PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Genetics and Regulation of nif and Related Genes in Klebsiella pneumoniae [and Discussion]

R. A. Dixon, S. Austin, M. Buck, M. Drummond, S. Hill, A. Holtel, S. MacFarlane, M. Merrick, S. Minchin, F. C. Cannon and R. Haselkorn

Phil. Trans. R. Soc. Lond. B 1987 **317**, 147-158
doi: 10.1098/rstb.1987.0053

Phil. Trans. R. Soc. Lond. B **317**, 147–158 (1987)
Printed in Great Britain

Genetics and regulation of *nif* and related genes in *Klebsiella pneumoniae*

By R. A. Dixon, S. Austin, M. Buck, M. Drummond, S. Hill, A. Holtel, S. MacFarlane, M. Merrick and S. Minchin

AFRC Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, East Sussex, U.K.

Seventeen genes specifically required for nitrogen fixation are clustered on the chromosome of *Klebsiella pneumoniae* and form a complex regulon that is organized into eight transcriptional units. The *nif* promoters are representative of a new class of promoter, the members of which lack the consensus sequences normally found in prokaryotic promoters. *nif* gene transcription is positively controlled and requires: (1) the *ntrA* gene product, which replaces the *rpoD*-encoded sigma subunit of RNA polymerase to allow recognition of *nif* promoter sequences; and (2) the product of either the nitrogen regulation gene *ntrC* or the specific *nif* regulatory gene, *nifA*, which are both transcriptional activators. Most *nif* promoters require an upstream activator sequence (UAS) for *nifA*-mediated activation. The UAS acts independently of orientation and can function when placed 2 kilobases upstream from the transcription start site. Current evidence suggests that activation requires an interaction between proteins bound at the UAS and at the downstream *nif* promoter consensus, possibly via a loop in the DNA structure.

Transcription of *nif* is modulated by the *ntrB* and *nifL* gene products. Both proteins can 'sense' environmental changes: *ntrB* prevents activation by *ntrC* in response to excess nitrogen whereas *nifL* prevents activation by *nifA* in response to fixed nitrogen and oxygen. The C-terminal end of *ntrB* shows clear homology at the amino acid level with a number of diverse control proteins involved in regulation or sensory transduction. Each member of this family interacts with another protein component showing homology to the N-terminal sequence of *ntrC*, but not to *nifA*. The significance of these protein homologies is discussed.

1. Introduction

From a genetic standpoint *Klebsiella pneumoniae* is relatively distinct among diazotrophs in that genes specifically required for the synthesis and activity of nitrogenase are clustered at a single location on the chromosome. The seventeen *nif* genes form a complex regulon that is linked to the histidine operon and is organized into eight transcriptional units (for reviews see Roberts & Brill 1981; Drummond 1984; Dixon 1984*a*). In addition to the seventeen designated genes, an open-reading frame (ORF) close to the structural genes has been identified (Shen *et al.* 1983). It is possible that further ORFs will be characterized once the nucleotide sequence of the 23 kb *nif* gene cluster is complete. The functions of the *nif* gene products can be approximately divided into five categories (Cannon *et al.* 1985 and the reviews listed above): (*a*) synthesis of molybdenum iron protein (Kp1) and FeMoco (*nifB*, *nifQ*, *nifE*, *nifN*, *nifV*, *nifK*, and *nifD*); (*b*) synthesis of iron protein (Kp2) (*nifH* and *nifM*); (*c*) electron transport to nitrogenase (*nifF* and *nifJ*); (*d*) regulation of *nif* transcription (*nifL* and *nifA*); and (*e*) unknown functions (*nifS*, *nifU*, *nifX*, *nifY*, and the ORF referred to above).

In recent years the complexities of regulation of *nif* transcription in *K. pneumoniae*, in response to external sources of fixed nitrogen and oxygen, have begun to be unravelled. It is now evident

that *nif* transcription is regulated by a cascade system that involves a general regulatory mechanism, mediated by genes outside the *nif* cluster, as well as a specific control mediated by the *nifL* and *nifA* genes referred to above. Regulation of *nif* in response to large levels of fixed nitrogen is mediated by a group of genes termed *ntr*, whose involvement in the control of many operons in response to nitrogen sources has been elucidated in a number of laboratories (see Merrick 1982; Magasanik 1982). The primary components of the *ntr* system are the *ntrA*, *ntrB* and *ntrC* genes (which are called *glnF*, *glnL* and *glnG* respectively in some publications). The *ntrA* gene product has a pleiotropic role and is required for the expression of many nitrogen-regulated operons. The *ntrB* and *ntrC* genes are unlinked to *ntrA* and form a regulon with *glnA*, the structural gene for glutamine synthetase (figure 1). The *ntrC* gene product (NTRC) is a bifunctional regulatory protein that can act either as an activator or as a repressor of transcription; the activity of this protein is modulated by the *ntrB* gene product (NTRB) in response to the nitrogen source. The *ntrB* and *ntrC* genes regulate transcription of operons that are subject to nitrogen control, including the *nif* genes, as well as autogenously regulating their own transcription and that of *glnA*. Under nitrogen-limiting conditions NTRB and NTRC activate transcription of *nif* genes, whereas under conditions of nitrogen excess NTRB prevents NTRC from activating transcription and hence *nif* is not expressed.

