JOURNAL OF BACTERIOLOGY, Mar. 1987, p. 1120–1126
0021-9193/87/031120-07$02.00/0

Vol. 169, No. 3

Conservation of Structure and Location of *Rhizobium meliloti* and *Klebsiella pneumoniae nifB* Genes

WILLIAM J. BUIKEMA,[1,2]† JOHN A. KLINGENSMITH,[1] SARAH L. GIBBONS,[2] AND FREDERICK M. AUSUBEL[1]*

Department of Genetics, Harvard Medical School, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114,[1] and Department of Cellular and Developmental Biology, Harvard University, Cambridge, Massachusetts 02138[2]

Received 28 July 1986/Accepted 25 November 1986

Using transposon Tn*5*-mediated mutagenesis, an essential *Rhizobium meliloti* nitrogen fixation (*nif*) gene was identified and located directly downstream of the regulatory gene *nifA*. Maxicell and DNA sequence analysis demonstrated that the new gene is transcribed in the same direction as *nifA* and codes for a 54-kilodalton protein. In *Klebsiella pneumoniae*, the *nifBQ* operon is located directly downstream of a gene which is structurally and functionally homologous to the *R. meliloti nifA* gene. The DNA sequences of the *K. pneumoniae nifB* and *nifQ* genes (which code for 51- and 20-kilodalton proteins, respectively) were determined. The DNA sequence of the newly identified *R. meliloti* gene was approximately 50% homologous to the *K. pneumoniae nifB* gene. *R. meliloti* does not contain a gene homologous to *nifQ* directly downstream of *nifB*. The *R. meliloti nifB* product shares approximately 40% amino acid homology with the *K. pneumoniae nifB* product, and 10 of the 12 cysteine residues of the *R. meliloti nifB* product are conserved with 10 of the 17 cysteine residues of the *K. pneumoniae nifB* product.

Klebsiella pneumoniae is a free-living nitrogen-fixing species closely related to *Escherichia coli*, and *Rhizobium meliloti* fixes atmospheric dinitrogen in symbiotic association with various legumes including *Medicago sativa* (alfalfa). In *K. pneumoniae*, 17 contiguous nitrogen fixation (*nif*) genes have been identified and characterized (2); in *R. meliloti*, several genes have been identified which are specifically required for symbiotic nitrogen fixation, some of which are clustered as in *K. pneumoniae* (1, 4, 7, 9, 12, 24, 29, 31, 36). According to convention, a *Rhizobium* symbiotic gene that is structurally or functionally homologous to a *K. pneumoniae nif* gene is given the same gene designation as the *K. pneumoniae* gene. Only four *R. meliloti* symbiotic genes have been paired with *K. pneumoniae nif* genes: *nifHDK* encode the three nitrogenase polypeptides (29), and *nifA* encodes a positive regulator for *nif* gene expression (7, 36). A genetic and physical map of the *R. meliloti nifHDK-nifA* region is shown in Fig. 1.

We performed Tn*5* mutagenesis to characterize the region downstream of the *R. meliloti* 1021 *nifA* gene (to the right in Fig. 1) and obtained a Fix$^-$ (inability to symbiotically fix nitrogen) mutation which contains an insertion approximately 500 base pairs (bp) downstream of *nifA*. In a different *R. meliloti* strain, insertion mutations in the corresponding region also resulted in a Fix$^-$ phenotype (9). Two observations suggested that the region downstream of *nifA* in *R. meliloti* contains a gene homologous to *K. pneumoniae nifB*, a gene required for the processing of the nitrogenase iron-molybdenum cofactor (FeMoco) (26, 27). First, in *Rhizobium leguminosarum*, a putative *nifB* homolog (*fixZ*) is located about 450 bp downstream of *R. leguminosarum nifA* (28). Second, we found DNA sequences highly homologous to *R. leguminosarum fixZ* gene sequences downstream of the *R. meliloti nifA* gene (7; W. Buikema, unpublished data).

To determine whether the newly identified *R. meliloti* 1021 symbiotic gene downstream of *nifA* is homologous to *K. pneumoniae nifB*, we sequenced the *K. pneumoniae nifB* gene and approximately 2 kilobases of *R. meliloti* DNA downstream of *nifA*. We also sequenced the *K. pneumoniae nifQ* gene which is directly downstream of *nifB*.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1.

DNA biochemistry. Plasmid DNA preparation, restriction enzyme analysis, agarose gel electrophoresis, sonication of DNA restriction fragments, preparation of M13 sequencing templates, and dideoxy DNA sequencing were performed as described previously (3, 7). When necessary, specific restriction fragments were cloned into M13mp18 and M13mp19 and sequenced as above.

Tn*5* mutagenesis. Site-directed mutagenesis with transposon Tn*5*, marker exchange of the mutant DNA into the wild-type *R. meliloti* genome, and symbiotic characterization of the mutants were performed as described previously (30). Prospective recombinants were verified by Southern blot analysis of total genomic DNA digested with an appropriate restriction endonuclease.

Construction of promoter fusions. A transcriptional fusion of the *R. meliloti nifB* gene to the *lacZ* promoter of pUC13 was constructed with a 1.9-kilobase *Nru*I-*Sst*I fragment that contains the entire *nifB* gene. The *Nru*I site was converted to a blunt end with the Klenow fragment of DNA polymerase I (18), and the modified *Nru*I-*Sst*I fragment was ligated into *Sma*I-*Sst*I-digested pUC13 (37). The resulting plasmid, pWB50, was transformed into the *E. coli* maxicell strain CSR603.

Protein biochemistry. Expression of the *nifB* polypeptide in maxicells was performed as described previously (10).

* Corresponding author.
† Present address: Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637.

