Further Analysis of Nitrogen Fixation (nif) Genes in *Azotobacter chroococcum*: Identification and Expression in *Klebsiella pneumoniae* of *nifS, nifV, nifM* and *nifB* Genes and Localization of *nifE/N*, *nifU*, *nifA*- and *fixABC*-like Genes

By DAVID EVANS,*† ROBERT JONES,‡ PAUL WOODLEY AND ROBERT ROBSON§

AFRC-IPSR Division of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, UK

(Received 6 April 1987; revised 23 November 1987)

The results presented extend previous investigations on the genetics of nitrogen fixation in *Azotobacter chroococcum* and indicate that nif- and fix-like DNA is located in at least five different regions of the genome. Region I contains functional copies of *nifS, V* and *M*, as well as *nifH, D* and *K*, all of which complemented mutants of *Klebsiella pneumoniae*. In addition, *nifE*- and/or *nifN*-like and *nifU*-like DNA is located in this region. The organization of the nif cluster in region I closely resembles that of *K. pneumoniae*, though spread over 22 kb as compared with 14 kb. Region II contains a functional *nifB* gene, which complemented a *K. pneumoniae* *nifB* mutant, and seems to be adjacent to a *nifA*-like gene. Region III harbours *nifH*, encoding a second nitrogenase Fe-protein. Region IV contains a reiteration of *nifE*- and/or *nifN*-like sequences, and DNA homologous to *Rhizobium meliloti fixABC* is present in region V. The apparent complexity of nif DNA in *A. chroococcum* is probably related to the two systems for N\textsubscript{2}-fixation present in this organism.

INTRODUCTION

Members of the genus *Azotobacter* are heterotrophic, obligately aerobic N\textsubscript{2}-fixing bacteria. Studies of the genetics of N\textsubscript{2} fixation in this genus should contribute to an understanding of their tolerance of air when fixing N\textsubscript{2} (see Robson & Postgate, 1980) and their long-recognized ability to use vanadium in place of molybdenum for N\textsubscript{2} fixation (Bortels, 1936).

The genes (*nif*) for N\textsubscript{2} fixation were defined in the facultative anaerobe *Klebsiella pneumoniae*, in which 17 genes are specific for N\textsubscript{2} fixation and are arranged as a contiguous cluster organized into seven or eight transcriptional units (see Dixon, 1984). Nitrogenase, the two-component enzyme complex which catalyses the reduction of N\textsubscript{2}, requires three genes: *nifH, nifD* and *nifK* encode respectively the Fe-protein (component 2) and the alpha- and beta-subunits of the MoFe-protein (component 1). Both components of nitrogenase require additional gene products for activity. *nifM* is required to give an active Fe-protein (Howard *et al.*, 1986). *nifE*, *nifN* and *nifB* are required for the synthesis or insertion of the Fe- and Mo-containing cofactor (FeMo-co). The *nifV* gene product subtly modifies the co-factor, altering the substrate specificity of the resulting

---

* † Present address: Department of Microbiology, University of Reading, London Road, Reading RG1 5AQ, UK.
‡ Present address: Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, USA.
§ Present address: Department of Biochemistry, University of Georgia, Athens GA 30602, USA.

Abbreviations: Restriction enzymes are abbreviated as follows: *BamHI, Ba*; *BgIII, Bg*; *Clal, C*; *EcoRI, E*; *HindIII, H*; *KpnI, K*; *PstI, P*; *SacI, Sc*; *SalI, Sl*; *SmaI, Sm*; *XhoI, X*.
nitrogenase. The \textit{nif}Q product is required in media low in molybdenum. An electron-transport chain specific for nitrogenase is encoded by the \textit{nif}I and \textit{nif}F gene products, which are respectively a pyruvate flavodoxin oxidoreductase and a flavodoxin. \textit{nif}A encodes a positive activator (see Dixon, 1984; Gussin \textit{et al.}, 1986) required for transcription of the other \textit{nif} operons, whilst \textit{nif}L codes for an anti-activator which, in response to fixed N or to O$_2$, antagonizes the activity of the \textit{nif}A gene product. The function of the \textit{nif}S gene is unknown, though mutations in this gene confer a Nif$^{-}$ phenotype, and the roles of the \textit{nif}X, \textit{nif}U and \textit{nif}Y genes have yet to be determined.

