JOURNAL OF BACTERIOLOGY, Apr. 1989, p. 1932–1941
0021-9193/89/041932-10$02.00/0

Vol. 171, No. 4

Identification of a Gene Linked to *Rhizobium meliloti ntrA* Whose Product Is Homologous to a Family to ATP-Binding Proteins

LISA M. ALBRIGHT, CLIVE W. RONSON,† B. TRACY NIXON,‡ AND FRED M. AUSUBEL*

Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

Received 4 November 1988/Accepted 25 January 1989

The *ntrA* gene of *Rhizobium meliloti* has recently been identified and shown to be required for a diverse set of metabolic functions (C. W. Ronson, B. T. Nixon, L. M. Albright, and F. M. Ausubel, J. Bacteriol. 169:2424–2431, 1987). As a result of sequencing the *ntrA* gene and its flanking regions from *R. meliloti*, we identified an open reading frame directly upstream of *ntrA*, ORF1, whose predicted product is homologous to a superfamily of ATP-binding proteins involved in transport, cell division, nodulation, and DNA repair. The homology of ORF1 to this superfamily and its proximity to *ntrA* led us to investigate its role in symbiosis by mutagenesis and expression studies. We were unable to isolate an insertion mutation in ORF1, suggesting that ORF1 may code for an essential function. We identified the start of transcription for the *ntrA* gene in vegetative cells and bacteroids and showed that ORF1 and *ntrA* are transcriptionally unlinked. ORF1 appears to be in an operon with one or more upstream genes.

Conservation throughout a region about 250 amino acids in length has been noted in a number of proteins involved in transport, cell division, nodulation, and DNA repair in bacteria (16, 23). This approximately 25-kilodalton domain includes two short stretches of amino acids which are also conserved in several ATP-binding proteins and which together are thought to form the nucleotide-binding pocket (47). We report here the discovery in *Rhizobium meliloti* of a new member of this superfamily.

Several members of the superfamily, MalK, HisP, PstB, RbsA, OppD, and OppF, are components of periplasmic binding-protein-dependent transport systems. These systems characteristically are composed of a periplasmic substrate-binding protein and an inner membrane complex. The membrane complex consists of MalK or its analog, which is hydrophilic, and two hydrophobic proteins (2). It has been proposed that this type of transport system is directly energized by hydrolysis of an intracellular high-energy phosphate bond, probably ATP or a related nucleotide (2, 6, 7); however, a role for the electrochemical proton gradient has not been ruled out (for discussions, see references in 2 and 24). Although the OppD, HisP, and MalK proteins have each been shown to bind ATP or an analog (23, 24, 27), no demonstration of ATP hydrolysis by any of these proteins has yet been made.

Several other members of the superfamily appear to function in transport processes distinct from periplasmic binding-protein-dependent systems. The HlyB protein functions in export of hemolysin in *Escherichia coli*. Unlike periplasmic binding-protein-dependent transport systems, only two proteins, HlyB and HlyD, are required directly for export of activated hemolysin (32). Recently, the products of the *ndvA* gene in *R. meliloti* and its functional homolog in *Agrobacterium tumefaciens*, *chvA*, have been shown to have homology to HlyB not only in its C-terminal ATP-binding domain but also in its N-terminal region (44; G. Cangelosi, personal communication). *ndvA* mutants are unable to produce extracellular β-(1,2)-glucan, leading to the proposal that *ndvA* functions in the export of β-(1,2)-glucan in *R. meliloti* (44). The HlyB subfamily has a eucaryotic member in the Mdr1 protein of multidrug-resistant mammalian cell lines. Resistance in these cells is associated with decreased intracellular drug accumulation resulting from increased energy-dependent drug efflux (for a discussion, see reference 12).

Two other members of the superfamily, NodI and FtsE, have unknown functions. FtsE is required for cell division in *E. coli* and is associated with the inner membrane (22). *nodI* mutants of *Rhizobium leguminosarum* are delayed in nodulation on pea plants (18).

While many of the members of the superfamily are involved in transport, at least one protein with homology to this superfamily, the product of the *uvrA* gene, is clearly not involved in transport (16). UvrA is a subunit of the ABC excision nuclease in *E. coli*, an ATP-dependent DNA repair enzyme. A complex series of insertions and duplications in the *uvrA* gene has resulted in separation of each half of the nucleotide-binding pocket by putative zinc fingers, implicated in DNA and metal binding by UvrA (16).

The apparent diversity of the functions in which this family of proteins is involved has led others to suggest that the role of the conserved region is to couple ATP hydrolysis to a variety of distinct biological processes, including transport (23).

The new member of the ATP-binding protein superfamily that we have discovered is encoded by an open reading frame, ORF1, directly upstream of the *ntrA* gene in *R. meliloti*. The *ntrA* gene product of *R. meliloti* is required for transcription of *nif* and some *fix* genes during symbiotic nitrogen fixation in bacteroids within root nodules of alfalfa. NtrA is also required for expression of C4-dicarboxylic acid transport genes and nitrate metabolism genes in free-living cells (39). In *E. coli* and *Salmonella typhimurium*, NtrA has been shown to be a sigma factor for RNA polymerase (26, 28).

ORF1 was identified after the *ntrA* gene and its flanking

* Corresponding author.
† Present address: BioTechnica International, Cambridge, MA 02140.
‡ Present address: Department of Molecular and Cellular Biology, Pennsylvania State University, University Park, PA 16802.

TABLE 1. Strains, bacteriophage, and plasmids used

Bacterial strains, bacteriophage, or plasmid	Relevant characteristics	Source or reference
Strains		
E. coli		
MC1061	*ara leu lac gal hsdR rpsL*	11
C2110	*polA* Nal^r	29
MM294	*pro thr endA hsdR supE*	M. Meselson
R. meliloti 1021	Wild-type Str^r	35
Bacteriophage λTn*5*	λ *b*221 *c*I857 Tn*5*	5
Plasmids		
GEM1	SP6, T7 RNA polymerase promoters; Amp^r	Promega Biotec
pLA140	Joint replicon of pRK290 and pNtr3.5BE:*ntrA*::Tn*5*; Tet^r Amp^r Kan^r	This study
pLA156	pNtrA10 with ORF1::Tn*5*-2; Tet^r Kan^r	This study
pLA162	3.5-kbp *Bam*HI-*Eco*RI fragment from pNtr3.5BE in GEM1; T7 transcript is antisense; Amp^r	This study
pLA172	*Stu*I fragment containing *ntrA* promoter in GEM1; T7 transcript is antisense; Amp^r	This study
pLA173	*Bam*HI-*Stu*I fragment containing 5′ end of ORF1 in GEM1; SP6 transcript is antisense; Amp^r	This study
pLA174	ORF1-β-galactosidase fusion in pMC1403; junction downstream of both GTGs	This study
pLA177	ORF1-β-galactosidase fusion in pMC1403; junction between two GTGs	This study
pLA178	ORF1-β-galactosidase fusion in pMC1403; junction upstream of both GTGs	This study
pMC1403	β-galactosidase fusion vector; Amp^r	10
pNtr3.5BE	3.5-kbp *Bam*HI-*Eco*RI containing ORF1 and *ntrA* in pUC8; Amp^r	39
pNtrA10	IncP, pLAFR1 cosmid clone containing ORF1 and *ntrA*; Tet^r	39
pPH1JI	IncP Spc^r Gen^r	8
pRK290	IncP, Tet^r	15
pRK2013	Conjugal helper; *rep*(ColE1) Kan^r	15

regions from *R. meliloti* were sequenced. The homology of ORF1 to the ATP-binding protein superfamily and its proximity to *ntrA* led us to investigate its role in symbiosis by mutagenesis and expression studies. In this work, we report the sequence of ORF1, our attempts to mutagenize the gene, and a transcriptional picture of the ORF1-*ntrA* region.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media. The strains and plasmids used in this study are described in Table 1. Rich media were LB for *E. coli* and LB supplemented with 2.5 mM $CaCl_2$ and 2.5 mM $MgSO_4$ (LBCaMg) or TY for *R. meliloti* (46). Defined medium contained RDM salts and vitamins and was buffered with 50 mM 2[*N*-morpholino]ethanesulfonic acid as described previously (39) and is referred to below as MES. Antibiotics were added to media at the following concentrations (μg/ml): for *E. coli*, tetracycline, 15; kanamycin, 50; ampicillin, 50; nalidixic acid, 20; gentamycin, 20; for *R. meliloti*, tetracycline, 5; neomycin, 50; spectinomycin, 50; gentamycin, 25; and streptomycin, 300.

