JOURNAL OF BACTERIOLOGY, June 1990, p. 3400–3408
0021-9193/90/063400-09$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 6

Nucleotide Sequences and Mutational Analysis of the Structural Genes for Nitrogenase 2 of *Azotobacter vinelandii*†

ROLF D. JOERGER,[1] TELISA M. LOVELESS,[1] RICHARD N. PAU,[2] LESLEY A. MITCHENALL,[2] BENJAMIN H. SIMON,[1] AND PAUL E. BISHOP[1]*

Agricultural Research Service, U.S. Department of Agriculture, and Department of Microbiology, North Carolina State University, Raleigh, North Carolina 27695-7615,[1] and Agricultural and Food Research Council, Institute of Plant Science Research, Nitrogen Fixation Laboratory, University of Sussex, Brighton BN1 9RQ, United Kingdom[2]

Received 21 December 1989/Accepted 23 March 1990

The nucleotide sequence (6,559 base pairs) of the genomic region containing the structural genes for nitrogenase 2 (V nitrogenase) from *Azotobacter vinelandii* was determined. The open reading frames present in this region are organized into two transcriptional units. One contains *vnfH* (encoding dinitrogenase reductase 2) and a ferredoxinlike open reading frame (Fd). The second one includes *vnfD* (encoding the α subunit of dinitrogenase 2), *vnfG* (encoding a product similar to the δ subunit of dinitrogenase 2 from *A. chroococcum*), and *vnfK* (encoding the β subunit of dinitrogenase 2). The 5′-flanking regions of *vnfH* and *vnfD* contain sequences similar to *ntrA*-dependent promoters. This gene arrangement allows independent expression of *vnfH*-Fd and *vnfDGK*. Mutant strains (CA80 and CA11.80) carrying an insertion in *vnfH* are still able to synthesize the α and β subunits of dinitrogenase 2 when grown in N-free, Mo-deficient, V-containing medium. A strain (RP1.11) carrying a deletion-plus-insertion mutation in the *vnfDGK* region produced only dinitrogenase reductase 2.

Azotobacter vinelandii can grow diazotrophically, using any of three distinct nitrogenases depending on the presence or absence of Mo or V in the growth medium. Nitrogenase 1 is synthesized in cells cultured under Mo sufficiency, nitrogenase 2 is expressed in cells grown in N-free media in which Mo has been replaced by V, and nitrogenase 3 is present in cells grown diazotrophically under Mo- and V-deficient conditions (14, 22). Each of the three nitrogenases is composed of two components; dinitrogenase reductase, a dimer of two identical subunits; and dinitrogenase, composed of two pairs of nonidentical subunits (α and β) (9, 10, 12, 17). The components of nitrogenase 2 from *A. vinelandii* (originally designated Av1′ and Av2′) have been purified (16, 17). Dinitrogenase reductase 2 (Av2′) was found to have an apparent M_r of approximately 31,000 (16). The M_r of dinitrogenase 2 (Av1′) was determined to be about 200,000 (17). This component was found to be a tetramer ($\alpha_2\beta_2$) with subunits of apparent M_rs of 52,000 and 55,000 (17). Dinitrogenase 2 contains V and Fe in a 1:13 ± 3 ratio (17). When dinitrogenase 2 (V nitrogenase) was isolated from the related organism *A. chroococcum*, it was observed that, in addition to the α and β subunits, two δ subunits were present (36). This subunit is encoded by a gene, *vnfG*, located between *vnfD*, encoding the α subunit of dinitrogenase 2, and *vnfK*, encoding the β subunit of dinitrogenase 2 (36). The *vnfG* gene is similar to a gene, *anfG*, found within the structural gene cluster for nitrogenase 3 (a nitrogenase complex that does not appear to contain either Mo or V) from *A. vinelandii* (23). The nucleotide sequence of the structural genes for nitrogenase 1 (*nifHDK*) and nitrogenase 3 (*anfHDGK*) from *A. vinelandii* have been determined (7, 23). The nucleotide sequence of the gene (originally designated *nifH2*) encoding dinitrogenase reductase 2 and of an open reading frame (ORF) (designated ORF2) encoding a ferredoxinlike protein has also been determined (34). In this study, we present the complete nucleotide sequence of the structural genes for nitrogenase 2. Strains of *A. vinelandii* that carry mutations in the structural genes for nitrogenase 2 are also described.

MATERIALS AND METHODS

Plasmids and bacterial strains. The *A. vinelandii* and *Escherichia coli* strains and plasmids used in this study are listed in Table 1. The *A. vinelandii* strains were grown at 30°C in modified Burk medium (39). For growth studies under Mo-deficient conditions and in the presence of V, precautions were taken to minimize contamination with metals as described previously (2). Na_2MoO_4 and V_2O_5 were each added to a final concentration of 1 μM. Fixed N was added as ammonium acetate (2.2-mg/ml final concentration). When required, kanamycin and spectinomycin were added to final concentrations of 10 and 20 μg/ml, respectively. The concentration of ampicillin was 50 μg/ml. *E. coli* HB101 was grown and maintained in TYE or LB medium. *E. coli* K-12 71-18 was maintained on M9 minimal medium (26). When necessary, kanamycin and ampicillin were added to final concentrations of 100 and 40 μg/ml of medium, respectively. Spectinomycin was added to a concentration of 5 μg/ml.

DNA manipulations. DNA isolation procedures, Southern hybridizations, ligations, and transformations of *E. coli* were carried out essentially as described by Maniatis et al. (25). *A. vinelandii* cells were made competent and transformed with DNA as described by Page and von Tigerstrom (29).

DNA sequencing and sequence analysis. Most of the sequence of the region containing the structural genes for nitrogenase 2 was determined by using a previously de-

* Corresponding author.
† Paper no. 12531 of the journal series of the North Carolina Agricultural Research Service.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Reference or source
E. coli		
HB101	F^- *hsd-20* (r_B^- m_B^-) *proA2 lacYI rpsL20 supE44*	25
K-12 71-18	Δ(*lac-pro*) F′ *lac*^{q} *lacZ* M15 *pro*$^+$ *supE*	27
A. vinelandii		
CA	Wild type	11
CA11	Δ*nifHDK*	5
CA80	*vnfH*::Kan^r	This study
CA11.80	Δ*nifHDK vnfH*::Kan^r	This study
RP1.11	Δ*nifHDK Bgl*II fragment (containing 3′ end of *vnfD*, *vnfG*, and 5′ end of *vnfK*) replaced by Spc^r gene	This study
Plasmids		
pUC9	Amp^r	40
pBS5	Amp^r, pUC9 containing *Sau*3A-*Eco*RI fragment (positions 1 to 1048 in Fig. 2)	This study
pBS13	Amp^r, pBS5 containing an insertion of Kan^r cartridge from pKIXX at *Bal*I site at position 707	This study
pKIXX	Amp^r Kan^r cartridge from Tn*5*	Pharmacia
pTZ19	Amp^r	Pharmacia
pLAM3	Amp^r, pTZ19 with *Bam*HI fragment containing *vnfHDGK* to position 5650 in Fig. 2	30
pLAM13	Amp^r, pTZ18 with 3.4-kilobase-pair *Eco*RI (5′ end *Eco*RI site at position 4648) fragment containing 3′ *vnfK*	This study
pJSM1	Amp^r, *Bgl*II fragment (positions 3292 to 4664 in Fig. 2) from pLAM3 replaced by Spc^c gene	This study

