REGULATION OF *AZORHIZOBIUM CAULINODANS* ORS571 NITROGEN FIXATION *(NIF/FIX)* GENES

Promotoren: dr. A. van Kammen, hoogleraar in de moleculaire biologie
dr. F.J. de Bruijn, professor of microbiology, Michigan State University, East Lansing.

STELLINGEN

1. In *Azorhizobium caulinodans* wordt de consensussequentie -24/-12 in de promoter van het *nifA* gen door een andere NtrA (RpoN) sigma factor herkend dan bij de overige *nif* genen.

   Dit proefschrift; hoofdstuk 3 en 4.

2. De methode van Kammann voor plaats gerichte mutagenese is niet juist.

   Kammann *et al.* 1989. NAR. 17:5404.

3. De biosynthese van haem wordt in *A. caulinodans* geregeld via het symbiontische zuurstof sensor/regulator systeem *fixLJK*.

   Dit proefschrift; hoofdstuk 6.

4. De werkelijke homologie tussen de *leuA* genen van *E. coli* en *S. typhimurium* is groter dan blijkt uit de door Ricca en Calvo gepubliceerde sequentie van het *leuA* gen van *S. typhimurium*.

   Ricca and Calvo 1990. NAR. 18:1290.

5. Het gen voor α-isopropylmalaat isomerase van *Mucor circinelloides* is verkeerd benoemd; het moet *leu1* (of *leuB*) worden genoemd.

   Roncero *et al.* 1989. Gene 84:335-343.

6. De genetische variatie in een bacteriekolonie wordt onderschat.

Drake, JW. 1991. PNAS 88:7160-7164.

7. De term "Miller units" dient in publicaties beter te worden gespecificeer

Miller, JH. 1972. Experiments in Molecular Genetics, p352-355. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

8. Spontane mutaties zijn waarschijnlijk niet altijd willekeurig.

Foster, PL. 1992. J. Bact. 174:1711-1716;
Lenski and Mittler 1993. Science 259:188-194.

9. Het uiteenvallen van de wereldschaakbond valt niet toevallig samen me het ineenstorten van het voormalige Sovjet-rijk.

10. Na de zetten 1. g4, e5 ontstaat na 2. f3 een stelling die niet te verdedige

Stellingen behorende bij het proefschrift:
"Regulation of *Azorhizobium caulinodans* ORS571
nitrogen fixation *(nifIfix)* genes"
door John Stigter,
te verdedigen op 14 januari 1994 te Wageningen.

John Stigter

REGULATION OF *AZORHIZOBIUM CAULINODANS* ORS571 NITROGEN FIXATION *(NIF/FIX)* GENES

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. C. M. Karssen
in het openbaar te verdedigen
op vrijdag 14 januari 1994
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Stigter, John

Regulation of Azorhizobium caulinodans ORS571 nitrogen fixation (Nif/fix) genes / John Stigter. - [S.l. : s.n.]
Thesis Wageningen. - With ref. - With summary in Dutch.
ISBN 90-5485-212-7
Subject headings: Azorhizobium caulinodans.

The front cover shows a sequencing gel (T-track) of the *A. caulinodans nifA* promoter (-24/-12) mutants: n23, n13, n12, n3, n2, n1 and wild-type (see chapter 3).

Ter nagedachtenis aan mijn moeder

Aan: mijn vader,
Els, Wilma en Ann

ACKNOWLEDGEMENTS

I would like to thank Dr. Frans J. de Bruijn for his support, trust and criticism and for providing the facilities in his laboratory to conduct the research described in this thesis. I also wish to thank Prof. dr. Ab van Kammen for his support.
I acknowledge the comforting support from various friends and colleagues during my stay in Cologne (Germany) and East-Lansing (Michigan, USA).

CONTENTS

page

CHAPTER 1

ABSTRACT AND OUTLINE

Biological nitrogen fixation is the microbial process by which atmospheric dinitrogen (N_2) is reduced to ammonia. In all microbes studied, dinitrogen reduction is catalyzed by a highly conserved enzyme complex, called nitrogenase. The nitrogenase subunits and functions required for nitrogenase assembly and activity are encoded by the nitrogen fixation *(nif/fix)* genes.

Nitrogen-fixing organisms can be roughly divided into two major groups: the free-living nitrogen fixing (diazotrophic) species, such as *Klebsiella pneumoniae*, *Azotobacter vinelandii* and *Rhodobacter capsulatus*, and symbiotically N_2 fixing organisms, such as *Rhizobium* species and *Bradyrhizobium japonicum*.

The fate of the ammonia produced by these two distinct groups of nitrogen fixing organisms is quite different. While free-living nitrogen-fixing organisms will assimilate the ammonia produced for their own growth (diazotrophy), the strictly symbiotically nitrogen-fixing organisms (predominantly) excrete the ammonia into the cytoplasm of infected plant cells, to be assimilated by the host. The regulation of nitrogen fixation *(nif/fix)* gene expression is quite different as well, and has been difficult to compare directly between diazotrophs and strictly symbiotic nitrogen fixing organisms. The discovery of *Azorhizobium caulinodans* ORS571, a tropical rhizobium capable of both diazotrophy and symbiotic nitrogen fixation in stem- and root nodules induced on its host, the tropical shrub *Sesbania rostrata*, changed this situation and has allowed a direct comparison of *nif/fix* gene regulation in the diazotrophic versus symbiotic state. This unusual organism was chosen for the research reported in this thesis, which focuses on the expression of the central *nif/fix* regulatory gene *nifA*, a gene which is essential for diazotrophic growth and symbiotic nitrogen fixation, and which responds to a complex regulatory circuit.

In chapter 2 a current review of the regulation of nitrogen fixation genes is presented. The review focuses on the regulation of *nifA* gene expression and NifA activity in different bacterial species.

In chapter 3 the involvement of *ntr*-mediated control of the *nifA* promoter via the -24/-12 element is analyzed by site-specific mutagenesis and chimeric *nifA-lacZ* reporter gene fusions integrated into the *A. caulinodans* chromosome.

The -24/-12 promoter element was shown to be important for *nifA* gene expression suggesting the involvement of a σ54 (NtrA; RpoN)-type transcription factor in *nifA* gene regulation.

Chapter 4 addresses the involvement of a *ntrA(rpoN)*-like sigma factor in *nifA* expression, and reports the cloning and analysis of an *A. caulinodans ntrA(rpoN)* gene. Although the *ntrA(rpoN)* gene identified in this study was shown to control the expression of some of the *A. caulinodans nif* genes *(like nifHDK)* and nitrate assimilation genes, it did not appear to regulate the expression of *nifA*, suggesting the presence of an additional (*nifA* specific) *ntrA(rpoN)*-equivalent gene.

To search for *trans*-acting factors involved in the regulation of the *nifA* expression, the *in vitro* interaction of proteins in *A. caulinodans* crude cell extracts with the *nifA* regulatory region was studied by gel retardation assays. Chapter 5 presents the binding of (a) sequence-specific protein(s) in extracts of *A. caulinodans* to the *nifA* 5' upstream region. However the nature of the protein(s) and the exact location of the binding site in the *nifA* 5' upstream region remains to be determined.

During the preparation of crude cell extracts from *A. caulinodans fixLJ* and *fixK* mutant strains, large amounts of a red/pink, UV fluorescent pigment was observed in the culture medium. Chapter 6 describes the isolation and characterization of this pigment.

Chapter 7 describes the isolation and characterization of an *leuA*-like gene of *A. caulinodans* which was incidentally identified with an *Escherichia coli fnr* probe. Comparison of the 5' upstream region of the *A. caulinodans leuA* gene with the *leu* operons of other organisms suggest a conserved regulation mechanism for the expression (transcription attenuation), found in many amino acid biosynthetic operons.

Finally in chapter 8 the current model for nitrogen fixation gene regulation in *A. caulinodans*, deduced from the studies presented in this thesis and previous data, is presented.

CHAPTER 2

REGULATION OF NITROGEN FIXATION GENES IN THE FREE-LIVING VERSUS SYMBIOTIC NITROGEN-FIXING STATE

2.1 INTRODUCTION

Biological nitrogen fixation is an extremely energy demanding process. Reduction of 1 molecule of N_2, catalyzed by the nitrogenase enzyme complex, requires 16 molecules of ATP, under ideal conditions. However, in vivo ATP requirements of up to 42 per N_2 reduced have been measured (O'Brian and Maier 1989).

Free-living nitrogen-fixing bacteria, such as *Klebsiella pneumoniae*, will therefore only reduce atmospheric N_2 when no other nitrogen sources are available. In the case of symbiotic nitrogen fixing bacteria such as rhizobia, the availability of fixed nitrogen is less important. Ammonia produced by the nitrogenase reaction generally does not enter the bacterial metabolism, but is exported to and assimilated by the plant (Miflin and Cullimore 1984). In return the fixing bacteroids are provided by the plant with (ample) carbon source(s) for the production of ATP and reducing equivalents.

The nitrogenase enzyme complex is extremely oxygen sensitive. The mechanism of oxygen inactivation is not known, but the extreme reducing environment might cause reduction of oxygen to superoxide and radicals which inactivates the enzyme. The environmental oxygen tension is therefore a major regulatory factor for both diazotrophs and symbiotic organisms.

Three nitrogenase systems, containing different metal centers have been discovered in *Azotobacter vinelandii* (Bishop and Premakumar 1992) and other diazotrophs (Newton 1993). In addition to nitrogen and oxygen control, metal regulation (availability of molybdenum and vanadium) plays a role in nitrogenase expression (Pau 1993).

Nitrogen fixation has not been found in eukaryotes, but it is very widely distributed among eubacteria and archaebacteria. Within these groups, N_2 fixation has been reported in almost 100 genera distributed over most of the major phylogenetic divisions (Young 1992).

Despite the considerable taxonomic diversity among N_2-fixing organisms, research over the last decade has revealed a surprising degree of uniformity in the mechanisms underlying the regulation of nitrogen fixation *(nif)* genes in many, though not all, diazotrophs. In the great majority of N_2-fixing organisms, a part of the pathway of *nif* gene regulation is similar. A majority of the *nif/fix* genes is activated through the action of the highly conserved central regulatory protein NifA. However the factors that influence both the activity and expression of NifA itself, vary considerably from one organism to another and reflect the physiological and environmental influences on N_2 fixation in the different nitrogen fixing species.

In section 2.2 I will review the common features of *nif* gene transcription and nitrogenase production found in the majority of nitrogen fixing organisms. In the main section (2.3) a review of the factors that influence both the expression and activity of the *nifA* gene product (NifA) in different species is presented.

2.2 THE MECHANISM OF NIFA MEDIATED ACTIVATION

The most detailed description of the regulation of the *nif* genes by NifA comes from studies on the facultative anaerobe *K. pneumoniae*, which has become a model system for genetic studies of the structure, function and regulation of nitrogen fixation and assimilation genes.

Under microaerobic conditions *K. pneumoniae* reduces atmospheric nitrogen to ammonia in response to low fixed nitrogen levels. The 20 *nif*-genes, needed to synthesize nitrogenase and to regulate its manufacture, are organized in 8 distinct operons, located on a 25kb chromosomal fragment. The DNA sequence of this entire *nif* gene cluster has been determined (Arnold *et al.* 1988). Two of the *nif* genes (*nifA* and *nifL*) are regulator genes. NifA is a transcriptional activator required for the expression of all *nif* operons, while NifL antagonizes the action of NifA in response to oxygen and low levels of fixed nitrogen (Dixon *et al.* 1980; Buchanan-Wollaston *et al.* 1981). Initiation of transcription by NifA requires another gene product NtrA(RpoN), which constitutes a novel RNA polymerase sigma factor (σ54; see fig. 1).

The 5' upstream (regulatory) regions of the *nif* genes have a special structure. They do not have the typical prokaryotic promoter sequence TTGACA-n17-TATAAT normally found in the -35/-10 region, but contain a highly conserved sequence between positions -27 and -11 relative to the start point of transcription.

This sequence CTGGCAC-n5-TTGCA (Gussin *et al.* 1986; Morett and Buck 1989), of which the $_{-24}$GG/$_{-12}$GC motif is absolutely essential, is diagnostic of promoters which utilize a minor form of RNA polymerase, in which the major vegatative sigma factor σ70 is replaced by the novel sigma factor σ54, encoded by the *ntrA(rpoN)* gene. Mutational analysis of -24/-12 promoters and studies of the interactions between both RNA polymerase containing σ54 (Eσ54) and σ54 alone indicate that the highly conserved nucleotides in this promoter represent specific sites of interaction between the promoter DNA and σ54, and the strict 10 bp spacing reflects the interaction of σ54 with one face of the DNA helix (Buck 1986; Buck and Cannon 1992).

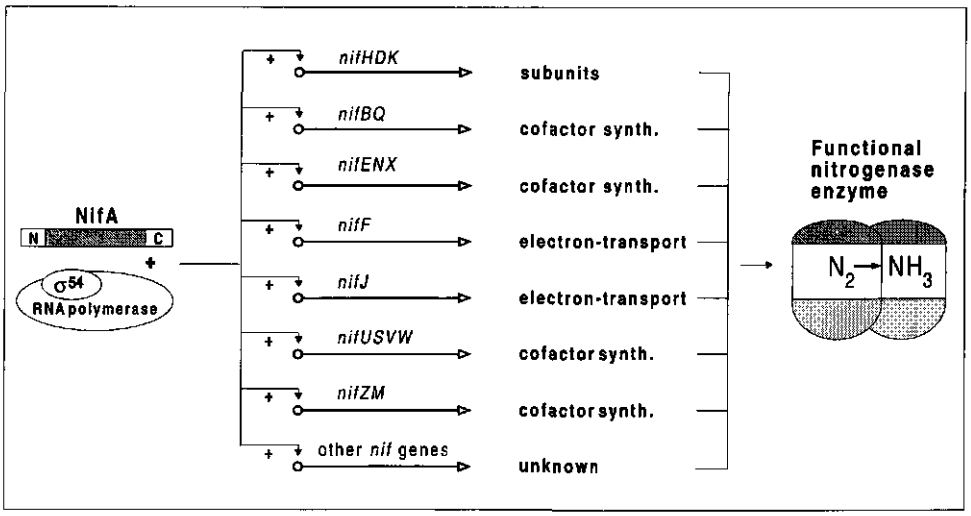

Figure 1. Common features of *nif* gene transcription and nitrogenase production. See text for details.

In addition another highly conserved DNA motif, TGT-n10-ACA, is present in almost all NifA activated promoters between 80 and 150 bp upstream of the transcription start. This motif, called the upstream activator sequence (UAS) has been shown to constitute the binding site for the NifA protein (Buck *et al.* 1986; Cannon *et al.* 1990).

While *nif* genes were some of the first σ54-dependent genes to be characterized, many other genetic systems are now known to utilize this form of RNA polymerase (Eσ54) and it has been suggested that in all these cases a common mechanism of transcription initiation exists (Kustu *et al.* 1989).

The following model has been proposed for the activation of transcription (see fig. 2): Eσ54 binds to the promoter in the -24/-12 region to form a closed complex, but is incapable of isomerizing to the open complex in the absence of a specific activator protein (NifA). The activator NifA, binds to a specific site or sites (UAS) about 100 bp upstream of Eσ54. The C-terminal domain of NifA contains a helix-turn-helix (HTH) motif (Drummond *et al.* 1986), which mediates DNA binding.

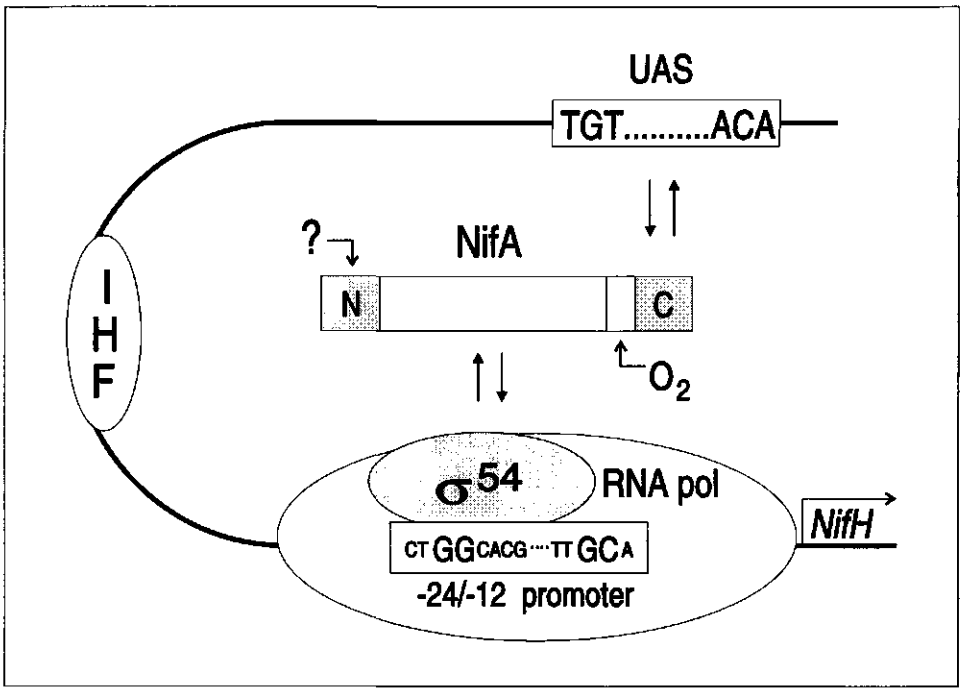

Figure 2. Model of NifA mediated *nif* activation.

See text for details.

The interaction between Eσ54 and the activator (NifA) is mediated by DNA loop formation (Buck *et al.* 1987), facilitated by the integration host factor (IHF) (Santero *et al.* 1989; Hoover *et al.* 1990). The IHF protein, which is involved in a variety of processes in the bacterial cell, introduces sharp bends in DNA (Friedman 1988; Freundlich *et al.* 1992). IHF binds to a site between the UAS for NifA and the promoter, and induce DNA bending which facilitates productive interactions between NifA bound at the UAS and Eσ54 bound in the -24/-12 region (Hoover *et al.* 1990; Santero *et al.* 1992).

The central domain of NifA is responsible for transcriptional activation. All members of the family of σ54 dependent activator proteins contain this highly conserved central domain, which is essential for their positive control function (Morett *et al.* 1988; Huala and Ausubel 1989; North *et al.* 1993). It consists of about 238 amino acids and contains a conserved nucleotide binding motif. It has been shown that hydrolysis of ATP or another nucleoside-triphosphate by the activator protein is necessary for DNA strand unwinding and initiation of transcription (Hoover *et al.* 1990; Santero *et al.* 1992). Mutations which modify the ATP-binding motif specifically impair the catalysis of open complex formation (Cannon and Buck 1992). So the rate of transcription is controlled by the ATPase activity of the activator protein, which is strongly stimulated by site-specific DNA binding (to the UAS), as shown for the σ54 dependent activator NtrC (see below; Weiss *et al.* 1991; Austin and Dixon 1992).

The NifA family can be divided into two groups on basis of the presence or absence of two extra cysteine residues in a Cys-XXXX-Cys motif just beyond the central domain. One group lacks the Cys motif and includes *K. pneumoniae*, *Azotobacter chroococcum*, and *A. vinelandii* NifA's, including VnfA and AnfA (see section 2.3.2). All these NifA proteins are likely to be oxygen insensitive. In contrast, all rhizobial NifA proteins (see section 2.3.5, 2.3.6 and 2.3.8) and those of *Rhodobacter capsulatus* (section 2.3.3), *Azospirillum brasilense* (section 2.3.4) and *Herbaspirillum seropedicae*, contain the Cys-X11-Cys-X19-Cys-XXXX-Cys motif, and three members of this group (*Bradyrhizobium japonicum*, *Rhizobium meliloti*, and *R. capsulatus*) have been shown to be oxygen sensitive (Fischer *et al.* 1988; 1989; Hennecke 1990).

2.3 FACTORS THAT AFFECT *NIFA* EXPRESSION AND NIFA ACTIVITY

Three major environmental factors are involved in the regulation of *nifA* expression and NifA activity, namely nitrogen, oxygen and metal availability. The importance of each of these factors varies from organism to organism and depends on their ecology and physiology. Whereas the nitrogen status is a major factor in free-living diazotrophs, the oxygen status is the dominant effector in symbiotic systems. Metal availability appears to be critical for organisms which synthesize alternative nitrogenases.

In the next sections I will review the most well studied regulatory systems. First in diazotrophs (sections 2.3.1 to 2.3.4), then in symbiotic nitrogen fixing organisms (sections 2.3.5 and 2.3.6), and finally in *Azorhizobium caulinodans* (section 2.3.8).

2.3.1 Regulation in *K. pneumoniae*

In *K. pneumoniae nif* gene expression is controlled by the nitrogen and oxygen status of the cells, and these factors affect both *nifA* transcription and NifA activity. The *nifA* gene is coordinately transcribed with *nifL*. The expression of the *nifLA* operon, in addition to being autoregulated by NifA, is under control of a general nitrogen regulation *(ntr)* system, which responds to fluctuating concentrations of combined nitrogen in the cell (Drummond *et al.* 1983; Ow and Ausubel 1983).

This general nitrogen regulation system (best studied in *E. coli;* see Magasanik 1982; Magasanik and Neidhardt 1987) controls the expression of many genes involved in nitrogen metabolism. When the supply of ammonia, the preferred nitrogen source, becomes restricted, enteric bacteria respond by activating the expression of a number of operons. These nitrogen-regulated *(ntr)* operons encode products that facilitate the assimilation of low concentrations of ammonia by the glutamine synthetase-glutamate synthase pathway and the utilization of alternative sources of nitrogen, such as amino acids. When all other N-sources in the medium are utilized, the diazotrophic organism will finally switch on the nitrogen fixation genes and use atmospheric N_2 as N-source.

Sensing of the intracellular nitrogen concentration occurs via the *gln*-system (see below), which in turn, activates the *ntr*-system. Under nitrogen starvation conditions the *ntr*-system first activates the genes involved in low ammonia assimilation such as *glnA*, which encodes glutamine synthetase, followed by the genes involved in the catabolism/transport of amino acids such as histidine (*hut)*, proline (*put)* and arginine (*aut)* via the NAC (nitrogen assimilation control) system (Bender 1991), and finally the genes involved the atmospheric nitrogen fixation via the NifA system (see fig. 4).

The *gln/ntr*-system

The intercellular concentration of ammonia (N) influences the glutamine (Gln) to 2-ketoglutarate (2KG) ratio, which are the 'signal metabolites' for the *gln* system. The conversion of 2KG to glutamate requires Gln or ammonia, while glutamate in turn is converted into Gln through the addition of an second ammonia molecule, by the enzyme glutamine synthetase.

A decrease in ammonia results in a rapid drop in the intracellular concentration of Gln, followed by a rise in that of 2KG. The resulting 'signal', low Gln and high 2KG, activates the primary sensor of the cellular N status, an uridylyl-transferase/uridylyl-removing enzyme (UTase/UR, encoded by *glnD*; see fig 3).

The UTase mediates the uridylylation of PII (a small regulatory protein, encoded by *glnB).* PII exists in the cell as a tetrameric polypeptide. During the activation reaction, a uridylyl-group from UTP is coupled to all four subunits of the PII protein (Holtel and Merrick 1988). The uridylylated form of PII (PII-UMP) promotes the phosphorylation of the transcriptional activator protein NtrC (NRI; *glnG*) by the kinase/dephosphorylase protein NtrB(NRII; *glnL*). NtrB phosporylates NtrC in a two-step process, involving an autophosphorylation step (in which a histidine residue in NtrB is phosphorylated), followed by transfer of the phosphate to an aspartate residue in the N-terminus of NtrC.

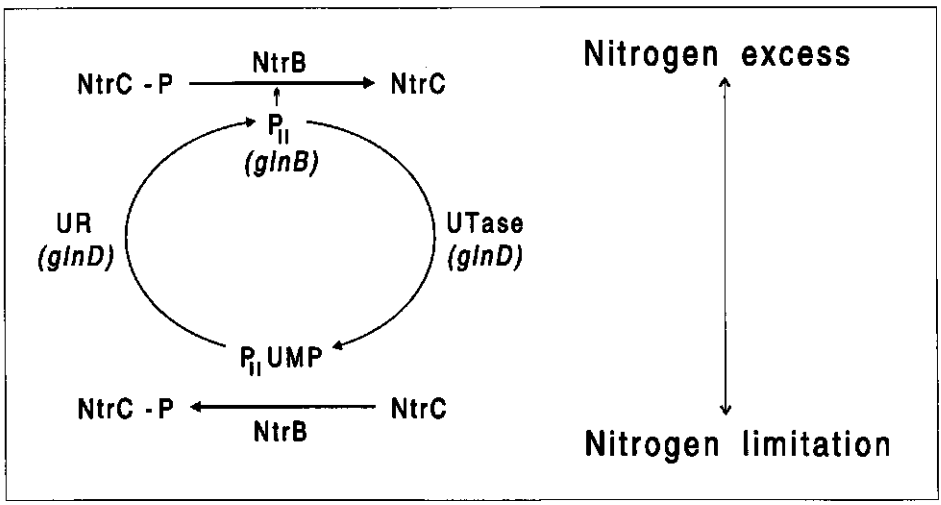

Figure 3. Cascade of covalent modification that regulates the activity of the transcriptional activator NtrC in response to changes in intracellular nitrogen status. See text for details.

NtrC, like NifA, is a σ54-transcriptional activator which binds to a specific sequence upstream of the promoter and activates transcription in concert which NtrA(RpoN; Wedel *et al.* 1990). The phosphorylation of NtrC increases its DNA binding properties and stimulates its ATPase activity, which is essential for open complex formation (Weiss *et al.* 1991; Austin and Dixon 1992).

Under conditions of N-excess, when Gln>>2KG, this cascade of events is reversed. GlnD now acts as a uridylyl-removing enzyme, converting PII-UMP to PII. PII no longer stimulates the kinase activity of NtrB, and NtrB now promotes the dephosphorylation of NtrC. As a result, the activator and DNA-binding properties of NtrC are diminished and expression from NtrC-dependent promoters is switched off (see figs. 3 and 4).

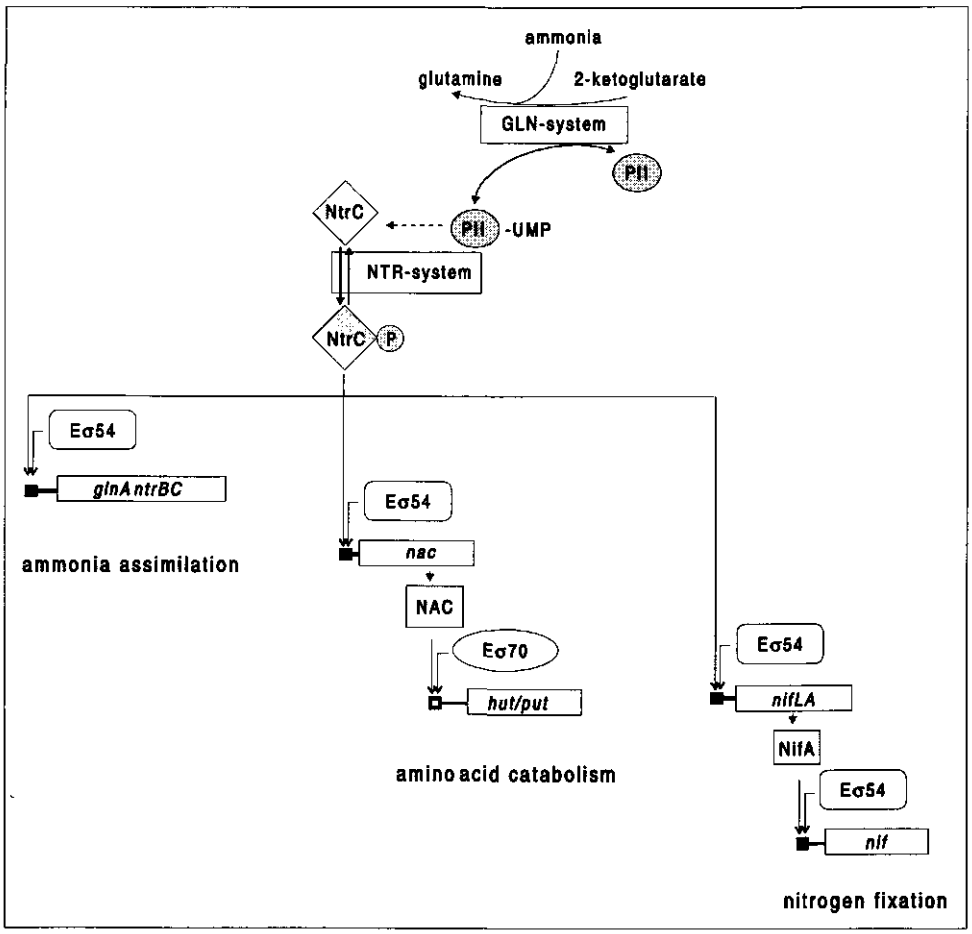

Figure 4. Nitrogen regulation in *K. pneumoniae*. See text for details.

The *ntrC/nifA* system

NtrC-phosphate (P), like NifA, is a transcriptional activator that stimulates the initation of transcription with NtrA(RpoN) from σ54 *(ntr)*-promoters. NtrC specifically interacts with a DNA motif: GCAC-n5-TGGTGC (Reitzer and Magasanik 1986; Sasse-Dwight and Gralla 1988) found upstream of the -24/-12 promoter regions of N-*(ntr)* regulated genes. The activity and amount of NtrC(-P), which is ultimately regulated by the N-status of the cell via the *gln* system (see above), determines whether *nifLA* are expressed.

When growing on a medium containing high concentrations of ammonia, a cell contains only about 5 molecules of NtrC (Reitzer and Magasanik 1983). This amount, when phosphorylated, is sufficient to activate transcription of only the *glnAntrBC* operon of the Ntr activated regulon. Activation is efficient at this operon because of the precence of multiple NtrC binding sites upstream of the promoter (Reitzer *et al.* 1989). Besides leading to synthesis of glutamine synthethase (GlnA; GS), this transcription elevates the levels of NtrB and NtrC. When the cellular NtrC concentration rises, the operons for the amino acid catabolism genes are activated via the transcriptional activator NAC (Bender 1991). When still N-starved, the NtrC concentration further increases, (approximately 70 molecules per cell) and the nitrogen fixation genes are activated via the transcriptional activator NifA.

The *K. pneumoniae nifLA* operon has two weak NtrC binding sites, which function cooperatively, 120-140 bp upstream of the -24/-12 consensus sequence (Minchin *et al.* 1988; 1989; Wong *et al.* 1987). Neither of these sites is highly homologous to the consensus NtrC binding site, and both have a very low affinity for NtrC. By comparison, the *glnA* promoter *(glnAp2)* contains 5 upstream NtrC binding sites, two of which are high affinity sites, highly homologous to the consensus sequence (Dixon 1984; Ninfa *et al.* 1987). The *nac* promoter contains two NtrC binding sites, one high affinity site, homologous to the NtrC consensus and one weak binding site (Schwacha and Bender 1993). *In vitro* experiments indicate that 10-fold lower concentrations of NtrC are required to activate *glnAp2* transcription than are required for *nifLA* transcription (Austin *et al.* 1987). The consequence of this difference in NtrC affinity between *pnifLA, pnac* and *glnAp2* is that a significantly greater concentration of NtrC-P must build up in the cell for activation of *pnifA* than for *pnac* and *glnAp2*. Hence, the *nif* regulon is only expressed under conditions of severe N-limitation (see fig. 4).

The strength of *ntr*-regulated promoters is not only determined by the amount and affinity of upstream activator binding sites. In addition, the properties of the Eσ54 binding site consensus determine strong or weak binding of NtrA(RpoN) to the promoter. Sequences from -17 to -14 are critical in the recognition and binding of Eσ54 to the -24/-12 element (Buck and Cannon 1989; 1992). Whereas a closed promoter complex with Eσ54 is formed at the *glnA* promoter in the absence of the activator, contacts between Eσ54 and the -24/-12 region of the *nifL* promoter are not observed in the absence of NtrC (Minchin *et al.* 1989). Similar closed complexes between Eσ54 and the -24/-12 *nifH* promoter are not observed *in vivo* in the absence of NifA (Morret and Buck 1989). These observations suggest that at promoters such as *nifL* and *nifH* where the affinity of Eσ54 is weak, the activator may play a role in stabilizing the closed complex in addition to catalyzing open complex formation (Austin *et al.* 1991).

The properties of the *nifLA* promoter determine its response not only to N-limitation but also to oxygen limitation. *nifLA* expression is controlled by the state of supercoiling of *pnifLA*, which is affected by anaerobiosis (Kong *et al.* 1986; Dixon *et al.* 1988; Whitehall *et al.* 1992).

In addition to transcriptional regulation of *nifA*, post-transcriptional regulation of NifA by the NifL protein has been documented. NifL inhibits NifA activity in the presence of oxygen, or if levels of fixed nitrogen exceed a certain threshold. However NifA activity is not impaired in the absence of NifL (Arnott *et al.* 1989).

Several mechanisms have been suggested, such as destabilizing of *nif* mRNA's in the presence of oxygen (Collins *et al.* 1986) and direct interaction with the central and/or C-terminal domain of NifA (Drummond *et al.* 1990). It is not known by which method NifL responds to nitrogen or oxygen, but NifL action appears to be independent of the *gln*-system (Holtel and Merrick 1989) and the redox environment (Contreras and Drummond 1991).

The NtrBC-system and two component regulatory systems

NtrB and NtrC are members of a large family of bacterial regulatory proteins involved in signal transduction, the so-called two-component systems. The two components consists of a sensor protein (e.g. NtrB), containing a conserved C-terminal domain and a response regulator protein (e.g. NtrC), containing a conserved N-terminal domain (Ronson *et al.* 1987b).

A general model for signal transduction for these systems has been proposed: The N-domain of the sensor protein, often located in the cytoplasmic membrane (however not for NtrB, since it responds to a cytoplasmic signal) receives an (environmental) signal, and transduces that signal to its C-domain by autophosphorylation. Next, the phosphate is transferred from the C-domain of the sensor protein to the N-terminal domain of the regulator protein. Phosphorylation of the regulator, which is often a transcriptional activator, alters the activity of its C-domain, which ultimately carries out the appropriate response (Nixon *et al.* 1986; Albright *et al.* 1989a).

Diverse environmental signals are transduced through this common mechanism to regulate a wide range of processes in bacteria, including response to nutrient deprivation, osmolarity changes, chemotaxis, sporulation, symbiosis, and bacterial pathogenesis (for review see Parkinson and Kofoid 1992).

2.3.2 Regulation in *Azotobacter*

Regulation of *nif* expression in the free-living diazotroph *Azotobacter* is complicated by the presence of three biochemical and genetically distinct nitrogenase enzymes, each of which is synthesized under different physiological conditions. A conventional molybdenum nitrogenase, a vanadium nitrogenase, and an alternative nitrogenase, which contains neither molybdenum nor vanadium, have been described (Bishop and Joerger 1990).

Expression of all three nitrogenase systems is repressed by high concentrations of fixed nitrogen and each of the alternative gene clusters has its own activator gene, consisting of a *nifA* homologue (i.e., *vnfA* for the vanadium containing enzyme and *anfA* for the iron containing enzyme; Joerger *et al.* 1989).

The regulation of the *nif* genes in *A. vinelandii* appears to be similar to that in *K. pneumoniae*. The *nif*, *vnf*, and *anf* genes contain conserved -24/-12 promoter sequences, characteristic of σ54-dependent operons, and upstream binding sites, for NifA, VnfA and AnfA respectively. As expected, the expression of all three systems is inactivated in an *ntrA(rpoN)* mutant (Bishop and Premakumar 1992).

Discrimination between activation of different systems occurs by selective expression and control of the activity of the different NifA homologues. Less well known are the factors that regulate this expression and activity. Best studied is the *nifLA* operon, involved in the activation of the conventional molybdenum nitrogenase genes.

Homologues of *ntrB*, *ntrC* and *glnD (nfrX)* are present in *A. vinelandii* but, unlike in *K. pneumoniae*, *ntrC* mutations do not affect expression of any of the three nitrogenase systems. While in *K. pneumoniae*, high ammonia concentrations prevent *nifLA* transcription, *A. vinelandii nifA* expression is not regulated by the N-status. It is independent of NtrBC, and NtrA(RpoN), and the *nifA* promoter does not contain a -24/-12 consensus sequence (Toukdarian and Kennedy 1986). The nitrogen control appears to occur primarily at the level of NifA activity and to be mediated solely through the *nifL* gene, which is located immediately upstream of *nifA* (Bennet *et al.* 1988; Bali *et al.* 1992; Raina *et al.* 1993).

In contrast to *K. pneumoniae*, the *A. vinelandii* NifL protein shows homology to the histidine protein kinase family, including NtrB and FixL, and seems to interact directly with NfrX (a GlnD homologue), regulating its activity (Contreras *et al.* 1991; Blanco *et al.* 1993).

2.3.3 Regulation in *Rhodobacter*

In the photosynthetic bacteria *R. capsulatus* the *nifAB* operon and *nifU* gene are duplicated. NifB is required for FeMo-cofactor biosynthesis. The function of NifU is not known, but appears to be involved in stabilization of the Fe protein (encoded by *nifH)*. The specific roles of these duplicate genes, which can substitute for each other, are not yet known (Masepohl *et al.* 1988; Preker *et al.* 1992).

Like in *K. pneumoniae*, expression of *nifA1* and *nifA2* requires the NtrB/NtrC homologues, encoded by *nifR2* and *nifR1*, and the *glnB* homologue, *nifR5*. The *glnB* product (PII) modifies NtrB activity in response to the nitrogen status (see fig. 3). However mutations in *ntrB* and *ntrC* do not have an Ntr$^-$ phenotype (i.e. they can still utilize proline, glutamine, or arginine as nitrogen sources). At least two more regions homologous to *ntrBC* have been identified in *R. capsulatus* (Kranz and Haselkorn 1988), and the genes identified in this study as *nifR1* and *nifR2* may be *nif*-specific in their action.

Another difference with *K. pneumoniae* is that the *nifA1* and *nifA2* promoter regions do not contain a -24/-12 consensus sequence, and that their expression is independent of the *ntrA(rpoN)* homologue *nifR4* (Preker *et al.* 1992). NifR4 seems different from other sequenced σ54's; it has no acidic domain, and lacks the two conserved putative leucine zipper domains (see chapter 4). NifR4 may be a nitrogen fixation-specific RNA polymerase sigma factor (Kranz *et al.* 1990). Expression of *nifR4*, like the *nifA* genes, is nitrogen controlled via NtrC, which probably interacts with an RNA polymerase sigma factor that is different from the NtrA-like (i.e., NifR4) sigma factor (Kranz *et al.* 1990).

It is likely that *R. capsulatus* has a second *ntrA(rpoN)*-like gene, since expression of hydrogenase (*hup*) genes in this organism is dependent on an activator protein, HupR1, which is homologous to other σ54-dependent activators (Richaud *et al.* 1991). However *hup* expression is not affected by *nifR4* mutations (Colbeau and Vignais 1992), indicating the existence of a possible NtrA2(RpoN2). Two copies of *ntrA(rpoN)* have recently been reported in *R. sphaeroides* (Meijer and Tabita 1992).

Since the *nifA*'s do not have a -24/-12 consensus sequence, this suggests involvement of a possible *ntrA2(rpoN2)* product and NtrC (NifR1) in the activation of a novel transcriptional activator which activates *nifA1* and *nifA2*. A situation analogous to the activation of *K. aerogenes hut* genes by the NtrC-regulated *nac* gene product (Bender 1991). Alternatively it has been suggested that in *R. capsulatus* the expression of the *nifA1,2* genes by NtrC (NifR1) occurs directly and requires a novel sigma factor not homologous to σ54 (Foster-Harnett and Kranz 1992).