A basic core of \textit{nif} genes are probably common to all diazotrophs (see in Evans \textit{et al.}, 1985), but different supplementary genes may be present in some of them. The \textit{fix}ABC genes are required for N$_2$ fixation in \textit{Rhizobium meliloti} (Ruvkun \textit{et al.}, 1982; Pühler \textit{et al.}, 1984) and, though present in all rhizobia, they are not homologous to any of the \textit{K. pneumoniae} \textit{nif} genes (Ausubel \textit{et al.}, 1985). Moreover, these genes may not be confined to symbiotic N$_2$ fixers, since \textit{fix}A-like DNA was detected in \textit{Azospirillum} sp. (Fogher \textit{et al.}, 1985), and the finding of \textit{jix}ABC-like DNA present in all rhizobia, they are not homologous to any of the \textit{K. pneumoniae} \textit{nif} genes (Ausubel \textit{et al.}, 1985). In previous studies of N$_2$ fixation in \textit{Azotobacter chroococcum} MCD1 we cloned a region of about 70 kb of the chromosome which contains functional copies of the \textit{nif}HDK genes, separated by approximately 15 kb from a region homologous to a \textit{K. pneumoniae} \textit{nif} probe bearing \textit{nif}V and parts of \textit{nif}M and \textit{nif}S (Jones \textit{et al.}, 1984). At least one other set of functional nitrogenase genes is present in \textit{A. chroococcum} since a strain deleted for the \textit{nif}HDK cluster was still capable of fixing N$_2$ in molybdenum-deficient medium (Robson, 1986). This is consistent with the presence of \textit{nif}H*, coding for a second nitrogenase Fe-protein (Robson \textit{et al.}, 1986a), and a second \textit{nif}K-like sequence in the genome of this organism (Robson, 1986). N$_2$ fixation in the deletion strain depends upon vanadium and is catalysed by a two-component nitrogenase complex with a typical Fe-protein (component 2) but in which the conventional molybdoprotein (component 1) is replaced by a vanadoprotein. Hence \textit{A. chroococcum} is capable of synthesizing genetically distinct V- or Mo-nitrogenases (Robson \textit{et al.}, 1986b).

In this work we have extended our investigations on \textit{nif} genes and their organization in \textit{A. chroococcum} by hybridization with heterologous probes, genetic complementation and DNA sequencing. The results, together with those of our previous studies, suggest that \textit{nif}- and \textit{fix}-like DNA is dispersed in at least five regions of the chromosome.

**METHODS**

\textit{Growth and maintenance of strains}. Bacteria used in this study are listed in Table 1. \textit{E. coli} or \textit{K. pneumoniae} strains were grown aerobically at 37°C on Luria–Bertoni medium (LB) or anaerobically at 30°C on nitrogen-free defined medium NFDM (Cannon \textit{et al.}, 1974) with appropriate growth factors added where required. Antibiotics were used at the following concentrations (\textmu g ml$^{-1}$): kanamycin, 25; carbenicillin, 50; tetracycline, 10; streptomycin, 50; chloramphenicol, 50. \textit{A. chroococcum} strains were grown under air at 30°C on RM medium (Robson \textit{et al.}, 1984) with antibiotics added when required at the following concentrations (\textmu g ml$^{-1}$): streptomycin, 20; nalidixic acid, 20; kanamycin, 0-25. For solidified medium, agar was added at 1.5\% (w/v). \textit{E. coli} and \textit{K. pneumoniae} strains were preserved at −20°C after suspension in 50\% (v/v) glycerol; \textit{A. chroococcum} was maintained on RM agar slopes at 20°C. Plasmids and vectors used in this study are also listed in Table 1.

\textit{Isolation of genomic and plasmid DNA}. Genomic DNA was prepared as described by Robson \textit{et al.} (1984), and plasmids by the method of Birnboim & Doly (1979), which was scaled up for large amounts.

\textit{Electrophoresis}. Plasmid or DNA fragments were electrophoresed in agarose gels (0.8\%, w/v; Sigma, type II) in TAE buffer (Maniatis \textit{et al.}, 1982). DNA fragments were recovered from low-melting-point agarose (BRL) by a 'freeze-squeeze' method (Thuring \textit{et al.}, 1975).

