The EMBO Journal vol.8 no.4 pp.1217 - 1224, 1989

Structural genes for the vanadium nitrogenase from *Azotobacter chroococcum*

R.L.Robson[1], P.R.Woodley, R.N.Pau and R.R.Eady

AFRC Institute of Plant Science Research, Nitrogen Fixation Laboratory, University of Sussex, Brighton BN1 9RQ, UK and [1]Department of Biochemistry, University of Georgia, Athens, GA 30602, USA

Communicated by R.Dixon

Structural genes for the VFe-protein ($Ac1^V$) of the vanadium nitrogenase from *Azotobacter chroococcum* were cloned and sequenced. The VFe-protein contains three subunit types with M_r of 53 793 (α), 52 724 (β) and 13 274 (δ). α and β subunits show 18 and 15% sequence identity respectively, with α and β subunits of the MoFe-protein of *A.chroococcum* molybdenum nitrogenase. The genes for the three subunits *vnfD* (α), *vnfG* (δ) and *vnfK* (β) are contiguous and form an operon whose transcription is repressed in response to ammonia. The Fe-protein component of the V-nitrogenase ($Ac2^V$) is the product of *nifH that we have previously cloned and sequenced. This gene was located 2.5 kb upstream of *vnfD*. A deletion in the *vnfD, G* and *K* gene cluster prevents V-dependent nitrogen fixation. A strain defective in both V-nitrogenase and Mo-nitrogenase structural genes showed no residual nitrogen fixing capacity arguing against the presence of a third nitrogen fixation system in this organism.**

Key words: Azotobacter/nitrogen fixation/vanadium nitrogenase/*vnf* genes

Introduction

The soil bacterium *Azotobacter chroococcum* can fix nitrogen by two pathways. One route involves a molybdenum-containing nitrogenase similar to that found in many other nitrogen fixing organisms. The other, recently discovered, route does not require molybdenum but depends on a vanadium-containing nitrogenase. Many of the biochemical properties of the two enzymes are similar yet they are encoded by different sets of structural genes (Robson *et al.*, 1986a).

Mo-nitrogenases from different bacteria are highly homologous and resolve into two essential components known as the Fe-protein and the MoFe-protein. The MoFe-protein is thought to contain the site of nitrogen reduction and the Fe-protein serves as its highly specific electron donor (reviewed in Eady, 1986). In *A.chroococcum*, the MoFe-protein is typical and is a tetramer with a native M_r of 227 000 which contains two atoms of Mo, 23 atoms of Fe and 20 acid-labile S atoms. The protein has an $\alpha_2\beta_2$ subunit structure (Yates and Planque, 1975). All the Mo and a part of the Fe can be extracted from MoFe-proteins in the form of a co-factor called FeMo-co (Shah and Brill, 1977).

The Fe-protein has a native M_r of 64 000 and is a homodimer (Yates and Planque, 1975). Fe-proteins generally contain four Fe atoms and four acid-labile S atoms probably as a [4Fe−4S] cluster The *nifH* gene encodes the subunit of the Fe-protein and the *nifD* and *nifK* genes encode the α and β subunits of the MoFe-protein.

The V-nitrogenase likewise resolves into two component proteins. One is an Fe-protein so similar to the Mo system that it can form a functional complex with the MoFe-protein *in vitro*. The other component appears analogous to the MoFe-protein but whereas it contains comparable amounts of Fe and acid-labile S atoms, Mo is absent and two V atoms are present instead. This protein, the VFe-protein, has a native M_r of 220 000 composed of at least two subunit types with apparent M_r of 50 000 and 55 000. A third small polypeptide reproducibly co-purifies with the VFe-protein though its significance is not known (Eady *et al.*, 1987).

To simplify the description of the components of the nitrogenase systems in *A.chroococcum* we propose to adapt the nitrogenase nomenclature of Eady *et al.* (1972) by incorporating a signifier indicating that the component is normally a part of the V- or Mo-nitrogenases. Thus we will refer to the Fe- and MoFe-proteins of the Mo system as $Ac2^{Mo}$ and $Ac1^{Mo}$ respectively and the Fe- and VFe-proteins of the V system as $Ac2^V$ and $Ac1^V$.

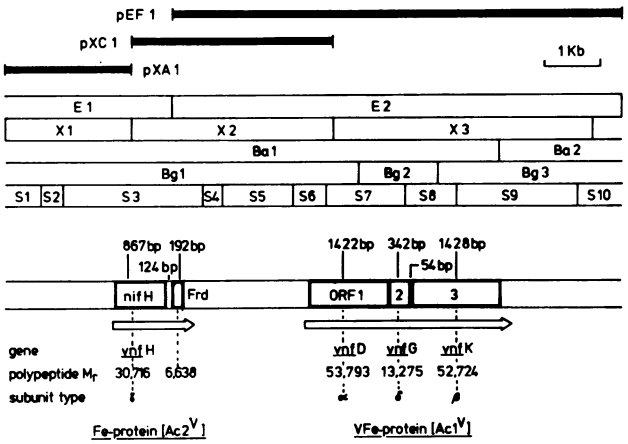

Fig. 1. Cloning and organization of structural genes for V-nitrogenase of *A.chroococcum*. The upper part shows cloned genomic fragments (represented as dark bars) used to construct a physical map for an 11 kb region of the chromosome of *A.chroococcum* carrying the structural genes for the V-nitrogenase. Plasmid pEF1 carrying the 8 kb *Eco*RI fragment (E1) was isolated in this work. Fragments X1 and X2 in plasmids pXA1 and pXC1 were cloned and characterized previously (Robson *et al.*, 1986b). Restriction fragments shown in the map below were generated with the following enzymes: E, *Eco*RI; X, *Xho*I; Ba, *Bam*HI; Bg, *Bgl*II; S, *Sal*I. The relative positions of the ORFs identified in the nucleotide sequence are shown below the restriction map. The direction of transcription for the two operons is shown by the arrows and the lower section shows the V-nitrogenase structural gene/gene product relationships.

