J. Mol. Biol. (1993) 229, 570–576

SEQUENCE NOTE

Nucleotide Sequence and Characterization of *Rhizobium meliloti* Nodulation Competitiveness Genes *nfe*†

María J. Soto, Adolfo Zorzano, Jesús Mercado-Blanco Viviana Lepek, José Olivares and Nicolás Toro‡

Departamento de Microbiología
Estación Experimental del Zaidín
CSIC, Profesor Albareda 1, 18008 Granada, Spain

(Received 16 June 1992; accepted 24 September 1992)

Rhizobium meliloti large plasmid pRmeGR4b carries the nodulation competitiveness locus *nfe* responsible for the nodulation efficiency and competitive ability of strain GR4 on alfalfa roots. We report here the nucleotide sequence and characterization of a 3345 base-pair DNA section of the *nfe* region. Sequence analysis revealed four open reading frames (ORFs), two of them with rightward polarity, termed *nfe1* and *nfe2*, are preceded by functional *nif* consensus sequences and NifA-binding motifs. An additional, NifA-independent, transcriptional start site for *nfe1* was also found. Two other ORFs with leftward polarity, designated ORFA and ORFB, were identified upstream from *nfe1* and *nfe2* but no *nif* consensus sequences were found. However, expression of ORFA might be indirectly coupled to the NifA–NtrA regulatory network. The gene products of *nfe1* and *nfe2* were identified using *in vitro* transcription/translation and bacteriophage T7 RNA polymerase/promoter system, respectively. A high degree of homology between the amino terminal domain of Nfe1 and the *nifH* gene product was found. In addition, *nfe1* shows homology with the upstream non-coding DNA region of the *fixABCX* operon. Furthermore, the putative ORFB encoded protein contains a helix-turn-helix motif that resembles the DNA-binding consensus sequence proposed for many prokaryotic regulatory proteins.

Keywords: DNA sequence; *R. meliloti*; competitiveness; *nif* promoters

Legume productivity in agricultural fields may be improved by inoculation with selected highly effective nitrogen-fixing root nodule bacteria. However, field legume inoculations with *Rhizobium* and *Bradyrhizobium* spp. have often been unsuccessful because of the presence in the soil of native strains that compete with the introduced strain in nodule formation on the host plant. This ability to dominate nodulation is termed competitiveness and is critical for the successful use of inoculants (Boonkerd *et al.*, 1978; Beattie *et al.*, 1989; Kamicker & Brill, 1986; Noel & Brill, 1980). However, the molecular mechanism of competitiveness is not understood. Two approaches are being used by geneticists to address the *Rhizobium* competition problem. First, by constructing *Rhizobium* inoculum strains with an altered host range and/or using legume genotypes with restricted *Rhizobium* infectivity. Second, by constructing strains of *Rhizobium* with an increased ability to compete for nodule occupancy (Triplett, 1990).

Previously, we reported that a DNA region named *nfe* (nodule formation efficiency) located on the large plasmid pRmeGR4b contains genes involved in nodulation efficiency and competitiveness of *Rhizobium meliloti* GR4 on alfalfa roots (Toro & Olivares, 1986; Sanjuan & Olivares, 1989). In addition, expression of the *nfe* genes was found to be dependent on the NifA–NtrA regulatory system (Sanjuan & Olivares, 1989, 1991). It was suggested that the *nfe* genes would transcribe from a *nifH* promoter-like DNA sequence located in the 905 bp§ *Pst*I DNA fragment, 33a, of pRmeGR4b (Sanjuan & Olivares, 1989). The aim of this work was to sequence and characterize the previously identified competitiveness *nfe* genes as a first step toward the understanding of competition in the *Rhizobium*–

† This manuscript is dedicated to A. Zorzano who died in a car accident on May 28, 1991.

‡ Author to whom all correspondence should be addressed.

§ Abbreviations used: bp, base-pair(s); kb, 10^3 base-pairs; ORF, open reading frame; HTH, helix-turn-helix motif; UAS, upstream activator sequence.