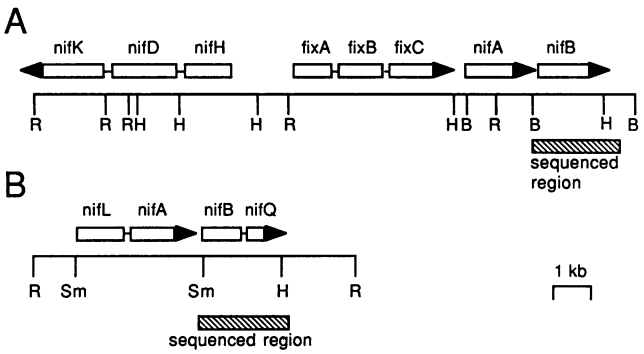

FIG. 1. Genetic and physical map of the *R. meliloti* (A) and *K. pneumoniae* (B) *nifB* regions. The orientation and position of the genes are shown above the lines. The hatched boxes show the regions sequenced for this study. Restriction nuclease sites: R, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; and Sm, *Sma*I. kb, Kilobase.

Computer analysis. Programs used for the input and assembly of sequence data, the analysis of DNA and amino acid sequence homologies, and restriction site mapping have been described previously (7, 11).

RESULTS

Genetic analysis. The *nifA* region of the *R. meliloti* 1021 megaplasmid that was mutagenized with Tn*5* is shown in Fig. 2. This region extends approximately 10 kilobases downstream of the *nifA* gene. Above the map are shown the positions of Tn*5* insertions examined. Only insertion no. 37 which maps about 500 bp 3′ to the *nifA* gene resulted in a mutant Fix$^-$ phenotype. In *R. meliloti* 102F34, Tn*5* insertions at a position comparable to insertion no. 37 also resulted in a Fix$^-$ phenotype (9).

DNA sequence analysis. In *R. leguminosarum*, a presumptive *nif* gene called *fixZ* was identified by Tn*5* mutagenesis; DNA sequence analysis showed that the 5′ end of *fixZ* is homologous to the 5′ end of *K. pneumoniae nifB* (28). The *R. leguminosarum fixZ* gene is situated adjacent to the 3′ end of *nifA*; in *K. pneumoniae*, *nifB* is situated directly downstream of *nifA*. To determine whether Fix$^-$ insertion no. 37 in *R. meliloti* 1021 inactivates a *nifB*-like gene, we determined the nucleotide sequence of the region shown in Fig. 1A. Because only a small portion of the *K. pneumoniae nifB* gene had been sequenced (J. Beynon, personal communication), we also determined the sequence of the *K. pneumoniae nifB* and *nifQ* genes (Fig. 1B). In each case, an appropriate restriction fragment was purified and sonicated, and the resulting fragments were cloned and randomly sequenced as previously described (7). In addition, specific restriction fragments were cloned and sequenced to complete small sections for which sequence data from only a single strand had been obtained by the random method. For both species, sufficient clones were sequenced to give completely overlapping, contiguous readings for both strands, averaging three to four readings at any position. The sequence data are displayed in Fig. 3 and 4.

For both *R. meliloti* and *K. pneumoniae*, a single unambiguous open reading frame (ORF) could be assigned to a presumptive *nifB* gene. The assignment of the AUG codon for the *K. pneumoniae nifB* ORF was based on the previously reported transcription initiation site (5) and was the next downstream AUG (which also had a suitable ribosome-binding site sequence preceding it) from this position. The AUG for the *R. meliloti nifB* ORF was chosen based on its position as the first AUG in the ORF, the immediate upstream placement of a likely ribosome-binding site sequence, the position of a potential *nif* promoter sequence 40 bases upstream, and its alignment with the *K. pneumoniae nifB* and the *R. leguminosarum fixZ* DNA sequences.

The *R. meliloti nifB* ORF measures 1470 bp and codes for a protein product of approximately 54 kilodaltons (kDa). The *K. pneumoniae nifB* ORF is 1,404 bp and codes for a polypeptide of 51 kDa. This latter value compares favorably with the published molecular sizes of 48 and 51.5 kDa for the *K. pneumoniae nifB* gene product (24, 33). A computer program that statistically determines and plots the probability that a reading frame is also a coding region (positional base preference method; 34), predicted the same ORFs as described above.

Within the region reported to contain the *K. pneumoniae nifQ* gene (16), we found an ORF of 501 bp. The UGA stop codon of *nifB* overlaps the AUG start codon of the presumptive *nifQ* gene by one base (Fig. 4). Finally, about 30 bp downstream of the UAG stop codon of the *nifQ* gene is a potential bidirectional transcription terminator. This sequence consists of an exact inverted repeat of seven G and C residues and five A and T residues, separated by four base pairs, and terminated in the downstream direction by four Ts and by six Ts in the upstream direction. This potential terminator has the characteristic structure of a rho-independent terminator and is similar in structure to other reported bidirectional terminators (8, 13).

Expression of *R. meliloti nifB* gene in maxicells. The size of the protein product of the *R. meliloti nifB* gene was determined in maxicells. A transcriptional fusion of the *nifB* gene was constructed in which the *lacZ* promoter of pUC13 was used to express the gene in the maxicell strain CSR603. The plasmid containing this construction, pWB50, expressed a protein product of 53 kDa which is presumably the product of the *R. meliloti nifB* gene (Fig. 5).

We attempted to use pWB50 to complement a *K. pneumoniae nifB* mutant (UN1712). However, in liquid culture under derepressing conditions, the presence or absence of pWB50 in *nifB* or wild-type strains had no effect on the levels of nitrogen fixation observed (data not shown).