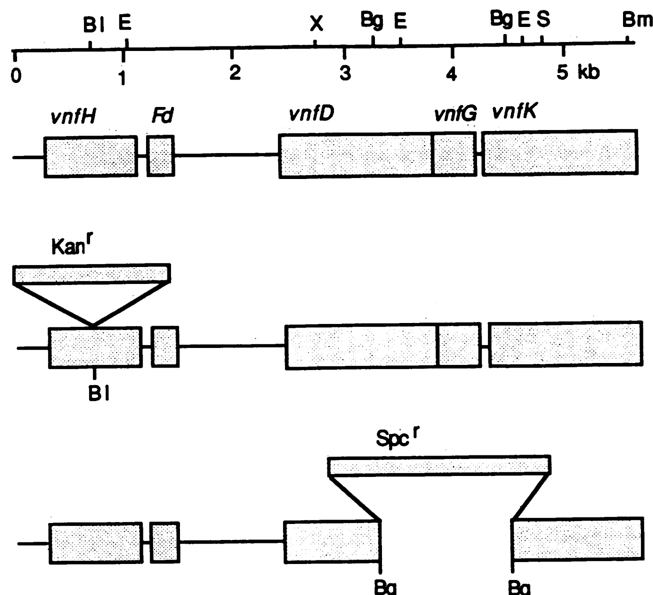

FIG. 1. Nitrogenase 2 structural gene cluster from *A. vinelandii* wild-type strain CA, Kan^r cartridge insertion strain CA80, and *vnfDGK* deletion/Spc^r cartridge insertion strain RP1. Restriction enzyme sites are indicated as follows: Bl (*Bal*I), E (*Eco*RI), X (*Xho*I), Bg (*Bgl*II), S (*Sma*I), Bm (*Bam*HI). Only one (position 707 in Fig. 2) of the eight *Bal*I sites is indicated. Strain *A. vinelandii* CA80 carries a Kan^r cartridge insertion at this *Bal*I site.

scribed strategy (23). Overlapping DNA fragments from the region containing the structural genes for nitrogenase 2 were isolated from a lambda EMBL3 recombinant clone containing insert DNA that hybridized to a *nifH* probe and from pLAM3. These overlapping fragments were cut with restriction enzyme *Alu*I, *Hae*III, *Rsa*I, *Sau*3A, or *Tha*I. The resulting fragments were ligated into M13 cloning vector mp18 or mp19 (42) followed by sequence determination, using the method of Sanger et al. (37). The sequence downstream from the *Bam*HI site at the 3′ end of *vnfK* was determined from ordered deletion derivatives of pLAM13 generated by exonuclease III (18).

Individual sequences were analyzed for overlaps and were organized into a contiguous sequence with the aid of a sequence alignment program (24). The DNA sequence was analyzed for base and codon preferences by using UWGCG computer programs (13). The GAP program in the UWGCG program package was used for alignments of nucleotide and amino acid sequences.

Construction of *A. vinelandii* mutants carrying insertion or deletion-plus-insertion mutations. Plasmid pBS5 containing an insert spanning the *Sau*3A site (position 1 in Fig. 2) to the *Eco*RI site (position 1048 in Fig. 2) was cleaved with *Bal*I (position 707), resulting in linearization of the plasmid. This linearized plasmid was combined with the Kan^r cartridge (released from pKIXX by *Sma*I digestion). After ligation and transformation into *E. coli* K-12 71-18, plasmid DNA was purified from kanamycin- and ampicillin-resistant colonies. *A. vinelandii* CA and CA11 (Δ*nifHDK*) were then transformed with this plasmid DNA, and kanamycin-resistant, but ampicillin-sensitive colonies were selected. Total genomic DNA was prepared from these strains (designated CA80 and CA11.80), and Southern hybridization analysis verified that the Kan^r cartridge was inserted into the *vnfH* gene of both of these strains. Strain RP1.11 was generated by transformation of strain CA11 with plasmid pJSM1 and selection for spectinomycin-resistant, but ampicillin-sensitive transformants. Plasmid pJSM1 was constructed from pLAM3 (pTZ19 with the *vnfHDGK*-containing *Bam*HI fragment from *A. vinelandii*) by replacement of the *Bgl*II fragment (position 3292 to 4664) with a spectinomycin resistance cartridge (33).

Two-dimensional gel electrophoresis. *A. vinelandii* CA11, CA11.80, and RP1.11 were derepressed for nitrogenase in N-free Mo-deficient Burk medium containing 1 μM V_2O_5 for 4 h. Cell-free protein extracts were then prepared as described previously (3). Isoelectric focusing and sodium dodecyl sulfate-polyacrylamide gel electrophoresis of proteins in the cell-free extracts were conducted by the method of O'Farrell (28) with modifications as described by Bishop et al. (3).