0022-2836/93/020570-07 $08.00/0

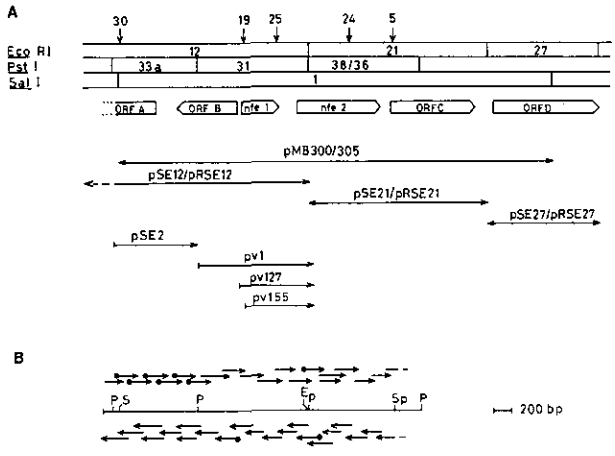

Figure 1. The *nfe* region. A, Physical map of the *nfe* region. The vertical arrows indicate the positions of Nfe$^-$ Tn*3*HoHo1 insertions based upon the mapping data of Sanjuan & Olivares (1989). Location of insertion 30 is according to our sequence data. The 4 ORFs identified in this work are indicated by open arrows below the map. A further downstream ORF (ORFC) identified as an ornithine cyclodeaminase gene (Soto, M. J., unpublished results) and the IS*Rm3* (Soto *et al.*, 1992) putative transposase (ORFD), are also shown. Note that the previously identified *Pst*I fragments 38 and 36 are shown here as a single *Pst*I fragment, designated 38/36. Cloned fragments used as primary source for DNA sequencing and/or *nfe* gene expression are shown below the coding regions. Double arrows indicate clones in both orientations. B Sequencing strategy of the *nfe* locus (E, *Eco*RI; P, *Pst*I; S, *Sal*I; Sp, *Spe*I). The nucleotide sequence presented in this work was determined by the chain termination method of Sanger *et al.* (1977) in both strands using Sequenase version 2.0 T7 DNA polymerase (United States Biochemical). Templates were generated by progressive unidirectional deletions of restriction fragments cloned in pUC18 using the Erase-a-Base system from Promega Biotech. The horizontal arrows indicate the extent and direction of DNA sequenced determined from individual clones; arrows marked with a bullet represent DNA regions from which the sequence was obtained using specific oligonucleotides.

legume symbiosis. In this report, we present the nucleotide sequence and characterization of a 3345 bp DNA section of the *nfe* region. Our results show that the former DNA region contains four ORFs: two of them, termed *nfe1* and *nfe2*, are controlled at the transcriptional level by the NifA–NtrA regulatory system, although *nfe1* could also transcribe on a NifA-independent manner. Two other ORFs, called ORFA and ORFB, were identified upstream from *nfe1* and *nfe2* but transcribing in opposite orientation. Expression of ORFA and the homology of ORFB to transposases and prokaryotic regulatory proteins are discussed.

(a) *The* nfe *region*

Upon genetic analysis of Tn*3* HoHo1 transposon insertion mutants, Sanjuan & Olivares (1989) showed that the *nfe* genes were mainly clustered in a DNA region of approximately 2·6 kb (Fig. 1A). The former region is delimited at the left site by insertion 30 within the *Pst*I fragment 33a internal to the *Eco*RI fragment 12, and on the right by insertion 5 located at the left site of the unique *Spe*I site within the *Eco*RI fragment 21. We determined the nucleotide sequence of 3345 bp DNA region from pRmeGR4b extending 118 nucleotides from the left *Pst*I site of *Pst*I fragment 33a to the unique *Spe*I site present in the *Eco*RI fragment 21. The sequencing strategy is outlined in Figure 1B.

(b) *Nucleotide sequence of the* nfe *region*

To sequence the *R. meliloti* strain GR4 *nfe* region, subclones derived from plasmid pRmNT40 (Toro & Olivares, 1986) presented in Figure 1A were used. Sequence analysis and homology search were done with the Genepro version 4·0 (Riverside Scientific Enterprises) and the GCG (University Research Park, Madison, Wisconsin) software packages. The programs FASTA (Pearson & Lipman, 1988), PEPTIDESTRUCTURE, PILEUP and PRETTY from the University of Wisconsin were used.