Oxygen control of *nif* expression in *R. capsulatus* is postulated to be mediated both by DNA supercoiling and by the oxygen sensitivity of the NifA protein itself. The *R. capsulatus nifA* gene product belongs to the class of oxygen-sensitive NifA proteins, which are independent of NifL (see above). Recently additional regulation of NifA activity by ammonia has been suggested by Masepohl *et al.* (1993).

2.3.4 Regulation in *Azospirillum*

Azospirillum can fix nitrogen under free-living conditions or in association with grasses. The genetics of nitrogen fixation is best studied in *A. brasilense* Sp7. The regulation of NifA synthesis and activity in *A. brasilense* differs from *K. pneumoniae*, in that the *nifA* promoter does not contain a -24/-12 consensus and is expressed both under conditions of nitrogen fixation and in the presence of oxygen and ammonia (Liang *et al.* 1991). However NifA appears to be in an inactive form under these conditions.

NtrB and NtrC are not essential for nitrogen fixation, although they are involved in the regulation of NifA synthesis and more importantly in the regulation of nitrogenase activity by the switch-off mechanism (de Zamaroczy *et al.* 1993a).

The *glnB* product (PII), is required for the regulation of nitrogen fixation and seems to be involved in post-translational modification of NifA, in response to the nitrogen status, a mechanism so far only found in *A. brasilense* (Liang *et al.* 1992; de Zamaroczy *et al.* 1993b).

In addition, NifA contains the conserved cysteine residues, correctly spaced (C-X11-C-X19-C-X5-C) which are probably involved in sensitivity to oxygen (de Zamaroczy *et al.* 1993a).

2.3.5 Regulation in *Rhizobium*

In contrast with *K. pneumoniae*, which fixes nitrogen in a free-living form, bacteria of the genus *Rhizobium* normally fix nitrogen only in symbiotic association with plants of the family *Leguminoseae*. Symbiotic nitrogen fixation occurs in root nodules, which are complex, highly differentiated structures formed by the interaction of bacteria and their host plant.

Rhizobia usually gain entry into the legume root by penetrating root hairs. In response to bacterial colonization and attachment to the root hairs, the plant produces a cellulose tube called an infection thread, which surrounds the bacteria and grows into the interior of the cell, where it ramifies. The infection threads serve as conduits through which *Rhizobium* penetrates the interior of the root. As the infection threads proliferate, cells in the inner root cortex, which are normally terminally differentiated, are induced to dedifferentiate, divide, and form a meristem (nodule primordium). The dividing meristematic zone of cells leads to the outgrowth of a nodule from the root surface. Finally, various branches of the infection thread release their bacteria into the host-cell cytoplasm. At this point the bacteria have differentiated into a new form called a bacteroid and are surrounded by a host-derived 'peribacteroid' membrane. The bacteroids express nitrogenase and specific cytochromes, and can be considered to be highly specialized nitrogen-fixing organelles (for review, see Hirsch 1992).

As rhizobia infect the root, and the nodule develops, the bacterial environment switches from aerobic to a microaerobic. The microaerobic environment in the infected cells of the nodule is not only essential for the oxygen sensitive nitrogen fixation process to take place, but it has, in fact been found to constitute the main developmental (environmental) signal for the induction of nitrogen fixation gene expression.

Three species of *Rhizobium* have been analyzed: *R. meliloti*, *R. leguminosarum*, and *R. trifolii*, of which *R. meliloti* has been studied in most detail. Therefore the latter system will be discussed in detail here.

Several genes homologous to *K. pneumoniae nif* genes have been found in *R. meliloti*. They map on a large plasmid called pSym (symbiotic plasmid), along with other genes involved in nodulation *(nod)* and in nitrogen fixation *(fix)*. In rhizobial species the designation *fix* has been used to identify those genes associated with N_2 fixation which do not have a homologue in *K. pneumoniae*; however the *fix* terminology is becoming increasingly blurred as homologues of *fixABC*, which have been proposed to be involved in electron transport to nitrogenase, are present in *A. vinelandii* and *A. brasilense* and have recently been identified in *E. coli* (Yura *et al.* 1992).

The *R. meliloti nif* and *fix* genes belong to two regulons: the *nifHDKE*, *fixABCX*, *nifAB*, *nifN* and *fdxN* genes, which are under the positive control of the transcriptional activator NifA and a second regulon, represented by the reiterated *fixNOQP* genes (David *et al.* 1987), which are activated by the product of the reiterated *fixK* gene (Batut *et al.* 1989; see fig. 5).

The *nifA* gene is located downstream of the *fixABCX* gene cluster and upstream of *nifB*, and can be transcribed either from the *fixABCX* promoter (P2) or from a promoter between *fixABCX* and *nifA* (*pnifA*; Kim *et al.* 1986). The *fixABCX* promoter is a typical σ54-dependent promoter, which is activated by NifA once a significant level of NifA has been produced from *pnifA*. Activation at P2 would allow readthrough transcription to boost NifA levels. No characteristic promoter features are present in the sequence upstream of *nifA* (see below). The *nif*-system in *R. meliloti* is not regulated by the *gln/ntr*-system as in *K. pneumoniae*. For example, while *K. pneumoniae nifA* is expressed under control of *ntrBC* in response to nitrogen limitation, a *R. meliloti ntrC* mutant is not affected in *nifA* expression and still able to fix nitrogen *in planta* (Szeto *et al.* 1987).

The *fixK* gene is located in the central part of the *fix* cluster, which is reiterated on pSym (Renalier *et al.* 1987). The amino acid sequence of the transcriptional activator, FixK shows homology with the *E. coli* regulator Fnr (fumarate and nitrate reduction). Fnr is a transcriptional regulatory protein essential for the oxygen-regulated expression of anaerobic respiratory processes in *E. coli* (Spiro and Guest 1990). In spite of its homology to Fnr, FixK lacks a 21 N-terminal amino acid stretch, which is thought to be responsible for oxygen sensitivity of Fnr transcriptional activity. Like Fnr, FixK can act positively and negatively; in addition to regulating *fixNOQP* expression positively, it has a negative effect on the expression of *nifA* and on its own expression (fig. 5).

Both *nifA* and *fixK* genes are induced asymbiotically to levels exceeding those in nodule bacteroids when the external O_2 concentration is reduced to microaerobic levels. This induction is independent of fixed nitrogen availability (Ditta *et al.* 1987; Hertig *et al.* 1989). This is logical since the physiological environment for *R. meliloti* inside the nodule is nitrogen-rich and microaerobic (Appleby 1984).

Two genes, *fixL* and *fixJ*, are responsible for sensing and transducing the low oxygen signal in *R. meliloti* (David *et al.* 1988; de Philip *et al.* 1990) and activating the expression of *nifA* and *fixK* in response to microaerobiosis. FixL and FixJ are homologous to a family of bacterial two-component regulatory proteins involved in signal transduction.

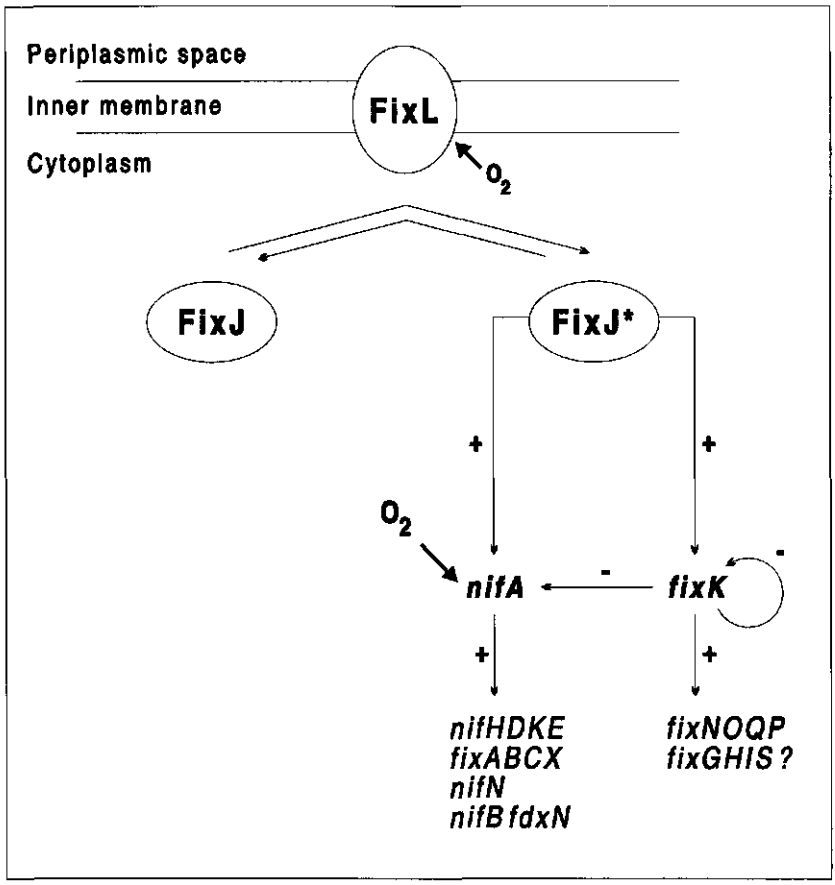

Figure 5. Cascade regulation of nitrogen fixation genes in *R. meliloti*.
See text for details.

FixL is a membrane-bound hemo-protein with (auto)kinase- and phosphatase activities (Gilles-Gonzales *et al.* 1991; Lois *et al.* 1993a). The heme moiety, located in the central domain is directly involved in oxygen binding and required for oxygen regulation of the kinase activity (Monson *et al.* 1992). In the absence of oxygen, the autophosphorylating activity of FixL is increased. FixL-phosphate then transfers the phosphate, by a direct phosphotransfer mechanism, as found for several other two-component system proteins, to the transcriptional activator FixJ (Hertig *et al.* 1989). FixJ consist of two modules: a carboxy-terminal module responsible for transcriptional activation, and an amino-terminal phosphoryl acceptor module that regulates the activity of the carboxy-terminal module (Kahn and Ditta 1991). Once phosphorylated, FixJ become an active transcription factor capable of inducing transcription from the *nifA* and *fixK* promoters (see fig. 5).

The phosphatase activity of FixL serves to maintain a very tight control over FixJ-phosphate levels. In addition, the FixL phosphatase activity, which is increased under atmospheric oxygen tension may prevent the accumulation of FixJ-phosphate levels, as a result of phosphorylation of FixJ by heterologous sensors (cross-talk; Ninfa *et al.* 1988). Phosphatase activity also provides a way for the system to efficiently revert to its original state if the oxygen concentration rises again (Lois *et al.* 1993a). Recently the complete FixLJ-*nifA* signal transduction pathway, from effector to its ultimate target was reconstituted *in vitro* (Agron *et al.* 1993).

Both in the case of the *fixK* and *nifA* genes, a region between -62 and -45 in the promoter region is essential for induction (Virts *et al.* 1988), suggesting that this region may be required for binding of a positive activator (FixJ). By mutational analysis critical residues were identified between -54 and -39 relative to the transcription start site and a common motif in this region with the *fixK* promoter region was suggested. However neither *pnifA* nor *pfixK* contains any apparent direct or inverted repeats suggesting a single asymmetric binding site. (Agron *et al.* 1992). Recent evidence suggest that for expression of *fixK* in *E. coli*, the major sigma factor (σ70, RpoD) is required (Batut *et al.* 1991).

2.3.6 Regulation in *Bradyrhizobium japonicum*

As in *R. meliloti*, in *B. japonicum* the expression of *nif* and *fix* genes is regulated predominantly, if not exclusively, in response to the cellular oxygen status. Activation of these genes occurs during symbiosis in root nodules or in free-living cells grown at low oxygen concentrations. NifA controls the expression of at least 13 *nif* and *fix* genes (Fischer *et al.* 1993).

The *B. japonicum nifA* gene is part of an operon, *fixRnifA*. The predicted *fixR* product is not homologous to any previously identified regulatory proteins in other diazotrophs, but shows strong homology to dehydrogenases involved in acetate metabolism. Mutations in *fixR* do not impair N_2 fixation (Thöny *et al.* 1989; Morett *et al.* 1993).

In contrast to *R. meliloti*, *B. japonicum nifA* is not only expressed anaerobically and in bacteroids but also aerobically. The level of aerobic expression is about one fourth of the microaerobic expression level. The microaerobic expression is dependent on NifA (auto-activation) and σ54. For the aerobic expression, a DNA element located 66bp from the transcription start point is essential. This element is the binding site for a protein present in crude extracts of *B. japonicum* and presumed to be a positive regulator. The promoter region of the *fixRnifA* operon shares homology with the -24/-12 class of promoters. However, in the absence of σ54 this operon is still transcribed (Kullik *et al.* 1991).

Recently it was found that the aerobic expression of *fixRnifA* is strongly repressed by acetic acid, indicating control by the carbon/redox state of the cell. In addition two overlapping promoters have been found (Morett *et al.* 1993). This suggests the existence of two different regulatory systems for the expression of the *fixRnifA* operon. One recognized by σ54 and regulated by oxygen via NifA, and the second recognized by σ70, the housekeeping sigma factor of *B. japonicum*, subject to the carbon/redox state of the cell (Morett *et al.* 1993).

Apart from the transcriptional regulation of the *nifA* gene, oxygen also controls the activity of the NifA protein. There is a basal level of expression in aerobic conditions from a -35/-10 promoter activated by an as yet unidentified *trans*-acting protein factor, however, the NifA synthesized under these conditions is inactive. Expression is elevated four to five-fold when O_2 limitation allows the formation of active NifA as a result of auto-activation form the -24/-12 promoter.

An *fixLJ* operon has recently been identified and shown to be essential for symbiotic nitrogen fixation and anaerobic nitrate respiration (Anthamatten and Hennecke 1991). However, unlike in *R. meliloti*, the *fixLJ* genes are not involved in the regulation of the *fixRnifA* operon.

As in most nitrogen fixing organisms *B. japonicum* NifA mediated activation does require σ54. Interestingly, *B. japonicum* has two functional *ntrA(rpoN)* genes, from which one is oxygen regulated by *fixLJ* and the other negatively autoregulated (Kullik *et al.* 1991).

A *fixK*-like gene was identified whose predicted product is homologous to the FixK protein of *R. meliloti* and the *E. coli* Fnr protein (Anthamatten *et al.* 1992). In contrast to FixK of *R. meliloti*, but similar to the *E. coli* Fnr protein, FixK of *B. japonicum* contains a cysteine-rich, putatively oxygen-responsive domain at its N-terminal end. Expression of *fixK* was induced at low oxygen tension and depended on the *fixLJ* gene products. However unlike mutations in *fixLJ*, a deletion in *fixK* did not affect symbiotic nitrogen fixation or anaerobic nitrate respiration, suggesting the existence of a second *fixLJ*-regulated *fixK* homologue (Fischer *et al.* 1993). So far two target genes for *fixLJ*/*fixK*-mediated control have been identified. These include the *ntrA1(rpoN1)* gene (Kullik *et al.* 1991) and a cluster of genes involved in microaerobic respiration, *fixNOQP* (Preisig *et al.* 1993).

2.3.7 Regulation in bacteria lacking σ54-dependent *nif* promoters

The great majority of nitrogen fixing organisms appear to have *nif* promoters that are dependent on NtrA(RpoN) and NifA for their activation (see above), however there are exceptions such as *Archaebacteria*, *Clostridium*, *Anabaena* and *Desulfovibrio* species. These systems will be briefly reviewed in the next paragraphs, although the mechanism of *nif* gene regulation in these organisms is not fully understood.

2.3.7.1 Archaebacteria

A number of species of archaebacteria, the methanogens, have been shown to be capable of fixing N_2. Nucleotide sequences have been determined for *nifH* homologues from three species, but no sequence motifs corresponding to *nif* consensus promoters have been identified.

In *Methanococcus themo-lithotrophicus* expression of ORF*nifH*1 is regulated by the nitrogen source, suggesting that a mode of nitrogen control distinct from the typical *ntr* system (see above) regulates *nif* transcription in these organisms (Souillard and Sibold 1989; Lobo and Zinder 1992).

2.3.7.2 *Clostridium*

In the anaerobic gram-positive bacterium, *Clostridium pasteurianum* six genes that hybridize to *K. pneumoniae nifH* and one gene homologous to *nifE*, have been cloned and sequenced. One of these genes (*nifH1)* encodes the Fe protein of

the nitrogenase enzyme. Four of the other five *nifH* copies are definitely transcribed in N-limiting conditions and the functions of these copies are presently unknown, although they could encode one or more alternative nitrogenase Fe proteins. The sequences upstream of these genes show homology to typical -35/-10 regions and no sequence homology to -24/-12 promoters. The genes do share common upstream sequences with a consensus ATCAATAT-N6-10-ATGGATTC in the -100 region, but the role of these sequences is not known (Chen *et al.* 1986; Wang *et al.* 1988).

2.3.7.3 *Anabaena*

In the cyanobacterium *Anabaena* nitrogen limitation causes heterocyst differentiation. In the heterocyst, nitrogenase and hydrogenase synthesis are co-induced with the enzyme glutamine synthetase, which is required for assimilation of fixed nitrogen. Nitrogenase synthesis is controlled by a novel mechanism of gene rearrangement that occurs coordinately with heterocyst differentiation (Haselkorn and Buikema 1992). Relatively little is known about the regulation of *nif* gene transcription. Both *nifH* and *glnA* (encoding glutamine synthetase) have promoter sequences different from the -35/-10 or σ54-specific -24/-12 sequence. It has therefore been suggested that these sequences may reflect an *Anabaena nif-ntr* consensus sequence that is recognized by a novel form of RNA polymerase σ factor that mediates nitrogen control in this organism (Turner *et al.* 1983).

2.3.7.4 *Desulfovibrio*

In the proteobacteria, *Desulfovibrio gigas* the *nifH* gene has been sequenced and a potential NtrA(RpoN)-dependent promoter sequence has been identified upstream of the coding sequence. An upstream activator sequence homologous to those in NifA-dependent promoters was not found and no significant expression of a p*nifH-lacZ* fusion was detected in *E. coli*. By contrast, the *D. gigas nifH* promoter was nitrogen regulated in *K. pneumoniae*, but this nitrogen control was not affected by *ntrA/rpoN*, *ntrC*, or *nifA* mutations. It is not known whether the genetic control observed in *K. pneumoniae* reflects a mechanism operative in *D. gigas*. It also remains to be elucidated how the nitrogen control is mediated in this organism (Kent *et al.* 1989).

2.3.8 Regulation in *Azorhizobium caulinodans*

As pointed out above, *A. caulinodans* ORS571 (Dreyfus *et al.* 1988; de Bruijn, 1989) has the unusual capacity to grow on N_2 in the free-living state, as well as to *fix* nitrogen in nodules induced on the stem and roots of its host, the tropical legume *Sesbania rostrata* (Dreyfus and Dommergues 1981; de Bruijn 1989).

The regulatory circuitry controlling the expression of the *A. caulinodans* nitrogen fixation *(nif/fix)* genes in culture and in symbiosis has been shown to have components in common with *nif*-regulatory elements characteristic of the free-living diazotroph *K. pneumoniae*, as well as for different *Rhizobium* species.

As in the case of other rhizobia and the free-living diazotroph *Klebsiella pneumoniae*, the expression of the structural genes for nitrogenase (*nifHDK*) and other *nif/fix* genes in *A. caulinodans*, is controlled by the central *nif*-specific gene *nifA*, both in culture and in nodules (*in planta*; Donald *et al.* 1986; Pawlowski *et al.* 1987; de Bruijn *et al.* 1990).

Expression of the *nifA* gene, in turn, is controlled by the cellular nitrogen and oxygen status and regulated by (at least) three distinct two-component regulatory systems, *ntrBC, ntrYX* (Ratet *et al.* 1989; Pawlowski *et al.* 1987; 1991) and *fixLJ* (Kaminski and Elmerich 1991), both in culture and *in planta*. The *A. caulinodans ntrBC* and *ntrYX* pairs of sensor-regulator genes share significant homology with one another and both appear to be involved in the cellular response to the concentration of combined nitrogen, facilitating *nif/fix* gene derepression under nitrogen-(N-) starvation conditions. However, while the *ntrB* gene product closely resembles its cytosolic counterpart in enteric bacteria (NR_{II}; Ninfa and Magasanik 1986), the *ntrY* gene product (NtrY) contains a distinct trans-membrane-like domain, and therefore may be involved in extra- versus intra-cellular N-sensing (Pawlowski *et al.* 1991).

The *A. caulinodans fixLJ* gene pair is involved in sensing the oxygen (O_2) concentration and facilitates the derepression of the *nif/fix* genes under microaerobic (O_2-limiting) conditions, through the product of the *fixK* gene (Kaminski and Elmerich 1991; Kaminski *et al.* 1991). Thus, *A. caulinodans nif/fix* gene expression appears to be controlled by the N-regulation *(ntr)* pathway, commonly found in free-living diazotrophs such as *K. pneumoniae*, as well as the O_2-regulation *(fixLJ)* pathway, found in strictly symbiotic nitrogen-fixing organisms, such as *R. meliloti* (see de Bruijn and Downie 1991; Merrick 1992).

This dual response pathway reflects the unusual chimeric nature of *A. caulinodans* as free-living diazotroph and symbiotic nitrogen-fixing organism and is a consequence of the need of this organism to respond to quite different physiological conditions for nitrogen fixation in culture, as well as *in planta* (see de Bruijn 1989; de Bruijn *et al.* 1990).

The *A. caulinodans nifA* 5' upstream region has been found to contain distinct DNA motifs found in the promoter regions of N- and O_2-regulated genes (Nees *et al.* 1988; Ratet *et al.* 1989), including a -24/-12 promoter element and a Fnr binding site consensus sequence, found in the promoter regions of genes anaerobically induced via the transcriptional activator Fnr (TTGAT-n4-ATCAA; see de Bruijn *et al.* 1988; Spiro and Guest 1990).

It is plausible that the Fnr binding site consensus sequence constitutes the target site for the FixK protein, which shares homology with Fnr (Kaminski *et al.* 1991). The -24/-12 element in the *A. caulinodans nifA* promoter is likely to be involved in mediating *ntr* regulation of *nifA* expression via the *ntrBC* and *ntrYX* systems (Ratet *et al.* 1989; Pawlowski *et al.* 1987; 1991) and to interact with the alternative σ factor NtrA(RpoN) (σ54; see Kustu *et al.* 1989), which has been identified in several rhizobial species (Ronson *et al.* 1987a; Stanley *et al.* 1989; Kullik *et al.* 1991).

In this thesis work, the involvement of the different *cis*-acting elements in the *A. caulinodans nifA* promoter region is analyzed using site directed mutagenesis and chimeric *nifA-lacZ* reporter gene fusions. To identify the *trans*-acting factor(s) involved in *ntr* control via the -24/-12 element in the *nifA* promoter an *ntrA(rpoN)*-like gene was cloned, sequenced and mutagenised. Gel retardation assays were carried out with the *nifA* 5' upstream region, to determine the role of previously determined *trans*-acting factors, and to identify additional *trans*-acting factors. Based on these and previous results a current model of *nif/fix* gene regulation in *A. caulinodans* is presented in chapter 8.

CHAPTER 3

MUTAGENESIS OF THE *A. CAULINODANS NIFA* -24/-12 PROMOTER ELEMENT

3.1 ABSTRACT

Using site-directed mutagenesis, mutations were introduced at three positions in the -24/-12 promoter element of the *Azorhizobium caulinodans nifA* gene and chimeric *nifA-lacZ* reporter gene fusions were constructed. Single basepair mutations in the conserved -25 or -13 G residues were found to reduce or abolish *nifA* promoter activity, respectively, demonstrating that the -24/-12 promoter element is important for *nifA* gene expression and suggesting the involvement of a σ54 (NtrA/RpoN)-type transcription factor in *nifA* gene regulation. A two-basepair mutation at positions -25 and -16 was found to create a relatively nitrogen-control-independent, highly expressed *nifA* promoter.

Part of this work has been published in: *Azorhizobium caulinodans* nitrogen fixation *(nif/fix)* gene regulation: Mutagenesis of the *nifA* -24/-12 promoter element, characterization of a *ntrA(rpoN)* gene, and derivation of a model. John Stigter, Maria Schneider, and Frans J. de Bruijn. Molecular Plant-Microbe Interactions (1993) Vol. 6, No. 2, pp. 238-252.

3.2 INTRODUCTION

The *A. caulinodans nifA* 5' upstream region has been found to contain distinct DNA motifs found in the promoter regions of N- and O_2-regulated genes (Nees *et al.* 1988; Ratet *et al.* 1989), including a -24/-12 promoter element ($_{-25}$GG-n10-GC$_{-12}$ relative to the transcriptional start site; see Thöny and Hennecke, 1989) and a Fnr binding site consensus sequence, found in the promoter regions of genes anaerobically induced via the transcriptional activator Fnr (TTGAT-n4-ATCAA; see Spiro and Guest, 1990; de Bruijn *et al.* 1990). It is plausible that the Fnr binding site consensus sequence constitutes the target site for the FixK protein, which shares homology with Fnr (Kaminski *et al.* 1991). The -24/-12 element in the *A. caulinodans nifA* promoter is likely to be involved in mediating *ntr* regulation of *nifA* expression via the *ntrBC* and *ntrYX* systems (Pawlowski *et al.* 1987; 1991; de Bruijn *et al.* 1990) and to interact with the alternative σ factor NtrA (σ54; also designated as RpoN; see Kustu *et al.* 1989).

The DNA sequence of the -24/-12 promoter element of 64 σ^{54} controlled promoters from 22 different species has been compared and a consensus sequence derived (-27 cTGGCACGgcctTTGCA -11; Morett and Buck 1989; Kustu *et al.* 1989). Three residues of this consensus sequence are completely invariant, namely the GG pair at positions -25/-24 and the G at position -13. The only exception to the latter appears to be the *glnH* promoter of *E. coli* (Claverie-Martin and Magasanik, 1991). The -12 position is almost invariant, with the exception of some rhizobial promoters, where it is replaced by an A residue (see Merrick and Chambers, 1992 for a discussion and references). Mutagenesis of the -24/-12 promoter element has been carried out in several σ54 regulated promoters in different organisms: *K. pneumoniae nifH* and *nifLA* (Ow *et al.* 1985; Buck *et al.* 1985; Khan *et al.* 1986), *B. japonicum nifH* and *fixRnifA* (Kaluza *et al.* 1985; Thöny *et al.* 1987), *Caulobacter crescentus flbG* (Mullin and Newton 1989) and *S. typhimurium argTr* (Schmitz *et al.* 1988).

In order to test the function of the *A. caulinodans nifA* -24/-12 promoter element, site-specific mutagenesis of critical basepairs in this element was carried out. Specific mutations in the -24/-12 element were found to have a drastic effect on *nifA* promoter activity, suggesting that the *A. caulinodans nifA* promoter is σ^{54} dependent and *nifA* activation in this organism requires the presence of a NtrA(RpoN)-like protein.

3.3 MATERIALS AND METHODS

3.3.1 Bacterial strains, plasmids, and growth media

Bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* strains were grown at 37°C in LB medium (Miller 1972). *A. caulinodans* ORS571 strains were grown at 37°C in TY (Beringer 1974), YLS (Elmerich *et al.* 1982) or in LSO medium (Elmerich *et al.* 1982), supplemented with 40 mg/l of nicotinic acid and 0.1% of the desired nitrogen sources.

Table 1. Bacterial strains and plasmids used in this study

Strain	Relevant characteristics	Source or reference
A. caulinodans		
ORS571	wild type, Cb^r; Nif^+, Nod^+, Fix^+	Dreyfus *et al.* 1988
E. coli		
DH5αF'	F', Δ(*lacZYA*)	Hanahan 1983
HB101	Sm^r, recA	Boyer and Roulland-Dussoix 1969
MC1061	Sm^r, Δ(*lacZYA*)	Casadaban *et al.* 1983
Plasmids		
pUC118/pUC119	Ap^r; used for cloning and sequencing	Vieira and Messing 1987
pAP118	Ap^r; ORS571 *nifA* promoter region in pUC118	this work
pRK2013	Km^r, IncN, Mob^+, Tra^+, helper plasmid in mobilizations	Figurski and Helinski 1979
pLRSC1	Tc^r, ORS571 *ntrBC-ntrYX* region in pLAFR1	Pawlowski *et al.* 1987
pRSA13	Cm^r, ORS571 *nifA* region in pACYC184	Pawlowski *et al.* 1987
pNM481	Ap^r, polycloning-sites fused to 8^{th} codon of *lacZ*	Minton 1984
pPR54	Ap^r, Gm^r, ORS571 *nifA*::MudIIPR46' fusion cloned in pJRD184 for cointegration into the ORS571 chromosome.	Ratet *et al.* 1989
pPR57	Ap^r, Gm^r, *SalI-XhoI* deletion derivative of pPR54	P. Ratet
p481	Ap^r, Gm^r, *Eco*RI-*Cla*I fragment from pNM481, containing the polycloning-sites fused to 8^{th} codon of *lacZ*, in *Eco*RI-*Cla*I digested pPR57	this work
pJS4812	Ap^r, Gm^r, 5.9kb *Sma*I fragment of pLRSC1 in *Xba*I site of p481, providing a region of homology for cointegration into the ORS571 chromosome	this work

Table 1 - continued

Plasmids	Relevant characteristics	Source or reference
pJSwp	Ap^r, Gm^r, wildtype ORS571 *nifA* promoter region fused to 8th codon of *lacZ* in pJS4812 for cointegration into the ORS571 chromosome	this work
pJSn1	as pJSwp but G -> T at -24	this work
pJSn2	as pJSwp but C -> T at -16	this work
pJSn3	as pJSwp but G -> T at -13	this work
pJSn1+2	as pJSwp but G -> T at -24 and C -> T at -16	this work
pJSn1+3	as pJSwp but G -> T at -24 and G -> T at -13	this work
pJSn2+3	as pJSwp but C -> T at -16 and G -> T at -13	this work

Antibiotics were used at the following concentrations: For *E. coli:* ampicillin (Ap; 100 µg/ml), chloramphenicol (Cm; 30 µg/ml), gentamycin (Gm; 5 µg/ml), kanamycin (Km; 20 µg/ml) and tetracycline (Tc; 10 µg/ml); for ORS571: carbenicillin (Cb; 500 µg/ml), Gm (50 µg/ml), Km (200 µg/ml) and Tc (10 µg/ml). Indicator medium for *E. coli* or ORS571 strains containing *lac* gene fusions was supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal; 40 µg/ml).

3.3.2 DNA manipulations

Plasmid DNA was prepared by the alkaline lysis method described by Sambrook *et al.* (1989). Chromosomal DNA was prepared as described by Meade *et al.* (1982). Conditions used for DNA manipulations and transformations were as described by Sambrook *et al.* (1989). The enzymes used in these analyses were used according to the specifications of the manufacturers (Boehringer, Mannheim; Bethesda Research Laboratories, Maryland; New England Biolabs, Massachusetts).

3.3.3 Growth and β-galactosidase assays

β-galactosidase activity was measured as described by Miller (1972). *A. caulinodans* strains carrying *nif-lacZ* gene fusions were grown with continuous shaking in YLS medium, supplemented with 0.05% L-glutamine, at 37°C. An 0.2 ml aliquot of a saturated YLS culture was re-inoculated in 10 ml LSO medium,

supplemented with 0.05% L-glutamine and 0.05% ammonium sulfate. An 1.2 ml aliquot of the LSO culture was re-inoculated into 10 ml of LSO medium, and the culture was grown aerobically for 6 hr. The OD_{600} of the culture was measured and adjusted to 0.2 with LSO medium. For ammonium repression studies, an 1 ml aliquot was re-inoculated into 9 ml of LSO medium supplemented with 0.1% ammonium sulfate. For microaerobic derepression studies cultures were incubated in the presence of a 97%N_2/3%O_2 gas mixture in a GasPak System (Becton Dickinson and Co, Cockeysville, Md).

An 0.2 to 0.6 ml aliquot of the final cultures was mixed in a 1.5 ml cuvette with 0.5 to 0.8 ml Z-buffer, in order to obtain a final OD_{600} of 0.3 - 0.6, and the exact OD_{600} was measured. An 0.8 ml aliquot of diluted cells was transferred to a 1.5 ml microcentrifuge test tube, vortexed in the presence of 0.01 ml 0.1% sodium dodecyl sulfate (SDS) and 0.01 ml chloroform for 1 min and incubated for 20 min at 30°C. The reaction was initiated by the addition of 0.16 ml o-nitrophenyl-β-galactoside (ONPG; 4 mg/ml). After 5-60 min incubation at 30°C the reaction was stopped by adding 0.4 ml of 1 M Na_2CO_3. This mixture was centrifuged and the OD_{420} of the supernatant was measured using a Pharmacia Ultrospec III spectophotometer. The β-gal units shown in the figures were derived as described by Miller (1972).

3.4 RESULTS

3.4.1 Construction of *lacZ*-translational fusion vector pJS4812

The 598 bp *Eco*RI/*Cla*I fragment from pNM481 (Minton 1984), containing unique cloning sites fused to the 8th codon of *lacZ*, was inserted into plasmid pPR57 (digested with *Eco*RI/*Cla*I), giving rise to plasmid p481. pPR57 is *Sal*I-*Xho*I deletion derivative of pPR54 (Ratet *et al.* 1989). A 5.9 kb *Sma*I fragment of pLRSC1 (Pawlowski *et al.* 1987) was cloned in the *Xba*I-site of p481, in order to provide a region of homology with the *A. caulinodans* genome for homologous recombination, giving rise to plasmid pJS4812 (see fig. 1).

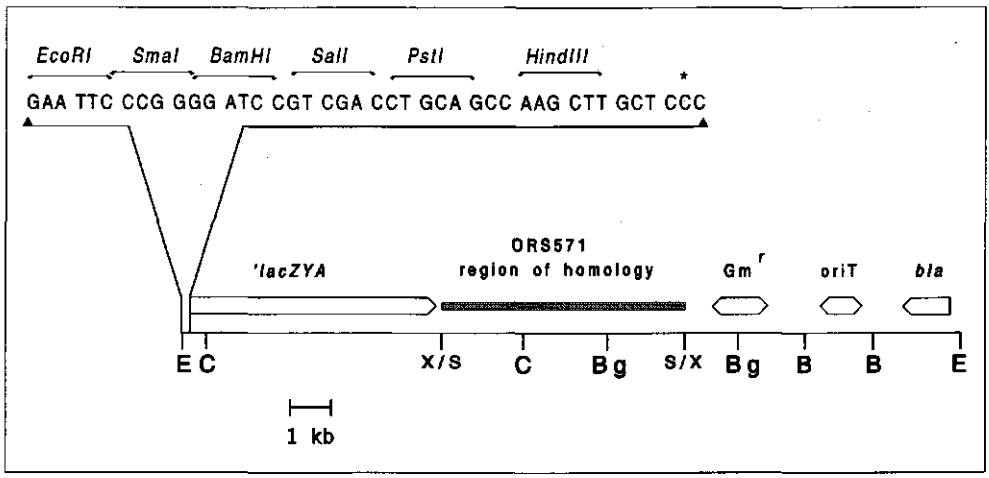

Figure 1. Structure of the integration vector pJS4812.

The positions of the Ap^r (*bla*) and Gm^r genes, the origin of transfer (oriT) and the truncated *lac* operon (*'lacZYA*) are indicated by open arrows. The region of DNA homology with the *A. caulinodans* chromosome (see text) is indicated by the hatched box. The translational phasing of the restriction sites in the polylinker (from pNM481) relative to the 8th codon of *lacZ* (indicated by an asterisk) is shown. The code used for restriction enzyme cleavage sites is as follows: E: *Eco*RI; S: *Sma*I; X: *Xba*I; B: *Bam*HI; Bg: *Bgl*II; C: *Cla*I.

3.4.2 Construction of *nifA:lacZ* translational fusion vector pJSwp

The *Xho*I/*Pst*I fragment from pRSA13 (Pawlowski *et al.* 1987), containing the *A. caulinodans nifA* 5' region and the coding region for the 43 N-terminal amino acid residues of NifA (Ratet *et al.* 1989), was isolated (see fig. 2). The *Xho*I sticky ends were rendered flush with the Klenow fragment of DNA Polymerase and the resulting fragment was cloned in the *Sma*I/*Pst*I site of pUC118, thereby reconstructing the *Xho*I-site, and giving rise to plasmid pAP118 (Table 1).

The *Eco*RI/*Hin*dIII fragment from pAP118 was cloned into pJS4812, thereby generating a translational fusion of *nifA* to the 8th codon of *lacZ*, under the control of the wild-type *nifA* promoter (see fig. 3). For expression studies in *A. caulinodans* ORS571, *nifA-lacZ* fusions were integrated into the ORS571 chromosome, via a single recombination event in the ORS571 DNA homologous region carried by the pJS4812 integration vector.

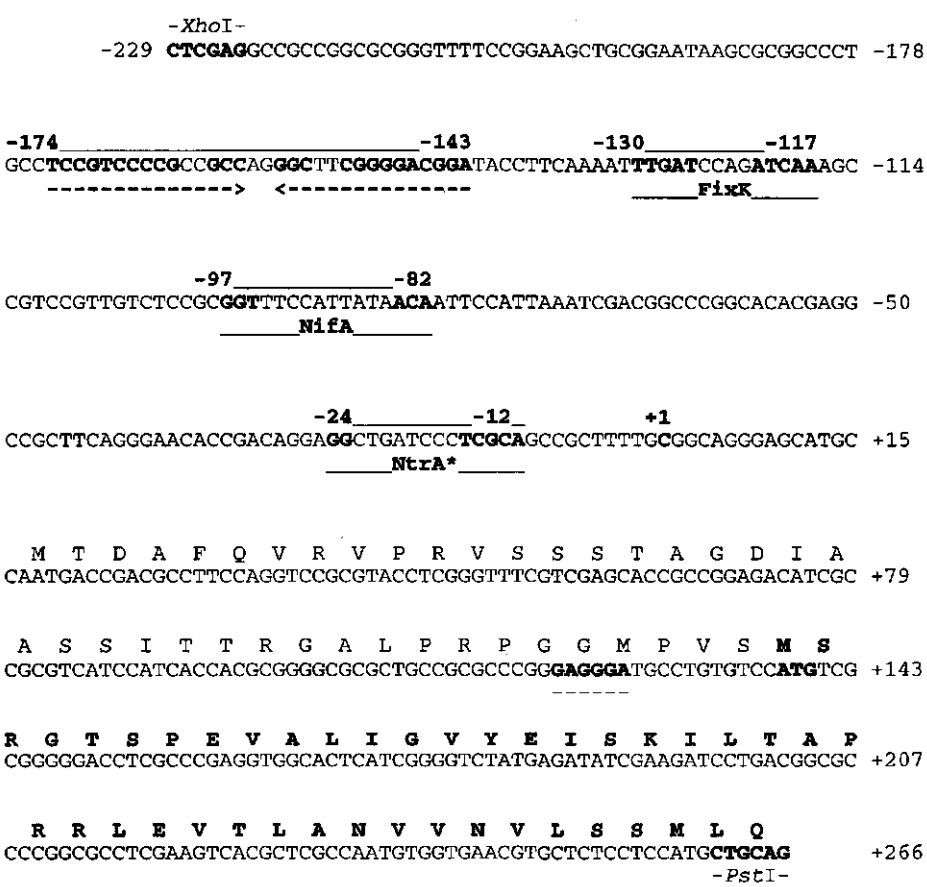

Figure 2. Structure of the *A. caulinodans nifA* promoter region.