RESULTS AND DISCUSSION

Nucleotide sequence analysis. The nucleotide sequence of a 6,559-base-pair region of the *A. vinelandii* genome (Fig. 1) (containing a previously identified *nifH*-hybridizing region [20]) has been determined (Fig. 2). Five complete ORFs designated *vnfH*, a ferredoxinlike ORF, *vnfD*, *vnfG*, and *vnfK* were found. The designation of these ORFs as *vnf* genes is based on the phenotype of mutants carrying lesions in these ORFs and on comparisons of these genes and their predicted products with counterparts from the related organism *A. chroococcum* (35, 36). Thus, the designation *vnfH*

```
                  1 GATCGCTTCCCGGCTGTACCTGCGGGTACGTCGACGGAGCG     41

  42 CACAGCATCCTGGCCTGGATTTATGGAGTCCAATAAACCTGCAAAAATTAAAAT     95

  96 AATTCACTAATTAAATATGTTTTTTTGATTTTATATTCCAAAAAATAGGCAATC    149

 150 ATCGACTTATCGATCCTTGTGGCACCGCCCTTGCTCAACTCTCGTCGGCACAAA    203
                                                  vnfH ->
 204 TCAAAACGCCAACGAATCAACGGAGGTTCCTAAG ATG GCA TTG CGT CAG    252
                                         M   A   L   R   Q

 253 TGT GCA ATT TAC GGC AAG GGT GGC ATC GGC AAG TCC ACC ACC    294
      C   A   I   Y   G   K   G   G   I   G   K   S   T   T

 295 ACC CAG AAC CTG GTC GCC GCC CTC GCC GAA GCC GGC AAG AAA    336
      T   Q   N   L   V   A   A   L   A   E   A   G   K   K

 337 GTG ATG ATC GTC GGT TGT GAC CCG AAA GCC GAC TCC ACC CGC    378
      V   M   I   V   G   C   D   P   K   A   D   S   T   R

 379 CTG ATC CTG CAC TCC AAG GCC CAG GGC ACC GTC ATG GAA ATG    420
      L   I   L   H   S   K   A   Q   G   T   V   M   E   M

 421 GCC GCG TCC GCC GGC TCG GTC GAA GAC CTG GAG CTG GAA GAC    462
      A   A   S   A   G   S   V   E   D   L   E   L   E   D

 463 GTG CTG CAG ATC GGC TTC GGC GGC GTC AAG TGC GTC GAA TCC    504
      V   L   Q   I   G   F   G   G   V   K   C   V   E   S

 505 GGT GGC CCG GAG CCG GGC GTC GGC TGC GCC GGC CGT GGC GTG    546
      G   G   P   E   P   G   V   G   C   A   G   R   G   V

 547 ATC ACC GCG ATC AAC TTC CTG GAA GAA GAA GGC GCC TAC AGC    588
      I   T   A   I   N   F   L   E   E   E   G   A   Y   S

 589 GAC GAC CTG GAC TTC GTG TTC TAT GAC GTG CTG GGC GAC GTG    630
      D   D   L   D   F   V   F   Y   D   V   L   G   D   V

 631 GTA TGC GGC GGC TTC GCC ATG CCG ATC CGC GAG AAC AAG GCC    672
      V   C   G   G   F   A   M   P   I   R   E   N   K   A

 673 CAG GAA ATC TAC ATC GTC TGC TCC GGC GAG ATG ATG GCC ATG    714
      Q   E   I   Y   I   V   C   S   G   E   M   M   A   M

 715 TAC GCC GCC AAC AAC ATC GCC AAG GGC ATC GTG AAA TAC GCC    756
      Y   A   A   N   N   I   A   K   G   I   V   K   Y   A

 757 CAC TCC GGC AGC GTG CGT CTG GGC GGC CTG ATC TGC AAC AGC    798
      H   S   G   S   V   R   L   G   G   L   I   C   N   S

 799 CGC AAG ACC GAC CGC GAA GAC GAG CTG ATC ATG GCC CTG GCC    840
      R   K   T   D   R   E   D   E   L   I   M   A   L   A

 841 GCG AAG ATC GGC ACC CAG ATG ATC CAC TTC GTG CCG CGC GAC    882
      A   K   I   G   T   Q   M   I   H   F   V   P   R   D

 883 AAC GTC GTG CAA CAC GCC GAA ATC CGC CGC ATG ACC GTG ATC    924
      N   V   V   Q   H   A   E   I   R   R   M   T   V   I

 925 GAA TAC GAT CCG AAG GCC GGA CAG GCC GAC GAG TAC CGT GCC    966
      E   Y   D   P   K   A   G   Q   A   D   E   Y   R   A

 967 CTG GCT CGC AAA ATC GTC GAC AAC AAG CTG CTG GTC ATC CCG   1008
      L   A   R   K   I   V   D   N   K   L   L   V   I   P

1009 AAC CCG GCC TCC ATG GAA GAA CTC GAA GAG CTG CTA ATG GAA   1050
      N   P   A   S   M   E   E   L   E   E   L   L   M   E

1051 TTC GGC ATC ATG GAA GTC GAA GAC GAG TCC GTC GTC GGC AAG   1092
      F   G   I   M   E   V   E   D   E   S   V   V   G   K

1093 GCC GCC GCC GAA GGC TGA TTCACCCAGCACAGCGTTTGCGGAGGAGCGT   1141
      A   A   A   E   G   *
                                    <-------.    ------->
1142 GCGCCGCGGGCTTTCGGAATGGCTTCTCGCGGCCGGCGGCGCACGCCGCCCTCCC   1196
                                     ferredoxin-like ORF ->
1197 TTCGAACAACCGACCTCAGGAGCTGACACC ATG GCC ATG GCC ATC GAC   1243
                                     M   A   M   A   I   D

1244 GGC TAC GAA TGC ACT GTC TGC GGC GAC TGC GAG CCG GTC TGC   1285
      G   Y   E   C   T   V   C   G   D   C   E   P   V   C

1286 CCG ACC GGT TCG ATC GTC TTC AGG GAC GAT CAC TAC GCG ATC   1327
      P   T   G   S   I   V   F   R   D   D   H   Y   A   I

1328 GAA GCC GAC AGT TGC AAC GAA TGC ACC GAC GTG GGC GAG CCG   1369
      E   A   D   S   C   N   E   C   T   D   V   G   E   P

1370 CGC TGT CTC GGC GTC TGC CCC GTC GAC TTC TGC ATC CAG CCG   1411