Nucleotide sequence analysis of the 3345 bp DNA region revealed the presence of four major open reading frames with high coding probability (Fig. 1A). The four ORFs were preceded by potential ribosome binding sequences with homology to the consensus binding site of *Escherichia coli* (Shine & Dalgarno, 1974). The two first ORFs had a leftward polarity while the other two transcribed in the opposite orientation. The first ORF, ORFA, was still open, predicting a protein of more than 205 amino acid residues. An ORF that overlaps ORFA but with rightward polarity was also found (data not shown). However, ORFA was chosen because the Tn*3*–*lacZ* fusion, pRmNT40-30 (Sanjuan & Olivares, 1989) maps in this ORF with the same orientation. The second ORF, ORFB, was 696 bp long, potentially coding for a 232 amino acid residue long protein. ORFB is separated by 251 bp of the translational start site of ORFA. At 32 bp upstream from ORFB a third ORF was found. This ORF had a rightward polarity and it was 423 bp long encoding a putative protein of 141 amino acid residues. At 211 bp downstream from the former ORF stop codon there was a fourth rightward ORF. This ORF was 933 bp long and codifies a putative protein of 311 amino acid residues. The latter two ORFs were designated *nfe1* and *nfe2*, respectively, because they were preceded by consensus *nif* promoter sequences and NifA-binding motifs. These specific signal structures characteristic of genes whose expression is controlled by the NifA–NtrA regulatory system were suggested to be responsible for the transciptional regulation of the *nfe* genes (Sanjuan & Olivares, 1989). At 128 bp downstream from *nfe2*, begins an ORF whose gene product has been identified as an ornithine cyclodeaminase; its complete nucleotide sequence and characterization will be published elsewhere (Soto, M. J., unpublished results). The four ORFs identified in the

3·3 kb region are not closely connected, which suggests that they could form single transcriptional units.

(c) *The* nfe nif-*like promoter elements*

The *nif* promoters have a characteristic primary structure in most nitrogen-fixing species with a consensus sequence CTGGYAYR-N4-TTGCA extending from position −26 to −10 relative to the transcription initiation site (Ausubel, 1984). In addition to the consensus *nif* promoter sequence, all NifA-activated promoters are preceded by an upstream element sequence (UAS) TGT-N10-ACA, which has been postulated to be a putative NifA-binding site (Buck *et al.*, 1986). Two *nif* consensus promoter sequences and two putative NifA-binding sites were located in the *Pst*I fragment 31 in front of *nfe1* and *nfe2*. The sequence CCGG-N8-TTGCA, at positions 1520 to 1536 upstream *nfe1* showed a single mismatch with the consensus *nif* sequence, a transition T→C at −26. Furthermore, the ACGA structure at −22 to −19, which is common in most *Rhizobium* NifA-regulated promoters was also present. In addition, a putative NifA-binding site at co-ordinates 1472 to 1487 was located 32 bp upstream from the *nif* promoter. The consensus *nif* sequence and UAS found in front of *nfe1* was located into the coding region of ORFB. A *nif* promoter sequence also precedes *nfe2*; the sequence CTGG-N8-TTGCA, at positions 2055 to 2071, perfectly fits the consensus. However, the ACGA structure at positions −22 to −19 is not absolutely conserved, showing a transition C→T at position −21. A perfect NifA-binding site at position 1953 to 1968 was found at 86 bp from the *nif* promoter sequence and located into the coding region of *nfe1*. How this structure may affect the transcription of *nfe1*, remains speculative. On the other hand, the putative NifA binding site preceding *nfe*1 is unusually closed to the *nif* promoter with a spacing of 32 bp, which raises the possibility that the activity of the *nfe1* promoter may be relatively reduced *in vivo*.

(d) *Functionality of the* nfe nif-*like promoters*

Primer extension experiments were performed to determine whether the *nif* promoter sequences located in front of *nfe1* and *nfe2* were functional. Using a specific oligonucleotide for *nfe1*, two major extension products were observed (Fig. 2A). The lower extension product comigrates with the nucleotide complementary to the G residue that is 11 nucleotides downstream from the *nif* promoter located in front of *nfe1*. The upper extension product (Fig. 2A) comigrates with the nucleotide complementary of the C residue at position −20 in the *nif* promoter. Whereas the lower extension product was dependent on NifA the upper one was stimulated but it was not dependent. It is possible that the specific interaction of the NifA activator with the UAS located in front of the *nif* promoter sequence of *nfe1* facilitates transcription from a second upstream promoter. When we used a specific primer complementary to *nfe2* only a single NifA-dependent extension product was observed (Fig. 2B). This extension product comigrates with a nucleotide complementary of the T residue that is 11 nucleotides downstream from the *nif* promoter sequence located in front of *nfe2*.

Genes controlled by *nif* promoter elements used to have 5′ leaders ranging between 20 to 150 bases (Buck *et al.*, 1986; Roelvink & Van Den Bos, 1989). The 5′ untranslated leader of *nfe1* is 59 bp long while the transcription start site for *nfe2* is 160 nucleotides from the ATG start codon (Fig. 3A). It would be interesting to see whether the different lengths of the 5′ untranslated mRNA of *nfe1* and *nfe2* transcripts reflects a distinctive expression pattern of former genes.