Enzymes and chemicals. Restriction enzymes were purchased from New England BioLabs, Inc., Beverly, Mass., or Boehringer Mannheim Biochemicals, Indianapolis, Ind. T4 DNA ligase, T4 polynucleotide kinase, and Klenow fragment of *E. coli* DNA polymerase I were from New England BioLabs. Calf intestinal alkaline phosphatase and RNasin were from Boehringer Mannheim. AMV reverse transcriptase was from Life Sciences, Inc. St. Petersburg, Fla. Ribonucleotides and deoxyribonucleotides were from Pharmacia, Inc., Piscataway, N.J.

DNA manipulations. Small- and large-scale plasmid preparations were made by alkaline lysis (9). Methods for isolation of *R. meliloti* genomic DNA (35) and labeling of DNA probes by random primer extension (19) have been described. Standard methods (3, 33) were used for restriction digestions, gel electrophoresis, purification of DNA fragments from agarose gels, ligations, 5′ end labeling with T4 kinase, fill-outs of 5′ overhangs with Klenow fragment and deoxyribonucleoside triphosphates, and Southern filter hybridization.

DNA sequencing. DNA sequencing was carried out by the dideoxy-chain termination method by using a set of nested deletions in M13mp18, as previously described (39). Compressions were resolved by replacing dGTP with 7-deaza-dGTP (Boehringer Mannheim) in equimolar amounts in all mixes (38) or by inclusion of 40% formamide in the sequencing gel (34). In some cases, synthetic oligonucleotides were designed and used to prime synthesis next to a region of compression.

Sequence analysis. DNA sequences were compiled by using the DBAUTO and DBUTIL programs (42) and were analyzed by using programs from the University of Wisconsin, Madison (14). Initial homology searches of the National Biomedical Research Foundation (NBRF) protein data base were made with the FASTP program with k-tuple values of 1 and 2; pairwise comparisons were made with the ALIGN program (13) using the mutation data matrix with a bias of +2 and a gap penalty of 16. The scores given represent the number of standard deviations separating the maximum score of the real sequences from the mean of the maximum scores of 100 random permutations of the two sequences (13).

Plasmid constructions. Templates for probe preparations were constructed as follows. The 3.5-kilobase-pair (kbp) *Bam*HI-*Eco*RI fragment from pNtr3.5BE, containing ORF1 and *ntrA* (39), was cloned into GEM1 to yield pLA162. The

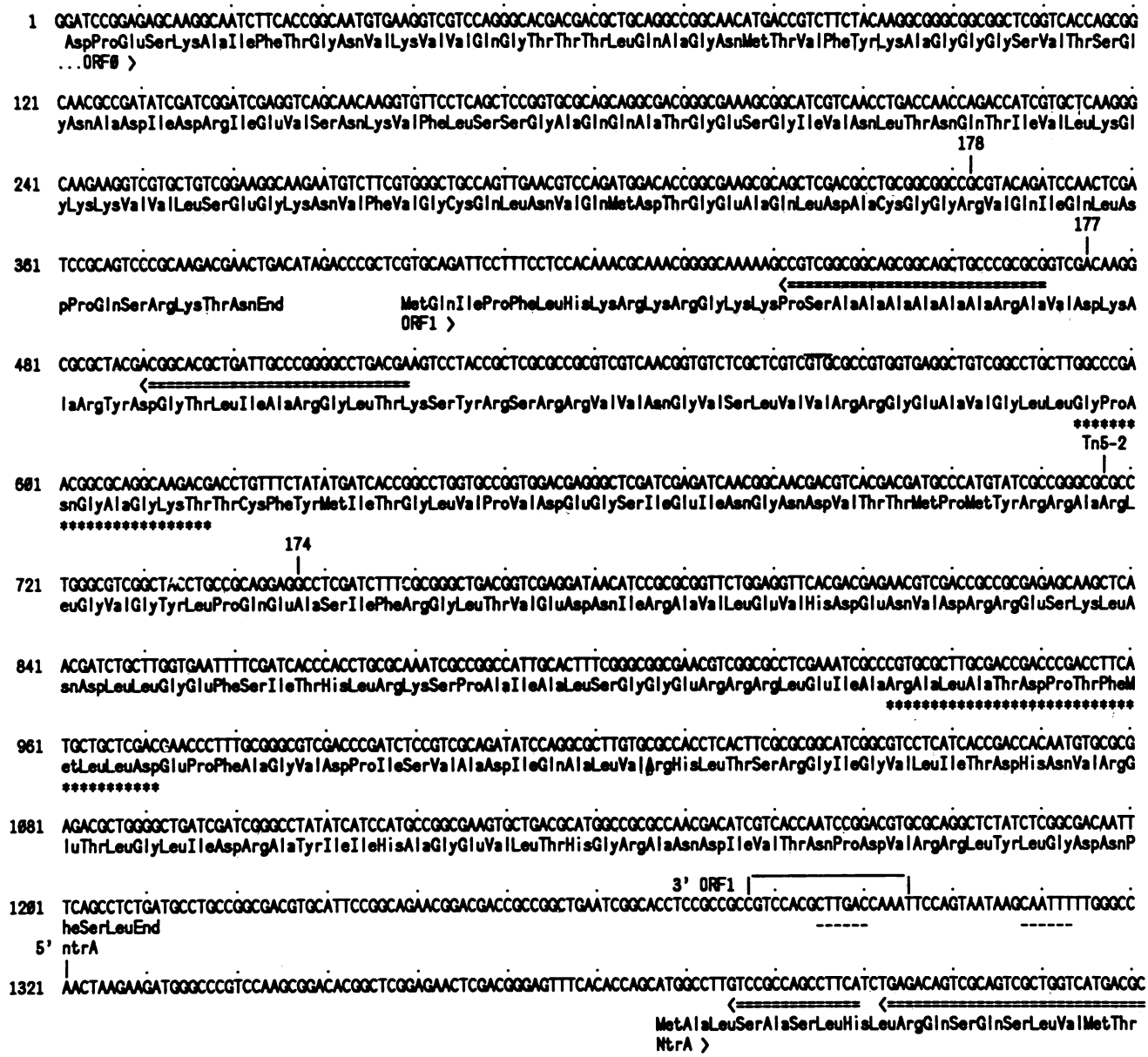

FIG. 1. Nucleotide sequence of ORF1 and flanking regions, with predicted translation products. The second GTG at nucleotide 564 in ORF1 is overlined. Regions of ORF1 containing nucleotide-binding pocket homology are underlined with asterisks. The numbers 178, 177, and 174 denote locations of fusions to *lacZ* as described in the text. The vertical bar is over the nucleotide at which the sequence diverges from that of ORF1. Tn*5*-2 denotes the location of the Tn*5* insertion. Vertical brackets indicate the 3′ end of the ORF1 transcript, with measurement error. Leftward arrows denote oligonucleotides used for transcript mapping. The 5′ end of the *ntrA* transcript at nucleotide 1321, mapped by primer extension, is indicated by the vertical bar. Potential −35 and −10 regions of the *ntrA* promoter are shown by a dashed underline.