\textit{Hybridizations}. DNA was transferred by electrophoresis from agarose gel to Genescreen (New England Nuclear) as described in the product handbook. Radioactive DNA probes were prepared by nick-translation (Rigby \textit{et al.}, 1977) using a commercial kit (BRL) with deoxyctydine 5'-\textit{P}[$\alpha$-\textit{32}P]triphosphate (3000 Ci mmol$^{-1}$, 111 TBq mmol$^{-1}$; Amersham). DNA was labelled to approx. 2 × 10$^8$ d.p.m. \mu g$^{-1}$. Hybridizations were carried out for 16 h at 42°C, usually in 50\% (v/v) formamide (high stringency) in Denhardt's solution (Denhardt, 1966) containing 10\% (w/v) dextran sulphate (\textit{M}, 500 000; Pharmacia) (Wahl \textit{et al.}, 1979) with 0.1\% sodium pyrophosphate, 0.1\% sodium dodecyl sulphate and 100 \mu g calf thymus DNA ml$^{-1}$. Stringency of hybridization was adjusted by varying the formamide concentration. Blots were washed free of unbound radioactivity first with
### Table 1. Azotobacter chroococcum nif genes

<table>
<thead>
<tr>
<th>Relevant genotypic or phenotypic characteristics</th>
<th>Reference or origin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Azotobacter chroococcum</strong></td>
<td></td>
</tr>
<tr>
<td>MCD1 Sm&lt;sup&gt;+&lt;/sup&gt; Na&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Robson et al. (1984)</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
</tr>
<tr>
<td>5K res&lt;sup&gt;K&lt;/sup&gt; thr lew-31 tonA supT 2&lt;sup&gt;+&lt;/sup&gt;</td>
<td>B. Spratt, Sussex University</td>
</tr>
<tr>
<td>71/18 Δ[lac-proF[lacI&lt;sup&gt;+&lt;/sup&gt; lacZ M15 proAB]]</td>
<td>Messing et al. (1977)</td>
</tr>
<tr>
<td><strong>Klebsiella pneumoniae</strong></td>
<td></td>
</tr>
<tr>
<td>UNF50231 His&lt;sup&gt;+&lt;/sup&gt; revertant of UNF5023 Nif&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Dixon et al. (1977)</td>
</tr>
<tr>
<td>UNF50581 His&lt;sup&gt;+&lt;/sup&gt; revertant of KP5058 nif/B213</td>
<td></td>
</tr>
<tr>
<td>UNF1145 nifJ2194::Tn7</td>
<td>Merrick et al. (1978)</td>
</tr>
<tr>
<td>UNF1146 nifJ2183::Tn7</td>
<td>Merrick et al. (1980)</td>
</tr>
<tr>
<td>UNF2050 nifM2104</td>
<td></td>
</tr>
<tr>
<td>UNF2142 nifS2442</td>
<td>Merrick et al. (1982)</td>
</tr>
<tr>
<td>UNF812 nifH2183: :Tn5</td>
<td></td>
</tr>
<tr>
<td>UNF854 nifH2104: :Tn5</td>
<td></td>
</tr>
<tr>
<td>UNF2161 nifU2461</td>
<td></td>
</tr>
<tr>
<td>UNF745 nifA2787: :Mu</td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
</tr>
<tr>
<td>1. Inserts of <em>A. chroococcum</em> DNA</td>
<td></td>
</tr>
<tr>
<td>pACB1</td>
<td></td>
</tr>
<tr>
<td>pACD30</td>
<td></td>
</tr>
<tr>
<td>pACD37</td>
<td></td>
</tr>
<tr>
<td>pLC11 Cosmid clone (pTBE) Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pER5 1·7 kb <em>SalI</em> fragment (pEMBL9&lt;sup&gt;+&lt;/sup&gt;) Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pEF3a/b 4·0 kb <em>EcoRI</em> fragment (pEMBL9&lt;sup&gt;+&lt;/sup&gt;) Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pSAR2a/b 5·1 kb Kpn1 fragment (pBR325) Cb&lt;sup&gt;+&lt;/sup&gt; Te&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>pSAR3a/b 3·6 kb <em>EcoRI/PsiI</em> fragment (pBR325) Cb&lt;sup&gt;+&lt;/sup&gt; Te&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pRR2/3 4·0 kb <em>EcoRI</em> fragment (pBR325) Cb&lt;sup&gt;+&lt;/sup&gt; Te&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pDE13 6·3 kb Kpn1 fragment (pEMBL18&lt;sup&gt;+&lt;/sup&gt;) Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>2. Inserts of <em>K. pneumoniae</em> DNA</td>
<td></td>
</tr>
<tr>
<td>pBCC12 niF</td>
<td></td>
</tr>
<tr>
<td>pMC5 niM</td>
<td></td>
</tr>
<tr>
<td>pMC11 niUX</td>
<td></td>
</tr>
<tr>
<td>pMC16 niI</td>
<td></td>
</tr>
<tr>
<td>pPC936 nifQBALF</td>
<td></td>
</tr>
<tr>
<td>pWK26 nifEN</td>
<td></td>
</tr>
<tr>
<td>pWF1 nifSVM</td>
<td></td>
</tr>
<tr>
<td>3. Inserts of <em>Rhizobium</em> DNA</td>
<td></td>
</tr>
<tr>
<td>pL1286 <em>R. leguminosarum</em> fixZ</td>
<td></td>
</tr>
<tr>
<td>pRm3 <em>R. meliloti</em> fixABC</td>
<td></td>
</tr>
<tr>
<td>4. Vectors and others</td>
<td></td>
</tr>
<tr>
<td>pBR325 Cb&lt;sup&gt;+&lt;/sup&gt; Te&lt;sup&gt;+&lt;/sup&gt; Cm&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pEMBL9&lt;sup&gt;+&lt;/sup&gt; Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pEMBL18&lt;sup&gt;+&lt;/sup&gt; Cb&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>pRK2013 Km&lt;sup&gt;+&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

