An open reading frame upstream from the *nifH* gene of *Klebsiella pneumoniae*

San-chiun Shen*, Zhong-tian Xue*, Qiu-tong Kong+ and Qi-long Wu+

*Division of Biology, California Institute of Technology, Pasadena, CA 91125, and +Laboratory of Molecular Genetics, Shanghai Institute of Plant Physiology, Academia Sinica, Shanghai, China

Received 21 March 1983; Revised 13 May 1983; Accepted 19 May 1983

ABSTRACT

An open reading frame upstream from nifHDK operon of Klebsiella pneumoniae had been described. The orientation of this open reading frame is opposite to that of nifHDK and sequence homology was found between the open reading frame promoter and the promoter of nifHDK operon.

A recombinant plasmid carrying the promoter region of the open reading frame fused to the β-galactosidase gene was constructed. Strains of E.coli were transformed with the plasmid containing this open reading frame promoter-lacZ fusion or co-transformed with it and a plasmid carrying the nifA gene. An appreciable activity of β-galactosidase was found in strains which received both plasmids, indicating that the promoter of the open reading frame can be activated by the product of nifA gene. Thus, the open reading frame found between nifHDK operon and nifJ behaves just like other nif genes of K.pneumoniae in requiring the product of nifA as the positive effector for expression.

INTRODUCTION

The nif gene cluster of Klebsiella pneumoniae consists of at least 17 genes organized into 5 polycistronic and 3 monocistronic transcriptional units near the promoter end of the histidine biosynthesis operon (1,2). The nif region is approximately 24 kilobases long and no non-nif genes seem to be interspersed within the nif region (3,4). A mutation nifC located between nifH and nifJ, genetically complemented mutations in each of the 17 known nif genes, and thus had been proposed as a new gene nifC (5). However, further investigation indicates that the observed complementation between the nifC mutation and nifJ mutation is due to the result of intragenic complementation between two inactive nifJ polypeptides (6). Therefore, nifC could not be defined as a separate gene from nifJ.

Mevarech et al. (7) sequenced the nifH gene coding for the

nitrogenase in cyanobacterium Anabaena, they found an open reading frame (ORF) preceding the nifH starting at residue -429 and running for 68 amino acids to -225. Whether this region is actually transcribed and translated in Anabaena is unknown.

Sundaresan and Ausubel (8,9) found that in K. pneumoniae a DNA sequence upstream from nifH gene can be transcribed in the opposite direction to nifH in vitro. These facts prompted us to investigate the nucleotide sequence between the nifH and nifJ in K.pneumoniae.

The results presented below show the existence of an ORF of 384 nucleotides upstream from nifH gene. The orientation of this ORF appears opposite to that of all nif genes. Its expression has been monitored by using an ORF-lacZ fusion carried on small plasmid.

MATERIALS AND METHODS

Sequencing of nif DNA

The plasmid pMC1 obtained from Ausubel's laboratory was constructed by subcloning from pMF6 the Bam H1 - Xho I fragment of K.pneumoniae DNA which contains sequences covering partial nifD through nifH to nifJ into the vector pACYC177 (10) (Fig. 1). It was used as the source of DNA for sequencing. The pMC1 DNA was prepared as described by Humphreys and Davis (11,12).

20 µg of pMC1 DNA was treated with different endonucleases which were chosen as being convenient for DNA sequencing on the basis of restriction mapping of the Bgl II - Bgl II fragment of the insert in pMC1. The resulting fragments were separated by electrophoresis in 0.8-1.2 % agarose gel or in 5 % acrylamide-bis-acrylylcystamine gel. In the former case the DNA was eluted and purified by passage over BD-cellulose column, in the latter case the DNA after eluted was mixed with β-mercaptoethanol and passed over DE-52 column.

The method used for labeling and sequencing were those of Maxam and Gilbert (13). After restriction of DNA fragment, 5'-phosphates were removed with alkaline phosphatase and the resulting 5'OH termini were labeled with γ-^{32}P, using γ-^{32}P ATP (1 mCi) and polynucleotide kinase.

