Results
RESULTS

The results of the present study are described in the following sections:

1. Characterization of the \textit{nifH} homologous region in the cosmid pMP1.

2. Mutagenesis of the ferredoxin-like gene (Av-ORF2) downstream of the \textit{nifH} hybridizing gene.

3. Construction of a \textit{lacZ}-fusion of the promoter of the \textit{nifH} homologous gene and study of its regulation.

4. Localization of the UAS (upstream activator sequence) for this promoter where the positive regulatory protein binds.

5. Isolation of a regulatory mutant in which the alternative pathway is not repressed by Mo.

1. Characterization of the \textit{nifH} Homologous Region in the Cosmid pMP1

From the genomic library of \textit{A. vinelandii}, three recombinant cosmids were isolated which hybridized to the \textit{nifHDKY} probe of \textit{K. pneumoniae} (Medhora \textit{et al.} 1983). A cosmid (pMP2) was found to contain the \textit{nifHDK} genes, structural genes for the enzyme nitrogenase (Reddy 1986). Another cosmid (pMP1) was found to have sequences homologous to \textit{nifH} only. The cosmid pMP1 is a 35 kb \textit{A. vinelandii} genomic DNA cloned in pHC79. The insert of pMP1 (30 kb \textit{BglII} fragment) was subcloned in a broad host range vector, pRK404 and the resultant plasmid pAM103 was used for the
complementation of the \textit{nifH} mutant of \textit{A. vinelandii} (Reddy 1986). This plasmid could not complement the mutation in \textit{nifH} gene. The restriction maps of pMP1 and pMP2 (Fig.9) suggest that the insert of pMP1 and pMP2 come from different regions of the \textit{A. vinelandii} genome (Reddy 1986). In the present work, the cosmid pMP1 has been characterized. It was speculated that pMP1 may have a \textit{nifH} equivalent gene for the alternative pathway of nitrogen fixation. Characterization of this \textit{nifH} homologous sequence of pAM103 involved following steps:

1.1 \textbf{Subcloning of the 6 kb \textit{BamHI-HindIII} fragment from pAM103}: The plasmid pAM103 on digestion with \textit{BamHI} and \textit{HindIII} gave five bands. A band of about 6 kb hybridized with the \textit{nifH} probe (dark rectangle, Fig.9). This 6 Kb DNA fragment was eluted from low melting agarose and cloned in pUC19. It was designated as pRR1.

The plasmid pRR1 on digestion with \textit{SalI} and \textit{PstI} gave 7 and 5 bands respectively. Out of these 7 \textit{SalI} bands, a single band of about 1 kb hybridized to the \textit{nifH} probe. With \textit{PstI}, two bands of 2.5 and 1.25 kb hybridized. It was difficult to map the \textit{SalI} and \textit{PstI} fragments on the plasmid pRR1, since there were no other convenient restriction enzyme sites. The restriction map of pRR1 was therefore made by transposon mutagenesis using the transposon Tn5.

1.2 \textbf{Tn5 mutagenesis}: Tn5 insertions were obtained on the insert of pRR1 for two purposes:

(a) To make the restriction map of the insert of pRR1.

(b) To get a clone which has Tn5 insert in the \textit{nifH} hybridizing \textit{SalI} fragment, which was to be used later for marker exchange with \textit{A. vinelandii} to get a Tn5 mutant.
Fig.9: Restriction map of pAM103, pMP1 and pMP2.

Abbreviations: Bg, BgII; H, HindIII; Bm, BamHI; R, EcoRI; P, PstI, Kp, KpnI; S, SalI; Sm, Smal. The dark rectangles represent nif sequences and the hatched rectangles represent pHCl79 (From Reddy 1986).
Since pUC derivatives cannot be mobilized, the 6 kb insert in plasmid pRR1 was cloned in the mobilizable plasmid vector pSUP102 (Fig.10). The plasmid pSUP102 is a mobilizable E. coli vector carrying a RP4-specific Mob-site (about 2 kb) cloned in the BclII site of pACYC184 (Simon et al. 1986).

The plasmid pRR1 was digested with BamHI and HindIII, the 6 kb insert was eluted from the low melting agarose and was ligated to pSUP102 digested with BamHI and HindIII. Competent E. coli TB1 cells were transformed with the ligated mixture, plated on L-agar medium with Cm (30 μg/ml) and incubated at 37°C overnight. About 50 transformants appeared. These were checked for Tet sensitivity (insertional inactivation of Tet gene due to cloning). The plasmid was isolated from a few Tet™ colonies and digested with BamHI and HindIII and electrophoresed on a mini gel. Two bands of nearly equal size were seen, one of insert and the other of vector. This plasmid was designated as pRR5 (Fig.10).

The competent E. coli HB101::Tn5 (Kan) cells were transformed with the plasmid pRR5 (Cm) and plated on L-agar medium with Cm (30 μg/ml) and Kan (50 μg/ml). After 36 hrs incubation at 37°C, about 3000 CmKan transformants appeared. All the transformants were pooled in 20 ml LB without antibiotics and incubated in a 37°C shaker for 2-3 hrs.

Equal volumes of (100 μl) each of pooled transformants, overnight grown E. coli culture containing helper plasmid pRK2073 (Spec) and overnight grown recipient E. coli C600 (Rif) culture were mixed on L-agar plate. The plate was incubated at 37°C for 12 hrs. Cells were scrapped out in 10 ml of normal saline. Serial dilutions (10⁻¹, 10⁻², 10⁻³, 10⁻⁴, 10⁻⁵) of this cell suspension was plated on three sets of media:
Fig. 10: Restriction map of pSUP102 and pRR5. Abbreviations: Cm, chloramphenicol; Tc, Tetracycline; E, EcoRI; C, Clal; H, HindIII, B, BamHI, S, SalI. Dark region indicates insert.
a) L-Agar plates containing Cm (30 μg/ml), Kan (50 μg/ml) and Rif (10 μg/ml).

b) L-Agar plates containing Cm (30 μg/ml) and Rif (10 μg/ml).

c) L-Agar plates containing Rif (10 μg/ml).

All the plates were incubated at 37°C. Colonies appeared on L-agar (Rif) plates after 12 hours, on L-agar (Rif, Cm) plates after 24 hrs and on L-agar (Rif, Cm, Kan) after 48 hrs of incubation. On L-agar (Rif) plates in 10⁻⁶ dilution, about 150 colonies appeared (1.5 x 10⁶ cells/ml); on L-agar (Cm, Rif) in 10⁻³ dilution, about 365 colonies appeared (3.6 x 10⁶ cells/ml) and on L-agar (Cm, Rif, Kan) in 10⁻¹ dilution, about 40 colonies appeared (4 x 10³ cells/ml). Based on these data, the conjugation frequency was calculated to be 10⁻⁶ per recipient and transposition frequency was calculated to be 10⁻⁶ per recipient. All the 40 colonies on L-agar (Cm, Rif, Kan) plates were pooled in 20 ml LB with Cm, Rif, and Kan and grown overnight. Plasmid DNA was isolated by the alkaline lysis method. About 0.5 μg of DNA was used to transform the E. coli S17.1 strain.

The E. coli S17.1 strain has portions of the self-transmissible antibiotic sensitive derivative of the plasmid RP4 integrated into its chromosome. Thus, the transfer functions of RP4 are used here to build the conjugation bridges to appropriate recipients to mobilize the vector plasmid e.g. pSUP102, which has its own mob gene (Simon et al. 1986).

The transformants were selected on L-agar plates containing Cm (30 μg/ml) and Kan (50 μg/ml). Approximately 1000 Cm Kan transformants appeared after 24 hrs incubation at 37°C. Plasmid DNA was isolated from 100 single colonies by the alkaline lysis method and digested with SalI to see which SalI band was hit by Tn5. Fig.11 shows a few pRR5::Tn5 clones digested with SalI.
Fig. 11: Restriction pattern of some pRR5::Tn5 mutants after digestion with SalI. Lane a, pRR25; b, pRR27; c, pRR29; d, pRR32; e, pRR5; f, λ DNA digested with HindIII; g, pRR68, h, pRR62; i, pRR67.
The transposon Tn5 is about 5.7 kb in size and has two HindIII sites in inverted repeats at position 1.19 kb and 4.5 kb. Fig. 12 shows the restriction map of Tn5. The plasmid pRR5 has only one HindIII site and therefore, when plasmid pRR5::Tn5 were digested with HindIII, they gave three bands. Out of these three bands, one (the central region of Tn5, about 3.3 kb in size), was a common band in all the mutants. The size of the other two bands varied depending on the position of the insertion of Tn5. The position of Tn5 on pRR5 was calculated by subtracting 1.2 kb (distance of HindIII site in inverted repeat). Fig. 13 shows the results of digestion of pRR5::Tn5 mutants with HindIII. The distance of Tn5 from the HindIII site of pRR5 is shown in Table 7. Based on these data, the position of the Tn5 insertions in pRR5 was mapped. Fig. 15 shows the position of Tn5 insertions in pRR5.

1.3 Restriction mapping of pRR1: Based on the distance of Tn5 insertion from the HindIII site (Table 7), the 7 SalI bands were arranged in pRR1. The position of other restriction enzyme sites were mapped after digesting the pRR1 with PstI, SalI, EcoRI, BamHI, HindIII. Fig. 14 shows the results of digestion of pRR1 with these enzymes. This gel was blotted onto a nitrocellulose membrane and hybridized with the probe made of the nifH gene to find out the hybridizing region in the different digests. Fig. 15 shows the restriction map of pRR5 (insert of pRR1 cloned in BamHI-HindIII site of pSUP102).

1.4 Marker exchange: In this technique, the transposon mutant obtained after transposon insertion in a plasmid is introduced into the wild type bacterium. The vector being unstable, is lost and the transposon insertion from the plasmid is transferred to the chromosome by homologous recombination. The position of the Tn5 on the chromosome is confirmed by hybridization of the blot of restriction enzyme digested genomic DNA with the 32P-labelled insert probe.
Fig. 12: Restriction map of transposon Tn5.

Abbreviations: B, BglII; H, HindIII; Bm, BamHI; R, EcoRI; P, PstI; Sm, SmaI; X, XhoI; S, SalI; Kan, kanamycin; Nm, Neomycin; Sn, Streptomycin.
Fig. 13: pRR5::Tn5 mutants digested with HindIII and BamHI.
Lane a, pRR25; b, pRR27; c, pRR29; d, pRR32; e, pRR68; f, pRR62; g, pRR67; h, λDNA; i, pRR5, with HindIII; j, pRR25; k, pRR27; l, pRR29; m, pRR32; n, pRR68; o, pRR62; p, pRR67; digested with BamHI.
### Table 7: Tn5 insertions and their position from HindIII site in pRR5

<table>
<thead>
<tr>
<th>Band number from origin</th>
<th>Sizes (kb) pRR5/SaI</th>
<th>Tn5 Insertion number</th>
<th>Designation of plasmid after Tn5 insertion</th>
<th>Position of Tn5 in pRR5 from HindIII site (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.7</td>
<td>25</td>
<td>pRR25</td>
<td>3.2</td>
</tr>
<tr>
<td>2</td>
<td>2.475</td>
<td>27</td>
<td>pRR27</td>
<td>4.55</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>62</td>
<td>pRR62</td>
<td>5.2</td>
</tr>
<tr>
<td>4</td>
<td>0.95</td>
<td>29</td>
<td>pRR29</td>
<td>3.1</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>32</td>
<td>pRR32</td>
<td>4.8</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
<td>67</td>
<td>pRR67</td>
<td>0.2-0.3</td>
</tr>
<tr>
<td>7</td>
<td>0.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>0.35</td>
<td>68</td>
<td>pRR68</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Fig. 14: Restriction mapping of pRR1.