*Stu*I fragment extending from nucleotides 747 to 1685 (Fig. 1) was cloned into the *Sma*I site of GEM1 to yield pLA172. The *Bam*HI-*Stu*I fragment from nucleotides 1 to 747 was cloned between the *Bam*HI and *Hin*cII sites of GEM1 to yield pLA173.

Three in-frame fusions between the N terminus of the ORF1 reading frame and β-galactosidase were constructed as follows. pLA173 has an *Eco*RI site from the vector polylinker next to the *Bam*HI site at nucleotide 1 and a *Hin*dIII site next to the *Stu*I-*Hin*cII junction in ORF1 at nucleotide 747. pLA173 was digested with *Hin*dIII, filled out, and digested with *Eco*RI. The approximately 800-bp *Eco*RI-*Hin*dIII fragment was ligated to pMC1403 which had been digested with *Bam*HI, filled out, digested with *Eco*RI, and treated with calf intestinal phosphatase. This plasmid, pLA174, encodes a protein composed of the first 117 amino acids of ORF1 fused in-frame to β-galactosidase at amino acid 7 through a 6-amino-acid linker encoded by polylinker sequences from GEM1.

pLA173 was digested with *Sal*I at nucleotide 472 in ORF1, partially filled out with dCTP and dTTP, and then digested with *Eco*RI. The approximately 485-nucleotide fragment was ligated to pMC1403 which had been digested with *Bam*HI, partially filled out with dATP and dGTP, digested with

*Eco*RI, and then treated with calf intestinal phosphatase. The partial fill out of the *Sal*I and *Bam*HI ends created 2-base complementary overhangs. This construction, pLA177, fused ORF1 at amino acid 25 to β-galactosidase at amino acid 7.

To create pLA178, pLA173 was digested with *Eag*I at nucleotide 338, partially filled out with dGTP, digested with *Eco*RI, and ligated to pMC1403 which had been digested with *Xma*I, partially filled out with dCTP, digested with *Eco*RI, and treated with calf intestinal phosphatase. The partially filled out *Eag*I and *Sma*I sites, again, were complementary. This created a fusion between the ORF1 reading frame upstream of the first GTG codon to β-galactosidase at amino acid 6. The nucleotide sequences of the junctions of each of these constructions were confirmed by sequencing the plasmid DNA by using a commercially available primer (−40; cat. no. 1212; New England BioLabs) which is complementary to pMC1403 downstream of the junction points.

Tn*5* mutagenesis. pNtr3.5BE was mutagenized in *E. coli* with λTn*5* essentially as described previously (40). A single insertion in ORF1, Tn*5*-2, was identified by restriction mapping. Its precise location was determined by cloning the kanamycin resistance-encoding *Bam*HI fragment (which contained only one inverted repeat of Tn*5* and flanking sequences) of pNtr3.5BE::Tn*5*-2 in pUC19cm (41) and nucleotide sequencing of this construction by using a Tn*5*-specific primer (39).

Tn*5*-2 was recombined into pNtrA10, the cosmid clone encompassing the *ntrA* region (39), in the following manner. MC1061 which had been transformed with pNtr3.5BE::Tn*5*-2 and pNtrA10 was mated with C2110, selecting for resistance to nalidixic acid (C2110), tetracycline, kanamycin, and ampicillin. A single transconjugant was then screened for loss of resistance to ampicillin. Since C2110 is *polA*$^-$ and is therefore unable to maintain pNtr3.5BE::Tn*5*-2 as an independent replicon, this scheme first selected for cointegrates between the two plasmids and then screened for loss of vector sequences of pNtr3.5BE while maintaining Tn*5*-2 on pNtrA10. Restriction mapping of the plasmid from this strain showed that exchange took place via homologous recombination on either side of Tn*5*-2. This pNtrA10:ORF1::Tn*5*-2 is called pLA156.

Matings and homogenotizations used pRK2013 as the conjugal helper, pPH1JI as the chaser (40), and MM294 as the *E. coli* host for all plasmids, unless otherwise stated.

RNA isolation. RNA was prepared from log-phase cultures of *R. meliloti* 1021, grown in LBCaMg or in MES with 10 mM succinate or 0.4% glucose as the carbon source and 5 mM NH_4Cl as the nitrogen source, either by boiling in sodium dodecyl sulfate and CsCl gradient centifugation (30) or by lysing spheroplasts in SDS and diethylpyrocarbonate (3). RNA made by the latter method was further treated with 50 U of RNase-free DNase (Promega Biotec, Madison, Wis.) per ml in the presence of 500 U of RNasin per ml for 15 min at 37°C and then extracted with phenol:$CHCl_3$:isoamyl alcohol (25:24:1) and precipitated with ethanol. Total RNA from alfalfa seedlings or 4-week-old nodules induced by *R. meliloti* 1021 on alfalfa was the generous gift of R. Dickstein (this laboratory), and its preparation has been described previously (17).

RNase protection. Uniformly labeled RNA probes for RNase protection were synthesized as previously described (1). Each probe is designated by the plasmid template, the restriction site defining its 3′ end, and the RNA polymerase used to synthesize the RNA. RNase protection was done as described elsewhere (1) by using 5 to 10 fmol of probe and 1 to 10 μg of RNA. The nucleic acid was digested with 4 μg of RNaseA per ml and 78 U of RNase T1 per ml at 30°C. RNase concentrations were titrated to give maximum signal to background ratios.

Oligonucleotides. Oligonucleotides for primer extension mapping and synthesis of single-stranded DNA probes are shown in Fig. 1. Each is described in the text by the location of its 5′ end and polarity with respect to the sense strand of the message (+). Each oligonucleotide was size fractionated on 16% acrylamide, eluted by diffusion into H_2O, and purified by affinity chromatography on DE52 (48) and precipitation with 15 mM $MgCl_2$ and 6 volumes of ethanol. Oligonucleotides were 5′ end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase.

Primer extensions. Approximately 20 fmol of 5′-end-labeled oligonucleotide was hybridized overnight at 45°C with up to 30 μg of RNA in 20 μl of either 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); pH 6.5]–0.3 M NaCl–1 mM EDTA–(1409$^-$, 15-mer) or 40 mM PIPES (pH 6.5)–0.4 M NaCl–80% formamide–1 mM EDTA (1440$^-$, 30-mer). Hybrids were extended in 120 μl of 50 mM Tris hydrochloride (pH 8)–5 mM $MgCl_2$–5 mM dithiothreitol–50 μg of bovine serum albumin–50 mM NaCl–0.5 mM all four deoxyribonucleotide triphosphates–400 U of RNasin per ml with 35 U of AMV reverse transcriptase for 1 h at 42°C. The reaction was stopped by the addition of EDTA to 10 mM and RNase A to 17 μg/ml and by incubation at 37°C for 15 min. After extraction with phenol-$CHCl_3$ (1:1) in the presence of 3.3 M ammonium acetate, nucleic acids were precipitated with ethanol and tRNA carrier, suspended in 80% formamide containing tracking dyes, and subjected to electrophoresis in a 6% acrylamide sequencing gel (3). Markers were sequencing-reaction products, primed by the same oligonucleotides on the appropriate template by using [α-^{32}P]dATP. The oligonucleotides were phosphorylated with unlabeled ATP for these reactions.