Nucleotide sequence of the *Xho*I-*Pst*I fragment of pRSA13 (Ratet *et al.* 1989) containing the *nifA* promoter region and part of the NifA protein coding sequence. The position indicated as +1 is the site of transcription initiation. The horizontal arrows (position -174 to -143) indicate the position of inverted repeat structure. Conserved DNA elements found in oxygen (Fnr, FixK) regulated genes (TTGAT-n4-ATCAA at position -130 to -117), in NifA activated genes (UAS; GGT-n10-ACA at position -97 to -82) and in *ntrA* (*rpoN*, σ^{54}) dependent genes (GG-n8-TCGCA at position -25 to -11) are underlined and in bold typeface. The putative *nifA* ribosome-binding site (GAGGGA; broken line at position 121-126) and probable start codon and deduced protein product are also in bold typeface (from Ratet *et al.* 1989).

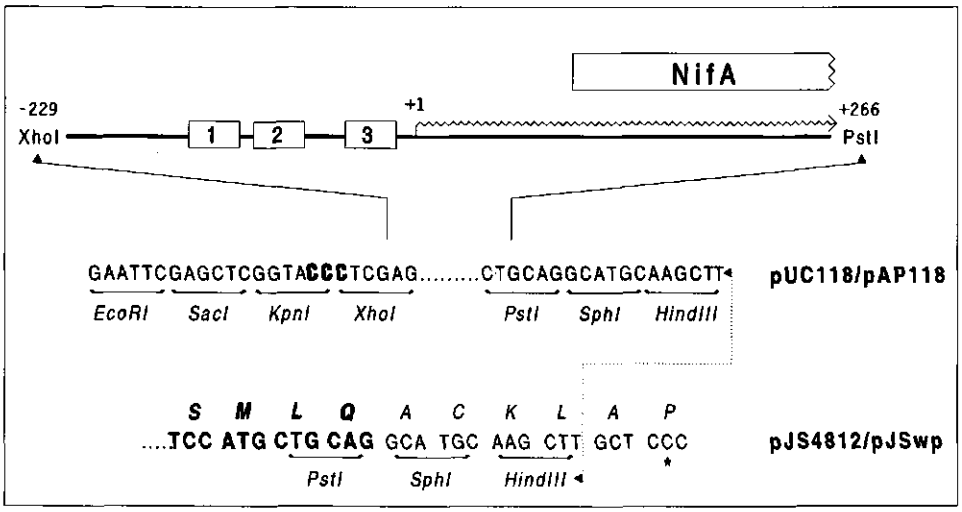

Figure 3. Construction of *nifA:lacZ* translational fusion.

The start point and direction of transcription of the *A. caulinodans nifA* gene are indicated by a wavy arrow and are based on data from Ratet *et al.* (1989). The small open boxes designate the Fnr binding-site consensus (box 1), NifA binding site consensus or Upstream Activating Sequence (box 2) and -24/-12 promoter element (box 3), as defined by Ratet *et al.* (1989). The *Sma*I half-site of pUC118 is shown in bold. The translational phasing of *nifA* (amino acids S, M, L, Q; shown in bold) relative to the 8th codon of *lacZ* (indicated by an asterisk) is as shown. See text for details.

3.4.3 Oligonucleotide-directed mutagenesis of the *A. caulinodans nifA* promoter and construction of mutant *nifA* promoter-*lacZ* fusions

A mixture of primers carrying the wild-type -24/-12 promoter sequence (5'-GCGGCTG$^{C}/_{A}$GA$^{G}/_{A}$GGATCAGC$^{C}/_{A}$TCCTGTCGGTG-3'), and derivatives thereof containing one or two mismatches, were synthesized and used together with the pUC sequencing primer in a primary PCR reaction with pAP118 (Table 1) as template DNA, in order to amplify mutant promoter fragments (PCRI). The amplified dsDNA-fragment, together with the reverse pUC sequencing primer of the vector, was used in a second PCR reaction to generate a complete *Xho*I-*Pst*I fragment containing the point mutations, as described by Kammann *et al.* (1989), flanked by an *Eco*RI and a *Hin*dIII site from the pUC118 polylinker (PCRII; see fig. 3).

The *Eco*RI/*Hin*dIII fragment of the PCRII product was cloned into pUC119. Mutants were analyzed by restriction digests, followed by sequencing analysis in order to verify the nature of the mutations. The mutant *Eco*RI/*Hin*dIII fragments were cloned into pJS4812, thereby generating a translational fusion of *nifA* to the 8th codon of *lacZ*, under the control of the wild-type or mutant *nifA* promoters (see fig. 2 and 3). For expression studies in *A. caulinodans* ORS571, these *nifA-lacZ* fusions were integrated into the ORS571 chromosome (see 3.4.2).

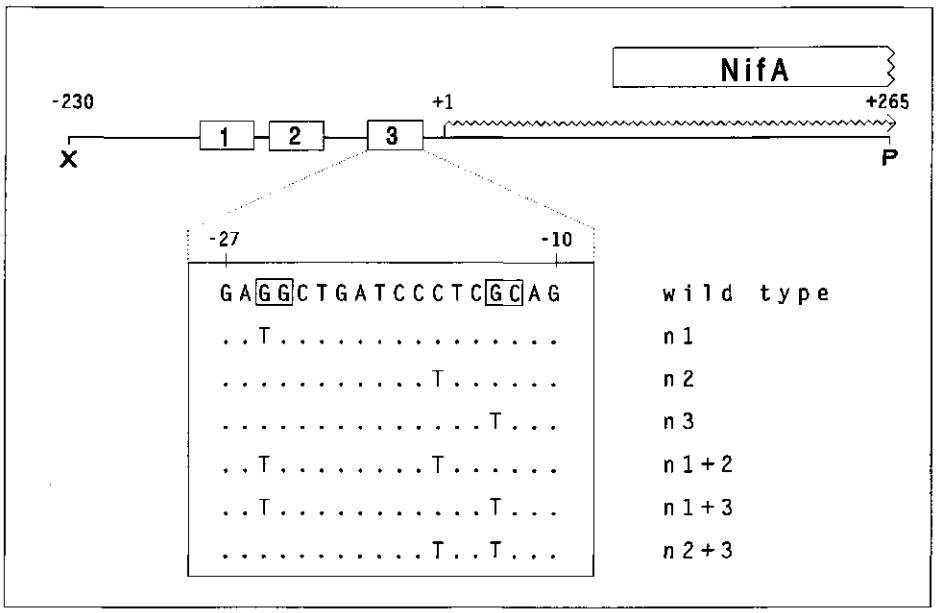

Figure 4. Structure of the *A. caulinodans nifA* locus and plasmids carrying mutations in the -24/-12 promoter element.

See figure 2. The nature of the mutations in the -24/-12 promoter element are shown in the large box. Restriction enzyme cleavage sites indicated are: X: *Xho*I and P: *Pst*I.

3.4.4 Effect of point mutations in the -24/-12 promoter element on *A. caulinodans nifA* gene expression

Specific single and double base-pair mutations were introduced in the -24/-12 element of the *nifA* promoter (fig. 4). The mutant *nifA* promoters were fused to the *lacZ* reporter gene and the resulting chimeric genes were integrated in the *A. caulinodans* chromosome using the integration vector pJS4812 (fig. 1).

The resulting strains were assayed for *nifA* promoter activity (β-galactosidase or β-gal activity) in cultures grown under different physiological conditions (aerobic or microaerobic; in the presence or absence of combined nitrogen sources; see Materials and Methods).

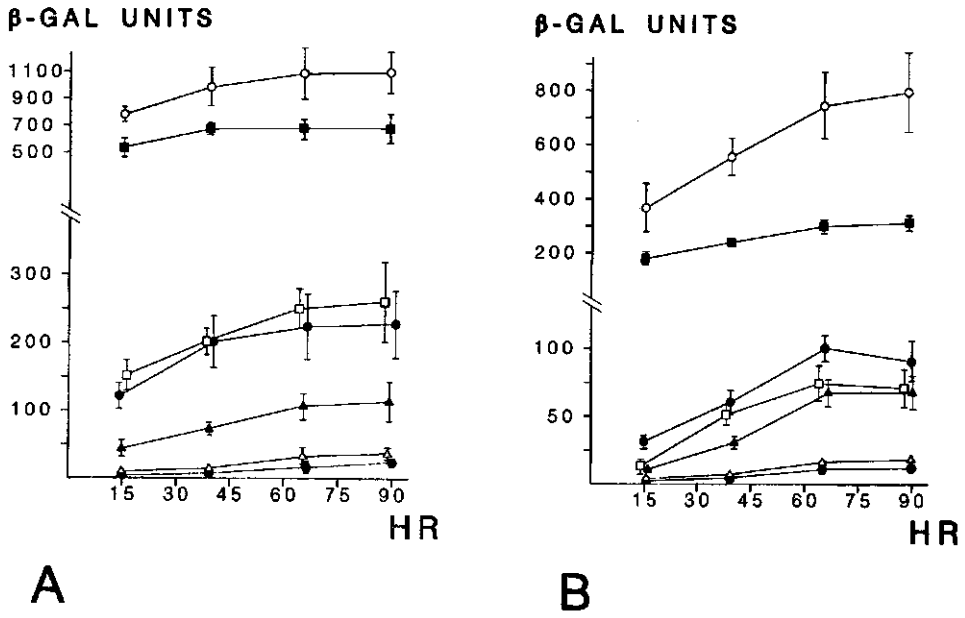

Figure 5. Expression of wild-type and mutant *nifA* promoters in *A. caulinodans* under different physiological conditions.

Panel A. β-gal levels of wild-type and mutant *nifA-lacZ* fusions in *A. caulinodans* under nitrogen fixing (derepressing) conditions (LSO medium without N-source; 3% O_2). The β-gal enzyme units (Miller 1972) are shown on the Y-axis and the number of hours after inoculation are shown on the X-axis.

Panel B. β-gal enzyme levels of wild-type and mutant *nifA-lacZ* fusions in *A. caulinodans* under ammonium-repressing conditions (LSO medium with 0.1% ammonium sulfate; 3% O_2). The open circles denote the strain carrying the n1+2 mutant fusion, the closed boxes the n2 mutant, the open boxes the wild-type promoter, the closed circles the n2+3 mutant, the closed triangles the n1 mutant, the open triangles the n3 mutant and the closed diamonds the n1+3 mutant. The results in Panels A and B represent an average of at least three independent experiments.

Under nitrogen-fixing conditions [microaerobic in the absence of ammonium ($NH4^+$)], a single base-pair change (G to T; Mutation n1) at position -25 or a G to T change at position -13 (Mutation n3; fig. 4) resulted in a substantial reduction (n1: 3 fold) or virtual abolishment (n3: 11 fold reduction) of *nifA* expression (fig. 5). The n1+3 double mutation lead to the reduction of *nifA* expression to a background level (20 fold reduction; fig. 5). Interestingly, a mutation at position -16 (C to T; Mutation n2; fig. 4) resulted in a considerable (3 fold) increase of *nifA* promoter activity relative to the wild-type construct (fig. 5). A similar phenotype was observed for a n1+2 double mutation (5 fold increase of *nifA* expression), while a n2+3 double mutation was observed to have a wild-type level of *nifA* promoter activity (fig. 5). These results suggested that the n2 and n1+2 mutations may have generated a (partially) constitutive, σ^{54} independent promoter.

In order to examine this question, the n1+2 mutant promoter-*lacZ* fusion (pJSn1+2) was also introduced into $NtrA^+$($RpoN^+$; DH5αF') and $NtrA^-$($RpoN^-$; TH1, Table 1) *E. coli* strains and β-gal levels were measured. The n1+2 mutant promoter was found to direct a significant level of reporter gene expression, in the presence or absence of a functional *ntrA(rpoN)* gene (data not shown).

The wild-type *nifA* promoter was found to be not expressed in *E. coli* and the n2 mutant was expressed to a low level (data not shown). The effect of repressing concentrations of ammonium on the expression of the wild-type and mutant *nifA* promoters was examined. The wild-type *nifA* promoter was expressed to a ~10 fold lower level at 15 hrs and a ~4 fold lower at 40 hrs of growth in medium containing ammonium (fig. 5). The n3 and n1+3 mutant *nifA* promoters, which already showed very low levels of expression under nitrogen-fixing conditions (-$NH4^+$), were expressed at even lower levels in the presence of ammonium (fig. 5). The activity of the n1, n2, n2+3 and n1+2 mutant promoters was repressed to a lesser degree than the wild-type promoter in the presence of ammonium. This was particularly clear at 15 hrs after inoculation, at which time the n1 and n2+3 mutant promoters were repressed ~4-5 fold (as compared to 10 fold in the case of the wild-type) and the n2 and n1+2 mutant promoters only ~2-3 fold (fig. 5). These results suggest that the conserved $_{-25}$GG-n10-GC$_{-12}$ motif in the *nifA* promoter is important for transcriptional activation, as well as ammonium repression, and that by the generation of a double mutation at positions -25 and -16 a highly expressed, relatively nitrogen-regulation-independent promoter was generated.

The effect of repressing concentrations of O_2 on *nifA* promoter activity was also examined. Incubating the cultures aerobically, in the absence of ammonium, resulted in a severe reduction of expression of the wild-type promoter, the n1, n2, n1+2 and n2+3 mutant promoters, as well as a further reduction of expression of the already poorly expressed n3 and n1+3 mutant promoters (data not shown). In the presence of ammonium under aerobic conditions, the activity of the wild-type, n3 and n1+3 mutant promoters was reduced to background levels (~ 30-fold reduction), the n1 and n1+2 mutant promoters showed a 3-5 fold reduction of activity at 40 hrs after inoculation (relative to the activity under nitrogen-fixing conditions). The n2 and n2+3 mutant promoter showed a 10-12 fold and reduction in activity (data not shown). The results of these experiments are summarized in Table 2 .

Table 2. Expression of wildtype (wp) and mutant (n) *nifA:lacZ* fusions in wild-type *A. caulinodans* at 40 hr after inoculation

	-NH_4 3% O_2	-NH_4 20% O_2	+NH_4 3% O_2	+NH_4 20% O_2
wp	++	--	-	--
n1	+-	--	-	-
n2	+++	-	++	+-
n3	-	--	--	--
n1+2	++++	-	+++	++
n1+3	--	--	--	--
n2+3	++	--	+-	-

-- : less than 10 units of β-gal activity
- : 10 - 50 units
+- : 50 - 100 units
++ : 200 - 500 units
+++ : 500 - 10000 units
++++ : 1000 or more units

3.5 DISCUSSION

3.5.1 Oligonucleotide-directed mutagenesis of the *A. caulinodans nifA* promoter

The first screening for mutants was carried out by restriction digest analysis. The n1 mutation created a new *Fok*I-site (and caused the loss of a *Mnl*I-site), the n2 mutation caused the loss of an *Mnl*I-site and the n3 mutation created a new *Dde*I-site. After selection, the mutants were analyzed by DNA sequencing to verify the mutation and check for additional changes. During this study it was found that a large fraction (80-90%) had acquired an additional bp change (usually an A) at the end of the primer sequence used in the first PCR reaction (PCR1). This is most likely due to the non-template directed nucleotide addition at the 3'end of the molecules during amplification (Clark 1988). In the second PCR reaction (PCR2) this additional base is likely be used as a template to incorporate the first base in elongation, resulting in an additional mutation. To avoid this problem the PCR1 products should have been treated with the Klenow fragment of DNA polymerase, to render the DNA molecules flush, as discussed by Hemsley *et al.* (1989).

3.5.2 Effect of mutations in the -24/-12 promoter element on *A. caulinodans nifA* expression

A mutation at position -25 (n1) in the *A. caulinodans nifA* promoter, results in a significant reduction of promoter activity ("down" promoter phenotype). However, the down phenotype is not as severe as found in the case of analogous mutations in the -24 and/or -25 position of the *K. pneumoniae nifH* (Ow *et al.* 1985), the *K. pneumoniae nifL* (Khan *et al.* 1986) and *B. japonicum nifH* (Kaluza *et al.* 1985) promoters or in the promoters of the *Caulobacter crescentus flbG* (Mullin and Newton 1989) and *S. typhimurium argTr* (Schmitz *et al.* 1988) genes. Interestingly, an analogous mutation in the -23 G residue of the *B. japonicum fixRnifA* promoter has no effect on *nifA* expression (Thöny *et al.* 1987).

A mutation in residue -13 of the *A. caulinodans nifA* promoter (n3) essentially abolishes *nifA* expression. This effect has also been observed for analogous mutations in the -13 and/or -12 position of the *K. pneumoniae nifH* and *nifLA* promoters (Ow *et al.* 1985; Buck *et al.* 1985; Khan *et al.* 1986), the *B. japonicum fixRnifA* promoter (Thöny *et al.* 1987) and the promoters of the *C. crescentus flbG* and the *S. typhimurium argTr* genes (Mullin and Newton 1989;

Schmitz *et al.* 1988). These results strongly suggest that the *A. caulinodans nifA* promoter is σ^{54} dependent and therefore supports the model that *nifA* activation in this organism requires the presence of a NtrA(RpoN)-like protein (NtrA*,σ^{54}*; see chapter 8).

A mutation in the -16 position of the *A. caulinodans* promoter (n2) results in a substantial increase of *nifA* expression ("up" promoter phenotype) and a severe decrease in repression by ammonia (partially constitutive phenotype). Based on results obtained with mutations (C to T transitions) in the -17 to -15 residues of the *K. pneumoniae nifH* promoter, it has been suggested that the presence of T-rich stretches in this region may result in the formation of a stronger NtrA(RpoN)-RNA polymerase-promoter complex (Buck and Cannon 1989; 1992a; Morett and Buck 1989; Cannon and Buck 1992), possibly due to a more efficient recognition of the target by the NtrA(RpoN)-RNA polymerase (Whitehall *et al.* 1992). The relevance of the -14 to -17 residues for σ^{54} (NtrA/RpoN) binding has also been demonstrated by Buck and Cannon (1992b), who have suggested that the methyl groups in the DNA major groove of this region are important for binding. This may also explain the up promoter phenotype of the n2 mutation.

The extreme up phenotype of the n1+2 double mutation suggests that the -24 part of the NtrA(RpoN)-RNA polymerase recognition site does not play a significant role in the "up" expression pattern observed with the n2 single mutation. However, the diminished up phenotype of the n2+3 double mutation suggests that the other half of the -24 to -12 motif (G residue at position -13) is involved in the n2 expression pattern. Since the increased expression levels of the n2 and n1+2 mutant promoters can no longer be fully repressed by nitrogen or oxygen, it is also possible that the *A. caulinodans nifA* promoter is subject to both activation under microaerobic, N-limited conditions (involving the $_{-25}$GG$_{-24}$ and the $_{-13}$GC$_{-12}$ residues), as well as repression under aerobic N-excess growth conditions (involving the $_{-16}$C residue). Another, perhaps more simple, explanation may be that the n2 and n1+2 mutations have generated a σ^{54} (NtrA/RpoN) independent promoter, since the TTcgCA motif around position -35 now resembles the DNA sequence recognized by a σ^{70}-RNA polymerase. This is supported by the finding that expression of the n2 and n1+2 mutant promoters in *E. coli* appears to be *ntrA(rpoN)* independent.

In conclusion the -24/-12 promoter element seems to be important for *nifA* gene expression and suggest the involvement of a σ54 (NtrA; RpoN)-type transcription factor in *nifA* gene regulation; This prompted a search for (a) *A. caulinodans ntrA*(*rpoN*)-like gene(s) (see chapter 4).

CHAPTER 4

CLONING AND CHARACTERIZATION OF AN *A. CAULINODANS NTRA(RPON; σ54)* GENE

4.1 ABSTRACT

Using a heterologous *ntrA(rpoN)* gene probe, an *A. caulinodans ntrA(rpoN)*-like gene was cloned and the DNA sequence of this gene and flanking regions was determined. The presence of 3 open reading frames (ORF1-3) was demonstrated. ORF2 was found to contain regions sharing a high degree of homology with all characterized bacterial *ntrA(rpoN)* genes. ORF1 was found to share homology with ORF's found upstream of other bacterial *ntrA(rpoN)* genes, which have been postulated to encode members of a superfamily of ATP-binding proteins. Transposon Tn*5* insertion mutations were introduced into the cloned *ntrA(rpoN)* gene and chromosomal *ntrA(rpoN)*::Tn*5* *A. caulinodans* mutants were created. The resulting mutants were found to be unable to fix nitrogen in the free-living state (Nif$^-$ in culture) or in stem- or root nodules induced on *Sesbania rostrata* (Fix$^-$ *in planta*), and to be unable to grow aerobically in the presence of nitrate as sole nitrogen source (Ntr$^-$). A *nifH-lacZ* gene fusion was found to be silent in *ntrA(rpoN)*::Tn*5* mutant strains, but a *nifA-lacZ* gene fusion was found to be expressed at a wild-type level, suggesting that the *ntrA(rpoN)* gene identified here controls the expression of some of the *A. caulinodans nif* genes, but not the central *nif*-regulatory gene *nifA*.

Part of this work has been published in: *Azorhizobium caulinodans* nitrogen fixation (*nif/fix*) gene regulation: Mutagenesis of the *nifA* -24/-12 promoter element, characterization of a *ntrA(rpoN)* gene, and derivation of a model. John Stigter, Maria Schneider, and Frans J. de Bruijn. Molecular Plant-Microbe Interactions (1993) Vol. 6, No. 2, pp. 238-252.

4.2 INTRODUCTION

The promoter specificity of bacterial RNA polymerases is determined by the sigma subunit of the holoenzyme. Binding of the RNA polymerase holoenzyme $\alpha_2\beta\beta'\sigma$ to a promoter sequence results in the formation of a closed complex, which is then converted to an open complex, characterized by local melting of the double-stranded DNA. Once transcription is initiated, the σ factor is released from the complex and the core enzyme continues with RNA synthesis (Helmann and Chamberlin 1988).

The identification, characterization, and sequence analysis of a large number of sigma factors have revealed that they fall into two broad classes. One family is similar to the originally identified *E. coli* σ70 subunit; the other is similar to the 54-kDa *E. coli* σ subunit (Sasse-Dwight and Gralla 1990; Kustu *et al.* 1989; Lonetto *et al.* 1992).

In *E. coli* σ70 (*rpoD*), the most abundant σ factor, is required for the expression of many housekeeping genes. It allows recognition of and transcription from -35/-10 promoters (TTGACA-n16/19-TATAAT). Several alternative factors, with a homologous structure to σ70, have been identified in enteric bacteria which enable the cells to transcribe specific sets of genes in response to environmental stimuli. Examples include sigma factors which are not essential for exponential cell growth in *E. coli* (*SigS*, *KatF*) and *Streptomyces coelicolor* (*HrdA*, *HrdC*, and *HrdD*); heat shock sigma factors from *E.coli* (*rpoH*) and *Citrobacter freundii* (*HtpR*); flagellar sigma factors from *S. typhimurium* (*FliA*) and *B. subtilis* (*SigD*); sporulation sigmas from *Bacillus* , *S. coelicolor* (*WhiG*) and *Myxococcus xanthus* (*SigB*) and are reviewed in Lonetto *et al.* (1992).

The other class of sigma factors represents σ54 (*rpoN*, *ntrA*, *glnF*), which is presently the only member, but has been found in many micro organisms. Many features of σ54 indicate that it is neither structurally nor functionally closely related to the σ70 family. While other alternative sigma factors direct transcription of genes whose products share a common physiological role, RNA polymerase containing σ54 (Eσ54) has been shown to direct the transcription of genes whose products have diverse physiological functions. *ntrA* (*rpoN*, σ54) was originally identified in enteric bacteria as a component of global regulation by nitrogen (Magasanik 1982). Several operons involved in nitrogen metabolism such as those containing the genes for nitrogen fixation in *K. pneumoniae*, and glutamine synthetase, are subject to control by *ntrA* (*rpoN*; σ54; Ausubel 1984; Dixon 1984).

Non-nitrogen regulated genes such as the genes for formate dehydrogenase and hydrogenase isoenzyme 3 of *E. coli* are also subject to *ntrA* (*rpoN*, σ54) control (Birkmann *et al.* 1987). *NtrA*(*rpoN*, σ54)-dependent genes with various physiological functions have been identified in other gram-negative bacteria of the genera *Azotobacter*, *Azospirillum*, *Rhizobium*, *Bradyrhizobium*, *Rhodobacter*, *Alcaligenes*, *Pseudomonas*, *Neisseria*, *Caulobacter* (reviewed in Kustu *et al.* 1989), *Thiobacillus* (Berger *et al.* 1990) and *Agrobacterium* (Wu *et al.* 1992) and the gram-positive bacterium *Bacillus* (Debarbouille *et al.* 1991).

Promoters recognized by RNA polymerase containing σ54 show characteristic sequence motifs around positions -24 and -12 relative to the start of transcription (5'-CT**GG**CAC-n5-TT**GC**A-3'; Beynon *et al.* 1983). In all cases investigated, transcription from these promoters was shown to depend on a specific activator protein whose activity was modulated by physiological signals and it has been suggested that the expression of σ54-dependent genes occurs by an common mechanism of transcription initiation (Kustu *et al.* 1989). Eσ54 bound to promoter DNA can isomerize from the closed to the open complex only in the presence of an activator protein that binds to specific sites upstream of Eσ54 and has an ATPase activity necessary for open complex formation (Austin and Dixon 1992; Popham *et al.* 1989; Weiss *et al.* 1991).

ntrA (*rpoN*, σ54)-like genes or the corresponding gene products have been identified genetically or functionally in the following bacteria: *E. coli* (Magasanik 1982; Sasse-Dwight and Gralla 1990), *S. typhimurium* (Hirschman *et al.* 1985; Popham *et al.* 1991), *K. pneumoniae* (de Bruijn and Ausubel 1983; Merrick and Gibbins 1985), *Azotobacter vinelandii* (Merrick *et al.* 1987), *Pseudomonas putida* (Inouye *et al.* 1989; Kohler *et al.* 1989a), *P. aeruginosa* (Ishimoto and Lory 1989; Totten *et al.* 1990), *P. facilis* (Romermann *et al.* 1989), *Alcaligenes eutrophus* (Romermann *et al.* 1989; Warrelmann *et al.* 1992), *Rhizobium meliloti* (Ronson *et al.* 1987a; Shatters *et al.* 1989), *Rhizobium* sp. strain NGR234 (Stanley *et al.* 1989; van Slooten *et al.* 1990), *Rhodobacter capsulatus* (Alias *et al.* 1989; Jones and Haselkorn 1989), *R. sphaeroides* (Meijer and Tabita 1992), *Thiobacillus ferrooxidans* (Berger *et al.* 1990), *Bradyrhizobium japonicum* (Kullik *et al.* 1991), *Agrobacterium tumefaciens* (Wu *et al.* 1992), *B. subtilis* (Debarbouille *et al.* 1991) and *Caulobacter crescentus* (Brun, unpublished; GenBank X68549).

Consistent with the finding that, within a given species, -24/-12 promoters are associated with functionally unrelated genes, many *ntrA* (*rpoN*, σ54) mutants show a pleotropic phenotype.

A. caulinodans nifH1, *nifH2* and *nifA* are preceded by -24/-12 promoters (Norel and Elmerich 1987; Nees *et al.* 1988; Ratet *et al.* 1989). The *nifA* promoter is partly regulated by NtrC/NtrX, whereas the *nifH1* gene is regulated by NifA (see section 2.3.7), indicating the involvement of a *ntrA* (*rpoN*, σ54)-like protein. In addition mutations in the -24/-12 promoter element of *A. caulinodans nifA* strongly suggested involvement of *ntrA* (*rpoN*, σ54) in transcriptional activation (see chapter 3).

To further analyze the regulation of *nif* genes in *A. caulinodans*, a *ntrA* (*rpoN*, σ54)-like gene was identified, characterized and sequenced. The predicted NtrA(RpoN) protein sequence was found to be highly homologous to the NtrA(RpoN) proteins of other bacteria. *A. caulinodans ntrA*(*rpoN*)::Tn*5* mutants showed a pleiotropic phenotype since they were not only unable to fix nitrogen in the free-living state or in nodules induced on *Sesbania rostrata*, but were also affected in nitrate assimilation, transport of C_4-dicarboxylates, and nodulation. Transcription of a *nifH-lacZ* fusion was shown not to be activated in the *ntrA* (*rpoN*) mutant, however a *nifA-lacZ* fusion was found to be expressed at wild-type level, suggesting the presence of a second *ntrA(rpoN)* gene in *A. caulinodans* with a distinct specificity, involved in *nifA* regulation (see chapter 8).

4.3 MATERIALS AND METHODS

4.3.1 Bacterial strains, plasmids, and growth media

Bacterial strains and plasmids used in this study are listed in Table 1. For growth media and antibiotics, see chapter 3.

4.3.2 DNA manipulations

See chapter 3.

Table 1. Bacterial strains and plasmids used in this study

Strain	Relevant characteristics	Source or reference
A. caulinodans		
ORS571	wild type, Cb^r; Nif^+, Nod^+, Fix^+	Dreyfus *et al.* 1988
ORS571N136-1c	*ntrA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	this work
ORS571N15-2c	*ntrA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	this work
ORS571N136-3c	*ntrA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	this work
R. meliloti		
1680	*ntrA1*::Tn*5*, Str^r, Nm^r	Ronson *et al.* 1987a
E. coli		
DH5αF'	F', Δ(*lacZYA*)	Hanahan 1983
HB101	Sm^r, recA	Boyer and Roulland-Dussoix 1969
TH1	Δ*lacU*169,Δ*glnF*	Hunt and Magasanik, 1985
Plasmids		
pNtr3.5BE	Ap^r, pUC8 derivative, *R. meliloti ntrA*	Ronson *et al.* 1987a
pUC118/pUC119	Ap^r; used for cloning and sequencing	Vieira and Messing 1987
pRK2013	Km^r, IncN, Mob^+, Tra^+, helper plasmid in mobilizations	Figurski and Helinski 1979
pRK290	Tc^r, IncP, Mob^+, Tra^-, broad host range cloning vector	Ditta *et al.* 1980
pWB5	Tc^r, Km^r, pRK290 derivative	W. Buikema and F. Ausubel
pPH1JI	Cm^r, Sm^r, Sp^r, Gm^r, IncP, Tra^+	Hirsch and Beringer 1984
pLAFR1	Tc^r, IncP, Mob^+, Tra^-, broad host range cosmid cloning vector	Friedman *et al.* 1982
pJRD184	Tc^r, Ap^r	Heusterspreute *et al.* 1985
pPR3408	Tc^r, Cm^r, Ap^r, ORS571 *nifHD:lacZ* in pRK290	Pawlowski *et al.* 1987
pBS71	Tc^r, ORS571 genomic DNA (*ntrA* region) in pLAFR1	this work
pBS714	Tc^r, 2.7kb *Sma*I fragment of pBS71 in pJRD184	this work
pNtrA2/pNtrA4	Ap^r, 2.7kb *Sma*I fragment of pBS714 in pUC119	this work
pJSwp	Ap^r, Gm^r, wildtype ORS571 *nifA* promoter region fused to 8th codon of *lacZ* in pJS4812 for cointegration into the ORS571 chromosome	this work, chapter 3

4.3.3 Southern blotting and colony hybridizations

Plasmid pNtr3.5EB (Ronson *et al.* 1987a) was used as the source for the *R. meliloti ntrA(rpoN)* DNA probe. The 3.5 kb *Eco*RI-*Bam*HI fragment was isolated and radioactively labeled by nick translation. Southern blotting and nick translations were carried out as described (Sambrook *et al.* 1989). Hybridizations were carried out in the presence of 50% formamide at 42°C and the blots washed at 68°C in 2x SSC (1x SSC contains 0.15M NaCl, 15mM Na-citrate, pH 7) buffer, containing 0.1% sodium dodecylsulfate (SDS). Colony hybridizations were carried out by using Whatman 541 filter paper, as described by Maas (1983).

4.3.4 Transposon Tn*5* mutagenesis of the *ntrA(rpoN)* gene

Transposon Tn*5* mutagenesis was carried out as described by de Bruijn and Lupski (1984) and de Bruijn (1987).

4.3.5 Conjugation and gene replacement experiments

Plasmids were mobilized from *E. coli* to *A. caulinodans* using the helper plasmid pRK2013, as described by Ditta *et al.* (1980). Gene replacement experiments with Tn*5* mutagenized regions were carried out as described (Ruvkun and Ausubel 1981; de Bruijn 1987). *NtrA(rpoN)*::Tn*5* mutants resulting from a double cross-over event were identified by examining the (absence of) proper growth of the transconjugants on LSO medium with nitrate as sole N-source. The pPH1JI plasmid was cured from the *ntrA(rpoN)*::Tn*5* strains by introducing pRK290 (selecting for Tc^r transconjugants) and loss of pRK290 was achieved by serial culture in the absence of Tc and selection for Tc^s colonies. The resulting strains were labelled with a "c" (e.g. ORS571N136-1c; Table 1).

4.3.6 Plant experiments

Nodulation and symbiotic nitrogen fixation ability was tested on the root and stem of *S. rostrata*, in test tubes and Leonard Jars. For test tube experiments *S. rostrata* seeds were surface sterilized by immersion in 96% sulfuric acid for 1 hr, rinsed in sterile water and transferred to slants of B + D agar (Broughton and Dilworth 1971) in 30 cm glass tubes. The seeds were germinated under artificial light at 21°C (16 hr day) and after 5-10 days infected with 1 ml of a 1:6 dilution in

H_2O of an overnight culture of bacteria grown in TY with antibiotics. Leonard Jar experiments were carried out as described by Broughton and John (1979) using 4 jars per strain. Plants were examined 7, 12 and 15 weeks after inoculation. Stem infections were performed by rubbing an overnight bacterial culture onto the stems of 6- to 8-week old *S. rostrata* plants using sterile cotton tips. The infected plants were grown in a conviron CMP2023 growth chamber under a 12 hr light cycle at 24°C. Nitrogenase activity (acetylene reduction assay) was determined as described by Pawlowski *et al.* (1987).

4.3.7 Growth and β-galactosidase assays

See chapter 3.

4.3.8 DNA sequencing

The dideoxynucleotide chain termination method using [^{35}S]dATP and Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio) was used for DNA sequence determination. Constructs used for DNA sequencing were derived from pNtrA2 or pNtrA4, which carry a 2.7 kb *Sma*I fragment cloned in both orientations in pUC119 (Table 1). The DNA sequence was obtained from nested deletion derivatives, generated with exonuclease III/mung bean nuclease (Ausubel *et al.* 1989) and from restriction fragments cloned into pUC119 and pUC118.

4.4 RESULTS

4.4.1 Cloning and characterization of an *A. caulinodans ntrA(rpoN)* locus

In order to identify the *trans*-acting factor(s) involved in *ntr* control via the -24/-12 element in the *nifA* promoter, we searched for the presence of (a) *ntrA(rpoN)*-like gene(s) in *A. caulinodans*, using the cloned *R. meliloti ntrA(rpoN)* locus (from plasmid pNtr3.5BE; Ronson *et al.* 1987a) as heterologous DNA probe. The 3.5 kb *Bam*HI-*Eco*RI insert of this plasmid was purified, labelled and used as a hybridization probe for a Southern blot carrying *Eco*RI digested *A. caulinodans* DNA and for a colony bank of *A. caulinodans* DNA constructed in the vector pLAFR1 (Pawlowski *et al.* 1987).

Three hybridizing *Eco*RI fragments of *A. caulinodans* DNA were observed in the Southern blot (~20, 3.6 and 1.5 kb in length) and 7 positive colonies were identified in the colony bank (out of 700 colonies screened; data not shown). The latter 7 colonies were shown to contain cosmids with overlapping segments of *A. caulinodans* DNA and one cosmid carrying all three (contiguous) hybridizing fragments was selected for further studies (pBS71). This cosmid was introduced into the *R. meliloti ntrA*(*rpoN*):Tn*5* mutant 1680 via conjugation and found to be able to complement the Ntr$^-$ phenotype of strain 1680 (Ronson *et al.* 1987a), as evidenced by restoration of growth on minimal LSO plates with potassium nitrate (0.2%) as sole nitrogen source (data not shown).

Using Southern blotting the region of homology with the *R. meliloti ntrA(rpoN)* locus was narrowed down to a 2.7 kb *Sma*I fragment of pBS71 and this fragment was subcloned into plasmid pJRD184 (Table 1) to form plasmid pBS714 (fig. 1).

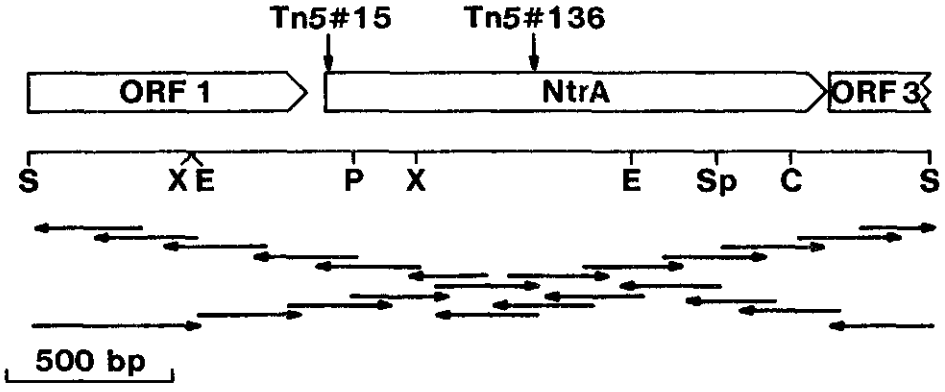

Figure 1. Physical and genetic map of the *A. caulinodans ntrA*(*rpoN*) locus and DNA sequencing strategy employed.

The extent and direction of the protein coding regions are shown by open arrows. The positions of the Tn*5* insertions are shown by vertical arrows. The restriction enzyme cleavage sites indicated are: S: *Sma*I; X: *Xho*I; E: *Eco*RI; P: *Pst*I; Sp: *Sph*I; C: *Cla*I. The horizontal arrows show the extent of the fragments sequenced on both strands.

4.4.2 Creation of transposon Tn*5* insertion mutants in the *A. caulinodans ntrA(rpoN)*-like locus.

Plasmid pBS714 was mutagenized with Tn*5*, as described by de Bruijn and Lupski (1984), and two independent Tn*5* insertions within the 1.5 kb *Eco*RI fragment, showing the highest degree of homology with the *R. meliloti ntrA(rpoN)* probe, were selected (Tn*5*#15 and Tn*5*#136; figure 1). These insertions were used for gene-replacement experiments (see de Bruijn 1987) and the position of the Tn*5*'s in the chromosomal *A. caulinodans ntrA(rpoN)*-like locus was verified by Southern blotting (fig. 2).

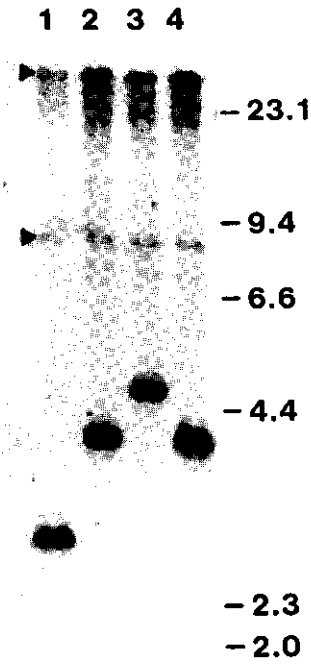

Figure 2. Mapping of the positions of Tn*5* insertions within the chromosomal DNA of *A. caulinodans*.

Southern blot of *Sma*I digested DNA, hybridized with a radioactive labeled synthetic oligonucleotide, corresponding to a highly conserved domain of NtrA(RpoN) proteins (RpoN-box; see 4.4.5).

Lane 1: *A. caulinodans* ORS571 wild-type

Lane 2: ORS571N136-1c (*ntrA(rpoN)*::Tn*5*)

Lane 3: ORS571N15-2c (*ntrA(rpoN)*::Tn*5*)

Lane 4: ORS571N136-3c (*ntrA(rpoN)*::Tn*5*)

Arrows indicate possible additional copies of *ntrA(rpoN)* in *A. caulinodans* (see below).