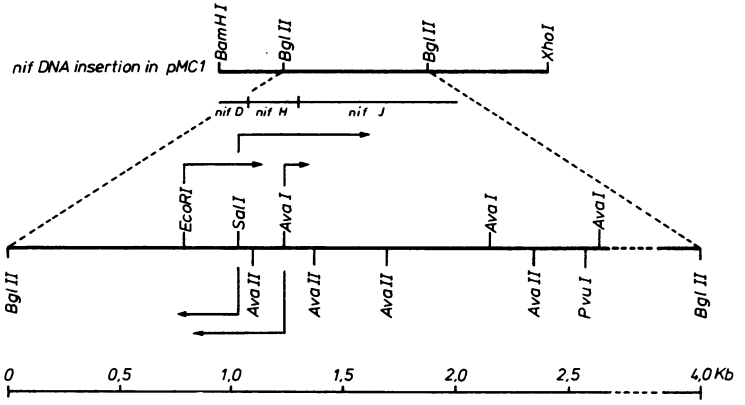

Fig 1. Physical map of BglII-BglII fragment of pMC1 insertion and the sequencing strategy. The arrows indicate the extent and direction of DNA fragment sequence determinations.

Construction of ORF-lac fusion in plasmid

Plasmid pRZ-5202 containing lacZ gene but devoid of promoter was kindly provided by Reznikoff, and plasmid pST1021 which carries nifA under the Tc^R promoter was constructed by Zhu (14) in this laboratory. The restriction endonuclease digestion, e.g. Bam H1, Bgl II, Eco RI or Sal I digestion was performed as described elsewhere (15). During ligation the restriction endonuclease digestion mixture was adjusted to contain T4 DNA ligase in 0.2 mM ATP, 10 mM $MgCl_2$, 10 mM DTT, 100mM NaCl, 50 mM Tris-HCl (pP 7.6) and 200 µg/ml bovine serum albumin. The reaction mixture was incubated for 14-18 hr. at 12°C.

Assay for β-galactosidase activity

Cultures to be assayed were grown anaerobically for 16-24 hr. in derepressing condition. β-galactosidase was assayed as described previously(16).

RESULTS AND DISCUSSION

Nucleotide sequence upstream from nifH

The Bgl II - Bgl II fragment from pMC1 DNA, 4 Kb in length which covers the region from the N-terminus coding region of nifH to part of nifJ was mapped with restriction enzymes as outlined in Fig. 1.

A region from the restriction site of Eco RI through Sal I -

```
5'..... AATTCGCGCTAACTCTTCTGTCATCCGCGAGCTGGCACAGGCTGTGCTTGAGG
        EcoRI                -100

CAACAACTGGGTTTGCCGCTTATTTGAGGCGGCGAAGATGCCTCTCACCTTCGAGGATTGAA
              -50

   met ser gly lys met lys thr met asp gly asn ala ala ala ala
CA ATG TCC GGA AAA ATG AAA ACA ATG GAT GGC AAC GCC GCG GCG GCC
   0

trp ile ser tyr ala phe thr glu val ala ala ile tyr pro ile thr
TGG ATC TCT TAT GCC TTT ACC GAG GTC GCG GCG ATT TAC CCC ATA ACC
      50

pro ser thr pro met ala glu asn val asp glu trp ala ala gln gly
CCC TCC ACG CCG ATG GCG GAA AAC GTC GAC GAG TGG GCG GCG CAG GGG
         100                      Sal I

lys lys asn leu phe gly gln pro val arg leu met glu met gln ser
AAA AAG AAC CTT TTT GGC CAG CCG GTG CGC TTA ATG GAG ATG CAG TCG
            150

glu ala gly ala ala gly ala val his gly ala leu gln ala gly ala
GAG GCC GGC GCG GCA GGC GCG GTC CAC GGC GCG CTG CAG GCC GGG GCG
                200          Ava II

leu thr thr thr tyr thr ala ser gln gly leu leu leu met ile pro
CTC ACC ACC ACC TAT ACG GCC TCC CAG GGG CTG CTG CTG ATG ATC CCC
                   250

asn met tyr lys ile ala gly glu leu leu pro gly val phe his val
AAC ATG TAC AAA ATC GCC GGT GAA CTG CTG CCG GGC GTC TTT CAC GTC
                      300

ser ala arg ala leu ala thr asn ser leu asn ile phe gly asp his
AGC GCC CGG GCG CTG GCG ACC AAT TCG CTG AAT ATT TTT GGC GAT CAC
      Ava I            350

TAG GATGTGATGGCGGTCCGCCAGAGCGGCTGCGCGATGCTGGCGGAGAACAA......3'
               AvaII
```

Fig. 2 Nucleotide sequence of an open reading frame and its flanking parts in the upstream of nifH. The sequence shown is complementary to the coding strand. The amino acids sequence deduced from it are shown.