S1 protection. pLA162 contains a *Hin*dIII site in polylinker sequences about 30 bp upstream of the *Bam*HI site at nucleotide 1. To make 5′-end-labeled DNA probes, DNA synthesis by Klenow fragment was primed by 5′-end-labeled oligonucleotides 470$^-$ or 519$^-$ by using *Hin*dIII-digested pLA162 as a template (3). The single-stranded DNA products were purified by electrophoresis in a 4% polyacrylamide sequencing gel (3, 30). Probe (2 to 2.5 fmol) was hybridized for 16 h at 37°C with up to 15 μg of RNA in 40 μl of 40 mM PIPES (pH 6.5)–0.4 M NaCl–1 mM EDTA–80% formamide. The remaining single-stranded nucleic acid was digested with 100 U of S1 nuclease for 1 h at 37°C in 250 μl of 30 mM sodium acetate (pH 5.2)–0.3 M NaCl–1.5 mM $ZnCl_2$–0.1 mM EDTA–20 μg of single-stranded salmon testes DNA per ml. The nucleic acids were extracted with phenol-$CHCl_3$ (1:1), precipitated with ethanol, and subjected to electrophoresis in a 4% polyacrylamide sequencing gel (3).

RESULTS

Nucleotide sequence of the region upstream of *ntrA*. The nucleotide sequence of DNA upstream of *R. meliloti ntrA* is presented in Fig. 1. This completes the sequence of the 3.5-kbp *Bam*HI-*Eco*RI fragment described by Ronson et al. (39), and the numbering system is the same as in that work. Using the positional base preference method, the program ANALYSEQ (43) predicted that two contiguous open reading frames upstream of the *ntrA* gene, ORF0 and ORF1, have a high likelihood of coding for protein (Fig. 1). ORF0 extends beyond the *Bam*HI site at nucleotide 1, and its sequence is thus incomplete.

TABLE 2. Pairwise alignment scores of ATP-binding family members

Gene product (amino acids)	Alignment score[a]									
	ORF1	MalK	HisP	OppD	OppF	PstB	RbsA (N)	RbsA (C)	NodI	HlyB
ORF1 (31–270)										
MalK (1–232)	17.11									
HisP (4–258)	17.26	19.21								
OppD (19–274)	9.49	13.95	13.93							
OppF (22–272)	10.47	16.32	17.41	29.98						
PstB (8–257)	13.69	13.86	21.06	14.83	13.81					
RbsA (N) (2–249)	15.13	14.19	11.76	8.77	11.71	10.48				
RbsA (C) (250–501)	13.99	10.46	12.48	12.45	12.00	11.52	16.43			
NodI (10–252)	20.11	13.87	17.24	10.79	9.75	10.93	16.15	12.16		
HlyB (467–707)	10.61	13.63	13.58	12.51	13.44	14.51	10.14	11.69	13.78	
FtsE (1–222)	12.15	15.70	20.44	11.10	14.47	12.04	12.95	13.11	13.07	15.37

[a] Pairwise alignment scores are in standard deviation units as described in Materials and Methods.

Similarity of ORF1 to a family of ATP-binding proteins. A search of the NBRF protein data base revealed that the predicted amino acid sequence of ORF1 is similar to the products of the *malK*, *hisP*, *oppD*, and *pstB* genes. These proteins function in periplasmic binding-protein-dependent transport systems for maltose, histidine, oligopeptides, and phosphate, respectively, in bacteria (2). The similarity of these gene products to a number of other proteins (which were not in the data base at the time of the search) involved in periplasmic binding-protein-dependent transport systems (*rbsA* and *oppF*) and in processes such as haemolysin export (*hlyB*), cell division (*ftsE*), and nodulation (*nodI*) was previously noted (2, 16, 23). Pairwise alignments of ORF1 to these gene products revealed statistically significant homology (Table 2). The similarity of this group of proteins extends over approximately 25 kilodaltons and includes two highly conserved regions which are thought to form the nucleotide-binding pocket in ATP-binding proteins (47; Fig. 2). In some proteins, the ATP-binding domain constitutes the entire protein, while in others, it forms a subunit of a substantially larger polypeptide (Fig. 2; for a summary, see reference 23).

ORF1 is conserved in location in at least two other procaryotic species. The DNA upstream of the *K. pneumoniae ntrA* gene (36) codes for a protein with homology to this family (Fig. 2). Partial sequence of the DNA upstream of the *ntrA* gene in *S. typhimurium* also indicates the presence of a homologous open reading frame (D. Popham and S. Kustu, personal communication).

A search of the NBRF protein data base and translated GenBank data base with the predicted partial amino acid sequence of ORF0 did not reveal statistically significant similarity to any member.

GTG initiation codon in ORF1. There is no ATG codon in the ORF1 reading frame upstream of the start of the homology to the family of ATP-binding proteins. There are two GTG codons in this region, which could function as translation initiation codons. Neither of these GTG codons has an appropriately spaced sequence with homology to ribosome-binding sites from *E. coli* (45). To identify which of these GTG codons is used to initiate translation, protein fusions of the N terminus of ORF1 to β-galactosidase were made with junctions downstream, in between, and upstream of the two GTG codons, as described in Materials and Methods (pLA174, pLA177, and pLA178, respectively).

In *E. coli* MC1061, the constructions with junctions in between or downstream of the two GTG codons showed low but reproducible β-galactosidase activity, while that with a fusion junction upstream of the GTG codons, in frame with ORF1 but out of frame with ORF0, showed no activity above the background (Table 3). This indicates that the first GTG codon can be used as a translation initiation codon in *E. coli*, and the start of ORF1 has been assigned on this basis, predicting a molecular size of 29,566 kDa. The ORF1 promoter is not located on the restriction fragments used to create the ORF1-β-galactosidase fusions (see below). Readthrough from vector sequences of pMC1403 is the likely source of transcription of the fusion genes and is one potential reason for the low activity of the fusion proteins. The lack of a strong ribosome-binding site suggests that translation of this protein in *E. coli* may be inefficient.

Mutagenesis of ORF1. To assess the function of ORF1 in free-living and symbiotic cells, pNtr3.5BE was targeted for mutagenesis by Tn*5* in *E. coli* as described in Materials and Methods, and a single insertion in ORF1 was identified, Tn*5*-2. A joint replicon of pNtr3.5BE::Tn*5*-2 and pRK290 was constructed by virtue of their single *Eco*RI sites, providing approximately 0.7 and 2.8 kbp of homology to the *R. meliloti* genome to each side of Tn*5*-2. By standard techniques (40), no recombinants which had exchanged Tn*5*-2 to the genome via homologous recombination were identified in several attempts. To provide more flanking homology, Tn*5*-2 was exchanged in *E. coli* from its location on pNtr3.5BE to that on pNtrA10 as described in Materials and Methods. This construction, pLA156, provided at least 15 kbp of homology to one side of Tn*5*-2 and at least 2.8 kbp to the other. pLA156 was mated into *R. meliloti* 1021, and marker exchange was attempted. *R. meliloti* resistant to neomycin (Tn*5*), spectinomycin, and gentamycin (pPH1JI, chaser) were found at a frequency of about 3×10^{-4}. Of 200 colonies tested, 199 were still resistant to tetracycline, indicating the continued presence of vector sequences of pLA156. Analysis of restriction digestions of the genomic DNA from the tetracycline-sensitive strain by filter hybridization indicated that Tn*5* had hopped elsewhere in the genome. Similar analysis of the tetracycline-resistant colonies indicated the presence of both wild-type and mutant copies of ORF1.