The *ntrA(rpoN)*::Tn*5* mutant strains were examined for their ability to fix nitrogen in culture and on plates (Nif phenotype), and in stem- and root nodules induced on *S. rostrata* (Nod and Fix phenotypes), growth on nitrate or aminoacids as sole N-source (Ntr phenotype), as well as growth on dicarboxylic acids as sole C-source (Dct phenotype). The results are summarized in Table 2. The insertion mutants #15 and #136 were Nif^- (see also below), Nod^+, Fix^-, Ntr^- (assimilatory nitrate reduction deficient) and Dct^-, but grew normally on arginine, histidine and proline as sole nitrogen sources. This phenotype closely resembled that found for *ntrA(rpoN)*::Tn*5* mutants of *R. meliloti* (Ronson *et al.* 1987a) and suggested that pBS714 carried an *A. caulinodans ntrA(rpoN)*- equivalent gene.

Table 2. Phenotypes of wild-type and *ntrA(rpoN)* mutant strains of *A. caulinodans*

	wild-type	*ntrA*::Tn*5*#15 *ntrA*::Tn*5*#136
Nif	+	-
Fix	+	-
Nod	+	+(a)
Ntr (0.1% KNO_3)	+	-
Dct	+(b)	-(b)
Aut	+/-(c)	+/-(c)
Hut	+(d)	+(d)
Put	+(e)	+(e)

(a) : numerous, light green, small nodules (see Pawlowski *et al.* 1987)
(b) : growth on 20 mM L-succinate, L-fumarate or L-malate as sole C-source.
(c) : growth on 15 mM L-arginine; +/- indicates poor but significant growth.
(d) : growth on 15 mM L-histidine
(e) : growth on 15 mM L-proline

4.4.3 DNA sequence analysis of the *A. caulinodans ntrA(rpoN)*-equivalent locus

The DNA sequence of the 2.7 kb *Sma*I fragment of pBS714 (fig. 1) was determined and the results are shown in fig 3. This analysis revealed the presence of three open reading frames (ORF1, position 7-849; ORF2, position 916-2457; partial ORF3, position 2470-2674; figures 1 and 3). The deduced polypeptide from ORF2 was found to share significant domain homology with NtrA(RpoN) proteins from other bacteria, including *R. meliloti* (see figure 4), and ORF2 was therefore designated as an *A. caulinodans ntrA(rpoN)* gene. Detailed mapping studies revealed that both Tn*5*#15 and #136 were located within ORF2 and thus were designated as *ntrA(rpoN)*::Tn*5* insertions (data not shown; figure 1). The ATG at position 916 was designated as the putative NtrA(RpoN) start codon, because of the presence of a Shine Dalgarno consensus sequence (position 901; GGGAGG versus AGGAGG; Ringquist *et al.* 1992), at a proper distance and because of the results from the sequence comparison studies (see fig. 4). ORF2 encodes a NtrA(RpoN) polypeptide of 514 aminoacids (predicted Mwt. 55,863 D) and is separated from ORF1 by an intergenic region of 62 bp. A DNA sequence analysis of this region failed to identify DNA sequence motifs characteristic of promoters or transcriptional terminators (data not shown), suggesting that the *A. caulinodans ntrA(rpoN)* gene described here may be part of an operon and co-transcribed with ORF1.

The deduced polypeptide corresponding to ORF1 (figures 1 and 3) shares significant homology with the polypeptides encoded by the corresponding ORF's upstream of the *ntrA(rpoN)* genes of *R. meliloti* (Ronson *et al.* 1987a), *Thiobacillus ferrooxidans* (Berger *et al.* 1990), *Pseudomonas putida* (Inouye *et al.* 1989), *K. pneumoniae* (Merrick and Gibbins 1985), *Rhizobium* sp. NGR234 (van Slooten and Stanley 1991) *E. coli* (Imaishi, unpublished; GenBank D12938) and *Salmonella typhimurium* (Popham *et al.* 1991). A limited sequence comparison of these ORF's is shown in figure 5. The predicted polypeptide encoded by the truncated ORF3, located immediately downstream of the *A. caulinodans ntrA(rpoN)* gene (figures 1 and 3) was not found to share significant homology with the products of ORF's found downstream of the *ntrA(rpoN)* gene in other bacteria (data not shown; see Discussion).

Figure 3. DNA sequence of the *A. caulinodans ntrA(rpoN)* locus and predicted protein products.

```
   1  CCCGGGATGAACGTCCTGTCCATGTTCGGTAGAAACGCCACACGCGAAACATCCAGTCCC  60
         M  N  V  L  S  M  F  G  R  N  A  T  R  E  T  S  S  P

  61  GCCGCCACCGCCGGCCGCTATGCGGACGAGGGCGACTGGGAGGGCGACGACCACCAGCCC 120
      A  A  T  A  G  R  Y  A  D  E  G  D  W  E  G  D  D  H  Q  P

 121  GCCACTGCGGAGGGCTCGCTGGCCGCCTTCGGCCTCGCCAAGTCCTATGGCGGTCGCAAG 180
      A  T  A  E  G  S  L  A  A  F  G  L  A  K  S  Y  G  G  R  K

 181  GTAGTGCGCGATGTGAGCCTGGACGTGCGCCGCGGCGAGGCCGTCGGCCTCCTTGGCCCG 240
      V  V  R  D  V  S  L  D  V  R  R  G  E  A  V  G  L  L  G  P

 241  AACGGCGCCGGCAAGACCACCGTCTTCTACATGATCACGGGCCTCGTGAAGGCCGATCAG 300
      N  G  A  G  K  T  T  V  F  Y  M  I  T  G  L  V  K  A  D  Q

 301  GGCCGCATCGAGCTCGACGGCCATGACGTGACGCCCATGCCCATGTACCGGCGCGCGCGG 360
      G  R  I  E  L  D  G  H  D  V  T  P  M  P  M  Y  R  R  A  R

 361  CTCGGCATCGGCTATCTGCCGCAGGAAGCCTCGATCTTCCGCGGCTTGTCGGTGGAGGAC 420
      L  G  I  G  Y  L  P  Q  E  A  S  I  F  R  G  L  S  V  E  D

 421  AATATCGGCGCGGTGCTGGAGATCACCGAGCCGAACAGGAAGCGCCGCGCCGAGGAACTC 480
      N  I  G  A  V  L  E  I  T  E  P  N  R  K  R  R  A  E  E  L

 451  GACGCGCTGCTCGAGGAATTCAAGATCACCCACGTGCGCAAGTCGCCCTCCATCGCCCTC 540
      D  A  L  L  E  E  F  K  I  T  H  V  R  K  S  P  S  I  A  L

 541  TCGGGCGGCGAGCGCCGGCGCGTGGAGATCGCGCGCGCGCTGGCGAGCCGCCCGGCCTTC 600
      S  G  G  E  R  R  R  V  E  I  A  R  A  L  A  S  R  P  A  F

 601  ATGCTGCTGGACGAGCCCTTCGCGGGCATCGACCCCATCGCCGTGGGCGACATCCAGGCG 660
      M  L  L  D  E  P  F  A  G  I  D  P  I  A  V  G  D  I  Q  A

 631  CTGGTGCGCCATCTGACCACCCGCGGCATCGGCGTGCTCATCACCGACCATAATGTGCGC 720
      L  V  R  H  L  T  T  R  G  I  G  V  L  I  T  D  H  N  V  R

 721  GAGACGCTGGGCCTGATCGACCGCGCCTACATCATCCATTCGGGCACCGTCCTCATGGAG 780
      E  T  L  G  L  I  D  R  A  Y  I  I  H  S  G  T  V  L  M  E

 781  GGCGACCCGGAGTCCATCGTCGCGAGCCCCGATGTGCGGCGGCTCTATCTCGGCGAAGAG 840
      G  D  P  E  S  I  V  A  S  P  D  V  R  R  L  Y  L  G  E  E

 811  TTCCGGCTGTGAGCACGGCGCTGCGCCTTCCCGCCTCGGCGAGGGCGCGGCGCGCGAGCG 900
      F  R  L  *

 901  GGGAGGGCTGACGTCATGGCGATGAGCCCAAAGATGGAGTTCCGCCAGAGCCAGTCTCTG 960
      ------           M  A  M  S  P  K  M  E  F  R  Q  S  Q  S  L

 961  GTGATGACCCCGCAGCTGATGCAGGCCATCAAGCTGCTGCAGCTCTCCAATCTCGAACTG 1020
      V  M  T  P  Q  L  M  Q  A  I  K  L  L  Q  L  S  N  L  E  L

 991  GTCGCCTATGTGGAGGCCGAGCTCGAACGCAATCCGCTGCTGGAGCGGGCGAGCGAGCCG 1080
      V  A  Y  V  E  A  E  L  E  R  N  P  L  L  E  R  A  S  E  P

1081  GAAAGCCCCGAGCTCGATCCGCCGAACCCGCAGGAAGAGGCACCCACCCCGCCTGACAGT 1140
      E  S  P  E  L  D  P  P  N  P  Q  E  E  A  P  T  P  P  D  S

1141  GGCGCGCCGGTGTCCGGCGACTGGATGGAAAGCGACATGGGCTCGAGCCGCGAGGCCATC 1200
      G  A  P  V  S  G  D  W  M  E  S  D  M  G  S  S  R  E  A  I
```

```
1201  GAGACCCGGCTGGACACCGACCTCGGCAATGTCTTTCCCGATGATGCGCCGGCCGAGCGC  1260
      E  T  R  L  D  T  D  L  G  N  V  F  P  D  D  A  P  A  E  R

1261  ATCGGCGCGGGCAGCGGCAGCGGCTCGTCCATCGAATGGGGCTCGGGCGGCGACCGGGGC  1320
      I  G  A  G  S  G  S  G  S  S  I  E  W  G  S  G  G  D  R  G

1321  GAGGACTACAATCCGGAAGCCTTCCTCGCTGCCGAGACGACGCTGGCCGACCATCTGGAA  1380
      E  D  Y  N  P  E  A  F  L  A  A  E  T  T  L  A  D  H  L  E

1381  GCCCAGCTCTCCGTGGCGGAGCCCGATCCGGCGCGCCGCCTCATCGGCCTCAACCTCATC  1440
      A  Q  L  S  V  A  E  P  D  P  A  R  R  L  I  G  L  N  L  I

1441  GGCCTCATCGACGAGACGGGTTATTTCTCCGGCGACCTCGATGCGGTGGCCGAGCAACTG  1500
      G  L  I  D  E  T  G  Y  F  S  G  D  L  D  A  V  A  E  Q  L

1501  GGCGCCACCCACGATCAGGTGGCCGACGTGCTGCGCGTCATCCAGAGCTTCGAGCCGTCC  1560
      G  A  T  H  D  Q  V  A  D  V  L  R  V  I  Q  S  F  E  P  S

1561  GGCGTCGGCGCACGGTCGCTCAGCGAATGCCTGGCCCTGCAATTGCGCGACAAGGATCGC  1620
      G  V  G  A  R  S  L  S  E  C  L  A  L  Q  L  R  D  K  D  R

1621  TGCGATCCCGCCATGCAGGCGCTGCTCGACAATCTGGAACTCCTCGCCCGCCACGACCGC  1680
      C  D  P  A  M  Q  A  L  L  D  N  L  E  L  L  A  R  H  D  R

1681  AACGCGCTGAAGCGCATCTGCGGGGTGGACGCGGAAGACCTCGCGGACATGATCGGCGAG  1740
      N  A  L  K  R  I  C  G  V  D  A  E  D  L  A  D  M  I  G  E

1741  ATCCGCCGCCTCGATCCGAAGCCCGGCCTCGCCTATGGCGGCGGCGTCGTCCACCCGCTG  1800
      I  R  R  L  D  P  K  P  G  L  A  Y  G  G  G  V  V  H  P  L

1801  GTGCCGGACGTGTTCGTGCGCGAGGGCTCCGACGGCAGCTGGATCGTGGAACTGAATTCC  1860
      V  P  D  V  F  V  R  E  G  S  D  G  S  W  I  V  E  L  N  S

1861  GAGACGCTGCCGCGCGTGCTGGTGAACCAGACCTATCACGCGACGGTGGCCAAGGCGGCG  1920
      E  T  L  P  R  V  L  V  N  Q  T  Y  H  A  T  V  A  K  A  A

1921  CGCTCGGCCGAGGAAAAGACCTTCCTCGCCGACTGCCTCCAGAGCGCCTCCTGGCTTACC  1980
      R  S  A  E  E  K  T  F  L  A  D  C  L  Q  S  A  S  W  L  T

1981  CGCTCGCTCGACCAGCGGGCTCGCACCATCCTCAAGGTGGCGAGCGAGATCGTGCGCCAG  2040
      R  S  L  D  Q  R  A  R  T  I  L  K  V  A  S  E  I  V  R  Q

2041  CAGGACGCCTTCCTCGTGCACGGCGTGCGGCACCTGCGCCCCCTGAACCTGCGCACGGTG  2100
      Q  D  A  F  L  V  H  G  V  R  H  L  R  P  L  N  L  R  T  V

2101  GCGGATGCCATCGGCATGCACGAATCCACCGTCTCGCGGGTGACCTCGAACAAGTACATC  2160
      A  D  A  I  G  M  H  E  S  T  V  S  R  V  T  S  N  K  Y  I

2161  TCCACCCCGCGCGGGGTGCTGGAGATGAAGTTCTTCTTCTCCTCCTCCATCGCTTCCTCG  2220
      S  T  P  R  G  V  L  E  M  K  F  F  F  S  S  S  I  A  S  S

2221  GGTGGTGGCGAGGCCCATGCGGCGGAGGCGGTGCGCCACCGCATCAAGAGCCTCATCGAG  2280
      G  G  G  E  A  H  A  A  E  A  V  R  H  R  I  K  S  L  I  E

2281  GCCGAGAGTGCGGACGACGTGCTGTCCGACGACACGCTGGTGCAGAAGCTGAAGGACGAC  2340
      A  E  S  A  D  D  V  L  S  D  D  T  L  V  Q  K  L  K  D  D

2341  GGCATCGATATCGCCCGCCGAACGGTCGCGAAATATCGCGAGAGCATGAACATCCCGTCC  2400
      G  I  D  I  A  R  R  T  V  A  K  Y  R  E  S  M  N  I  P  S

2401  TCGGTCCAGCGCCGCCGCGAAAAGCAGGCCCTGCGCAGCGACGCCGCCGCCGCCGGCTGA  2460
      S  V  Q  R  R  R  E  K  Q  A  L  R  S  D  A  A  A  A  G  *--

2461  GAGGCGAAGATGACCGATCACGAATGTGAAATCCGGTCCGTGGATGGTAGTTGGGAGGCC  2520
      ----     M  T  D  H  E  C  E  I  R  S  V  D  G  S  W  E  A
```

```
2521  GTCGGGGTCGAGGAGGCCTTGGGCTTACCGTCCAGCCTTCTCAAGCGTTGTCCCGAATGT  2580
      V  G  V  E  E  A  L  G  L  P  S  S  L  L  K  R  C  P  E  C

2581  CACGGACGGGTCCGGGTGCATCGGGCCAGCGTGAACGGTATGCGTGCGCATTTCGAGCAC  2610
      H  G  R  V  R  V  H  R  A  S  V  N  G  M  R  A  H  F  E  H

2641  ATGGAGGCGCACCGCGGGTGCTCACTGTCCCGGG  2674
      M  E  A  H  R  G  C  S  L  S  R
```

The nucleotide sequence of the 2,674 bp *Sma*I fragment of pBS714 and the deduced aminoacid sequence are shown. Stop codons are indicated with asterixes. Putative ribosome binding sites are underlined. The nucleotide sequence reported here has been submitted to GenBank and has the accession number X69959.

Figure 4. Aminoacid sequence alignment of the NtrA(RpoN) proteins.

```
Ac      MAMSPKMEFRQSQSLVMTPQLMQAIKLLQLSNLELVAYVEAELERNPLLE..RASEPESP.ELDPPNP..   65
Bj2     MALTQRLEFRQSQSLVMTPQLMQAIKLLQLSNLDLTTFVEEELERNPLLE..RANDEASG.GEAPAEAGQ   67
Bj1     MALTQRLEFRQSQSLVMSPQLMQAIKLLQLSNLDLMTFVEEELECNPLLE..RASDDAAG.AEAPTEVDQ   67
NGR     MALSASLHLRQSQSLVMTPQLMQSIQLLQMTHLELGQFIAQEVEKNPLLEVPIADEAVSV.RIGPSVMRH   69
Rm      MALSASLHLRQSQSLVMTPQLMQSIQLLQMNHLELSHFIAQEVEKNPLLEVQ.PADEPTI.SDREDAGPH   68
RmA     MALSASLHLRQSQSLVMTPQLMQSIQLLQMNHLELTQFIAQEIEKNPLLEVQSPSDEAST.AERGDSGPQ   69
Cc      LALSHRLELRQGQGLVITPQLQQAIKLLQLSNIELDAFVEAELERNPLLQRDEGDHEPAVEVEAERDASL   70
Ae        MKPSLHVRLSQHLALTPQLQQSIRLLQLSTLELQQEVEQALTENPLLERENDWIESPLRVAADGSVNL   68
Pp        MPKSLVLKMGQQLTMTPQLQQAIRLLQLSTLDLQQEIQEALESNPMLERQEDGEDFDNSDPMADNAEN   68
Av        MPKSLVLKMGQQLTMTPQLQQAIRLLQLSTLDLQQEIQEALDSNPMLERQEDAEDYDSPDMLGEHGDQ   68
Kp        MKQGLQLRLSQQLAMTPQLQQAIRLLQLSTLELQQELQQALDSNPLLEQTDLHDEVETKE.....AED   63
Ec        MKQGLQLRLSQQLAMTPQLQQAIRLLQLSTLELQQELQQALESNPLLEQIDTHEEIDTRE.....TQD   63
St        MKQGLQLRLSQQLAMTPQLQQAIRLLQLSTLELQQELQQALENNPLLEQTDLHEDEIDTQQ.....PQD   63
Tf        MKQGLELKLGQHLAMTPQLQQAIRLLQLSTVDLQQEVQGMLESNPLLD.EETGDEGGGGP.....IPE   62
Bs        MDMKLQQVQVLKPQLTQELRQAITLLGYHSAELAEYIDELSLENPLIERKETDTPPLS.........   58
Rc        MELAQTLSQRQTMQMAGQMLRSLAILGMSSQDLSEHLTEQATSNPFLTYRAP.............   52
Rs        MDMMQFQRQTTQLAMTQRMQESLRILQMSNADLADYLTAQALENPCLEVRVPEGASVA.........   58
----------------------------------------------------------------------------
CON       M   L LR SQ L MTPQL QAI LLQLS L L       LE NPLLE
```

```
Ac      ........QEEAPTPPDSGAPVSGDWMESDMGSSREAIETRLDTDLGNVFPDDAPAERIGAGSGSGSSI.  126
Bj2     FSDSDGGHNDEPGGGPGEAFEPGQEEWMSKDLGTRAEIEQTLDTGLDNVFSEEPAEAAARNAQDAAPTTY  137
Bj1     VSG.....................................................DQLAEAQVRDARDGAMTTY   89
NGR     VTAEAGGAEESVDHG.........DLYDSATTSPGERLRSELDADFANVFQDDTAP.....QR.DAPELL  124
Rm      PAETGGETDEAAGQS.........DLYDSAMSRSGERLSEGLDADFANVFPDDTAP.....QRADAPELL  123
RmA     PEEAGSEIDEGAGEG.........DVYDSATSRSGERLSDGLDSDFANVFPDDTTP.....QRADAPELL  125
Cc      TQVDAVADTTAGRELDTPVDDVSPGERATGEGTDAEHAGGQIDWSRA.....................  117
Ae      QSAPAPAPAEPQGNGEARADGAADDDSYGDSGNGDDYGS..SDWSLDDF...................  115
Pp      KPAAEVQDNSFQESTVS......ADNLEDGEWSERIPNELPVDTAWEDIYQTSA...............  116
Av      S.TLDTTPGSYQEGYESGAASEDGGTLEEGDWHERIPSELPVDTAWEDIYQTSA...............  121
Kp      RESLDT.................VDALEQKE....MPEELPLDASWDEIYTAGT...............   96
Ec      SETLDT.................ADALEQKE....MPEELPLDASWDTIYTAGT...............   96
St      NDPLDT.................ADALEQKE....MPEELPLDASWDEIYTAGT...............   96
Tf      TVELPS.................EERQLDLAAENILPDELPVDSQWDDIFDMGT...............   99
Bs      ........................................................................
Rc      ........................................................................
Rs      ........................................................................
```

```
Ac       .EWGS..GGD...RGEDYNPEAFLAAETTLADHLEAQLSVAEPDPARRLIGLNLIGLIDETGYFSGDLDA  190
Bj2      TEWGG..GAS...GDEDYNLEAFVAAEVTLGDHLAEQLSVAFTAPAQRMIGQYLIDLVDEAGYLPPDLGQ  202
Bj1      TEWGG..GGS...GDEDYNLEAFVASETTLSDHLAEQLSVAFTAPAQRMIGQYLIDLVDEAGYLPPDLGQ  154
NGR      GQWKSMPGAGGSNDGEGYDLDDSSPVGK.LRETLIEQVAFAFAAAADRLIAQHLIDQLDEAGYLHAEIAE  193
Rm       GQWKSMPGAG...DAEGYDLDDFVGGRKTLRETLAEQLPFALSAVSDRLIARYFIDQLDDAGYLHADLAE  190
RmA      GQWKSMPGAS...DGESYDLDDFVASRKTLREALIEQLPFALESGSHRLIAQYFIDQLDDAGYLHADLAE  192
Cc       ........GGGGSFESDEGYERALTDSPTLAAHLRTQLVQAALSPAHNAIAEILIDAVDEGGYLRLDIVE  179
Ae       ....ARRPQG...DEDEKTPMQLREAEPTLREYLMEQLTPLKISARDKGLAIFLIESLDDDGYLSASLEE  178
Pp       ...SSLPSND...DDEWDFTT.RTSAGESLQSHLLWQLNLAPMSDTDRLIAVTLIDSINGQGYLEDTLEE  179
Av       ...SNLPSTD...EDEWDFTT.RTSTGESLQSHLLWQLNLTPMSDTDRLIAVTLIDSINSDGYLEAALEE  184
Kp       ...PSGNGVD...YQDDELPVYQGETTQSLQDYLMWQVELTPFTDTDRAIATSIVDAVDDTGYLTISVED  160
Ec       ...PSAPAVT...YIDDELPVYQGETTQTLQDYLMWQVELTPFSDTDRAIATSIVDAVDETGYLTVPLED  160
St       ...PSGPSGD...YIDDELPVYQGETTQSLQDYLMWQVELTPFSDTDRAIATSIVDAVDDTGYLTVSLDE  160
Tf       ...SGSGNGS...DEDLPDFESRNSRTQSLQDYLRWQADMTHFTADERNMAELIIDAIDERGYLADSLED  163
Bs       ..YHKTNKNR...MMAQEAGLQLSNPQKTLQDALKQQSLDMNLTNTEKKIFNYLIHSLDSNGYLEEDIEE  123
Rc       ...PAFIARG...GEDFDAVGAVAAHKPSLMAHVVDQIEMAFTETPDRLLALRFAEALEPSGWLGQSLDS  116
Rs       PALPSRGIQA...GLDRDAFATVEGQPPSLLAHVEAQIDLAFFDPGDRRTALAFAEALEPSGWLGQFVSE  125
---------------------------------------------------------------------------
CON                                    L   L  Q          DR IA    D  D  GYL   L

Ac       VAEQLG....ATHDQVADVLRVIQSFEPSGVGARSLSECLALQLRDKDRC.....DPAMQALLDNLELLA  251
Bj2      AAERLG....ASQQEVEDVLAVLQKFDPPGVCARNLSECLAIQLRELDRY.....DPAMQALVEHLDLLA  263
Bj1      AAERLG....ATQEDVEHVLAVLQEFDPPGVCARNLRECLAIQLRELDRY.....DPAMQALVEHLDLLA  215
NGR      TAARLA....ASAADVTRVLVVLQQFDPPGVFARTLSECLAIQLRLRNGL.....DPAMEALVANLELLG  254
Rm       TAETLG....AAGEDVARVLHVLQQFDPPGVTARTLGECLAIQLRARNRL.....DPAMEALVANLELLA  251
RmA      TAQRLA....SASEDVTRVLDVLQQFDPPGVTARTLGECLGLQLRARNRL.....DPAMEALVGNLDLLA  253
Cc       LADRLG....CALALVEETLSVLQQFEPVGVFARDVRECLALQLKDLNRY.....DPAMAAMLDHLELLA  240
Ae       ICTELPEELEFEIEEVHAILTLLQSFDPPGVGARNAAECLALQLRRLTHP...QRELALNIVTNHLELLA  245
Pp       ICAGFDPELDIELDEVEAVLHRIQQFEPAGVGARNLGECLLLQLRQLPATT.PWMTEAKRLVTDFIDLLG  248
Av       ILASLDPELGVELDEVEMVLRRIQQFEPAGIAARDLSESLLLQLRQLPPDT.PWLEEAKRLAKDYLDLLG  253
Kp       IVESIG.DDEIGLEEVEAVLKRIQRFDPVGVAAKDLRDCLLVQLSQFAKET.PWIEEARLIISDHLDLLA  228
Ec       ILESIG.DEEIDIDEVEAVLKRIQRFDPVGVAAKDLRDCLLIQLSQFDKTT.PWLDDARLIISDHLDLLA  228
St       IRESMG.DVEVDLDEVEAVLKRIQRFDPVGVAAKDLRDCLLIQLSQFDKST.PWLEEARLIICDHLDLLA  228
Tf       LAATMN....VQEDALLAVLLRVQDFDPPGVGARNLSECLLLQLKQMVEKDDAHVLLAQRIVKDHLQALG  229
Bs       AARRLS....VSAKEAEAVLAKLQSLEPAGIGARSLQECILLQLQRLPNRN....EQAEMLVSAHFDAFA  185
Rc       IALAAG....VSLSRAESMLAVLQGFEPTGLFARDLSDCLILQAREADILT.....WEVETLIRNIRLIA  177
Rs       VAAAAE....VEEEEALVILERLQALEPAGLFARSLAECLALQLEDLGLLT.....WELRTMLDHLPLLA  186
---------------------------------------------------------------------------
CON       A              V  VL   Q FDP GV AR L ECL  QLR             A        L LLA

Ac       RHDRNALKRICGVDAEDLADMIGEIRRLDPKPGLAYGGGVVHPLV.PDVFVREGSDGSWIVELNSETLPR  320
Bj2      KRDIAGLRKVCGVDDEDIADMIGEIRRLNPKPGMKFGAARLQTMV.PDVYVRPGPDGGWHVELNSDTLPR  332
Bj1      KRDIASLRKLCGVDDEDIADMIDELRRLSPKPGMKFGSARLQTMV.PDVYVRPAPDGGWHVELNSDTLPR  284
NGR      RRDFASLKKICGVDEEDLIEMLAEIRKLDPKPGTSFETSVTEAII.PDVVVRSRPDGGWLVELNPDALPR  323
Rm       RRDFASLKKICGVDEEDLIDMLAEIRKLDPKPGTSFETGVFEAII.PDVVVRAAPDGGWLVELNPDALPR  320
RmA      RRDFASLKKICGVDEEDLIDMFAEIRKLDPKPGTSFETGSFETII.PDAVVRTAPDGGWLVELNPDALPR  322
Cc       KRDMAGLRRICGVDDEDLREMIGEIRSLNPRPGAAYHSEPAETLV.PDVMVREGLGGMWHVELNTDTLPR  309
Ae       VRDYTRLKKALQVDEAALKSAHELIRSLAPYPGHAYSRPEADFVV.PDVFVRKGGGG.WIAQLNPDVMPR  313
Pp       SRDYSQLMRRMKIKEDELRQVIELVQSLNPRPGSQIESSEPEYVV.PDVIVRKDSDR.WLVELNQEAIPR  316
Av       NRDFTQLMRRMKLKEEELRPVIELIQSLNPRPGAQIESSEPEYVV.PDVIVRKHNDR.WLVELNQEAVPR  321
Kp       NHDFRSLMRVTRLKEEVLKEAVNLIQSLDPRPGQSIQTGEPEYVI.PDVLVRKVNDR.WVVELNSDSLPR  296
Ec       NHDFRTLMRVTRLKEDVLKEAVNLIQSLDPRPGQSIQTGEPEYVI.PDVLVRKHNGH.WTVELNSDSIPR  296
St       NHDFRTLMRVTRLKEEVLKEAVNLIQSLDPRPGQSIQTGEPEYVI.PDVLVRKHNGR.WTVELNSDSIPR  296
Tf       RHDYPRLCTVLGVDEAALRAAMALISALNPKPGEDVGTESTEYVI.PDVIVRWAGSR.LRTDLNPEAMPK  297
Bs       QKKWKTLSVETGIPLHTIQDISDDIAALHPRPGLLFARPEQDVYIEPDIFITVKNGH.IAAELNTRSFPE  254
Rc       ENRLSDLADLCDCDIGDIPEIIKQIRHLNPKPGLAFDHQPTPVFP.PDLIAVRGAEG.WTVELNRATSPT  245
Rs       ERGIADLARRCDCEPEHIRENLALIRSLSPKPGEAFAADRTPIQP.PDVRVLRGPEG.WEVELTRAQLPR  254
---------------------------------------------------------------------------
CON        D   L        E  L       I  L PKPG              PDV VR       W VELN    PR
```

```
Ac      VLVNQTYHATVAKAARSA.EEKTFLADCLQSASWLTRSLDQRARTILKVASEIVRQQDAFLVHGVRHLRP  389
Bj2     VLVNQTYYSELSKKIGKD.GDKSYFTDALQNATWLVRALDQRARTILKVATEIVRQQDGFFTHGVAHLRP  401
Bj1     VLVNQTYYSKLSKKIGKD.VDKSYFNDALQNATWLVRALDQRARTILKVATEIVRQQDGFFTLGVAHLRP  353
NGR     VLVNHDYFAEISRHSQKNSAEQAFLSECMQNANWLTRSLDQRARTILKVASEIVRQQDAFLVHGVDHLRP  393
Rm      VLVNHDYFTEISRSSRKNSGEQAFLNECLQNANWLTRSLDQRARTIMKVASEIVRQQDAFLIHGVGHLRP  390
RmA     VLVNHEYFAEISRSCRKSSGEQIFLNECLQNANWLTRSLDQRARTIMKVASEIVRQQDAFLMHGVGHLRP  392
Cc      VLVDQRYHARVSKG.ARSDQEKTFVADCMASANWLVKSLDQRARTILKVASEIVRQQDGFLAFGVEHLRP  378
Ae      LRINDMYAQILRG..AKGESGTAGLQQKLQEARWLIKNIQQRFDKILVRSQAIVERQKNFFSHGEIAMRP  381
Pp      LRVNPQYAGFVRR..ADTSADNTFMRNQLQEARWFIKSLQSRNETLMKVATQIVEHQRGFLDHGDEAMKP  384
Av      LRINPHYAGFIRR..ADASADNTFMRNQLQEARWFIKSLQSRNETLMKVSTQIVEHQRGFLDYGEEAMKP  389
Kp      LKINQQYAAMGNS..TRNDADGQFIRSNLQEARWLIKSLESRNDTLLRVSRCIVEQQQAFFEQGEEFMKP  364
Ec      LQINQHYASMCNN..ARNDGDSQFIRSNLQDAKWLIKSLESRNDTLLRVSRCIVEQQQAFFEQGEEYMKP  364
St      LQINQHYAAMCNS..ARNDADSQFIRSNLQDAKWLIKSLESRNDTLLRVSRCIVEQQQAFFEQGEEYMKP  364
Tf      LRINRHYADMAGGKDAAHK....YIQDQLNEARWFIKSLQSRQDTILKVARAIVERQKDFFANGPESMRP  363
Bs      IDLHPQYRTLLSSGSCQDTV..SYLSAKYQEWRWLSRALRQRKQTITRIINELITRQKDFFLKGRSAMKP  322
Rc      ITVREDRFADGTADAKAR.....AERRRRGRGPGAGEALERRRDTLLRTAAVLVARQSAFLDKGPAHLVP  310
Rs      IRVSEA....GDTGDRQADA...WLARARSQARWLERAVERRQATLLRTAVCLVRHQADFLDQGPRALRP  317
        ----------------------------------------------------------------------
CON          N  Y                          LQ A WL  SL  R  T L VA  IV  Q  F   G     P
```

```
Ac      LNLRTVADAIGMHESTVSRVTSNKYISTPRGVLEMKFFFSSSIASSGGGEAHAAEAVRHRIKSLIEAESA  459
Bj2     LNLKAVADAIQMHESTVSRVTANKYMATNRGTFELKYFFTASIASADGGEAHSAEAVRHHIKQLIDSEAP  471
Bj1     LNLKAVAEAIQMHESTVSRVTANKYMATNRGTFELKYFFTASIPSADGGEAHSAEAVRHRIKQLIESEEP  423
NGR     LNLRIVADAIKMHESTVSRVTSNKYRLTPRGLFELKYFFTVSIGSAENGDAHSAESVRHRIRTMISQESA  463
Rm      LNLRIVADAIKMHESTVSRVTSNKYMLTPRGLFELKYFFTVSIGSAENGDAHSAESVRHRIRTMINQESA  460
RmA     LNLRTVADAIKMHESTVSRVTSNKYMLTPRGLYELKYFFTASIGSAENGDAHSAESVRHRIRTMVNQESA  462
Cc      LNLKTVADAIGMHESTVSRVTSNKYIATPRGVFELKFFFTSAIQSSEGGEAHSAASVRHKIKGLVDAEKC  448
Ae      LVLRETADTLGLHESTISRVTTNKYMATPMGTFELKYFFGSHVSTETGGAASST.AIRALIKQLIGAEDP  450
Pp      LVLHDIAEAVGMHESTISRVTTQKYMHTPRGIYELKYFFSSHVSTSEGGECSST.AIRAIIKKLVAAENQ  453
Av      LVLHDIAEAVGMHESTISRVTTQKYMHTPRGIYELKYFFSSHVSTAEGGECSST.AIRAIIKKLIAAENP  458
Kp      MVLADIAQAVEMHESTISRVTTQKYLHSPRGIFELKYFFSSHVNTEGGGEASST.AIRALVKKLIAAENP  433
Ec      MVLADIAQAVEMHESTMSRVTTQKYLHSPRGIFELKYFFSSHVNTEGGGEASST.AIRALVKKLIAAENP  433
St      MVLADIAQAVEMHESTISRVTTQKYLHSPRGIFELKYFFSSHVNTEGGGEASST.AIRALVKKLIAAENP  433
Tf      MVLRHIADAVEMHESTVSRVTNQKYMITPRGLYEFKYFFSSHVGTDSGGSASAT.AIRALLIKMTQAEDA  432
Bs      LTLREVADCLSLHESTVSRAIKGKTIQTPYGLFEMKLFFSAKAEASGDGDASNY.AVKTHLENLINQEDK  391
Rc      LTLEDVASELGLHASTISRAVSGRMIQTQTRALPLRAFFSRAVSTQGGGEAVSRDSLDF.VQRTWAAKIR  379
Rs      LSMEEVALELDLHPSTISRATATRLIETPRGLIPLRAFFSRSVSSDGPEAPQSQDALMA.LVEIIARERDR  386
        ----------------------------------------------------------------------
CON     L L   A A  MHEST SRVT  KY  TPRG  ELKYFFS        GG A S  A R     LI  E
```

```
Ac      DDVLSDDTLVQKLKDD.GIDIARRTVAKYRESMNIPSSVQRRREKQALRSDAAAAG*               514
Bj2     AAILSDDTIVERLRAS.GIDIARRTVAKYREAMRIPSSVQRRRDKQSALGNVLSTAMSDRSRNPEPA*   537
Bj1     SAVLSDDAIVERLRVS.GIDIARRTVAKYREAMRIRSSVQRRRDMMW......STMNSRASGGTGLDK*  484
NGR     DAVLSDDDIVDILKRA.GVDIARRTVAKYREAMNIPSSVQRRREKRALPRQPASDCGFFAAAN*       525
Rm      DAVLSDDDIVDVLKQA.GVDIAARTVAKYREAMSIPSSVQRRREKRALPRPRDSERCRQAASA*       523
RmA     DAVLSDDDIVDILKKA.GVDIAARTVAKYREAMHIPSCVQRRREKRALARVG*                  513
Cc     EADVHSDDRIVEILKAA.GVDIARRTVAKYREAMRIPSSVERRRLMKEAV*                     497
Ae      RNPLSDSRIAELLGEQ.GFVVARRTVAKYREALKIPAVNLRKSL*                          493
Pp      KKPLSDSKIAGLLEAQ.GIQVARRTVAKYRESLGIAPSSERKRLM*                         497
Av      KKPLSDSKIAGLLEEQ.GIQVARRTVAKYRESLSIAPSSERKRLM*                         502
Kp      AKPLSDSKLTTMLSDQ.GIMVARRTVAKYRESLSIPPSNQRKQLV*                         477
Ec      AKPLSDRKLTSLLSEQ.GIMVARRTVAKYRESLSIPPSNQRKQLV*                         477
St      AKPLSDSKLTSLLSEQ.GIMVARRTVAKYRESLSIPPSNQRKQLV*                         477
Tf      QHPLSDAEIARVLADQ.GIQIARRTVAKYREAANVPPASQRRRL*                          475
Bs      TKPLSDQKLVDLLYEQHGIQISRRTVAKYRDQMNIPSSAARKRYK*                         436
Rc      QNPLSDDAIVTLAERA.GLRIARRTVAKYRSTLGLASSYERRRAAAAR*                      426
Rs      TKPFSDDAIVKQAKLA.GAVLARRTVTKYRETLGIPSSYDRKRAAAAA*                      434
        ----------------------------------------------------------------------
CON       PLSD  I   L    GI  ARRTVAKYRE   IP S QR R
```

The abbreviations for the species and references for the sequences presented are: Ac, *Azorhizobium caulinodans* ORS571 (this work), Bj1/2, *Bradyrhizobium japonicum* (Kullik *et al.* 1991), NGR, *Rhizobium* sp. NGR234 (van Slooten *et al.* 1990), Rm, *Rhizobium meliloti* 1021 (Ronson *et al.* 1987a), RmA, *Rhizobium meliloti* 104A14 (Shatters *et al.* 1989), Cc, *Caulobacter crescentus* (Brun, unpublished; GenBank: X68549; S50955), Ae, *Alcaligenes eutrophus* (Warrelmann *et al.* 1992), Pp, *Pseudomonas putida* (Inouye *et al.* 1989: Kohler *et al.* 1989a), Av, *Azotobacter vinelandii* (Merrick *et al.* 1987) Kp, *Klebsiella pneumoniae* (Merrick and Gibbins 1985), Ec, *Escherichia coli* (Sasse-Dwight and Gralla 1990), Sm, *Salmonella typhimurium* (Popham *et al.* 1991), Tf, *Thiobacillus ferrooxidans* (Berger *et al.* 1990), Bs, *Bacillus subtilis* (Debarbouille *et al.* 1991), Rc, *Rhodobacter capsulatus* (Alias *et al.* 1989; Jones and Haselkorn 1989), Rs, *Rhodobacter sphaeroides* (Meijer and Tabita 1992). The consensus line (CON) shows the position where at least ten NtrA sequences contained identical amino acid residues.

Figure 5. Aminoacid sequence comparison of the deduced protein products of ORF1.