      R   C   L   G   V   C   P   V   D   F   C   I   Q   P

1412 CTC GAT GAC TGA ACACTGAACGACTCCGCACCCCGTTGCCGGCGGCAGGACA 1464
      L   D   D   *

1465 TTCCGCGCCGTCCTGCCGCCGGACCCAGAACGGCGATCGCTTTTCCTCAGGGGCGA 1520

1521 TCGCCGTTTTACTTTTCCCCGCTCCGCTAGCGCCCGCGGACAACAGCCGTCCGTAT 1576

1577 CCGCGCCGTTCGCCGTCCTCGACCGGCCACGATCGCGGCACAGCTCGCCAAGCTAC 1632

1633 AGCCCGCTCCACAAGCTGACCATCGGCATCCAGACCTTCGCCAAGATCCGCGAAGA 1688

1689 CTTCTTACGTCGACAAGACCGGCTTCGCCCGGCGACCGATCGAACGGACACACTA 1744

1745 CTTCCTCTCCCGGCCATGCCGCTTCGGCAAGAGCCCGTTCCTCGATACCCTGGCCG 1800

1801 AACCGTTCGCCGGCGACGAACCTTCGTTCCGCAGGCTGCAGATGCATGACCGATGG 1856

1857 GACTGGCGCCGGCGCCATCCGGCATTGCGGATCGGCTTCGGCGGCGCGCCGATCCG 1912

1913 GGACGGCGGATAAGCTCGCCGAGCCGCTGCTCCCAAAGCCACGCCCACCTCCACGT 1968

1969 CCCGGCGCGGAAGTGGCGATCCCCAGCCGCACGTCCCCATGGTTTCCGGGTCTGGA 2024

2025 ACCGTCCCCCGGACTTCCCGTACGGACGCCCGGAAGCCACCGCCCCGACGGCCACG 2080

2081 CACAGCCACGTACAGACCCCCCGCTCTCCGTACTTGGCCGAACGAAAAACCTTTTA 2136

2137 CGTAACCCGCCGGCCGTCCACCCGGTTGTCCCGAACAGGAAAAAAAAGCCCGGAAA 2192

2193 AGGCTTGCCCCGGCGCTTTTCCAAAACCTCGAAAATGCGCACAAATAATTGATTCG 2248

2249 AAAGGATTAATCTGAGACAGCGGCGATGCCGCGAAAAATTCCTGCAAGCGCTGCAA 2304

2305 GGACATATGGCACGCATCCTGCCCTTACCTCTTCGCCAACCGGTTTTTCGGTCCCG 2360

2361  GTCGAGTGCCGTCAGGGGACTCGATCCACGCATAGAGCCGAGGAGACTTCAATC  2414
     vnfD ->
2415 ATG CCA ATG GTA TTG CTG GAA TGT GAC AAG GAC ATA CCC GAG   2456
      M   P   M   V   L   L   E   C   D   K   D   I   P   E

2457 CGC CAG AAA CAC ATC TAT CTG AAG GCG CCC AAC GAG GAC ACC   2498
      R   Q   K   H   I   Y   L   K   A   P   N   E   D   T

2499 CGC GAG TTC CTG CCG ATC GCC AAC GCG GCG ACC ATC CCC GGC   2540
      R   E   F   L   P   I   A   N   A   A   T   I   P   G

2541 ACC CTG TCC GAA CGC GGC TGC GCC TTC TGC GGC GCC AAG CTG   2582
      T   L   S   E   R   G   C   A   F   C   G   A   K   L

2583 GTG ATC GGC GGT GTG CTC AAA GAC ACC ATC CAG ATG ATT CAC   2624
      V   I   G   G   V   L   K   D   T   I   Q   M   I   H

2625 GGC CCG CTC GGC TGT GCC TAC GAC ACC TGG CAC ACC AAG CGC   2666
      G   P   L   G   C   A   Y   D   T   W   H   T   K   R

2667 TAC CCG ACC GAC AAC GGC CAC TTC AAC ATG AAG TAC GTC TGG   2708
      Y   P   T   D   N   G   H   F   N   M   K   Y   V   W

2709 TCG ACC GAC ATG AAG GAA AGC CAT GTG GTC TTC GGC GGC GAG   2750
      S   T   D   M   K   E   S   H   V   V   F   G   G   E

2751 AAA CGC CTC GAG AAG AGC ATG CAC GAA GCC TTC GAC GAA ATG   2792
      K   R   L   E   K   S   M   H   E   A   F   D   E   M

2793 CCC GAC ATC AAG CGG ATG ATC GTC TAC ACG ACC TGC CCG ACC   2834
      P   D   I   K   R   M   I   V   Y   T   T   C   P   T

2835 GCG CTG ATC GGC GAC GAC ATC AAG GCC GTG GCC AAG AAG GTG   2876
      A   L   I   G   D   D   I   K   A   V   A   K   K   V

2877 ATG AAG GAC CGT CCG GAC GTG GAC GTC TTC ACC GTC GAA TGC   2918
      M   K   D   R   P   D   V   D   V   F   T   V   E   C

2919 CCC GGC TTC TCC GGT GTG TCC CAG TCC AAG GGC CAC CAC GTC   2960
      P   G   F   S   G   V   S   Q   S   K   G   H   H   V

2961 CTG AAC ATC GGC TGG ATC AAC GAG AAA GTC GAG ACG ATG GAG   3002
      L   N   I   G   W   I   N   E   K   V   E   T   M   E

3003 AAG GAA ATC ACC AGC GAA TAC ACC ATG AAC TTC ATC GGT GAC   3044
      K   E   I   T   S   E   Y   T   M   N   F   I   G   D

3045 TTC AAT ATT CAA GGT GAT ACC CAA CTG CTG CAA ACC TAC TGG   3086
      F   N   I   Q   G   D   T   Q   L   L   Q   T   Y   W

3087 GAC CGC CTG GGC ATC CAG GTC GTC GCC CAC TTC ACC GGC AAC   3128
      D   R   L   G   I   Q   V   V   A   H   F   T   G   N

3129 GGC ACC TAC GAC GAC CTG CGC TGC ATG CAC CAG GCC CAG CTC   3170
      G   T   Y   D   D   L   R   C   M   H   Q   A   Q   L

3171 AAC GTG GTG AAC TGC GCC CGT TCC TCC GGC TAC ATC GCC AAC   3212
      N   V   V   N   C   A   R   S   S   G   Y   I   A   N

3213 GAG CTG AAG AAG CGC TAC GGC ATC CCG CGT CTG GAC ATC GAC   3254
      E   L   K   K   R   Y   G   I   P   R   L   D   I   D

3255 TCC TGG GGC TTC AAC TAC ATG GCC GAG GGC ATC CGC AAG ATC   3296
      S   W   G   F   N   Y   M   A   E   G   I   R   K   I

3297 TGC GCC TTC TTC GGC ATC GAG GAG AAG GGC GAG GAG CTG ATC   3338
      C   A   F   F   G   I   E   E   K   G   E   E   L   I

3339 GCC GAG GAA TAC GCC AAG TGG AAG CCG AAG CTC GAC TGG TAC   3380
      A   E   E   Y   A   K   W   K   P   K   L   D   W   Y

3381 AAG GAG CGT CTG CAA GGC AAG AAA ATG GCG ATC TGG ACC GGC   3422
      K   E   R   L   Q   G   K   K   M   A   I   W   T   G

3423 GGC CCG CGC CTG TGG CAC TGG ACC AAG TCG GTC GAG GAC GAC   3464
      G   P   R   L   W   H   W   T   K   S   V   E   D   D
```