A) pRR1 digested with lane: a, Xhol; b, Smal; c, BamHI-HindIII; d, PstI-SalI; e, PstI-EcoRI; f, PstI-BamHI; g, PstI-HindIII; h, PstI; k, SalI-EcoRI; l, SalI-HindIII; m, SalI-BamHI; n, SalI; b, HindIII-EcoRI; p, EcoRI; i, φX174 DNA digested with Haelll (size marker); lane j, 1 kb ladder (size marker).

B) Autoradiogram of the Southern blot of the gel after hybridization with nifH gene of A. vinelandii.
Fig. 15: Restriction map of pRR5. Abbreviations: H, HindIII; P, PstI; S, SalI; R, EcoRI; Bm, BamHI; C, CiaI. The dark rectangle represents the nifH homologous sequence. The broken lines represent the vector pSUP102. The numbered arrows above the map show the position and number of the Tn5 insertions obtained in pRR5 after Tn5 mutagenesis. The arrows below the map show the region sequenced by Sanger's dideoxy method. The broken arrows show the region sequenced by Maxam and Gilbert's method. The dark circle with horizontal arrow above the map shows the position and direction of transcription of nifH promoter. The horizontal arrows below the map show the sequencing strategy.
The plasmid pRR29, in which the Tn5 insertion is in the region which hybridizes with \textit{nifH} of \textit{A. vinelandii} and plasmid pRR62, which has Tn5 insertion outside the \textit{nifH} hybridizing region, were mobilized into \textit{A. vinelandii} CA12 by conjugation between \textit{E. coli} S17.1 containing these plasmids and \textit{A. vinelandii} CA12. \textit{A. vinelandii} CA12 is a \textit{nifHDK} deletion mutant of wild type \textit{A. vinelandii} and can fix nitrogen only under molybdenum deficient conditions. The exconjugants were selected on BNF agar plates supplemented with 0.11% ammonium acetate (BNF) and Kan (50 µg/ml). Single colonies were replica plated by patching on BNF, Cm (200 µg/ml) and BNF Kan (50 µg/ml) to check for a Cm colony (loss of vector). High concentration of Cm was used, because \textit{A. vinelandii} CA12 itself is resistant to Cm upto 175 µg/ml. One colony out of 100 colonies screened, was found to be Cm, in case where pRR29 was mobilized, indicating two point cross over and thus loss of vector. Similarly, two colonies out of 300 colonies screened were found to be Cm in cases where pRR62 was mobilized. These two mutants were designated \textit{A. vinelandii} RR29 and \textit{A. vinelandii} RR62 respectively.

1.5 Hybridization pattern of chromosomal DNA from \textit{A. vinelandii} RR29 and \textit{A. vinelandii} RR62: To confirm that these mutants were the result of homologous recombination leading to marker exchange, total genomic DNA of these mutants was isolated, digested with \textit{SalI} and the Southern blots were hybridized with the \textit{nifH} probe. \textit{A. vinelandii} CA12 chromosomal DNA was used as control (Fig.16A,B). A \textit{SalI} fragment of 1 kb from \textit{A. vinelandii} CA12 (lane 2) hybridized to the probe. This corresponds to the 1 kb \textit{SalI} fragment of \textit{A. vinelandii} pRR5 (lane 6). With \textit{A. vinelandii} RR29 chromosomal DNA (lane 1), the two bands that hybridized, corresponded to 3.5 and 2.75 kb bands from pRR29 (lane 5).
Fig. 16: Hybridization analysis of genomic DNA from *A. vinelandii* Tn5 mutants. All DNA samples were digested with *SalI*. (A) Gel photograph showing electrophoretic separation of (i) genomic DNA of: RR29, lane 1; CA12, lane 2, (ii) plasmid DNA of: pRR29, lane 5; pRR5, lane 6, (iii) DNA size markers: φX174 DNA digested with *HaeIII*, lane 3; λDNA digested with *HindIII*, lane 4. (B) Autoradiogram of the blot of gel 'A' after the blot was hybridized with *nifH* gene probe of *A. vinelandii*. 
The SalI-digested chromosomal DNA from both A. vinelandii RR29 and A. vinelandii RR62 was also hybridized with the central HindIII fragment of Tn5 containing the KanR gene. The autoradiogram (Fig.17A,B) showed a band of 3.3 kb in A. vinelandii RR62 DNA (lane 2,3; these lanes contain the chromosomal DNA isolated from 2 Cm$^+$ colonies which had the same mutation as in pRR62) and two bands of 3.5 kb and 2.75 kb in A. vinelandii RR29 DNA (lane 1). These fragments correspond to those in pRR62 (Fig.17A, lane 6) and pRR29 (Fig.16A, lane 5). These results show that Tn5 is inserted into the chromosome by homologous recombination in the position as it was on the plasmid.

1.6 Characterization of mutants: Growth of A. vinelandii CA12 and its Tn5 mutants under different growth conditions is shown in Table 8.

The results of Table 8 show that A. vinelandii CA12 and its mutants grew equally well in BNF medium. Mo-deficient BNF, supplemented with 50 nM V$_5^+$ (under nitrogen fixing conditions) both A. vinelandii CA12 and A. vinelandii RR62 grew equally well, but A. vinelandii RR29 failed to grow.

Acetylene reduction assay (a measure of nitrogen fixation) of A. vinelandii CA12, RR29 and RR62 was also done. Fig.18 shows the extent of reduction of acetylene by these three strains. Thus A. vinelandii CA12 and RR62 reduced acetylene (57% and 62% of the total acetylene injected) in the Mo-deficient conditions in the presence of 50 nM V$_5^+$ whereas A. vinelandii RR29 could exhibit a maximum of 1.5% reduction.

It can thus be concluded that the Tn5 insertion in the A. vinelandii RR62 does not affect the growth of A. vinelandii under nitrogen fixing conditions in the presence of vanadium, whereas the insertion in A. vinelandii RR29 makes it NIF under similar
Hybridization analysis of genomic DNA from *A. vinelandii* Tn5 mutants. All DNA samples were digested with *SalI*. (A) Gel photograph showing electrophoretic separation of (i) genomic DNA of: RR29, lane 1; RR62 (Clone 1), lane 2; RR62 (Clone 2), lane 3; CA12, lane 4; (ii) plasmid DNA of pRR5, lane 5; pRR62, lane 6; (iii) DNA size markers: φX174 DNA digested with *HaeIII*, lane 7; λDNA digested with *HindIII*, lane 8. (B) Autoradiogram of the blot of gel 'B' after the blot was hybridized with the central *HindIII* fragment of Tn5 as probe.
Table 8: Growth of *A. vinelandii* CA12 and its mutants under different growth conditions after 3 days of inoculation

<table>
<thead>
<tr>
<th>Strain</th>
<th>Ammonium acetate (0.11%)</th>
<th>V$_2$O$_5$ (50 nM)</th>
<th>Kan (50 µg/ml)</th>
<th>O.D. at 620 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA12</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>2.280</td>
</tr>
<tr>
<td>CA12</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>1.426</td>
</tr>
<tr>
<td>RR62</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>2.195</td>
</tr>
<tr>
<td>RR62</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>1.398</td>
</tr>
<tr>
<td>RR62</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0.986</td>
</tr>
<tr>
<td>RR29</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>2.205</td>
</tr>
<tr>
<td>RR29</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0.206</td>
</tr>
<tr>
<td>RR29</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0.095</td>
</tr>
</tbody>
</table>
Fig. 18: Acetylene reduction assay of CA12, RR29 and RR62 in presence of 2 mM urea and 50 nM V$_2$O$_5$. 'X' indicates point of injection of gas sample.
conditions. *A. vinelandii* RR29 was NIF even in the absence of Kan (Table 8) suggesting that it is a stable mutant. Since, this Tn5 insertion is in the region of DNA homologous to *nifH*, it is proposed that this copy of *nifH* is involved in the vanadium dependent alternative pathway of nitrogen fixation.

1.7 DNA sequencing: The sequencing strategy is shown in Fig.15. Initially the 1 kb *SalI* fragment, which hybridized to the *nifH* probe (dark rectangle, Fig.15) was sequenced by Sanger's dideoxy chain termination method. The *SalI* and *PstI* sites in this fragment were used to clone smaller *SalI-PstI* fragments in M13mp18 and M13mp19 to sequence both the strands. Because of high GC content in the bigger *SalI-PstI* fragment (marked by broken arrows, Fig.15), there were compression of bands in certain regions. Since these compressions could not be resolved by Sanger's method, this region was sequenced by Maxam-Gilbert's chemical cleavage method also. This fragment was subcloned in pUC8 and the *SalI* and *HindIII* site (from vector pUC8) were used to label the 3'-end of DNA with Klenow fragment of DNA polymerase I in order to read the sequence of both the strands. Once the sequence of this 1 kb *SalI* fragment was completed, the three forward frames were checked for an open reading frame (ORF). An ORF which was extending into the next *SalI* fragment was found in one of the reading frames. To complete the sequence, the next *SalI* fragment (having the *EcoRI* site) was cloned in pUC19. This fragment was sequenced directly in the plasmid. The single stranded templates were obtained by denaturing the plasmid by alkali treatment. Using the universal primer and the reverse primer, both the strands were sequenced in the same clone. Once the sequence of this *SalI* fragment was completed, the sequence was aligned with the 1 kb *SalI* fragment based on the orientation of the *EcoRI* site. The ORF of the 1 kb *SalI* fragment continued into this small *SalI* fragment. The *EcoRI* site which was mapped in the restriction map, was also found. About 120 bases after this site, this ORF terminated.
When the sequence was checked further downstream, a second ORF started, after few more bases, which extended into the next SalI fragment. The next SalI fragment was therefore, cloned in pUC19 and the strands were again sequenced as earlier. After 30 bases in this SalI fragment, this ORF also terminated. No other ORF could be found in any frame downstream of this second ORF in this SalI fragment.

1.8 Sequence analysis: The complete nucleotide sequence of 1650 bp is shown in Fig.19. The analysis of the sequence shows that it has two open reading frames. The first open reading frame (ORF1) is 870 bases long and codes for a protein of 290 amino acids. This gene which is homologous to the nifH gene and is its equivalent for the vnf pathway, will be hereafter called vnfH gene. Nine bases upstream of the translation start site (AUG), there is the sequence 5'-GGAGG-3', which corresponds well with the consensus ribosomal binding site (Shine and Dalgarno, 1975). Also present upstream, is the sequence 5'-GTGGACACGCCCTTGCT-3', which corresponds well with the -24 to -12 region of ntrA-dependent promoters, including the nif promoters (Beynon et al. 1983; Dixon 1984). Upstream of this promoter, is an A/T rich region, which is a typical feature of these promoters. The G+C content of the region upstream of the promoter is approximately 10% (The G+C content of vnfH as a whole is 61%). However, no upstream activator sequence (UAS), thought to be involved in NIFA binding, was present (cf. Brigle et al., 1985).

Downstream of vnfH, separated by 116 bp, there is a second ORF (ORF2). This ORF2 is 195 bp long and code for a protein of 65 amino acids. The arrangement of the cysteine residues (Cys-X2Cys-X2Cys-X3Cys-Pro; Arnold et al. 1988) suggests that this could be a ferredoxin-like protein (marked in box, Fig.20B). The presence of a similar sequence downstream of the vnfH gene has also been observed in A. chroococcum.
5' - GTCGACGGAGCGCACAGCATTGCTGGCTATTATGGAGTCCAATAAAACC
TGCAAAAAATTTAAATATTTCACCTAAATTATATGTTTTTTGATTTTATAATCCCCAAAAA
ATAGGCAAATCATCGACTTTATCGATCTCTTGTTGCAACGCCCTTTGCTCAACTCTCGTCGGC
ACAAAATCAAACGCAACGAATCAACCGGAGGTCTTCTAAG ATG GCA TTG CGT CAG 15
TGT GCA ATT TAC GGC AAG GGT GGC ATC GGC AAG TCC ACC ACC ACC 60
CAG AAC CTG GTC GCC GCC CTC GCC GAA GCC GGC AAG AAA GGT ATG 105
ATC GTC GGT TGT GAC CCG AAA GCC GAC TCC ACC CGC CTG ATC CTG 150
CAC TCC AAG GCC CAG GGC ACC GTC ATG GAA ATG GCC GCG TCC GGC 195
GGC TCG GTC GAA GAC CTG GAG CTG GAA GAC GTG CTG CAG ATC GGC 240
TTC GCC GGC GTC AAG TGC GTC GAA TCC GGT GGC CCG GAG CCG GGC 285
GTC GGC TGC GCC GCC CTG GCC GTG ACC ACC CCG ATC AAG TCC GTC 330
GAA GAA GAA GCC GCC TAC AGC GAC GAC CTA GAC TCC GTG TCC TAT 375
GAC GTG CTG GCC GAC GTG GTA TGC GCC GCC TCC GCC ATG CCG ATC 420
CGC GAG AAC AAG GCC CAG GAA ATC TAC ATC GTC TGC GCC GCC GAG 465
ATG ATG GCC ATG TAC GCC GCC AAC ACC ATC GCC AAG GCC ATC GTG 510
AAA TAC GCC CAC TCC GCC AGC GTG CTG GCC GCC CTG ATC TGC 555
AAC AGC CGC AAG ACC GAC CGC GAA GAC GAG CTG ATC ATG GCC CTG 600
GCC GCG AAG ATC GCC ACC CAG ATG ATC CAC TCC GTG CCG CGC GAC 645
AAC GTC GTG CAA CAC GCC GAA ATC CGC CGC ATG ACC GTG ATC GAA 690
TAC GAT CCG AAG GCC GGA CAG GCC GAC GAG TAC CTG GCC CTG GCT 735
CGC AAA ATC GTC GAC AAC AAG CTG CTG GTC GTC ATC ATC CCG AAC CCG GCC 780
TCC ATG GAA GAA CTC GAA GAG CTG ATG GAA TCC GCC ATC ATG 825
GAA GTC GAA GAC GAG TCC GTC GTC GCC GAG GCC GCC GCC GAA GCC 870
contd...
TGATTCACCCAGCACAGCGTTTGCGGAGGAGCGTGCGCCGCGGGCTTTCGGAATGGCTTCTCGCGGCCGGCGGCGACGCCGCCCTCCCTTCGAACAACCGACCTCAGGAGCTGACACCATGCCATG GCC ATG GCC ATC GAC GGC TAC GAA TGC ACT GTC TGC GGC GAC 15
TGC GAG CCG GTC TGC CCG ACC GGT TCG ATC GTC TTC AGG GAC GAT 60
CAC TAC GCG ATC GAA GCC GAC AGT TGC AAC GAA TGC ACC GAC GTG 105
GGC GAG CCG CGC TGT CTC GGC GTC TGC CCC GTC GAC TTG TGC ATC 150
CAG CCG CTC GAT GAC TGAACACTGAACGACTCCGCACCCCGTTGCCGGCGCAGG 165
ACATTCCGCGCCGTCCTGCGCGCGACCCGAACCGCGATCGCTTTTCTCAGGGGCGA
TCGGCGTTTTACTTTTCCCCGCTCCGCTAGCGCCCGCGGACAACAGCCGTCCGTATCCG
CGCCGTTCCGCGTCCTGACCACCGGACGATCGCGGCACAGCTGCGCCAAAGCTACAGCCCG
CTCCACAAGCTGACCATCGGCATCCAGACCTTTCGCAAGATC -3'  

Fig. 19: Sequenced nucleotides of A. vinelandii DNA. Only one strand is shown ORF1 is 870 nucleotides long and is numbered from 'A' of the ATG in line 4 from top. At position -13 to -9 is GGAGG, the potential ribosome binding site. Further upstream is the potential nif promoter sequence GG-N10-GC, underlined. ORF2 is 195 nucleotides long and is numbered from 'A' of the ATG in the line 9 from bottom. In this case, at position -12 to -8 is GGAGGC, the potential ribosome binding site.
It has a 5'GGAGC-3' ribosome binding site, 8 bases upstream of the likely translation start site. No consensus promoter sequence between ORF1 and ORF2 could be found and thus the ORF1 promoter probably controls the transcription of ORF2.

The predicted amino acid sequence of the VNFH protein was compared with the predicted amino acid sequence of NIFH (cf. Brigle et al. 1985) and ANFH (cf. Joerger et al. 1989) proteins (Fig.20A). The VNFH and NIFH have 290 amino acids each, while ANFH has only 275 amino acids. The VNFH protein is more homologous to the NIFH protein (91% identity) than to the ANFH protein (63.5% identity). The homology of the \textit{vnfH} gene at DNA sequence level with the \textit{nifH} gene is 88.5% and with the \textit{anfH} gene is 70.1%. A potential ATP binding site (Gly-\textit{X}_2\textit{Gly-Lys-Ser}, Arnold et al. 1988) at the extreme N-terminus was found to be conserved in all the three proteins (showed in box, Fig.20A).

The hydrophobicity profile of the three proteins (Fig.21) was determined according to Sweet and Eisenberg (1983) using the program of peptide structure from the UWGCG program package. The overall hydrophobicity profile of the three proteins looks to be similar. All the three proteins have a dominant hydrophilic character. The region between residues 104-140 are rather hydrophobic in all the three proteins. The hydrophobicity profile of NIFH and VNFH are very similar, ANFH differs from both these especially towards the N-terminus. While at the N-terminus NIFH and VNFH proteins are hydrophobic, ANFH appears to be more hydrophilic. Rest part of all the three proteins is quite similar.

The secondary structure of the three proteins (Table 9) was predicted according to Chau and Fasman (1974) using the program of peptide structure from the UWGCG program package. The secondary structure of NIFH and VNFH are very similar,
<table>
<thead>
<tr>
<th>NIFH</th>
<th>MAMRQCAIYGGKGGIGKSTTTQNLVAAL 27</th>
</tr>
</thead>
<tbody>
<tr>
<td>VNFH</td>
<td>MALRQCAIYGGKGGIGKSTTTQNLVAAL 27</td>
</tr>
<tr>
<td>ANFH</td>
<td>MTRKVAIYGGKGGIGKSTTTQNTAAAL 26</td>
</tr>
<tr>
<td>NIFH</td>
<td>AEEM.GKMKVMIVGCDCPDKADSTRLLHSK 53</td>
</tr>
<tr>
<td>VNFH</td>
<td>AEAEM.GKMKVMIVGCDCPDKADSTRLLHSK 53</td>
</tr>
<tr>
<td>ANFH</td>
<td>AYFDKVKFTHGCDCPDKADSTRLLIGGK 53</td>
</tr>
<tr>
<td>NIFH</td>
<td>AQTIMEMAAEAAGTVEDLELEDVKLAG 80</td>
</tr>
<tr>
<td>VNFH</td>
<td>AQGTVEMAAASAGSVEDLELEDVLOIG 80</td>
</tr>
<tr>
<td>ANFH</td>
<td>PEETLMMDMVRDKG.AEKITNDDVIKKG 79</td>
</tr>
<tr>
<td>NIFH</td>
<td>YGGVKCVESGGPEPVGCGAGRGVITAI 107</td>
</tr>
<tr>
<td>VNFH</td>
<td>FGGVKCVESGGPEPVGCGAGRGVITAI 107</td>
</tr>
<tr>
<td>ANFH</td>
<td>FLDIQCVCESGGPEPVGCGAGRGVITAI 106</td>
</tr>
<tr>
<td>NIFH</td>
<td>NFLLEEGAYEDDDLDFVYDVLGDVVCGB</td>
</tr>
<tr>
<td>VNFH</td>
<td>NFLLEEGAYSDDDLDFVYDVLGDVVCGB</td>
</tr>
<tr>
<td>ANFH</td>
<td>DLMEENGYTDDDLDFVFFDDLGDVVCGB</td>
</tr>
<tr>
<td>NIFH</td>
<td>GFAMPIRENKAQEIYIVCSEMMAMAYA161</td>
</tr>
<tr>
<td>VNFH</td>
<td>GFAMPIRENKAQEIYIVCSEMMAMAYA161</td>
</tr>
<tr>
<td>ANFH</td>
<td>GFAMPIRDGKAQEVYYIVASGSEMMAMAYA160</td>
</tr>
<tr>
<td>NIFH</td>
<td>ANNISKGIVKYNANSGSVRLGGGLICNSR188</td>
</tr>
<tr>
<td>VNFH</td>
<td>ANNIAKGIVKYAHSGSVRLGGGLICNSR188</td>
</tr>
<tr>
<td>ANFH</td>
<td>ANNICKGLVKYAKQSAVGLGGGLICNSR187</td>
</tr>
<tr>
<td>NIFH</td>
<td>NTDREDELIAALANKLGTQMIMHVFPDR215</td>
</tr>
<tr>
<td>VNFH</td>
<td>KTDREDELIMALAAKITQMIHVFPDR215</td>
</tr>
<tr>
<td>ANFH</td>
<td>KVGDGERSVEEFYTAIGTKMIMHVFPDR214</td>
</tr>
</tbody>
</table>

contd...
### Fig. 20:

(A) Comparison of amino acid sequence of Fe-protein for the vanadium dependent alternative pathway (ANFH) as predicted from the *vnfH* base sequence, with that of the Fe-protein for the conventional pathway (NIFH) and the second alternative pathway (ANFH). The conserved ATP binding site is shown in the box.

(B) Comparison of the amino acid sequence of the ferredoxin-like protein (Av-ORF2) as predicted from the ORF2, with that of the ferredoxin-like protein (Ac-ORF2) from *A. chroococcum*. The arrangement of cysteine residues (shown within the box) suggest these protein to be ferredoxin like proteins.
Fig. 21: Hydrophobicity profile of NIFH, VNFH and ANFH proteins, as predicted from amino acid sequence.
<table>
<thead>
<tr>
<th>α-helix (amino acid residues)</th>
<th>β-sheet (amino acid residues)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIFH</td>
<td>VNFH</td>
</tr>
<tr>
<td>1-8</td>
<td>1-8</td>
</tr>
<tr>
<td>57-79</td>
<td>57-75</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>102-111</td>
<td>102-111</td>
</tr>
<tr>
<td>137-147</td>
<td>137-147</td>
</tr>
<tr>
<td>155-162</td>
<td>155-162</td>
</tr>
<tr>
<td>173-179</td>
<td>173-179</td>
</tr>
<tr>
<td>192-201</td>
<td>192-204</td>
</tr>
<tr>
<td>226-239</td>
<td></td>
</tr>
<tr>
<td>259-290</td>
<td>261-290</td>
</tr>
</tbody>
</table>
however, both these proteins have differences with ANFH. The more important differences are as follows: (a) amino acid residues 1-8 in NIFH and VNFH from α-helix, whereas in ANFH they form β-sheet; (b) amino acid residues 103-110 in NIFH and VNFH form β-sheet, whereas these residues in ANFH from α-helix; (c) amino acid residues from 192-201 in both NIFH and VNFH form α-helix, whereas in ANFH they form β-sheet. (d) the amino acid residues 256-290 in NIFH and VNFH form α-helix, whereas the amino acid residues 267-271 at C-terminus of ANFH form β-sheet. However, all the three proteins are overall essentially helical.

The codon usage in all the three proteins was calculated (Table 10) using the program on codon frequency from the UWGCG program package. The bias of codons used for all amino acids is towards those which have higher G+C ratio. This is consistent with their being the genes of *A. vinelandii*, which has about 65% G+C content. All the three genes show marked codon bias for Asp (GAC), Asn (AAC), Ile (ATC), Tyr (TAC), Phe (TTC), His (CAC) and Pro (CCG), all of which are consistent with efficient expression (Grosjean and Fiers, 1982). This indicates that all the three genes are strongly expressed. Modulatory codons Gly (GGG, GGA), Arg (AGG, AGA, CGG, CGA), Ile (ATA) or leu (CTA) are found in genes that are weakly expressed (Grosjean and Fiers 1982). Only Gly (GGG, GGA) are used by *vnfH* and *anfH* (only 3% in both cases) and no other modulatory codon is used by any of these genes. This again suggests that these genes are highly expressed.

The deduced amino acid sequence from the *vnfH* gene of *A. vinelandii* was compared with that of the deduced amino acid sequence of *vnfH* gene of *A. chroococcum* (cf. Robson *et al.* 1986). These two proteins are highly homologous (91%). The small
Table 10: Codon usage (%) in *A. vinelandii* *nifH*, *vnfH*, *anfH*, *A. vinelandii* ferridoxin-like gene Av-ORF2 and *A. chroococcum* ferridoxin-like gene Ac-ORF2

<table>
<thead>
<tr>
<th>Amino Codon</th>
<th>Codon usage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>nif genes</em></td>
</tr>
<tr>
<td></td>
<td><em>nifH</em></td>
</tr>
<tr>
<td>+Gly</td>
<td>GGG</td>
</tr>
<tr>
<td>+Gly</td>
<td>GGA</td>
</tr>
<tr>
<td>Gly</td>
<td>GGT</td>
</tr>
<tr>
<td>Gly</td>
<td>GGC</td>
</tr>
<tr>
<td>Glu</td>
<td>GAC</td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
</tr>
<tr>
<td>Asp</td>
<td>GAT</td>
</tr>
<tr>
<td><em>Asp</em></td>
<td>GAC</td>
</tr>
<tr>
<td>Val</td>
<td>GTG</td>
</tr>
<tr>
<td>Val</td>
<td>GTA</td>
</tr>
<tr>
<td>Val</td>
<td>GTT</td>
</tr>
<tr>
<td>Val</td>
<td>GTC</td>
</tr>
<tr>
<td>Ala</td>
<td>GCG</td>
</tr>
<tr>
<td>Ala</td>
<td>GCA</td>
</tr>
<tr>
<td>Ala</td>
<td>GCT</td>
</tr>
<tr>
<td>Ala</td>
<td>GCC</td>
</tr>
<tr>
<td>+Arg</td>
<td>AGG</td>
</tr>
<tr>
<td>+Arg</td>
<td>AGA</td>
</tr>
<tr>
<td>+Arg</td>
<td>CGG</td>
</tr>
<tr>
<td>+Arg</td>
<td>CGA</td>
</tr>
<tr>
<td>Arg</td>
<td>CGT</td>
</tr>
<tr>
<td>Arg</td>
<td>CGC</td>
</tr>
<tr>
<td>Ser</td>
<td>AGT</td>
</tr>
<tr>
<td>Ser</td>
<td>AGC</td>
</tr>
<tr>
<td>Ser</td>
<td>TCG</td>
</tr>
<tr>
<td>Ser</td>
<td>TCA</td>
</tr>
<tr>
<td>Ser</td>
<td>TCT</td>
</tr>
<tr>
<td>Ser</td>
<td>TCC</td>
</tr>
<tr>
<td>Lys</td>
<td>AAG</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Lys</td>
<td>AAA</td>
</tr>
<tr>
<td>Asn</td>
<td>AAT</td>
</tr>
<tr>
<td>*Asn</td>
<td>AAC</td>
</tr>
<tr>
<td>Met</td>
<td>ATG</td>
</tr>
<tr>
<td>+Ile</td>
<td>ATA</td>
</tr>
<tr>
<td>Ile</td>
<td>ATT</td>
</tr>
<tr>
<td>*Ile</td>
<td>ATC</td>
</tr>
<tr>
<td>Thr</td>
<td>ACG</td>
</tr>
<tr>
<td>Thr</td>
<td>ACA</td>
</tr>
<tr>
<td>Thr</td>
<td>ACT</td>
</tr>
<tr>
<td>Thr</td>
<td>ACC</td>
</tr>
<tr>
<td>Trp</td>
<td>TGG</td>
</tr>
<tr>
<td>End</td>
<td>TGA</td>
</tr>
<tr>
<td>End</td>
<td>TAG</td>
</tr>
<tr>
<td>End</td>
<td>TAA</td>
</tr>
<tr>
<td>Cys</td>
<td>TGT</td>
</tr>
<tr>
<td>Cys</td>
<td>TGC</td>
</tr>
<tr>
<td>Tyr</td>
<td>TAT</td>
</tr>
<tr>
<td>*Tyr</td>
<td>TAC</td>
</tr>
<tr>
<td>Leu</td>
<td>TTG</td>
</tr>
<tr>
<td>Leu</td>
<td>TAA</td>
</tr>
<tr>
<td>Leu</td>
<td>CTG</td>
</tr>
<tr>
<td>+Leu</td>
<td>CTA</td>
</tr>
<tr>
<td>Leu</td>
<td>CTT</td>
</tr>
<tr>
<td>Leu</td>
<td>CTC</td>
</tr>
<tr>
<td>Phe</td>
<td>TTT</td>
</tr>
<tr>
<td>*Phe</td>
<td>TTC</td>
</tr>
<tr>
<td>Gln</td>
<td>CAG</td>
</tr>
<tr>
<td>Gln</td>
<td>CAA</td>
</tr>
<tr>
<td>His</td>
<td>CAT</td>
</tr>
<tr>
<td>*His</td>
<td>CAC</td>
</tr>
<tr>
<td>*Pro</td>
<td>CCG</td>
</tr>
<tr>
<td>Pro</td>
<td>CCA</td>
</tr>
<tr>
<td>Pro</td>
<td>CCT</td>
</tr>
<tr>
<td>Pro</td>
<td>CCC</td>
</tr>
</tbody>
</table>

*Highly expressed codons.  +Weakly expressed codons (Modulatory codons
ferridoxin like protein of A. vinelandii is also very similar to that of A. chroococcum (Fig.20B), differing in only 10 amino acids.

Codon usage of the A. vinelandii ferridoxin-like gene (Av-ORF2) and the A. chroococcum ferredoxin-like gene (Ac-ORF2) was determined and compared with that of nifH, vnfH and anfH genes (Table 10). Both these genes had similar codon bias as that of nifH, vnfH and anfH genes. This would mean that these two genes are also highly expressed genes.

2. Mutagenesis of Av-ORF2

2.1 Tn5 mutagenesis: The results of the sequence analysis suggest that Av-ORF2 is a ferredoxin-like gene and is controlled by the promoter upstream of the vnfH gene. The next attempt was to get a Tn5 insertion in this ORF and to study the phenotypic characters of the resulting mutant. This would let us find out if this ORF has any role to play in the vanadium dependent alternative pathway of nitrogen fixation. It was therefore attempted to get a clone with Tn5 insertion in the 0.4 kb Sall fragment (7th band from the top, lane h; Fig.22) by screening DNA of more clones of pRR5::Tn5 by restriction digestion as described in Results (section 1.2). This clone was to be used for marker exchange with the wild type A. vinelandii to get a Tn5 mutant as described for the vnfH gene. Plasmid DNA samples from more than 1000 single colonies of pRR5::Tn5 were screened by restriction digestion with Sall to get a Tn5 insertion in the 0.4 kb Sall fragment. Tn5 insertions could be obtained in all the seven other fragments, but never in the 0.4 kb Sall fragment. Tn5 is the most non-specific transposon known and is considered to be the best transposon for generalized mutagenesis (deBrujin and Lupski 1984; Ruvkun and Ausubel 1981). The number of Tn5 insertions in a Sall fragment
Fig. 22: Restriction pattern of some pRR5::Tn5 mutants after digestion with SalI. Lane a,b,c,d,e,f,g,j,k, l,m,n of pRR5::Tn5 mutants; lane h, pRR5; lane i, λDNA digested with HindIII; lane o, φX174 digested with HaeIII.
should have therefore depended only on the size of the fragment. Tn5 insertions on the SalI fragments of pRR5::Tn5 in different clones were therefore analysed to find out the randomness of transposition by the transposon Tn5 (Table 11). The results in Table 11 suggests that the small size of the fragment was not a limiting factor to obtain a Tn5 insertion, as the insertions could be obtained in a smaller fragment (0.35 kb; 8th band from the top, lane n, Fig.22). The results of Table 11 also suggest that the Tn5 prefers certain sequences (e.g. sequence of DNA in the 2nd band where 329 insertions were obtained against expected 176) and avoids others (e.g. sequences of DNA in 3rd and 8th band where only 20 and 13 insertions were obtained against expected 84 and 23 respectively, when normalized with respect to size based on band no.1). In the band no.7, where 31 insertions were expected, none could be obtained. Thus, Tn5 is not strictly a random transposon.

2.2 Site directed mutagenesis: Since the attempts to obtain a Tn5 insertion in Av-ORF2 failed, site directed mutagenesis was done to create a unique BglII site in the cloned Av-ORF2. One base each in the codons for the 18th (second C → A) and 19th amino acids (G → A) were changed. The oligonucleotide of the following sequence was synthesized and used for this purpose:

\[ 3'\text{-}\text{CTGACGCTCGTCTAGACCGGGCTGG} - 5'\]

The 2.5 kb EcoRI-HindIII fragment from pRR5 (Fig.15), containing the Av-ORF2, was cloned in EcoRI-HindIII digested pUC19 (pRR600). The EcoRI-HindIII fragment from pRR600 was eluted from the gel and cloned in the RF form of M13mp19 and mutagenesis was done using the Bio-Rad kit to generate a unique BglII site. This mutagenised EcoRI-HindIII fragment was cloned in EcoRI-HindIII digested pUC19 (pRR601). The generation of the BglII site in the insert of pRR601 was confirmed by
Table 11: Analysis of the Tn5 insertions onto the SaI fragments of the plasmid pRR5, to find out the random transposition by the transposon Tn5

<table>
<thead>
<tr>
<th>Band No. from top</th>
<th>Size (kb)</th>
<th>Tn5 insertion per 1000 mutants</th>
<th>Expected number if normalized with respect to size based on #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.7</td>
<td>437</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.3</td>
<td>329</td>
<td>176</td>
</tr>
<tr>
<td>3</td>
<td>1.1</td>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>1.0</td>
<td>77</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>0.8</td>
<td>67</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
<td>57</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>0.4</td>
<td>-</td>
<td>31</td>
</tr>
<tr>
<td>8</td>
<td>0.35</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1000</td>
<td></td>
</tr>
</tbody>
</table>
restriction digestion. Fig. 23 shows the restriction pattern of pRR600 and pRR601 with different enzymes. The plasmid pRR600 has no BgIII site and hence cannot be cleaved with BgIII. The plasmid pRR601, in which a BgIII site has been created in the insert, becomes linear with BgIII. On double digestion with HindIII-BgIII, the plasmid pRR600 becomes linear and in case of the plasmid pRR601, a fragment of about 2.5 kb comes out.

The mutagenised EcoRI-HindIII fragment from pRR601 was next cloned in EcoRI-HindIII digested pRR5 as shown in Fig.24. The plasmid pRR5 has two EcoRI sites and one HindIII site. The mutagenised EcoRI-HindIII fragment has to be replaced with the EcoRI-HindIII fragment of the insert. For this, first pRR5 was partially digested with EcoRI. The linear 12 kb fragment (in which EcoRI has cut only once) was eluted from the gel. This DNA was then completely digested with HindIII. This would generate 4 fragments. The mixture of two bigger fragments (about 9.5-10 kb) were eluted from the gel. The EcoRI-HindIII fragment from pRR601 was ligated to this mixture, E. coli S17.1 cells were transformed and transformants were selected on L-agar, Cm. The DNA from a few transformants was digested with BgIII. The correct clones gave linear fragments of 12 kb. These clones were also digested with SalI the pattern was compared with that of pRR5. One of them, which had similar restriction pattern as that of pRR5, was designated pRR650 (Fig.24).

An interposon (the BamHI fragment, β-lactamase gene) from the plasmid, pKT253-Amp (Fig. 25a) (Fellay et al. 1987) was ligated to the BgIII digested pRR650. This would create a BamHI-BgIII fusion. The E. coli S17.1 cells were transformed with the ligated mixture and the transformants were selected on L-agar Cm, Amp. The plasmid DNA from a few transformants was isolated and digested with BamHI-HindIII.
Fig. 23: Restriction pattern of plasmids pRR600 and pRR601. Plasmid DNA pRR600 uncut, lane b; digested with BgIII, lane c; digested with BgIII-EcoRI, lane d; digested with HindIII-EcoRI, plasmid DNA pRR601 uncut, lane f; digested with BgIII, lane g; digested with BgIII-EcoRI; lane h, digested with HindIII-EcoRI, lane i; λDNA digested with HindIII lane a.
Fig. 24: Strategy adopted for construction of plasmid pR650
Fig. 25: Restriction map of (a) pKT254- Amp (Amp interposon); (b) pUC-4-KISS (Kan-cartridge); (c) pUC7 (polylinker) of pUC7.

Abbreviations: RI, EcoRI; Sm, SmaI; BHI, BamHI; HIII, HindIII; P, PstI; Bs, BstEII; k, KpnI; Sp, SphI; X, XbaI; Ac, AccI; S, SalI; HIII, HincII; Kan, kanamycin; Amp, ampicillin.
The pattern was compared with pRR650 digested with *BamHI-HindIII*. The size of the insert of the transformants had increased by about 2 kb as compared to the insert in pRR650 due to the cloning of the interposon.

The cloning of the interposon would create a selectable insertional mutation in the Av-ORF2. One of these clones was designated pRR651.

2.3 Marker exchange: The plasmid pRR651 was transferred into *A. vinelandii* CA12 by conjugation and the exconjugants were selected on BN^+F, Amp (50 µg/ml). About 500 exconjugants were replica plated by patching on BN^+F, Cm (200 µg/ml) and BN^+F Amp (50 µg/ml) to select for a colony that is Amp^r^ and Cm^s^ (a two point crossover event resulting in the loss of vector). About 15 Amp^r^, Cm^s^ clones were obtained. When these colonies were again replica plated by patching on BN^+F, Cm and BN^+F Amp, some of them started growing slowly on BN^+F Cm (this would mean the presence of the vector). To confirm the loss of vector from these clones, colony hybridization of these 15 clones was done using the vector pSUP102 as the probe (Fig.26). No signal was observed in some colonies, whereas the intensity of signal varied in other colonies. To confirm that there was no signal in few colonies (*i.e.* colony No.1,6,7,12,13 Fig.26), because of the loss of vector, the genomic DNA from these clones from *A. vinelandii* CA12 and plasmid DNA pRR650 and pRR651, were digested with *SalI*, fractionated on 0.8% agarose gel, Southern blotted and hybridized with the probe made of the *EcoRI-SalI* fragment (Fig.15) containing most of the Av-ORF2. Signals were observed in all the samples at the positions equivalent to the wild type (Lane d, e, Fig.27), as well as the mutagenized copy of the gene (lane c, Fig.27). Thus, all the clones had a single point cross over.
Fig. 26: Colony hybridization of a few Amp<sup>r</sup> exconjugants using pSUP102 as probe.
Hybridization analysis of genomic DNA of colonies from Fig. 26. All samples were digested with SalI. (A) DNA in lane f, g, h, i, j, k, of colonies, 8, 1, 6, 7, 12, 13, respectively; lane e, CA12; lane d, pRR650; lane c, pRR651; lane b, λDNA digested with HindIII, lane a, φX174 digested with HaellI. (B) Autoradiogram of the blot of gel 'A' after the blot was hybridized with 0.4 kb SalI fragment having ferridoxin-like gene.
2.4. Cloning of Kan-resistance gene cartridge in pRR651: Since the attempts to screen for a two point cross over by colony hybridization failed, a kanamycin-resistance gene cartridge from the plasmid pUC4-KISS (Pharmacia cloning vectors) was cloned in the EcoRI site of the chloramphenicol-resistance gene of the plasmid pRR651. This was done because, as discussed earlier, *A. vinelandii* CA12 is resistant to high a dose of Cm (175 μg/ml) and hence, it is difficult to screen for a real Cm^R colony. If the Cm-marker of the vector is replaced by a Kan-marker, the screening would be much easier. *A. vinelandii* CA12 is sensitive to very low concentration of Kan (1 μg/ml). It was for this reason that the Kan resistance gene cartridge was cloned in the Cm-resistance gene of vector. The strategy for cloning is shown in Fig.28. The plasmid pRR651 was partially digested with EcoRI and the 14 kb linear fragment was eluted from the gel. The Kan-resistance gene cartridge was taken out from pUC4-KISS as an EcoRI fragment (1.3 kb). This cartridge was ligated to the eluted 14 kb linear DNA fragment. *E. coli* S17.1 cells were transformed with the ligated mixture and the transformants were selected on L-agar Amp, Kan plates. The plasmid DNA from a few transformants was isolated and digested with EcoRI. Three bands of expected size were observed *i.e.* 7, 5 and 1.3 kb. One of these clones was designated as pRR800.

The plasmid pRR800 was transferred into *A. vinelandii* CA12 by conjugation and the exconjugants were selected on BN^F Amp (50 μg/ml). About 1000 colonies were replica patched on BN^F Kan (1 μg/ml) and BN^F Amp (50 μg/ml). About 200 colonies were found to be Kan^S, thus indicating the loss of the vector. When these Kan^S colonies from the BN^F Amp plate were repatched on BN^F Kan to confirm kanamycin sensitivity, about 150 colonies started growing slowly. When the remaining 50 colonies were repatched from this second subculture to ensure kanamycin sensitivity, they also started growing on BN^F Kan. Thus no Kan^S (result of two point crossover event) colony
Fig. 28: Strategy adopted for construction of plasmid pRR800.
could be obtained by this method also. Since attempts to screen for a two point crossover event in this case also failed, even after changing the vector marker, the only thing which could be suspected as the cause of the difficulty was the interposon Ω-Amp. It was suspected that the β-lactamase gene (responsible to impart resistance against Amp) of Ω-Amp is responsible for a gene dosage effect, which results in exhibiting resistance to Cm or Kan, by clones which were initially sensitive to respective drugs. (see discussion, section 2). In the next step the Amp marker of the interposon was changed to Kan in the plasmid pRR650, and this new construct was used for marker exchange.

### 2.5 Cloning of the Kan-resistance gene cartridge in pRR650:

The Kan-resistance gene cartridge from the plasmid pUC4-KISS (Fig 25b) was to be cloned in the BgIII site of pRR650. Since there are no BgIII or BamHI sites on either side of this cartridge, it was taken out as PstI fragment (about 1.3 kb) and cloned in the PstI site of pUC7 (construct pKUC7). The plasmid pUC7 (Fig. 25c) has BamHI sites on both sides of the PstI site. The Kan-cartridge was taken out from pKUC7 as a BamHI fragment and cloned in BgIII digested pRR650. The *E. coli* S17.1 cells were transformed with the ligated mixture and the transformants were selected on L-agar Cm, Kan. Plasmid DNA from few transformants was isolated and digested with *SalI*. The restriction pattern was compared with pRR5 digested with *SalI*. One band of 1.3 kb size (Kan cartridge) was extra in the plasmid DNA of transformants and a 0.4 kb fragment was missing. Instead, two smaller fragments of 0.25 kb and 0.15 kb fragments were seen which were not present in pRR5. The Kan cartridge has *SalI* sites on both sides, they get cut and it results in disappearance of a 0.4 kb *SalI* fragment (where it was cloned as BamHI fragment) and generates two smaller fragment of 0.25 kb and 0.15 kb. One of the transformants was designated as pRR900.
The plasmid pRR900 was transferred into *A. vinelandii* CA12 by conjugation and the exconjugants were selected on BN^{+}F Kan (1 μg/ml). About 500 colonies were replica patched on BN^{+}F Cm (200 μg/ml) and BN^{+}F Kan (1 μg/ml). Seven colonies were found to be Cm^{s}, Kan^{r}, thus indicating the loss of vector. When these Cm^{s} colonies were repatched on BN^{+}F Cm to confirm Cm sensitivity, they again started growing slowly. On repatching from the second patching, they started growing well on Cm. This time the conjugation of pRR900 was done with UW (wild type *A. vinelandii*) also as a control to rule out the possibility of any problem with the *A. vinelandii* CA12 strain (*nifHDK* deletion derivative of UW). The exconjugants with this strain also behaved the same way as with *A. vinelandii* CA12. Hence, there was no problem with the host strain *A. vinelandii* CA12. Thus this attempt to screen for a two point crossover also failed.

Finally, even though the efforts were made in all possible ways to design a plasmid for doing a marker exchange, none of the three constructs used could yield a two point crossover and hence a mutant. There was no scope of changing the construct any further, so no more experiments could be done with this plasmid. There is no other mobilizable vector which is unstable in *A. vinelandii* which could have been used for doing the marker exchange. The transformation of *A. vinelandii* CA12 strain does not work well in our hands. The results are surprising, but interesting because the plasmid (pRR29) used for marker exchange to get a *vnfH*:Tn5 mutant and the plasmids (pRR651 and pRR900) used to get a Av-ORF2::Tn5 mutant are the same, except that in pRR29 a Tn5 is inserted in the *vnfH* gene and in pRR651 and pRR900, the interposon Ω-Amp and Kan-cartridge respectively are inserted in the Av-ORF2. A two point crossover could be easily obtained in case of pRR29, but no two point crossover could be obtained in case of pRR651, pRR800 or pRR900. With the date in hand, it is difficult to speculate any definite reason as to why this region of DNA (Av-ORF2 and possibly some flanking
region) is behaving so differently. The possibility of Av-ORF2 being an essential gene can be ruled out, since the vnfh::Tn5 mutant could be obtained. The mutation in Av-ORF2 can not be lethal, because vnfh and Av-ORF2 are in the same operon and a Tn5 insertion in vnfh would have shown polar effect on Av-ORF2.

3. Construction of vnfh::lacZ Fusion and the Study of the Expression of the Promoter of the vnfh Gene

To study the regulation of expression of the vnfh promoter, it was fused to the lacZ gene by cloning in the transcriptional fusion vector, pGD499 (Fig.29) (Ditta et al. 1985). The plasmid pGD499 is a derivative of pRK290, in which lacZY operon is fused to a truncated kanamycin resistance gene (including its promoter) from which β-galactosidase expresses constitutively. Replacement of the BamHI-HindIII fragment with a fragment of interest, which has been modified so as to have properly oriented BamHI-HindIII sites, allows the monitoring of transcription originating from within the new fragments and crossing the HindIII site.

3.1 Construction of vnfh::lacZ fusion: The expression of the vnfh gene was studied in different growth conditions and NIF\textsuperscript{-} backgrounds. This would give us an insight into the regulation of the vanadium dependent alternative pathway of nitrogen fixation. For this, the promoter of the vnfh gene was cloned in the lac-fusion vector pGD499. The activity of the promoter was assessed by monitoring the activity of the lacZ gene.

The base sequence of the vnfh gene (Fig.19) shows that the promoter is in a 450 bp SalI-PstI fragment and the direction of the promoter is from SalI to PstI site (Fig.15). For cloning in pGD499 this SalI-PstI fragment was cloned in SalI-PstI digested pUC19
Fig. 29: Restriction map of plasmid pGD499 and pRR50.2
Abbreviations: RI, EcoRI; Bm, BamHI; H, HindIII; S, SalI, Pr, Promoter; Kan, kanamycin; Tet, tetracyclin.
The poly linker of pUC19 has BamHI and HindIII sites on the either side of SalI and PstI sites respectively. Since the insert has no BamHI and HindIII sites, the insert from pRR30 was taken out as a BamHI and HindIII fragment and cloned in the BamHI-HindIII digested pGD499 (pRR30.2). This construct was mobilized into A. vinelandii CA12 strain by conjugation with the E. coli S17.1 strain containing pRR30.2. A. vinelandii CA12 containing pRR30.2 was grown in BNF supplemented with 2 mM urea (derepressing nitrogen source) and 50 nM V2O5. The β-galactosidase activity of these cells was assayed. Not more than background levels of β-galactosidase activity was observed in this construct (Table 12). The only explanation for this could be that some upstream activator sequence (UAS), required for the expression of the vnfH promoter is missing in this SalI-PstI fragment.

A bigger PstI fragment having vnfH promoter and about 2.4 kb of upstream sequence (Fig.15) was cloned in the PstI site of pUC8. The orientation of the promoter with respect to the BamHI and HindIII sites of the polylinker was decided by digesting these clones with SalI. In the correct orientation, where the direction of transcription of the promoter is towards the lacZ gene, a 2.1 kb SalI fragment will come out (one SalI site is in the insert and the other is in the polylinker of the vector). This clone was designated as pRR50. The clone having the fragment and therefore the promoter in the opposite orientation to this, was designated as pRR51. The inserets from these clones were taken out as BamHI-HindIII fragments and cloned in pGD499-digested with BamHI-HindIII. These clones were designated as pRR50.2 and pRR51.2 respectively (Fig.29). These clones were assayed for β-galactosidase activity as described earlier. The clone with pRR50.2 showed good β-galactosidase activity, whereas pRR51.2 failed to show any activity (as expected because the promoter in this clone was in opposite orientation with respect to the lacZ gene) (Table 12). To make sure that the observed activity of pRR50.2
**Table 12:** Expression of different *vnfH::lacZ* fusion plasmids in *A. vinelandii* CA12 strain

<table>
<thead>
<tr>
<th>Strain (plasmid)</th>
<th><em>Growth conditions</em></th>
<th>β-Galactosidase activity (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA12 (pGD499)</td>
<td>BNF (Urea, V)</td>
<td>160</td>
</tr>
<tr>
<td>CA12 (p830.2)</td>
<td>BNF (Urea, V)</td>
<td>170</td>
</tr>
<tr>
<td>CA12 (pRR50.1)</td>
<td>BNF (Urea, V)</td>
<td>165</td>
</tr>
<tr>
<td>CA12 (pRR50.2)</td>
<td>BNF (Urea, V)</td>
<td>8974</td>
</tr>
<tr>
<td>CA12 (pRR45.2)</td>
<td>BNF (Urea, V)</td>
<td>110</td>
</tr>
<tr>
<td>CA12 (pRR46.2)</td>
<td>BNF (Urea, V)</td>
<td>132</td>
</tr>
</tbody>
</table>

*BNF (Urea, V): Burk's nitrogen free medium supplemented with 2 mM Urea (derepressing nitrogen source), 50 nM V$_2$O$_5$ and Tet (5 μg/ml).
was because of the vnfH promoter and not because of any other promoter, the 2.1 kb SalI-PstI fragment (upstream of the vnfH promoter) was separately cloned in pGD499 as described earlier and the clones with correct and opposite orientations were designated as pRR45.2 and pRR46.2 respectively. The β-galactosidase activity of both the constructs was assayed. None of these clones could show more than background levels of β-galactosidase activity. It was thus clear that the insert of pRR50.2 had only the vnfH promoter and the β-galactosidase activity shown by it was because of the activity of the vnfH promoter. All further studies regarding the expression of the vnfH promoter were carried out with this construct.

3.2 Regulation by ammonia: It is known that the A. vinelandii does not fix nitrogen when grown in the presence of fixed nitrogen source (e.g. ammonium acetate). It is also known that the expression of the genes for the conventional pathway is repressed in the presence of ammonium acetate.

In the present study, the effect of fixed nitrogen source on the expression of pRR50.2 has been studied. The expression of pAM60 (lac-fusion of the promoter of the nifH gene, for Mo dependant conventional pathway) was also studied under similar conditions and compared with pRR50.2. Table 13 shows the effect of fixed nitrogen source on the expression of vnfH and nifH promoters. It is clear from these results that the expression of the vnfH gene is repressed poorly by fixed nitrogen source. In contrast, the expression of the nifH gene is repressed strongly by the fixed nitrogen source. Even at double concentrations of ammonium acetate (24 mM) as compared to the normal growth conditions (12 mM), the repression in case of vnfH promoter was only 7 fold, whereas in case of the nifH promoter, it was 21 fold.
**Table 13:** Effect of fixed nitrogen source on the expression of *vnfH* and *nifH* promoters

<table>
<thead>
<tr>
<th>Strain</th>
<th><em>BNF</em> supplemented with</th>
<th>β-Galactosidase activity (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA12 (pGD499)</td>
<td>Urea (2 mM)</td>
<td>155</td>
</tr>
<tr>
<td>CA12 (pGD499)</td>
<td>Ammonium acetate (12 mM)</td>
<td>160</td>
</tr>
<tr>
<td>CA12 (pRR50.2)</td>
<td>Urea (2 mM)</td>
<td>7652</td>
</tr>
<tr>
<td>CA12 (pRR50.2)</td>
<td>Ammonium acetate (12 mM)</td>
<td>1562</td>
</tr>
<tr>
<td>CA12 (pRR50.2)</td>
<td>&quot;</td>
<td>(24 mM)</td>
</tr>
<tr>
<td>CA12 (pAM60)</td>
<td>Urea (2 mM)</td>
<td></td>
</tr>
<tr>
<td>CA12 (pAM60)</td>
<td>Ammonium acetate (12 mM)</td>
<td></td>
</tr>
<tr>
<td>CA12 (pAM60)</td>
<td>&quot;</td>
<td>(24 mM)</td>
</tr>
</tbody>
</table>

*BNF* is Mo-free Burk's nitrogen free medium. In case of *A. vinelandii* (CA12 (pRR50.2), it was supplemented with 100 nM V$_2$O$_5$ and in case of *A. vinelandii* CA12 (pAM60), it was supplemented with 1 μM sodium molybdate. In all cases Tet (5 μg/ml) was added to medium.
3.3 Regulation by metals: Three metals *i.e.* molybdenum (Mo), vanadium (V) and tungsten (W) are known to play role in nitrogen fixation in *A. vinelandii*. In the present study the role of these metals in the regulation of expression of the promoter of the *vnfH* gene has been studied and compared to that of the promoter of the *nifH* gene.

*A. vinelandii* CA12 (pRR50.2) and *A. vinelandii* CA12 (pAM60) were grown in molybdenum free BNF, supplemented with 2 mM urea, 5 μg/ml Tet and different concentrations of sodium molybdate, vanadium pentaoxide and sodium tungstate. The β-galactosidase activity was determined in each case.

Table 14 shows the effect of increasing concentration of these metals on the expression of the *vnfH* and *nifH* promoters. The results can be summarized as follows:

(a) The expression of *vnfH* gene is independent of V. The addition of V does not bring about much change in the level of expression. At higher concentration of V₂O₅ (100 nM) it does enhance the transcription, but that is just two fold. This is in contrast to the *nifH* gene, where the expression is Mo dependent. Concentration of Na₂MoO₄, as low as 12.5 nM, is enough to enhance the expression of the *nifH* gene several fold. The optimal concentration of Na₂MoO₄ appears to be 1 μM. Whatever β-galactosidase activity was observed with *A. vinelandii* CA12 (pAM60) in the absence of added metal, is suspected to be due to the traces of contaminating Mo present in the growth medium. It is extremely difficult to make the medium completely metal free.

(b) Expression of the *vnfH* gene is strongly repressed by Mo. Concentration of Na₂MoO₄ as low as 12.5 nM is enough to repress the expression of the *vnfH* by about ten fold. Maximum repression is achieved at about 1 μM of Na₂MoO₄. In
Table 14: Effect of Mo, V and W on the expression of the \textit{vnfH} and \textit{nifH} promoters. \textit{A. vinelandii} CA12 (pRR50.2) and \textit{A. vinelandii} CA12 (pAM60) were grown in Mo-deficient medium supplemented with 2 mM urea and 5 \( \mu \)g/ml Tet. \( \beta \)-Galactosidase activity was assayed at different concentrations of Mo, V and W in the growth medium.

\begin{center}
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Conc. of Metal} & \textbf{Strain} & \textbf{\( \text{VO}_5 \)} & \textbf{\( \text{Na}_2\text{MoO}_4 \)} & \textbf{\( \text{Na}_2\text{WO}_4 \)} \\
\hline
0 & CA12 (pRR50.2) & 3762 & - & - \\
 & (No metal) & & & \\
 & CA12 (pAM60) & 1086 & - & - \\
 & (No metal) & & & \\
12.5 nM & CA12 (pRR50.2) & 4375 & 462 & 2963 \\
 & CA12 (pAM60) & 1572 & 4269 & 1627 \\
25 nM & CA12 (pRR50.2) & 5293 & 392 & 1963 \\
 & CA12 (pAM60) & 1740 & 5012 & 2527 \\
50 nM & CA12 (pRR50.2) & 6723 & 362 & 392 \\
 & CA12 (pAM60) & 1770 & 5920 & 3752 \\
100 nM & CA12 (pRR50.2) & 7396 & 258 & 269 \\
 & CA12 (pAM60) & 1544 & 5960 & 4962 \\
1 mM & CA12 (pRR50.2) & 6925 & 248 & 242 \\
 & CA12 (pAM60) & 1924 & 6750 & 5692 \\
100 mM & CA12 (pRR50.2) & 7534 & 256 & 292 \\
 & CA12 (pAM60) & 1867 & 6570 & 5429 \\
\hline
\end{tabular}
\end{center}
contrast to this, V has no effect on the expression of the nifH gene, it neither enhance nor represses the expression of the nifH gene. Whatever little effect of higher concentration of V$_2$O$_5$ is seen on the expression of the nifH gene, can be attributed to the contaminating Mo present in V$_2$O$_5$, which may be responsible for the little (approximately 2 fold) enhanced expression of this gene.

(c) W shows similar effect on the expression of the vnfH and the nifH genes as that exhibited by Mo, but the effect is less pronounced at lower concentrations. At higher concentration, W strongly represses the vnfH gene and enhance the expression of the nifH gene.

3.4 Expression of the vnfH gene in some nif mutants: The expression of the vnfH gene was studied in some nif backgrounds to see if these genes have any role to play in the regulation of the vnf pathway of nitrogen fixation.

The plasmid pRR50.2 was conjugated into the nif mutants listed in Table 15. These strains, carrying pRR50.2, were grown in Mo-free BNF, supplemented with 2 mM urea, 5 µg/ml Tet and 100 nM V$_2$O$_5$. The β-galactosidase activity was assayed as described earlier.

Table 15 shows the effect of these nif mutations on the expression of the vnfH gene. The results can be summarized as follows:

a. NTRA is essential for the expression of the vnfH gene because pRR50.2 gave background levels of β-galactosidase activity in A. vinelandii MV700 (ntrA-).
Table 15: Effect of some *nif* mutations on the regulation of *vnfH* gene expression. All the strains were grown in Mo-free BNF media supplemented with 2 mM urea, 100 nM V$_2$O$_5$. In CA12 (pRR50.2) 5 µg/ml Tet was added and in rest all strain Kan (1 µg/ml) and Tet (5 µg/ml) was added in growth medium.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristics</th>
<th>β-Galactosidase activity of strain (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA12 (pRR50.2)</td>
<td>Δ<em>nifHDK</em></td>
<td>6529</td>
</tr>
<tr>
<td>MV700 (pRR50.2)</td>
<td><em>ntrA::Tn5</em></td>
<td>263</td>
</tr>
<tr>
<td>MV551 (pRR50.2)</td>
<td><em>ntrC::Tn5</em></td>
<td>5923</td>
</tr>
<tr>
<td>CA11.46 (pRR50.2)</td>
<td>Δ<em>nifHDK, vnfA::Tn5</em></td>
<td>896</td>
</tr>
<tr>
<td>RR29 (pRR50.2)</td>
<td>Δ<em>nifHDK, vnfH::Tn5</em></td>
<td>203</td>
</tr>
</tbody>
</table>
b. NTRC is not required for the expression of the \textit{vnfH} gene, because the β-galactosidase activity of pRR50.2 was not affected in \textit{A. vinelandii} MV551 (\textit{ntrC}^+) \\

c. VNFA is essential for the expression of the \textit{vnfH} gene, because pRR50.2 gave reduced levels of β-galactosidase activity in \textit{A. vinelandii} CA11.46 (\textit{ΔnifHDK}, \textit{vnfA}^-).

d. VNFH is essential for the expression of its own promoter, because pRR50.2 gave background levels of β-galactosidase activity in \textit{A. vinelandii} RR29 (\textit{Δnif HDK}, \textit{vnfH}^-).

4. Localization of the UAS

Transcription of all the nitrogen fixation genes is activated by a positive regulatory protein. This positive regulatory protein binds to the upstream sequence of the \textit{nif} genes and activates the transcription from the promoter. For the conventional molybdenum dependent (\textit{nif}) pathway, the positive regulatory protein is NIFA. Consensus NIFA binding site or upstream activator sequence (UAS) for all the \textit{nif} operons is TGT-N_{10}-ACA. No TGT-N_{10}-ACA sequence could be found upstream of the \textit{vnfH} promoter (Fig.19). Since the \textit{nif} and \textit{vnf} pathways operate under different growth conditions, it is expected that genes for the two pathways will have two different positive regulatory proteins and thus different UAS for their binding. This speculation has the support from the observation mentioned earlier that no consensus UAS required for the binding of NIFA could be found upstream of the \textit{vnfH} operon. The UAS for the genes of \textit{vnf} pathway is not known.
The aim of the present work is to localize the position of the UAS for the vnfH gene. The vnfH::lacZ fusion construct (Results, section 3.1), pRR30.2 (containing the vnfH promoter and 130 bases upstream) did not show any β-galactosidase activity. This would mean that the UAS required for the activation of the vnfH promoter is either not present in these 130 bases or only a part of it is within these 130 bases and rest is further upstream of it.

The standard technique to find out the protein binding site on the DNA is DNaseI footprinting. For this experiment purified or at least partially purified protein is required. Since nothing is known about the putative VNFA protein, this experiment could not be done. The other possible strategy, which was adopted in the present work, involves creating a set of deletions in the upstream sequence, assaying of the β-galactosidase activity of the lac-fusion clones carrying these deleted fragments and looking for a smallest fragment which retains the activity. Sequencing of this small fragment would localize a small region in the upstream sequence, which contains the UAS. Careful analysis of this region can give an idea of the possible UAS. The final confirmation of the UAS can be done by changing the bases in the suspected UAS by site directed mutagenesis and assaying the β-galactosidase activity of the lac-fusions carrying these mutagenised fragments.

The plasmid pRR50, having the vnfH promoter and about 2.4 kb upstream sequence cloned in pUC8, was used to create a set of deletions by the exonuclease BAL-31. Following are the steps involved in each cycle of creating the deletions, cloning the deleted fragments in pGD499 and mobilizing the pGD499 derivatives in A. vinelandii for β-galactosidase assay (Fig.30):
Fig. 30: Steps involved in one complete cycle of BAL-31 deletion, cloning in plasmid pGD499 and β-galactosidase assay in A. vinelandii.
a. The plasmid pRR50 (purified by passing through CsCl gradient twice to get rid of RNA completely, because RNA inhibits the BAL-31 digestion reaction) was linearized with the restriction enzyme BamHI, phenol extracted, precipitated and dissolved in sterile distilled water to a final concentration of 1 µg/µl.

b. The linear plasmid was treated with the exonuclease BAL-31 for different time intervals to generate several sets of fragments having deletions to different extent (Fig.31). All the DNA samples in these sets were phenol extracted separately, precipitated and dissolved in water. Since the enzyme BAL-31 would digest the linear DNA from both the ends and thus deleting the portions of the vector also, these linear fragments cannot be recircularized. Also insert has to be finally cloned into BamHI-HindIII digested pGD499, so the insert from these deleted fragments was cloned against in fresh pUC19.

d. Since the enzyme BAL-31 does not digest both the strands of DNA with equal efficiency, many of the ends remain staggered. These staggered ends have to be repaired before doing a blunt end ligation. The DNA was treated with all the four dNTPs and the Klenow fragment. The repaired DNA samples were extracted with phenol, precipitated and dissolved in 20 µl of water.

e. The repaired DNA samples were next digested with HindIII to separate the deleted insert from the vector. These digested samples were fractionated on a low melting agarose gel and the inserts were eluted from the gel.

f. The eluted fragments were ligated to the HindIII-HincII (HincII generates a blunt end) digested pUC19. E. coli DH5α was transformed with the ligated mixture and plated on L-agar, Cm, X-gal medium. The plasmid DNA was isolated from all the
Fig. 31: Gel photograph showing successive digestion of linearized pRR50 with BAL-31 exonuclease for increasing amount of time (left to right). Lane a, linearized pRR50; lane b to i and k to s treated with BAL-31; lane j, λDNA digested with HindIII.
white colonies, digested with *BamHI-HindIII*, electrophoresed on an agarose gel
to determine the size of the insert in each clone.

g. Inserts of different sizes taken out as *BamHI-HindIII* fragment from different
clones, were ligated into *BamHI-HindIII* digested pGD499. *E. coli* S17.1 cells
were transformed with the ligated mixture. Analysis of a few clones was done and
one clone from each set was used for conjugation with *A. vinelandii* CA12.

h. Exconjugants were selected on BN^+F^−Tet (5 µg/ml).

i. The extracts of exconjugants were assayed for β-galactosidase activity.

Fig.30 shows the clone with the smallest insert (pRR60.2) obtained in the first
cycle of deletion experiment, which had β-galactosidase activity equal to the control
(pRR50.2). In one cycle of BAL-31 deletion not more than 1 kb from each end (total 2
kb) can be deleted, because as the BAL-31 digestion proceeds after 12-15 minutes,
uniformity of deletion is lost resulting in a smear (lane p-s, Fig.31) which becomes very
difficult to handle for further manipulations. The BAL-31 deletion experiment was
repeated with the plasmid pRR60 in a second cycle of deletion to get an insert of smaller
size as described above. Four such cycles were performed to get inserts of different sizes.
The smallest insert with full β-galactosidase activity had 210 bases upstream of the
promoter (pRR72.2). The biggest insert in which β-galactosidase activity was lost had
142 bases upstream of the promiter (pRR76.2). The results clearly suggest that if not
complete, at least a part of the UAS lies within this region of 68 bases (Table 16).

In a second set of experiments the plasmid pRR70 (Fig.32) the insert of which is
in plasmid pRR70.2), was linearized with *SalI* (*SalI* site of the vector pUC19 is lost
during cloning of insert in *HincII-HindIII* sites of pUC19), filled up with the Klenow
Table 16: β-Galactosidase activity of different BAL-31 deletion and other clones

<table>
<thead>
<tr>
<th>Upstream sequence</th>
<th>SalI site</th>
<th>yn/H promoter</th>
<th>Construct</th>
<th>β-Galactosidase activity (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-24 -12</td>
<td>pRR50.2</td>
<td>7962</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--About 2.4 kb--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR60.2</td>
<td>7276</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--About 1.4 kb--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR69.2</td>
<td>7590</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--About 0.5 kb--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR70.2</td>
<td>6950</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--140 bases--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR72.2</td>
<td>7232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--70 bases--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR76.2</td>
<td>476</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--ACGTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR30.2</td>
<td>390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--GTCGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>pRR75.2</td>
<td>569</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SalI site disturbed</td>
<td>pRR333.2</td>
<td>365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'--CCGGCTGTACCTGCGGTTGAC--134 bases--GGN$_{10}$GC--3'</td>
<td>569</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 32: Nucleotide sequence of the insert of pRR70. The vnfH promoter is shown in the box. The first G and the last G in the box from 5'-end represent -24 and -12 sequences respectively. The sequence underlined has the UAS or at least a part of UAS. ↓ indicates the 5'-end of the insert of pRR72.2.
fragment and ligated. The insert was taken out as a \textit{BamHI-HindIII} fragment and cloned in the \textit{BamHI-HindIII} digested pGD499 (pRR75.2). This construct failed to show \(\beta\)-galactosidase activity more than the background levels (Table 16). The difference between pRR70.2 and pRR75.2 is only at the \textit{SalI} site. In pRR75.2 four extra bases have been added, during filling up of the \textit{SalI} site with the Klenow fragment. The failure of pRR75.2 to show \(\beta\)-galactosidase activity can be explained by one of the following reasons:

a. The \textit{SalI} site which has been disturbed may have a part of UAS.

b. The most persuasive general model for modulation of transcription by a protein bound at a remote site, involves looping of the intervening DNA to bring the regulatory protein close to the promoter proper (Ptashine 1986). Consistent with such a model, it has been shown in \textit{K. pneumoniae} that introducing half a turn of the DNA helix between the \textit{nifH} promoter and the UAS in an upstream position, reduces activation by NIFA, but a near integral number of turns can be inserted with little effect. This suggests that activation of the \textit{nifH} promoter by the upstream bound NIFA is face-of-the-helix dependent, indicating a requirement for stereospecific positioning of the activator with respect to the -24, -12 region (Buck \textit{et al.} 1987a). If it is considered that the UAS of the \textit{vnfH} gene is further upstream of the \textit{SalI} site, then the loss of \(\beta\)-galactosidase activity in pRR75.2 can be explained to be due to the change in the stereospecific positioning of the activator with respect to the promoter. In pRR75.2 there are 4 extra bases as compared to pRR70.2 (which shows full \(\beta\)-galactosidase activity) and these extra bases are in between the putative UAS and the promoter. Each turn of DNA helix has 10 bases. Addition of 4 bases means addition of about half a turn, thus
changing the stereospecific positioning of the UAS with respect to the promoter, resulting in the loss of β-galactosidase activity.

To find out due to which of the above two given reasons, pRR75.2 showed loss of β-galactosidase activity, the sequence of the insert in pRR70 was searched for a convenient restriction enzyme site upstream of SalI. Fig.32 shows 4 HpaII sites in pRR70. A HpaII fragment of 320 bp has 20 bases upstream of the SalI site. Ebright (1986) proposed a consensus symmetrical motif TNTNAN(N)_n-NTNANA (n=0-6; the Ebright box) postulated to be involved in the helix-turn-helix binding of transcription regulators to DNA. In case of the nif genes, the consensus is TGT-N^10-ACA. If the consensus proposed by Ebright is true for vnfH gene also, and SalI site in pRR70 is involved in the binding, then a maximum of 16 bases upstream of SalI are required for the full activity of β-galactosidase. The 320 bp HpaII fragment in pRR70 has 20 bases upstream of the SalI site. Thus if the bases in this SalI site are involved in UAS then this 320 bp HpaII fragment cloned in pGD499, in correct orientation should show full β-galactosidase activity. To check this the insert of pRR70 was eluted from the gel. This eluted fragment was digested with HpaII, 320 bp fragment eluted from the gel and cloned in the AccI site of pUC19 (HpaII site is compatible with AccI but once ligated it can not be cut with either enzymes). The orientation of the promoter with respect to BamHI and HindIII sites was checked by digesting the clones with SalI and PstI (SalI site of pUC19 is lost on cloning in AccI). The clone with correct orientation (pRR333) was digested with BamHI-HindIII, the fragment was eluted and cloned in BamHI-HindIII digested pGD499 (pRR333.2). This plasmid was mobilized into A. vinelandii CA12 and β-galactosidase activity was assayed. This clone failed to give β-galactosidase activity more than the background levels (Table 16).
These results suggest that the bases in the SalI site of pRR70 are not a part of UAS and the loss of β-galactosidase activity in pRR75.2 was due to the addition of 4 bases resulting in the change in the stereospecific positioning of the activator with respect to the promoter. The results presented in the Table 16 strongly suggests that if not complete, at least a part of the UAS lies in the 50 bases within the HpaII site and 5'-end of the insert of pRR72.2 (underlined, Fig.32). Thus the UAS of vnfH gene has been localized within 50 bases upstream of SalI site. The exact UAS can now be found out by doing more precise deletion and site directed mutagenesis within these 50 bases.

5. Isolation of a Regulatory Mutant for the Alternative Pathway

It is now clear that Mo represses the vanadium dependent alternative pathway of nitrogen fixation and the repression is at the transcription level. Nothing is known about the mechanism of this repression. No gene or protein involved in this repression has been identified yet. A mutant of A. vinelandii CA12 which can fix nitrogen even in the presence of Mo will be very useful to identify the gene(s) involved in the repression by Mo. It is for this purpose, attempt was made to isolate a spontaneous mutant in which the alternative pathway of nitrogen fixation is not repressed by Mo.

Ten ml of freshly grown A. vinelandii CA12 (ΔnifHDK) culture (10^8 cells/ml) was centrifuged, resuspended in 1 ml of BNF and 100 μl each was plated on BNF agar containing 1 μM sodium molybdate. The plates were incubated at 30°C for 3 days. One colony appeared on one of the 5 plates.

The colony was streaked on a BNF plate containing 1 μM Na₂MoO₄ along with A. vinelandii CA12 (as a control) and incubated at 30°C. A. vinelandii CA12 failed to grow on this plate, while the mutant showed good growth. This mutant was designated as
CARR (Fig. 33). Both *A. vinelandii* CA12 and *A. vinelandii* CARR were inoculated in BNF supplemented with 12 mM ammonium acetate and shaken at 30°C for 24 hours. The cells were centrifuged, washed with BNF and resuspended in BNF supplemented with 2 mM urea. These cultures were checked for their ability to reduce acetylene in presence of 100 nM V<sub>2</sub>O<sub>5</sub> and 1 μM Na<sub>2</sub>MoO<sub>4</sub> (Fig. 34). The results in Fig. 34 show that *A. vinelandii* CA12 could not reduce acetylene in the presence of Mo, but *A. vinelandii* CARR reduced acetylene in the presence of Mo at the same rate as in the presence of V.

To check if this mutant is a regulatory one, the *vnfH::lacZ* fusion (pRR50.2) was mobilized into this strain by conjugation and the β-galactosidase activity was assayed at different concentrations of Mo, V and W (Table 17). The results in Table 17 show that neither Mo nor W, at concentrations as high as 1 mM, could inhibit the activity of the *vnfH* promoter in *A. vinelandii* CARR as they do in *A. vinelandii* CA12. These results suggests that this is a regulatory mutant in which a gene, which is involved in the repression of the *vnf* pathway of nitrogen fixation, has got mutated and thus this mutant does not respond to Mo.

To check if this mutation has any effect on the regulation of the transcription of the genes of the conventional pathway, *nifH::lacZ* fusion (pAM60) was mobilized into the *A. vinelandii* CARR strain by conjugation and β-galactosidase activity was assayed in presence of both Mo and V (Table 18). The results in the table show that pAM60 failed to express in *A. vinelandii* CARR while it could express at wild type rates in *A. vinelandii* CA12.
Fig. 33: Growth of CA12 and CARR on a BNF agar plate supplemented with 1 μM sodium molybdate.
Fig. 34: Acetylene reduction assay of CA12 and CARR. Cells were grown in BNF supplemented with 2 mM urea, and either 100 nM V or 1 μM Mo. 'X' indicates point of injection of gas sample.
Table 17: β-Galactosidase activity of *A. vinelandii* CARR (pRR50.2) at different concentrations of Mo, V, W. The cultures were grown in BNF supplemented with 2 mM urea, 5 μg/ml Tet and the following metal contents.

<table>
<thead>
<tr>
<th>Conc. of Metal</th>
<th>Strain</th>
<th>( \text{VO}_2^+ )</th>
<th>( \text{NaMoO}_4 )</th>
<th>( \text{Na}_2\text{WO}_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>CARR (pRR50.2)</td>
<td>6950</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>50 nM</td>
<td>CARR (pRR50.2)</td>
<td>6987</td>
<td>6394</td>
<td>6629</td>
</tr>
<tr>
<td>100 nM</td>
<td>CA12 (pRR50.2) (control)</td>
<td>3695</td>
<td>214</td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>CARR (pRR50.2)</td>
<td>6869</td>
<td>6793</td>
<td>6768</td>
</tr>
<tr>
<td>1 μM</td>
<td>CARR (pRR50.2)</td>
<td>7039</td>
<td>6529</td>
<td>7323</td>
</tr>
<tr>
<td>100 μM</td>
<td>CARR (pRR50.2)</td>
<td>6263</td>
<td>6923</td>
<td>6213</td>
</tr>
<tr>
<td>1 mM</td>
<td>CARR (pRR50.2)</td>
<td>7219</td>
<td>6329</td>
<td>6639</td>
</tr>
</tbody>
</table>
Table 18: β-Galactosidase activity of *A. vinelandii* CA12 (pAM60) and *A. vinelandii* CARR (pAM60). The strains were grown in BNF supplemented with 2 mM urea, 5 μg/ml Tet and the following metal contents

<table>
<thead>
<tr>
<th>Strain</th>
<th>Metal conditions</th>
<th>β-Galactosidase activity (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA12 (pAM60)</td>
<td>1 μM NaMoO₄</td>
<td>6320</td>
</tr>
<tr>
<td>CA12 (pAM60)</td>
<td>100 nM V₂O₅</td>
<td>1270</td>
</tr>
<tr>
<td>CARR (pAM60)</td>
<td>1 μM NaMoO₄</td>
<td>360</td>
</tr>
<tr>
<td>CARR (pAM60)</td>
<td>100 mM V₂O₅</td>
<td>210</td>
</tr>
</tbody>
</table>