The *A. eutrophus* (Ae) sequence was derived from Warrelmann *et al.* (1992), *Rhizobium* sp. NGR234 (NGR) from van Slooten and Stanley (1991), *R. meliloti* (Rm) from Albright *et al.* (1989b), *E. coli* (Ec) from Imaishi (unpublished; GenBank D12938), and the *T. ferroxidans* (Tf) sequence from Berger *et al.* (1990).

```
Ac      MNVLSMFGRNATRETSSPAATAGRYADEGDWEGDDHQPATAEGSLAAFGLAKSYGGRK   58
Ae       MRPAPETPERQTVPMTDTATIADKPSVEASTVLAG.....GSTLVVRHLKKRYGSRT   52
NGR                MQIPFLPKRKRVKKPSAAAAAAARAVDKARYDGTLIARGLTKSYRSRR   47
Rm                 MQIPFLHKRKRGKKPSAAAAAAARAVDKARYDGTLIARGLTKSYRSRR   47
Ec                                                MATLTAKNLAKAYKGRR   17
Tf                                               MSELLQAQSLFKSYRRRV   18
---------------------------------------------------------------------
CON                                                 L A  L KSY  R

                          **********
Ac      VVRDVSLDVRRGEAVGLLGPNGAGKTTVFYMITGLVKADQGRIELDGHDVTPMPMYRR  116
Ae      VVKDVSLDVKSGEVVGLLGPNGAGKTTSFYMIVGLVALDEGDIVLDGDHISGLPIHER  110
NGR     VVNGVSLVVRRGEAVGLLGPNGAGKTTCFYMITGLVPVDEGSIEINGNDVTTMPMYRR  105
Rm      VVNGVSLVVRRGEAVGLLGPNGAGKTTCFYMITGLVPVDEGSIEINGNDVTTMPMYRR  105
Ec      VVEDVSLTVNSGEIVGLLGPNGAGKTTTFYMVVGIVPARCGNIIIDDDDISLLPLHAR   75
Tf      VVRDVSVQVATGEVVGLLGPNGAGKTTTFYMMVGLVRPDRGHIFLQQRDITALPMHER   76
---------------------------------------------------------------------
CON     VV DVSL V  GE VGLLGPNGAGKTT FYMI GLV  D G I L G D T  PM  R


Ac      ARLGIGYLPQEASIFRGLSVEDNIGAVLEITEPN.....RKRRAEELDALLEEFKITH  174
Ae      APMGLSYLPQEASVFRKLNVEENIRAVLELQVSNGKPLPKAEIERRLDSLLDDLQIAH  168
NGR     ARLGVGYLPQEASILRGLTVEENIRAVLEIHDKN.....VDRRESKLNDLLGEFSITH  163
Rm      ARLGVGYLPQEASIFRGLTVEDNIRAVLEVHDEN.....VDRRESKLNDLLGEFSITH  163
Ec      ARRGIGYLPQEASIFRRLSVYDNLMAVLQIRDDL....SAEQREDRANELMEEFHIEH  133
Tf      ARMGLGYLPQEPSVFRQMSAADNVLAVLETLPLS....PVERQERQ.EQLLSELHLHA  134
---------------------------------------------------------------------
CON     AR G GYLPQEASIFR L VEDNI AVLE    N       RRE  L  LL E  I H
```

```
                                  **************
Ac      VRKSPSIALSGGERRRVEIARALASRPAFMLLDEPFAGIDPIAVGDIQALVRHLTTRG  227
Ae      LRNNPALSLSGGERRRVEIARALASSPRFILLDEPFAGVDPIAVGEIQRIVSFLKARN  226
NGR     LRKSPAIALSGGERRRLEIARALATDPTFMLLDEPFAGVDPISVADIQALVRHLTSRG  216
Rm      LRKSPAIALSGGERRRLEIARALATDPTFMLLDEPFAGVDPISVADIQALVRHLTSRG  216
Ec      LRDSMGQSLSGGERRRVEIARALAANPKFILLDEPFAGVDPISVIDIKRIIEHLRDSG  187
Tf      LRDTKGHSLSGGERRRVEIARALAMSPRFILLDEPFAGIDPISVLEIQRLIRDLRARG  187
--------------------------------------------------------------------
CON     LR SP   LSGGERRRVEIARALA  P F LLDEPFAGVDPISV DIQ LVRHL  RG

Ac      IGVLITDHNVRETLGLIDRAYIIHSGTVLMEGDPESIVASPDVRRLYLGEEFRL*     281
Ae      IGVLITDHNVRETLGICDHAYIISEGTVLAAGQPEEIIANDAVRRVYLGENFRM*     280
NGR     IGVLITDHNVRETLGLIDRAYIIHAGEVLTHGRANDIVTNPDVRRLYLGDNFSL*     270
Rm      IGVLITDHNVRETLGLIDRAYIIHAGEVLTHGRANDIVTNPDVRRLYLGDNFSL*     270
Ec      LGVLITDHNVRETLAVCERAYIVSQGHLIAHGTPTEILQDEHVKRVYLGEDFRL*     241
Tf      IGVLITDHNVRETLGICERAYILHDGKVLTAGSPQEIVDDPMVRQVYLGDQFQI*     241
--------------------------------------------------------------------
CON     IGVLITDHNVRETLG  DRAYIIH G VL  G P  IV  P VRR YLG  F L
```

Aminoacid residues conserved in at least 4 out of 6 proteins are shown in the line labelled "CON". The regions of homology with nucleotide-binding pockets of ATP-requiring enzymes (Walker *et al.* 1982) are indicated by asterixes.

4.4.4 Effect of *A. caulinodans ntrA(rpoN)*::Tn*5* mutants on *nifHD-lacZ* and *nifA-lacZ* gene expression

In order to examine the effect of the *ntrA(rpoN)*::Tn*5* mutations #15 and #136 on *A. caulinodans nif* gene expression, plasmid pPR3408 (Pawlowski *et al.* 1987), carrying an *A. caulinodans nifHD-lacZ* gene fusion, was introduced into strains ORS571N15-2c and ORS571N136-1c, as well as into the wild-type *A. caulinodans.* The transconjugants were cultured under different physiological conditions (aerobic versus microaerobic; in the presence and absence of ammonium) and the reporter gene activity (β-gal) was measured. No β-gal activity was detected in the *ntrA(rpoN)*::Tn*5* mutant strains under any physiological conditions examined, while a high level of β-gal activity was found in the wild-type strain harboring the *nifHD-lacZ* fusion under nitrogen-fixing conditions (data not shown).

Plasmid pJSwp, carrying a fusion of the wild-type *A. caulinodans nifA* promoter to *lacZ* (Table 1), was introduced into the same *ntrA(rpoN)*::Tn*5* strains and integrated into the chromosome (see Materials and Methods). *NifA-lacZ* expression was examined in these strains by measuring β-gal activity. In addition, the growth curves of the *nifA-lacZ* containing strains were determined under different physiological conditions. The results are shown in figure 6.

The wild-type strain harboring the *nifA-lacZ* fusion grew well in the presence of ammonium under aerobic conditions, but could not grow aerobically in the absence of combined nitrogen (ammonium); growth under microaerobic conditions in the presence of ammonium was proficient (but less good than aerobically in the presence of ammonium), while growth under nitrogen-fixing conditions (microaerobically in the absence of ammonium) was intermediate (fig. 6A).

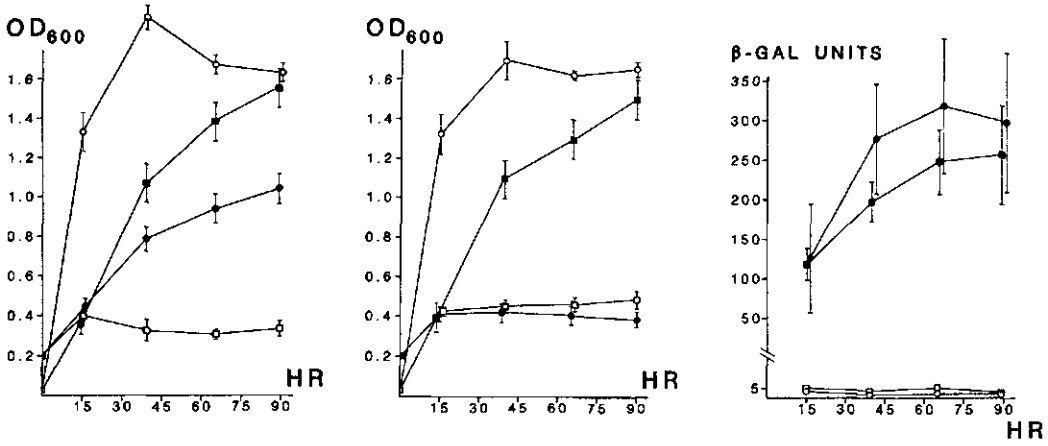

Figure 6. Growth and β-gal activity of wild-type and *ntrA(rpoN)*::Tn*5* strains harboring a wild-type *nifA-lacZ* gene fusion under different physiological conditions.

Panel A. Growth of strain *A. caulinodans* harboring the *nifA-lacZ* gene fusion under nitrogen fixing (derepressing) conditions (LSO medium lacking a N-source; 3% O_2; solid circles), ammonium repressing conditions (LSO medium supplemented with 0.1% ammonium sulfate; 3% O_2; solid boxes), O_2 repressing conditions (LSO medium lacking a N-source; air; open boxes) and aerobically in the presence of 0.1% ammonium sulfate (open circles). The OD_{600} is shown on the Y-axis and the hours after inoculation on the X-axis. The results shown represent an average of more than ten independent experiments.

Panel B. Growth of strain ORS571N15-2c *(ntrA(rpoN)*::Tn*5)* harboring the wild-type *nifA-lacZ* gene fusion under the same physiological conditions as described for panel 6A. The results represent the average of at least five independent experiments.

Panel C. β-galactosidase activity of *A. caulinodans* wild-type (squares) and *ntrA(rpoN)*::Tn*5* mutant strains (circles) harboring the wild-type *nifA-lacZ* gene fusion under nitrogen-fixing (derepressing) conditions (closed circles/squares) or under ammonium plus oxygen repressing conditions (open circles/squares). The results represent an average of at least three independent experiments.

The *ntrA(rpoN)*::Tn*5* mutant strain ORS571N15-2c, harboring the *nifA-lacZ* fusion, grew equally well as the wild-type strain aerobically or microaerobically in the presence of ammonium and did not grow aerobically in the absence of combined nitrogen (ammonium); in contrast to the wild-type, it could not grow under nitrogen-fixing conditions (microaerobically in the absence of combined nitrogen/ammonium; fig. 6B).

The β-gal enzyme assays on the same cultures under nitrogen-fixing conditions (microaerobically in the absence of ammonium) revealed that, in spite of the fact that *ntrA(rpoN)*::Tn*5* strains could not grow or fix nitrogen, the *nifA* gene was expressed at a wild-type level and fully repressed by ammonium. In fact, *nifA-lacZ* expression appeared to be elevated in the *ntrA(rpoN)*::Tn*5* versus wild-type strains (fig. 6C).

4.4.5 Search for additional *ntrA(rpoN)*-like loci in *A. caulinodans*

The results shown above and previous observations (Pawlowski *et al.* 1987; 1991; Ratet *et al.* 1989) revealed that the *A. caulinodans nifA* promoter is *ntr*-controlled, in response to the N-status of the cells, and contains a functional -24/-12 element, normally responsible for interaction with NtrA(RpoN). However, the *A. caulinodans ntrA(rpoN)* gene described here, while controlling nitrogen fixation and nitrate assimilation, does not appear to be involved in *nifA* promoter activity (fig. 6C), suggesting the presence of an additional *ntrA(rpoN)*-equivalent gene in this organism, as has been observed in *Bradyrhizobium japonicum* (Kullik *et al.* 1991) and *Rhodobacter* (Meijer and Tabita 1992).

In order to examine this hypothesis, a synthetic oligonucleotide (3'C$^{G}/_{T}$GCGTGCCAGCGCTT$^{C}/_{T}$AT$^{G}/_{A}$GCGCT5'), corresponding to a highly conserved domain of NtrA(RpoN) proteins (RRTVAKYRE; fig. 4) was prepared and used as a DNA hybridization probe for a Southern blot carrying chromosomal DNA of *A. caulinodans, B. japonicum, R. meliloti* and *E. coli.* The hybridization results confirmed the presence of two *ntrA(rpoN)* copies in *B. japonicum* (Kullik *et al.* 1991), and suggested the presence of at least 2 additional *ntrA(rpoN)*-homologous regions in *A. caulinodans* (see fig. 2).

4.5 DISCUSSION

4.5.1 Phenotype of the *A. caulinodans ntrA(rpoN)* mutant

The inability of the *A. caulinodans ntrA(rpoN)*::Tn*5* mutant to assimilate nitrate or utilize dicarboxylic acids constitute typical phenotypes of *ntrA(rpoN)* mutants of other (nitrogen fixing) bacteria (see Kustu *et al.* 1989; de Bruijn *et al.* 1990; Merrick 1992 and references cited therein; Kullik *et al.* 1991) and supports the designation of this locus as *ntrA(rpoN)*.

The Nif^-/Fix^- phenotype of the *A. caulinodans ntrA(rpoN)*::Tn*5* mutant described here resembles that observed with corresponding single *ntrA(rpoN)* mutants of other diazotrophs such as *K. pneumoniae* (de Bruijn and Ausubel 1983; Merrick and Gibbins 1985), *Rhodobacter capsulatus* (Jones and Haselkorn 1989), and *Azotobacter vinelandii* (Toukdarian and Kennedy 1986; Merrick *et al.* 1987), as well as single *ntrA(rpoN)* mutants of symbiotic nitrogen fixing organisms, such as *R. meliloti* (Ronson *et al.* 1987a) and *Rhizobium* sp. NGR234 (Stanley *et al.* 1989), and a double *ntrA(rpoN)* mutant of *B. japonicum* (Kullik *et al.* 1991). In the case of *K. pneumoniae*, both the *nifA* promoter and the promoters of the other *nif* genes require *ntrA(rpoN)* for their expression (see chapter 2), while in the other cases cited above *nifA* expression appears to be independent of *ntrA(rpoN*; see chapter 2).

The expression of the *A. caulinodans nifA* promoter is also independent of the *ntrA(rpoN)* gene described here, in spite of the fact that the *nifA* 5' upstream region contains a functional -24/-12 promoter element and is controlled by the N-status of the cell, *ntrBC* and *ntrYX* (see chapter 2 and 3). In this respect, the regulatory circuit controlling *A. caulinodans nifA* expression resembles that proposed for *R. capsulatus*, where the presence of another (NtrA(RpoN)-like; NtrA*; see chapter 8) alternative σ factor (specifically) involved in *nifA* promoter activity has also been proposed (Hübner *et al.* 1991; Preker *et al.* 1992; Foster-Hartnett and Kranz 1992). Whether one of the additional fragments showing homology with the oligonucleotide corresponding to the conserved NtrA(RpoN) domain encodes the postulated *ntrA**(σ54*) locus is presently under investigation.

4.5.2 Conservation of the *ntrA(rpoN)* locus

Analysis of the aminoacid sequence deduced from ORF2 (figures 3 and 4) further confirms that the gene inactivated by the Tn*5* insertion #136 encodes an alternative σ factor of the σ^{54} family, as defined by Hirschmann *et al.* (1985), Merrick *et al.* (1987) and Kustu *et al.* (1989), and therefore can be designated as *ntrA(rpoN)*. The alignment of the *A. caulinodans* NtrA(RpoN) aminoacid sequence with the sequences of NtrA(RpoN) proteins from 14 different bacterial species (fig. 4) reveals that the *A. caulinodans* NtrA(RpoN) protein is most closely related to the NtrA(RpoN) proteins from *B. japonicum* (Kullik *et al.* 1991). *A. caulinodans* NtrA(RpoN) contains the three major regions defined by Merrick *et al.* (1987).

Region I (Met-1 to Gln-50) represents a strongly conserved domain, rich in leucine and glutamine residues. This domain has been postulated to be involved in contacting the -12 region of the promoter, to facilitate the interaction between σ^{54} and activator proteins, and to play a role in positioning the σ^{54}-RNA polymerase complex near the DNA region to be melted upon activation, since deletions in this domain are unable to progress from a closed complex to a transcriptionally active open complex (Sasse-Dwight and Gralla 1990; Merrick 1992). However, recently Merrick and Chambers (1992) have proposed that the helix-turn-helix motif in domain III (see below) plays a role in interaction with the -12 region and have suggested an alternative explanation for the results by Sasse-Dwight and Gralla (1990).

Region II (Arg-51 to Thr-148) of the *A. caulinodans* NtrA(RpoN) protein shows no significant homology to other NtrA(RpoN) proteins, which is consistent with the fact that this is the least conserved region in these proteins. This region contains multiple negatively charged residues, which have been postulated to play a role in melting DNA in the promoter, since mutating these residues in the *E. coli* NtrA(RpoN) protein (residues 51-77) results in an inability to form open complexes, while leaving the DNA binding activity of the σ^{54}-RNA polymerase complex relatively unaffected (Sasse-Dwight and Gralla, 1990). However, the absence of this region in the NtrA(RpoN) proteins of *Rhodobacter spaeroides*, *R. capsulatus* and *Bacillus subtilis* (Meijer and Tabita 1992; Alias *et al.* 1989; Jones and Haselkorn 1989; Debarbouille *et al.* 1991) suggests that this region is not essential in all bacteria.

Within region III, four conserved subdomains have been identified (Merrick *et al.* 1987; van Slooten *et al.* 1990). Subdomain IIIA (Leu-188 to Pro-214) shares homology with other σ factors, such as RpoD and SpoIIAC (Gribskov and Burgess 1986; Helmann and Chamberlin 1988) and has been proposed to be involved in interacting with the core RNA polymerase (see Merrick *et al.* 1987). It may also play a role in protein-DNA interactions, since a deletion of the analogous *E. coli* NtrA(RpoN) protein reduces contact formation at the -12 region of the promoter (Sasse-Dwight and Gralla 1990). Subdomain IIIB (Trp-355 to His-382) shares homology with an aminoacid sequence near the N-terminus of the β'-subunit of *E. coli* RNA polymerase (RpoC) and may play a role in protein-protein interactions (Merrick *et al.* 1987). Subdomain IIIC (Asn-391 to Ser-411) contains the α-helix (residues 391-399)--β-turn (400 and 401)--α-helix (402-411) motif, characteristic of DNA-binding proteins (Dodd and Egan 1990; Coppard and Merrick 1991), which has been shown to be involved in recognition of the -13/-12 residues (Merrick and Chambers 1992). The adjacent aminoacid sequences (residues 412-429) are also completely conserved in NtrA(RpoN) proteins. Subdomain IIID (Ala-480 to Arg-488) is also conserved in all NtrA(RpoN) proteins and has been designated as the RpoN-box (Van Slooten *et al.* 1990). Its function is unknown, but the oligonucleotide synthesized to screen for the presence of additional *ntrA(rpoN)*-like sequences in rhizobial genomes was derived from this region. It is also interesting to note that rhizobial NtrA(RpoN) proteins carry a 14-25 aminoacid 'tail' at the C-terminal end of subdomain IIID, the function of which remains to be determined.

In a variety of bacterial species, including *A. caulinodans* (fig. 1), an ORF (designated ORF1) has been found immediately upstream of the *ntrA(rpoN)* gene. The *A. caulinodans* ORF1 shares significant homology (fig. 5) with the corresponding ORF's of *R. meliloti* (Ronson *et al.* 1987a; Albright *et al.* 1989b), *Rhizobium* sp. NGR234 (van Slooten *et al.* 1990; van Slooten and Stanley 1991), *S. typhimurium* (Popham *et al.* 1991), *E. coli* (Imaishi, unpublished; GenBank D12938), *T. ferrooxidans* (Berger *et al.* 1990), *P. putida* (Inouye *et al.* 1989), *K. pneumoniae* (Merrick and Gibbins 1985) and *Alcaligenes eutrophus* (Warrelmann *et al.* 1992). The biological function of the protein encoded by ORF1 and the reason for the direct linkage of ORF1 to the *ntrA(rpoN)* gene in these diverse bacteria are unknown. Albright *et al.* (1989b) have shown that the predicted ORF1 product shares homology with a superfamily of ATP-binding proteins involved in transport, cell division, nodulation and DNA repair (Higgins *et al.* 1986) and have attempted, unsuccessfully, to introduce an insertion mutation in ORF1. These authors

concluded that ORF1 may encode an essential housekeeping function. However, *in vitro* transcription studies using σ^{54}-dependent promoters have shown that purified activator protein and σ^{54}-RNA polymerase are sufficient for promoter activity (Hirschman *et al.* 1985; Hunt and Magasanik 1985; Wong *et al.* 1987), suggesting that the protein encoded by ORF1 may not be important for NtrA(RpoN) action.

The partial ORF downstream of the *A. caulinodans ntrA(rpoN)* locus (ORF3; fig. 1) does not share any significant homology with corresponding ORF's downstream of the *ntrA(rpoN)* genes of other bacteria (data not shown). This is interesting in view of the fact that the corresponding ORF's in *R. meliloti* (Ronson *et al.* 1987a; Albright *et al* 1989b), *T. ferrooxidans* (Berger *et al.* 1990), *P. putida* (Inouye *et al.* 1989), *K. pneumoniae* (Merrick and Coppard 1989), *A. vinelandii* (Merrick *et al.* 1987), *Rhizobium* sp. NGR234 (van Slooten *et al.* 1990), *B. japonicum* (Kullik *et al.* 1991), *A. eutrophus* (Warrelmann *et al.* 1992) and *S. typhimurium* (Popham *et al.* 1991) do share significant homologies. No mutations in the ORF3's of these organisms have been reported, and therefore no biological role for the ORF3 product has been established. A homology search of the NBRF/PIR and SWISS.PROT protein banks has revealed interesting similarity matches for a stretch of residues (LKR**C**PE**C**H**G**RVR; position 31 to 42) of the deduced ORF3 product, to a region (VRR**C**PQ**C**H**G**DML; 83% similarity) of the *hypA* gene product of *E. coli*, postulated to be a transcriptional activator of hydrogenase genes (Lutz *et al.* 1991), and the (covalent) heme-binding domain (IMK**C**SQ**C**H**T**VEK; 75% similarity) of human heart cytochrome-c (Schroeder 1968).

In conclusion the *ntrA*(*rpoN*) gene identified here controls the expression of some of the *A. caulinodans nif* genes, but not the central *nif* regulatory gene *nifA*, suggesting that another *ntrA(rpoN)* copy may exist in *A. caulinodans* and stimulated the interest in searching for this gene or its gene product. One approach was to search for NtrA(RpoN)-like factors (and other *trans*-acting factors) interacting with the *nifA* 5' upstream region by *in vitro* protein DNA interactions (see chapter 5).

CHAPTER 5

SPECIFIC BINDING OF *A. CAULINODANS* PROTEIN(S) TO THE *NIFA* REGULATORY REGION

5.1 ABSTRACT

Expression of the *A. caulinodans nifA* gene is regulated by the cellular oxygen and nitrogen concentration. In order to elucidate the role of *cis*-acting sequences in the *A. caulinodans nifA* 5' upstream region and their interaction with previously identified *trans*-acting gene products involved in *nifA* expression, gel mobility shift assays were performed. A stable DNA-protein complex was observed when a radioactive DNA fragment containing the wild-type *nifA* promoter region was incubated with the crude cell lysate of *A. caulinodans*. The complex was found to have an absolute requirement for Mg^{2+} ions and to be resistant to competition by a large excess (1000-fold by weight) of non-specific DNA, demonstrating that the complex was formed by the binding of (a) sequence-specific protein(s) to the *nifA* 5' upstream probe. In order to identify which sequences were involved in complex formation, binding experiments were performed with DNA fragments carrying a variety of point-mutations in the promoter region. The stability of the DNA-protein complex was found to be identical using wt and mutant fragments, even though some of these point-mutations had previously been found to abolish transcriptional initiation completely. A complex of the same mobility was also found when extracts of *A. caulinodans nifA*, *ntrC*, *ntrY*, *ntrX*, *ntrA*, *fixLJ* and *fixK* mutants were used, suggesting that the protein(s) responsible for the observed complex formation is(are) encoded by (a) yet unidentified gene(s).

5.2 INTRODUCTION

A deletion analysis of the *A. caulinodans nifA* 5'-regulatory region has previously shown that a 229 bp fragment upstream from the start of transcription is sufficient for *nifA* gene activation and its regulation by oxygen and nitrogen (Ratet *et al.* 1989). A number of DNA sequence motifs are present in this region, which may play different roles in *nifA* gene expression. At position -174 to -143 relative to the transcription start (see fig. 1) an extensive inverted repeat (palindrome) structure is present, which may serve as a terminator preventing readthrough transcription into the *nifA* gene. If this is the case, one may assume that the essential *cis*-acting elements required for *nifA* expression are located between positions -142 and +1. In fact, several potentially important DNA motifs have been identified in this region, which may mediate the observed nitrogen- and oxygen control, as well as *nifA* autoregulation (Ratet *et al.* 1989).

A sequence motif at position -130 to -117 TTGAT-n4-ATCAA (see fig. 1), identical to the Fnr binding site consensus, has been identified (de Bruijn *et al.* 1988; Ratet *et al.* 1989). From recent experiments on the oxygen regulation of the *A. caulinodans nifA* gene it has been concluded that the FixK product positively regulates the transcription of *nifA* (Kaminski *et al.* 1991). This suggest that the Fnr motif might be the target-site for the FixK protein, since in *R. meliloti* genetic evidence for the interaction of FixK with an Fnr-binding site has been provided (Waelkens *et al.* 1992).

A DNA motif gGT-n10-ACA has been identified at position -97 to -82 (see fig. 1) sharing homology with the NifA upstream activator sequence (UAS; TGT-n10-ACA), which is usually found about 100 bp upstream of *nif* promoters and has been shown to be essential for NifA-mediated gene regulation (Buck *et al.* 1986). Autoregulation of the *A. caulinodans nifA* gene by NifA + NtrA has been suggested (Ratet *et al.* 1989). Binding of the NifA protein to UASs of the *nifH*, *nifJ* and *nifU* promoters in *K. pneumoniae* has been demonstrated using *in vivo* dimethylsulphate protection experiments (Morett and Buck 1988; Cannon *et al.* 1990) and binding by intact NifA from *K. pneumoniae* to the UAS of *nifH* and *nifJ* has recently been demonstrated *in vitro* by DNase I footprinting (Lee *et al.* 1993). Based on the position and homology in the *A. caulinodans nifA* promoter, the -97 to -82 sequence in the *nifA* promoter may therefore be involved in binding of NifA.

As discussed above (Chapter 3) a -24/-12 promoter element is present in the *nifA* 5' upstream region and its function has been confirmed by site-directed mutagenesis. NtrA (RpoN, σ54)-RNApolymerase have been shown to bind *in vivo* and *in vitro* to this element (Sasse-Dwight and Gralla 1988; Ninfa *et al.* 1987; Cannon *et al.* 1993).

The presence of these DNA motifs in the *nifA* 5' upstream region suggests that they may serve as binding sites for proteins involved in transcriptional regulation. One line of experiments designed to shed some light on this question is the examination of *in vitro* protein-DNA interactions by gel shift assays (Lane *et al.* 1992). Here we report that the -138 to +9 region of the *A. caulinodans nifA* 5' upstream region indeed interacts with (a) proteinaceous factor(s) from *A. caulinodans* extracts, but that this protein-DNA interaction appears to be independent of NifA, NtrC, NtrX, NtrA(RpoN) and FixK proteins. Moreover, it is shown that the binding of *A. caulinodans* protein to the -138 to +9 *nifA* 5' upstream region is not affected by mutations in the Fnr/FixK binding site consensus or -24/-12 promoter element.

5.3 MATERIALS AND METHODS

5.3.1 Bacterial strains, plasmids, and growth media

Bacterial strains and plasmids used in this study are listed in Table 1.

Table 1. Bacterial strains and plasmids used in this study

Strain	Relevant characteristics	Source or reference
A.caulinodans		
ORS571	wild type, Cb^r; Nif^+, Nod^+, Fix^+	Dreyfus *et al.* 1988
ORS571N15-2c	*ntrA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	this work, chapter 4
ORS571N136-3c	*ntrA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	this work, chapter 4
ORS571A5	*nifA*::Tn*5*, Cb^r, Km^r, Nif^-, Nod^+, Fix^-	Pawlowski *et al.* 1987
ORS571C6	*ntrC*::Tn*5*, Cb^r, Km^r	Pawlowski *et al.* 1987
ORS571YX1215	*ntrY*::Tn*5*, Cb^r, Km^r	Pawlowski *et al.* 1991
ORS571Y5	*ntrY*::Tn*5*, Cb^r, Km^r	Pawlowski *et al.* 1991
ORS571X2816	*ntrX*::Tn*5*, Cb^r, Km^r	Pawlowski *et al.* 1991
ORS571/57601	*fixK*$^-$, Cb^r, Km^r, Nif^-, Fix^-	Kaminski *et al.* 1991
ORS571/57592	*fixLJ*$^-$, Cb^r, Km^r, Nif^-, Fix^-	Kaminski and Elmerich 1991

Table 1 - continued

Plasmids	Relevant characteristics	Source or reference
pAP118	Ap^r; ORS571 *nifA* promoter region in pUC118 used as template for PCR	this work, chapter 3
pJSwp	Ap^r, Gm^r, wildtype ORS571 *nifA* promoter region fused to 8th codon of *lacZ* in pJS4812 for cointegration into the ORS571 chromosome	this work, chapter 3
pJSn-mutants	-24/-12 region	this work, chapter 3
pJSF4	as pJSwp but A -> C at -118 (in FixK/Fnr binding site consensus)	this work
pJSF1	as pJSwp but T -> G at -130 (in FixK/Fnr binding site consensus)	this work

Azorhizobium caulinodans ORS571 strains were grown at 37°C in TY (Beringer 1974), YLS (Elmerich *et al.* 1982) or in LSO medium (Elmerich *et al.* 1982), supplemented with 40 mg/l of nicotinic acid and 0.1% of the desired nitrogen sources. Antibiotics were used at the following concentrations: carbenicillin (Cb; 500 µg/ml) and kanamycin (Km; 200 µg/ml).

5.3.2 Labeling of DNA fragments for binding studies

A 152 bp fragment containing the *nifA* promoter region, spanning from bp -138 to +9 (see fig. 1) was prepared by the polymerase chain reaction (PCR). Radioactive labeled fragments were generated by including 3 µl [α-^{32}P]-dATP (10 µCi/µl; 6000 Ci/mmol) in the PCR reaction (section 5.3.3). The PCR-generated fragments were separated from unincorporated ^{32}P-labeled nucleotides by using a Pharmacia NICK Column (2x) and directly used in DNA-binding studies. Specific competitor PCR-generated fragments were extracted with phenol/chloroform/isoamyl alcohol (25/24/1), chloroform/isoamyl alcohol (24/1), and precipitated with 1/10 vol 3M Na-acetate pH 4.8 and 2 vol 100% Ethanol (30 min at -70°C). The precipitated DNA was collected by 15 min centrifugation at 13000 rpm at 4°C. The pellet was washed with 70% Ethanol, dried and dissolved in water. The DNA concentration was measured spectrophotometrically at OD_{260}. Non-specific competitor DNA (pUC118, salmon sperm DNA) was digested with *Msp*I and treated as described above for the PCR-generated specific competitor DNA fragments.

5.3.3 Polymerase chain reaction (PCR) conditions

Each 100µl PCR reaction contained 50 pmol each of 2 opposing primers (5'-GAATTCAAAATTTGATCCAGATC-3' and 3'-CGGCGAAAACGCCGTCCCTAGG-5'; see fig. 1), 5-25 ng of template (plasmid pAP118) DNA, 15 µM of each of 4 dNTPs, 2U AmpliTaq DNA polymerase (Perkin-Elmer/Cetus) in Gitschier buffer (16.6 mM $[NH_4]_2SO_4$; 67 mM Tris-HCl pH 8.8; 6.7 mM $MgCl_2$; 6.7 µM EDTA; 30 mM β-mercaptoethanol and 170 µg/ml BSA; Kogan *et al.* 1987) with 10% DMSO (v/v). PCR amplifications were performed in an automated thermal cycler (Perkin-Elmer/Cetus DNA Thermal Cycler) with an initial denaturation (94°C, 5 min) followed by 30 cycles of denaturation (94°C, 1 min), annealing and extension (49°C, 40 sec), and a single final extension (72°C, 5 min).

5.3.4 Preparation of lysates from cultured cells

Azorhizobium caulinodans ORS571 cells were grown with continuous shaking in 10 ml LSO medium supplemented with 0.05% L-glutamine and 0.05% ammonium sulfate at 37°C. An 1 ml aliquot of a saturated culture was re-inoculated in 100-200 ml LSO medium, containing 0.05% ammonium sulfate (3% oxygen), for 2 days. The cells were collected by centrifugation, (Sorvall RC-5B Refrigerated superspeed centrifuge; GSA rotor) and the bacterial pellet was resuspended in ice-cold extraction buffer [NXB: 20 mM Hepes pH 7.9, 12% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF (freshly added), 0.5 mM DTT (freshly added)]. Cells were lysed by three passages through a French press at 20,000 psi (French pressure cell press, American Instrument Company, Silver Spring, Maryland 20910; Aminco French pressure cell). Unbroken cells were removed by centrifugation in a Sorvall SS34-rotor (15K rpm, 30 min, 4°C), the supernatant fluid (cleared lysate) was stored in small aliquots at -70°C and used for binding studies. The protein concentrations of cleared lysate were estimated using the Bicinchoninic acid protein determination kit (Sigma), using bovine serum albumin as the standard.

5.3.5 DNA-protein binding assay

DNA-binding ability was measured by gel retardation (gel mobility shift) assays (Lane *et al.* 1992). About 6 ng of a radiolabeled DNA fragment (10,000 cpm) was incubated with 4 µl (16-20 µg protein) and competitor DNA, in a total volume of 30 µl, 1x binding buffer (25 mM HEPES pH 7.8, 1 mM DTT, 10% glycerol). The mixture

was incubated for 30 min at room temperature and then loaded onto a 5-6% nondenaturing polyacrylamide gel (acrylamide/N,N'-methylene bisacrylamide, 29:1) in Tris-borate-EDTA buffer, pH 8.3 (10x TBE; 500 mM TRIS, 500 mM Boric Acid, 10 mM EDTA). Electrophoresis was performed at room temperature at 8 V/cm (25mAmp; 100 V). The gels were dried and exposed to Kodak X-ray film, using intensifying screens.

5.4 RESULTS

5.4.1 Complex formation of the *nifA* 5'upstream region with *A. caulinodans* extracts

A 152 bp *A. caulinodans* DNA fragment, containing the regulatory region of the *nifA* gene spanning from bp -138 to +9 (see fig. 1), was prepared and labeled using the polymerase chain reaction (PCR; see Materials and Methods). This fragment was incubated with *A. caulinodans* cleared cell lysate.

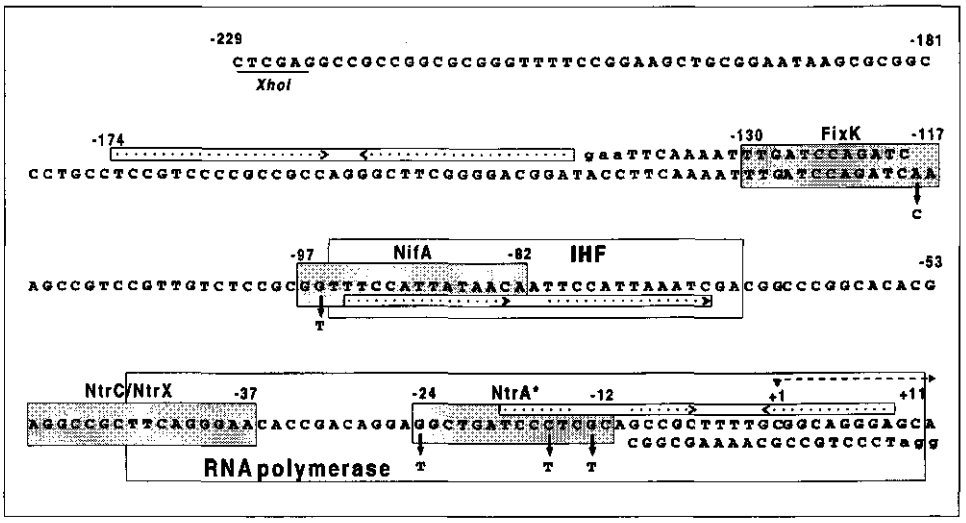

Figure 1. DNA sequence and potential protein binding motifs of the 5' upstream region of the *A. caulinodans nifA* gene.

For details see text.

No significant complex were initially found, although several very weak bands could be seen (data not shown). In order to optimize the binding conditions, various concentrations of ions were included in the binding assay. A high molecular weight complex was observed when the binding reaction was performed in the presence of 5-20 mM $MgCl_2$ (fig. 2).

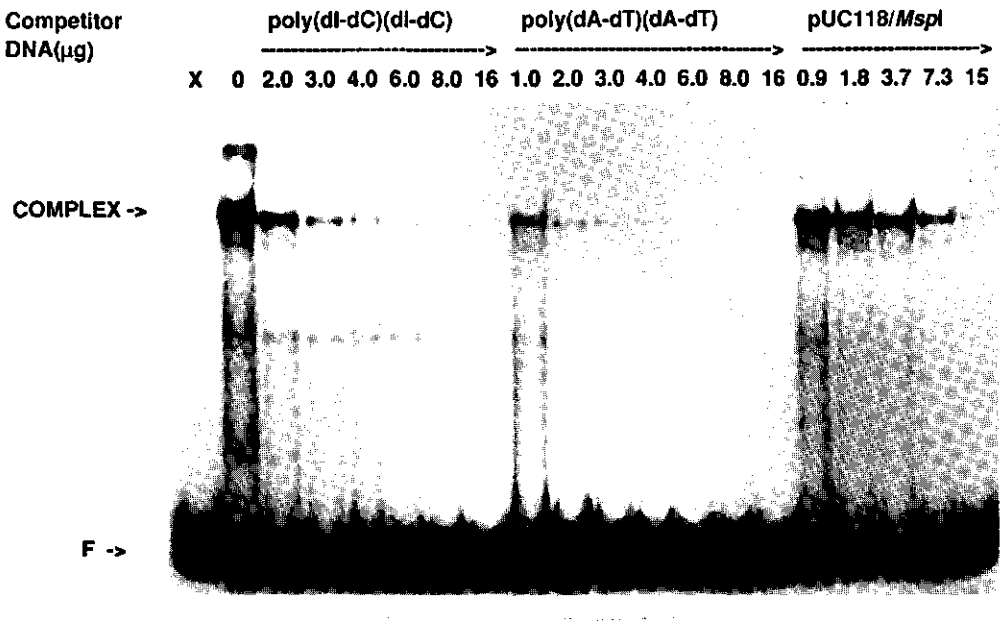

Figure 2. Gel mobility shift assay of the *nifA* 5' upstream fragment with crude protein extracts from *A. caulinodans* ORS571.

The autoradiograph of a polyacrylamide gel is shown. Each lane contains 6 ng radioactively labeled 152 bp *nifA* 5' upstream fragment.

Lane 1: no bacterial extract.

Lane 2: no non-specific competitor DNA.

Lanes 3-8: competition with increased amounts of poly(dI-dC)(dI-dC).

Lanes 9-15: competition with increased amounts of poly(dA-dT)(dA-dT).

Lanes 16-20: competition with increased amounts of pUC118/*Msp*I DNA.

The positions of free DNA (F) and protein-DNA complex are as indicated. See text for details.