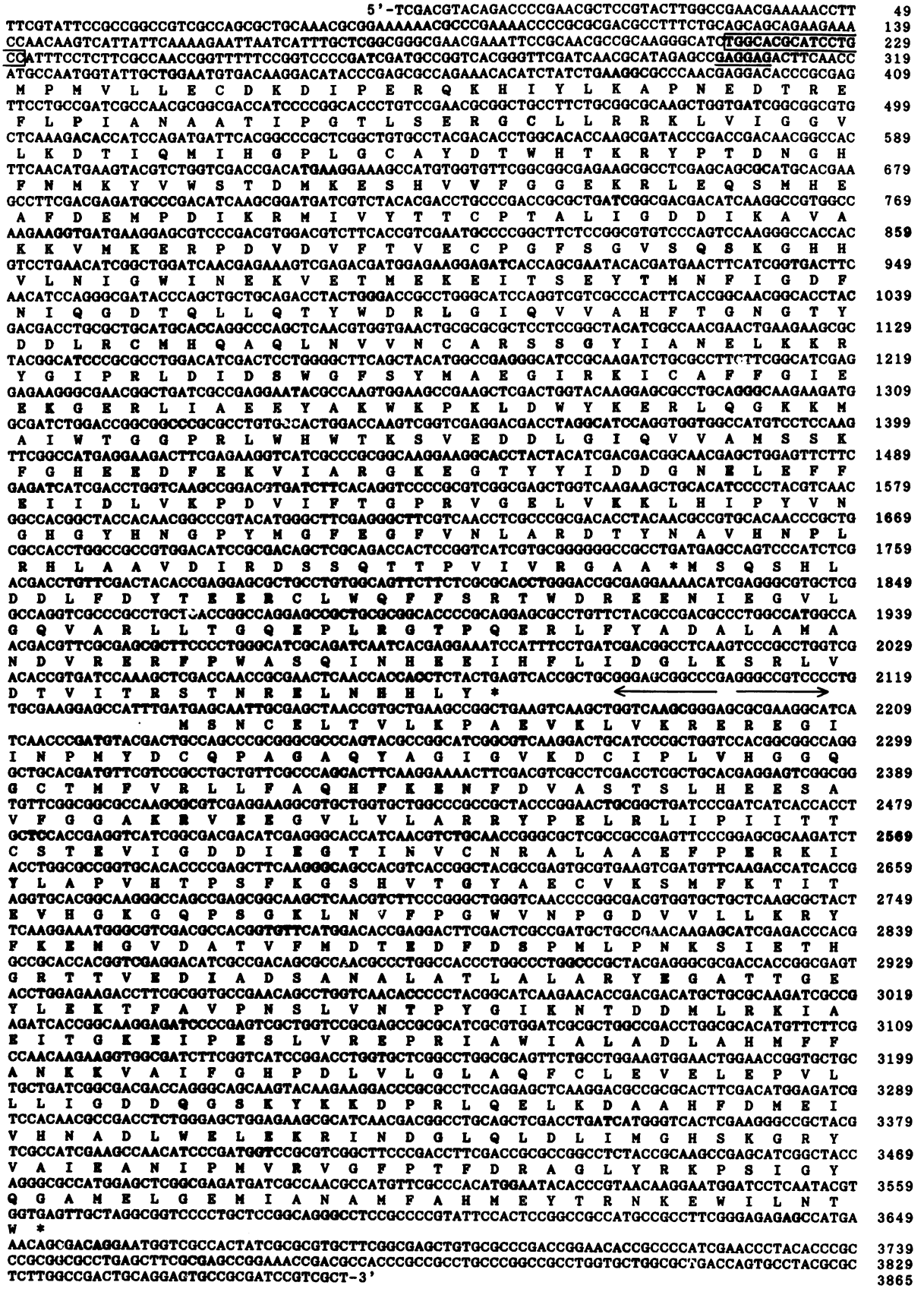

Fig. 2. Nucleotide sequence for the structural genes of the VFe-protein ($Ac1^V$) from *A.chroococcum*. The DNA sequence for 3865 bases extending to the right of the *Sal*I restriction site between fragments S3 and S4 (Figure 1) is shown together with the deduced amino acid sequences (in single letter code) of the three ORFs discussed in the text. Features in the sequence also discussed are marked as follows: the possible −12 to −24 sequence (*ntrA* binding site) is boxed; potential ribosome binding sites are underlined; the inverted repeat between ORF2 and 3 is marked by arrows.