Ava II - Ava I - Ava II site (the 2nd Ava II site from the left end) of the Bgl II - Bgl II fragment was selected for sequencing. About 87 % of the nucleotide sequence presented here was determined on both strands. Only the small fragment Ava I - Ava II was sequenced on one strand, though performed twice, starting from different restriction sites. The complete nucleotide sequence of the Eco RI - Sal I - Ava II - Ava I - Ava II region complementary to the coding strand is shown in Fig. 2. From the Eco RI site at 117 bp, an initiation codon ATG is followed by an ORF of 384 bp. Preceding the ORF, a sequence homologous to the "Shine-Dalgarno" ribosome binding sequence, A-G-G-A is present at nucleotides -11

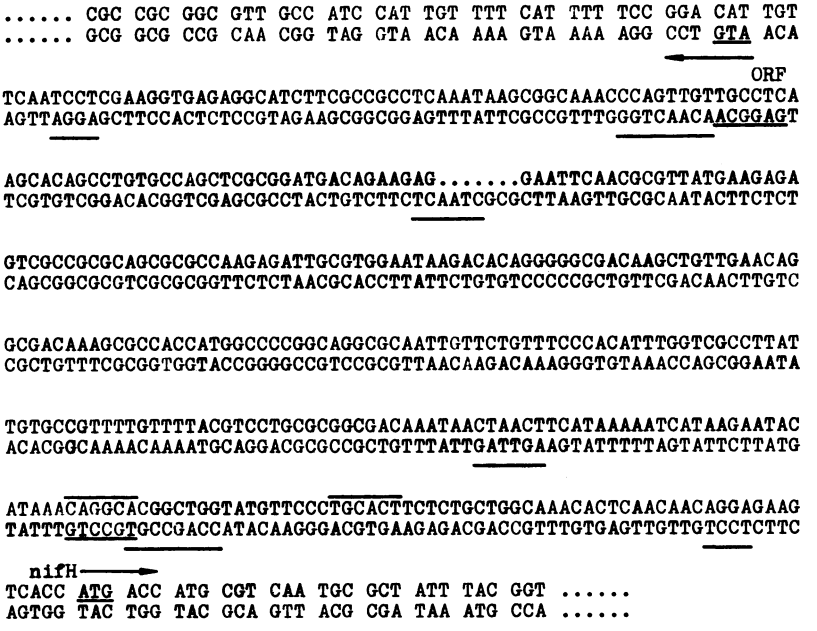

Fig. 3 Nucleotide sequence between nifH and the open reading frame upstream from nifH. The -30 and -10 sequences of nifHDK promoter are overscored and the regions underlined represent the sequence of homology shared by the putative promoter of the open reading frame and the nifHDK promoter.

to -8. There is no sequence that corresponds closely to the consensus promoter TATAAT found about 10 bp upstream from the transcription starts in E.coli (17). The closest fit is the sequence TTATTT found at -47 to -40 nucleotides upstream from the coding region. The sequence TGG..... starting at nucleotide 392 and ending .....CCA at nucleotide 404 can be folded to form a stem structure that might indicate a termination sequence.

The following facts emerged from a comparison of the DNA sequence of the putative ORF promoter and the promoter of nifHDK operon: (1) the sequence located at -10 in nifHDK promoter, according to Sundaresan et al. (18) is different from the sequence in the ORF promoter, while the sequence located about -30 in nifHDK promoter is in good homology to the corresponding sequence in the ORF promoter, (2) an 8-nucleotide sequence ACAACTGG closely homologous to the sequence ACGGCTGG in nifHDK promoter, which