This complex was not present when Mg^{2+} was substituted by another divalent ion, such as Ca^{2+} (5-33 mM). 50 mM KCl was included in the assay since monovalent ions has shown to increase the formation of DNA/protein complexes in gel mobility shift assays (Robidoux *et al.* 1992). However, no complex was observed when only K^{+}-ions (50-200 mM) were included in the assay. When the reaction was carried out in the presence of 5 mM $MgCl_2$ plus 8.3 mM EDTA, no complex formation was observed, suggesting that the DNA binding activity requires Mg^{2+} ions (data not shown).

The DNA/protein complex was specific, since it was resistant to out competition by a large excess of non-specific competitor DNA's (6 µg poly(dI-dC)(dI-dC), ~1000-fold excess by weight; 4 µg poly(dA-dT)(dA-dT), ~660-fold excess; 15 µg pUC118/*Msp*I, ~2500-fold excess; see fig. 2).

The DNA-protein complex was not only present in *A. caulinodans* ORS571 cells grown under nitrogen fixing conditions (low nitrogen, 3% oxygen), but a complex of identical mobility was observed when cells were grown in the presence of nitrogen (NH_4^{+}) and 20% oxygen (repressing conditions; data not shown).

5.4.2 Effect of mutations in conserved DNA motifs on complex formation

To identify which sequences in the *nifA* 5' upstream 152 bp fragment were involved in the observed complex formation, gel-shift experiments were performed in the presence of DNA fragments carrying a variety of point-mutations (see table 2) as specific competitor DNA's.

Table 2. Location of point-mutations in the *nifA* 5' upstream DNA motifs

wt	GGCTGATCCCTCGC	**-24/-12 promoter element**
n1	T.............	
n2	.........T....	
n3	............T.	
n12	T........T....	
n13	T...........T.	
n23	.........T..T.	
wt	TTGATCCAGATCAA	**Fnr binding site consensus**
F1	............C.	
wt	GGTTTCCATTATAACA	**NifA UAS**
U1	.T..............	

Although mutations n3, n13 (chapter 3) and F1 (data not shown) have been found to abolish *nifA* expression completely, the stability of the DNA-protein complex was identical in all cases (300 ng, about 50-fold excess of cold fragment reduced complex formation significantly; see fig. 3).

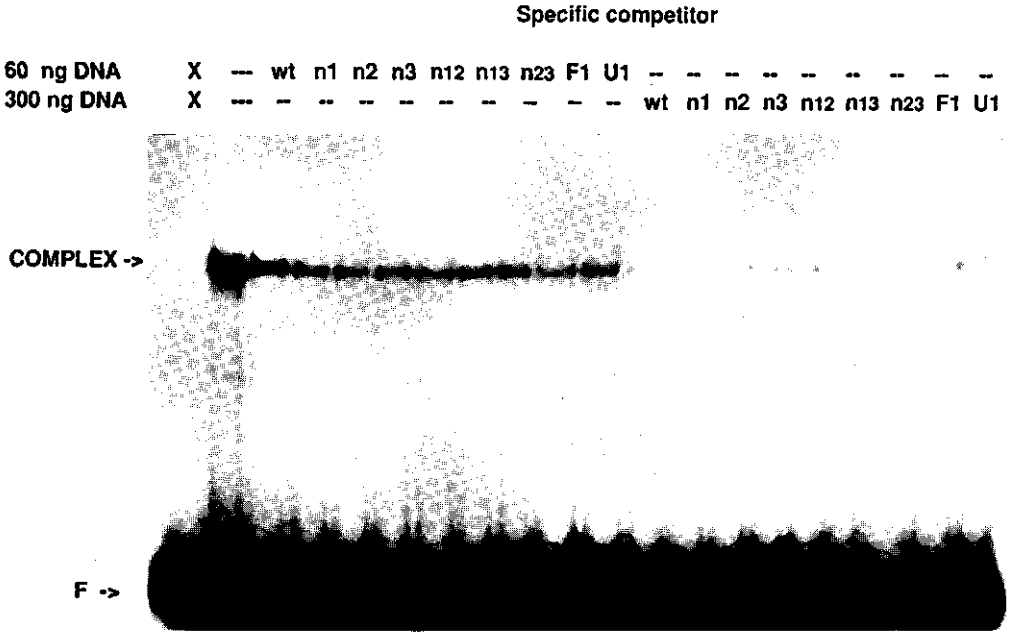

Figure 3. Gel mobility shift assay of the *nifA* 5' upstream fragment with crude protein extracts from *A. caulinodans* ORS571 in the presence of mutant DNA fragments.

The autoradiograph of a polyacrylamide gel is shown. Each lane contains 6 ng radioactively labeled 152 bp *nifA* 5' upstream fragment.

Lane 1: no bacterial extract.

Lane 2: 2 µg pUC118/*Msp*I DNA (non-specific competitor).

Lane 3-11: 2 µg pUC118/*Msp*I DNA + 60 ng of specific competitor DNA.

Lane 12-20: 2 µg pUC118/*Msp*I DNA + 300 ng of specific competitor DNA.

The positions of free DNA (F) and protein-DNA complex are as indicated. See text for details.

5.4.3 Effect of mutations in *trans*-acting regulatory genes on complex formation

To further examine the DNA-protein complex formation, binding assays were performed with cleared cell extracts from various (Tn*5)* mutants of *A. caulinodans*, known to be affected in the regulation of *nifA* expression, such as *nifA*, *ntrC, ntrY, ntrX, ntrA, fixLJ* and *fixK*. In all mutant extracts the complex was present, suggesting that the protein(s) responsible for the observed gel shift (protein-DNA interaction) is (are) encoded by (an) yet unidentified gene(s) (data not shown).

5.4.4 DNA search for additional binding motifs

Since neither the experiments using extracts of the *A. caulinodans* strains carrying mutations in various regulatory loci or the mutant oligo competition studies revealed (an) obvious candidate(s) for the protein(s) interacting with the *nifA* 5' upstream region, a search was initiated for alternative putative target sites for *trans*-acting factors on the -138 to +9 fragment.

A DNA motif TAaCAA-n4-aTt at position -86 to -74 (see fig. 1) sharing homology with the IHF binding consensus (A/TATCAA-n4-TTA/G; Craig and Nash 1984; Kur *et al.* 1989), was identified. It has been shown that IHF binds adjacent to the σ54 promoter regions from a variety of (nitrogen-fixing) organisms, mediating the formation of a specific bend that brings the upstream-bound regulatory protein into proper proximity to RNA polymerase and enhances transcriptional activation. IHF stimulates NifA-mediated activation of several *nif* promoters (Hoover *et al.* 1990), NtrC-mediated activation at the *glnH* promoter (Claverie-Martin and Magasanik 1991) and XylR-mediated activation of the Pu promoter of the upper catabolic operon of the *Pseudomonas putida* TOL plasmid (de Lorenzo *et al.* 1991). Recently it has been found that a very similar mechanism may also be present in the σ70-dependent *nar* operon promoter in *E. coli* (Rabin *et al.* 1992).

Since this motif is located between the FixK binding site and the -12/-24 promoter it may be possible that *A. caulinodans* IHF plays a role in positioning FixK into the proper proximity to σ54-RNA polymerase and facilitate protein-protein interaction.

A direct repeat sequence (TTCCATTAtAaC-n2-TTCCATTAaAtC) at position -94 to -69 (fig. 1) was also identified, which may be involved in the formation of a secondary structure and/or binding of an (yet) unidentified regulatory protein, thereby generating the observed complex formation.

A DNA motif AgGCCgcTT-n2-GGGaA was found at position -52 to -37 (fig. 1) which shares homology with the weak NtrC binding site 5 (AcGCCttTT-n2-GGGcA) found at the exact same position (-52 to -37) in the *E. coli* and *S. typhimurium glnA* promoter regions. *A. caulinodans nifA* expression has been shown to be controlled by the *ntr* system, mediated by the nitrogen regulatory genes *ntrC* and *ntrX* (see section 2.3.7). Although no high affinity NtrC binding-sites (TGCACC-n4-TGGTGCA; Ames and Nikaido 1985; Reitzer and Magasanik 1986, Sasse-Dwight and Gralla 1988) could be found in the *nifA* 5' upstream region, the weak site could, nevertheless, be involved in NtrC/NtrX binding, under high concentrations (*in vivo*). Binding assays with extracts from strains which over-produce these transcriptional factors may adress this question.

5.5 DISCUSSION

The *in vitro* interaction of proteins present in *A. caulinodans* crude cell extracts with the *nifA* regulatory region was studied by gel retardation assays. A stable DNA-protein complex was observed, which was resistant to competition by a large excess (about 1000-fold by weight) of non-specific DNA, demonstrating that the complex was formed by the binding of (a) sequence-specific protein(s) to the probe.

The complex formation was found to have an absolute requirement for Mg^{2+} ions; no complex was observed when $MgCl_2$ was omitted, or when other divalent ions were included in the binding reaction. In *Streptomyces griseus* binding of an A-factor-dependent protein to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis, has been shown to be greatly enhanced by 5 mM Mg^{2+} (Vujaklija *et al.* 1993). A possibility is that phosphorylation of the DNA binding protein (requires ATP and Mg^{2+}) increase its binding affinity, as found for the NtrC protein, where phosphorylation increased the affinity for DNA by a factor of 2 (Ninfa *et al.* 1987).

The observation that the 152 bp promoter fragment undergoes a large mobility shift can be explained in several ways.

A. It could be due to interactions of several binding proteins with several binding sites, generating a multi-protein complex. These multi-protein complexes could contain one kind of protein or several different proteins. Also binding could be occurring to several sites on the 152 bp promoter fragment. A in depth analysis of deletion derivatives of the fragment and competition studies using oligonucleotides containing the proposed binding sites, and using extracts of various mutants will be necessary to answer this question.

B. It is possible that the observed gel shift of the -138 to +9 fragment is due to interaction with a small protein, which induce a strong bend in the DNA (like IHF). Binding of IHF to DNA has been shown to induce a sharp bend of greater than 140^{o}, which results in a strongly retarded migration of the fragment (Prentki *et al.* 1987; Stenzel *et al.* 1987). Binding of IHF to the 287 bp Mu operator region has been shown to result in a strongly retarded band, which migrated with an apparent size of 1250 bp (Alazard *et al.* 1992). Mg^{2+} has been shown to be necessary for IHF function however, it is not known if Mg^{2+} is also essential or increases IHF binding activity (Friedman, pers. comm.). Therefore, it is possible that IHF mediated bending is responsible for the complex migration, observed with the *A. caulinodans nifA* promoter fragment since a DNA motif TAaCAA-n4-aTt at position -86 to -74 (see fig. 1) sharing homology with the IHF binding consensus sequence (Craig and Nash 1984; Kur *et al.* 1989), is indeed present.

It has been reported previously that IHF binds to *nif(H)* promoter regulatory regions from *K. pneumoniae*, *A. vinelandii*, *R. meliloti*, *B. japonicum*, *R. rubrum*, *R. capsulatus* and *T. ferrooxidans* (Hoover *et al.* 1990). These binding sites were defined by the consensus $^{A}/_{T}$ATCAA motif surrounded by AT-rich sequences (55-78% AT). The putative IHF binding site in the *A. caulinodans nifA* promoter region contains one mismatch with the IHF consensus sequence. In addition the 29 bp region surrounding this motif (position -95 to -57) is very AT-rich (76%).

It is also possible that the *A. caulinodans nifA* fragment undergoes a large shift due to the interaction with a large protein (like NtrA*-RNA polymerase). Recently, it has been shown by gel mobility shift assays that σ54-RNApolymerase binds specifically to the *R. meliloti nifH* -24/-12 promoter element, and that it forms a high molecular weight complex (Cannon *et al.* 1993). Binding assays were conducted in the presence of 8 mM magnesium acetate, however no data are available about a specific role for Mg^{2+} in complex formation (Buck, pers. comm.). Therefore, the high molecular weight complex found in gel retardation assays with extracts of *A. caulinodans* and the *nifA* 5' upstream region might be due to NtrA(RpoN)-RNA polymerase binding to the -24/-12 element. Since this complex was present in extracts of all mutants, including the *ntrA* mutant, the NtrA(RpoN) protein in this complex might be encoded by another yet unidentified copy of *ntrA* (*ntrA**,σ54*; see chapter 8).

To determine which DNA sequences on the *nifA* 5' upstream fragment were involved in complex formation, binding experiments were performed using DNA fragments carrying a variety of point-mutations in the promoter region as specific competitor DNA. However, the stability of the DNA-protein complex was identical in all cases, even when the mutated fragment carried basepair changes in the -24/-12 region, which abolished transcriptional initiation completely. This could be explained by the fact that the mutations only abolish transcriptional activation but do not affect binding (or that binding occurs at a different site of the mutations; for example, at IHF binding site).

The complex was present in extracts of *A. caulinodans* grown under nitrogen fixing (3% oxygen, no $NH4^+$) and repressing (20% oxygen, 0.1% $NH4^+$) conditions. This suggest that the protein(s) is(are) constitutive expressed and is(are) bound to the *nifA* 5' upstream region under both conditions and may, therefore, need to be modified/activated under nitrogen fixing conditions.

It was surprising that no differences in complex formation were observed with extracts of the various *trans*-acting regulatory mutants. It might be that the regulatory proteins are present in too low concentrations in the cell free extract to identify a complex under these assay conditions. Another reason may be that some regulatory proteins are inactive under aerobic conditions. Although the extracts were made from cultures grown under microaerobic conditions, the preparation of extracts was carried out aerobically. It might be necessary to prepare the cell

extracts under anaerobic conditions. Anaerobically isolated cell extract of wild-type *B. japonicum* cells grown under anaerobic conditions was shown to contain a protein (FixK?) which could bind to the Fnr/FixK binding site consensus in the *B. japonicum rpoN1* promoter region (this complex was not observed with a fragment containing a deletion of the Fnr/FixK binding site). It was shown that aerobic conditions during the preparation of anaerobic or aerobic grown cell extract abolished the binding abilities and complex formation (I. Kullik, PhD thesis 1992). The same phenomenon may apply here.

In summary, the specific protein-DNA complex identified here is not caused by previously characterized *A. caulinodans trans*-acting factors, its binding ability depends on Mg^{2+} ions and is not affected by point mutations in the Fnr-binding site consensus sequence, NifA UAS or -24/-12 promoter element. Therefore, these experiments did not yield further insights in *A. caulinodans nifA* promoter regulation, and lead to the pursuit of alternative strategies, including a search for additional *trans*-acting genes by using heterologous DNA probes (e.g. *fnr*, see chapter 7).

CHAPTER 6

COPROPORPHYRIN EXCRETION BY *A. CAULINODANS FIXLJ* AND *FIXK* MUTANTS

6.1 ABSTRACT

fixLJ and *fixK* mutant strains of *A. caulinodans* were found to excrete large amounts of a red/pink, UV fluorescent, pigment into the culture medium when grown under microaerobic conditions (3% oxygen). The predominant pigment was isolated and identified on the basis of its spectrophotometric properties. The absorption spectrum of the neutral methyl ester of the compound suggested that the pigment could be coproporphyrin III, the oxidation product of coproporphyrinogen, which is an intermediate compound in heme-biosynthesis. The excretion was found to be independent of N-sources, but to be strictly dependent on microaerobic conditions. These results suggest that *A. caulinodans* has an anaerobic coproporphyrinogen oxidase enzyme, the expression of which may be regulated by the oxygen sensing regulatory pathway (*fixLJK*). This leads to the suggestion that nitrogen fixation (nitrogenase production) and heme biosynthesis (leghemoglobin production in the nodule) may be coordinately regulated by the *fixLJK* regulatory system.

6.2 INTRODUCTION

6.2.1 Heme biosynthetic pathway and designation of the corresponding genes

The biochemistry of heme synthesis is well established and, with the exception of the initial reactions leading to δ-aminolevulinic acid (ALA), the heme pathway (see fig. 1) is conserved among many different species. In different organisms, ALA is made either by a C5 route from glutamate or by a C4 route from succinyl coenzyme A and glycine (Avissar *et al.* 1989). The heme-biosynthetic pathway can be divided into three parts: (i) the synthesis of ALA; (ii) the assembly of eighth ALA molecules to give uroporphyrinogen III (the first tetrapyrrole); and (iii) the modification of tetrapyrrole side chains, oxidation, and iron (Fe^{2+}) insertion to yield heme, the prosthetic group of proteins such as myoglobin, catalase, peroxidase, and cytochrome c.

In the first step of the heme biosynthetic pathway, two molecules of δ-aminolevulinic acid condense and form the basic pyrrole building block, porphobilinogen. This reaction is catalyzed by δ-aminolevulinic acid dehydratase (EC 4.2.1.24; also called ALAD, porphobilinogen synthase, ALA dehydrase, and PBG synthase) encoded by the *hemB* gene.

In the next step four molecules of porphobilinogen are polymerized into a linear tetrapyrrole chain, to generate a highly unstable hydroxymethylbilane called preuroporphyrinogen. This step is catalyzed by porphobilinogen deaminase (EC 4.3.1.8; also called PBGD, hydroxymethylbilane synthase, and uro I synthase) encoded by the *hemC* gene.

Then, uroporphyrinogen III synthase (EC 4.2.1.75; also called uro III synthase and uro III cosynthase) encoded by the *hemD* gene, carries out an reaction in which the fourth pyrrole-ring of preuroporphyrinogen is rearranged and the molecule is cyclized to give the key intermediate uroporphyrinogen III. Uroporphyrinogen III is the universal precursor from which porphyrins, hemes, chlorophyll's, corrins, and all other tetrapyrroles are derived.

The next step is the decarboxylation of all four acetic acid side chains of uroporphyrinogen III to methyl groups yielding coproporphyrinogen III. This reaction

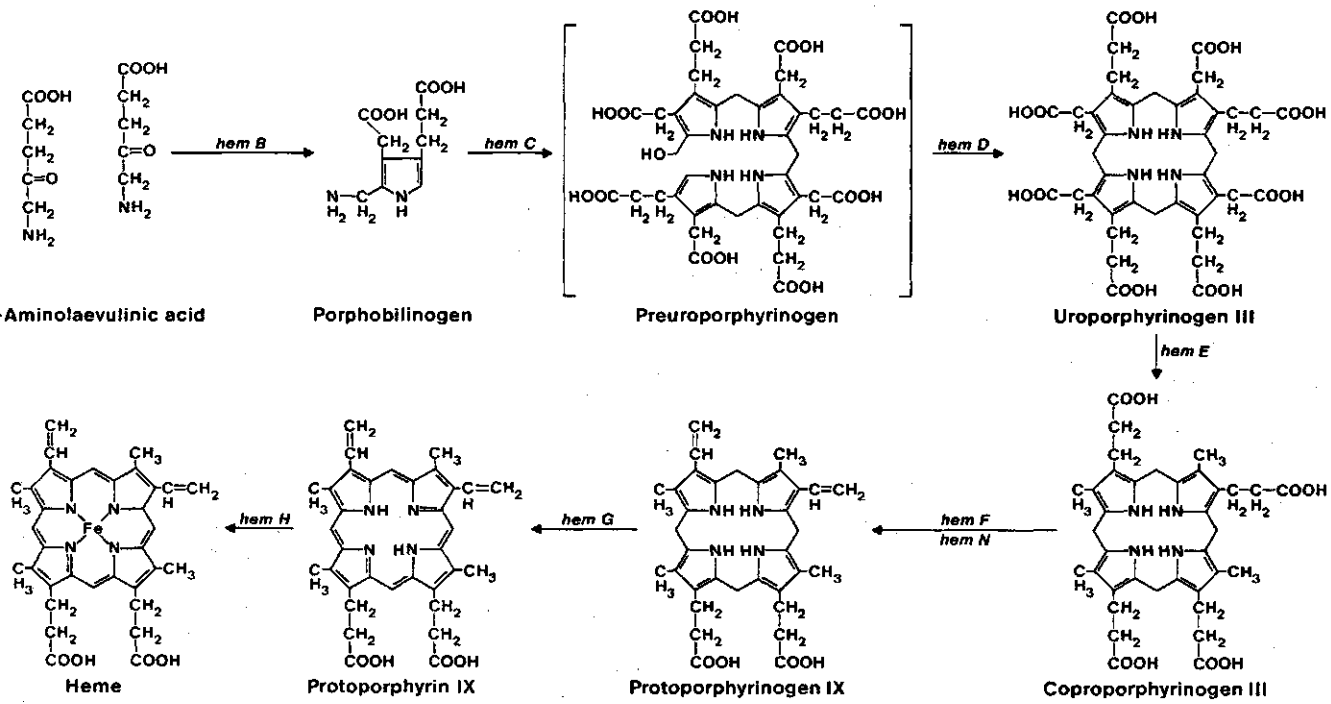

Figure 1. Heme biosynthetic pathway and designation of corresponding genes.

See text for details.

is catalyzed by uroporphyrinogen decarboxylase (EC 4.1.1.37; also called Uro D and uro'gen decarboxylase) encoded by the *hemE* gene (Jordan 1990).

Coproporphyrinogen III is converted into protoporphyrinogen IX by the oxidative decarboxylation of two propionate side chains into vinyl groups. The enzyme responsible for this reaction is coproporphyrinogen oxidase (EC 1.3.3.3; also called CPO, copro'gen oxidase, and coproporphyrinogenase). Two types of coproporphyrinogen oxidase have been described previously (Dailey 1990). One enzyme, encoded by the *hemF* gene requires oxygen as a substrate. The second enzyme, encoded by the *hemN* gene is oxygen independent, and its activity accounts for the ability of many bacteria to synthesize heme under anaerobic conditions (Xu *et al.* 1992).

The next step is the (six-electron) oxidation of the porphyrinogen to porphyrin, catalyzed by protoporphyrinogen oxidase (EC 1.3.3.4; also called PPO and proto'gen oxidase), encoded by the *hemG* gene.

The final step, the insertion of iron, is catalyzed by ferrochelatase (EC 4.99.1.1; also called Fc, protoheme ferrolyase, heme synthase, and heme synthetase) encoded by the *hemH* gene (Dailey 1990).

6.2.2 Heme biosynthesis in rhizobia

Of particular interest is the regulation of heme production in the obligate aerobic rhizobia. Inside the legume nodule an oxygen-binding protein, leghemoglobin (legume hemoglobin) is found in very high concentrations; it facilitates the diffusion of oxygen to the rapidly respiring bacteroids (differentiated, N_2-fixing bacteria) and concomitantly buffers the free O_2 concentration at an extremely low tension, to prevent inactivation of the oxygen-labile nitrogenase enzyme (Appleby 1984).

Leghemoglobin appears to be a truly symbiotic protein; The apoprotein is a plant product, whereas the heme prosthetic group may be synthesized by the bacterium (Nadler and Avissar 1977, Sangwan and O'Brian 1991). Bacterial heme synthesis has been shown to be essential for leghemoglobin expression in soybean nodules (O'Brian *et al.* 1987), alfalfa nodules (de Bruijn *et al.* 1989) and *Sesbania rostrata* nodules (Pawlowski *et al.* 1993).

Studies on heme biosynthesis in rhizobia has raised questions about how the plant and the bacterium interact at the molecular level and it has been suggested that bacterial heme, or a heme precursor, induces leghemoglobin apoprotein synthesis by the plant (Appleby 1984; O'Brian *et al.* 1987).

Heme synthesis in rhizobia is also interesting from another point of view. The late steps in heme biosynthesis require oxygen as a substrate in the two oxidative steps (see fig. 1); however inside the legume nodule, the O_2 concentration is only about 10 nM. This suggests that two mechanisms for heme biosynthesis exist, one operating under aerobic free-living conditions and one present under micro- or anaerobic conditions during symbiosis. Keithly and Nadler (1983) found that *B. japonicum* bacteroids contain a $NADP^+$-dependent coproporphyrinogen oxidase activity under anaerobic conditions. It is not yet known whether rhizobial species have an anaerobic protoporphyrinogen oxidase. It is also unclear how anaerobic (microaerobic) induction of such a pathway may be mediated.

As described above, the microaerobic induction of *nif/fix* gene expression in rhizobia is dependent on the *fixLJ/fixK* regulatory system (chapter 2). FixL and FixJ belong to a family of two-component sensor-regulator proteins which respond to microaerobic conditions. FixL, which has features of a transmembrane protein, is the oxygen sensor and FixJ the transcriptional activator. It is interesting to note that fixL is a heme containing protein (Gilles-Gonzalez *et al.* 1991). FixK is homologous to Fnr, the transcriptional regulator controlling the expression of genes involved in anaerobic respiration in *E. coli*. FixK is involved in activating the nitrogen fixation genes via *nifA*, while the activity of FixK itself is regulated by the *fixLJ* genes (see chapter 2 and 8).

FixLJ and *fixK* mutants of *A. caulinodans* have been constructed (Kaminski and Elmerich 1991; Kaminski *et al.* 1991). While analyzing the interaction of DNA binding proteins to the *nifA* 5' upstream region (see chapter 5), cultures of *A. caulinodans fixLJ* and *fixK* mutants were grown under nitrogen fixing conditions, for cell-free extract preparations. It was found that these cultures excreted large amounts of coproporphyrin into the medium, but only when grown 2-3 days under low oxygen concentrations. This observation suggests that the *fixLJ* and *fixK* genes involved in the regulation of nitrogen fixation in *A. caulinodans* ORS571 may be involved in the regulation of heme biosynthesis as well.

6.3 MATERIALS AND METHODS

6.3.1 Bacterial strains

Azorhizobium caulinodans ORS571 (wt, Cb^r; Nif^+, Nod^+, Fix^+) has been described by Dreyfus *et al.* (1988), ORS571/57592 (*fixLJ*$^-$, Cb^r, Km^r, Nif^-, Fix^-) has been described by Kaminski and Elmerich (1991), and ORS571/57601 (*fixK*$^-$, Cb^r, Km^r, Nif^-, Fix^-) has been described by Kaminski *et al.* (1991).

6.3.2 Isolation of pigment from culture medium

Azorhizobium caulinodans cells were grown with continuous shaking in 100 ml LSO medium, supplemented with 0.05% ammonium sulfate under 3% oxygen at 37°C for three days. The cells were collected by centrifugation, (Sorvall RC-5B Refrigerated superspeed centrifuge; GSA rotor) and the supernatant containing the red/pink fluorescent pigment produced by the *fixLJ* and *fixK* mutants was acidified to pH 3.5 with acetic acid and extracted with diethyl ether. To obtain the acid absorbtion spectrum, the pigment was extracted from the ether solution with 0.1 N HCl. To obtain a neutral absorption spectrum the solution was evaporated to dryness and the pigment was dissolved in methanol-sulfuric acid to esterify the carboxylic side-chains (Fuhrhop and Smith 1975).

6.3.3 Esterification of porphyrins using methanol with sulfuric acid

The carboxylic side-chains of the excreted pigment were esterified with pure, dried (with anhydrous Na_2SO_4), filtered methanol containing 5% w/v concentrated sulfuric acid for 20 hr at room temperature in the dark, since some porphyrins, especially those with unsaturated side-chains, are light-sensitive. After esterification, the porphyrin ester was extracted into chloroform, washed with 2N Na_2CO_3, followed by water, and evaporated to dryness (Fuhrhop and Smith 1975).

6.4. RESULTS

6.4.1 Identification of red pigment in *A. caulinodans fixLJ* and *fixK* mutants

It was observed that liquid cultures of *A. caulinodans fixLJ* and *fixK* mutants, turned pink/red and were fluorescent under UV-light after growing for at least two days under 3% oxygen. Separation of the cells from the growth medium revealed that a red/pink fluorescent pigment was excreted into the medium. When *A. caulinodans* cells were grown under 3% oxygen, the supernatant (growth medium) remained colorless; however the cell pellet possessed a light red/pink color. In contrast, the cell pellets from the pigment excreting mutants were white.

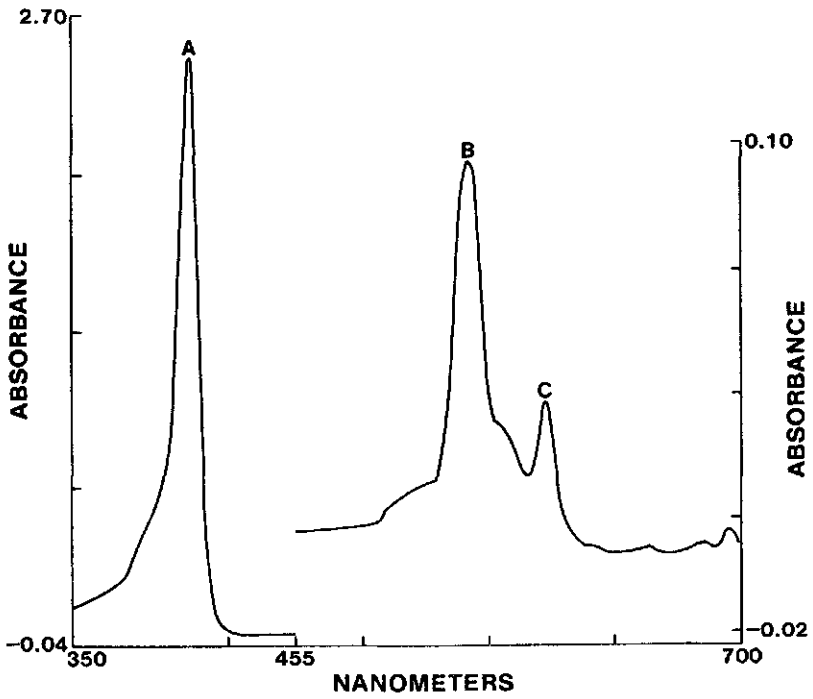

Figure 2a. Absorption spectrum in 0.1 M HCl.

The absorption spectrum from 350 to 700 nm of the pigment excreted by the *fixLJ* and *fixK* mutants of *A. caulinodans* grown in liquid culture under nitrogen fixing conditions, is shown.

It was found that the pigment started to accumulate after two days of growth under 3% oxygen. The excretion of the pigment was not dependent on N-limitation in the medium, since it was also found to occur in cultures grown in full, YLS medium. However the excretion seemed to be controlled by oxygen availability, since no accumulation was found in mutants grown under 20% oxygen.

The visible absorption spectra of the supernatant of *fixLJ* and *fixK* mutants was found to have a strong absorption band around 400nm (Soret band), which is characteristic of the aromatic porphyrin macrocycle. The absorption spectrum of the isolated pigment in aqueous acid (0.1 M HCl; fig. 2a) showed a sharpened Soret band (maximum at 401.5 nm), and a two-banded absorption spectra, typical of porphyrin dications, in the 540 to 600 nm region (maxima at 547.5 and 590.0 nm).

The absorption spectrum of the methyl ester of the pigment in neutral solvent such as chloroform (fig. 2b), showed a strong absorption band at 401.5 nm (Soret band) and a four banded spectra in the 500 to 640 region (maxima at 499.5, 532.5, 567.0, and 622.0). On basis of these absorption spectra (Smith 1975) the pigment was identified as coproporphyrin.

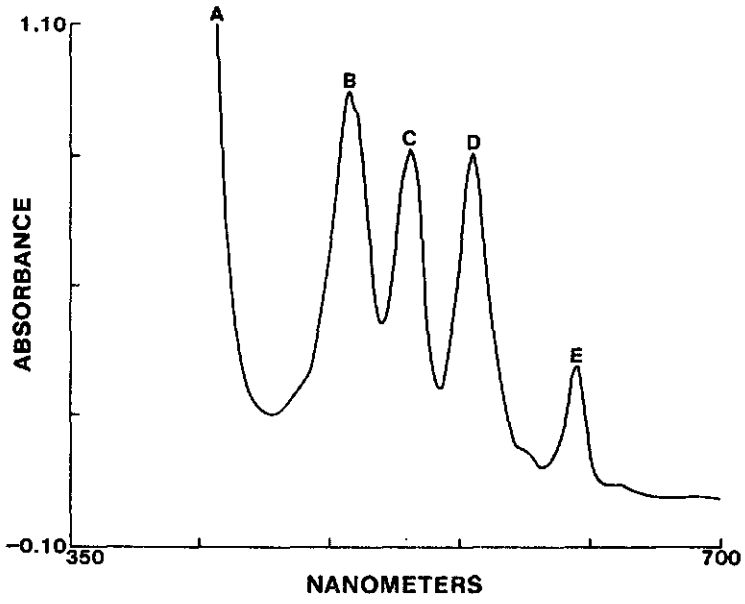

Figure 2b. Absorption spectrum of the neutral methyl ester in chloroform.

The absorption spectrum from 350 to 700 nm of the pigment excreted by the *fixLJ* and *fixK* mutants of *A. caulinodans* grown in liquid culture under nitrogen fixing conditions, is shown.

DISCUSSION

6.5.1 Heme biosynthesis in (Azo)rhizobia

(Azo)rhizobia can live either free in the soil and in laboratory culture or endosymbiotically in infected host cells of the central nodule tissue. The oxygen concentrations in these habitats may differ by a factor of up to 10^5, ranging from ~250 μM in O_2-saturated environments, to 3-30 nM in endosymbiosis (Layzell *et al.* 1990). It seems that aerobic rhizobia have adapted to this very low oxygen environment in several ways.

To generate sufficient ATP for their own maintenance and particularly for the energy-demanding N_2-fixation process, rhizobia(bacteroids) express a micro-aerobically induced symbiosis-specific oxidase complex, with a terminal oxidase with a high affinity for oxygen (Appleby 1984). Recently a symbiotically essential *fixNOQP* gene cluster in *B. japonicum* was shown to encode a membrane-bound, cytochrome c-containing heme/copper oxidase which may constitute the bacteroid oxidase operating at extremely low free O_2 tensions (Preisig *et al.* 1993).

To synthesize heme under different O_2 tensions, rhizobia seem to use two mechanisms for coproporphyrinogen oxidation. The first is an oxygen dependent process, carried out by an aerobic coproporphyrinogenase, with an absolute requirement for molecular oxygen as the electron acceptor, and the second is an oxygen independent process, carried out by an anaerobic coproporphyrinogenase, which transfers electrons to $NADP^+$ and has an allosteric requirement for S-adenosyl-L-methionine. Both enzyme activities have been identified in *B. japonicum* (Keithly and Nadler 1983).

The *fixLJ* and *fixK* mutant strains of *A. caulinodans* were found to excrete large amounts of coproporphyrinogen III into the culture medium, when grown under microaerobic conditions (3% oxygen) for several days. No pigments were excreted under aerobic conditions or in wild-type *A. caulinodans*. Excretion of coproporphyrinogen in the *A. caulinodans fixLJ* and *fixK* mutants suggest a defect in the enzyme coproporphyrinogenase, required for the conversion of coproporphyrinogen III to protoporphyrinogen IX (fig. 1). Since excretion of this heme precursor was not observed under aerobic conditions, the defect might be in a specific anaerobic coproporphyrinogenase.

The gene for a specific anaerobic coproporphyrinogenase, *hemN*, was recently identified in *S. typhimurium*, and *hemN* mutants were reported to (require heme and) accumulate uroporphyrin and coproporphyrin under anaerobic conditions but to have a wild-type phenotype under aerobic conditions (Xu *et al.* 1992).

A strain of *Rhodobacter sphaeroides* mutated in a gene which encodes a putative anaerobic copro-porphyrinogen III oxidase was shown to excrete large quantities of coproporphyrin into the culture medium. Under highly aerated culture conditions, however, the mutant produced extremely low levels of coproporphyrin (Coomber *et al.* 1992).

It is possible that *A. caulinodans fixK* positively regulates the expression of a *hemN*-like gene, which encodes the oxygen-independent form of copro-porphyrinogen oxidase, resulting in the accumulation of coproporphyrin, only under microaerobic or anaerobic conditions. Assays of coproporphyrinogen oxidase activity in cell extracts of the *A. caulinodans* ORS571 wild type strain vs *fixLJ* and *fixK* mutants, under low and high oxygen, might address this question.

It is interesting to note that the promoter region of the putative *hemF* (*hemN)* gene of *R. sphaeroides* contains a sequence (TTGAT-n4-cgCAA, at -34 from the transcription start) sharing distinct homology with the Fnr/FixK binding site consensus (TTGAT-n4-ATCAA) found in *E. coli* genes, which are (positively) regulated by Fnr under anaerobic conditions (Spiro and Guest 1990).

It has been shown that under restricted aeration *B. japonicum* cells have elevated levels of ALA synthase and ALA dehydratase activities (about 10-fold), and excreted porphyrins into the growth medium (Avissar and Nadler 1978). These observations suggest that heme synthesis is negatively regulated by oxygen. *A. caulinodans* wild-type does not excrete porphyrins under low oxygen; however it would be interesting to compare ALAS and ALAD enzyme activities in an *A. caulinodans fixLJ*, *fixK* mutant vs wt under low oxygen concentrations and high aeration.

The ALA synthase structural genes (*hemA*) of *B. japonicum* (McClung *et al.* 1987) and *R.meliloti* (Leong *et al.* 1982, 1985) have a sequence homologous to the FixK/Fnr binding site consensus (Spiro and Guest 1990) in their 5' upstream region. However *R. meliloti fixK* did not appear to regulate a *R. meliloti hemA-lacZ* fusion (Batut *et al.* 1989). Somerville and Chelm (1988), reported a 10-fold increase in the level of *B. japonicum hemA* mRNA in response to oxygen limitations; however Kim

et al. (1991), reported that a *B. japonicum hemA-lacZ* fusion is not regulated by oxygen. This suggest the existence of post transcriptional regulation. Post transcriptional control specifically by heme has been suggested for chimaeric soybean *leghemoglobin-CAT* (cloramphenicol acetyl transferase) and *leghemoglobin-NPTII* (neomycin phosphotransferase) genes in yeast (Jensen *et al.* 1986).

6.5.2 Is bacterial heme involved in plant leghemoglobin production?

The coordination of rhizobial heme and plant hemoglobin synthesis is not understood, but it has been speculated that bacterial heme induces leghemoglobin apoprotein synthesis by the plant (Appleby 1984; Jensen *et al.* 1986; O'Brian *et al.* 1987). Heme has a variety regulatory functions in eukaryotic cells. Heme regulates protein synthesis via the heme-regulated inhibitor (HRI), which is a protein kinase which specifically phosphorylates the α subunit (eIF-2α) of the eukaryotic initiation factor 2 (Chen *et al.* 1989). Heme controls the DNA binding capabilities and activity of some transcription factors in yeast, like CYP1 (also called HAP1) and the HAP2/3/4 complex (Pfeifer *et al.* 1989; Fytlovich *et al.* 1993; Zhang *et al.* 1993). And recently it was reported that heme also regulates mitochondrial protein transport (Lathrop and Timko 1993). The presence of a conserved amino acid motif (HRM) involved in heme binding, found in several heme regulated proteins, has been proposed to mediates heme control of synthetic and regulatory processes, via a general mechanism (Lathrop and Timko 1993).

One way to connect these observations is to hypothesize that bacterial heme or heme intermediates are involved in the regulation of plant leghemoglobin synthesis, and the finding that bacterial heme production is regulated by *fixK* via *fixLJ* suggest the involvement of *fixLJK* in the coordinate development of a mature functional nodule.

By inserting an O_2 micro-electrode into nodules, Tjepkema and Yocum (1974) found a very sharp decline in O_2 concentration in the inner cortex. This decline was correlated with the morphology of the nodule parenchyma (inner cortex) and the synthesis of Enod2, which is a (hydroxy)proline-rich (glyco)protein, expressed exclusively in the nodule parenchyma cell layers (van de Wiel *et al.* 1990). It has been postulated that this tissue constitute an oxygen barrier to limit oxygen diffusion into the central (infected) nodule tissue (Witty *et al.* 1986).

The increased oxygen demand by the proliferating prebacteroid rhizobia will cause microaerobic conditions. This low oxygen environment of the nodule triggers *nifA* expression via the *fixLJ*, *fixK* system. In addition *fixLJ* and *fixK* activates the expression of a specific (symbiotic) respiratory chain, containing a terminal oxidase with a high affinity for O_2 and regulates/increase the biosynthesis of heme. Concomitant heme (or heme precursors) may activate the production of leghemoglobin in the plant, while NifA activates the *nif/fix* genes in the bacteroids necessary for the synthesis of the nitrogenase enzyme. It has been shown that leghemoglobin is generally first detectable just before nitrogenase activity can be measured. After functional leghemoglobin has been assembled; it allows the vigorous flux of oxygen necessary for bacteroid oxidative phosphorylation and nitrogen fixation.