FIG. 2. Nucleotide sequence of the nitrogenase 2 structural gene cluster (GenBank accession no. M32371). Deduced amino acid sequences are shown in single-letter code below the ORFs. Possible −12 to −24 sequences (NtrA-binding sites) and potential ribosome-binding sites are underlined. Inverted repeats are indicated by arrows.

```
3465 CTG GGC GTG CAA GTG GTG GCC ATG TCC TCC AAG TTC GGC CAT  3506
      L   G   V   Q   V   V   A   M   S   S   K   F   G   H

3507 GAG GAA GAC TTC GAG AAG GTC ATC GCC CGC GGC AAG GAA GGC  3548
      E   E   D   F   E   K   V   I   A   R   G   K   E   G

3549 ACC TAC TAC ATC GAC GAC GGC AAC GAG CTG GAA TTC TTC GAG  3590
      T   Y   Y   I   D   D   G   N   E   L   E   F   F   E

3591 ATC ATC GAC CTG GTC AAG CCG GAC GTG ATC TTC ACC GGC CCG  3632
      I   I   D   L   V   K   P   D   V   I   F   T   G   P

3633 CGC GTC GGC GAA CTG GTC AAG AAG CTG CAC ATC CCC TAC GTC  3674
      R   V   G   E   L   V   K   K   L   H   I   P   Y   V

3675 AAC GGC CAC GGC TAC CAC AAC GGC CCG TAC ATG GGC TTC GAA  3716
      N   G   H   G   Y   H   N   G   P   Y   M   G   F   E

3717 GGC TTC GTC AAC CTG GCC CGC GAC ATG TAC AAC GCC GTG CAC  3758
      G   F   V   N   L   A   R   D   M   Y   N   A   V   H

3759 AAC CCG CTG CGC CAC CTG GCC GCC GTG GAT ATC CGC GAC AAG  3800
      N   P   L   R   H   L   A   A   V   D   I   R   D   K
                                                    vnfG ->
3801 TCG CAG ACT ACC CCG GTC ATC GTG CGG GGG GCC GCC TG ATG   3841
      S   Q   T   T   P   V   I   V   R   G   A   A   *  M

3842 AGC CAG TCC CAT CTC GAC GAT CTG TTC GCC TAT GTC GAG GAG  3883
      S   Q   S   H   L   D   D   L   F   A   Y   V   E   E

3884 CGC TGC CTG TGG CAG TTC TTC TCG CGC ACC TGG GAC CGC GAG  3925
      R   C   L   W   Q   F   F   S   R   T   W   D   R   E

3926 GAA AAC ATC GAG GGC GTG CTC AAT CAG GTC GGC CGC CTG CTG  3967
      E   N   I   E   G   V   L   N   Q   V   G   R   L   L

3968 ACC GGC CAG GAG CCG CTG CGC GGC ACC CCG CAG GAG CGC CTG  4009
      T   G   Q   E   P   L   R   G   T   P   Q   E   R   L

4010 TTC TAC GCC GAC GCC CTG GCC ATG GCC AAC GAT GTC CGC GAG  4051
      F   Y   A   D   A   L   A   M   A   N   D   V   R   E

4052 CGT TTC CCC TGG GCT TCG CAG GTC AAC AAG GAA GAG ATC GAG  4093
      R   F   P   W   A   S   Q   V   N   K   E   E   I   E

4094 TTC CTG CTC GAC GGT CTC AAG TCC CGT CTG GTC GAC GTG ACC  4135
      F   L   L   D   G   L   K   S   R   L   V   D   V   T

4136 ATC ACC CGC TCG ACC AAC CGC GAA CTC AAC CAC CAC CTC TAC  4177
      I   T   R   S   T   N   R   E   L   N   H   H   L   Y

            <------------      ------------>.
4178 TGA GTCATCGGCGCGGAGACGGCCGGAACGCCGTCTGCGCGCGAAGGAGCCA  4231

           vnfK ->
4232 TTTTC ATG AGC AAT TGC GAA CTG ACC GTG CTG AAG CCG GCA   4272
            M   S   N   C   E   L   T   V   L   K   P   A

4273 GAA GTC AAG CTG AGC CCG CGT GAC CGC GAG GGC ATC ATC AAC  4314
      E   V   K   L   S   P   R   D   R   E   G   I   I   N

4315 CCG ATG TAC GAC TGC CAG CCG GCC GGC GCC CAG TAC GCC GGC  4356
      P   M   Y   D   C   Q   P   A   G   A   Q   Y   A   G

4357 ATC GGC ATC AAG GAC TGC ATC CCG CTG GTC CAC GGC GGC CAG  4398
      I   G   I   K   D   C   I   P   L   V   H   G   G   Q

4399 GGC TGC ACG ATG TTC GTC CGC CTG CTG TTC GCC CAG CAC TTC  4440
      G   C   T   M   F   V   R   L   L   F   A   Q   H   F

4441 AAG GAA AAC TTC GAC GTC GCC TCC ACC TCG CTG CAC GAG GAG  4482
      K   E   N   F   D   V   A   S   T   S   L   H   E   E

4483 TCG GCG GTG TTC GGC GGC GCC AAG CGC GTC GAG GAA GGC GTG  4524
      S   A   V   F   G   G   A   K   R   V   E   E   G   V

4525 CTG GTC CTC GCC CGC CGC TAC CCG AAC CTG CGC GTC ATC CCG  4566
      L   V   L   A   R   R   Y   P   N   L   R   V   I   P

4567 ATC ATC ACC ACC TGC TCC ACC GAA GTC ATC GGC GAC GAC ATC  4608
      I   I   T   T   C   S   T   E   V   I   G   D   D   I

4609 GAG GGC AGC ATC CGG GTC TGC AAC CGG GCA CTC GAA GCC GAA  4650
      E   G   S   I   R   V   C   N   R   A   L   E   A   E

4651 TTC CCG GAT CGC AAG ATC TAC CTG GCG CCG GTA CAC ACC CCG  4692
      F   P   D   R   K   I   Y   L   A   P   V   H   T   P

4693 AGC TTC AAG GGC AGC CAC GTC ACC GGC TAC GCC GAG TGC GTG  4734
      S   F   K   G   S   H   V   T   G   Y   A   E   C   V

4735 AAG TCG GTG TTC AAG ACC ATC ACC GAC GCG CAC GGC AAG GGC  4776
      K   S   V   F   K   T   I   T   D   A   H   G   K   G

4777 CAG CCG AGC GGC AAG CTC AAC GTG TTC CCG GGC TGG GTC AAC  4818
      Q   P   S   G   K   L   N   V   F   P   G   W   V   N

4819 CCC GGC GAC GTG GTG CTG CTC AAG CGC TAC TTC AAG GAA ATG  4860
      P   G   D   V   V   L   L   K   R   Y   F   K   E   M

4861 GAC GTC GAA GCC AAC ATC TAC ATG GAC ACC GAG GAC TTC GAC  4902
      D   V   E   A   N   I   Y   M   D   T   E   D   F   D

4903 TCG CCG ATG CTG CCG AAC AAG AGC ATC GAG ACC CAC GGC CGC  4944
      S   P   M   L   P   N   K   S   I   E   T   H   G   R

4945 ACC ACC GTC GAG GAC ATC GCC GAC AGC GCC AAC GCC CTG GCC  4986
      T   T   V   E   D   I   A   D   S   A   N   A   L   A

4987 ACC CTG TCC CTG GCC CGC TAC GAG GGC AAC ACC ACC GGC GAG  5028
      T   L   S   L   A   R   Y   E   G   N   T   T   G   E

5029 TTG CTG CAG AAG ACC TTC GCG GTG CCG AAT GCC CTG GTC AAC  5070
      L   L   Q   K   T   F   A   V   P   N   A   L   V   N

5071 ACC CCC TAC GGC ATC AAG AAC ACC GAC GAC ATG CTG CGC AAG  5112
      T   P   Y   G   I   K   N   T   D   D   M   L   R   K

5113 ATC GCC GAA GTC ACC GGC AAG GAG ATC CCC GAG TCG CTG GTC  5154
      I   A   E   V   T   G   K   E   I   P   E   S   L   V

5155 CGC GAG CGC GGC ATC GCC CTC GAC GCC CTG GCC GAC CTG GCG  5196
      R   E   R   G   I   A   L   D   A   L   A   D   L   A

5197 CAC ATG TTC TTC GCC AAC AAG AAA GTG GCG ATC TTC GGC CAC  5238
      H   M   F   F   A   N   K   K   V   A   I   F   G   H

5239 CCG GAC CTG GTG CTC GGC CTG GCC CAG TTC TGC ATG GAA GTC  5280
      P   D   L   V   L   G   L   A   Q   F   C   M   E   V

5281 GAG CTG GAA CCC GTA CTG CTG CTG ATC GGC GAC GAC CAG GGC  5322
      E   L   E   P   V   L   L   L   I   G   D   D   Q   G

5323 AAC AAG TAC AAG AAG GAC CCG CGC ATC GAG GAG CTG AAG AAC  5364
      N   K   Y   K   K   D   P   R   I   E   E   L   K   N

5365 ACC GCG CAC TTC GAC ATC GAG ATC GTC CAC AAC GCC GAC CTC  5406
      T   A   H   F   D   I   E   I   V   H   N   A   D   L

5407 TGG GAA CTG GAG AAG CGC ATC AAC GCC GGC CTC CAG CTC GAC  5448
      W   E   L   E   K   R   I   N   A   G   L   Q   L   D

5449 CTG ATC ATG GGT CAC TCG AAG GGC CGC TAC GTC GCC ATC GAG  5490
      L   I   M   G   H   S   K   G   R   Y   V   A   I   E

5491 GCC AAC ATC CCG ATG GTC CGC GTC GGC TTC CCG ACC TTC GAC  5532
      A   N   I   P   M   V   R   V   G   F   P   T   F   D

5533 CGC GCC GGT CTC TAC CGC AAG CCC TCG ATC GGC TAC CAG GGC  5574
      R   A   G   L   Y   R   K   P   S   I   G   Y   Q   G

5575 GCC ATG GAA CTG GGC GAG ATG ATC GCC AAC GCC ATG TTC GCC  5616
      A   M   E   L   G   E   M   I   A   N   A   M   F   A

5617 CAC ATG GAA TAC ACC CGT AAC AAG GAG TGG ATC CTC AAT ACG  5658
      H   M   E   Y   T   R   N   K   E   W   I   L   N   T

5659 TGG TGA GTTGAGGTGCCGGAGCGGTTTCCACGCACTCCGGCTGTCGAGCCGAC  5711
      W   *
                                  <---------   .--------->
5712 CGAGATGACGGCATGCACGGATCTCCCCTCGCCCATGCATGGGCGAGGGAGTTGC  5766

5767 GAGCCGTCACGCCCGAGGTCGTTCCGACGGCAGGCCGATGCACCGGAAAACCGCT  5821

5822 CCCGCCCCCGGCGGCTCCCGGCTCCGGCCGGGGCCTCCGCCCCGTATTCCAGATC  5876

5877 CGGCCGCCATGCCGCATTCCCGGGAGTGATGCCATGAAACAGCGACAGGAAATGG  5931

5932 TCGCCCACTACCGCGCCTGCTTCGGCGAGCTGTGCGCCCGACCGGAACACCGTCC  5986

5987 TATCGAACCCTACACCCGCCCCCGGCGCCTGAGCTTCGCGGAGCCGGAAAGCGCG  6041

6042 ACCGCCCGTCAGGTGCCGGCGCGCCTGGTACTGGCCCTGACCAGTGCCTACGCCC  6096

6097 TGCTCGCCGACTGGCAGGAATCCCGCGACCCGTCGCTGGCCGACCTGGGCAGTTG  6151

6152 GCAACGCTACCTCGCCCTGCCCCGCCGCACCCCGGCGGAAAAGCTGATGGCCGAG  6206

6207 ATCTTCCGCATCCTGCGCGTGTTCCGCGCCGCGGCGATCCAGCAGAACGGCGCCA  6261

6262 TCGAAATCCGCGACGACGGCCTGATCCGCGCCAGTTGCACCTACAACCGCTGCGC  6316

6317 GCTGAGCCTGCTGATCACCCAGGCCGGCCTCGAACTGCTCACCGCCTGCGTCGCC  6371

6372 TGCTACCTCGAATCCTTCGACCAGCCCTACAGCGATGCCTACGAGGAACTGCTGC  6426

6427 TCGGCCAGTACTACGCCGACATCGTCGCCGAGATCCGCTCCTTCGCCGACGACGA  6481

6482 CCGGGTGCTGTTCCAGTTCCGCCAGAAAGCCTGGTTCAACCGCCATGTCCGCCTG  6536

6537 GACTGCGACAACCCGCGCCTGCA 6559
```

```
NIFH  AAATCAATGGTTTATTTATGTGTTGCGGGTGCTGGCACAGACGCTGCATTACCT
      :::  :     : ::  :  : : :    :  :::::: : :  :::   ::
VNFH  AAAATAGGCAATCATCGACTTATCGATCCTTGTGGCACCGCCCTTGCTCAACTC
       ::: :     :    :   :         : ::::::    : ::: :    :
VNFD  GAAAAAATTCCTGCAAGCGCTGCAAGGACATATGGCACGCATCCTGCCCTTACC
           : ::  :    :    :   ::  :  ::::::      :::   ::
ANFH  TCGCGACTTGTTTGGTGGCGCGGCCGGTAAACTGGCACATGCATTGCTTATATA
```