Interspecies conservation of *nifB*. A comparison of the sequences of the *R. meliloti* and *K. pneumoniae nifB* genes revealed about 50 and 40% homology at the DNA and amino acid levels, respectively (Fig. 6). The most highly conserved

TABLE 1. Bacterial strains and plasmids

Strains and plasmids	Genotype and phenotype[a]	Source or reference
Strains		
Klebsiella pneumoniae KP1	Wild type	17
Klebsiella pneumoniae KP5614	*recA56 hisD hsdR* Δ*srl*	23
Klebsiella pneumoniae UN1712	*recA56 nifB srl*::Tn*10*	16
Rhizobium meliloti 1021	*str*	19
Escherichia coli CSR603	*recA1 uvrA6 phr-1*	32
Plasmids		
pWB50	*nifB* Ap^r	This work
pRMB8.3R	Δ*nifA nifB* Ap^r Tc^r	This work
pGR112	*nifLABQ* Tc^r	25
pUC13	Ap^r	37

[a] Abbreviations: Ap, ampicillin; Tc, tetracycline; r, resistance.

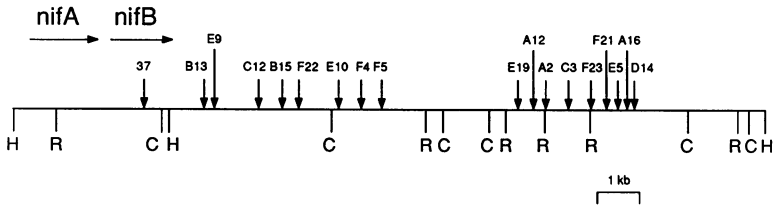

FIG. 2. Region of the *R. meliloti* megaplasmid that was mutagenized with Tn*5*. Positions of Tn*5* insertions are shown above the line. R, *Eco*RI; C, *Cla*I; and H, *Hin*dIII. kb, Kilobase.

region was found near the 5′ end of the genes, about 200 bp from the beginning of the *R. meliloti* gene. The *K. pneumoniae nifB* gene product contains a total of 17 cysteine residues, while the *R. meliloti nifB* gene product contains 12, of which 10 are conserved between the two genes. Several of these conserved cysteine residues are clustered near the amino ends of the proteins.

Comparison of the DNA sequences of the *R. meliloti* and *K. pneumoniae nifB* genes with the published sequence of the *R. leguminosarum fixZ* gene (28) revealed about 67 and 48% homology, respectively. When the *R. leguminosarum fixZ* gene amino acid sequence was compared with the amino acid sequences of the two *nifB* genes, however, it was apparent that there was less amino acid homology than

```
-198 GCGTCGCGAGCGTCGCCTCTTTTCTCCGTCCTTTCGAAACACGACCGGATGCAATTCAACTTTGCCCTTGTCGGAAGTCCGACAAGCCTGTGAGCAAAG -100

 -99 AGACAGGACCGAATCCAAAGACGCCAATCCATTGACGAAAAGGATTTGGCATAGCTGTTGCTGGTTGAATTGCAAAATCACCCTCCACTGGAGTTGTTC   -1

   1 ATGTCCACACCCATGATTTTGCGTGAGAGCCGGACCAGCACTACATTCTCTGACCAGTTGCTGGAGAACGCTAAATCGGTTGGCTGCTCACCCCCATCG   99
     M  S  T  P  M  I  L  R  E  S  R  T  S  T  T  F  S  D  Q  L  L  E  N  A  K  S  V  G  C  S  P  P  S

 100 ACGGCGCCGGGCGACATAGATCCTGGAACTTGGGACAAGATTAAGAATCACCCCTGTTTTTCAGAGGAGGCGCATCACTATTTCGCGCGCATGCACGTG  198
     T  A  P  G  D  I  D  P  G  T  W  D  K  I  K  N  H  P  C  F  S  E  E  A  H  H  Y  F  A  R  M  H  V

 199 GCGGTCGCGCCTGCCTGCAACATCCAATGTAACTATTGCAATCGCAAATACGATTGCGCCAACGAAAGTCGGCCCGGTGTTGCCTCGGAAAAGCTCACA  297
     A  V  A  P  A  C  N  I  Q  C  N  Y  C  N  R  K  Y  D  C  A  N  E  S  R  P  G  V  A  S  E  K  L  T