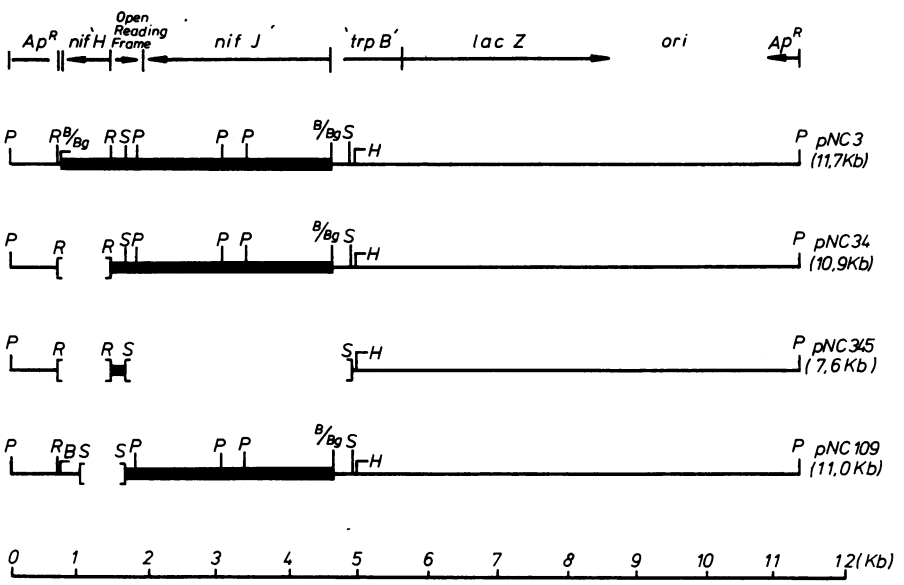

Fig 4. Construction and physical maps of plasmids with nif ORF promoter-lacZ fusions. Thick line indicates nif DNA; thin line vector DNA; parenthesis, deletion.
P denotes PstI;R,EcoRI;S,SalI;H,HindIII;B,BamHI;Bg,BglII.

is regarded as the recognition site of the activation protein is shown in ORF promoter, (3) the homologous sequences appear to be located in corresponding regions of both promoters. Fig. 3 shows the nucleotide sequence between nifH and the ORF, indicating the different orientation between these two DNA sequences and the homology regions shared by the ORF promoter and the promoter of nifHDK operon.

Construction of nif-ORF promoter-lacZ fusion

In order to substantiate the activity of this ORF, we constructed a gene fusion between the promoter region of ORF and the lacZ to test whether the ORF promoter can initiate the transcription of lacZ in this fusion. The schedule for making such a gene fusion is illustrated in Fig.4.

The plasmid pRZ5202 which contains lacZ gene but not its promoter was used as a vector for constructing ORF promoter-lacZ fusion. A nif DNA fragment containing the ORF sequence was obtained from the plasmid pMC1 by Bgl II restriction. The Bgl II - Bgl II fragment which covers the region from the N-terminus coding

region of nifH to the structure of part of nifJ gene was isolated and cloned into the Bam H1 site of the plasmid pRZ5202. Its orientation was checked by examining the restriction sites of Eco RI and Sal I of the insert related to the restriction map of the plasmid DNA (results are not shown here). A clone pNC3 thus obtained contains nif'H-J' DNA fused with the lacZ at the same direction of transcription. It was then subject to Eco RI restriction to eliminate a small Eco RI fragment, 0.71 Kb in length, which contains the promoter region of nifH gene. The remaining large Eco RI fragment was then self-circled, generating a deletion derivative pNC34 which contains the nif DNA spanning the region of ORF and part of nifJ gene upstream from lacZ. This plasmid was further restriction with Sal I. After removal of the small Sal I restriction fragment, 3.3 Kb in length, a resulting plasmid pNC345 containing only the promoter region of ORF fused to lacZ was obtained. In this nif-lac fusion, lacZ is directly under the control of the promoter of ORF.

A plasmid which deleted the ORF promoter was also constructed. We started with the plasmid pNC3 and restricted it with Sal I, the Sal I fragment, 3.3 Kb in size was selected and subcloned into the Sal I site of the plasmid pRZ5202, maintaining an orientation as it was originally in plasmid pNC3. The recombinant plasmid pNC109, thus constructed contains only the C-terminus coding region of nifJ and the region of ORF devoid of its promoter. This nif-lac fusion was used as a control in the experiment of testing the activity of ORF promoter to initiate the lacZ expression.

Activity of nif-ORF promoter

For examining the activity of the promoter of ORF (pORF), the above constructed plasmids, pNC34, pNC345 and pNC109, or their combination with the nifA carrying plasmid pST1021 (14) were introduced respectively to the lac-deletion mutant of E.coli, SY203. The transformants were grown under anaerobic derepressing conditions and β-galactosidase was measured. As shown in table 1, little β-galactosidase activity was found in strain SY2031 and SY2032 with the pORF::lac fusions. However, an appreciable activity of β-galactosidase was demonstrated if nifA was present. It indicates that ORF promoter is activated to initiate the trans-

Table 1 Activation of *nifORF* promoter by *nifA* product

Strain	*Relevant genotype	β-galactosidase units	
		$-NH_4^+$	$+NH_4^+$
SY 2031 (pNC 34)	ORF-*nifJ*'-*lacZ*	8	10
SY 2031-1 (pNC 34) (pST 1021)	ORF-*nifJ*'-*lacZ* *nifA*	51	22
SY 2032 (pNC 345)	ORF promoter-*lacZ*	10	15
SY 2032-1 (pNC 345) (pST 1021)	ORF promoter-*lacZ* *nifA*	180	64
SY 2033 (pNC 109)	'ORF-*nifJ*'-*lacZ*	0.9	1.5
SY 2033-1 (pNC 109) (pST 1021)	'ORF-*nifJ*'-*lacZ* *nifA*	1.6	1.9