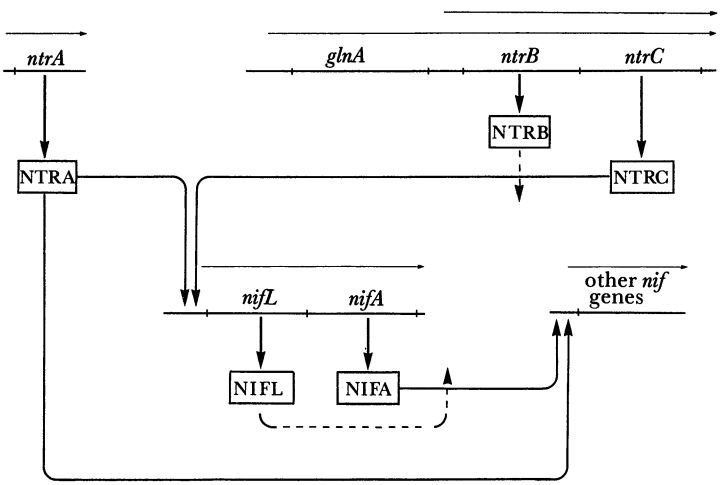

FIGURE 1. Model for *nif* regulation in *K. pneumoniae*. Boxed characters represent gene products, with arrows representing regulatory functions. The thin horizontal arrows represent transcripts.

The target for *ntr*-mediated regulation of *nif* transcription is the promoter of the regulatory *nifLA* operon (figure 1). The *nifA* gene product (NIFA), is a positive activator protein required for transcription at the remaining *nif* promoters. Both *nifA*- and *ntrC*-mediated activation require a functional *ntrA* gene and the *nifA* and *ntrC* products show some functional homology (for reviews see Ausubel (1984) and Dixon (1984*a*)). The *nifL* gene product (NIFL) has an analogous role to that of NTRB and apparently modulates the activity of NIFA in response to both fixed nitrogen and oxygen (Hill *et al.* 1981; Merrick *et al.* 1982). Transcription of *nif* genes other than *nifL* and *nifA* is therefore prevented when levels of oxygen that will inactivate nitrogenase are present or when alternative sources of fixed nitrogen are available.

In this paper we shall highlight recent advances in our understanding of the mechanism of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

transcriptional control of nitrogen fixation genes. These findings have important implications for *nif* regulation in other diazotrophs as well as elucidating a novel mode of gene regulation in prokaryotes.

2. *nif* Promoter structure

Comparative sequence analysis of *K. pneumoniae nif* promoters revealed that they were atypical, lacking the consensus −35 and −10 elements found in most prokaryotic promoters. Instead, these promoters contain conserved features upstream of the transcription initiation site with the consensus CTGGCAC around −24, and TTGCA around −12 respectively (Beynon *et al.* 1983). The two consensus elements contain an invariant GG dinucleotide (at −24) and an invariant GC dinucleotide (at −12). The spacing of ten base pairs between these invariant dinucleotides is also conserved and is critical for promoter function, because single base pair deletions within the spacer region perturb promoter activity (Buck 1986). Other promoters in enteric bacteria that are dependent on *ntrA* for transcriptional activity show considerably homology to the *nif* promoter consensus (Dixon 1984*b*); these include the *glnA* (RNA1) promoter (Dixon 1984*b*; Reitzer & Magasanik 1985; Hirshman *et al.* 1985) and the *arg*Tr promoter from *Salmonella typhimurium* (Ames & Nikaido 1985). Two *Pseudomonas* promoters, *xylCAB* (OP1) (Inouye *et al.* 1984) and CPG2 (Minton & Clarke 1985) also show homology to the *nif* promoter consensus. In the case of *xylCAB* (OP1), transcription from this promoter in *E. coli* has been shown to be dependent on *ntrA* (Dixon 1986).

Mutations have been isolated in both invariant and non-conserved nucleotides in the *nifL* and *nifH* promoters (Brown & Ausubel 1984; Buck *et al.* 1985 ; Ow *et al.* 1985; Khan *et al.* 1986 Kaluza *et al.* 1985). All *nif* promoters are activated by *nifA* but most (e.g. the *nifH* promoter) are only weakly activated by *ntrC*. The *nifL* promoter, however, can be efficiently activated by either *ntrC* or *nifA* (Drummond *et al.* 1983; Ow & Ausubel 1983). Point mutations in each of the invariant dinucleotides in the *nifL* promoter at −25, −24, −13, and −12 give a strong down phenotype with respect to both *ntrC*- and *nifA*-mediated activation whereas mutations in semi-conserved nucleotides have a much weaker down phenotype (Buck *et al.* 1985; Khan *et al.* 1986). Comparable base changes in the *nifH* promoter also give a down phenotype when *ntrC*-mediated activation is examined; however, *nifA*-mediated activation of the *nifH* promoter is far less sensitive to base changes in conserved residues (Buck *et al.* 1985). This suggests that additional sequence elements present in the *nifH* promoter are involved in *nifA*-mediated activation and that these sequences are absent in the *nifL* promoter (see below). C to T transitions in the *nifH* promoter at −17 and −15 increased activation by *ntrC* (Ow *et al.* 1985) as did a G to A mutation at −18 in the *nifL* promoter (Khan *et al.* 1986), consistent with the suggestion that some activator specificity is conferred by nucleotides in the −18 to −14 regions of these promoters (Ow *et al.* 1983).

3. Role of the *ntrA* product in promoter recognition

All promoters examined so far that contain the canonical −24 and −12 consensus sequences are dependent on the *ntrA* gene product for functional activity. The unique structure of these promoters indicated a role for *ntrA* in modifying the transcriptional specificity of RNA polymerase and it was suggested that the *ntrA* product might be an alternative sigma factor that

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

could allow recognition of such promoters (de Bruijn & Ausubel 1983). The *ntrA* gene is transcribed constitutively and is not subject to nitrogen control (de Bruijn & Ausubel 1983; Castano & Bastarrachea 1984). Transcription from *nif* promoters is increased in strains carrying a multicopy *ntrA* plasmid, and is decreased in strains carrying multiple copies of *rpoD* that encodes the 'standard' sigma factor σ^{70} (Merrick & Stewart 1985). This suggests that NTRA protein is limiting in cells and competes with σ^{70} for binding to RNA polymerase core enzyme.