 298 CCTGACCAGGCGGTGCGAAAGGTGATTGCCGTTGCCAACGAAGTGCCTCAGCTGTCAGTGCTTGGCATCGCTGGGCCTGGCGATGCCTGTTACGACTGG  396
     P  D  Q  A  V  R  K  V  I  A  V  A  N  E  V  P  Q  L  S  V  L  G  I  A  G  P  G  D  A  C  Y  D  W

 397 AAGAAAACAAGGGCGACGTTCGAACGAGTGGCTAGGGAAATTCCCGACATAAGACTCTGCATCTCCACGAACGGGCTCTCGCTGCCGGACCATGTCGAT  495
     K  K  T  R  A  T  F  E  R  V  A  R  E  I  P  D  I  R  L  C  I  S  T  N  G  L  S  L  P  D  H  V  D

 496 GAGCTTGCCGAAATGAACGTCGATCACGTGACGATCACCATCAACATGGTCGATCCGCGTGTCGGCGTAAAGATCTACCCCTGGATTTACTATGGTCAG  594
     E  L  A  E  M  N  V  D  H  V  T  I  T  I  N  M  V  D  P  R  V  G  V  K  I  Y  P  W  I  Y  Y  G  Q

 595 CGCCGCCACACTGGTATCGACGCTGCGAGAATCCTGCACGAACGGCAGATGTTGGGCCTGGAGATGCTAGCCGAACGCGGCATCCTCACCAAGGTCAAC  693
     R  R  H  T  G  I  D  A  A  R  I  L  H  E  R  Q  M  L  G  L  E  M  L  A  E  R  G  I  L  T  K  V  N

 694 TCGGTAATGATCCCCGGCGTCAATGATGAGCACCTGATCGAAGTCAACAAAGTTGTGAAAGGAAGAGGCGCGTTGCTGCACAACGTAATGCCGCTAATT  792
     S  V  M  I  P  G  V  N  D  E  H  L  I  E  V  N  K  V  V  K  G  R  G  A  L  L  H  N  V  M  P  L  I

 793 TCAAACCGCATACACGGGACCTATTACGGACTGACAGGGCAGCGCGGCCCGGAGGCCTTCGAACTGCAGGCCCTTCAGGACCGTCTAGAAGGAACCAAA  891
     S  N  R  I  H  G  T  Y  Y  G  L  T  G  Q  R  G  P  E  A  F  E  L  Q  A  L  Q  D  R  L  E  G  T  K

 892 CTGATGCGTCATTGTCGACATTGCCGGGCCGATGCCATAGGCTTGCTCGGCGATGATCGTGGTCACGAGTTCACGCTCGCTGAAATCCCCGACGAGATA  990
     L  M  R  H  C  R  H  C  R  A  D  A  I  G  L  L  G  D  D  R  G  H  E  F  T  L  A  E  I  P  D  E  I

 991 ACCTACGATGCCAGCAAGCGACAGGCCTATCGCCAGTTGGTCGCGCGCGAACGCGGGGACCACCTAGTGGCCAAGAACGAGGCGAACAGAACGGTAATG 1089
     T  Y  D  A  S  K  R  Q  A  Y  R  Q  L  V  A  R  E  R  G  D  H  L  V  A  K  N  E  A  N  R  T  V  M

1090 TCGGTGGAATATGGCGGATCGCTTCTCATTGCCGTGGCGACCAAAGGCGGGGGCCGGATCAACGAACATTTTGGACACGCGAAAGAATTTCACGTTTAT 1188
     S  V  E  Y  G  G  S  L  L  I  A  V  A  T  K  G  G  G  R  I  N  E  H  F  G  H  A  K  E  F  H  V  Y

1189 ACCGTCTCCCAGAGAGGGATCAAGCTGGCAGGCCGCCGCAGGGTTGAGCAGTATTGCCTCGGCGGTTGGGGCGAGGTCGCCACCCTCGATCACATCGTC 1287
     T  V  S  Q  R  G  I  K  L  A  G  R  R  R  V  E  Q  Y  C  L  G  G  W  G  E  V  A  T  L  D  H  I  V

1288 GTTGCGCTTGAAGGAATAGACATCCTGCTCTGCGTCAAAATCGGAGATTACCCAAGGAAACAGCTGACACAGGCCGGGCTTCGAGCGACGGAAGCTTAT 1386
     V  A  L  E  G  I  D  I  L  L  C  V  K  I  G  D  Y  P  R  K  Q  L  T  Q  A  G  L  R  A  T  E  A  Y

1387 GGCCATGACTACATCGAGAGTGCGCTCGAAAGCTCTACGCCGCCGAGTTTGGCTATCGAACCACCGGTAAAGACGGCGACACGTTGAGCTGCCCTCCGA 1485
     G  H  D  Y  I  E  S  A  L  E  S  S  T  P  P  S  L  A  I  E  P  P  V  K  T  A  T  R  *

1486 CTGAATAAGGAGTTTAAAATGGCCTTCAAGATTATTGCA 1524
```