*ORF-*nifJ*'-*lacZ* denotes the DNA fragment which contains the whole ORF and partial *nifJ* fused to *lacZ*, 'ORF-*nifJ*'-*lacZ*, the promoter deleted ORF fused to *lacZ*. Cultures were grown in LB media containing ampicillin 100 μg/ml at 37°C, for about 7 hrs. and then were inoculated to NFDM (20) supplemented with biotin 5 μg/ml, L-arginine 25 μg/ml, L-proline 25 μg/ml and casamino acids 50 μg/ml, incubating under N_2 anaerobically for 20 hrs. For testing the repressive effect of ammonium, 1 mM of $(NH_4)_2SO_4$ was added. Bacterial suspensions were used for β-galactosidase assay. The enzyme activity units are defined in Miller (16).

cription of *lacZ* gene by the product of *nifA* gene. The lower activity of β-galactosidase in a strain with pNC34 and *nifA* is probably due to a distance between the ORF promoter and *lacZ* in the fusion.

Excess ammonium exerts some repressive effect on the expression of the ORF promoter in the presence of *nifA*. Since the *nifL* product is absent in *E.coli*, so ammonia repression of ORF transcription remains to be elucidated.

In conclusion, an open reading frame in opposite orientation to that of nifHDK operon was demonstrated. The ORF promoter can be activated by nifA product as are promoters of other nif operons (15,19). Nevertheless, we still cannot be sure whether this reading frame represents a real nif gene and nor its role in vivo. Genetic investigations are underway to answer these questions.

ACKNOWLEDGEMENTS

We would like to thank Dr. F. Ausubel for kindly providing us the pMC1 plasmid, unpublished nifH sequence data and for information concerning the in vitro transcription of a DNA sequence upstream from nifH. We are also grateful to Dr. J. Bonner and Dr. E. M. Meyerowitz for their valuable suggestions. S S-C and X Z-T were supported by E.S.Gosney fund.

REFERENCES

1. MacNeil, T., MacNeil, D., Roberts, G. P., Supiano, M. A. and Brill, W. J. (1978) J. Bact. 136, 253-266.
2. Merrick, M., Filser, M., Dixon, R., Elmerich, C., Sibold, L. and Houmard, J. (1980) J. Gen. Microbiol. 117, 509-520.
3. Cannon, F. C., Riedel, G. E. and Ausubel, F. M. (1979) Mol. Gen. Genet. 174, 56-66.
4. Puhler, A. and Klipp, W. (1981) Boethe H, Trebst A(eds) Biology of inorganic nitrogen and sulphur. Springer-verlag, Berlin, Heidelberg and New York, 276-286.
5. Jin, R. Z., Huang, Y. D., Shen, M. C., Chen, J. B., Go, D. A., Shen, B. F., Li, W. J. and Shen, S. C. (1980) Scientia Sinica, XXIII, No. 1, 108-118.
6. Stacey, G., Zhu, J. B., Shah, V. K., Shen, S. C. and Brill, W. J. (1982) J. Bact. 150, 293-297.
7. Mevarech, M., Rice, D. and Haselkorn, R. (1980) Proc. Natl. Acad. Sci. USA 77, No. 11, 276-286.
8. Sundaresan, V. and Ausubel, F. M. (1981) J. biol. Chem. 256, No. 6, 2808-2812.
9. Ausubel, F. M. Personal communication.
10. Cannon, M. unpublished results.
11. Humphreys, G. O., Willshaw, G. A. and Anderson, E. S. (1975) B.B.A. 282, 457-463.
12. Davis, R.W., Roth, J. R. and Botstein, D. (1980) Experiments in Genetic Engineering, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
13. Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
14. Zhu, J. B., Yu, Q. C., Jiang, Q. Y., Wang, L. W. and Shen, S. C., Scientia Sinica, in press.
15. Kong, Q. T., Wu, Q. L., Jia, L. B., Syvanen, M., Lin, E. C. C. and Shen, S. C. (1982) Scientia Sinica, XXV, No. 10, 1061-1070.
16. Miller, J. H. (1972) Experiments in Molecular Genetics., Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y. 352-355.
17. Rosenberg, M., Court, D. (1979) Annu. Rev. Genet. 13, 319-353.

18.Sundaresan, V., Jones, J. D. G., Ow, D. W. and Ausubel, F. M., in press.
19.Buchanan, W. V., Cannon, M. C., Beynon, J. and Cannon, F. C. (1981) Nature, 294, 776-778.
20.Cannon, F. C., Dixon, R.A. and Postgate, J. R. (1974) J. Gen. Microbiol. 80, 227-239.