In vitro studies have confirmed that the *ntrA* gene encodes a protein with properties similar to those of a sigma factor. NTRA protein was partly purified from *Salmonella typhimurium* by using stimulation of transcription from the nitrogen-regulated *glnA* promoter as an assay for NTRA activity. NTRA co-purified with RNA polymerase during the early stages of purification but could be separated from core enzyme and σ^{70} by heparin agarose chromatography (Hirschman *et al.* 1985). *glnA* transcription was dependent on core polymerase and the NTRA-containing fraction but did not require σ^{70}. *E. coli* NTRA has been purified to homogeneity; it forms a complex with core RNA polymerase and allows transcription initiation from the nitrogen-regulated *glnA* promoter (Hunt & Magasanik 1985); no transcription was detected when σ^{70} was substituted for NTRA. It has been suggested that *ntrA* should be renamed *rpoN* and its product designated σ^{60} (Hunt & Magasanik 1985).

The nucleotide sequence of *K. pneumoniae ntrA* reveals that the gene product is an acidic 54 kDa polypeptide with an overall amino acid composition similar to that of σ^{70}, although the amino acid sequence of NTRA is not homologous with other sigma factors (Merrick & Gibbins 1985). However, most sigma factors contain at least one potential DNA-binding domain at the C-terminal end of the molecule (Merrick & Gibbins 1985; Stragier *et al.* 1985; or Gribskov & Burgess 1986) and NTRA contains two analogous DNA-binding regions that could possibly contact the −24 and −12 consensus elements found in NTRA-dependent promoters.

4. Activation of *nif* transcription

(*a*) *Activator binding sites enhance transcriptional activation*

Deletion analysis of the *nifL* promoter demonstrated that sequences as far as 150 base pairs upstream of the transcription initiation site were necessary for maximum promoter activity, although activation by either *ntrC* or *nifA* was still detectable in deletions removing sequences upstream of −28 (Drummond *et al.* 1983). The retention of some positive control in deletions removing the −35 region would tend to preclude binding of regulatory proteins to an upstream site as the sole mechanism of promoting transcriptional activation and suggests that an interaction of RNA polymerase, NTRA and activator proteins occurs close to the −24 and −12 sequences in this promoter.

Other *nif* promoters, particularly those which are efficiently activated by *nifA* (and only weakly by *ntrC*) appear to have a much higher affinity for activator proteins than the *nifL* promoter. For example, multiple copies of the *nifH* promoter inhibit chromosomal *nif* expression in *K. pneumoniae* resulting in Nif$^-$ phenotype. This 'multicopy effect', which is not shown by the *nifL* promoter, is thought to result from titration of activator molecules by excess promoter copies, thus preventing activation of chromosomal *nif* promoters (Buchanan-Wollaston *et al.* 1981; Riedel *et al.* 1983). Mutations that suppress the multicopy effect of a

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

nifH plasmid were obtained by selection of Nif$^+$ derivatives from a Nif$^-$ strain carrying multicopy *nifH* (Brown & Ausubel 1984). Most of these mutations were located in the −12 region of the promoter, but two mutations in upstream sequences were also identified; a point mutation at −136 and a deletion extending from −184 to −72. Subsequently, it was shown that all mutations in invariant nucleotides in the *nifH* promoter relieved the multicopy effect, although most of these mutations did not prevent activation by *nifA*. However, deletion of upstream sequences from −184 to −72 decreased *nifA*-mediated activation to 5% of the wild-type level as well as relieving the multicopy effect (Buck *et al.* 1985). Further deletion analysis of the *nifH*, *nifU* and *nifB* promoters revealed that sequences upstream of −100 were required both for *nifA* mediated activation and multicopy inhibition. Upstream sequences were also required for *nifA*-mediated activation of the *nifB* promoter; this promoter is not strongly expressed and does not show the multicopy effect. Because multicopy inhibition is thought to result from activator titration and upstream sequences are required for the multicopy response, it might be expected that the upstream sequences alone, when cloned on a multicopy plasmid, would give rise to a Nif$^-$ phenotype. However, this is not the case; multicopy inhibition requires the presence of both of the upstream sequences and the downstream −24 and −12 elements (Buck *et al.* 1986). Moreover, both the upstream and downstream elements must be *in cis* to observe the multicopy response. These results suggest that *nifA*-mediated transcriptional activation requires binding of regulatory proteins to both the upstream and downstream elements of these promoters. The upstream sequences apparently enhance the affinity of NIFA for the promoter but activator titration does not occur in the absence of the downstream elements.

These observations are further complicated by a recent finding that sequences downstream of the *nifH* transcription start also play a role in multicopy inhibition and that activator titration requires ongoing transcription. The introduction of a transcription terminator or a frameshift mutation (which presumably gives rise to transcription termination via transcriptional polarity) in 3′ sequences close to the promoter resulted in relief of multicopy inhibition (Buck & Cannon 1987). These results are difficult to explain in mechanistic terms although it seems plausible that a 'pile-up' of non-transcribing RNA polymerase molecules upstream of the terminator might prevent access of proteins to the transcription initiation site, thus preventing activator titration.

Comparison of the nucleotide sequences upstream of −100 in the *nifH*, *nifU*, *nifB* and ORF (*nifJ*) promoters reveals a conserved sequence characterized by an invariant TGT–N_{10}–ACA motif (Buck *et al.* 1986). This sequence conforms to a consensus sequence for protein-binding sites on DNA and is most probably a NIFA-binding site although, as explained above, activator titration requires interaction with both the upstream and downstream consensus elements. The upstream element is not unique to *K. pneumoniae* and is also found in a similar location in the *nif* promoters of many diazotrophs (Buck *et al.* 1986; Alvarez-Morales *et al.* 1986). Mutations in conserved nucleotides in this sequence affect multicopy inhibition and *nifA*-mediated activation, and the spacing of ten nucleotides between the TGT and the ACA motifs is also critical for its activity (Buck *et al.* 1985; M. Buck, unpublished results).