As a control for recombination in this region of the genome, pNtr3.5BE containing a Tn*5* insertion in the *ntrA* gene was joined to pRK290 at its single *Eco*RI site, yielding pLA140. This construction provided about 1.6 and 1.9 kbp of homology to the genome on either side of the Tn*5* insertion. In matings done in parallel with the one described for pLA156, *R. meliloti* resistant to neomycin, spectinomycin,

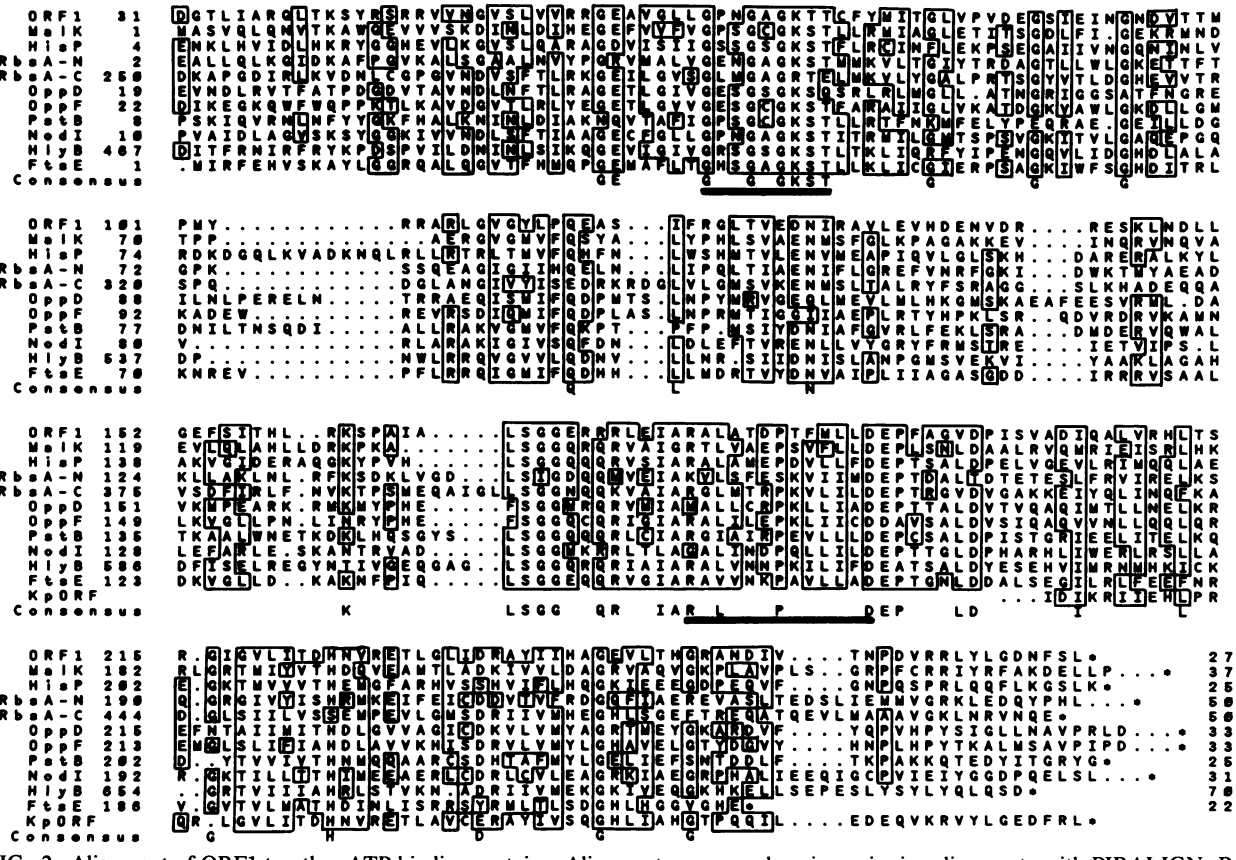

FIG. 2. Alignment of ORF1 to other ATP-binding proteins. Alignments were made using pairwise alignments with PIRALIGN. Regions with lower similarities were adjusted manually to minimize gapping. Consensus amino acids are those which were identical in at least eight proteins. Boxed amino acids show conservative substitutions in at least seven proteins according to the following amino acid groupings: HKR, QNED, PAGST, C, ILVM, YFW (20). Regions with similarity to the nucleotide-binding pocket are underlined (47). Sources of protein sequences are as follows: MalK, HisP, and OppD, NBRF protein data base (accession codes MMECMK, QREBPT, and QREBOT, respectively); RbsA, PstB, NodI, HlyB, and *K. pneumoniae* ORF, GenBank sequences under "Bacterial" (accession codes Ecorbs, Ecophowtu, Rhlnodij, Ecohly, and Kpnntra, respectively). The sequences for OppF (25) and FtsE (21) were entered by hand.

and gentamycin were obtained at a frequency of about 10^{-6}. Of 38 colonies tested, 33 were tetracycline sensitive, as expected of double recombinants. Two of the three analyzed by filter hybridization had the expected restriction digestion pattern of replacement of the wild-type copy of the *ntrA* gene with that containing the Tn*5* mutation. This control indicates that recombination in this region is not limiting, since a Tn*5* insertion in *ntrA*, provided with much less flanking DNA than Tn*5*-2 in ORF1, recombined readily into the genome by a double crossover.

TABLE 3. Activity of ORF1-β-galactosidase fusion proteins in MC1061[a]

Plasmid	Location of fusion	β-Galactosidase (U)
		0.2 ± 0.1 (2)
pMC1403		0.5 ± 0.0 (1)
pLA174	Downstream	2.5 ± 0.7 (2)
pLA177	In between	12.0 ± 1.4 (2)
pLA178	Upstream	0.35 ± 0.15 (2)

[a] Cultures were grown in LB, with ampicillin when appropriate, overnight at 37°C. The location of the fusion of ORF1 to *lacZ* is with respect to the two GTGs in ORF1 (see Materials and Methods and Fig. 1). The average units from duplicate samples, with the range of values, are according to Miller (37). The number of experiments is in parentheses.

Transcript mapping in the *ntrA*-ORF1 region. The proximity of ORF1 and *ntrA* prompted an investigation of transcription in that region. RNA was prepared from free-living *R. meliloti* 1021 or from nodules on alfalfa and analyzed by RNase protection as described in Materials and Methods. Two probes that differed in length at their 3′ ends but had the same 5′ ends were used to examine RNA derived from the region around the 3′ end of ORF1 and the start of the *ntrA* gene (Fig. 3b).

After hybridization to RNA from free-living cells, a strong signal of approximately 550 nucleotides in length was derived from the probe pLA172/HindIII/T7, while a strong signal of approximately 298 nucleotides in length was derived from the probe pLA172/SalI/T7 (Fig. 3a, lanes 2 to 5 and 10 to 13; Fig. 3b). These results indicate that the 3′ end of an RNA reading from ORF1 is located approximately at nucleotide 1290 ±10 in the sequence shown in Fig. 1. These 3′ ORF1 signals were virtually absent after hybridization with total nodule RNA (Fig. 3a, lanes 6 to 8 and 14 to 16).

Four signals derived from the 5′ ends of probes pLA172/HindIII/T7 and pLA172/SalI/T7 were seen in RNA from free-living cells (Fig. 3a, lanes 2 to 5 and 10 to 13; Fig. 3b). The highest-molecular-weight signal mapped a 5′ end for the *ntrA* transcript approximately at nucleotide 1310 ±10. No difference in the amount or ratio of 5′ *ntrA* and 3′ ORF1