Data indicate that both *nif*-like promoters are functional and that transcription of downstream genes, *nfe1* and *nfe2*, is dependent of NifA. Our results further strengthened previous reports (Sanjuan & Olivares, 1989, 1991), indicating that expression of the *nfe* genes is subject to a complex regulatory network involving the *nifA*-dependent circuit. However, the fact that *nfe1* can be expressed uncoupled from the NifA–NtrA regulatory system suggests that *R. meliloti* GR4 competitiveness may be triggered before the onset of nitrogen fixation.

(e) *Expression of ORFA*

DNA sequence analysis revealed two *nif* consensus promoter sequences and two putative NifA-binding sites located in the *Pst*I fragment 31 of 1·4 kb. On the other hand, sequence data did not reveal any consensus *nif* promoter in the *Pst*I fragment 33a, which is in contradiction of previous data reported by Sanjuan & Olivares (1989). In our sequence, the *Pst*I fragment 33a is 966 bp long while Sanjuan & Olivares (1989) reported the size of the fragment at 905 bp. The sequence shown in this study diverges from the latter mainly in the DNA region where it was proposed that a consensus *nif* promoter sequence could be found. Nevertheless, Sanjuan & Olivares (1989) located the transposon insertion 30 affecting the competitive ability of strain GR4 in the *Pst*I fragment 33a. Expression of the former Tn*3–lacZ* fusion was dependent on the NifA/NtrA regulatory system (Sanjuan & Olivares, 1989, 1991). By using the 24-mer oligonucleotide, 5′-AGGCAGAAAACGTTGCTTAACGTG-3′, specific for the Tn*3*HoHo1 transposon (Stachel *et al.*, 1985), we mapped insertion 30 into the coding region of ORFA between co-ordinates, position 146 and 147. To rule out the presence of degenerated *nif* consensus sequences in front ORFA primer extension analyses were carried out. Using the identical RNA samples that we used for mapping the transcriptional start sites of *nfe1* and *nfe2* we failed to detect any NifA-dependent extension products from ORFA (data not shown). Our results suggest that expression of the former ORF is not directly