One of the most interesting properties of the upstream activator sequence (UAS) is its ability to act at a distance. The optimal position for the element is around −136 but *nifA*-mediated activation of the *nifH* promoter is still detectable when the element is placed two kilobases upstream of the transcription initiation site (Buck *et al.* 1986). Multicopy inhibition is apparently more sensitive to changes in spacing than is activation, and is not detectable when the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

element is moved more than 200 nucleotides upstream of the start site. The element is able to act independent of its orientation although it is not active when placed 3′ to the promoter. Its activity is reduced considerably when placed closer than 100 nucleotides upstream from the transcription start site (Buck *et al.* 1987). These properties resemble those of the yeast upstream activator sequences (Guarente 1984; Giniger *et al.* 1985).

Recently it has been shown that *ntrC*-mediated activation of the *glnA* promoter can also occur at a distance (Reitzer & Magasanik 1986). The 5′ regulatory region of enteric *glnA* promoters contains tandem promoters separated by about 100 nucleotides, the upstream promoter resembles a typical prokaryotic promoter and is subject to repression by *ntrC*, whereas the downstream promoter resembles the *nif* promoter consensus and is activated by *ntrC* in the presence of *ntrA* (Dixon 1984*b*; Reitzer & Magasanik 1985). Purified NTRC protein binds to a site (or sites) close to the upstream promoter (Ames & Nikaido 1985; Hawkes *et al.* 1985) thus repressing transcription initiation. NTRC also represses transcription from the *ntrBC* promoter by binding to a homologous site (Reitzer & Magasanik 1983; Ueno-Nishio *et al.* 1984). Using higher concentrations of NTRC, Hirschman *et al.* (1985) identified three weak binding sites located between the upstream and downstream *glnA* promoters as well as confirming the presence of two high affinity sites overlapping the upstream *glnA* promoter. The consensus sequence for the high affinity NTRC-binding sites contains the invariant motif 5′ GCAC–N_7–GTGC 3′ whereas the weak binding sites contain a less well conserved inverted repeat of the motif 5′ GGTGC 3′. Deletion of the high affinity sites in the *E. coli glnA* promoter does not prevent transcriptional activation in response to high levels of NTRC, but does decrease activation when low levels of this protein are present. The high affinity sites appear to function when moved 1400 base pairs upstream of the promoter because activation by low levels of NTRC is still detectable (Reitzer & Magasanik 1986). As in the case of *nifA*-mediated activation of the *nifH* promoter, the high affinity NTRC-binding site did not apparently function when placed close to the *glnA* promoter.

Both NIFA and NTRC can therefore activate transcription at a distance, provided that a high affinity binding site is present (figure 2). The sequence upstream of the *nifL* promoter does not contain a recognisable NIFA- or NTRC-binding site although there are several candidates for potential half-sites. This may explain the absence of activator specificity observed with *nifL* upstream sequences (Drummond *et al.* 1983) and it is plausible that a high concentration of activator is required for transcriptional activation of the *nifL* promoter. Purified

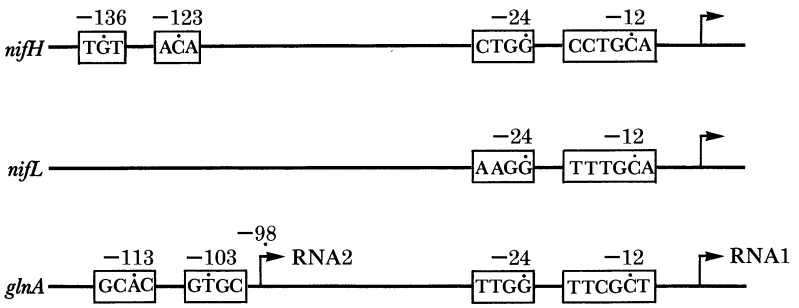

FIGURE 2. Schematic representation of *ntrA*-dependent promoters and upstream activator sequences (UAS) in *K. pneumoniae*. All numbering is with reference to the transcription initiation site (+1). The *glnA* promoter contains two transcription initiation sites, forming RNA1 at +1 and RNA2 at −98. Sequence motifs in the UAS and the −24, −12 elements are boxed. The *nifL* promoter does not contain a recognizable UAS. For the *glnA* promoter, only the high-affinity binding site for NTRC is shown.

NTRC does not bind this promoter at concentrations adequate to allow binding to the *glnA* and *ntrBC* promoters (Hawkes *et al.* 1985) and *in vitro* transcription experiments indicate that a higher concentration of NTRC is required for activation of *nifL* versus *glnA* transcription (Austin *et al.* 1987). When the *nifA*-specific activator sequence (derived from the upstream region of the *nifH* promoter) was placed 120 nucleotides upstream of the *nifL* transcription start site, *nifA*-specific activation was enhanced as expected, whereas *ntrC*-mediated activation did not increase (Buck *et al.* 1986). Moreover, the presence of the *nifA*-specific UAS conferred multicopy inhibition on plasmids carrying this hybrid promoter.