FIG. 3. Nucleotide sequence of the *R. meliloti nifB* gene and its amino acid translation product. The putative ribosome-binding site (nucleotide positions −7 to −10) and putative transcriptional regulatory sequences (see references 2 and 6) are underlined.

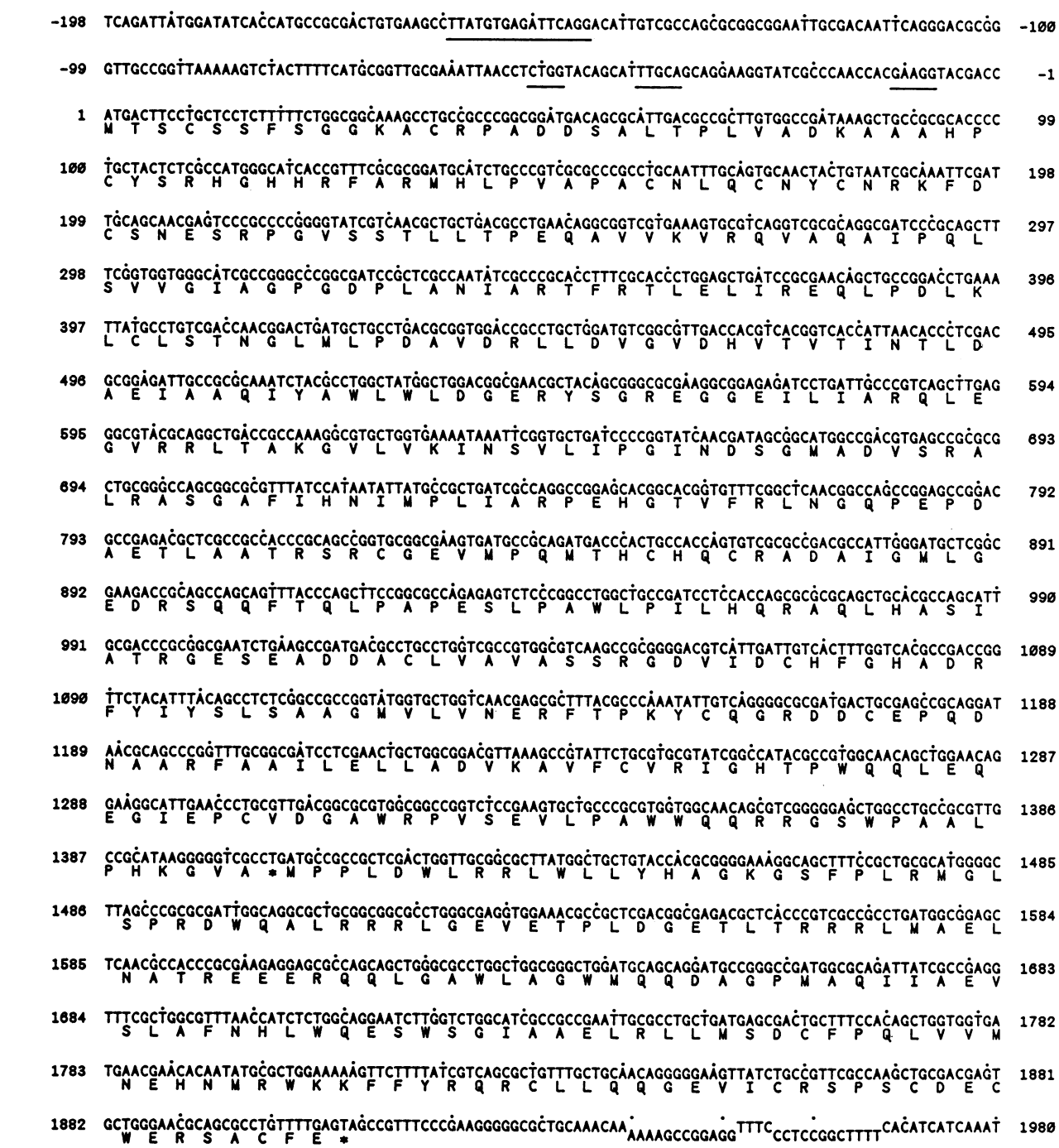

FIG. 4. Nucleotide sequences of the *K. pneumoniae nifB* and *nifQ* genes and their amino acid translation products. The putative ribosome-binding site (nucleotide positions −8 to −12) and putative transcriptional regulatory sequences (see references 2 and 6) for the *nifBQ* operon are underlined, and a putative transcriptional terminator structure is shown subscripted. The putative initiation codon for the *nifQ* gene is at nucleotide position 1407.

would be predicted from the level of DNA homology. Moreover, the *R. leguminosarum fixZ* gene ORF predicted by Rossen et al. (28) codes for a 40-kDa protein, whereas the *R. meliloti* and *K. pneumoniae* genes code for 54- and 51-kDa proteins, respectively. Finally, there is extensive DNA homology between DNA sequences both 5′ and 3′ of the *R. leguminosarum fixZ* ORF and DNA sequences within both the *R. meliloti* and *K. pneumoniae nifB* ORFS. These observations suggest that the published *R. leguminosarum fixZ* sequence (28) contains frameshift sequencing errors, and in fact, reexamination of the *fixZ* sequence data has revealed the presence of errors in the published sequence (A. W. B. Johnston, personal communication). The correction of these errors, although not yet complete, combined with the extensive DNA homology of *fixZ* and *R. meliloti nifB* genes allowed us to derive a tentative revised version of

the *fixZ* amino acid sequence which is the same size as the *R. meliloti nifB* ORF. A comparison of the predicted amino acid sequences of the *R. meliloti* and *K. pneumoniae nifB* genes and the revised sequence of the *R. leguminosarum fixZ* gene is shown in Fig. 6.

DISCUSSION

Using Tn*5* mutagenesis, maxicell analysis, and DNA sequence analysis, we showed that *R. meliloti* contains a homolog of the *K. pneumoniae nifB* gene directly downstream of *nifA*. *R. leguminosarum* also appears to contain a *nifB* homolog directly downstream of *nifA* (28). Therefore, in three distantly related nitrogen-fixing species which have been examined, *nifB* is directly downstream of *nifA* in a separate transcription unit, and *nifA* and *nifB* are transcribed in the same direction. The reason for the conservation of map position is not clear, especially since *nifA* (a regulatory gene) and *nifB* (a gene required for FeMoco biosynthesis) have significantly different roles in the nitrogen fixation process. Moreover, other features of the ordering and clustering of *nif* operons are not conserved between *K. pneumoniae* and *Rhizobium* species.

Because we could readily detect production of the *R. meliloti nifB* product in a maxicell experiment in which the *nifB* gene was transcribed from the *lacZ* promoter (Fig. 5),

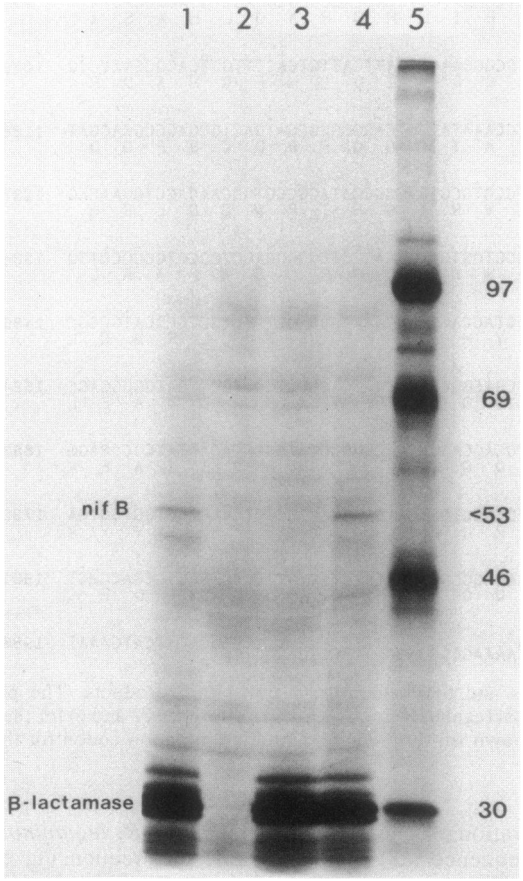

FIG. 5. Identification of the *R. meliloti nifB* gene product in maxicells. Lane 5 shows the ^{14}C-labeled protein molecular weight standards (numbers on right are $\times 10^3$). Lanes 1 to 4 show the polypeptides synthesized by maxicells harboring pWB50, no plasmid, pUC13, and pWB50, respectively. The band corresponding to the putative 53-kDa *nifB* protein is indicated.