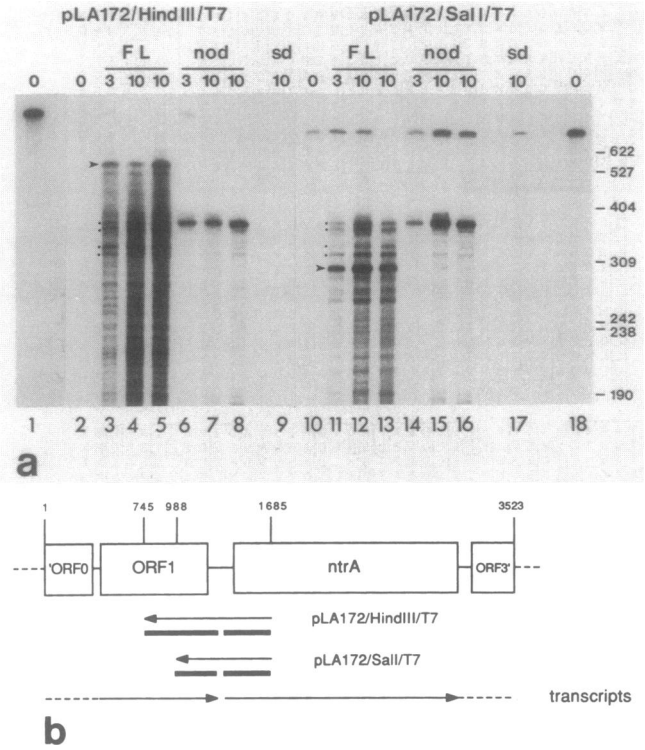

FIG. 3. Mapping ORF1 and *ntrA* transcripts by RNase protection. (a) Approximately 5 fmol of probes pLA172/HindIII/T7 (approximately 1.4×10^5 cpm/fmol of transcript; lanes 1 to 9) or pLA172/SalI/T7 (approximately 1.1×10^5 cpm/fmol of transcript; lanes 10 to 18) was hybridized with no RNA (lanes 1, 2, 10, and 18) or to RNA from vegetative *R. meliloti* 1021 grown in LBCaMg (FL; lanes 3 to 5 and 11 to 13), from nodules (nod; lanes 6 to 8 and 14 to 16), or from alfalfa seedlings (sd; lanes 9 and 17), digested with RNase, and subjected to polyacrylamide gel electrophoresis as described in Materials and Methods. The amount of RNA used in each reaction is shown (in micrograms) at the top of each lane. Samples in lanes 1 and 18 received no RNase treatment and were diluted 1,000-fold before electrophoresis. Samples in lanes 5, 8, 13, and 16 were treated with three times the usual amounts of RNases A and T1. The arrowhead points to the 3′ ORF1 signal; the four dots emphasize the four 5′ *ntrA* signals. Lengths of single-stranded DNA markers (in nucleotides) are shown on the right. (b) Schematic diagram of the experiment. Leftward arrows represent the uniformly labeled antisense RNA probes. Heavy lines represent the regions protected. Rightward arrows depict deduced transcripts reading through the region covered by the probes.

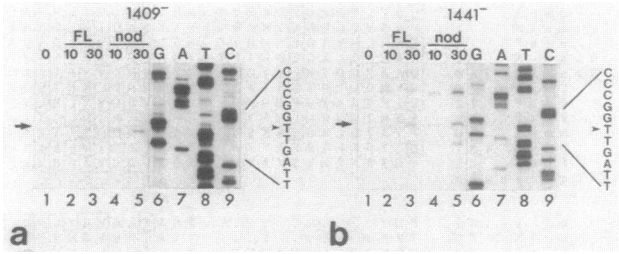

FIG. 4. Mapping of the 5′ end of the *ntrA* transcript by primer extension. Approximately 20 fmol of oligonucleotides 1409^- (3.7×10^4 cpm/fmol; panel a) and 1441^- (3.8×10^4 cpm/fmol; panel b) were hybridized with no RNA (lane 1) or with RNA from vegetative *R. meliloti* 1021 grown in LBCaMg (FL; lanes 2 and 3) or from nodules (nod; lanes 4 and 5) and extended with reverse transcriptase as described in Materials and Methods. The amount of RNA used in each reaction is shown (in micrograms) at the top of each lane. The same oligonucleotides were used to prime sequencing reactions on an appropriate template, yielding the sequence, shown to the right of each panel, of the strand complementary to that shown in Fig. 1. The arrow on the left of each panel points to the primer extension product given by both oligonucleotides. Lanes 1 to 5 are from a 10-day exposure; lanes 6 to 9 are from a 16-h exposure of the same autoradiogram.

signals was seen in RNA from cells grown in LBCaMg, TY, or MES using glucose or succinate as the carbon source and ammonia as the nitrogen source as described in Materials and Methods (data not shown). The highest-molecular-weight *ntrA* signal was predominant in nodule RNA; only trace amounts of the lower-molecular-weight *ntrA* signals were seen (Fig. 3a, lanes 6 to 8 and 14 to 16). No protection was seen by nonnodulated alfalfa seedling RNA (Fig. 3a, lanes 9 and 17), indicating that protection by the nodule RNA was from bacterial RNA.

If transcription were to proceed from ORF1 through *ntrA*, then protection of each probe pLA172/HindIII/T7 and pLA172/SalI/T7 along the entire length of its homology to the genome would occur. Because the RNA probes were nonhomologous to the genome at their 5′ ends, the resulting signal would be slightly shorter than the entire probe length. Although some full-length protection of the pLA172/SalI/T7 probe was evident after hybridization with free-living RNA (Fig. 3a, compare lanes 11 to 13 with lane 10), the entire length of the probe was resistant, indicating that it was not homology dependent. No full-length protection of the pLA172/HindIII/T7 probe was seen after hybridization with free-living RNA (Fig. 3a, compare lanes 3 to 5 with lane 2). Trace amounts of homology-dependent full-length protection of each probe were seen in samples hybridized with nodule RNA (not resolved from the homology-independent signal in the exposure shown in Fig. 3a). We conclude that there is no readthrough transcription to *ntrA* from ORF1 in free-living cells and that only a tiny fraction of *ntrA* transcripts in bacteroids is due to readthrough transcription from ORF1.

To confirm the presence and location of the 5′ end of the *ntrA* transcript, two oligonucleotide primers complementary to the beginning of the *ntrA* coding sequence (Fig. 1) were used to prime DNA synthesis by using free-living and nodule RNA as templates as described in Materials and Methods. One primer extension product, whose length indicated that the 5′ end of the *ntrA* transcript is located at nucleotide 1321, was seen in reactions primed by either oligonucleotide 1409^- or 1441^- and was present in both free-living and nodule RNA (Fig. 4a and b, lanes 1 to 5). Its location is consistent with that indicated by the 5′ signals mapped by RNase protection. The use of two different oligonucleotides served as a control for hybridization of either oligonucleotide to RNA from other regions of the genome. This is likely to be the cause of the other primer extension products which were unique to each oligonucleotide (Fig. 4a and b, lanes 1 to 5).

Since there was no evidence of more than one 5′ end for the *ntrA* transcript by primer extension analysis, it seems likely that the multiple signals mapping to the 5′ end of the probes in the RNase protection experiment are due to heterogeneity at the 3′ end of the RNA rather than at the 5′ end. Other RNA preparations from free-living cells which showed less rRNA degradation also showed four 5′ signals, but the lowest-molecular-weight signal was much weaker (data not shown). Therefore, the most likely explanation at this point for the four signals is RNA degradation, either