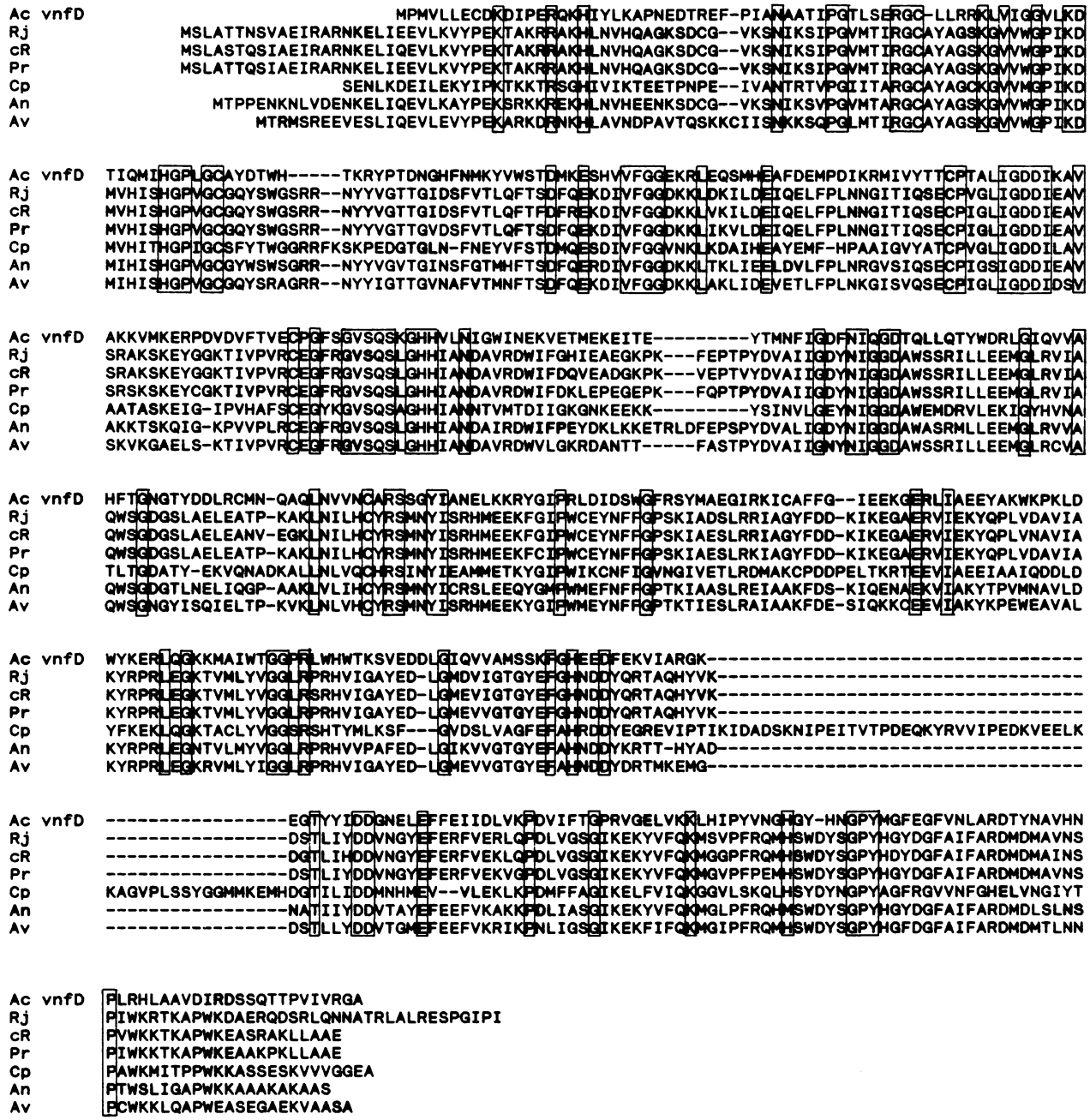

Fig. 3. Comparison of predicted amino acid sequences for ORF1 with α subunits of Mo-nitrogenases. The predicted amino acid sequence of ORF1 (*vnfD*) is compared with α subunits of Mo-nitrogenases from different bacteria. Apparent sequence identities are boxed. α subunits (*nifD* gene products) are from the following organisms: Rj, *Rhizobium japonicum* (Kaluza and Hennecke, 1984); cR, Cowpea *Rhizobium* (Yun and Szalay, 1984); Pr, *Parasponia rhizobium* ANU289 (Weinmann *et al.*, 1984); Cp, *Clostridium pasteurianum* (Hase *et al.*, 1984); An, *Anabaena* 7120 (Lammers and Haselkorn, 1983) but modified by the gene rearrangement (Golden *et al.*, 1985); Av, *Azotobacter vinelandii* (Dean and Brigle, 1985).

The objective of this work was to isolate and characterize the structural genes for the V-nitrogenase. Previously we cloned the structural gene operon (*nifH, nifD* and *nifK*) for the Mo-nitrogenase from *A.chroococcum* (Jones *et al.*, 1984). Their deletion prevents Mo-dependent but not V-dependent nitrogen fixation showing that the V-nitrogenase is genetically distinct (Robson *et al.*, 1986a). We previously suggested that additional *nifH*- and *nifK*-like sequences in the genome of this organism might be associated with a second nitrogenase system (Jones *et al.*, 1984; Robson, 1986). A second *nifH* gene (*nifH**) was cloned and sequenced though its role was not determined (Robson *et al.*, 1986b).

Here we describe the cloning and sequencing of the genes for the VFe-protein. We show that this protein contains three subunit types encoded by three contiguous genes organized into an operon. We also show that *nifH** is the structural gene for the Fe-protein of the V-nitrogenase and that it is closely linked to the genes for the VFe-protein.

Results

Similarity between the two enzyme systems at the biochemical level suggested that the second *nifK*-like sequence in the genome (on an 8 kb *Eco*RI genomic fragment) of