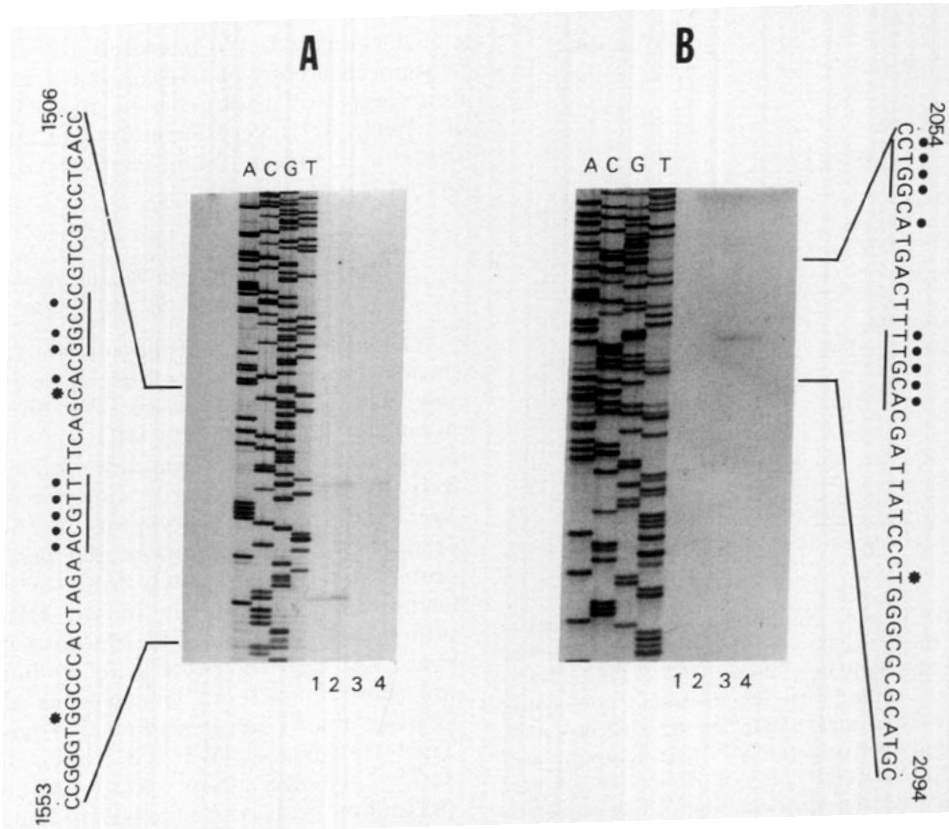

Figure 2. Primer extension mapping of *nfe* promoters. A, Primer extension mapping of *Pnfe1*. Primer extension products were obtained using total RNA from *E. coli* cells harbouring plasmid pCK3, *K. pneumoniae nifA* cloned in pRK290 (Buchanan-Wollaston *et al.*, 1981) and pSE12, lanes 1 (3 μl) and 2 (5 μl); or only pSE12, lanes 3 (3 μl) and 4 (5 μl). B, Primer extension mapping of *Pnfe2*. Primer extension products were obtained using total RNA from *E. coli* cells harbouring plasmid pSE12, lanes 1 (3 μl) and 2 (5 μl); or pCK3 and pSE12, lanes 3 (3 μl) and 4 (5 μl). On the left side of panel A and on the right side of panel B are indicated the DNA region containing the *nif* consensus sequences (underlined). The transcription start points are shown by an asterisk above the sequence. A 24-mer oligonucleotide 5'-CGACCTGACGCACAGCAGCCATCT-3' (complementary to *nfe1*) co-ordinates position 1604 to 1627 and a 20-mer oligonucleotide 5'-AAGCTCATATGGAAGTTGAC-3' (complementary to *nfe2*) co-ordinates position 2199–2218 were 5' end-labelled by [γ^{32}-ATP] using T4 polynucleotide kinase. The labelled primers (150 ng) were allowed to anneal at 65°C and with 6 μg RNA for 2 min and then cooled down slowly to room temperature. Extension by AMV reverse transcriptase (6 U, Promega Biotech.) took place at 42°C for 1 h. A dideoxy sequencing reaction of pSE12 using the same oligonucleotides primers was run in parallel on a 6% sequencing gel.

controlled by NifA/NtrA regulatory system. One possible explanation that conciliates the sequence and the genetic data is that the expression of ORFA might be indirectly coupled to the NifA/NtrA regulatory network. This coupled transcription of RNA polymerase-σ^{70}-dependent operons to regulation by the RNA polymerase-σ^{54}-dependent NTR (nitrogen regulation) system has been described to *Klebsiella aerogenes*; a protein called NAC (nitrogen assimilation control) is involved in the process (Bender, 1991). Further investigations are required to elucidate whether this activation model has any significance in the expression of some *nfe* genes.

(f) *DNA sequence homologies to* nfe1 *and* nfe2

Comparison of the DNA sequence of *nfe1* and *nfe2* with sequences in the GenBank and EMBL data bases showed extensive homologies with the *R. meliloti nifH* and upstream non-coding sequences of the *fixABCX* operon (Fig. 3A). *nfe1* showed 79·5% identity in a 132 bp overlap with the promoter region (P1) and coding sequence of the *nifH* gene. The rest of the coding sequence of *nfe1* showed homology with the upstream non-coding sequence of the *fixABCX* operon, 61·4% identity in a 316 bp overlap, including the NifA-binding motif of the former operon (Fig. 3A). The promoter region and part of the untranslated 5' leader of *nfe2* also showed homology with the P1 promoter region 80% identity in a 92 bp overlap (Fig. 3A). The rest of *nfe2* did not show any significant homologies with sequences in the data bases. The DNA sequence of *nfe1* and the promoter element of *nfe2* showed a striking homology with the *R. meliloti nifH* promoter and its 5' end coding region as well as the upstream non-coding region of the *fixABCX* operon. This observation leads to the question of how the competitiveness *nfe* genes have evolved? The presence of different IS (insertion sequence) elements, with a multiplicity of copies in *Rhizobium* (Simon *et al.*, 1991; Wheatcroft & Laberge, 1991) is

