters in the anaerobically expressed gene cluster coding for hydrogenase 3 components of *Escherichia coli*. Mol. Microbiol. **4:**13–20.
25. **Lutz, S., A. Jacobi, V. Schlensog, R. Böhm, G. Sawers, and A. Böck.** 1991. Molecular characterization of an operon (*hyp*) necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*. Mol. Microbiol. **5:**123–135.
26. **Menon, A. L., L. W. Stults, R. L. Robson, and L. E. Mortenson.** 1990. Cloning, sequencing and characterization of the [NiFe]hydrogenase-encoding structural genes (*hox*K and *hox*G) from *Azotobacter vinelandii*. Gene **96:**67–74.
27. **Menon, N. K., J. Robbins, H. D. Peck, Jr., C. Y. Chatelus, E.-S. Choi, and A. E. Przybyla.** 1990. Cloning and sequencing of a putative *Escherichia coli* [NiFe] hydrogenase-1 operon containing six open reading frames. J. Bacteriol. **172:**1969–1977.
28. **O'Brian, M. R., and R. J. Maier.** 1985. Expression of cytochrome *o* in hydrogen uptake constitutive mutants of *Rhizobium japonicum*. J. Bacteriol. **161:**507–514.
29. **O'Brian, M. R., and R. J. Maier.** 1985. Role of ubiquinone in hydrogen-dependent electron transport in *Rhizobium japonicum*. J. Bacteriol. **161:**775–777.
30. **O'Brian, M. R., and R. J. Maier.** 1988. Hydrogen metabolism in *Rhizobium*: energetics, regulation, enzymology and genetics. Adv. Microb. Physiol. **29:**1–52.
31. **Palacios, J. M., J. Murillo, A. Leyva, and T. Ruiz-Argüeso.** 1990. Differential expression of hydrogen-uptake (*hup*) genes in vegetative and symbiotic cells of *Rhizobium leguminosarum*. Mol. Gen. Genet. **221:**363–370.
32. **Richaud, P., P. M. Vignais, A. Colbeau, R. L. Uffen, and B. Cauvin.** 1990. Molecular biology studies of the uptake hydrogenase of *Rhodobacter capsulatus* and *Rhodocyclus gelatinosus*. FEMS Microbiol. Rev. **87:**413–418.
33. **Rousset, M., Z. Dermoun, C. E. Hatchikian, and J. P. Bélaich.** 1990. Cloning and sequencing of the locus encoding the large and small subunits of the periplasmic [NiFe] hydrogenase from *Desulfovibrio fructosovorans*. Gene **94:**95–101.
34. **Ruiz-Argüeso, T., F. J. Hanus, and H. J. Evans.** 1978. Hydrogen production and uptake by pea nodules as affected by strains of *Rhizobium leguminosarum*. Arch. Microbiol. **116:**113–118.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
37. **Sayavedra-Soto, L. A., G. K. Powell, H. J. Evans, and R. O. Morris.** 1988. Nucleotide sequence of the genetic loci encoding subunits of *Bradyrhizobium japonicum* uptake hydrogenase. Proc. Natl. Acad. Sci. USA **85:**8395–8399.
38. **Schantz, R., and L. Bogorad.** 1988. Maize chloroplast genes *ndh*D, *ndh*E, and *psa*C. Sequences, transcripts and transcript pools. Plant Mol. Biol. **11:**239–247.
39. **Schneider, C. G., H. J. Schmitt, C. Schild, H. V. Tichy, and W. Lotz.** 1990. DNA sequence encoding the two structural genes for uptake hydrogenase of *Rhizobium leguminosarum* bv. *viciae* B10. Nucleic Acids Res. **18:**5285.
40. **Shinozaki, K., M. Ohme, T. Wakasugi, N. Hayashida, T. Matsubayashi, N. Zaita, J. Ghunwongse, J. Obokata, K. Yamaguchi-shinozaki, C. Ohto, K. Torazawa, B. Y. Meng, M. Sugita, H. Deno, T. Kamogashira, K. Yamada, J. Kusuda, F. Takaiwa, A. Kato, N. Tohdoh, H. Shimada, and M. Sugiura.** 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. **5:**2043–2049.
41. **Short, J. M., J. M. Fernández, J. A. Sorge, and W. D. Huse.** 1988. Lambda ZAP: a bacteriophage lambda expression vector with *in vivo* excision properties. Nucleic Acids Res. **16:**7583–7600.
42. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
43. **Stults, L. W., F. Moshiri, and R. J. Maier.** 1986. Aerobic purification of hydrogenase from *Rhizobium japonicum* by affinity chromatography. J. Bacteriol. **166:**795–800.
44. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
45. **Toussaint, B., C. Bosc, P. Richaud, A. Colbeau, and P. M. Vignais.** 1991. A mutation in a *Rhodobacter capsulatus* gene encoding an integration host factor-like protein impairs in vivo hydrogenase expression. Proc. Natl. Acad. Sci. USA **88:**10749–10753.
46. **Tran-Betcke, A., U. Warnecke, C. Böcker, C. Zaborosch, and B. Friedrich.** 1990. Cloning and nucleotide sequences of the genes for the subunits of NAD-reducing hydrogenase of *Alcaligenes eutrophus* H16. J. Bacteriol. **172:**2920–2929.
47. **von Hejne, G.** 1985. Signal sequences. The limits of variation. J. Mol. Biol. **184:**99–105.
48. **Wada, K., Y. Wada, H. Doi, F. Ishibashi, T. Gojobori, and T. Ikemura.** 1990. Codon usage tabulated from the GenBank genetic sequence data. Nucleic Acids Res. **19:**1981–1986.
49. **Widger, W. R., W. A. Cramer, R. G. Herrmann, and A. Trebst.** 1984. Sequence homology and structural similarity between cytochrome *b* of mitochondrial complex III and the chloroplast b_6-*f* hemes in the membrane. Proc. Natl. Acad. Sci. USA **81:**674–678.