At least two washes in STE buffer (NaCl, 0·3 M; Tris/HCl pH 8·0, 0·06 M; EDTA, 0·002 M) for 5 min each at room temperature, then two washes of 30 min each at 60 °C in STE containing SDS (0·5%, w/v) and finally two washes of 30 min each in 0·1 x STE at room temperature. DNA size markers were either *HindIII* or *HindIII/EcoRI* restriction digests of phage 4 DNA or a commercial 1 kb ladder (BRL). Radioactive DNA size markers were prepared by end-labelling fragments with DNA polymerase Klenow fragment (Amersham) with [α-<sup>32</sup>P]dATP. Autoradiographs were produced on Kodak XAR-5 film.

Hybridization probes are listed in Table 1 and illustrated in Figs 1 and 3.

**DNA sequencing.** This was done by the chain-termination method (Sanger et al., 1977) using [α-<sup>35</sup>S]dATP (Biggin et al., 1983). Preparation of templates is described elsewhere (Robson et al., 1986a). DNA sequences were
analysed using the University of Wisconsin Genetics Computer Group BESTFIT program run at the suggested default settings. Figs 5 and 6 were produced using the program PUBLISH.

Transformations and conjugations. DNA was transformed into E. coli and K. pneumoniae strains by the method of Kushner (1978) except that K. pneumoniae was ‘freeze-thawed’ three times after addition of the DNA and no ‘heat shock’ was used (J. Gibbins, M. Merrick and J. Postgate, unpublished data). The Mob* cosmid pLC11 was introduced into K. pneumoniae strains in tripartite crosses with the helper plasmid pRK2013.

Complementation studies. Genetic complementation of K. pneumoniae Nif" mutants by cloned A. chroococcum DNA was tested in two ways. For growth on N₂, strains were patched onto NFDM agar plates containing antibiotics appropriate for the clones examined. Plates were incubated under anaerobic conditions at 30°C and growth was assessed after 4 d. Complementation was considered successful when growth was comparable to that of UNF50231, a Nif* strain of K. pneumoniae. Complementation was also determined by measuring acetylene-reducing activity of strains in liquid culture by the method described in Dixon et al. (1977). In mutants for which positive results were obtained by the growth test, acetylene-reducing activities reached rates of the same order as those typical for UNF50231, whereas in strains not complemented for growth on N₂, the activities were not significantly changed.

RESULTS

Localization of nif- and fix-like DNA in A. chroococcum

DNA fragments containing nif genes from K. pneumoniae and fix genes from R. leguminosarum and R. meliloti (Table 1, Fig. 1) were tested for homology to A. chroococcum genomic DNA and to the recombinant cosmids pACB1, pACD30 and pACD37 (Table 1, Fig. 3), which contain DNA surrounding the nifHDK cluster from A. chroococcum (Jones et al., 1984). Hybridization was observed under stringent conditions to cosmids pACB1 and pACD37, but not to pACD30, with the K. pneumoniae nif‘EN” probe (probe 2, Fig. 2a), and to pACB1, but not to pACD30, with a K. pneumoniae nif‘UX” probe (probe 3, Fig. 2b). We also located more precisely the region homologous to the K. pneumoniae nif‘SVM” probe (probe 4, Fig. 2c) that we reported previously.