we attempted to complement a *K. pneumoniae nifB* mutation with the *lacZ-nifB* fusion carried on plasmid pWB50. There are at least two explanations for the failure to obtain complementation. First, the *R. meliloti* and *K. pneumoniae nifB* genes exhibit only 40% amino acid homology overall. In *K. pneumoniae*, the *nifB* product may interact with other gene products that have been shown to be involved in the synthesis of FeMoco, for example, the *nifN* or *nifE* product (22, 27); sufficient homology may not exist for this interaction to occur with the *R. meliloti nifB* product. Second, the dosage level of the *nifB* gene product may be important; this was not addressed in either the construction of the fusion or its expression in the mutant.

In the course of sequencing the *K. pneumoniae nifB* gene, we also sequenced the region downstream of *nifB*, previously identified as containing the *nifQ* gene (16), and identified a 501-bp ORF. The presence of a presumptive rho-independent transcription terminator directly downstream of the ORF suggests that it is transcribed from the *nifB* promoter. The sequence of the presumptive *nifQ* gene suggests that it is involved in metal binding based on the number and clustering of six cysteine residues (see below); this is consistent with the previously published suggestion that the *nifQ* gene is involved in the processing of molybdenum (14, 15). Within a strongly conserved region about 20 bp downstream of both the *R. meliloti nifB* and *R. leguminosarum fixZ* ORFS are potential ribosome-binding sequences, AUG codons, and single ORFs which show about 85% homology to each other and which extend to the end of the available sequences (100 bp) for both species (W. Buikema, unpublished data; J. A. Downie, personal communication). This region corresponds in position to the *nifQ* gene in *K. pneumoniae*, but we were not able to detect any homology to *nifQ* with the limited DNA sequence that was available to us.

The two *nifB* protein sequences from *K. pneumoniae* and *R. meliloti* and the *R. leguminosarum fixZ* protein sequence demonstrate a high level of conservation of cysteine residues which is characteristic of metal-binding proteins and consistent with the proposed role of the *nifB* gene product in FeMoco processing. The spacing of the cluster of cysteine residues near the amino ends of the *nifQ* protein and has the following consensus: Cys-X_4-Cys-X_2-Cys-X_5-Cys. The spacing of cysteines in other iron-sulfur (Fe-S)-binding proteins such as ferredoxins (20, 21) shows the following typical pattern: Cys-X_2-Cys-X_2-Cys-X_3-Cys-Pro. Rossen et al. (28) noted the same clustering of cysteine residues in the sequence of the *R. leguminosarum fixZ* gene. They also noted the presence of several residues adjacent to the cysteines which contain free amido (asparagine [N] and glutamine [Q]) or guanidino (arginine [R]) groups; these same residues are also present in high proportions adjacent to the cysteine clusters in the *R. meliloti* and *K. pneumoniae nifB* genes. The cysteine clusters in the *nifB*, *fixZ*, and *nifQ* genes are within a very hydrophilic region that contains both polar and charged residues, suggesting that these regions are solvent exposed (35) and that the FeMoco precursor binds either to the outside or within a hydrophilic pocket of these proteins.

Several lines of evidence argue that the *R. meliloti nifB* gene is preceded by a promoter that is activated by the *nifA* gene product. First, in both *K. pneumoniae* and *R. meliloti*, the *nifB* gene is preceded by a rho-independent transcription terminator (7) and by a sequence which conforms to the consensus *nifA*-activated promoter (2). Second, *nifB* transcription in *R. meliloti* nodules requires the *nifA* product (W. W. Szeto, B. T. Nixon, C. W. Ronson, and F. M. Ausubel, J. Bacteriol., in press). Third, the *R. meliloti nifB*

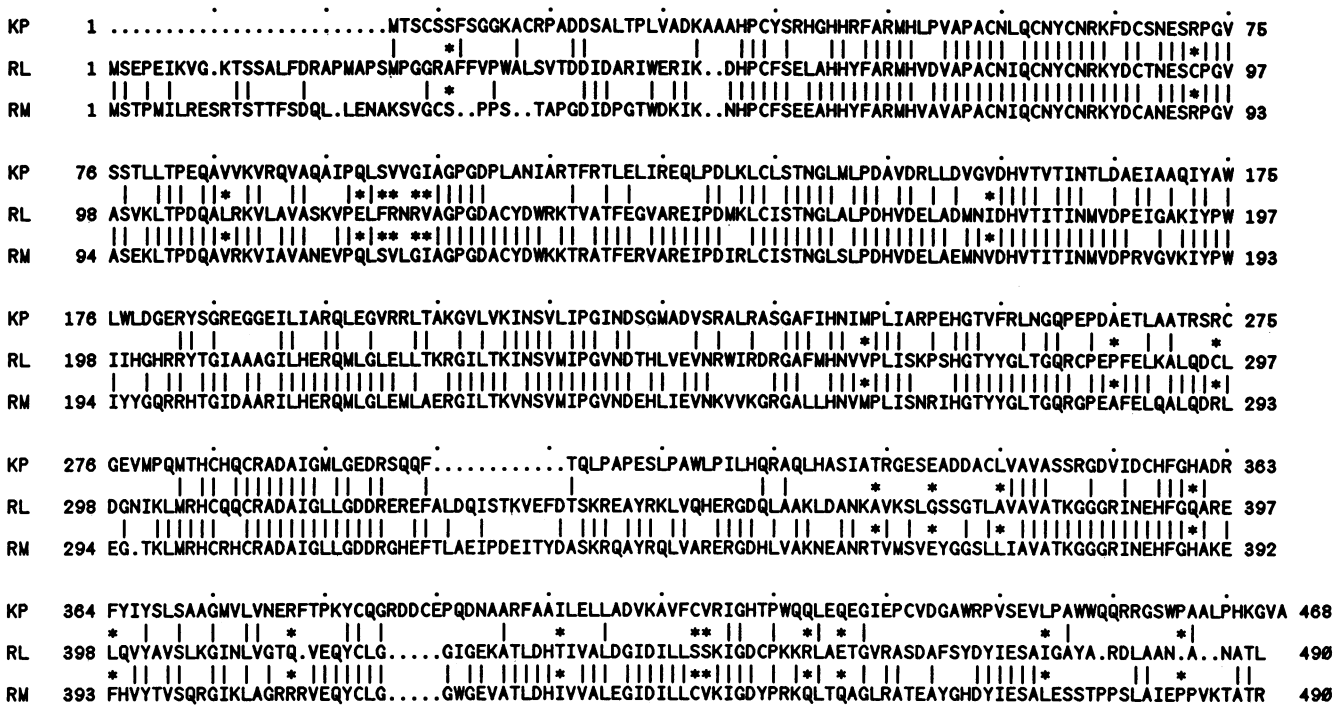

FIG. 6. Comparison of the expected amino acid sequences of the *K. pneumoniae* (KP) and *R. meliloti* (RM) *nifB* and the *R. leguminosarum* (RL) *fixZ* genes. The sequences were aligned for maximum matching by using the GAP program of the University of Wisconsin Genetic Computer Group set to the default parameters (11), and the published *fixZ* sequence has been altered as explained in Results. Vertical bars and stars denote identity between adjacent and nonadjacent sequences, respectively.

promoter region contains the sequence TGT-N_{10}-ACA which has recently been identified as a putative upstream *nifA* binding sequence (6). Using the nomenclature of Buck et al. (6) in which the position of the C residue in the invariant GC dinucleotide in the downstream consensus *nifA* promoter element is designated −12 and the position of the upstream consensus sequence refers to the position of the G in the TGT motif, two upstream elements are found at positions −82 and −102 of the *R. meliloti nifB* gene. We found that the position of the upstream element in the *K. pneumoniae nifB* gene is at position −131, rather than at −143 as stated by Buck et al. (6).

ACKNOWLEDGMENTS

We thank J. A. Downie and A. W. B. Johnston for unpublished *R. leguminosarum nifB* sequence data, J. Beynon for unpublished *K. pneumoniae nifB* sequence data, W. Szeto and D. Ow for unpublished data, C. Earl, J. Golden, and M. Mulligan for critically reading the manuscript, P. McLean for helpful discussions, and R. Hyde for assistance with the references.

This work was supported by a grant from Hoechst AG to Massachusetts General Hospital.

ADDENDUM IN PROOF

Recent additional sequence analysis has shown that the ORF downstream of the *R. meliloti nifB* gene shares 49 and 42% homology at the amino acid level to putative bacterial ferredoxins from *Chlorobium limicola* and *Azotobacter chroococcum*, respectively, including conservation of nine cysteine residues (27a). In addition, it shows 34% homology to a recently identified ORF from *Anabaena* 7120, which also lies immediately downstream of a *nifB* gene (M. Mulligan, personal communication).

LITERATURE CITED

1. **Aguilar, O. M., D. Kapp, and A. Puhler.** 1985. Characterization of a *Rhizobium meliloti* fixation gene (*fixF*) located near the common nodulation region. J. Bacteriol. **164**:245–254.
2. **Ausubel, F. M.** 1984. Regulation of nitrogen fixation genes. Cell **37**:5–6.
3. **Bankier, A. T., and B. G. Barrell.** 1983. Shotgun DNA sequencing, p. 1–34. *In* R. A. Flavell (ed.), Techniques in life sciences, B5: nucleic acid biochemistry. Elsevier/North Holland, Limerick, Ireland.
4. **Batut, J., B. Terzaghi, M. Gherardi, M. Huguet, E. Terzaghi, A. M. Garnerone, P. Boistard, and T. Haguet.** 1985. Localization of a symbiotic *fix* region on *Rhizobium meliloti* pSym megaplasmid more than 200 kilobases from the *nod-nif* region. Mol. Gen. Genet. **199**:232–239.
5. **Beynon, J. L., M. C. Cannon, V. Buchanan-Wollaston, and F. C. Cannon.** 1983. The *nif* promoters of *Klebsiella pneumoniae* have a characteristic structure. Cell **34**:665–671.
6. **Buck, M., S. Miller, M. Drummond, and R. Dixon.** 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature (London) **320**:374–378.
7. **Buikema, W. J., W. W. Szeto, P. V. Lemley, W. H. Orme-Johnson, and F. M. Ausubel.** 1985. Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. Nucleic Acids Res. **13**:4539–4555.
8. **Carlomagno, M. S., A. Riccio, and C. B. Bruni.** 1985. Convergently functional, rho-independent terminator in *Salmonella typhimurium*. J. Bacteriol. **163**:362–368.
9. **Corbin, D., L. Barran, and G. Ditta.** 1983. Organization and expression of *Rhizobium meliloti* nitrogen fixing genes. Proc. Natl. Acad. Sci. USA **80**:3005–3009.
10. **de Bruijn, F. J., and F. M. Ausubel.** 1983. The cloning and characterization of the *glnF* (*ntrA*) gene of *Klebsiella pneumoniae*: role of *glnF* (*ntrA*) in the regulation of nitrogen fixation (*nif*) and other nitrogen assimilation genes. Mol. Gen. Genet.

192:342–353.
11. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
12. **Earl, C. D., C. W. Ronson, and F. M. Ausubel.** 1986. Genetic and structural analysis of the *Rhizobium meliloti fixA, fixB, fixC,* and *fixX* genes. J. Bacteriol. **169**:1127–1136.
13. **Holmes, W. M., T. Platt, and M. Rosenberg.** 1983. Termination of transcription in *E. coli*. Cell **32**:1029–1032.
14. **Imperial, J., R. A. Ugalde, V. K. Shah, and W. J. Brill.** 1984. Role of the *nifQ* gene product in the incorporation of molybdenum in nitrogenase in *Klebsiella pneumoniae*. J. Bacteriol. **158**:187–194.
15. **Imperial, J., R. A. Ugalde, V. K. Shah, and W. J. Brill.** 1985. Mol$^-$ mutants of *Klebsiella pneumoniae* requiring high levels of molybdate for nitrogenase activity. J. Bacteriol. **163**:1285–1287.
16. **MacNeil, T., D. MacNeil, G. P. Roberts, M. A. Supiano, and W. J. Brill.** 1978. Fine structure mapping and complementation analysis of *nif* (nitrogen fixation) genes in *Klebsiella pneumoniae*. J. Bacteriol. **136**:253–266.
17. **Mahl, M. C., P. W. Wilson, M. A. Fife, and W. H. Ewing.** 1965. Nitrogen fixation by the tribe Klebsielleae. J. Bacteriol. **89**: 1482–1487.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual, pp. 113–114. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Meade, H. M., S. R. Long, S. E. Brown, W. J. Buikema, 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.
20. **Newton, W. E.** 1984. Sulfide and other sulfur-containing ligands in metalloproteins and enzymes, p. 409–477. *In* A. Muller and E. Krebs (ed.), Studies in inorganic chemistry, vol. 5. Elsevier/North-Holland Publishing Co., Amsterdam.
21. **Orme-Johnson, W. H.** 1973. Iron-sulfur proteins: structure and function. Annu. Rev. Biochem. **42**:159–204.
22. **Orme-Johnson, W. H.** 1985. Molecular basis of nitrogen fixation. Annu. Rev. Biophys. Biophys. Chem. **14**:419–459.
23. **Ow, D. W., and F. M. Ausubel.** 1980. Recombinant P4 bacteriophages replicate as viable lytic phage particles or as autonomous plasmids in *Klebsiella pneumoniae*. Mol. Gen. Genet. **180**:165–175.
24. **Puhler, A., M. O. Aquilar, M. Hynes, P. Muller, W. Klipp, U. Preifer, R. Simon, and G. Weber.** 1984. Advances in the genetics of free-living and symbiotic nitrogen fixation bacteria, p. 609–619. *In* C. Veeger and W. E. Newton (ed.), Advances in nitrogen fixation research. Nijhoff/Junk, The Hague.
25. **Riedel, G. E., S. E. Brown, and F. M. Ausubel.** 1983. Nitrogen fixation in *Klebsiella pneumoniae* is inhibited by certain multicopy hybrid *nif* plasmids. J. Bacteriol. **153**:45–56.
26. **Roberts, G. P., and W. J. Brill.** 1980. Gene-product relationships of the *nif* (nitrogen fixation) regulon of *Klebsiella pneumoniae*. J. Bacteriol. **144**:210–216.
27. **Roberts, G. P., T. MacNeil, D. MacNeil, and W. J. Brill.** 1978. Regulation and characterization of protein products coded by the *nif* (nitrogen fixation) genes of *Klebsiella pneumoniae*. J. Bacteriol. **136**:267–279.
27a. **Robson, R., P. Woodley, and R. Jones.** 1986. Second gene (*nifH**) coding for a nitrogenase iron protein in *Azotobacter chroococcum* is adjacent to a gene coding for a ferredoxin-like protein. EMBO J. **5**:1159–1163.
28. **Rossen, L., Q.-S. Ma, E. A. Mudd, A. W. B. Johnston, and J. A. Downie.** 1984. Identification of DNA sequences of *fixZ*, a *nifB*-like gene from *Rhizobium leguminosarum*. Nucleic Acids Res. **12**:7123–7134.
29. **Ruvkun, G. B., and F. M. Ausubel.** 1980. Interspecies homology of nitrogenase genes. Proc. Natl. Acad. Sci. USA **77**:191–195.
30. **Ruvkun, G. B., and F. M. Ausubel.** 1981. A general method for site directed mutagenesis in prokaryotes: construction of mutations in symbiotic nitrogen fixation genes of *Rhizobium meliloti*. Nature (London) **289**:85–88.
31. **Ruvkun, G. B., V. Sundaresan, and F. M. Ausubel.** 1982. Directed transposon mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. Cell **29**: 551–559.
32. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**: 692–693.
33. **Sibold, L., B. Quiviger, N. Charpin, A. Paquelin, and C. Elmerich.** 1983. Cloning and expression of a DNA fragment carrying a *his-nifA* fusion and the *nifBQ* operon from a *nif* constitutive mutant of *Klebsiella pneumoniae*. Biochimie **65**: 53–63.
34. **Staden, R.** 1984. Graphic methods to determine the function of nucleic acid sequences. Nucleic Acids Res. **12**:521–567.
35. **Sweeney, W. V., and J. C. Rabinowitz.** 1980. Proteins containing 4Fe-4S clusters: an overview. Annu. Rev. Biochem. **49**:139–161.
36. **Szeto, W. W., J. L. Zimmerman, V. Sundaresan, and F. M. Ausubel.** 1984. A *Rhizobium meliloti* symbiotic regulatory gene. Cell **36**:1035–1043.
37. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.