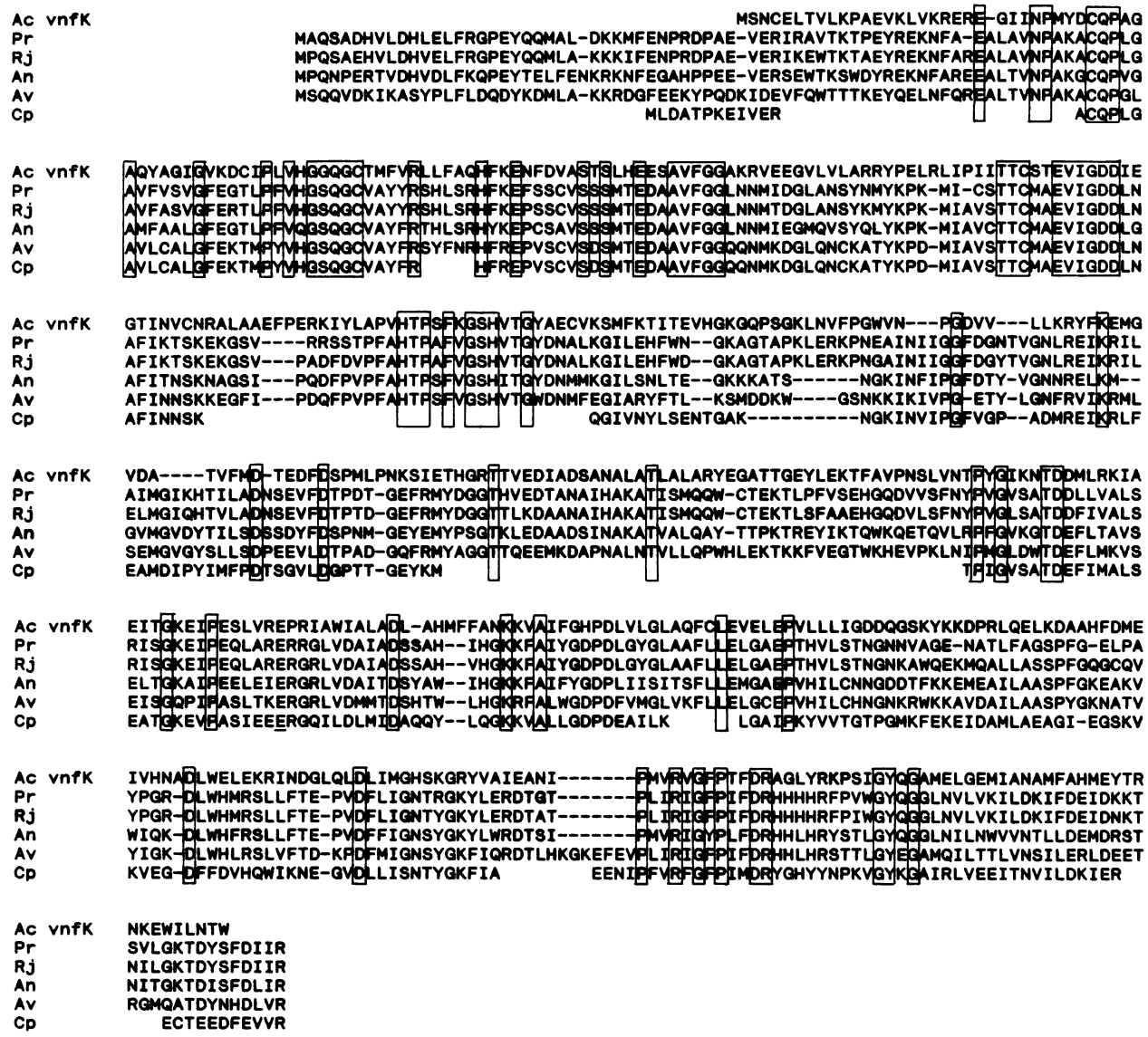

Fig. 4. Comparison of predicted amino acid sequence for ORF3 (*vnfK*) with β subunits of Mo-nitrogenases. The predicted amino acid sequence of ORF3 (*vnfK*) is compared with β subunits of Mo-nitrogenases. Apparent sequence identities are boxed. β subunits (*nifK* gene products) are from the following organisms: Pr, *Parasponia rhizobium* (Weinmann *et al.*, 1984); Rj, *Rhizobium japonicum* (Thöny *et al.*, 1985); An, *Anabaena* 7120 (Mazur and Chui, 1982); Av, *Azotobacter vinelandii* (Dean and Brigle, 1985); Cp, *Clostridium pasteurianum* cysteinyl-containing tryptic peptides (Hase *et al.*, 1984).

A.chroococcum might correspond to a gene for $Ac1^V$ (Robson, 1986). We cloned this reiterated sequence by first constructing a gene library of *Eco*RI genomic fragments in the size range 7−9 kb in the plasmid pEMBL18+. An *A.chroococcum nifK* specific DNA fragment, a 1.4 kb *Hin*dIII fragment (Robson, 1986), was used to probe the gene library and hybridizing clones contained apparently similar plasmids (represented by pEF1) containing 8 kb inserts. After mapping the restriction sites within the insert we localized the *nifK* hybridizing sequence on a 990 bp *Sal*I fragment (S8) within the insert DNA (Figure 1).

The nucleotide sequence for almost 4 kb around the region of homology to *nifK* was determined (Figure 2). Three open reading frames (ORFs) with codon preferences similar to other *Azotobacter* genes were identified in the sequence. The position of the ORFs with respect to the physical map is shown in Figure 1. ORF1, 2 and 3 potentially encode polypeptides with M_r of 53 793, 13 275 and 52 724, respectively. ORF1 and 3, though not ORF2, are preceded by good potential ribosome binding sites. The translation initiation codon of ORF2 overlaps the translation termination codon of ORF1 and suggests that the translation of ORF1 and ORF2 could be tightly coupled.

The deduced amino acid sequences of ORF1 and ORF3 show homology to *nifD* and *nifK* gene products of Mo-nitrogenases (Figures 3 and 4). ORF2 does not appear to be closely related to any protein previously described.

ORF1, 2 and 3 encode the structural genes for $Ac1^V$ as shown by comparing the predicted amino acid sequences of the ORFs to the NH_2-terminal amino acid sequences of the polypeptides isolated from the protein. The α and β polypeptides are the products of ORF1 and ORF3. The third small polypeptide which co-purifies with $Ac1^V$ corresponds to ORF2 (Figure 5). The ORF2 (δ) subunit is likely to be a third subunit type in the VFe-protein because it co-purifies reproducibly with $Ac1^V$ and appears to be present in

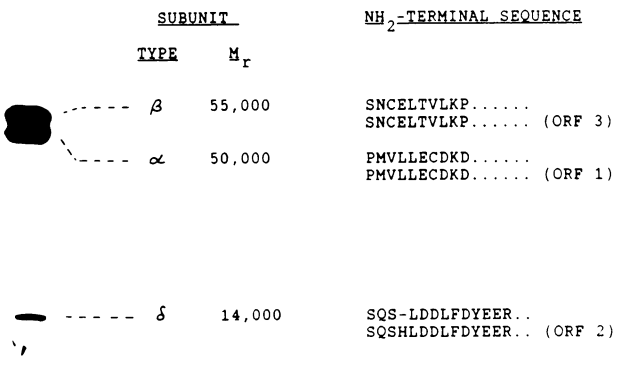

Fig. 5. Assignment of polypeptides of the VFe-protein ($Ac1^V$) as the gene products of ORF1, 2 and 3. The polypeptides of purified VFe-protein were blotted onto an Immobilon filter (Millipore Corp.) according to the manufacturers' instructions. The Coomassie Blue stained blot is shown on the left. Mol. wts were interpolated from a standard curve obtained from protein standards (Biorad, low mol. wt markers). For the determination of NH_2-terminal amino acid sequences the 55 and 50 kd polypeptides were electroeluted from Coomassie Blue stained gels and the 14 kd protein electroblotted onto trifluoroacetic acid-activated glass paper.