The ability of NIFA and NTRC to act at a distance raises interesting mechanistic possibilities and introduces a novel mode of positive control in prokaryotes. It is possible that proteins bound at a UAS could make contact with molecules bound at the downstream −24, −12 consensus via a loop in the DNA structure or by sliding along a linear DNA molecule towards proteins located at the downstream elements (Buck *et al.* 1986). Similar models have been proposed for action at a distance in a number of regulatory systems, both eukaryotic and prokaryotic (for review see Ptashne 1986). The DNA looping model may require binding of proteins to sequences located on the same side of the DNA helix; this would involve a strict spatial relation between the upstream and downstream sequences because these sites would have to be separated by an integral number of helical turns of DNA. This does indeed appear to be the case for the *nifH* promoter; activation is not affected by the addition of full helical turns but is hindered by the introduction of half-integral turns between the *nifH* UAS and the −24, −12 elements. Moreover, placing the *lac* operator between the upstream and downstream elements does not significantly affect activation, either in the presence or absence of the *lac* repressor, thus again favouring the DNA folding model rather than a sliding mechanism (Buck *et al.* 1987).

(*b*) *Structure of NIFA and NTRC*

The nucleotide sequences of *nifA* and *ntrC* genes show that their gene products possess considerable homology at the amino acid level (Buikema *et al.* 1985). Three domains have been identified in both NIFA and NTRC on the basis of secondary structure predictions (Drummond *et al.* 1986). The N-terminal domains of these proteins are not homologous but the remaining two thirds of each molecule show blocks of homology representing two folding domains. The central portion of each protein contains a structure representative of an interdomain linker and a block of more extensive homology indicative of a common function in both proteins. This large central domain has been implicated in positive control and may interact with NTRA or RNA polymerase. The C-terminal end of each protein contains a clear helix–turn–helix motif that is homologous with the DNA binding motifs found in repressors, activators and resolvases (Pabo & Sauer 1984). The first proposed helix in the motif (analogous to α-helix E of CAP and α-2 of lambda repressor) is almost identical in both NTRC and NIFA whereas the amino acid sequence of the proposed second helix (analogous to α-helix F in CAP and α-3 of lambda repressor) is not homologous. This is an expected observation because of recent proposals that the second α-helix constitutes a recognition helix that makes direct contact with DNA, whereas the first helix determines the general affinity of the protein for DNA (Ho *et al.* 1986; Ebright 1987). Because the nucleotide sequences required for NTRC and NIFA binding are dissimilar, it is not surprising that these proteins contain non-homologous second α-helices. Mutations in the C-terminal end of NTRC disrupt negative control, as expected for loss of DNA-binding

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

function. Interestingly, some of these mutants are still able to activate *glnA* expression (MacNeil *et al.* 1982). These results suggest that DNA binding is not an absolute pre-requisite for positive control, although it is possible that such mutations increase the affinity of the protein for the weak binding sites in the *glnA* regulatory region and decrease the affinity for tight binding sites.

The N-terminal domain of NTRC is surprisingly homologous to a number of diverse control proteins including OMPR, CHEY, CHEB, PHOB and DYE from *E. coli* and SPO0A and SPO0F from *Bacillus subtilis*. The significance of this homology is not clear but it is possible that this domain is involved in modulating the biological activity of the protein. This domain is not shared by NIFA, which may indicate that it is required for a specific interaction with NTRB, whereas the N-terminal domain of NIFA may be required to interact with NIFL.

5. Modulation of transcriptional activation

Data from a number of laboratories indicate that *ntrB* is required to prevent activation by *ntrC* in response to excess nitrogen and that *nifL* prevents activation by *nifA* in response to fixed nitrogen and oxygen. Both NTRB and NIFL can therefore 'sense' environmental changes and modulate transcription accordingly.

Although activation of *nif* transcription can occur in strains lacking NTRB, current evidence suggests that NTRB optimizes both the positive and negative functions of NTRC in nitrogen-limiting conditions, in addition to its role in deactivating NTRC under conditions of nitrogen excess. NTRB is required for optimal regulation of *glnA* promoters in *K. pneumoniae* (Alvarez-Morales *et al.* 1984; Dixon 1984*b*). Although there is no absolute requirement for *ntrB* for transcriptional activation *in vivo*, strains carrying a defined deletion removing the entire *ntrB* sequence show only weak regulation of the *glnA* and *ntrBC* promoters and do not fully activate transcription from the *nifLA* promoter (MacFarlane & Merrick 1987). These results show that NTRB modulates the activity of NTRC and suggest that there is a factor present in these strains that can partly substitute for NTRB. In agreement with this, crude S30 extracts prepared from strains lacking *ntrB*, activate transcription from the downstream *glnA* promoter *in vitro* (Hirschman *et al.* 1985) although activation of *glnA* transcription in a defined *in vitro* system requires the presence of core RNA polymerase and purified NTRA, NTRC and NTRB proteins (Hunt & Magasanik 1985). A mutant form of NTRC that activates *glnA* expression *in vivo* under conditions of nitrogen excess can activate *glnA* transcription in a defined *in vitro* system in the absence of NTRB (Hirschman *et al.* 1985). Both the *in vivo* and *in vitro* data therefore indicate that NTRB is required for conversion of NTRC into an active form and that an unknown factor present in enteric bacteria (and absent in defined *in vitro* systems) can substitute for this activity.