FIG. 3. Comparison of the sequences of proposed promoter regions for the structural genes of nitrogenases 1, 2, and 3.

rather than *nifH2* (34) is used for the first ORF. Analysis of presumed noncoding regions upstream from *vnfH* (positions 167 to 184) showed that a sequence similar to those of *ntrA*-dependent promoters (1) is present (Fig. 3). This potential promoter is preceded by a very AT-rich region (positions 84 through 142) with a G+C content of approximately 10%. (The G+C content of the nitrogenase 2 structural gene region as a whole is 61%.)

A 116-base-pair region separates *vnfH* from a ferredoxinlike (Fd) gene. The nucleotide sequence of the region containing *vnfH* and the Fd gene differs from that reported by Raina et al. (34) at the following positions (numbers for nucleotides are those from Fig. 2). Five additional nucleotides were reported to be present at a position corresponding to nucleotides 54 and 55 upstream from *vnfH*. Within *vnfH*, nucleotide 437 (C) was reported to be a G. Thus, our sequence predicts a serine residue where a tryptophan residue was reported previously for the predicted amino acid sequence of *vnfH*. The nucleotide corresponding to nucleotide 743 (T) was reported to be a G residue. Therefore, a serine residue was predicted previously, whereas our data predict an isoleucine residue. In the region between *vnfH* and the ferredoxinlike ORF, nucleotide 1176 (G) has not been reported previously. Raina et al. (34) reported an additional nucleotide (C) following nucleotide 1206. Downstream from the ferredoxinlike ORF, nucleotide 1457 (G) was not present in the previously reported sequence. It should be noted that the DNA sequences reported by Raina et al. (34) were from the same *A. vinelandii* strain as those reported in this study. The ferredoxinlike gene (whose function is still unknown) is separated from *vnfD* by about 1 kilobase pair of DNA. This intergenic region apparently does not contain any coding sequences as judged by base preference, codon usage, and lack of potential initiation codons preceded by possible ribosomal binding sites. Transcription initiated from the *vnfH* promoter appears to terminate within this region because the largest *nifH*-hybridizing transcript is only large enough to accommodate *vnfH* and the ferredoxinlike ORF (20). A potential NtrA-dependent promoter site is located upstream of *vnfD* at positions 2311 to 2327. This site is preceded by an AT-rich region (33% G+C content at positions 2209 through 2259). The presence of these AT-rich sequences near *ntrA*-dependent promoter sites is a feature shared by the majority of proposed promoter sites for *nif*, *vnf*, and *anf* operons from *A. vinelandii* (19, 23) and *A. chroococcum* (35, 36). It does not appear that a consensus sequence can be derived from these AT-rich sequences, and their distances from the proposed promoter sites vary. Therefore, it has been suggested that these regions serve as generalized RNA polymerase recognition sites (1). Differences in location and nucleotide sequence may then influence the initiation of transcription from adjacent promoters (19).

Genes *vnfD*, *vnfG*, and *vnfK* are located downstream from this putative promoter. With the exception of *vnfG*, all ORFs are preceded by sequences that resemble ribosome-binding sites with respect to nucleotide content and spacing from the

TABLE 2. Comparison of genes and gene products for nitrogenases from *A. vinelandii*

Gene (gene product)	% Identity with corresponding gene (gene product)[a]	
	nif (NIF)	*anf* (ANF)
vnfH (VNFH)	88.5 (91.0)	70.1 (63.5)
vnfD (VNFD)	52.6 (33.0)	65.8 (54.4)
vnfG (VNFG)		55.9 (39.8)
vnfK (VNFK)	51.4 (31.1)	69.8 (57.4)

[a] Percent identity refers to identity at the DNA sequence level for the gene (indicated by lowercase italics) and the identity at the level of the amino acid sequence for the gene product (in parentheses).

initiation codon. There is a sequence, GGGGGG, 7 bases upstream from the proposed initiation codon for *vnfG*, and it is possible that this sequence acts as part of a ribosome-binding site. Alternatively, the translation machinery might not dissociate from the RNA template upon completion of translation of *vnfD* and, following appropriate repositioning, might proceed to translate *vnfG*. Robson et al. (36) found that in *A. chroococcum* the product of *vnfG* is present in equimolar amounts to the *vnfD* and *vnfK* products. This indicates that *vnfG* is translated as efficiently as the adjacent genes.

mRNA from the transcriptional unit containing *vnfDGK* may form a stem-loop structure in the intergenic region between *vnfG* and *vnfK*. The significance of such potential structures, which can also be predicted for transcripts from intergenic regions in the *nifHDK* operon (7), is currently not understood. Also, 72 base pairs downstream from the stop codon of *vnfK* there is a GC-rich dyad symmetry immediately followed by consecutive thymidine residues, which corresponds to a rho-independent terminator of transcription (6). This and the maximum size (3.4 kilobases) of *vnfDGK* transcripts (36), indicate that *vnfDGK* transcription terminates a short distance after the end of the *vnfK* coding sequence. Analysis of the sequence 826 base pairs downstream from *vnfK* did not reveal any potential genes.

Comparison of the three regions encoding nitrogenase structural proteins. In Table 2, the similarity of the genes present in the three regions is expressed as a percentage of identical residues. The 5′ and 3′ noncoding regions of these genes do not share any significant nucleotide sequence identity. However, the proposed binding site for an RNA polymerase sigma-54 transcription factor (1) is conserved (Fig. 3). As pointed out above, AT-rich regions are found upstream from all of these binding sites. A possible NifA-binding site (5′ TGT-N_{10}-ACA 3′) (8) is only found upstream from the *nifH* promoter (19). It is safe to assume that regulatory sites are also present upstream of *vnfH*, *vnfD*, and *anfH*, because the expression of the respective products is suppressed by ammonia and Mo (3, 4, 20). However, the location and nucleotide sequence of these presumed regulatory sites are unknown.

The structural genes for nitrogenases 1 (*nifHDK*) and 3 (*anfHDGK*) are apparently transcribed from a single promoter preceding the genes encoding dinitrogenase reductase subunits (19, 23). In contrast, the structural genes for nitrogenase 2 are organized in two separate transcriptional units. This has also been observed for the structural genes of V-nitrogenase from *A. chroococcum* (35, 36). In this organism the region separating the ferredoxinlike ORF from *vnfD* is approximately twice as long as the corresponding region in

FIG. 4. Alignment of the predicted amino acid sequences of the subunits of the three nitrogenases from *A. vinelandii*. (A) Dinitrogenase reductases; (B) α subunits; (C) δ subunits; (D) β subunits.

A. vinelandii. The nucleotide sequence of the region downstream from the ferredoxinlike ORF from *A. chroococcum* has been determined recently, and it appears to contain an ORF of still unknown function (R. L. Robson, personal communication). It is not completely understood why the dinitrogenase reductase 2 and dinitrogenase 2 structural genes are organized as two separate transcriptional units. However, it has been observed previously that the product of *vnfH*, but not those of *vnfD* and *vnfK*, was found not only in cells grown under Mo-deficient diazotrophic conditions in the presence of V, but also in cells grown under Mo and V deficiency (4, 31). In fact, the *vnfH* product (R. D. Joerger et al., submitted for publication), but not the products of *vnfDGK*, is required for expression of nitrogenase 3 (30). *A. chroococcum*, on the other hand, does not appear to contain a nitrogenase corresponding to nitrogenase 3 from *A. vinelandii*. Therefore, independent expression of the *vnfH*-Fd operon may be required for other reasons in this organism.

Comparison of products of the three nitrogenase structural

TABLE 3. Predicted molecular weights and pIs of nitrogenase structural gene products

Gene	Total no. of amino acids	Calculated mol wt of predicted products	Estimated pI
vnfH	290	31,026	4.58
nifH	290	31,514	4.48
anfH	275	29,883	4.72
vnfD	474	53,874	5.94
nifD	492	55,285	5.93
anfD	518	58,409	6.34
vnfG	113	13,337	4.83
anfG	132	15,342	5.00
vnfK	475	52,772	5.45
nifK	523	59,455	5.98
anfK	462	51,176	4.95

gene clusters. The translation products of structural genes for the three nitrogenases from *A. vinelandii* are compared in Fig. 4A to D, and the molecular weights and pIs are listed in Table 3. The differences in molecular weight and pI from those reported earlier (19, 23) are due to the use of a different version (2.0) of the IBI Pustell programs. The values listed in Table 3 agree with those obtained with the PeptideMap program from the UWGCG program package.

Comparisons of the predicted amino acid sequences of the nitrogenase 2 subunits from *A. vinelandii* and *A. chroococcum* (35, 36) show that these subunits retain a high degree of similarity (91.1 to 98.3% identity). However, mutations have accumulated to a different degree in the individual structural genes since the two *Azotobacter* species diverged. There are 36 differences between the predicted amino acid sequences of the *vnfK* gene products, but only 7 differences between the *vnfD* products. Detailed comparisons of predicted products of the structural genes of nitrogenases 1 and 3 from *A. vinelandii* (23) and of nitrogenase 2 from *A. chroococcum* (35, 36) have been presented previously (23). Due to the close relationship of the predicted products of the nitrogenase 2 structural genes in *A. vinelandii* and *A. chroococcum*, the conclusions drawn in the earlier comparisons remain unchanged. The previously reported (36) sequence for *vnfD* from *A. chroococcum* has been revised, and the predicted amino acid sequence of the *vnfD* gene product is now identical to that from *A. vinelandii* at residues 61 to 69 (R. Robson, personal communication).

The most interesting result from the comparisons of the structural gene products is the high degree of similarity between *vnfH* and *nifH*, but not between these two genes and *anfH*. As pointed out earlier (23), the *anfH* gene product is more closely related to the products of *nifH1* from *Methanococcus thermolithotrophicus* (38) and *nifH3* from *Clostridium pasteurianum* (41). On the other hand, the products of *vnfD* and *vnfK* share a higher degree of sequence identity with the *anfD* and *anfK* products than either pair does with the products of *nifD* and *nifK*.

Analysis of mutant strains CA80, CA11.80, RP1, and RP1.11. Two-dimensional gels of protein extracts obtained from cells derepressed for nitrogenase 2 in medium containing 1 μM V_2O_5 are shown in Fig. 5. Three prominent protein spots, which had been observed previously in two-dimensional gels (4, 32), could be observed in gels of extracts from strain CA11 (Δ*nifHDK*) (Fig. 5A). Extracts from strain CA11.80 (Δ*nifHDK vnfH*::Kan^r) (Fig. 5B) did not contain a

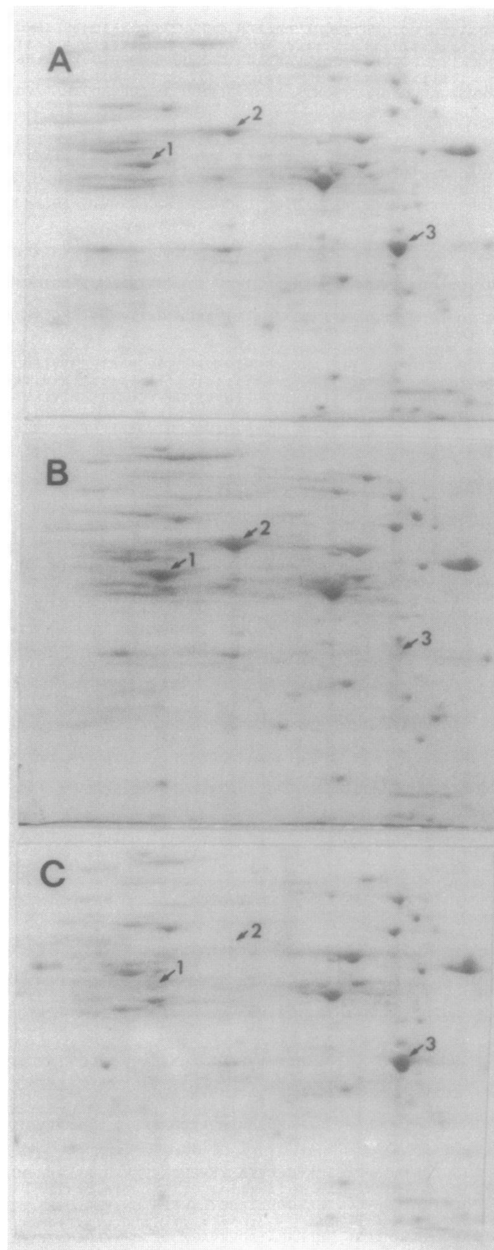

FIG. 5. Two-dimensional gels of protein extracts from cells derepressed for nitrogenase 2 in medium containing 1 μM V_2O_5. Numbered arrows indicate positions where previously described spots representing NH_4^+-repressible proteins are found (4, 32). (A) *A. vinelandii* CA11. Spot 1 (*vnfD* gene product), spot 2 (*vnfK* gene product), and spot 3 (*vnfH* gene product) are present. (B) Strain CA11.80. Only spots 1 and 2 are present. (C) Strain RP1.11. Only spot 3 is present.

protein corresponding to spot 3, a spot that has been attributed previously to dinitrogenase reductase 2 subunits (4, 32). Extracts from strain RP1.11 (Δ*nifHDK* Δ*vnfDGK*) did not contain proteins corresponding to spots 1 and 2 (Fig. 5C). Thus, these spots represent the products of *vnfD* and *vnfK*. The calculated pI values suggest that spot 1, as the more basic protein, represents the *vnfD* gene product, while spot 2 originates from the product of *vnfK*.

A. vinelandii CA80 (*vnfH*::Kan^r) and RP1 (Δ*vnfDGK*) grow as well as the wild-type strain CA under N_2-fixing

conditions in the presence of 1 μM Na_2MoO_4. These strains also grow in N-free medium deficient in both Mo and V, but as described previously for strain RP114 (30), they grow slowly in medium containing 1 μM V_2O_5. Strain RP1.11 (Δ*nifHDK* Δ*vnfDGK*) does not grow diazotrophically in Mo-sufficient medium, but grows in Mo- and V-deficient medium. Growth in the presence of V is very slow and comparable to that reported previously for strain RP206 (30). Strain CA11.80 (Δ*nifHDK vnfH*::Kan^r) does not grow diazotrophically whether or not Mo or V is present. A strain comparable to strain CA11.80, *A. vinelandii* RR29 (Δ*nifHDK vnfH*::Tn*5*), has been described previously (34). This strain is unable to grow in the presence of either Mo or V. Acetylene reduction activity by whole cells incubated in medium containing 50 nM V_2O_5 was only a small fraction of the activity by the parental strain CA12 (34). Growth and acetylene reduction activity, however, were not determined for Mo- and V-deficient conditions. The results obtained with the different mutant strains described above demonstrate that the genes designated *vnfH*, *vnfD*, *vnfG*, and *vnfK* encode the structural genes for nitrogenase 2. Furthermore, these results may also provide some insight into the regulation of the nitrogenase systems present in *A. vinelandii*. Thus, strains that carry deletions in *vnfDGK* or in both *vnfDGK* and *nifHDK* grow in Mo-deficient N-free medium or even in Mo-deficient V-containing medium, albeit at a slow rate. This growth is due to the expression of nitrogenase 3 and the *vnfH*-Fd operon under these conditions. Strains carrying mutations in *vnfH* require the presence of the structural genes for nitrogenase 1 (*nifHDK*) for the expression of nitrogenase 3, since only strain CA80, and not strain CA11.80, expressed nitrogenase 3. These strains, and others, will provide a basis for further studies on the regulatory role of the proposed products of the transcriptional unit containing *vnfH* and the ferredoxinlike ORF.

ACKNOWLEDGMENTS

This study was supported by U.S. Department of Agriculture competitive grant 88-37120-3872 and by NATO travel grant no. 0532/88. This was a cooperative study of the Agricultural Research Service, U.S. Department of Agriculture and the North Carolina Agricultural Research Service.

LITERATURE CITED

1. **Beynon, J., M. Cannon, V. Buchanan-Wollaston, and F. Cannon.** 1983. The *nif* promoters of *Klebsiella pneumoniae* have a characteristic primary structure. Cell **34**:665–671.
2. **Bishop, P. E., M. E. Hawkins, and R. R. Eady.** 1986. Nitrogen fixation in Mo-deficient continuous culture by a strain of *Azotobacter vinelandii* carrying a deletion of the structural genes for nitrogenase (*nifHDK*). Biochem. J. **238**:437–442.
3. **Bishop, P. E., D. M. L. Jarlenski, and D. R. Hetherington.** 1980. Evidence for an alternative nitrogen fixation system in *Azotobacter vinelandii*. Proc. Natl. Acad. Sci. USA **77**:7342–7346.
4. **Bishop, P. E., D. M. L. Jarlenski, and D. R. Hetherington.** 1982. Expression of an alternative nitrogen fixation system in *Azotobacter vinelandii*. J. Bacteriol. **150**:1244–1251.
5. **Bishop, P. E., R. Premakumar, D. R. Dean, M. R. Jacobson, J. R. Chisnell, T. M. Rizzo, and J. Kopczynski.** 1986. Nitrogen fixation by *Azotobacter vinelandii* strains having deletion in structural genes for nitrogenase. Science **232**:92–94.
6. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential procaryotic terminators. Nucleic Acids Res. **12**:4411–3327.
7. **Brigle, K. E., W. E. Newton, and D. R. Dean.** 1985. Complete nucleotide sequence of the *Azotobacter vinelandii* nitrogenase structural gene cluster. Gene **37**:37–44.
8. **Buck, M., S. Miller, M. Drummond, and R. Dixon.** 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature (London) **320**:374–378.
9. **Bulen, W. A., and J. R. LeComte.** 1966. The nitrogenase system from *Azotobacter*: two enzyme requirement for N_2 reduction, ATP-dependent hydrogen evaluation, and ATP hydrolysis. Proc. Natl. Acad. Sci. USA **56**:979–989.
10. **Burgess, B. K., D. B. Jacobs, and E. I. Stiefel.** 1980. Large-scale purification of high activity *Azotobacter vinelandii* nitrogenase. Biochim. Biophys. Acta **614**:196–209.
11. **Bush, J. A., and P. W. Wilson.** 1959. A non-gummy chromogenic strain of *Azotobacter vinelandii*. Nature (London) **184**: 381–382.
12. **Chisnell, J. R., R. Premakumar, and P. E. Bishop.** 1988. Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii*. J. Bacteriol. **170**: 27–33.
13. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
14. **Eady, R. R., R. L. Robson, R. N. Pau, P. Woodley, D. J. Lowe, R. W. Miller, R. N. F. Thornley, B. E. Smith, C. Gormal, K. Fisher, M. Eldridge, and J. Bergstrom.** 1988. The vanadium nitrogenase of *Azotobacter chroococcum*, p. 81–86. *In* H. Bothe, F. J. de Bruijn, and W. E. Newton (ed.), Nitrogen fixation: hundred years after. Gustav Fischer Verlag, Stuttgart, Federal Republic of Germany.
15. **Eady, R. R., R. L. Robson, T. H. Richardson, R. W. Miller, and M. Hawkins.** 1987. The vanadium nitrogenase of *Azotobacter chroococcum*. Purification and properties of the VFe protein. Biochem. J. **244**:197–207.
16. **Hales, B. J., D. J. Langosch, and E. E. Case.** 1986. Isolation and characterization of a second nitrogenase Fe-protein from *Azotobacter vinelandii*. J. Biol. Chem. **261**:15301–15306.
17. **Hales, B. J., E. E. Morningstar, M. F. Dzeda, and L. A. Mauterer.** 1986. Isolation of a new vanadium-containing nitrogenase from *Azotobacter vinelandii*. Biochemistry **25**:7251–7255.
18. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155**:156–165.
19. **Jacobson, M. R., K. E. Brigle, L. T. Bennett, R. A. Setterquist, M. S. Wilson, V. L. Cash, J. Benyon, W. E. Newton, and D. R. Dean.** 1989. Physical and genetic map of the major *nif* gene cluster from *Azotobacter vinelandii*. J. Bacteriol. **171**:1071–1027.
20. **Jacobson, M. R., R. Premakumar, and P. E. Bishop.** 1986. Transcriptional regulation of nitrogen fixation by molybdenum in *Azotobacter vinelandii*. J. Bacteriol. **167**:480–486.
21. **Joerger, R. D., and P. E. Bishop.** 1988. Nucleotide sequence and genetic analysis of the *nifB-nifQ* region from *Azotobacter vinelandii*. J. Bacteriol. **170**:1475–1487.
22. **Joerger, R. D., and P. E. Bishop.** 1988. Bacterial alternative nitrogen fixation systems. Crit. Rev. Microbiol. **16**:1–14.
23. **Joerger, R. D., M. R. Jacobson, R. Premakumar, E. D. Wolfinger, and P. E. Bishop.** 1989. Nucleotide sequence and mutational analysis of structural genes (*anfHDGK*) for the second alternative nitrogenase from *Azotobacter vinelandii*. J. Bacteriol. **171**:1075–1086.
24. **Johnston, R. E., J. Mackenzie, Jr., and W. G. Dougherty.** 1986. Assembly of overlapping DNA sequences by a program written in BASIC for 64K CP/M and MS-DOS IBM-compatible microcomputers. Nucleic Acids Res. **14**:517–527.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–79.
27. **Messing, J., B. Gronenborn, B. Mueller-Hill, and P. H. Hofschneider.** 1977. Filamentous coliphage M13 as a cloning vehicle: insertion of a *Hind*II fragment of the *lac* regulatory region in M13 replicative form *in vitro*. Proc. Natl. Acad. Sci. USA **74**:3642–3646.
28. **O'Farrell, P. H.** 1975. High resolution two-dimensional electro-

phoresis of proteins. J. Biol. Chem. **250**:4007–4021.

29. **Page, W. J., and M. von Tigerstrom.** 1979. Optimal conditions for transformation of *Azotobacter vinelandii*. J. Bacteriol. **139**: 1058–1061.
30. **Pau, R. N., L. A. Mitchenall, and R. L. Robson.** 1989. Genetic evidence for an *Azotobacter vinelandii* nitrogenase lacking molybdenum and vanadium. J. Bacteriol. **171**:124–129.
31. **Premakumar, R., J. R. Chisnell, and P. E. Bishop.** 1989. A comparison of the three dinitrogenase reductases expressed by *Azotobacter vinelandii*. Can. J. Microbiol. **35**:344–348.
32. **Premakumar, R., E. M. Lemos, and P. E. Bishop.** 1984. Evidence for two dinitrogenase reductases under regulatory control by molybdenum in *Azotobacter vinelandii*. Biochim. Biophys. Acta **797**:64–70.
33. **Prentki, P., and H. M. Krisch.** 1984. *In vitro* mutagenesis with a selectable DNA fragment. Gene **29**:303–313.
34. **Raina, R., M. R. Reddy, D. Ghosal, and H. K. Das.** 1988. Characterization of the gene for the Fe-protein of the vanadium dependent alternative nitrogenase of *Azotobacter vinelandii* and construction of a Tn*5* mutant. Mol. Gen. Genet. **214**:121–127.
35. **Robson, R. L., P. R. Woodley, and R. Jones.** 1986. Second gene (*nifH**) coding for a nitrogenase iron-protein in *Azotobacter chroococcum* is adjacent to a gene coding for a ferredoxin-like protein. EMBO J. **5**:1159–1163.
36. **Robson, R. L., P. R. Woodley, R. N. Pau, and R. R. Eady.** 1989. Structural genes for the vandium nitrogenase from *Azotobacter chroococcum*. EMBO J. **8**:1217–1224.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Souillard, N., and L. Sibold.** 1989. Primary structure, functional organization snd expression of nitrogenase structural genes of the thermophilic archaebacterium *Methanococcus thermolithotrophicus*. Mol. Microbiol. **3**:441–552.
39. **Strandberg, G. W., and P. W. Wilson.** 1968. Formation of the nitrogen-fixing enzyme system in *Azotobacter vinelandii*. Can. J. Microbiol. **14**:25–31.
40. **Viera, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
41. **Wang, S.-Z, J.-S. Chen, and J. L. Johnson.** 1988. The presence of five *nifH*-like sequences in *Clostridium pasteurianum*: sequence divergence and transcriptional properties. Nucleic Acids Res. **16**:439–454.
42. **Yanish-Perron, C., J. Viera, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13 mp18 and pUC19 vectors. Gene **33**:103–110.