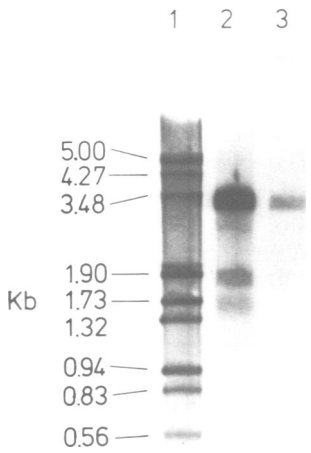

Fig. 6. Transcription of VFe-protein ($Ac1^V$) structural genes. RNA prepared from *A.chroococcum* MCD1155 was fractionated by electrophoresis on a denaturing agarose gel and a Northern blot prepared. The blot was probed with an internal VFe-protein structural gene probe (fragment Bg2, Figure 1). RNA isolated from a culture growing under nitrogen fixing conditions was loaded in **lane 2**. RNA isolated from the same culture 5 h following the addition of NH_4Cl to 10 mM was loaded in **lane 3**. Size markers loaded in **lane 1** were ^{32}P-end-labelled denatured DNA fragments generated by *Hin*dIII digestion of bacteriophage lambda.

equivalent proportions to the other subunits based on laser scanning of Coomassie Blue stained gels and NH_2-terminal amino acid ratio analysis of the native protein. Furthermore, ORF2 should be expressed at equimolar levels with ORF1 and ORF3 considering the organization of the three genes.

The three ORFs are probably co-transcribed. Northern blots of RNA isolated from cultures of *A.chroococcum* fixing

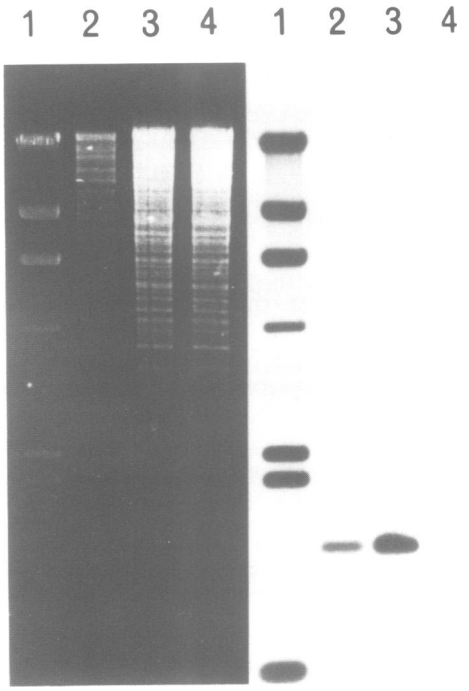

Fig. 7. Chromosomal deletion of VFe-protein structural genes. A deletion in the *A.chroococcum* MCD1155 chromosome of the 1.4 kb *Bgl*II fragment internal to the VFe-protein structural genes was made by a gene replacement technique. The *Bgl*II fragment in pEF1 (Bg2, Figure 1) was first replaced by a spectinomycin gene resistant cassette *in vitro*. The mutagenized insert DNA was recloned into the suicide vector pCU101 which was introduced into MCD1155 by conjugation from *E.coli*. Strain MCD1301 was selected originally as a spectinomycin-resistant clone arising after the conjugation. Genomic DNA was prepared from MCD1, MCD1155 and MCD1301, digested with *Bgl*II, Southern blotted and probed with radioactively labelled 1.4 kb *Bgl*II fragment. The expected fragment was detected in MCD1 (**lane 2**) and MCD1155 (**lane 3**) but was missing from MCD1301 (**lane 4**). The ethidium bromide stained gel used to make the blot is shown on the left and the hybridization data on the right. **Lane 1** contains ^{32}P-end-labelled *Hin*dIII fragments of bacteriophage lambda.

nitrogen with V were probed with an internal structural gene probe (fragment Bg2, Figure 1). The largest and most abundant species was 3.4 kb in length (Figure 6, lane 2) corresponding in size to a transcript originating just 5′ to ORF1 and terminating just 3′ to ORF3. Less abundant species at 1.9 and 1.7 kb could arise from premature termination or from processing of the largest transcript. An inverted repeat in the intergenic region between ORF2 and 3 (Figure 2) may be a potential site for transcript termination or processing. All the transcripts were repressed in response to ammonia (Figure 6, lane 3) as expected since the V-nitrogenase activity cannot be detected when the organism is supplied with this nitrogen source.

We showed that the region containing ORF1, 2 and 3 was required for V-dependent nitrogen fixation by constructing a chromosomal deletion of the 1.4 kb *Bgl*II fragment (Bg2; Figure 1) which spans all of ORF2 and the C-terminal and NH_2-terminal portions of ORF1 and 3, respectively. The *Bgl*II fragment in pEF1 was first replaced with a spectinomycin resistance gene cassette *in vitro*. The mutagenized *Eco*RI fragment was recloned into the conjugatable suicide vector pCU101 which was then used as suicide vehicle to introduce the mutation into *A.chroococcum*. The deletion was constructed in *A.chroococcum* MCD1155 which lacks the structural genes for the Mo-nitrogenase and can therefore

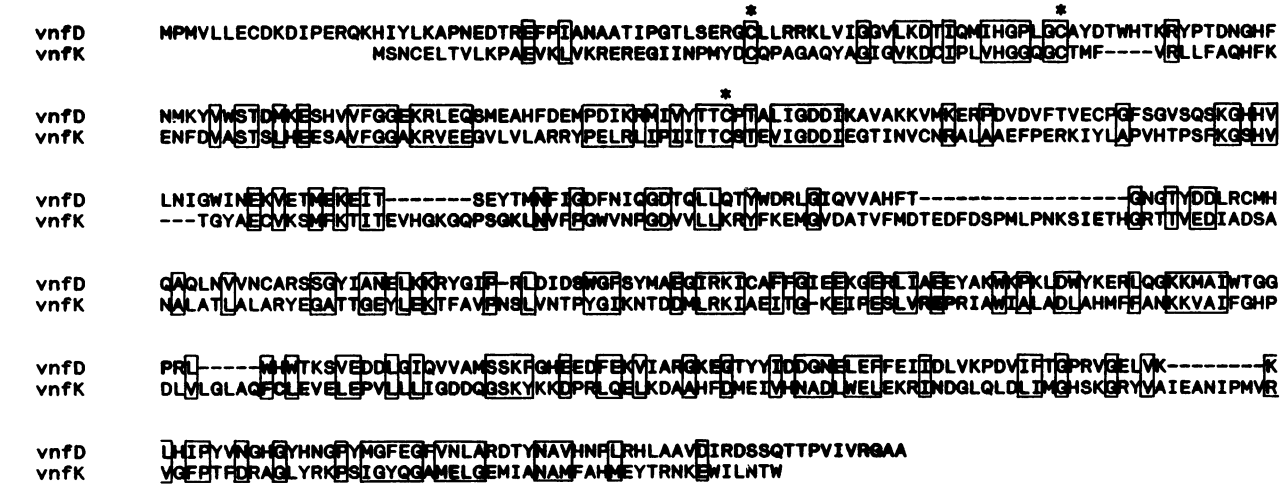

Fig. 8. Comparison of the predicted amino acid sequences of α and β subunits of the VFe-protein ($Ac1^V$) from *A.chroococcum*. Predicted amino acid sequences for the α subunit (*vnfD* gene product) and the β subunit (*vnfK* gene product) of the VFe-protein were aligned using the University of Wisconsin Genetics computer group program GAP. The boxed areas mark sequence identities or conservative replacements. The asterisks mark conserved cysteinyl residues.

only fix nitrogen when provided with V. Deletion mutants were first selected as spectinomycin resistant colonies since the acquisition of stable resistance to spectinomycin only occurs after the recombinational exchange of the marker into the chromosome. The deletion was confirmed in the resultant strain (MCD1301) by hybridization using the Bg2 fragment as a probe (Figure 4). MCD1301 was completely incapable of growth in N_2 irrespective of the presence or absence of Mo or V although growth was normal when the strain was supplied with a source of combined nitrogen such as ammonium (Figure 7).

We previously cloned the second *nifH*-like sequence (*nifH**) from *A.chroococcum* and suggested that it might be part of a second nitrogenase system in this organism (Robson *et al.*, 1986b). In order to demonstrate whether this sequence encodes the Fe-protein of the V-nitrogenase we determined the NH_2-terminal sequence of the $Ac2^V$ which matched exactly that predicted from the DNA sequence of *nifH**.

The restriction map for the 8 kb *Eco*RI fragment in pEF1 overlaps that for one of the *Xho*I fragments (X2; see Figure 1) we had previously isolated in the cloning of *nifH**. The overlap was confirmed by hybridization and the combined map shows that the structural genes for the two V-nitrogenase components are separated by ~2.5 kb.

Discussion

We propose the genotype *vnf* (V-dependent nitrogen fixation) for genes uniquely required for the V-nitrogenase system to distinguish them from *nif* genes as originally defined for the Mo-nitrogenase in *Klebsiella pneumoniae*. We assign the α and β subunits of $Ac1^V$ as the products of *vnfD* and *vnfK* respectively and the small subunit (δ) as the product of *vnfG*. *NifH**, which encodes the Fe-protein component of the V-nitrogenase, we redesignated as *vnfH*. At present it is not known whether the putative ferredoxin gene which is co-transcribed with *vnfH* is required for expression of the V system though it seems likely that it is a specific electron donor to the V enzyme (Robson *et al.*, 1986b).

The V- and Mo-nitrogenases show significant homology. This is greatest in the case of the Fe-proteins ($Ac2^V$ and $Ac2^{Mo}$) which are 89% homologous at the amino acid level (Robson *et al.*, 1986b). The α and β subunits of $Ac1^V$ and MoFe-proteins are much less homologous. They are both significantly shorter at their NH_2-terminal ends than all examples of their *nif* gene homologues; 85 residues (18%) of *vnfD* are conserved in *nifD* gene products and 71 residues (15%) of the *vnfK* are conserved in *nifK* gene products. This includes five cysteinyl residues in the α subunits and three in the β subunits which probably provide ligands to comparable metal centres in these proteins. Biochemical evidence suggests that the two protein types probably contain similar numbers and types of clusters. First estimates of the Fe content of $Ac1^V$ fall within the range reported for MoFe-proteins (Eady *et al.*, 1987). Furthermore, it is known that V in $Ac1^V$ is present in an environment similar to that of Mo in MoFe-proteins (Arber *et al.*, 1987) such that it can be extracted in a co-factor (FeVa-co) analogous to FeMo-co (Smith *et al.*, 1988). Glycine residues which might be important in the secondary structure constitute 22 of the 85 conserved residues amongst *vnf/nifD* gene products and 13 of the 71 conserved amongst *vnf/nifK* gene products.

In view of the similarity between the α and β subunits of the VFe- and MoFe-proteins the finding of a third subunit type in the VFe-protein is unexpected. The role of this subunit is not known at present but it would be surprising if it were essential for activity at least *in vitro*. The α and β polypeptides of the MoFe-protein do not contain a region homologous to the *vnfG* product indicating that *vnfG* has not been incorporated into *nifDK*. However, small proteins which co-purify with Mo-nitrogenases are known. The Mo-nitrogenase in *Azotobacter* can be isolated as a complex containing the Fe-protein, the MoFe-protein and a third Fe−S protein having a M_r of 14 500. This protein, though not essential for activity, protects the enzyme from damage by O_2 ('conformational protection') and may also regulate activity *in vitro* (Scherings *et al.*, 1977; Robson, 1979). The *vnfG* gene product may play a similar role but, whilst its size (13 275) is comparable, it contains only one cysteine residue and is therefore unlikely to be an Fe−S protein. Also, in crude extracts the V-nitrogenase, unlike the Mo-enzyme, is highly sensitive to inactivation by O_2 and

therefore does not appear 'conformationally protected' (Eady *et al.*, 1987).

It has been suggested that the α and β subunits of the MoFe-protein evolved from a common ancestral gene since they show sequence homology (Lammers and Haselkorn, 1983; Thöny *et al.*, 1985) and overall structural similarity (Yamane *et al.*, 1982). The similarity is more marked in the case of the $Ac1^V$ subunits (Figure 8). In this, 89 residues are conserved and the two polypeptides differ in size by only two residues. This suggests that if the genes for these subunits have arisen from a common ancestor, their divergence may have occurred more recently than those in the Mo-nitrogenase. We conclude that the two nitrogenases have probably arisen from a common ancestral gene.

Structural genes for $Ac2^V$ and $Ac1^V$ are organized into separate operons, whereas their counterparts from the Mo-nitrogenase of *A.chroococcum* form a single operon (Jones *et al.*, 1984). This indicates a potential for independent modulation of the levels of the V-nitrogenase components.

Transcription of *nif* genes in Gram-negative organisms is well understood and requires the products of the *ntrA* gene (which encodes the alternative sigma factor σ^{54}, Hirschman *et al.*, 1985; Hunt and Magasanik, 1985) and that of *nifA*, which encodes a *nif* specific positive regulator (see Gussin *et al.*, 1987 for review). Two conserved sequences have been recognized in *nif* promoters. The sequence 5′-CTGGPyAPyPuNNNNTTGCA-3′ lies 26 bp upstream of the transcription initiation site (Beynon *et al.*, 1983) and is recognized by core RNA polymerase containing σ^{54} (Gussin *et al.*, 1986). The consensus upstream activator sequence (UAS) 5′-TGT-N_{10}-ACA-3′, probably a binding site for NIFA, usually lies between 120 and 140 bp further upstream (Buck *et al.*, 1986). Though the sequence 5′-CTGGCACGCATCCTGCA-3′ (Figure 2; boxed) located 104 bp upstream of ORF1 conforms well to a consensus −24, −12 sequence we can find no evidence for a UAS immediately upstream of ORF1. This suggests that expression of this operon requires σ^{54} and that if it requires an activator analogous to NIFA, this activator has a different DNA recognition sequence. Comparison of the promoter regions of *vnfH* (Robson *et al.*, 1986b) and *vnfD* do not show significant homology other than the potential binding site for a σ^{54}-like transcription factor (Kustu *et al.*, 1986).

Inactivation of both sets of nitrogenase structural genes by deletion renders our strain of *A.chroococcum* incapable of nitrogen fixation and seems to rule out the existence of a third system in this organism. This situation may be contrasted with that in the closely related organism *A. vinelandii* which contains three different nitrogenase systems. Two correspond with the Mo- and V-nitrogenases in *A. chroococcum* and the third is a system requiring neither Mo nor V (Chisnell *et al.*, 1988; Pau *et al.*, 1989).

Materials and methods

Bacterial strains, plasmids, bacteriophage and media

A.chroococcum MCD1 (Robson *et al.*, 1984) and its derivative MCD1155 (deleted for *nifHDK*) (Robson, 1986) were routinely grown and maintained on RM medium (Robson *et al.*, 1984) or, when growth on N_2 was required, a simple defined medium (Nil medium, Robson, 1986) supplemented with either V or Mo (to 0.1 μM) was used. Spectinomycin and kanamycin were added to media at 2 and 0.25 μg/ml when required. *Escherichia coli* strains HB101 (Boyer and Roulland-Dussoix, 1969), 5K (Jones *et al.*, 1984) and 71/18 (Messing *et al.*, 1977) were routinely grown in Luria Bertani agar or broth. Carbenicillin, spectinomycin and kanamycin were added when required at 50, 20 and 25 μg/ml, respectively. Bacteriophage M13K07 used to raise single-stranded templates for DNA sequencing was obtained from Pharmacia and propagated according to the manufacturers' instructions. Cloning vector pEMBL18+ was described by Dente *et al.* (1983) and pTZ18R and pTZ19R were described by Mead *et al.* (1986). pCU101 was described by Thatte and Tyler (1983) and its use as a suicide vector for *A.chroococcum* was demonstrated earlier (Robson, 1986). The plasmid pHP45Ω was the source of the spectinomycin resistance cassette (Prentki and Krisch, 1984).

Genetic techniques

The procedure for construction of chromosomal deletions in *A.chroococcum* is decribed in the text and in an earlier report (Robson, 1986). Conjugations between *E.coli* and *A.chroococcum* were performed as described by Jones *et al.* (1984) and transformation of *E.coli* was done as described in Kushner (1987).

Molecular biology techniques

Preparation of *Azotobacter* chromosomal DNA was as in Robson *et al.* (1984) and RNA was prepared as described in Krol *et al.* (1982). General recombinant DNA techniques and radioactive labelling of DNA were as given in Jones *et al.* (1984). DNA sequencing was done by the chain termination method (Sanger *et al.*, 1977) using [^{35}S]dATP (Biggin *et al.*, 1983) with templates produced from inserts cloned in pEMBL19+ or pTZ18R. Both strands were sequenced using an ordered deletion strategy. Gaps were filled and uncertainties resolved using synthetic oligonucleotide primers. The sequence was assembled using the DB system of Staden (1982). The likely validity of ORFs was tested using the computer program CODONPREFERENCE from the University of Wisconsin Computer Group (Deveraux *et al.*, 1985) using data for codon usage compiled from other *Azotobacter* genes. Protein comparisons were made using the UWCG program GAP.

Protein and polypeptide analysis

$Ac1^V$ and $Ac2^V$ were purified as described previously (Robson *et al.*, 1986a; Eady *et al.*, 1987). Polypeptides were separated by SDS−PAGE, stained with Coomassie Blue and extracted by electroelution. The δ subunit was recovered by electroblotting into trifluoroacetic acid-activated glass paper (Aebersold *et al.*, 1986). NH_2-terminal analysis was performed using an Applied Biosystems 470A gas phase sequenator equipped with a 120A on-line analyser.

Acknowledgements

We would like to thank Martin Buck for making oligonucleotides and P.Barker of the AFRC Institute of Animal Physiology and Genetics Research, Cambridge Research Station, Babraham, Cambridge for the gas phase amino acid sequence analysis. We would like to thank Barry Smith for helpful discussions, and Beryl Scutt and Janet Walker for typing the manuscript. We acknowledge the support of a NATO grant no. 04-0050-88, for international collaboration in research.

References

Aebersold,R.H., Teplow,D.B., Hood,L.E. and Kent,S.B.H. (1986) *J. Biol. Chem.*, **261**, 4229−4238.

Arber,J.M., Dobson,R.R., Eady,R.R., Stevens,P., Hasnain,S.S., Garner,C.D. and Smith,B.E. (1987) *Nature*, **325**, 372−374.

Beynon,J., Cannon,M., Buchanan-Wollaston,V. and Cannon,F. (1983) *Cell*, **34**, 665−671.

Biggin,M.D., Gibson,J.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963−3965.

Boyer,H.B. and Roulland-Dussoix,D. (1969) *J. Mol. Biol.*, **41**, 459−472.

Buck,M., Miller,S., Drummond,M. and Dixon,R. (1986) *Nature*, **320**, 374−378.

Chisnell,J.R., Premakumar,R. and Bishop,P.E. (1988) *J. Bacteriol.*, **170**, 27−33.

Dean,D.R. and Brigle,K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5720−5723.

Dente,L., Cesareni,G. and Cortese,R. (1983) *Nucleic Acids Res.*, **11**, 1645−1655.

Deveraux,J., Haeberli,P. and Smithies,O. (1985) *Nucleic Acids Res.*, **12**, 387−395.

Eady,R.R. (1986) In Broughton,W.J. and Pühler,A. (eds), *Nitrogen Fixation*. Clarendon Press, Oxford, pp. 1−49.

Eady,R.R., Smith,B.E., Cook,K.A. and Postgate,J.R. (1972) *Biochem. J.*, **128**, 655–675.
Eady,R.R., Robson,R.L., Richardson,T.H., Miller,R.W. and Hawkins,M. (1987) *Biochem. J.*, **244**, 197–207.
Golden,J.W., Robinson,S.J. and Haselkorn,R. (1985) *Nature*, **314**, 419–423.
Gussin,G.N., Ronson,C.W. and Ausubel,F.M. (1986) *Annu. Rev. Genet.*, **20**, 567–591.
Hase,T., Wakabayashi,S., Nakano,T., Zumft,W.F. and Matsubara,H. (1984) *FEBS Lett.*, **166**, 39–43.
Hirschman,J., Wong,P.K., Sei,K., Keener,J. and Kustu,S. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7525–7529.
Hunt,T. and Magasanik,B. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8453–8457.
Jones,R., Woodley,P.R. and Robson,R.L. (1983) *Mol. Gen. Genet.*, **197**, 318–327.
Kaluza,K. and Hennecke,H. (1984) *Mol. Gen. Genet.*, **196**, 35–42.
Krol,A.J.M., Hontelez,J.G.J., Roozendaal,B. and van Kammen,A. (1982) *Nucleic Acids Res.*, **10**, 4147–4157.
Kushner,S.R. (1987) In Boyer,H.W. and Nicosia,S. (eds), *Genetic Engineering*. Elsevier/North Holland Biomedical Press, Amsterdam, pp. 17–23.
Kustu,S., Sei,K. and Keener,J. (1986) In Booth,I. and Higgins,C.F. (eds), *Regulation of Gene Expression—25 Years On*. Cambridge University Press, Cambridge, pp. 139–154.
Lammers,P.J. and Haselkorn,R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4723–4727.
Mazur,B.J. and Chui,C.F. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6782–6786.
Mead,D.A., Szczesna-Skorupa,E. and Kemper,B. (1986) *Protein Eng.*, **1**, 67–74.
Messing,J., Gronenborn,B., Müller-Hill,B. and Hofschneider,P.H. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 3642–3646.
Pau,R.N., Mitchenall,L.A. and Robson,R.L. (1989) *J. Bacteriol.*, **171**, 124–129.
Prentki,P. and Krisch,H.H. (1984) *Gene*, **29**, 303–313.
Robson,R.L. (1979) *Biochem. J.*, **181**, 569–575.
Robson,R.L. (1986) *Arch. Microbiol.*, **146**, 74–79.
Robson,R.L., Chesshyre,J.A., Wheeler,C., Jones,R., Woodley,P.R. and Postgate,J.R. (1984) *J. Gen. Microbiol.*, **130**, 1603–1612.
Robson,R.L., Eady,R.R., Richardson,T.H., Miller,R.W., Hawkins,M. and Postgate,J.R. (1986a) *Nature*, **32**, 388–380.
Robson,R.L., Woodley,P.R. and Jones,R. (1986b) *EMBO J.*, **5**, 1159–1163.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Scherings,G., Haaker,H. and Veeger,C. (1977) *Eur. J. Biochem.*, **77**, 621–630.
Shah,V.K. and Brill,W.J. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 3249–3253.
Smith,B.E., Eady,R.R., Lowe,D.J. and Gormal,C. (1988) *Biochem. J.*, **217**, 299–302.
Standen,R. (1982) *Nucleic Acids Res.*, **10**, 4731–4751.
Thatte,V. and Tyler,V.N. (1983) *Gene*, **21**, 227–236.
Thöny,B., Kaluza,F. and Hennecke,H. (1985) *Mol. Gen. Genet.*, **198**, 441–448.
Weinmann,J.J., Felows,F.F., Gresshoff,P.M., Shine,J. and Scott,K.F. (1984) *Nucleic Acids Res.*, **12**, 8329–8344.
Yamane,T., Weininger,M.S., Mortenson,L.E. and Rossman,M.G. (1982) *J. Biol. Chem.*, **275**, 1221–1223.
Yates,M.G. and Planque,K. (1975) *Eur. J. Biochem.*, **607**, 467–476.
Yun,A.C. and Szalay,A.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7358–7362.

Received on December 2, 1988