Under conditions of nitrogen excess, NTRB apparently converts NTRC into an inactive form. This response to nitrogen status involves a complex metabolic cascade, mediated by the products of *glnB* (P_{II} protein) and *glnD* (uridylyltranferase) (Bueno *et al.* 1985). Recent *in vitro* experiments show that NTRC undergoes covalent modification in response to NTRB. In the presence of ATP and wild-type NTRB, NTRC is phosphorylated and is then capable of activating transcription (Ninfa & Magasanik 1986). When purified P_{II} protein is added, NTRC is dephosphorylated and is inactive as a transcriptional activator. A mutant form of NTRB, which fails to respond to the nitrogen status *in vivo* continues to phosphorylate NTRC when P_{II} protein is added. The obvious interpretation of these results is that NTRB has a

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

protein kinase activity that is directly antagonized by P_{II} protein. Under nitrogen-limiting conditions P_{II} is mainly uridylylated, a form which presumably does not interact with NTRB. A series of defined mutations have been constructed in the *K. pneumoniae ntrB* gene by site-directed mutagenesis. Some of these mutations confer loss of response to nitrogen status, resulting in lack of repression at the *glnA* (RNA2) and the *ntrBC* promoters, as well as activation of the *glnA* (RNA1) and *nifLA* promoters (MacFarlane & Merrick 1987). Hence NTRB apparently modulates both the activator and repressor functions of NTRC. It is not yet known whether the NTRB-mediated covalent modification merely increases the affinity of NTRC for DNA or whether some other property of NTRC required for positive control is also affected.

The nucleotide sequence of *ntrB* (MacFarlane & Merrick 1985) reveals that the *ntrB* product does not have the typical features expected of a protein kinase; for example it does not contain a typical adenine nucleotide-binding pocket. However, a sequence closely resembling that of a nucleotide-binding site is found in both the NTRC and NIFA, leading to the speculation that both of these proteins have an autocatalytic kinase activity which is modulated by NTRB and NIFL respectively (Drummond & Wootton 1987). The C-terminal end of NTRB shows significant homology with comparable regions of a diverse family of regulatory proteins, including ENVZ, PHOR, CPXA and CHEA. Each of these proteins participates in a control system that involves pairs of proteins, the other members of the pair, OMPR, PHOB, DYE, and CHEB, being related by sequence homology to the N-terminal end of NTRC (Merrick *et al.* 1987; Drummond & Wootton 1987). The significance of the sequence conservation among these protein families is not yet clear, although there is an indication that these protein pairs functionally interact and in most cases the NTRB homologues appear to modulate the activate of the NTRC homologues. It is possible that the sequence homologies reflect either protein–protein interactions or covalent modification via a common phosphorylation event.

The C-terminal region of NIFL shows some homology with the comparable region of NTRB, but as mentioned above, the N-terminal ends of NIFA and NTRC are not homologous. This suggests that the mechanism of the modulation of NIFA activity by NIFL is different from that done by NTRB with NTRC. Multiple copies of the *nifL* gene inhibit transcriptional activation of *nif* operons (Buchanan-Wollaston *et al.* 1981). This may indicate that overproduction of NIFL results in titration of an effector required for maintaining NIFA in an active form in the presence of NIFL. Current genetic evidence suggests that NIFL, unlike NTRB, does not interact with P_{II} in response to nitrogen-status. *K. pneumoniae* strains carrying insertion mutations in or close to *glnB*, are constitutive with respect to *ntrC*-mediated activation, but *nifA*-mediated activation is still responsive to the nitrogen source (A. Holtel and M. Merrick, unpublished results). In such mutants NTRB is presumably unable to 'sense' the nitrogen signal because the P_{II} protein is absent, whereas P_{II} is apparently unnecessary for NIFL to respond to the presence of fixed nitrogen. The response of NIFL to oxygen might also result from interaction with an effector molecule or from the direct redox sensitivity of the NIFL protein itself. Interestingly, the NIFL sequence contains a cysteine pair with flanking amino acid sequences structurally similar to those found in the haem binding sites of C-type cytochromes (Drummond and Wootton 1987). It is therefore possible that NIFL 'senses' oxygen via a bound haem moiety.

Clearly, further genetic and biochemical analysis is necessary to define precisely the mechanism of *nif* transcriptional regulation, in response to oxygen and fixed nitrogen.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References

Alvarez-Morales, A., Betancourt-Alvarez, M., Kaluza, K. & Hennecke, H. 1986 Activation of the *Bradyrhizobium japonicum nifH* and *nifDK* operons is dependent on promoter upstream sequences. *Nucl. Acids Res.* **14**, 4207–4227.

Ames, G. F-L. & Nikaido, K. 1985 Nitrogen regulation in *S. typhimurium*. Identification of an *ntrC* protein-binding site and definition of a consensus binding sequence. *EMBO J.* **4**, 539–547.

Austin, S., Henderson, N. & Dixon, R. 1987 Requirements for transcriptional activation *in vitro* of the nitrogen-regulated *glnA* and *nifLA* promoters from *Klebsiella pneumoniae*: dependence on activator concentration. *Molec. Microbiol.* (In the press.)

Ausubel, F. M. 1984 Regulation of nitrogen fixation genes. *Cell* **37**, 5–6.

Beynon, J., Cannon, M., Buchanan-Wollaston, V. & Cannon, F. C. 1983 The *nif* promoters of *Klebsiella pneumoniae* have a characteristic primary structure. *Cell* **34**, 665–671.

Brown, S. E. & Ausubel, F. M. 1984 Mutations affecting regulation of the *Klebsiella pneumoniae nifH* (nitrogenase reductase) promoter. *J. Bact.* **157**, 143–147.

Buchanan-Wollaston, V., Cannon, M. C. & Cannon, F. C. 1981 The use of cloned *nif* (nitrogen fixation) DNA to investigate transcriptional regulation of *nif* expression in *Klebsiella pneumoniae*. *Molec. gen. Genet.* **184**, 102–106.

Buck, M. 1986 Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around −12 and −24 for activation by the *nifA* and *ntrC*(*glnF*) products. *J. Bact.* **116**, 545–551.

Buck, M., Khan, H. & Dixon, R. 1985 Site-directed mutagenesis of the *Klebsiella pneumoniae nifL* and *nifH* promoters and *in vivo* analysis of promoter activity. *Nucl. Acids Res.* **13**, 7621–7638.