K. pneumoniae nif DNA

R. meliloti fix DNA

R. leguminosarum fix DNA

Fig. 1. DNA hybridization probes. Probes are defined by number. DNA was prepared by digesting the appropriate plasmid with the restriction enzymes shown and the fragment purified by electrophoresis on an agarose gel from which the DNA was recovered. Fragments used are shown by the solid lines above the position of the various genes. Probe 7 was the HindIII-SmaI fragment and probe 8 was the adjacent SmaI-SmaI fragment. Information about the position of genes with respect to restriction sites was taken from published data or provided by personal communication (see Table 1). See also Fig. 3.
Fig. 2. Detection of *nif*- and *fix*-like DNA in *A. chroococcum* DNA. Parts (a) to (c) represent restriction digests of cosmids pACB1 (lane 1), pACD37 (lane 2) and pACD30 (lane 3). Parts (d) to (f) represent restriction digests of *A. chroococcum* genomic DNA. Restriction enzymes are abbreviated above the lanes. In each figure, the left side shows the ethidium-bromide-stained gel photographed under UV illumination. The right side shows the autoradiograph of the blot obtained from the gel following hybridization to the radioactive probes as defined in Figs 1 and 3. (a) *nif*EN, probe 2; (b) *nifUX*, probe 3; (c) *nif*SV, probe 4; (d) *fixZ*, probe 10; (e) *fixABC*, probe 9; (f) *nif*EN-like DNA from *A. chroococcum*, probe 11. All experiments were done under stringent conditions (50%, v/v, formamide at 42°C) except in (d) and (e), where the formamide concentration was lowered to 40%, (v/v). DNA size markers are shown in kb to the left of each figure.
Fig. 3. Organization of nif genes and nif-like sequences in region I of the chromosome of A. chroococcum. The upper section shows a map of restriction sites in that region of the chromosome of A. chroococcum obtained from cosmids pACB1, pACD37 and pACD30. The relative positions of the cosmid clones and other plasmids used in this study are shown below. DNA fragments used to prepare radioactive DNA probes are indicated by the circled numbers. The lower part summarizes the relative positions of nif genes or nif-like sequences in this region as compared to the order in K. pneumoniae taken from Beynon et al. (1983). Where the positions of the genes are uncertain they are boxed by discontinuous lines. Regions showing homology to the K. pneumoniae nif probes used in this study are denoted by the hatched boxes.

(Jones et al., 1984). The positions of the homologous DNA are shown relative to a map of the three cosmids in Fig. 3. We did not detect homology to these cosmids or to genomic DNA digests with K. pneumoniae DNA containing nif'J' (probe 1), nif'M' (probe 5), nif'QB' (probe 7), or nif'BAL' (probe 8). Weak homology was observed to all three cosmids over a wide region (Fig. 3) with K. pneumoniae nif'F' (probe 6) at 40% (w/v) formamide (data not shown).

Hybridizations to R. leguminosarum fixZ (nifB) DNA (probe 10) revealed homologous sequences in A. chroococcum genomic DNA (Fig. 2d), though not in the cosmids. A probe
Azotobacter chroococcum nif genes

Fig. 4. Cloning of the nifB gene from A. chroococcum. The upper section shows a map of restriction sites for the 4.0 kb EcoRI fragment in pEF3a/b containing A. chroococcum DNA which hybridizes to fixZ (nifB) from R. leguminosarum. The heavy arrows below indicate DNA sequencing reactions used to locate the nifB- and nifA-like sequences and to determine the likely positions (boxes) and direction of transcription (open arrows) of the genes.

containing the R. meliloti fixABC genes (probe 9) hybridized well to A. chroococcum genomic DNA (Fig. 2e). Again the homologous sequence was not located on the cosmids.

In view of the dual N₂-fixation systems in A. chroococcum we tested for reiteration of nif DNA other than that coding for nitrogenase polypeptides. A. chroococcum DNA fragments were used to prepare probes since, as in the case of nifK (Jones et al., 1984; Robson, 1986), reiteration may not be detectable with the K. pneumoniae DNA. The A. chroococcum nifEN-like sequence (probe 11, Fig. 3) not only detected the fragments expected in the genome (6.3 kb KpnI, 10-7 kb SmaI, 5.4 kb SalI, 10 kb and 3 kb EcoRI (Fig. 2, Fig. 3) but also revealed a second homologous region, even at high stringency (Fig. 2f), located on 4 kb SmaI, 3-3 kb SalI and 6 kb EcoRI fragments. The sizes of the latter fragments preclude the nifEN-like reiteration from lying within the region cloned in the cosmids pACB1, pACD37 and pACD30. However, reiteration was not apparent when DNA from the region homologous to the K. pneumoniae nif'SVM' DNA (probe 12, Fig. 3) was used (data not shown).

Cloning of the fixZ (nifB)-like gene from A. chroococcum

A 4.0 kb EcoRI fragment from A. chroococcum which hybridized to the fixZ gene (probe 10, see Fig. 2d) was cloned as follows. EcoRI-restricted genomic DNA fragments of between 3 and 5 kb were cloned into pEMBL9 +, and recombinant plasmids containing DNA homologous to fixZ were identified by colony screening. Two types of plasmids containing identical 4.0 kb EcoRI inserts cloned in either orientation were obtained. A restriction map of the insert in one of the plasmids (pEF3a) is shown in Fig. 4. The location and orientation of the fixZ (nifB)-like gene was determined by limited DNA sequencing from defined restriction sites in the region containing the homologous sequence (Fig. 5), as predicted from genomic hybridization results. The nifB-like sequence was identified by computer alignment (Fig. 5) against the published R. meliloti nifB sequence (Buikema et al., 1987).

Computer analysis of the sequence at one end of pEF3a identified good homology to the C-terminal of the K. pneumoniae nifA gene. Alignment of the DNA, and predicted amino acid sequences for the region present in pEF3, are shown in Fig. 6. This implies that nifB and the adjacent nifA-like genes are transcribed in the same direction (Fig. 4).

Expression of presumptive A. chroococcum nifS, V, M and B genes in K. pneumoniae

Complementation of defined K. pneumoniae nif mutants was performed to determine whether the A. chroococcum nif-like DNA encodes functional genes. The cosmid pLC11 contains a 21 kb
Fig. 5. Location of nifB-like sequence in pEF3a/b by DNA sequence analysis. Comparison of the DNA sequence for 841 bp spanning the unique BamHI, XhoI and SmaI sites in pEF3a (see Fig. 4) with the first 861 bp of R. meifotii nip coding sequence (Buikema et al., 1987). The sequences were aligned with the University of Wisconsin Genetics Computer Group program BESTFIT run at the suggested default settings. Numbering of the R. meifotii sequence is taken from Buikema et al. (1987); the nip coding sequence starts at 199 bp.

The A. chroococcum nifS, I/ and M genes were more precisely located using subcloned fragments of pACB1. K. pneumoniae nifM (UNF2050, UNF828) and nifV (UNF812) mutants were complemented for growth on N₂ by the plasmid pSAR2a (Table 1, Fig. 3), containing the insert of A. chroococcum DNA and complements a postulated A. vinelandii nifM mutant (Kennedy et al., 1986). The insert lies within the region cloned in pACB1 – spanning the nifU-like and nifSVM-like regions – but does not contain the entire nifEN-like region (Fig. 3). pLC11 complemented nifU (UNF2161), nifS (UNF2142, UNF866), nifV (UNF812) and nifM (UNF828) mutants for growth on N₂ and restored wild-type levels of acetylene-reducing activity, but did not complement nifJ (UNF1145), nifE (UNF864), nifF (UNF1146), nifA (UNF745) or nifB (UNF50581) mutants.

The A. chroococcum nifS, V and M genes were more precisely located using subcloned fragments of pACB1. K. pneumoniae nifM (UNF2050, UNF828) and nifV (UNF812) mutants were complemented for growth on N₂ by the plasmid pSAR2a (Table 1, Fig. 3), containing the
Azotobacter chroococcum nif genes

Kp nifA
CAGTTCCTTCGCTGAAAGAAATCGCCACAGCGCCAGGGCGAACGCTGCGCATCAGCGATGGGGCGATTCGCCTGCTGATGGAGTACAGCTGG

Ac nifA
GAATCTCACACAGGACAGGGCGACAGCGCCAGCCCCGTCAGCAACGACGGCCAGCCATGCTCGGATCTGCG

EFLLTKIGRQQGRPLTVTDSAIRLLMSHRW

** ** *** * * ****** *

PGNVRDVENCLERSAESLGDVILF

***** ******** ** * * ***

PGNVRDVENCLERSAESLGDVILF

** ** ****

TGVDNESPPLAAPLP-EVNLADENTLDDRER

** ** ****

VIAALEQAGWVQAIARLLGMTPQIAYRI

** ** ****

CAGACCCTCAACATCCACATGCGCAAGATCTAGGACTCGCATTGATGCGCTTCCTCCGTCCGATCGTCGCAGGCCTCAGCGG

NHRDNPPKALASSGPAEDGMDLNSLDERQR

** ** ****

LIAALEKAGWYQAKAAHERLLGMTPRQVAYRI

** ** ****

VIAALEQAGWYQAKAAHERLLGMTPRQIAYRI

** ** ****

CAGATCTGCCCATCGGAAAGCGGCTGAGGCGCACGGGCGAACGCTGCGCATCAGCGATGGGGCGATTCGCCTGCTGATGGAGTACAGCTG

QIMDITMPRL*

** ** ****

QIMDITMPRL*

** ** ****

QTLNHEIRKI*

** ** ****

CAGATTATGYATACATCCATGACCAGGTGTCATTAATGATAGGACAGTAGCTAACTGCGCGAAGCGCCGCGGATTTG

QIMDITMPRL*

** ** ****

QIMDITMPRL*

** ** ****

CAGACCCTCAACATCCACATGCGCAAGATCTAGGACTCGCATTGATGCGCTTCCTCCGTCCGATCGTCGCAGGCCTCAGCGG

NifI

Fig. 6. Identification and location of the nifA-like sequence in pEF3a/b. The DNA sequence and predicted amino acid sequence for the C-terminal end of nifA from K. pneumoniae (Buikema et al., 1985; Drummond et al., 1986) are compared with those for 438 bp of DNA located at one end of the cloned A. chroococcum DNA in pEF3a/b. Restriction sites shown correlate with those located to the right in the map shown in Fig. 4. Alignment of the two sequences was achieved using the University of Wisconsin Genetics Computer Group program BESTFIT, which introduces a 3 bp space in the A. chroococcum sequence at a position equivalent to nucleotide 226 in the K. pneumoniae DNA sequence.

5·1 kb KpnI fragment carrying part of the nifSVM-like DNA under the control of the promoter for chloramphenicol acetyltransferase (CAT) in the EcoRI site of pBR325. The plasmid pSAR3a (Table 1, Fig. 3), containing the 3·6 kb EcoRI–PstI fragment from the nif'SVM'·like region (Fig. 3), cloned under CAT promoter control complemented K. pneumoniae nifS (UNF2142, UNF866) and nifV (UNF812) mutants well for growth on N2. Plasmids containing the same inserts cloned in the opposite orientation with respect to the CAT promoter (pSAR2b, pSAR3b) did not complement K. pneumoniae nifS, V, or M mutants for growth on N2. N2 fixation in UNF50231 (Nif+ control) was not prevented by pSAR2a/b, pSAR3a/b or pBR325. These results indicate that the A. chroococcum nifS, V, and M genes are closely grouped, and are organized and transcribed in the order nifS, nifV, nifM.

The EcoRI fragment from pEF3a containing the nifB gene from A. chroococcum was likewise cloned into the EcoRI site of pBR325. The resultant plasmids pRR2 and 3, containing the insert in the two possible orientations, both complemented a K. pneumoniae nifB mutant (UNF50581) for growth with N2 equally well. This implies that the A. chroococcum nifB gene is expressed from its own promoter in K. pneumoniae, which presumably is situated between the putative nifA gene and nifB.

The good complementation of nifU, S, V, M, and B mutants of K. pneumoniae by A. chroococcum DNA demonstrates a high degree of functional homology of these genes in the two organisms. This is in contrast to the nitrogenase structural genes of A. chroococcum, which complemented Nif− mutants of K. pneumoniae for acetylene reduction, but not for growth on N2 (Jones et al., 1984).
DISCUSSION

All eight genes known to code for the Mo-nitrogenase polypeptides and products required for their processing in *K. pneumoniae* appear to be present in *A. chroococcum*. The *A. chroococcum* nifH, D and K genes (Jones et al., 1984) and nifV, nifM, and nifB genes complement *K. pneumoniae* mutants. A region showing homology to *K. pneumoniae* nifE and/or nifN is located adjacent to the *A. chroococcum* nifH, D and K genes. Although we have yet to demonstrate that this region contains functional genes this seems likely in view of recent results for *A. vinelandii*, in which nifE and nifN have been identified in a comparable genomic location (Brigle et al., 1987). Based on our complementation results nifS is also present in *A. chroococcum*, and both our complementation and hybridization data suggest that nifU (whose function in *K. pneumoniae* is unknown) may also be conserved. A region showing weak homology to *K. pneumoniae* nifF was located, although as hybridization covered a wide region other corroborative evidence is needed to confirm that a nifF-like gene maps in this position.

As previously postulated (Kennedy & Robson, 1983; Jones et al., 1984) *A. chroococcum* seems to contain a nifA-like gene, which is apparently located adjacent to nifB. A nifJ-like gene was not detectable by heterologous hybridization. This may not be surprising since a Nif-related polypeptide of equivalent size to the nifJ product of *K. pneumoniae* seems to be absent in *A. chroococcum* (Robson, 1979). However, in *A. vinelandii* there is evidence that electrons for nitrogenase may be derived from NADPH since an NADPH-dehydrogenase activity is induced upon N-starvation, which coincides with the appearance of two membrane-bound polypeptides, one apparently associated with the NADPH dehydrogenase complex (Klugkist et al., 1986). The genetic determinants of this system have not yet been described, though Gubler & Hennecke (1986) and Earl et al. (1987) have suggested that one or more of the fixABC gene products may be involved in electron transport. Our data indicate that one or more of the fixABC genes may be present in *A. chroococcum*.

*nif*- or *fix*-like DNA is dispersed in five regions of the chromosome of *A. chroococcum*. The organization of *nif* DNA in one region (region I) of the chromosome of *A. chroococcum* is strikingly similar to that of the *K. pneumoniae* nif cluster between and including nifH and nifM. Region I contains genes for at least the synthesis and maturation of a functional Mo-nitrogenase, as suggested by our complementation studies in *K. pneumoniae*, which as far as is known cannot synthesize a vanadium nitrogenase. Although the order of *nif* genes or *nif*-like sequences is similar, the genes are spread over 22 kb in *A. chroococcum* compared to 14 kb in *K. pneumoniae*. By comparison to the *K. pneumoniae* nif cluster it appears that an additional 6 kb sequence separates the nifE/N-like and nifU-like regions in the *A. chroococcum* nif cluster. At present we cannot say whether this region harbours genes required for *N₂* fixation in *A. chroococcum*.

Other *nif*- and *fix*-like DNA is relatively dispersed in the genome of *A. chroococcum*. We define the other regions as follows. Region II contains the nifB gene, probably closely linked to nifA; region III contains nifH* and its adjacent ferredoxin gene; region IV and region V contain the second nifEN-like sequence and the fixABC-like sequence respectively. At present, we cannot deduce linkage between these five regions.

The relationship at the genetic level between the Mo- and V-based *N₂* fixation systems is particularly interesting. Considering the broad similarities in the two sets of nitrogenase proteins (Robson et al., 1986a; Robson, 1986) it is probable that the two systems have arisen from one or more gene duplication events. Not all genes may be duplicated; some may be common to both systems, others unique to each system. There was no clear evidence, in our experiments using the *R. leguminosarum* fixZ probe, for multiple nifB-like sequences. 'Back-hybridization' with a fragment from the nifSV complementing region suggests that similar sequences are also not reiterated in the genome of *A. chroococcum*. In *A. vinelandii* nifM appears common to both conventional and alternative systems (Kennedy et al., 1986) and if the alternative system in this organism is analogous to that of *A. chroococcum* then presumably this gene processes both Fe-proteins. By contrast, nifN appears not to be required for *N₂* fixation under Mo-deficient conditions in *A. vinelandii* (Kennedy et al., 1986). This might not be unexpected since this gene is involved in FeMo-co synthesis or insertion into the MoFe-protein (Roberts et al., 1978). However, our hybridization results may mean that the second nifEN-like
region specifies analogous genes for the V-nitrogenase, which would imply that a cofactor similar to FeMo-co yet containing V in place of Mo is present in the V-nitrogenase.

We thank Drs Mike Merrick, Ray Dixon, Frank Cannon, Maura Cannon, Werner Klipp, Andy Johnston, Claudine Elmerich and Christina Kennedy for providing bacterial strains, plasmids and information in some cases prior to publication. We thank Professor John Postgate for helpful criticism of the manuscript and Beryl Scutt for typing. We also thank Media Services of the University of Sussex for photographic work.

REFERENCES