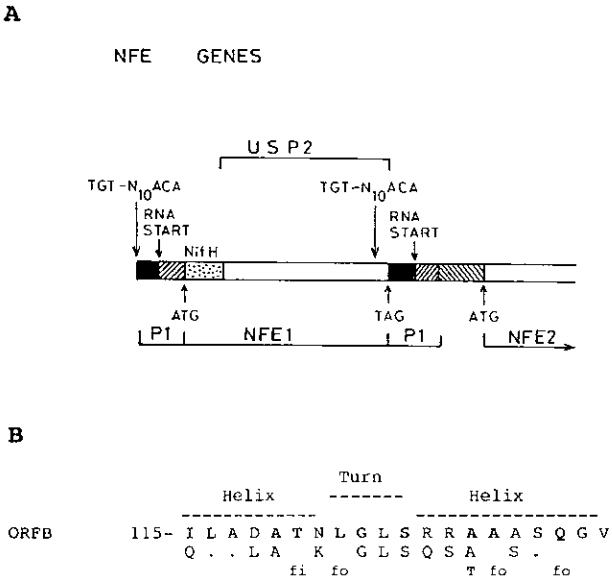

Figure 3. DNA and protein sequence alignments. A, Structure of the *nfe* genes. Filled and leftward hatched boxes represent the homology among *nfe1* (1478 to 1668) and *nfe2* (1041 to 2136) with the upstream non-coding sequence of *R. meliloti nifH* (P1). The stippled box indicates the homology between the Nfe1 first 30 amino acid residues and the NifH ferroprotein. Homology of *nfe1* coding sequence with the upstream non-coding sequences of *R. meliloti fixABCX* operon (USP2) is also shown. Untranslated leader of *nfe2* with no homology to *nif/fix* sequences is indicated by a rightward hatched box. B, Conservation of amino acid residues in the helix-turn-helix element of ORFB and other regulatory proteins. Conserved residues in the HTH of ORFB with the consensus (below) are in bold. fi, hydrophilic residues; fo, hydrophobic residues.

well documented. We have recently described the presence of a copy of *ISRm3* linked to the nodulation competitiveness genes in strain GR4 (Soto *et al.*, 1992). It is very suggestive of the possibility that this IS element might be involved in the evolution of the *nfe* genes.

(g) *Amino acid sequence homologies to putative proteins of the* nfe *region*

We looked for proteins homologous to putatively encoded proteins of the *nfe* region in the Swissprot and PIR protein data bases. We did not find any significant homology with ORFA, or Nfe2-encoded products. On the other hand, we found strong similarity between the first 30 amino acid residues of Nfe1 and the amino terminus of the *nifH* gene product (Fig. 3A). However, the rest of the Nfe1 protein diverges completely from the *nifH*-encoded ferroprotein. Comparison of the ORFB putative protein with sequences in the data bases showed homology with the *tnpA* gene product of some of the class II bacterial transposable elements. The identity in a 51 amino acid residue overlap with the Tn*21* transposase (Ward & Gransted, 1987) was 42·1% whereas there was a 31·3% identity (83 amino acid residue overlap) with the Tn*501* transposase (Brown *et al.*, 1985). This region comprises a helix-turn-helix (HTH) motif (Chou & Fasman, 1978) that resembles the proposed DNA-binding consensus sequence (Kondorosi *et al.*, 1991) of a large number of prokaryotic regulatory proteins (Fig. 3B). Whether ORFB actually codifies for a transcriptional regulator or a transposase remains unknown. Mutations in this ORF are not still available and its involvement in competitiveness, if any, has to be demonstrated.

