Nucleotide sequence of *nifK* and partial sequence of *nifD* from *Frankia* species strain FaC1James M. Ligon and James P. Nakas¹Ciba-Geigy Corporation, PO Box 12257, Research Triangle Park, NC 27709 and ¹SUNY, College of Environmental Science and Forestry, Syracuse, NY 13210, USA

Submitted October 18, 1988

Accession no. X12649

We have previously reported the cloning of the *nifD* and *nifK* genes from *Frankia* (1) that encode the α - and β -subunits of dinitrogenase, respectively. Here we present the complete nucleotide sequence of the *Frankia nifK* gene and partial sequence of the 3' portion of the *nifD* gene. The partial *nifD* reading frame begins at base 3 and ends at 814 and the *nifK* gene starts at base 883 and ends at 2452 (underlined). The deduced NifK polypeptide has 523 amino acid residues, a calculated M_r of 58,700 and shares 66, 46 and 49% similarity with NifK polypeptides of *A. vinelandii*² (2), *B. japonicum* (3) and *Anabaena* (4), respectively. The *nifK* gene is followed by two potential stem loop structures (underlined).

```

GGATCTGAACAATCGCGAAGGACAGCCGTTTGAACCAACCCCTTACGATGTGGCGATCATCGGCGACTACAACATCGGCGCGAAGCGCTGGCGCTCGCGCATCTGCTG
200
GAAGAGATGGGGCTACGGGTAGTCGCGCAGTGGTCCGGCAGCGCAGCTGGTGGGAGATGGAGATACCCCATTCGTAAAGCTGAACCTGGTCTACTGCTACCGTTCGAT
300
GAACATATATCGCCGCCCATATGGAGGACAAACATCAGATTCCGTGGATGGAGTACAACCTTCTCGGGCGGACCAAAATCGCCGAATCGCTGCGCAAAATCGCGACCACT
400
TCGACGATACCATTCGCGCGAAGCGGATCGCCGGTATGAGGGGAGATGGCGCGGATTATCGCAATATCGCCCGCGCTGGAGGGCTAAGGTCTGTC
500
TCTATATGGCGCTCGCGCAGCTTATTGGCGCTATGAGGATCTCGGGATGGAGATCATCGCGCGCGCTACGAGTTTGGCCCAATACGATGATTACGACCGCACCTCGCC
600
GGATCTGAAGAGGGCAGCTGCTGTTGATGACGCGCAGCTACGAGCTGGAAGCGTTCGTCAAGCGCTGAAGCCGACCTTATCGGCTCCGGCATCAAGGAAAAAT
700
ATATCTTCAGAAAAATGGCGTGGCTGCCGCTCCGCGAGATGCACTCGTGGGACTATTCCGGCCGCTACCGGCTACGATGGTTTCGCCATTTTCGCCCGGATATGATATG
800
ACCCCTGAACACCCCGGCGTGAAGCAACTGACCGCTCCGTGGCTGAAGTCTGCGTGATTCGCCACTCACTGTCCCGCTCTGTCACCGATTGTTGGCGCGGGAGAGAACA
900
CCATGAGCCAAACGATTGATAAAATTAATAGCTGTATTCGCTATTTCGAACAGGATGAATACCGAGAGCTGGTTCGCAATAAGCGCGAGCTGGAAGAGCGCCACGATGCG
1000
CAGCGCGTGCGAGGAGCTTCGCGTGACCAACCGCGGATGATGAAGCGCTGAATTCACGAGCGCTGACCGTTGACCGCGGAAAGCGTGGCCACGCTTGGGCGC
1100
GGTCTTTGCTCGCTGGGATTTCGCAACCCCTCGCGTATGTGACGCGCTCTCAGGGTTGGGTGGCTACTTCGCGACCTATTTTAAACGCCATTTCAAAGAGCGGATCG
1200
CCTGCGTCTCGACTCGATGACGCAAGACGCGCGGCTTCGCGCGCAACAACATATGAACCTGGGCTGCAGAACGCGCAGCGCGTGTACAAACCGGAGATCATTTGGC
1300
GTGTCCACCACTGCATGCGGAAGTTATCGCGGATGACCTGCAGCGCTTATTCGCCAACGCTAAAAAAGATGGCTTCGTGACAGCAGCATGCCTGGCCCCACGCCCCA
1400
TACGCCAAGCTTTATCGCGCAGCCAGCTCACCAGCTGGGATAACATGTTTGAAGGCTTCGCCAAACCTTCCTACTGCGGACTACCGAGGCGAGCGGGCAATTCGCCGAAGC
1500
TCAAGCTCAATCTGTCGCCCGCTTTGAAACCTATCTCGGCACTGGCAACTTCGCGTATTAAGCGGATGATGAACAGATGGCGGTGCGTGCAGCTGCTCTCCGAT
1600
CGTCGGAAGTTCTCGACACGCCCGCGCGGATCGGATGATTTCGCGCGCACCGCAGCAGGAGATGAAAGAGGCCCTGACGCCATCGATACGCTGCTCTCT
1700
GCAGCGTGCGAGCTGCTGAAGAGCAAAAAAGTGGTGACGAGATGTGAACACGCCGCCACCGAGGTGCCATTCGCTGGGGCTGGCGCCACCGATGAACCTGCTGA
1800
TGACCGTCAGCGAGCTTAGCGCAAGCCGATTTCGCGACGCCCTCACCTTTAGCGCGCGCGGCTGGTTGACATGATGCTCGACTCCACACCTGGCTGCACGCAAGAAG
1900
TTTGGCTGTAGCGCGATCCGACTTCGTGATGGGCTCACCCGCTTCCTGCTGAGAGCTGGGCTGCGAGCCAAACGCTGATCTGAGCCATAACGCCAAACAGCTGGCA
2000
AAAAGCGATGAACAAATCTCGATGCTCGCGCAATCGGCGCGATACAGCGTGTATTAACATCGGATTTGTGCGACTTCGCTGCTGATGTTCAACCGTCAGCGG
2100
ACTTTATGATCGGCAACTCTACGCGAAGTTTATCCAGCGGTAAACCTGGCGAAGGTTAAAGCTTTGAAGTGCCGCTTATCGCGCTCGGCTTTCCGCTGTTTCGACCGC
2200
CACCATCTGCACGCCAGACAACTGGGGTTATGAAGGGCGATGAACATTGTGACGAGCTGGTGAACCGCTGCTGGAGAACTGGATAGCATACCGCAGCTGGC
2300
CAAAACCGATTACGCTTCGATCTGCTCGTTAATCATCAGTGAACGCTGGGGCGGCGGATTCGCTGGAGGGGAGCGCGCTACTACGTGCATCGCCATGCGGGGGC
2400
GTCCCGCTCTTTCGATAAGCTCCGACGCGCAATACCTTGATATAAGGAGCTAGTGATGGTCCGACACGATACCCCTATTCTGCGGTATCGTGCGCTGTTTCACTCTC
2500
TCCCGGACTACAGCGCGCAAAATCGTCACTGGCTGGCGCAGGAGAGCGCGAGACCGTGACGCCAGAGCGTCTGGCGACCCAGCCGCGCAGCTGGCGCGCAGC

```

References: (1) Ligon, J. and Nakas, J. (1987) Appl. Environ. Microbiol. 53, 2321-2327. (2) Brigle, K. et al. (1985) Gene 37, 37-44. (3) Thöny, B. et al. (1985) Mol. Gen. Genet. 198, 441-448. (4) Hazur, B. and Chui, C. (1982) Proc. Natl. Acad. Sci. USA 79, 6782-6786.