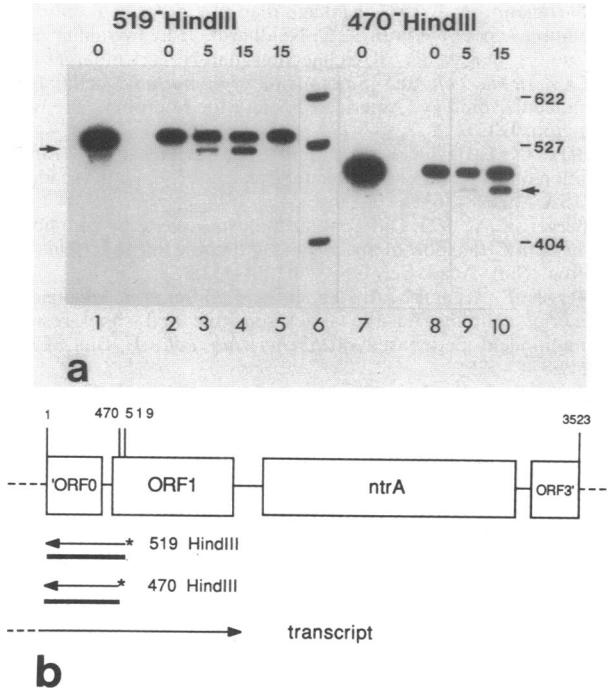

FIG. 5. S1 protection analysis of the ORF1 transcript. (a) Approximately 3 fmol of 519⁻HindIII (8.2×10^4 cpm/fmol; lanes 1 to 5) or 2 fmol of 470⁻HindIII (7.5×10^4 cpm/fmol; lanes 7 to 10) were hybridized with no RNA (lanes 1, 2, 7, and 8) or to RNA from vegetative *R. meliloti* grown in LBCaMg (lanes 3 to 5 and 9 to 10) and digested with S1 nuclease as described in Materials and Methods. The amount of RNA used in each reaction is shown (in micrograms) at the top of each lane. No S1 was added to the samples in lanes 1 and 7. The sample in lane 5 was treated with RNaseA before hybridization. Lane 6 contains DNA molecular weight markers; sizes (in nucleotides) are shown on the right. The arrows point to protection of each probe to the full extent of its homology with the *Rhizobium* RNA. (b) Schematic diagram of the experiment. The leftward arrows represent the 5′-end-labeled DNA probes. Heavy lines represent the region protected. The rightward arrow represents deduced transcript, which ends at the location previously mapped by RNase protection (Fig. 3).

intracellular or extracellular, creating molecules with heterogeneous 3′ ends.

The start of the ORF1 coding sequence begins 13 nucleotides downstream of the end of ORF0. To determine whether transcription of ORF0 and ORF1 are coupled, 5′-end-labeled, single-stranded DNA probes whose 5′ ends are located at nucleotides 519 and 470 were hybridized to free-living RNA and digested with S1 nuclease as described in Materials and Methods. Because each probe was not homologous to the *R. meliloti* genome for 30 nucleotides at its 3′ end due to polylinker sequences from the vector, transcription reading from ORF0 to ORF1 should give rise to a signal 30 nucleotides shorter than the full length of the probe. This was seen for each probe (Fig. 5, lanes 3, 4, 9, and 10). This protection was due to RNA rather than contaminating genomic DNA (which would give the same kind of signal) because a sample treated with RNase before hybridization did not show the homology-dependent protection (Fig. 5, lane 5). Only very minor amounts of shorter signals were seen, indicating that readthrough transcription from ORF0 is the major source of expression of ORF1.

DISCUSSION

We have identified a gene linked to *R. meliloti ntrA*, ORF1, which is homologous to a superfamily of ATP-binding proteins, many of which are involved in transport processes. ORF1 is unusual in that it begins with a GTG codon. In characterizing transcription in the region of ORF1 and *ntrA*, we identified the start of transcription for *ntrA* in vegetative cells and bacteroids. We also showed that ORF1 appears to be in an operon with one or more upstream genes. The ORF1 promoter remains unmapped at this time.

We have been unable to assign a function to ORF1, due to our inability to marker exchange an ORF1::Tn*5* insertion into the *R. meliloti* genome. This was not due to limiting recombination in the region, since an *ntrA*::Tn*5* insertion with much less flanking homology was homogenotized at an appreciable frequency. It is unlikely that the cosmid from which the ORF1::Tn*5* marker exchange was attempted had experienced a gross rearrangement, preventing recombination, since Southern analysis of the genome in the region of pNtrA10 indicated that they were colinear (data not shown). These results suggest, but do not prove, the ORF1 has an essential function in *R. meliloti*. Its conservation in both *S. typhimurium* and *Klebsiella pneumoniae* is consistent with a function which is not specifically symbiotic. Since ORF1 is conserved in location in *S. typhimurium*, it is likely to be conserved in location in *E. coli*. The one member of the superfamily which has been proven to be essential, FtsE, is located at 76 min on the *E. coli* K-12 linkage map, while *ntrA* is at 70 min (4); therefore it seems unlikely that ORF1 is the FtsE homolog of *R. meliloti*. In addition, *R. meliloti* ORF1 is no more homologous to FtsE than to any other members of the superfamily, with the exception of ORF1 from *K. pneumoniae* (Fig. 2 and Table 2), suggesting that ORF1 function is distinct from that of FtsE. Since the genetic system of *S. typhimurium* is better developed than that of *R. meliloti*, it would probably be easier to try to determine ORF1 function by isolating mutants in *S. typhimurium*, conditional or otherwise. At this point, we favor the notion that ORF1 encodes a well-conserved, probably essential, function in a number of bacterial species.

The tight linkage of ORF1 and *ntrA* in at least three bacterial species prompted us to determine if their expression is coupled transcriptionally. RNase protection experiments which detected two RNA molecules, one corresponding to the 5′ end of the *ntrA* gene and one corresponding to the 3′ end of ORF1, indicated that these genes are expressed independently (Fig. 4). We have not unequivocally shown that the 5′ end of the *ntrA* transcript arises from transcription initiation rather than processing of a longer message reading through from ORF1. Previous data indicated that DNA downstream of nucleotide 1263 is sufficient for complementation of *ntrA* function in free-living cells (39), in good agreement with the 5′ end mapping shown here. This indicates that processing of a larger transcript is not necessary to yield sufficient NtrA for free-living functions. Essentially all of the transcripts in bacteroids which read into the *ntrA* gene start at the same location as in free-living cells, which additionally argues that this promoter is sufficient for full symbiotic expression of *ntrA*. Moreover, the large difference in the ratio of the ORF1 3′ and *ntrA* 5′ signals in free-living versus nodule RNA argues against RNA processing as the source of the two separate RNAs. While ORF1 and *ntrA* expression are uncoupled, their light linkage and its conservation remain intriguing observations which may indicate some involvement of their functions.

The region upstream of the 5′ end of the *ntrA* transcript includes a sequence around −35 with respect to the start which is similar to consensus −35 sequences for *E. coli* promoters, whereas the region around −10 has poor homology to *E. coli* −10 sequences (Fig. 1). The promoters for the two δ-aminolevulinic acid synthetase transcripts in *R. meliloti* also each have good homology to −35 sequences but not to −10 sequences. The −10 regions in these promoters are purine rich, in contrast to that for the *ntrA* transcript, which is AT-rich (31).

The same start site for transcription of *ntrA* was seen in RNA from free-living cells and from nodules (Fig. 4 and 5). Since the RNA from nodules was approximately one-third bacterial in origin (based on rRNA band intensities [data not shown]), a given amount of nodule RNA represents proportionately fewer bacterial equivalents than the corresponding amount of free-living RNA. Thus, it appears that the steady-state level of *ntrA* transcripts is somewhat higher in nodule bacteria than in free-living cells (Fig. 3, compare lanes 14 to 16 with lanes 11 to 13). It is unknown whether this is due to induction of the *ntrA* promoter or to stabilization of the *ntrA* message.

The ratio of the 3′ ORF1 signal to the 5′ *ntrA* signal was much lower in nodule RNA than in RNA from free-living cells (Fig. 3, compare lanes 14 to 16 with lanes 11 to 13). This suggests that the relative level of ORF1 expression to *ntrA* expression is substantially lower in nodule bacteria than in free-living cells. The 3′ ORF1 transcript was barely detectable in nodule RNA. If our estimate of the bacterial content of the nodule RNA preparation is correct, then this implies that ORF1 expression is reduced in bacteroids. If ORF1 encodes an essential free-living function, it may represent a class of housekeeping genes whose expression is diminished or turned off in bacteroids.

Essentially all of the transcription reading into ORF1 initiates upstream of the *Bam*HI site at nucleotide 1 (Fig. 5). This indicates that ORF1 is in an operon with ORF0 and perhaps other genes. Many of the members of the superfamily are found in operons with other genes involved in the same process (for a summary, see reference 23). Further sequence analysis of the DNA upstream of ORF1 may provide insight into ORF1 function if those gene products are homologous to other proteins of known function.