CHAPTER 7

CLONING AND CHARACTERIZATION OF AN *A. CAULINODANS LEUA*-LIKE GENE

7.1 ABSTRACT

Using a DNA fragment containing *Escherichia coli fnr* gene sequences as heterologous hybridization probe, an *Azorhizobium caulinodans* cosmid carrying a region of DNA homology was isolated. An Open Reading Frame (ORF), encoding a polypeptide sharing a high degree of homology with bacterial α-isopropyl malate synthase enzymes (*leuA* gene products), was identified on the cosmid. The cloned *A. caulinodans* region was found to be sufficient to complement a *leuA* mutant of *E. coli*, suggesting that it carries a functional *leuA* gene. A transposon Tn*5* insertion mutation in the 5' upstream region of the cloned *A. caulinodans leuA* gene was found to abolish complementation of the *E. coli leuA* mutant strain, but the corresponding Tn*5* insertion in the *A. caulinodans* chromosomal *leuA* locus did not result in a detectable phenotype.

7.2 INTRODUCTION

A DNA sequence motif has previously been identified in the *A. caulinodans nifA* 5' upstream region, that is highly homologous to the Fnr binding site consensus sequence (TTGAT-n4-ATCAA; see chapter 5 and figure 1: de Bruijn *et al.* 1988; Ratet *et al.* 1989).

Fnr is a transcriptional regulator mediating the activation or repression of a variety of *E. coli* genes in response to low oxygen tension. The Fnr protein resembles CRP (the cyclic-AMP receptor protein; de Crombrugghe *et al.* 1984), except for the presence of a cysteine-rich N-terminal segment which may form part of an iron-binding redox-sensing domain (Shaw and Guest 1982; Spiro and Guest 1990). The Fnr family contains several other structurally related transcriptional regulators which control a variety of physiological functions (see Irvine and Guest 1993). The amino acid homology between members of the Fnr family is particularly strong in the helix-turn-helix motif which is the DNA-binding domain. The target promoters for activation (and repression) by Fnr are characterized by the presence of a specific inverted repeat sequence, the Fnr-binding site TTGAT-n4-ATCAA.

Using site-directed mutagenesis, and chimeric *nifA-lacZ* reporter gene fusions integrated into the *A. caulinodans* chromosome, it was found that the T in the conserved TTGAT and the A in the ATCAA motif of the *A. caulinodans nifA* 5' upstream region are absolutely required for *nifA* activation (J. Stigter and F.J. de Bruijn, unpublished observation). These results suggested that the *A. caulinodans nifA* promoter could harbor a functional Fnr-binding site, involved in microaerobic induction.

In order to search for the corresponding *A. caulinodans fnr* gene, we attempted to use a heterologous DNA hybridization probe, carrying *E. coli fnr* gene sequences (pGS24; Shaw and Guest 1982). A cosmid of an *A. caulinodans* library constructed in our laboratory (Pawlowski *et al.* 1987) was identified which showed limited DNA homology to the *E. coli* pGS24 probe. The region of weak DNA homology was partially delimited, subcloned and subjected to DNA sequence analysis. Surprisingly, an Open Reading Frame (ORF) was found, with a very low degree of homology with *fnr* genes, but sharing a high degree of homology with bacterial α-isopropyl malate synthase (*leuA)* and homocitrate synthase (*nifV)* genes.

This *A. caulinodans leuA*-like region was found to be able to complement a *leuA* mutant strain of *E. coli*, suggesting that the cloned region carries a functional *leuA* gene.

An Fnr-like activator FixK, in *Rhizobium meliloti* (Batut *et al.* 1989), involved in symbiotic nitrogen-fixation has recently also been identified in *A. caulinodans* (Kaminski *et al.* 1991), *B. japonicum* (Anthamatten *et al.* 1992), and *R. leguminosarum* (FnrN; Colonna-Romano *et al.* 1990).

7.3 MATERIALS AND METHODS

7.3.1 Bacterial strains, plasmids, and growth media

Bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown at 37°C in LB medium (Miller 1972) or M9 medium (Sambrook *et al.* 1989) supplemented with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1µg/ml thiamine and 0.4% glucose. ORS571 strains were grown at 37°C in TY (Beringer 1974), YLS (Elmerich *et al.* 1982) or in LSO medium (Elmerich *et al.* 1982), supplemented with 40 mg/l of nicotinic acid and 0.1% of the desired nitrogen sources. Antibiotics were used at the following concentrations: For *E. coli:* ampicillin (Ap; 100 µg/ml), chloramphenicol (Cm; 30 µg/ml), gentamycin (Gm; 5 µg/ml), kanamycin (Km; 20 µg/ml) and tetracycline (Tc; 10 µg/ml); for ORS571: carbenicillin (Cb; 500 µg/ml), Gm (50 µg/ml), Km (200 µg/ml) and Tc (10 µg/ml).

7.3.2 DNA manipulations

See chapter 3.

7.3.3 Southern blotting and colony hybridizations

Plasmid pGS24 (Shaw and Guest 1982) containing the *E. coli fnr* gene was used as the source for the DNA probe. A 1.65 kb *Bam*HI-*Hind*III fragment was isolated and radioactively labeled by nick translation, using ^{32}P-dATP and non-radioactively via random primed labeling using Dig-dUTP (Boehringer, Mannheim). Southern blotting and nick translations were carried out as described (Sambrook *et*

al. 1989). Hybridizations were carried out overnight at 68°C. The blots were washed at room temperature in 2x SSC (1x SSC contains 0.15M NaCl, 15mM Na citrate, pH 7) buffer, containing 0.1% sodium dodecylsulfate (SDS) and at 68°C with 0.1x SSC buffer containing 0.1% SDS. Non-radioactive DIG-labeled DNA was detected with Lumi-Phos 530 (Boehringer, Mannheim).

Table 1. Bacterial strains and plasmid used in this study

Strain	Relevant characteristics	Source or reference
A. caulinodans		
ORS571	wild type, Cb^r; Nif^+, Nod^+, Fix^+	Dreyfus *et al.* 1988
ORS571*leuA*::Tn*5*	*LeuA*::Tn*5*, Cb^r, Km^r, Nif^+, Nod^+, Fix^+	this work
E. coli		
DH5αF'	F', Δ(*lacZYA*)	Hanahan 1983
HB101	Sm^r, recA	Boyer and Roulland-Dussoix 1969
CV512	*leuA*	Somers *et al.* 1973
Plasmids		
pUC118/pUC119	Ap^r; used for cloning and sequencing	Vieira and Messing 1987
pRK2013	Km^r, IncN, Mob^+, Tra^+, helper plasmid in mobilizations	Figurski and Helinski 1979
pRK290	Tc^r, IncP, Mob^+, Tra^-, broad host range cloning vector	Ditta *et al.* 1980
pWB5	Tc^r, Km^r, pRK290 derivative	W. Buikema and F.M. Ausubel
pPH1JI	Cm^r, Sm^r, Sp^r, Gm^r, IncP, Tra^+	Hirsch and Beringer 1984
pLAFR1	Tc^r, IncP, Mob^+, Tra^-, broad host range cosmid cloning vector	Friedman *et al.* 1982
pJRD184	Tc^r, Ap^r	Heusterspreute *et al.* 1985
pGS24	Ap^r, *E. coli fnr* in pBR322	Shaw and Guest 1982
pACleuA	Tc^r, Ap^r, ORS571 *leuA* in pJRD184	this work
pACleuA::Tn*5*	Tc^r, Ap^r, Km^r, *leuA*::Tn*5*	this work

7.3.4 Transposon Tn*5* mutagenesis of the *leuA*-like gene

Transposon Tn*5* mutagenesis was carried out as described by de Bruijn and Lupski (1984) and de Bruijn (1987).

7.3.5 Conjugation and gene replacement experiments

Plasmids were mobilized from *E. coli* to *A. caulinodans* using the helper plasmid pRK2013, as described by Ditta *et al.* (1980). Gene replacement experiments with Tn*5* mutagenized regions were carried out as described (Ruvkun and Ausubel 1981; de Bruijn 1987).

7.3.6 DNA sequencing

The dideoxynucleotide chain termination method, using ^{35}S-dATP and Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio), was used for DNA sequence determination. The DNA sequence was obtained from restriction fragments cloned into pUC119 and pUC118 and synthetic oligonucleotide primers.

7.4 RESULTS

7.4.1 Cloning of the *A. caulinodans leuA*-like locus

To search for the *A. caulinodans fnr* gene, a DNA fragment containing the *E. coli fnr* gene (1.65 kb *Bam*HI-*Hind*III fragment of pGS24) was initially used as heterologous DNA hybridization probe. When chromosomal DNA of *A. caulinodans* was hybridized with the *E. coli fnr* probe, one strong hybridizing *Eco*RI fragment (13 kb), two weaker hybridizing *Eco*RI fragments (8 kb and 3 kb) and one very weak hybridizing *Eco*RI fragment was identified. However only one strongly hybridizing *Cla*I/*Kpn*I fragment (6 kb) was observed (see fig. 4).

Pooled cosmid miniprep DNA preparations of a cosmid bank of *A. caulinodans* (Pawlowski *et al.* 1987) were screened using this probe and two weakly hybridizing cosmids were isolated. These cosmids were initially called pLfnr1 and pLfnr2. A 4.7 kb *Eco*RI fragment with a low level of homology to the *E. coli fnr* probe was found to be present in both cosmids (data not shown) and subcloned into the vector pJRD184 (Table 1) to form a plasmid initially labeled as pACfnr and later pACleuA (see below).

7.4.2 DNA sequence analysis of the *A. caulinodans leuA*-like locus

The left hand portion of the *Eco*RI fragment cloned in pAC*fnr* (pACleuA) was subjected to DNA sequence analysis (see figures 1 and 2). The DNA sequence analysis revealed the presence of one open reading frame (ORF), from position 539 to 2117 (figs. 1 and 2). The deduced polypeptide from this ORF was found to share significant homology with LeuA proteins from *E. coli*, *Salmonella typhimurium*, *Lactococcus lactis* subsp. *lactis* and *Saccharomyces cerevisiae* (see fig. 3).

Interestingly significant amino acid homology was also found to the nitrogen fixation *nifV* gene product (fig. 3). However, no homology with Fnr products was detected (data not shown). The ATG at position 539 was designated as the putative start codon of the ORF, because of the presence of a Shine-Dalgarno consensus sequence (position 525; GAAGGA; Ringquist *et al.* 1992), at the expected distance and because of the results from the sequence comparison studies (see fig. 3).

The *A. caulinodans* ORF encodes a polypeptide of 526 amino acids (predicted molecular weight 56,829 Kdaltons). Downstream of the ORF at position 2144-2169, a putative factor-independent transcription terminator site which include a GC-rich stem structure, followed by a stretch of T residues (Brendel *et al.* 1986) could be identified (see figure 2).

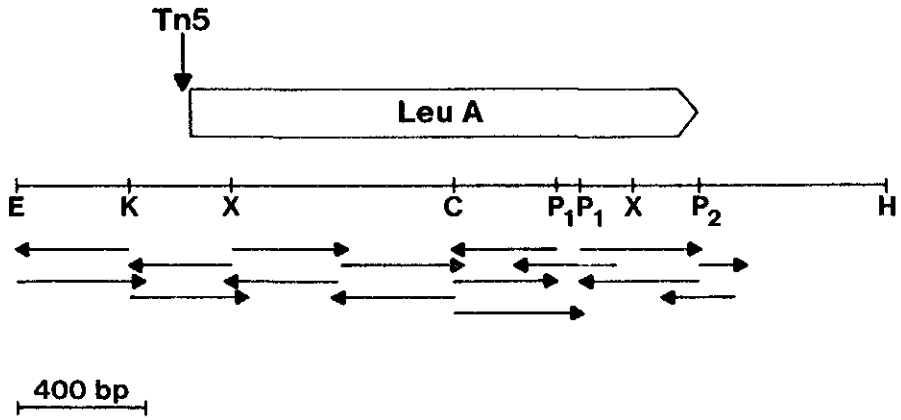

Figure 1. Physical and genetic map of the *A. caulinodans leuA*-like locus and the DNA sequencing strategy employed.

The extent and direction of the protein coding region is shown by open arrow. The position of the Tn*5* insertion is shown by vertical arrow. The restriction enzyme cleavage sites indicated are: E: *Eco*RI; K: *Kpn*I; X: *Xho*I; C: *Cla*I; P1: *Pvu*I; P2: *Pvu*II, H: *Hin*dIII. The horizontal arrows show the extent of the fragments sequenced on both strands.

Figure 2. DNA sequence of the *A. caulinodans leuA*-like locus and predicted protein products.

The nucleotide sequence of the *leuA* gene, 5' and 3' region carried by a 2,279 bp fragment of plasmid pACleuA, starting at the *EcoR*I site is shown.

```
1    GAATTCCCGCGCGCCCACTTCGATGACGCTCACGCCGAGATCGTTGCAGAAGTGAGGAAT 60

61   GCCTGTGTCCGCCATTGGCCTTGTCTCTGTCCCATCGGGTTGCGCGCGAGCTTAGAGCAT 120

121  GTTCCGGCGCGCTTGAACACCTCCCTGCCGCACCGCGCCGCCATGTCGCGCGGGGACATC 180

181  GCGCACGAAGCGGCCGGATCGGCGCTGCACGCGCGGCTTTCTCTTCCGAGCCGCAAAGCA 240

          *** *                       *** *
241  GACCTTGCCTCCATCGCACCTTTGCGGGTATGAGCGGACCAGAGCGCAATTTTTTGGTGA 300

       .....      V  S  S  R  P  T  N  I  A  G  T  S  P  R  I  E
301  GCGAGCAGCAGCCCGTGTCGTCCCGTCCGACGAACATCGCTGGTACCTCCCCCCGGATTG 360
                                                    -------->

       T  N  P  G  N  A  L  V  L  R  A  L  L  L  L  S  R  P  *
361  AGACCAATCCGGGCAACGCTCTTGTTCTGCGCGCGCTCCTTCTCCTTAGCCGCCCCTGAG 420
        <---------              ---------->

        ---------->  <---------
421  GGCCGGCCGGGCGCACGCGCCTGGGCACTCAGGGGATGCGGAAAAGCCAGACCGACGCAC 480
             <----------
                                                        Tn5
                                                .....  .        M
481  CCGAGGCCCTCTCTCCTGACAATCGGATCGAGACGGCCCGTCCCGAAGG ACCCGTGACA 539
                              --------> <--------

       T  E  T  T  A  A  T  P  A  V  S  T  E  A  T  D  R  V  I  I
540  TGACCGAGACCACCGCCGCTACGCCCGCCGTCTCCACCGAGGCGACCGACCGCGTCATCA 599

       F  D  T  T  L  R  D  G  E  Q  C  P  G  A  S  M  T  F  E  E
600  TCTTCGACACCACCCTGCGCGACGGCGAACAGTGCCCCGGCGCCTCCATGACCTTCGAGG 659

       K  L  E  V  A  E  L  L  E  T  M  G  V  D  V  I  E  A  G  F
660  AGAAGCTCGAGGTCGCGGAATTGCTGGAAACCATGGGCGTCGATGTCATCGAGGCCGGCT 719

       P  I  A  S  I  G  D  F  E  S  V  A  E  I  A  R  R  S  K  T
720  TCCCCATCGCCTCCATCGGCGATTTCGAGTCGGTGGCCGAGATTGCCCGCCGCTCCAAGA 779

       A  I  I  A  G  L  S  R  A  A  L  N  D  I  D  R  C  A  E  A
780  CGGCGATCATCGCCGGCCTCTCCCGCGCCGCCCTCAACGACATCGACCGCTGCGCGGAAG 839

       V  K  Q  A  K  R  G  R  I  H  T  F  L  S  T  S  P  V  H  M
840  CGGTGAAGCAGGCCAAGCGCGGCCGCATCCACACCTTCCTGTCCACCTCGCCGGTGCACA 899

       K  Y  K  L  Q  K  E  P  H  Q  V  L  E  M  I  V  A  S  V  T
900  TGAAGTACAAGCTCCAGAAGGAGCCCCACCAGGTGCTGGAGATGATCGTCGCCTCGGTGA 959

       R  A  R  N  H  V  E  D  V  E  W  S  A  E  D  G  T  R  T  E
960  CGCGCGCCCGGAACCATGTGGAGGATGTGGAGTGGTCGGCGGAGGACGGTACGCGCACCG 1019

       M  D  F  L  C  R  C  V  E  A  A  I  K  A  G  A  T  T  I  N
1020 AGATGGACTTCCTCTGCCGCTGCGTGGAGGCCGCCATCAAGGCCGGCGCAACCACAATCA 1079

       I  P  D  T  V  G  Y  T  T  P  Q  E  Y  E  A  L  F  R  T  V
1080 ACATCCCCGACACCGTGGGCTACACCACGCCGCAGGAATACGAGGCGCTCTTCCGGACGG 1139

       R  E  R  V  P  N  S  D  K  A  I  F  S  V  H  C  H  N  D  L
1140 TCCGCGAGCGCGTACCAAATTCCGACAAGGCCATCTTCTCGGTCCATTGCCATAATGACC 1199
```

```
      G  M  A  V  A  N  S  L  A  G  L  A  G  G  A  R  Q  I  E  C
1200 TCGGCATGGCGGTCGCCAACTCGCTGGCCGGCCTCGCCGGCGGCGCGCGGCAGATCGAAT 1259

      T  I  N  G  I  G  E  R  A  G  N  A  A  L  E  E  V  V  M  A
1260 GCACCATCAACGGCATCGGCGAGCGGGCCGGCAATGCGGCGCTGGAAGAGGTGGTGATGG 1319

      I  N  T  R  R  D  V  L  P  Y  R  T  G  I  D  A  T  L  L  T
1320 CCATCAACACCCGCCGCGACGTGCTGCCCTATCGCACCGGCATCGATGCCACCCTGCTCA 1379

      R  A  S  K  M  V  S  G  V  T  S  F  P  V  Q  Y  N  K  A  I
1380 CCCGCGCCTCGAAGATGGTCTCGGGCGTGACGTCCTTCCCGGTGCAGTACAACAAGGCCA 1439

      V  G  R  N  A  F  A  H  E  S  G  I  H  Q  D  G  M  L  K  H
1440 TCGTCGGCCGGAATGCCTTCGCCCATGAGAGCGGCATCCATCAGGACGGCATGCTCAAGC 1499

      T  Q  T  Y  E  I  M  T  P  E  S  V  G  V  T  K  T  S  L  V
1500 ACACCCAGACCTACGAGATCATGACCCCCGAGAGCGTGGGCGTGACCAAGACCTCGCTGG 1559

      M  G  K  H  S  G  R  A  A  F  R  D  K  L  K  A  L  G  Y  E
1560 TGATGGGCAAGCACTCCGGCCGCGCTGCCTTCCGCGACAAGCTGAAGGCGTTGGGCTACG 1619

      L  G  E  N  A  L  N  D  A  F  T  R  F  K  D  L  A  D  R  K
1620 AGCTGGGTGAGAATGCGCTGAACGACGCCTTCACCCGCTTCAAGGACCTCGCCGATCGCA 1679

      K  V  I  Y  D  E  D  I  E  A  L  V  D  Q  G  I  A  A  A  Y
1680 AGAAGGTCATTTACGACGAGGACATTGAGGCGCTGGTGGATCAGGGCATCGCCGCTGCCT 1739

      D  R  V  K  L  V  S  L  S  V  I  A  G  T  R  G  P  Q  R  A
1740 ACGATCGCGTGAAGCTCGTGTCGCTCTCGGTGATCGCCGGCACCCGCGGCCCGCAGCGCG 1799

      T  M  R  I  E  V  D  G  Q  P  R  I  E  E  A  E  G  N  G  P
1800 CCACGATGCGCATCGAAGTGGATGGCCAGCCCCGCATTGAGGAAGCCGAGGGCAACGGCC 1859

      V  D  A  T  F  N  A  I  K  A  L  I  P  H  T  A  K  L  E  L
1860 CGGTGGATGCCACCTTCAACGCCATCAAGGCGCTGATCCCGCACACCGCCAAGCTCGAGC 1919

      Y  Q  V  H  A  V  T  E  G  T  D  A  Q  A  E  V  S  V  R  L
1920 TCTATCAGGTGCATGCCGTGACCGAGGGCACGGACGCGCAGGCCGAGGTCTCGGTGCGTC 1979

      A  E  D  G  K  V  V  T  A  R  A  A  D  P  D  T  L  V  A  S
1980 TCGCAGAGGACGGCAAGGTGGTGACCGCCCGCGCGGCCGACCCCGACACGCTGGTCGCGT 2039

      A  Q  A  Y  I  T  A  L  N  K  L  S  V  K  R  Q  S  V  N  A
2040 CCGCGCAGGCCTACATCACGGCGCTCAACAAGCTCTCCGTGAAGCGCCAGAGCGTGAACG 2099

      Q  A  A  A  S  *
2100 CGCAGGCCGCCGCCAGCTGAGCGGTGCCCGACACGACACATCCGAAGCCCTCGCCCTCCG 2159
                                                  ---------->    <

                          .....      M  H  R  T  G  V  L  L  A  L  L
2160 GCGGGGGCTTTTTTGTCTGAGGACGGTCCATGCACCGCACCGGTGTTCTGCTCGCGCTCC 2219
     ----------

      S  A  V  L  F  G  A  S  T  P  L  A  K  L  L  L  G  A  V  D
2220 TCTCCGCTGTGCTGTTCGGCGCCAGCACGCCGCTCGCCAAGCTGCTGCTGGGCGCGGTGG 2279
```

The potential -35 and -10 region is indicated by underlined asterisks. The deduced aminoacid sequence of the leader peptide, *leuA* and part of ORF2 is shown above the sequence. The putative leader-peptide ribosome binding site (GAGCA, position 303), *leuA* ribosome binding site (GAAGGA, position 525) and ORF2 ribosome binding site (GAGGA, position 2178) are indicated by dots. The stop codon is indicated with an asterisk. The inverted repeats are indicated by horizontal arrows and the location of Tn*5* insertion is shown by a vertical arrow.

Figure 3. Aminoacid sequence comparison of the deduced protein products of *leuA* and *nifV*.

```
Ac LeuA   1 MTETTAATPAVSTEATDRVIIFDTTLRDGEQCPGASMTFEEKLEVAELLETMGVDVIEAGFPIA 64
St LeuA                 1 MSQQVIIFDTTLRDGEQALQASLSAKEKLQIALALERMGVDVMEVGFPVS 50
Ec LeuA                 1 MSQQVIIFDTTLRDGEQALQASLSVKEKLQIALALERMGVDVMEVGFPVS 50
Ll LeuA                 1  MRKIEFFDTSLRDGEQTPGVSFSISEKVTIAKQLEKWRISVIEAGFSAA 49
Sc Leu4                57 ITRAPRWLSTDLRDGNQSLPDPMSVEQKKEYFHKLVNIGFKEIEVSFPSA 106
CON                               VIIFDTTLRDGEQ.L.AS.S..EKL.IA..LE.MGVDVIEVGFP.A
RspV      1 MSRQQPRASFLPESPLAPVALCDTTLRDGEQTAGVAFTRAEKRAIAEALQAAGVAEVEVGVPAM 64
AvV                     1 MASVIIDDTTLRDGEQSAGVAFNADEKIAIARALAELGVPELEIGIPSM 49
CpVa                    1  MGINIVDTTLRDGEQKAGIALSVQDKVEIAKIISEMGVHQIEAGIPAM 48
KpV                     1 MERVLINDTTLRDGEQSPGVAFRTSEKVAIAEALYAAGITAMEVGTPAM 49
AcV                     1 MASVIIDDTTLRDGEQSAGVAFNADEKIAIRRALAELGVPELEIGIPSM 49
CON                       MA.V.I.DTTLRDGEQSAGVAF...EK.AIA.AL.E.GV.E.E.GIPAM

Ac LeuA                65 SIGDFESVAEIARRSKTAI.IAGLSRAALNDIDRCAEAVKQAKRGRIHTFL 114
St LeuA                51 SPGDFESVQTIARTIKNSR.VCALARCVEKDIDVAAQALKVADAFRIHTFI 100
Ec LeuA                51 SPGDFESVQTIARQVKNSR.VCALARCVEKDIDVAAESLKVAEAFRIHTFI 100
Ll LeuA                50 SPDSFEAVKQIADSLNDTA.VTALARCVISDIDKAVEAVKGAKYPQIHVFI 99
SC Leu4               107 SQTDFDFTRYAVENAPDDVSIQCLVQSREHLIKRTVEALTGAKKATIHTYL 157
CON                       SPGDFESV..IAR..K....V.ALARCVE.DID.AAEALK.AK..RIHTFI
RspV                   65 GEEERADIRAVA.AVLKTAAPVVWCRLRAEDLAAAQRTGVV..RLHIGV.. 110
AvV                    50 GEEEREVMHAIA.GLGLSSRLLAWCRLCDVDLAAARSTGVT..MVDLSL.. 95
CpVa                   49 GGDEKISVSKIA.ALGLPSKIAAWNRMSTKDIDTSIECGVD..IVHISS.. 94
KpV                    50 GDEEIARIQLVR.RQLPDATLMTWCRMNALEIRQSADLGID..WVDISI.. 95
AcV                    50 GEEEREVMRAIA.GLGLSSRLLAWCRLCDFDLSAARSTGVT..MVDLSL.. 95
CON                       GEEER.....IA..LGL.S.L.AWCRL...DL.AA..TGV....VDIS...

Ac LeuA               115 STSPVHMKYKLQKEPHQVLEMIVASVTRARNHVEDVEWSAEDGTRTEMDF 164
St LeuA               101 ATSPMHIATKLRRTLDEVIERAVYMVKRARNYTDDVEFSCEDAGRTPVDD 150
Ec LeuA               101 ATSPMHIATKLRSTLDEVIERAIYMVKRARNYTDDVEFSCEDAGRTPIAD 150
Ll LeuA               100 ATSPIHMKYKLKISPEEVLKNIDKCVRYARERVEVVEFSPEDATRTELNF 149
SC Leu4               158 ATSDMFREIVFNMSREEAISKAVEATKLVRKLTKDD..PSQQATRWSY 203
SC Leu4                                                  204 EFSPECFSDTPGEF 217
CON                       ATSPMH...KL.....EVIE.AV..VKRARN.T.DVEFS.EDATRTP..F
RspV                  111 PVSERQISAKLGKDAAWVRDKVEKLVRAASWAGHKVSVGAEDASRADPFF 160
AvV                    96 PVSDLMLHHKLNRDRDWALREVARLVGEARMAGLEVCLGCEDASRADLEF 145
CpVa                   95 PVSDLQIKTKLEKDRKWVAENLKRTVIYALEKDCEVTVGLEDSSRADLNF 144
KpV                    96 PASDKLRQYKLREPLAVLLERLAMFIHLAHTLGLKVCIGCEDASRASGQT 145
AcV                    96 PISDLMLRHKLNRDRDWALGEVARLVSEARMAGLEVCLGCEDASRADQDF 145
CON                       PVSDL....KL..DR.W.L..VARLV..A..AGLEVCLGCEDASRAD..F

Ac LeuA               165 LCRCVEAAIKAGATTINIPDTVGYTTPQEYEALFRTVRERVPNSDKAIFS 214
St LeuA               151 LARVVEAAINAGARTINIPDTVGYTMPFEFAGIISGLYERVPNIDKAIIS 200
Ec LeuA               151 LARVVEAAINAGATTINIPDTVGYTMPFEFAGIISGLYERVPSIGKAIIS 200
Ll LeuA               150 LLEAVQTAVDAGATYINIPDTVGYTTPEEYGKIFKFLIDNTKSDREIIFS 199
SC Leu4                                   242 PATVEVASPNVVYADQIEYFATHITEREKVCIS 273
SC Leu4               218 AVEICEAVKKAWEPTEENPIIFNL 241
CON                       L.R.VEAAI.AGATTINIPDTVGYT.P.EYA.II..L.ERVP...KAIIS

RspV                  161 LAEIAHVAAEAGAIRFRISDTLGVLDPFAAHELVGRVVTRCPLPVE     206
AvV                   146 VVQVGEVAQAAGARRLRFADTVGVMEPFGMLDRFRFLSRRLDMELE     191
CpVa                  145 LIQLCEMIFALGVKRVRYADTVGIMEPKELYSQIKKIRDKVPIDIE     190
KpV                   146 LRAIAEVAQNAPAARLRYADTVGLLDPFTTAAQISALRDVWSGEIE     191
AcV                   146 IVRVGAVAQAARPPP.AFADTVGVMEPFGMLDRFRFLRQRLDVELE     190
CpVw                  126 YLSEDNKIFCIFNIVCLRINNLSKLDFIDEDFRIKDLKSKFNVLVD     171
CON                       L....EVAQAAGA.R.R.ADTVGVMEPF....RI..LR.R...E.E
```

```
Ac LeuA      215 VHCHNDLGMAVANSLAGLAGGARQIECTINGIGERAGNAALEEVVMAINT  264
St LeuA      201 VHTHDDLGIAVGNSLAAVHAGARQVEGAMNGIGERAGNCALEEVIMAIKV  250
Ec LeuA      201 VHTHDDLGLAVGNSLAAVHAGARQVEGAMNGIGERAGNCSLEEVIMAIKV  250
Ll LeuA      200 PHCHDDLGMAVANSLAAIKAGAGRVEGTVNGIGERAGNAALEEIAVALHI  249
SC Leu4      274 THCHNDRGCGVAATELGMLAGADRVEGCLFGNGERTGNVDL..VTVAMNM  321
CON              VHCHDDLG.AVANSLAA..AGARQVEG..NGIGERAGN.ALEEV.MAI..

RspV         207 FHGHNDLGMATANSLAAARAGASHLSVTVNGLGERAGNAALEEVAAALEA  256
AvV          192 VHAHDDFGLATANTLAAVMGGATHINTTVNGLGERAGNAALEECVLALKN  241
CpVa         191 IHVHNDFGMAISNSFAAFKAGAKFADCTITGMGERAGNCDFLKFVKVIQE  240
KpV          192 MHAHNDLGMATANTLAAVSAGATSVNTTVLGLGERAGNAALETVALGLER  241
AcV          191 VHAHDDFGLATANTLAAVMGGATHINTTVNGLGERAANAALEECVLALKN  240
CpVw         172 FCASNKYNMATAIIINAFLNGSDIITTEFNS....NDYAAMEEVIIALKS  217
CON              .HAHNDFGMATANTLAAV.AGATHINTTVNGLGERAGNAALEEVVLALK.

Ac LeuA      265 RRDVLPYRTGIDATLLTRASKMVSGVTS.FPVQYNKAIVGRNAFAHESGIH 314
St LeuA      251 RKDIMNVHTNINHHEIWRTSQTVSQICN.MPIPANKAIVGSGAFAHSSGIH 300
Ec LeuA      251 RKDILNVHTAINHQEIWRTSQLVSQICN.MPIPANKAIVGSGAFAHSSGIH 300
Ll LeuA      250 RKDFYQAQSPLKLSETAATAELISQFSG.IAIPKNKAIVGANAFAHESGIH 299
SC Leu4      322 YTQ..GVSPNLDFSDLTSVLDVV.ERCNKIPVSQRAPYGGDLVVCAFSGSH 369
CON              RKD...V.T.I...E..RTS..VSQ.CN..PIP.NKAIVG..AFAH.SGIH

RspV         257 ....AGRATGVALGQLCALSELVARASGR.PLSPQKPIVGEGVFTHECGIH 302
AvV          242 ...LHGIDTGIDTRGIPAISALVERASGR.QVAWQKSVVGAGVFTHEAGIH 288
KpV          242 ...CLGVETGVHFSALPALCQRVAEAAQR.AIDPQQPLVGELVFTHESGVH 288
AcV          241 ...LHGIDTGIDTRGIPAISALVERASGR.QWPGRRAWL.APVFTHEAGIH 286
CpVw         218 ...IRNIEIRGDLKLISKLTRIYEKITSE.RVYSMKPILGEDIFKYESGIH 264
CpVa         241 ...LTGEKIYTGDFEDIIEKENEIKKILR.LNW*                   269
CON              ...L.GI.TG.D...IPALS.LVERASGR.....KP.VGE.VFTHE.GIH

Ac LeuA      315 QDGMLKHTQTYE..IMTPESVGVTKTSLVM..GKHSGRAAFRDKLKALGYE 361
St LeuA      301 QDGVLKNRENYE..IMTPESIGLNQIQLNL..TSRSGRAAVKHRMEEMGYK 347
Ec LeuA      301 QDGVLKNRENYE..IMTPESIGLNQIQLNL..TSRSGRAAVKHRMDEMGYK 347
Ll LeuA      300 QDGVLKNAETYE..IITPELVGIKHNSLPL..GKLSGRHAFSEKLTELNIA 346
SC Leu4      270 QDAIKKGFNLQNKKRAQGETQWRIPYL 296
SC Leu4                  297 PLDPKDIGRDYEAVIRVNSQSGKGGAAWVILRSLGLD 333
CON              QDGVLKN.E.YE..IMTPESIG.....L.L.....SGRAAV..R..ELGY.

RspV         303 VDGLMKDRATYESADLRPERFGRSH.RIAI..GKHSSAAGLARALAEAGLP 350
AvV          289 VDGLLKHRRNYE..GLNPDELGRSH.SLVL..GKHSGAHMVRNTYRDLGIE 334
KpV          289 VAALLRHSESYQ..SIAPSLMGRSY.RLVL..GKHSGRQAVNGVFDQMGYH 334
AcV          287 VDGLLKHRRNYE..GLNPDELGRSH.SLVL..GKHSGAHMVRNSYRELGIE 332
CpVw         265 ADGIAKNPKNYE..PFNPELIGTNR.KLYI..GKHSGKAALVVKFKELNLN 310
CON              VDGLLKHR.NYE....NP...GRSH..LVL..GKHSGA..V.....ELG..

Ac LeuA      362 LGENALNDAFTRFKDLADRKKVIYDE....DIEALVDQGIAAA.YDRVKLV 407
St LeuA      348 DTDYNMDHLYDAFLKLADKKGQVFDY....DLEAL..AFINKQQEEPEHFR 392
Ec LeuA      348 ESEYNLDNLYDAFLKLADKKGQVFDY....DLEAL..AFIGKQQEEPEHFR 392
Ll LeuA      347 YDDESLAILFEKFKKLADKKKEITDA....DIHALFTGETVKN...LAGFI 390
SC Leu4      334 LPR.NMQIEFSSA..VQDHADSLGRELKSDEISKLFKEAYNYNDEQYQAIS 381
CON              ....NL..LF..F.KLADKK....D.....DIEAL.........E......

RspV         351 ADAATLAALMPALRDWAAITKRAAAPEDLAALLAAQTETAR*            391
AvV          335 LADWQSQALLGRIRAFSTRTKRRSPQPAELQDFYRQLCEQGNPELAAGGMA*385
CpVw         311 CNNIDMNLFLQDIREKSIQEKRNVLDNEIIEMYKEYNKSYQR*          352
KpV          335 LNAAQINQLLPAIRRFAENWKRSPKDYELVAIYDELCGESALRARG*      380
AcV          333 LADWQSQALLGRIRAFSTRTKRSPQAAELEDFYRQLCEQGTAELAAGGMA* 382
CON              ....Q..ALL..IR.FS...KR.....EL...Y
```

```
Ac LeuA          408 SLSVIA..GTRGPQRATMRIEVDGQPRIEEAEGNGPVDATFNAIKALIPHT 456
St LeuA          393 LDYFSVQSGSSDIATASVKLACGEEIKAEAANGNGPVDAIYQAINRITGYD 443
Ec LeuA          393 LDYFSVQSGSNDIATAAVKLACGEEVKAEAANGNGPVDAVYQAINRITEYN 443
Ll LeuA          391 LDNV....QIDGHKALVQLKNQEEEIYVSQGEGSGSVDAIFKAIDKVFNHQ 437
SC Leu4          382 LVNYNVEKFGTERRVFTGQVKVGDQIVDIEGTGNGPISSLVDALSNLLNVR 432
CON                  LD...V..G......A......GEEI..E.A.GNGPVDA...AI.......

Ac LeuA          457 AKLELYQVHAVTEGTDAQAEVSVRLAEDGKVVT......ARAADPDTLVAS 501
St LeuA          444 VELVKYDLNAKGRGKDALGQVDIVVNHHGRRFH......GVGLATDIVESS 488
Ec LeuA          444 VELVKYSLTAKGHGKDALGQVDIVANYNGRRFH......GVGLATDIVESS 488
Ll LeuA          438 LKLISYSVDAVTDGIDAQATTLVSVENLSTGTIFNAK....GVDYDVLKGS 484
SC Leu4          433 FAVANYTEHSLCSGSSTQAASYIHLSYRRNADNEKAYKWGVGVSEDVGDSS 483
CON                  ..L..Y...A.G.G.DAQA.V.I.....G..........GVG...D...SS

Ac LeuA          502 AQAYITALNKLSVKRQSVNAQAAAS*              526
St LeuA          489 AKAMVHVLNNIWRAAEVEKELQRKAQNKENNKETV*    522
Ec LeuA          489 AKAMVHVLNNIWRAAEVEKELQRKAQHNENNKETV*    522
Ll LeuA          485 AIA..YMNANVLVQKENLQGKVEQISAHDGI*       513
SC Leu4          484 VRAIFATINNIIHSGDVSIPSLAEVEGKNAAASGSA*  519
CON                  A.A....LNNI
```

Ac denotes *A. caulinodans leuA* (this work), St, *S. typhimurium leuA* (Ricca and Calvo, 1990, modified, see discussion), Ec, *E. coli leuA* (Yura *et al.* 1992), Ll, *Lactococcus lactis* subsp. *lactis leuA* (Godon *et al.* 1992), SC, Saccharomyces cerevisiae LEU4 (Beltzer *et al.* 1986), RspV, *Rhodobacter sphaeroides nifV* (Meijer and Tabita, 1992), AvV, *A. vinelandii nifV* (Beynon *et al.* 1987), CpVa/CpVw, *Clostridium pasteurianum nifVα/nifVω* (Wang *et al.* 1991), KpV, *K. pneumoniae nifV* (Beynon *et al.* 1987; Arnold *et al.* 1988), AcV, *Azotobacter chroococcum nifV* (Evans *et al.* 1991).

7.4.3 Creation of a transposon Tn*5* insertion mutant in the *A. caulinodans leuA*-like region

Site directed Tn*5* mutagenesis was performed on pACleuA. However all Tn*5* insertions in the cloned DNA were found to be located in the 5' upstream region of the *leuA*-like locus. One Tn*5* insertion was used for gene-replacement experiments (see de Bruijn 1987), and the position of the Tn*5* in the chromosomal *A. caulinodans leuA* locus was verified by Southern blotting (fig. 4). The resulting strain was examined for its ability to fix nitrogen in culture and on plates (Nif phenotype), and in stem and root nodules on *S. rostrata* (Nod and Fix phenotypes). However no significant phenotype differences with the wild-type strain could be identified.
Also growth of the Tn*5* containing strain with and without leucine (data not shown) was found to be identical to wild-type. Therefore, no conclusive data about the functionality of the cloned *leuA*-like locus could be derived from these experiments.