Buck, M., Miller, S., Drummond, M. & Dixon, R. 1986 Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature, Lond.* **320**, 374–378.

Buck, M. & Cannon, W. 1987 Frameshifts close to the *Klebsiella pneumoniae nifH* promoter prevent multicopy inhibition by hybrid *nifH* plasmids. *Molec. gen. Genet.* **207**, 492–498.

Buck, M., Cannon, W. & Woodcock, J. 1987 Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve DNA loop formation. *Molec. Microbiol.* (In the press.)

Bueno, R., Pahel, G. & Magasanik, B. 1985 Role of *glnB* and *glnD* gene products in regulation of *glnALG* operon of *Escherichia coli*. *J. Bact.* **164**, 816–822.

Buikema, W. J., Szeto, W. W., Lemley, P. V., Orme-Johnson, W. H. & Ausubel, F. M. 1985 Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. *Nucl. Acids Res.* **12**, 4539–4555.

Cannon, F., Beynon, J., Buchanan-Wollaston, V., Burghoff, R., Cannon, M., Kwiatsowski, R., Lauer, G. & Rubin, R. 1985 Progress in understanding organisation and expression of *nif* genes in *Klebsiella*. In *Nitrogen fixation research progress* (ed. H. J. Evans, P. J. Bottomley & W. E. Newton), pp. 453–460. Dordrecht, Boston and Lancaster: Martinus Nijhoff.

Castano, I. & Bastarachea, F. 1984 *glnF–lacZ* fusions in *Escherichia coli*: studies on *glnF* expression and its chromosomal orientation. *Molec. gen. Genet.* **195**, 228–233.

de Bruijn, F. J. & Ausubel, F. M. 1983 The cloning and characterisation of the *glnF*(*ntrA*) gene of *Klebsiella pneumoniae*: role of *glnF*(*ntrA*) in the regulation of nitrogen fixation (*nif*) and other N-assimilation genes. *Molec. gen. Genet.* **192**, 342–343.

Dixon, R. 1984*a* The genetic complexity of nitrogen fixation. *J. gen. Microbiol.* **130**, 2745–2755.

Dixon, R. 1984*b* Tandem promoters determine regulation of the *Klebsiella pneumoniae* glutamine synthetase (*glnA*) gene. *Nucl. Acids Res.* **12**, 7811–7830.

Dixon, R. 1986 The *xylABC* promoter from the *Pseudomonas putida* TOL plasmid is activated by nitrogen regulatory genes in *Escherichia coli*. *Molec. gen. Genet.* **203**, 129–136.

Drummond, M. 1984 The nitrogen fixation genes of *Klebsiella pneumoniae*: a model system. *Microbiol. Sci.* **1**, 29–32.

Drummond M., Clements, J., Merrick, M. & Dixon, R. 1983 Positive control and autogenous regulation of the *nifLA* promoter in *Klebsiella pneumoniae*. *Nature, Lond.* **301**, 302–307.

Drummond, M., Whitty, P. & Wootton, J. 1986 Sequence and domain relationships of *ntrC* and *nifA* from *Klebsiella pneumoniae*: homologies to other regulatory proteins. *EMBO J.* **5**, 441–447.

Drummond, M. & Wootton, J. 1987 Sequence of *nifL* from *Klebsiella pneumoniae*: mode of action and relationship to two families of regulatory proteins. *Molec. Microbiol.* (In the press.)

Ebright, R. H. 1987 Proposed amino acid-base pair contacts for thirteen sequence-specific DNA binding proteins. In *Protein structure folding and design* (ed. D. Oxender). New York: Liss. (In the press.)

Giniger, F., Varnum, S. M. & Ptashne, M. 1985 Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* **40**, 767–774.