ACKNOWLEDGMENTS

This work was supported by Hoechst AG. L.M.A. was a recipient of a fellowship from the Damon Runyon-Walter Winchell Cancer Fund during a portion of this work.

We thank Dan Voytas and John Smith of the Department of Molecular Biology at Massachusetts General Hospital for synthesis of the oligonucleotides; Andrew McKee for isolation of Tn*5* mutants in pNtr3.5BE; David Popham, Sydney Kustu, Gerry Cangelosi, and Gene Nester for communicating results prior to publication; and Eva Huala for thoughtful discussions and a careful and critical reading of the manuscript.

LITERATURE CITED

1. **Albright, L. M., M. F. Yanofsky, B. Leroux, D. Ma, and E. W. Nester.** 1987. Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. J. Bacteriol. **169**:1046–1055.
2. **Ames, G. F.-L.** 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. Annu. Rev. Biochem. **55**:397–425.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl (ed).** 1987. Current protocols in molecular biology, John Wiley & Sons, Inc., New York.
4. **Bachmann, B. J.** 1987. Linkage map of *Escherichia coli* K-12, edition 7, p. 807–876 *In* E. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
5. **Berg, D., J. Davies, B. Allet, and J.-D. Rochaix.** 1975. Transposition of R factor genes to bacteriophage. Proc. Natl. Acad. Sci. USA **72**:3628–3631.
6. **Berger, E. A.** 1973. Different mechanisms of energy coupling for the active transport of proline and glutamine in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **70**:1514–1518.
7. **Berger, E. A., and L. A. Heppel.** 1974. Different mechanisms of energy coupling for the shock-sensitive and shock-resistant amino acid permeases of *Escherichia coli*. J. Biol. Chem. **249**:7747–7755.
8. **Beringer, J. E., J. L. Beynon, A. V. Buchanon-Wollaston, and A. W. B. Johnston.** 1978. Transfer of the drug-resistance transposon Tn5 to *Rhizobium*. Nature (London) **276**:633–634.
9. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
10. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
11. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
12. **Choi, K., C. Chen, M. Kriegler, and I. B. Roninson.** 1988. An altered pattern of cross-resistance in multidrug-resistant human cells results from spontaneous mutations in the *mdr*1 (P-glycoprotein) gene. Cell **53**:519–529.
13. **Dayhoff, M. O., W. C. Barker, and L. T. Hunt.** 1983. Establishing homologies in DNA sequences. Methods Enzymol. **91**: 524–545.
14. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
15. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77**:7347–7351.
16. **Doolittle, R. F., M. S. Johnson, I. Husain, B. Van Houten, D. C. Thomas, and A. Sancar.** 1986. Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. Nature (London) **323**:451–453.
17. **Dunn, K., R. Dickstein, R. Feinbaum, B. K. Burnett, T. K. Peterman, G. Thoidis, H. M. Goodman, and F. M. Ausubel.** 1988. Developmental regulation of nodule-specific genes in alfalfa root nodules. Mol. Plant-Microbe Interact. **1**:66–74.
18. **Evans, I. J., and J. A. Downie.** 1986. The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins; nucleotide sequence analysis of the *nodI* and *nodJ* genes. Gene **43**:95–101.
19. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
20. **French, S., and B. Robson.** What is a conservative substitution? 1983. J. Mol. Evol. **19**:171–175.
21. **Gill, D. R., G. F. Hatfull, and G. P. C. Salmond.** 1986. A new cell division operon in *Escherichia coli*. Mol. Gen. Genet. **205**:134–145.
22. **Gill, D. R., and G. P. C. Salmond.** 1987. The *Escherichia coli* cell division proteins FtsY, FtsE and FtsX are inner membrane-associated. Mol. Gen. Genet. **210**:504–508.
23. **Higgins, C. F., I. D. Hiles, G. P. C. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell, and M. A. Hermodson.** 1986. A family of related ATP-binding subunits coupled to many distinct processes in bacteria. Nature (London) **323**:448–450.
24. **Higgins, C. F., I. D. Hiles, K. Whalley, and D. J. Jamieson.** 1985.

Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems. EMBO J. **4**:1033–1040.

25. **Hiles, I. D., M. P. Gallagher, D. J. Jamieson, and C. F. Higgins.** 1987. Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. J. Mol. Biol. **195**:125–142.
26. **Hirschmann, J., P.-K. Wong, K. Sei, J. Keener, and S. Kustu.** 1985. Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a sigma factor. Proc. Natl. Acad. Sci. USA **82**:7525–7529.
27. **Hobson, A., R. Weatherwax, and G. F.-L. Ames.** 1984. ATP-binding sites in the membrane components of histidine permease, a periplasmic transport system. Proc. Natl. Acad. Sci. USA **81**:7333–7337.
28. **Hunt, T. P., and B. Magasanik.** 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc. Natl. Acad. Sci. USA **82**:8453–8457.
29. **Kahn, M., D. Ow, B. Sauer, A. Rabinowitz, and R. Calendar.** 1980. Genetic analysis of bacteriophage P4 using P4-plasmid ColEl hybrids. Mol. Gen. Genet. **177**:399–412.
30. **Kassavetis, G. A., and E. P. Geiduschek.** 1982. Bacteriophage T4 late promoters: mapping 5′ ends of T4 gene 23 mRNAs. EMBO J. **1**:107–114.
31. **Leong, S. A., P. H. Williams, and G. S. Ditta.** 1985. Analysis of the 5′ regulatory region of the gene for δ-aminolevulinic acid synthetase of *Rhizobium meliloti*. Nucleic Acids Res. **13**:5965–5976.
32. **Mackman, N., J.-M. Nicaud, L. Gray, and I. B. Holland.** 1986. Secretion of hemolysin by *Escherichia coli*. Curr. Top. Microbiol. Immuno. **125**:159–181.
33. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Martin, R.** 1987. Overcoming DNA sequencing artifacts: stops and compressions. Focus (Bethesda Res. Lab.) **9**:8–10.
35. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. **149**:114–122.
36. **Merrick, M. J., and J. R. Gibbins.** 1985. The nucleotide sequence of the nitrogen regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial RNA polymerase sigma factors. Nucleic Acids Res. **13**: 7607–7620.
37. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
39. **Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel.** 1987. *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J. Bacteriol. **169**:2424–2431.
40. **Ruvkun, G. B., and F. M. Ausubel.** 1981. A general method for site-directed mutagenesis in prokaryotes. Nature (London) **289**: 85–88.
41. **Schmidhauser, T. J., and D. R. Helinski.** 1985. Regions of broad-host-range RK2 involved in replication and stable maintenance in nine species of gram-negative bacteria. J. Bacteriol. **164**:446–455.
42. **Staden, R.** 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. **10**:4731–4751.
43. **Staden, R.** 1984. Graphic methods to determine the function of nucleic acid sequences. Nucleic Acids Res. **12**:521–538.
44. **Stanfield, S. W., L. Ielpi, D. O'Brochta, D. R. Helinski, and G. S. Ditta.** 1988. The *ndvA* gene product of *Rhizobium meliloti* is required for β-(1,2)glucan production and has homology to the ATP-binding export protein HlyB. J. Bacteriol. **170**:3523–3530.
45. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2971–2996.
46. **Szeto, W. W., B. T. Nixon, C. W. Ronson, and F. M. Ausubel.** 1987. Identification and characterization of the *Rhizobium meliloti ntrC* gene: *R. meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. J. Bacteriol. **169**:1423–1432.
47. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gray.** 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **8**: 945–951.
48. **Wallace, R. B., M. Schold, M. J. Johnson, P. Dembek, and K. Itakura.** 1981. Oligonucleotide directed mutagenesis of the human β-globin gene: a general method for producing specific point mutations in cloned DNA. Nucleic Acids Res. **9**:3647–3656.