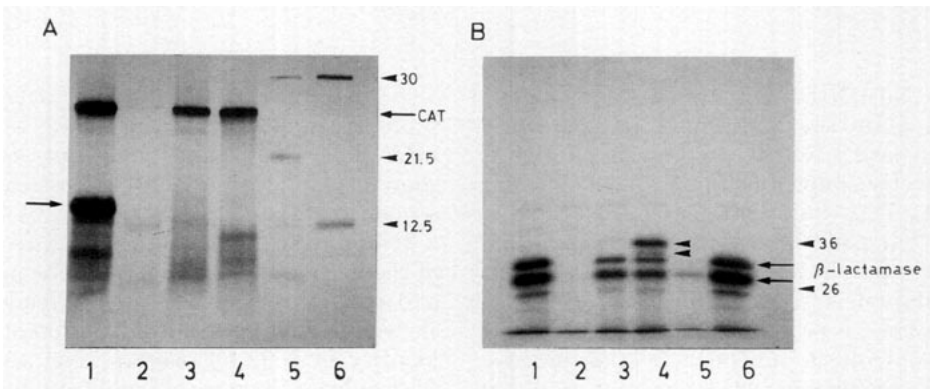

Figure 4. SDS/polyacrylamide gel of *nfe* region proteins synthesized *in vitro* and *in vivo*. A, Protein synthesis by *in vitro* coupled transcription/translation with an *E. coli* S-30 linear system (Promega Biotech.) of *Pvu*II-digested PUC18 clones pV155 (lane 1), pV127 (lane 2), pV1 (lane 3) and pUC18 (lane 4). Relative molecular masses ($\times 10^{-3}$) of protein standards (lanes 5 and 6) are given and proteins encoded by the *Pvu*II digested plasmids are indicated by arrows. CAT (chloramphenicol acetyl transferase). B, Protein synthesis using the *in vivo* coupled T7 RNA polymerase/T7 promoter system performed as described by Tabor & Richardson (1985). *E. coli* K38 (Russel & Model, 1984) cells containing the T7 promoter-expression vector pT7-3 (lane 1 and 6), or the T7 polymerase vector pGp1-2 (lane 2 and 5), and cells containing pGp1-2/pMB300 (lane 3), or pGp1-2/pMB305 (lane 4). Plasmids pMB305/300 are pT7-3 clones containing the 4·9 kb *Sal*I fragment from pRmeGR4b in both orientations. Specific insert polypeptides are indicated by small arrows. Protein samples obtained by either *in vitro* or *in vivo* synthesis were electrophoresed on 12% polyacrylamide gels according to Laemmli (1970). Gels were fluorographed with Amplify (Amersham), dried and exposed to autoradiography.

(h) *Protein products of* nfe *clones in* E. coli

To identify proteins encoded by the *nfe* region a number of expression vector clones of this region were analysed by *in vitro* coupled transcription/translation and *in vivo* with the coupled T7 polymerase/promoter system.

The proteins produced *in vitro* by *Pvu*II linear fragments of pUC18 clones spanning *nfe1* regions are shown in Figure 4A. *Pvu*II linear fragments from plasmid pV155 (Fig. 4A, lane 1) produce an insert-specific protein with a relative molecular mass (M_r) of 15,000, corresponding to the predicted size of Nfe1 ($M_r = 15{,}652$). The 15,000-M_r protein did not appear when the *in vitro* synthesis was performed with *Pvu*II linear fragments of plasmids pV1, pV127 or pUC18 used as a control (Fig. 4A; lanes 2, 3 and 4). Plasmid pV115 contains most of the coding region of *nfe1* but it lacks 17 codons of its 5′ coding region. Sequence data revealed that the 15,000-M_r protein resulted from a translational fusion between 124 codons of *nfe1* and the first five codons of the *lacZ*-encoded α-peptide. On the other hand, *in vitro* transcription/translation of *Pvu*II linear fragments of plasmids pSE12/pRSE12 spanning ORFA and ORFB did not reveal any insert-specific polypeptide (data not shown).

The existence of the *nfe2* gene product was demonstrated by *in vivo* synthesis with the coupled T7 polymerase/promoter system. Using the plasmid clone pMB305, we detected an approximately 34,000-M_r protein (Fig. 4B, lane 4) corresponding to the predicted size of Nfe2 ($M_r = 34{,}223$). The former polypeptide appears as a faint band, which suggests a poor translational initiation or that the protein is unstable. The higher relative molecular mass protein (more intense band) also observed in Figure 4B, lane 4, corresponded to the product of an ORF located downstream from *nfe2*, identified as an ornithine cyclodeaminase gene (Soto, M. J., unpublished results). The 34,000-M_r protein was not observed with plasmid clone pMB300, which contains the same insert as pMB305 but in the opposite orientation (Fig. 4B, lane 3). In addition, *in vitro* transcription/translation of plasmid clones containing upstream or downstream sequences of *nfe2* (pSE12/pRSE12 and pSE27/pRSE27), or plasmids that contain most of the *nfe2* coding region but lack its 5′ end (pSE21/pRSE21), did not show the presence of the 34,000-M_r protein (data not shown).

We are grateful to J. Sanjuan who provided some of the sequencing templates and to S. Tabor for the T7 polymerase/T7 promoter plasmids vector. This work was supported by Comision Asesora de Investigación Científica y Técnica, Grant BIO90-0740 and by the E.E.C. Bridge Contract BIOT-0159-C. M.J.S., J.M. and V.L. were supported by M.E.C. fellowships. The nucleotide sequence presented in this work is available on the EMBL GenBank Nucleotide Sequence Database under the accession number X66124.

References

Ausubel, F. M. (1984). Regulation of nitrogen fixation genes. *Cell*, **37**, 5–6.

Beattie, G. A., Clayton, M. K. & Handelsman, J. (1989). Quantitative comparison of the laboratory and field competitiveness of *Rhizobium leguminosarum* biovar *phaseoli*. *Appl. Environ. Microbiol.* **55**, 2755–2761.

Bender, R. A. (1991). The role of NAC protein in the nitrogen regulation of *Klebsiella aerogenes*. *Mol. Microbiol.* **5**, 2575–2580.

Boonkerd, N. D., Weber, D. F. & Bezdicek, D. F. (1978). Influence of *Rhizobium japonicum* strains and inoculation methods on soybeans grown in rhizobia-populated soil. *Agron. J.* **70**, 547-549.

Brown, N. L., Winnie, J. N., Fitzinger, D. & Pridmore, R. D. (1985). The nucleotide sequence of the *tnpA* gene completes the sequence of the *Pseudomonas transposon* Tn*501*. *Nucl. Acids Res.* **15**, 5657–5663.

Buchanan-Wollaston, A. V., Cannon, M. C., Beynon, J. L. & Cannon, F. C. (1981). Role of the *nifA* gene product in the regulation of *nif* expression in *Klebsiella pneumoniae*. *Nature (London)*, **294**, 776–778.

Buck, M., Miller, S., Drummond, M. & Dixon, R. (1986). Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature (London)*. **320**, 374–378.

Chou, P. Y. & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Av. Enzymol. Relat. Areas Mol. Biol.* **47**, 45–148.

Kamicker, B. J. & Brill, W. J. (1986). Identification of *Bradyrhizobium japonicum* nodule isolates from Wisconsin soybean farms. *Appl. Environ. Microbiol.* **51**, 487–492.

Kondorosi, E., Pierre, M., Cren, M., Hauman, U., Buiré, M., Hoffman, B., Schell, J. & Kondorosi, A. (1991). Identification of NolR, a negative transacting factor controlling the *nod* regulon in *Rhizobium meliloti*. *J. Mol. Biol.* **222**, 885–896.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)*, **227**, 680–685.

Noel, K. D. & Brill, W. J. (1980). Diversity and dynamics of indigenous *Rhizobium japonicum* populations. *Appl. Environ. Microbiol.* **40**, 931–938.

Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequences comparison. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 2444–2448.

Roelvink, P. W. & Van den Bos, R. C. (1989). Regulation of nitrogen fixation in diazotrophs: the regulatory *nifA* gene and its characteristics. *Acta Bot. Neerl.* **38**, 233–252.

Russel, M. & Model, P. (1984). Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned in a plasmid. *J. Bacteriol.* **159**, 1034–1039.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sanjuan, J. & Olivares, J. (1989). Implication of *nifA* in regulation of genes located on a *Rhizobium meliloti* cryptic plasmid that affects nodulation efficiency. *J. Bacteriol.* **171**, 4154–4161.

Sanjuan, J. & Olivares, J. (1991). NifA–NtrA regulatory system activates transcription of *nfe*, a gene locus involved in nodulation competitiveness of *Rhizobium meliloti*. *Arch. Microbiol.* **155**, 543–548.

Shine, J. & Dalgarno, L. (1974). The 3′ terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 1342–1346.

Simon, R., Hötte, B., Klauke, B. & Kosier, B. (1991). Isolation and characterization of insertion sequence

elements from Gram-selection vectors. *J. Bacteriol.* **173**, 1502–1508.

Soto, M. J., Zorzano, A., Olivares, J. & Toro, N. (1992). Nucleotide sequence of *Rhizobium meliloti* GR4 insertion sequence IS*Rm3* linked to the nodulation competitiveness locus *nfe*. *Plant. Mol. Biol.* **20**, 307–309.

Stachel, S. E., An, G., Flores, D. & Nester, E. W. (1985). A Tn*3* *lacZ* transposon for the random generation of β-galactosidase gene fusions: application to the analysis of gene expression in *Agrobacterium*. *EMBO J.* **4**, 891–898.

Tabor, S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 1074–1078.

Toro, N. & Olivares, J. (1986). Characterization of a large plasmid of *Rhizobium meliloti* involved in enhancing nodulation. *Mol. Gen. Genet.* **202**, 331–335.

Triplett, E. W. (1990). The molecular genetics of nodulation competitiveness in *Rhizobium* and *Bradyrhizobium*. *Mol. Plant-Microbe Interact.* **3**, 199–206.

Ward, E. & Grinsted, J. (1987). The nucleotide sequence of the *tnpA* gene of Tn*21*. *Nucl. Acids Res.* **15**, 1799–1806.

Weatcroft, R. & Labarge, S. (1991). Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm3*: similarity between the putative transposase encoded by IS*Rm3* and those encoded by *Staphylococcus aureus* IS*256* and *Thiobacillus ferrooxidans* IS*T2*. *J. Bacteriol.* **173**, 2530–2538.

Edited by N.-H. Chua