Gribskov, M. & Burgess, R. R. 1986 Sigma factors from *E. coli*, *B. subtilis* phage SP01 and phage T4 are homologous proteins. *Nucl. Acids Res.* **14**, 6745–6763.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Guarente, L. 1984 Yeast promoters: positive and negative elements. *Cell* **36**, 799–800.
Hawkes, T., Merrick, M. & Dixon, R. 1985 Interaction of purified NtrC protein with nitrogen regulated promoters from *Klebsiella pneumoniae*. *Molec. gen. Genet.* **201**, 492–498.
Hill, S., Kennedy, C., Kavanagh, E., Goldberg, R. & Hanua, R. 1981 Nitrogen fixation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *Klebsiella pneumoniae*. *Nature, Lond.* **290**, 424–426.
Hirschman, J., Wong, P-K., Sei, K., Keener, J. & Kustu, S. 1985 Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a σ factor. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7525–7529.
Ho, Y-S., Wulf, D. & Rosenberg, M. 1986 Protein–nucleic acid interactions involved in transcription: activation by the phage lambda regulatory protein cII. In *Regulation of gene expression – 25 years on* (ed. I. R. Booth & C. F. Higgins), pp. 79–103. Cambridge: Cambridge University Press.
Hunt, T. P. & Magasanik, B. 1985 Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG* and *glnL*. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8453–8457.
Inouye, S., Nakazawa, A. & Nakazawa, T. 1984 Nucleotide sequence surrounding the transcription initiation site of the *xylABC* operon on TOL plasmid of *Pseudomonas putida*. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1688–1691.
Kaluza, K., Alvarez-Morales, A. & Hennecke, H. 1985 Oligonucleotide directed mutagenesis of the *Rhizobium japonicum nifH* promoter. Functional evidence for the involvement of the −24 region in *ntrA* plus *nifA*-mediated activation. *FEBS Lett.* **188**, 37–42.
Khan, H., Buck, M. & Dixon, R. 1986 Deletion loop mutagenesis of the *nifL* promoter from *Klebsiella pneumoniae*: role of the −26 to −12 region in promoter function. *Gene* **45**, 281–288.
MacFarlane, S. A. & Merrick, M. 1985 The nucleotide sequence of the nitrogen regulation gene *ntrB* and the *glnA-ntrBC* intergenic region of *Klebsiella pneumoniae*. *Nucl. Acids Res.* **13**, 7591–7606.
MacFarlane, S. A. & Merrick, M. J. 1987 Analysis of the *Klebsiella pneumoniae ntrB* gene by site-directed *in vitro* mutagenesis. *Molec. Microbiol.* (In the press.)
MacNeil, T., Roberts, G. P., MacNeil, D. & Tyler, B. 1982 The products of *glnL* and *glnG* are bifunctional regulatory proteins. *Molec. gen. Genet.* **188**, 325–333.
Magasanik, B. 1982 Genetic control of nitrogen assimilation in bacteria. *A. Rev. Genet.* **16**, 135–168.
Merrick, M. 1982 A new model for nitrogen control. *Nature, Lond.* **297**, 362–363.
Merrick, M. & Gibbins, J. R. 1985 The nucleotide sequence of the nitrogen regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial RNA polymerase sigma factors. *Nucl. Acids Res.* **13**, 7607–7620.
Merrick, M., Hill, S., Hennecke, H., Hahn, M., Dixon, R. & Kennedy, C. 1982 Repressor properties of the *nifL* gene product in *Klebsiella pneumoniae*. *Molec. gen. Genet.* **185**, 75–81.
Merrick, M. J. & Stewart, W. D. P. 1985 Studies on the regulation and function of the *Klebsiella pneumoniae ntrA* gene. *Gene* **35**, 297–303.
Merrick, M. J., Austin, S., Buck, M., Dixon, R., Drummond, M., Holtel, A. & MacFarlane, S. 1987 Regulation of nitrogen assimilation in enteric bacteria. In *Proceedings of the International Symposium on Phosphate Metabolism in Micro-organisms*. Washington: ASM publications. (In the press.)
Minton, N. P. & Clarke, L. E. 1985 Identification of the promoter of the *Pseudomonas* gene coding for carboxy-peptidase G2. *J. mol. appl. Genet.* **2**, 26–35.
Ninfa, A. J. & Magasanik, B. 1986 Covalent modification of the *glnG* product NR_I, by the *glnL* product NR_{II}, regulates the transcription of the *glnALG* operon in *Escherichia coli*. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5909–5913.
Ow, D. W. & Ausubel, F. M. 1983 Regulation of nitrogen metabolism genes by *nifA* gene product in *Klebsiella pneumoniae*. *Nature, Lond.* **301**, 307–313.
Ow, D. W., Sundaresan, V., Rothstein, D. M., Brown, S. E. & Ausubel, F. 1983 Promoters regulatd by the *glnG* (*ntrC*) and *nifA* gene products share a heptameric consensus sequence in the −15 region. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2524–2528.
Ow, D. W., Xiong, Y., Gu, Q. & Shen, S. C. 1985 Mutational analysis of the *Klebsiella pneumoniae* nitrogenase promoter: sequences essential for positive control by *nifA* and *ntrC* (*glnG*) products. *J. Bact.* **161**, 868–874.
Pabo, C. O. & Sauer, R. T. 1985 Protein-DNA recognition. *A. Rev. Biochem.* **53**, 293–321.
Ptashne, M. 1986 Gene regulation by proteins acting nearby and at a distance. *Nature, Lond.* **322**, 697–701.
Reidel, G. E., Brown, S. E. & Ausubel, F. M. 1983 Nitrogen fixation by *Klebsiella pneumoniae* is inhibited by certain multicopy hybrid *nif* plasmids. *J. Bact.* **153**, 45–56.
Reitzer, L. J. & Magasanik, B. 1983 Isolation of the nitrogen assimilation regulatory NR_I the product of the *glnG* gene of *Escherichia coli*. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5554–5558.
Reitzer, L. J. & Magasanik, B. 1985 Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1979–1983.
Reitzer, L. J. & Magasanik, B. 1986 Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites from the promoter. *Cell* **45**, 785–792.
Roberts, G. P. & Brill, W. J. 1981 Genetics and regulation of nitrogen fixation. *A. Rev. Microbiol.* **35**, 207–235.
Shen, S. C., Xue, Z-t, Kong, Q-t, & Wu, Q-l. 1983 An open reading frame upstream from the *nifH* gene of *Klebsiella pneumoniae*. *Nucl. Acids Res.* **11**, 4241–4250.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Stragier, P., Parsot, C. & Bouvier, J. 1985 Two functional domains conserved in major and alternate bacterial sigma factors *FEBS lett.* **187**, 11–15.
Ueno-Nishio, S., Mango, S., Reitzer, L. J. & Magasanik, B. 1984 Identification and regulation of the *glnL* operator-promoter of the complex *glnALG* operon of *Escherichia coli*. *J. Bact.* **160**, 379–384.

Discussion

F. C. Cannon (*Biotechnica International Inc., Cambridge, Massachusetts, U.S.A.*). How does Dr Dixon know that NIFA bound to the UAS does not exert its influence through sequences other than the −24 and −12 consensus elements, for example at the A–T-rich sequences located between the two elements?

R. A. Dixon. Deletion analysis reveals that these A–T-rich sequences are not essential for *nifA*-mediated activation or activator titration. However, we cannot discount the possibility that sequences between the UAS and the downstream elements contribute to loop formation.

R. Haselkorn (*Department of Biophysics and Theoretical Biology, University of Chicago, Illinois, U.S.A.*). Have Dr Dixon's footprinting experiments revealed whether covalent modification of NTRC alters its DNA-binding properties?

R. A. Dixon. We have not done such experiments with fully phosphorylated NTRC protein.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES