

A gene upstream of the *Rhizobium trifolii* *nifA* gene encodes a ferredoxin-like proteinSiiri E. Iismaa^{1,2} and John M. Watson¹¹CSIRO Division of Plant Industry and ²Centre for Recombinant DNA Research, RSBS, Australian National University, Canberra City, ACT 2601, Australia

Submitted March 2, 1987

We have identified a short open reading frame (ORF) between the nodulation and nitrogenase genes of *Rhizobium trifolii* strain ANU843. The 294-bp ORF spans two adjacent *Hind*III fragments (578bp and 2.2kb) and terminates 225bp upstream of the proposed translation start of *nifA* (1). The predicted amino acid sequence of this ORF contains the characteristic Cys--Cys--Cys---Cys motif (circled in figure), present in ferredoxins from 15 different bacterial species (2). One possible function for this ferredoxin-like protein is that it may be involved in electron transport to the *Rhizobium* nitrogenase complex. In this context, it is significant that this gene (termed *fixX*) is immediately downstream from another ORF that is homologous to the *Rhizobium meliloti* *fixC* gene (3) which is also thought to be involved in electron transport to nitrogenase (4,3). A gene encoding an homologous, ferredoxin-like protein has also recently been located very close to the 3' end of *fixC* in *R. meliloti* (3).

```

      . . . . .
L F V S Y P K L I S K A A Q S F V R V D G T S K I D K E K A
TTTGTTTGTTTCATATCCAAAGTTAATATCTAAAGCTGCGCAAAGTTTGTCCGTGGATGGACTTCAAATCGACAAAGAAAAGGC 90
      . . . . .
T S A A F L K A R S R W G L I S D A V R L A R A W R *
CACGTCTGCCGCTTCTTGAAGCACGATCCCGGTGGGGATTAATTCAGATGCAGTCCGGCTCGCCCGCGCTGGCCGCTAAAAGGAGAT 180
      . . . . .
M K A I V K R R V E D K L Y Q N R Y L V D P G R P H I S V
CCAGATGAAGGCCATCGTTAAGCGCGAGTTGAGGACAAGCTTACCAGAACCGATATCTGGTTGACCCGGCCGCTCCACACATATCGGT 270
      . . . . .
R K H L F P T P N L I A L T Q V C P A K C Y Q L N D R R Q V
CAGAAAACACCTGTTCCAAACCCAAACCTAATCGCCTTGACGCAAGTGTGCCGGCAAATGCTATCAGCTGAACGACGACAGCTCAAGT 360
      . . . . .
I I V S D G (C) L E (C) G T (C) N V L (C) G P D G D I E W T Y P R G
GATCATCGTTTCCGATGGCTGCTGGGAATCGCGACATGCAACGTGTTATGCGGACCCGACGGCGACATGAGTGAGGACGTATCCACGAGG 450
      . . . . .
G F G V L F K F G *
TGGTTCCGCGTTCCTTCAAGTTTGGATGA 481

```

Legend: DNA sequence of the 3' end of *fixC*, and the *fixX* gene of *R. trifolii*. Predicted amino acids are shown above the DNA sequence; dots indicate residues which are conserved between *R. trifolii* and *R. meliloti* (3).

ACKNOWLEDGEMENTS We thank F.M. Ausubel for communication of results prior to publication and J.A. Downie and K.F. Scott for constructive comments. This work was supported in part by an Agrigenetics Sponsored Research Program. SEI is the recipient of a Commonwealth Postgraduate Award.

REFERENCES

- 1) Iismaa, S.E., Ealing, P.M. and Watson, J.M., manuscript in preparation
- 2) Protein sequence database of the Protein Identification Resource (NBRF)
- 3) Earl, C.D., Ronson, C.W. and Ausubel, F.M. (1987) *J. Bacteriol.* (in press)
- 4) Gubler, M. and Hennecke, H. (1986) *FEBS Lett.* 200, 186-192