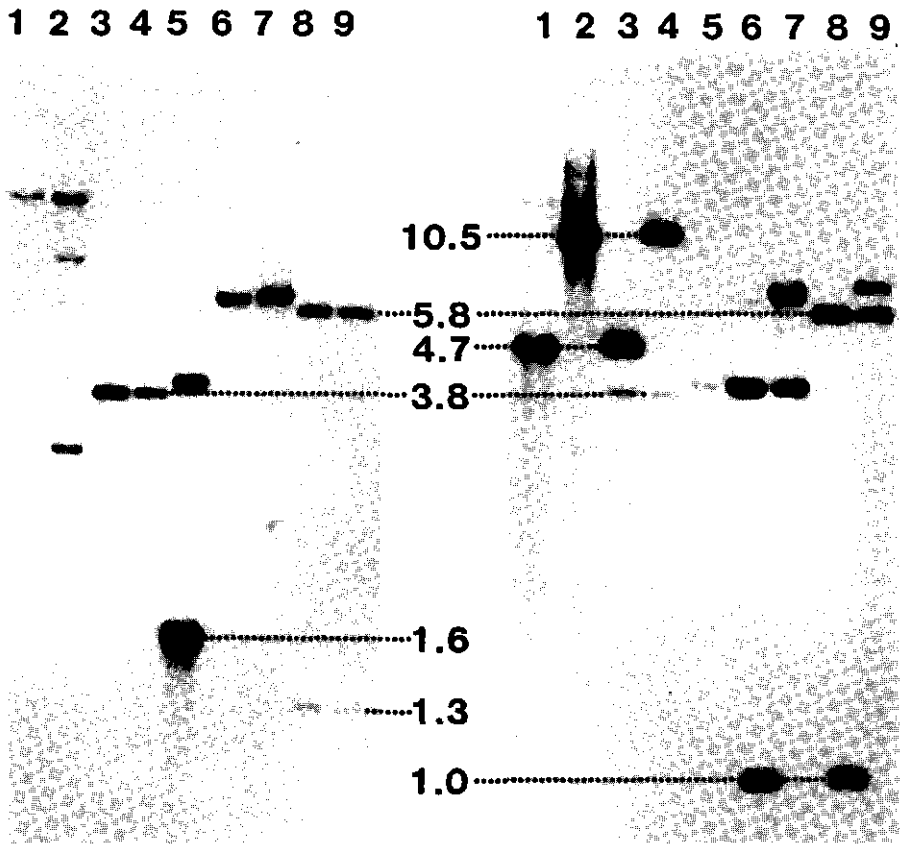

Figure 4. Mapping of the position of Tn*5* insertion within the chromosomal DNA of *A. caulinodans*.

Southern blot hybridized with the 1.64 kb intragenic probe from *E. coli fnr* (*Hind*III-*Bam*HI fragment of pGS24) **[A]** and with the 1.25 kb intragenic probe from *A. caulinodans leuA* (*Xho*I-fragment of pACleuA) **[B]**.

Lane 1: *A. caulinodans* ORS571 wt genomic DNA (*EcoR*I digest)

Lane 2: ORS571 *leuA*::Tn*5* genomic DNA (*EcoR*I digest)

Lane 3: pACleuA plasmid DNA (*EcoR*I digest)

Lane 4: pACleuA::Tn*5* plasmid DNA (*EcoR*I digest)

Lane 5: pGS24 plasmid DNA (*Hind*III-*Bam*HI digest)

Lane 6: *A. caulinodans* ORS571 wt genomic DNA (*Cla*I-*Kpn*I digest)

Lane 7: ORS571 *leuA*::Tn*5* genomic DNA (*Cla*I-*Kpn*I digest)

Lane 8: pACleuA plasmid DNA (*Cla*I-*Kpn*I digest)

Lane 9: pACleuA::Tn*5* plasmid DNA (*Cla*I-*Kpn*I digest)

7.4.4 Complementation of an *E. coli LeuA* mutant with the *A. caulinodans LeuA*-like locus

In order to further investigate the functionality of the *A. caulinodans leuA*-like locus in leucine biosynthesis, an *E. coli* strain CV512, which carries a defect in α-isopropylmalate synthase (*leuA*), was transformed with the plasmid harboring the *A. caulinodans leuA*-like gene, pACleuA and the Tn*5* mutagenized plasmid, pACleuA::Tn*5*.

CV512 was not able to grow on minimal M9 medium plates without leucine, however transformants harboring pACleuA were able to grow on M9 medium. Plasmid pACleuA::Tn*5*, containing a Tn*5* insertion in the *leuA* 5' upstream regulatory region was not able to complement the leucine defect (fig. 5). The ability of plasmid pACleuA to complement the *E. coli leuA* mutant strain CV512 was verified by determining the growth curves of CV512 with and without the wildtype and Tn*5* containing plasmids, in the presence and absence of leucine. This experiment is shown in figure 6. These experiments suggests that the plasmid pACleuA carries a functional *leuA* gene and supports the designation of the cloned locus as the *A. caulinodans leuA* gene.

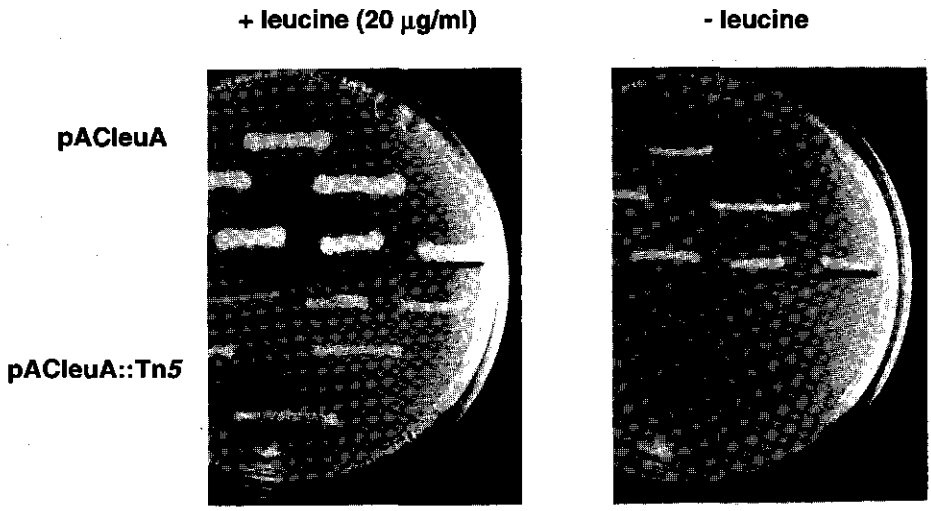

Figure 5. Complementation of *E. coli LeuA* mutant with *A. caulinodans LeuA*-like gene on M9 minimal medium plates.

pACleuA = *A caulinodans leuA* region in pJRD184

pACleuA::Tn*5* = *A. caulinodans leuA*::Tn*5* in pJRD184

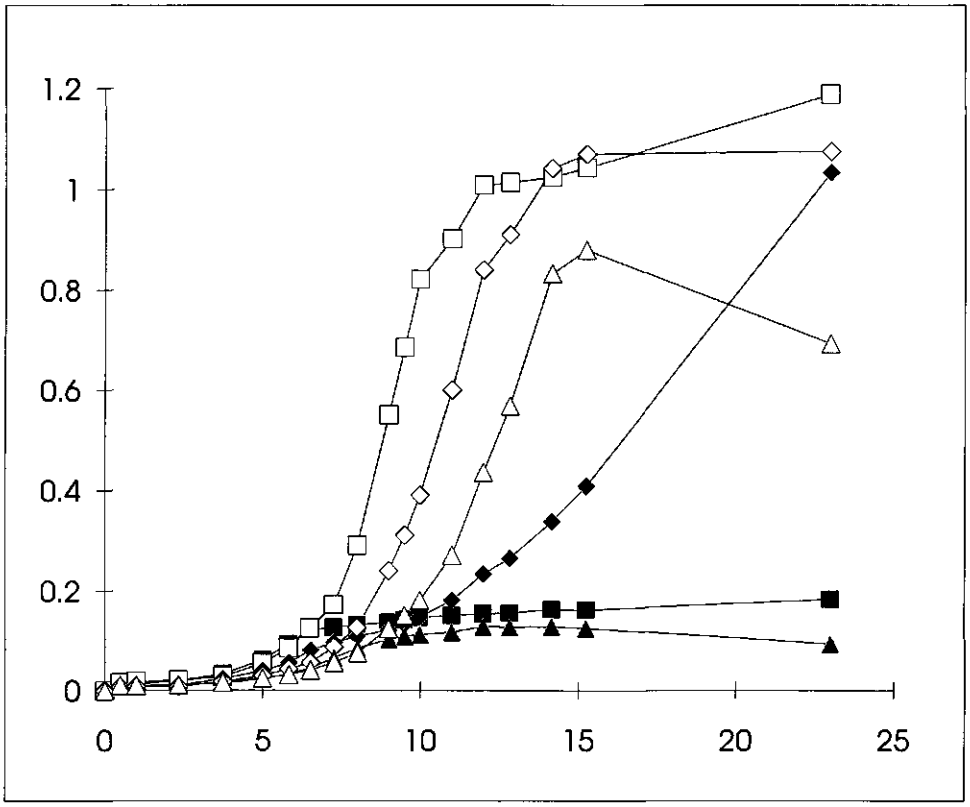

Figure 6. Growth curve of *E. coli* CV512 harboring pACleuA and pACleuA::Tn*5* in liquid M9 minimal medium.

Open symbols denote growth with leucine (20 μg/ml); closed symbols without leucine.

Squares denote absence of plasmids; diamonds the presence of pACleuA; triangles the presence of pACleuA::Tn*5*.

7.5 DISCUSSION

7.5.1 Cloning of the *A. caulinodans leuA*-like locus

To search for the *A. caulinodans fnr* gene, a DNA fragment containing the *E. coli fnr* gene (1.65 kb *Bam*HI-*Hind*III fragment of pGS24) was used as heterologous probe. However a *leuA*-like gene was isolated.

How could a *leuA* gene have been picked up with a *fnr* probe? The homology between *E. coli fnr* and *A. caulinodans leuA* is not significant (29.4%) and no cross hybridization between the cloned genes can be seen on a Southern Blot. This can be seen in figure 5. Were the same blot is probed with the *E. coli fnr* gene (panel A) and the *A. caulinodans leuA* gene (panel B). Lane 5, containing the *fnr*-gene (1.6 kb *Hind*III-*Bam*HI band; see panel A), shows no hybridization to the *A. caulinodans leuA* probe (panel B; the weak 4.0 kb band is due to hybridization of pBR322 to pJRD184 vector DNA contamination in the *leuA*-probe]. In addition, lane 3, containing the 4.7 kb *leuA*-gene, and lane 4 containing the 10.5 kb *leuA*::Tn*5* insert in pJRD184 (see panel B) show no hybridization to the *E. coli fnr* probe (panel A; the 3.8 kb band is due to hybridization of pJRD184 to pBR322 vector DNA contamination in the *fnr*-probe).

However there are some stretches (36-70 bp) of higher homology between *E. coli fnr* and *A. caulinodans leuA*, (57.1-69.4%), and it is possible that a degraded probe might have shown initial cross-hybridization. In any case, our unusual screen clearly identified a leucine biosynthetic (*leuA*) gene, based on the following criteria: The deduced ORF of the cloned *A. caulinodans* region shows a high degree of homology to the *leuA* gene products from *E. coli*, *S. typhimurium*, *Lactococcus lactis* and *Saccharomyces cerevisiae*, and the cloned *A. caulinodans* region is able to complement a *leuA* mutant of *E. coli*.

7.5.2 Leucine biosynthesis in bacteria: The *leuA* gene

The biosynthesis of leucine, isoleucine, and valine (also called the branched-chain amino acids BCAA) has been well studied in *E. coli*, *S. typhimurium* and *B. subtilis* (Umbarger 1987). These three amino acids, which have branched aliphatic R groups, are synthesized via similar pathways. The biosynthetic pathways of isoleucine and valine are catalyzed by the same set of enzymes.

The formation of leucine (fig. 7) starts with condensation of α-ketoisovalerate (which is also the precursor of valine) with acetyl-CoA to yield α-isopropylmalate (α-IPM). This reaction is catalyzed by α-IPM synthase (EC 4.1.3.12), the product of the *leuA* gene.

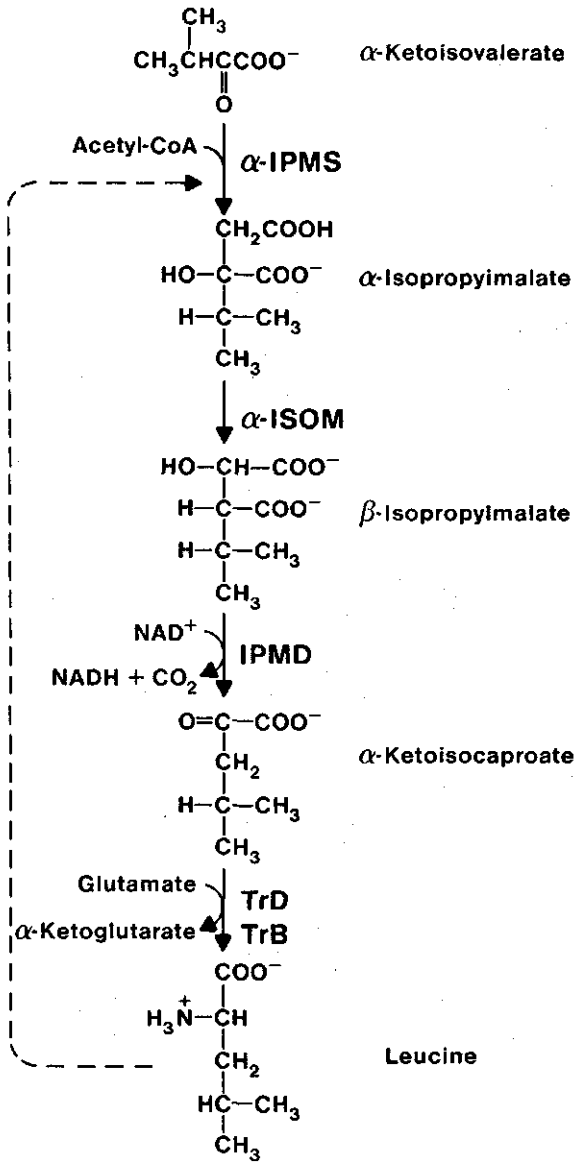

Figure 7. Biosynthesis of leucine.

The enzymes catalyzing the biosynthetic steps are abbreviated, and the corresponding structural genes are indicated, as follows: αIPMS (*leuA*), α-isopropyl malate synthase; αISOM (*leuC* and *leuD*), α-isopropyl malate isomerase; βIPMD (*leuB*), β-isopropyl malate dehydrogenase; TrD (*tyrB*), transaminase D; TrB (*ilvE*), transaminase B. In bacteria the biosynthesis of valine, isoleucine, and leucine is subject to feedback inhibition at the first step by the end products.

Interestingly, as shown in figure 3, LeuA exhibits strong similarity to the *nifV* gene product of nitrogen fixing organisms (*Rhodobacter sphaeroides nifV* (Meijer and Tabita, 1992), *A. vinelandii nifV* (Beynon *et al.* 1987), *Clostridium pasteurianum nifVα/nifVω* (Wang *et al.* 1991), *K. pneumoniae nifV* (Beynon *et al.* 1987; Arnold *et al.* 1988), and *Azotobacter chroococcum nifV* (Evans *et al.* 1991).

The *nifV* gene encodes a homocitrate synthase which carries out the condensation between acetyl-CoA and α-ketoglutarate. Homocitrate is a component of FeMo-cofactor which is incorporated into the FeMo-protein (Hoover *et al.* 1987). The two major blocks of identity between NifV and LeuA, which are placed at approximately equivalent positions in the two enzymes, could be binding sites for acetyl-CoA or for the structural common portion of α-ketoglutarate and α-ketoisovalerate.

In the second step of leucine biosynthesis, α-IPM is converted to β-IPM in a complex two-step reaction by isopropyl malate isomerase (α-ISOM), encoded by the *leuC* and *leuD* genes. The final pathway-specific step is the oxidation of β-IPM to α-ketoisocaproate, catalyzed by the NAD-dependent β-isopropyl malate dehydrogenase (EC 1.1.1.85), the product of the *leuB* gene. Finally the terminal step in leucine biosynthesis, transamination of α-ketoisocaproate into leucine, is catalyzed by a non-specific aminotransferase, either the aromatic transaminase (*TyrB)* or the branched-chain amino acid transaminase (*IlvE)* to convert a-ketoisocaproic acid to leucine (Umbarger 1987).

7.5.3 Organization of leucine biosynthetic *(leu)* genes

In *E. coli* and *S. typhimurium*, the enzymes responsible for leucine biosynthesis are encoded by four contiguous genes. These genes, *leuABCD*, are transcribed as a single unit and regulated by an adjacent control region (Umbarger 1987). The stop codon for *leuA* overlaps the start codon of *leuB* by one nucleotide, and the start codon for *leuC* begins two nucleotides downstream of the *leuB* stop codon, making this a very compactly organized operon (Andreadis and Rosenthal 1992).

In *B. subtilis,* 7 of the 10 genes required for synthesis of leucine, isoleucine, and valine form a single operon called the *ilv-leu* operon (*ilvBNCleuACBD).* However Ward and Zahler (1973) have named the *leu* genes differently, namely on

basis of the order of the enzymatic steps in leucine synthesis which they control (see fig. 1). The IPM isomerase which is the second enzyme in the pathway (encoded by *leuCD* in *E. coli* and *S. typhimurium*), is named *leuBD* and the β-IPM dehydrogenase which is the thirth enzyme in the pathway (encoded by *leuB* in *E. coli* and *S. typhimurium*), is named *leuC*. So the arrangement is similar but the nomenclature is used different.

The *leu* genes in *L. lactis* subsp. *lactis* are organized in a large cluster divided into two units, *leuABCD*orf2 and *ilvDBNCA*. Both units are necessary for leucine biosynthesis, while only the second is required for the synthesis of isoleucine and valine. Both the *leu* and *ilv* gene clusters are preceded by a putative promoter. However, they are not separated by a rho-independent transcription terminator, which suggest that they might form a single operon (Godon *et al.* 1992).

In contrast to the leucine biosynthetic genes in other bacteria, which are part of an operon consisting of *leuA* followed by *leuB*, no homology was found in the *A. caulinodans leuA* downstream region with a *leuB*-like gene. Instead the deduced amino acid sequence of the partial ORF2 shows some homology to an ORF downstream of the *E. coli* genes for ribosomal protein L21 and L27 (Kitakawa, M., Jeong, J., Isono, S., and Isono, K. 1992. unpublished; genbank D13267).

LeuA ORF2	MHRT**GVLLALLSAV**LF**GA**S**T**PLA**K**L**LL**GA**VDP**
ORF	MKQQAGI**GILLALTTAI**CW**GA**L**P**IAM**K**Q**VL**EV**MEP**

Recently it was found that the organization of the leucine biosynthetic genes in the pathogen *Campylobacter jejuni* (Labigne *et al.* 1992), in the spirochete *Leptospira interrogans* serovar *pomona* (Ding and Yelton 1993) and the intracellular pathogen *Brucella abortus* (Essenberg and Sharma 1993) also differs from that found in *E. coli*, *S. typhimurium* and *B. subtilis* (based on functionality).

It would be interesting to identify also the *leuB* gene of *A. caulinodans*, and map it in respect to this *leuA*-locus. The isolation could easily be achieved by complementation of *E. coli* HB101, *E. coli* JA221 or *E. coli* RE521, which all have a *leuB* defect, with the *A. caulinodans* cosmid bank. Many *leuB* genes has been identified recently in this fashion (Ding and Yelton 1993; Essenberg and Sharma 1993; Labigne *et al.* 1992).

7.5.4 Regulation of *leu* gene expression

The leucine operon in *E. coli* and *S. typhimurium* is controlled by a transcription attenuation mechanism (Kolter and Yanofsky 1982; Landick and Yanofsky 1987). Operons known to be regulated by transcription attenuation share a number of common features. They contain a long leader RNA of about 200 bp preceding the first major structural gene. These leader transcripts can be divided into three functionally important segments: a ribosome binding site, a segment coding for a short leader peptide containing several control codons for the relevant regulatory amino acid(s) and overlapping regions of dyad symmetry that are capable of forming stem-and-loop structures including a transcriptional termination signal (Gemmill *et al.* 1979; Wessler and Calvo 1981).

According to a model proposed by Gemmill *et al.* (1979), the formation of alternative secondary structures in the leader RNA determines whether premature termination of transcription occurs at the site 160 base pair downstream from the point of transcription initiation. Formation of a terminator stem-and-loop structure (called attenuator site) located at the end of the leader RNA causes premature transcription termination. Formation of an alternative stem and loop, the preemptor, is thought to preclude formation of the terminator, thereby allowing transcription to continue past the attenuator and through the structural genes of the operon. Which of these stem-and-loop structures form is dependent upon the progress of a ribosome translating the leader transcript. The preemptor stem and loop can form only when a translating ribosome stalls at one of the four tandem leucine control codons in the leader RNA, as in the case when cells are grown under conditions of leucine limitation. Therefore the structural genes of the *leu* operon are transcribed at a high rate during limitation for charged leucine tRNA (Gemmill *et al.* 1979; Bartkus *et al.* 1991).

Analysis of the *A. caulinodans leuA* 5' regulatory region revealed the presence of all these features. An σ70-like promoter sequence is present, located 265 nucleotides upstream of the initiation codon of the *leuA* gene, suggesting a leader RNA of about 258 bp. The suggested σ70 promoter region of the *A. caulinodans leuA* gene TTGCCT-n18-TATGAG has a spacing of 18 base pairs. The same is found in the tryptophan operon of *B. subtilis* TTGACA-n18-TACGAT (Shimotsu *et al.* 1986). The 18 bp spacing is not unusual for *E. coli* promoter regions. The preferred spacing is 17 bp (43.2%) but spacing of 16 and 18 occurs in a relative high frequency (17.4% and 17.1% respectively; Lisser and Margalit 1993).

In addition, several regions of dyad symmetry are readily observed, which are able to form potential stem-and-loop structures, and an open reading frame starting with an GUG initiation codon, containing six leucine codons (including four adjacent Leu-codons; see Table 2). From the six leucine codons, three are rarely used codons (CTT) in *A. caulinodans*, similar to the situation in *E. coli* and *S. typhimurium*. Whether the *A. caulinodans leuA* gene (operon) is controlled by a mechanism of attenuation remains to be experimentally determined.

Table 2. Distribution of putative control codons in *leu* operons.

Organism	Amino acid sequence of polypeptide encoded in leader RNA	reference
E.c , S.t	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-**Leu-Leu-Leu-Leu**-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His-stop	Gemmill *et al.* 1979
L. l	Met-Thr-Tyr-Thr-Gln-Phe-Ser-**Leu-Leu-Leu**-Ile-Lys-Val-Asp-**Leu**-His-stop	Godon *et al.* 1992
*B.s**	Val-Asn-**Leu**-Ile-His-Ser-**Leu-Leu**-Lys-Arg-Ile-Glu-Ile-Ser-Met-Arg-Arg-Pro-Ile-Asn-Arg-Pro-stop	Grandoni *et al.* 1992
A.c	Val-Ser-Ser-Arg-Pro-Thr-Asn-Ile-Ala-Gly-Thr-Ser-Pro-Arg-Ile-Glu-Thr-Asn-Pro-Gly-Asn-Ala-**Leu**-Val-**Leu**-Arg-Ala-**Leu-Leu-Leu-Leu**-Ser-Arg-Pro-stop	this work

*Preliminary results of translational fusion of *lacZ* to this open reading frame indicate that translation does not occur (Grandoni *et al.* 1992). However an unlinked mutation affecting *ilv-leu* expression located in *LeuS*, which encodes a leucine tRNA synthetase, implies that either leucyl tRNA synthetase is directly involved in the regulatory mechanism or that it is indirectly involved by attaching leucine to its cognate tRNA (Vandeyar and Zahler, submitted).

7.5.5 Amino acid homology among Leu proteins

The deduced protein sequence of *S. typhimurium leuA* (Ricca and Calvo 1990) differs in several parts from the *E. coli leuA* protein sequence (Yura *et al.* 1992). However investigation of the *S. typhimurium leuA* DNA sequence and translation of all three reading frames, suggest several sequencing errors that lead to frameshifts and difference from the *E. coli leuA* sequence. In the comparison of the *leuA* and *nifV* genes, these frameshifts were taken into account and resulted an almost identical protein sequence of *S. typhimurium leuA* and *E. coli leuA*. The following sequencing errors are assumed to have taken place in the published sequence of *S. typhimurium leuA* (Ricca and Calvo 1990) which is also deposited in GenBank: P15875; EMBL X51583; STYLEUA; X00059; STLEUA; PIR S08431.

Basepair numbers are taken from sequence in GenBank:
At position 796-797 ATC is present, instead of AC, at 822-823 C instead of CC, at 836-837 T instead of TT, and at 882-883 CCC instead of CC, resulting in two frameshifts:

published *St LeuA*	266	TGAPARPSVQICNIADPSQQSDCRQRRFR	294
published *Ec LeuA*	266	IWRTSQLVSQICNMPIPANKAIVGSGAFA	294
St after frameshifts	266	IWRTSQTVSQICNMPIPANKAIVGSGAFA	294

At position 960-961 T instead of TT and at position 1093 GG instead of G, resulting in one long frameshift:

pub *St LeuA*	321	SEPDTAEPDLPLWPCRRETSHGRDGLQGHRLQHGPPVRRVPEAG	364
pub *Ec LeuA*	321	LNQIQLNLTSRSGRAAVKHRMDEMGYKESEYNLDNLYDAFLKLA	364
St after fs	321	LNQIQLNLTSRSGRAAVKHRMEEMGYKDTDYNMDHLYDAFLKLA	364

At position 1366-1367 CAC or CAT is present, instead of AG, at 1371-1373 AAGG or AAAG is present, instead of AAG and at 1430-1431 GTC instead of GC, resulting in one frameshift and the addition of one amino acid:

published *St LeuA*	456	RARRAGSGRYRREPSWSPLPRR	477
published *Ec LeuA*	456	HGKDALGQVDIVANYNGRRFHGV	478
St after frameshift	456	HGKDALGQVDIVVNHHGRRFHGV	478

7.5.6 Phenotype of the Tn*5* mutant and complementation of the *E. coli leuA* mutant

When site directed Tn*5* mutagenesis was performed on pACleuA, all Tn*5* inserts in the insert were found to be located in the middle of the region sharing a low level of homology with the *E. coli fnr* probe. This region corresponds to the 5' upstream region of the *leuA* gene, and seems to be a hot spot for Tn*5* insertion.

Insertion of a Tn*5* in the Shine Dalgarno consensus sequence (see fig. 2), that separates the regulator region from the structural *leuA* gene, did not have an apparent effect for leucine biosynthesis in *A. caulinodans*. This observation suggests that a low level of expression from the Tn*5* promoter is occurring, sufficient to allow leucine biosynthesis in minimal LSO medium and in symbiosis. However a Tn*5* at the identical position on the plasmid carrying the *A. caulinodans leuA* gene, leads to an abolishment of complementation of the *E. coli leuA* mutant by the pACleuA::Tn*5* plasmid, suggesting that the *A. caulinodans* promoter and regulator mechanism is operative in *E. coli*, while the Tn*5* readthrough is not sufficient to overcome the *leuA* defect. New insertion mutants in the *A. caulinodans* chromosomal *leuA* coding region are presently being constructed to investigate its functionality in leucine biosynthesis in the homologous system.

CHAPTER 8

CONCLUDING REMARKS AND A NITROGEN FIXATION GENE REGULATION MODEL FOR *A. CAULINODANS* ORS571

The primary aim of the studies described in this thesis was the further elucidation of the regulatory pathway responsible for control of the nitrogen fixation (*nif/fix)* genes in *A. caulinodans,* both in culture and *in planta.* The *A. caulinodans* - *S. rostrata* symbiotic system was chosen for the analysis, since it represents the best system for a systematic comparison of *nif/fix* gene regulation in the free-living versus symbiotic state (de Bruijn 1989). The approach which was taken involved focusing on the promoter region of the *A. caulinodans* central *nif/fix* regulatory gene *nifA*, studying the *cis*-acting elements responsible for *nifA* expression and trying to identify *trans*-acting regulatory products/genes interacting with them.

A model for the regulation of the *A. caulinodans* nitrogen fixation (*nif/fix)* genes, deduced from the studies presented in this thesis and previous data from our laboratory and from other groups, is schematically presented in Figure 1.

As observed in all other nitrogen fixing systems (de Bruijn *et al.* 1990; Merrick 1992; see chapter 2), the *A. caulinodans* NifA protein is responsible for activating the other *nif/fix* genes (Pawlowski *et al.* 1987), and this activation requires the product of the *ntrA(rpoN;* σ54*)* gene described here. The promoter regions of *nifHDK* and *fixABCX* have been shown to contain a NifA binding site consensus (UAS; box 2) and a -24/-12 promoter element (box 3; Norel and Elmerich 1987, Kaminski *et al.* 1988).

The *A. caulinodans nifA* promoter mediates the regulatory response to fluctuating nitrogen- and oxygen concentrations (Ratet *et al.* 1989). Under N-limiting conditions, *nifA* expression appears to be controlled by two bi-component regulator systems (see chapter 2), consisting of the regulator proteins NtrC and NtrX, and their sensor partners NtrB and NtrY (Pawlowski *et al.* 1987; 1991). The NtrY protein contains transmembrane domains, suggesting it is membrane-bound and may be sensing the extracellular (periplasmic) concentration of nitrogen, while NtrB may be responding to the intracellular N-concentration (Pawlowski *et al.* 1991).

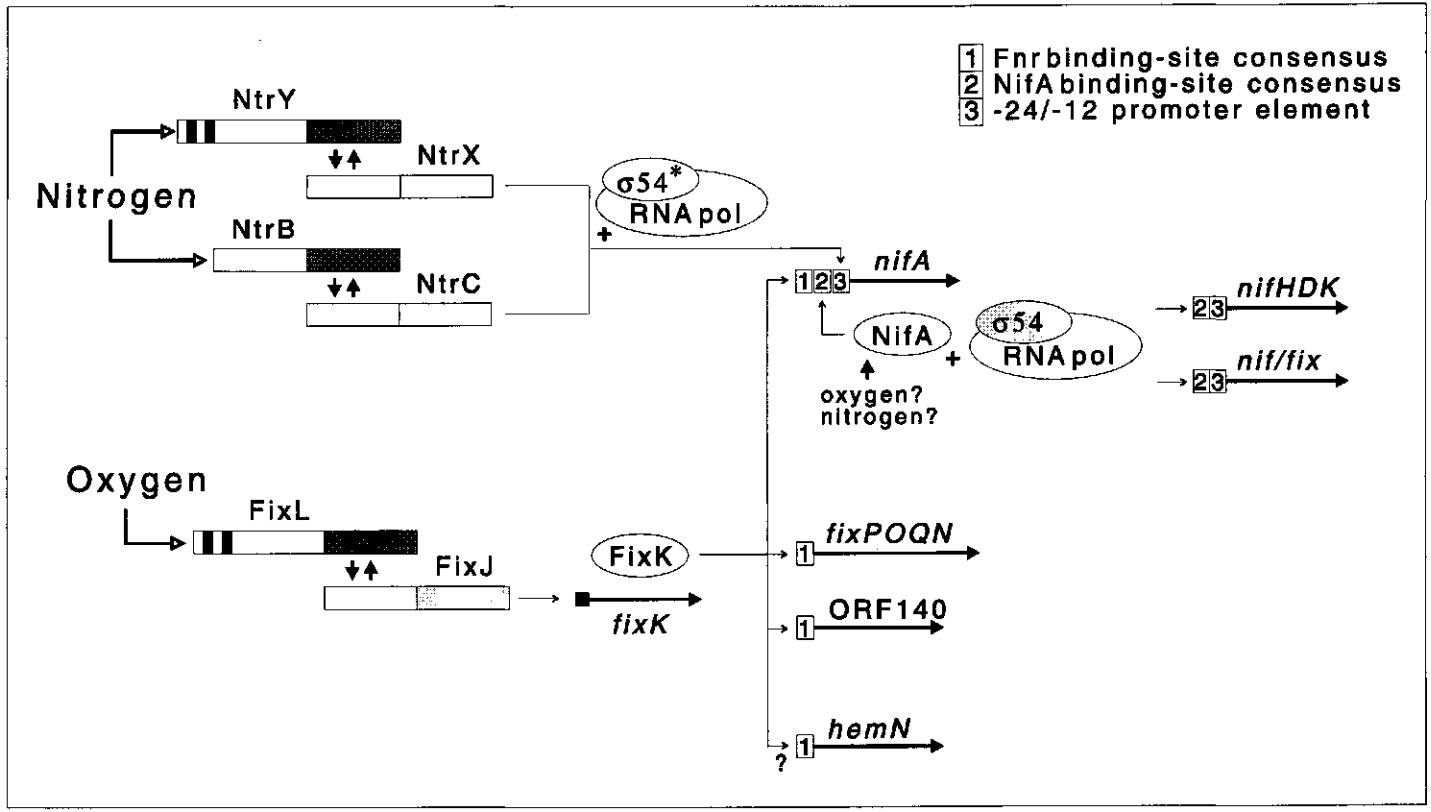

Figure 1. Model for the regulation of nitrogen fixation genes in *A. caulinodans* ORS571.

For details, see text.

The expression of the *ntrYX* operon appears to be affected by *ntrC*, therefore it cannot be excluded that *ntrBC* act through *ntrYX* in regulating the expression of the *nifA* promoter (Pawlowski *et al.* 1991).

Based on DNA homology a very weak NtrC binding site has been identified in the *nifA* promoter region (chapter 5). This site might be involved in *nifA* activation at high NtrC concentrations. Two weak binding sites with very low homology have been identified in the *K. pneumoniae nifLA* promoter region, and have been shown to act cooperatively in activating *nifLA* gene expression (Minchin *et al.* 1989).

Like in *K. pneumoniae*, the *A. caulinodans nifA* promoter contains a -24/-12 promoter element (Box 3; Nees *et al.* 1988; Ratet *et al.* 1989), that is essential for *nifA* expression, as demonstrated in chapter 3, and may be mediating the *ntr* response by *ntrBC* and *ntrYX*. However, the "normal" *ntrA(rpoN*; σ54*)* gene, described in chapter 4, does not appear to be involved in this process, suggesting the presence of an alternative *ntrA(rpoN)*-like gene (σ54*) in *A. caulinodans*, with a distinct specificity for the *nifA* promoter. The latter could be reflected in the differences in DNA sequence (denoted by bold italics) surrounding the -24/-12 regions of the *nifH* (Norel and Elmerich 1987), *nifO* and *fixA* promoters (Kaminski *et al.* 1988) versus the *nifA* promoter of *A. caulinodans* ORS571.

	NifA UAS	**-24/-12**
nifO	No UAS	cTGGCACGggcaTTGCT
fixA	TGTggcaTgccagACA	tTGGTACGacacTTGCT
nifH1	TGTcgcgTttgaaACA	cTGGCACAcccgTTGCA
nifH2	TGTttccTttccaACA	aTGGCACGgtcgTTGCT
	TGTcaggTtcgagACA	
consensus	TGT....T.....ACA	cTGGCACGgcctTTGCA
nifA	GGTttccattataACA	g***A***GGC***TG***AtcccT***C***GCA

Oxygen control of *nifA* expression is mediated by the *fixLJ* genes, acting through *fixK* (Kaminski *et al.* 1991; Kaminski and Elmerich 1991). FixL is likely to be the oxygen sensor, by analogy with *R. meliloti* (Gilles-Gonzalez *et al.* 1991) and activates FixJ (chapter 2), which controls the expression of the *fixK* gene. FixK, an Fnr-like transcriptional activator (Spiro and Guest 1990) has been identified in several other rhizobia like *R. meliloti* (Batut *et al.* 1989), *B. japonicum* (Anthamatten *et al.* 1992) and *R. leguminosarum* (FnrN; Colonna-Romano *et al.* 1990).

In *A. caulinodans fixK* has been found to positively regulate the transcription of *nifA* (Kaminski *et al.* 1991).

The target for the FixLJK system in the *nifA* promoter region is unknown, but it is plausible that the Fnr binding site consensus sequence, identified in the *nifA* promoter region (Box 1; Nees *et al.* 1988; Ratet *et al.* 1989) and found to be important for *nifA* expression, may be involved in mediating the oxygen response. Genetic evidence for the interaction of FixK in *R. meliloti* with an Fnr-binding site has recently been shown by Waelkens, *et al.* (1992).

Like in *R. meliloti*, some *A. caulinodans* operons have recently been shown to be activated directly by FixK, namely *fixPOQN* and an ORF140 located downstream of *fixJ* (Mandon *et al.* 1993). Moreover, a Fnr-binding site consensus is present in the *fixPOQN* promoter region. In addition, FixK may not only activate the central *nif* regulator *nifA* and other *fix*-genes, but also genes involved in 'housekeeping' functions (like *hemN*, required for heme biosynthesis), under microaerobic conditions (see chapter 6).

The *A. caulinodans nifA* gene may also be subject to autoregulation, which may be mediated through the NifA Upstream Activation Sequence (UAS; Box 2), found in the *nifA* promoter region (Nees *et al.* 1988; Ratet *et al.* 1989). Autoregulation of *nifA* has been found in *K. pneumoniae*, *Rhizobium* species and *B. japonicum*.

The *A. caulinodans* NifA protein itself is probably regulated by nitrogen and oxygen. As observed in other systems (see chapter 2), NifA contains the four conserved cysteine residues, correctly spaced (C-X11-C-X19-C-X5-C) which are probably involved in sensitivity to oxygen (Ratet *et al.* 1989; Fischer *et al.* 1988; Hennecke *et al.* 1988). In addition it has been show that the N-terminal domain of NifA modulates NifA activity in response to changes in the nitrogen status (Loroch and Ludwig, 9th Int. Congress on N-fixation, Cancun, Mexico 1992).

FixK activation of a σ54 (-24/-12) promoter has not been reported. Since FixK is very homologous to Fnr which activates σ70 genes carrying a Fnr binding site (around -40) nearby the binding site of the RNA polymerase. The mechanism here is likely to be different, activation of *nifA* might involve loop formation as found for other σ54 regulated promoters. The Fnr/FixK binding site is located about 100 bp upstream of the transcription start point and an potential IHF binding site has been identified between the Fnr binding site consensus and the -24/-12 promoter element, which may bring these sites in close proximity for transcription activation to occur.

The analysis of *A. caulinodans nifA* gene expression has clearly revealed that this unusual nitrogen fixing organism appears to combine regulatory circuitry's specific to strict diazotrophs with those of strict symbionts. Although substantial progress has been made in elucidating the signal transduction pathways mediating nitrogen- and oxygen-control of *A. caulinodans nif/fix* gene expression, several pieces of the puzzle are still missing. It is to be expected that further in depth studies on the *nif* and *ntr* regulatory genes and their targets, using some of the tools and approaches described in this thesis, will fill the gaps and lead to a more complete picture of *nif/fix* gene regulation.

REFERENCES

Alazard, R., Betermier, M., and Chandler, M. 1992. *Escherichia coli* integration host factor stabilizes bacteriophage Mu repressor interactions with operator DNA *in vitro.* Mol. Microbiol. 6:1707-1714.

Albright, L. M., Huala, E., and Ausubel, F. M. 1989a. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. Annu. Rev. Genet. 23:311-336.

Albright, L. M., Ronson, C. W., Nixon, B. T., and Ausubel, F. M. 1989b. Identification of a gene linked to *Rhizobium meliloti ntrA* whose product is homologous to a family of ATP-binding proteins. J. Bacteriol. 171:1932-1941.

Alias, A., Cejudo, F. J., Chabert, J., Willison, J. C., and Vignais, P. M. 1989. Nucleotide sequence of wild-type and mutant *nifR4 (ntrA)* genes of *Rhodobacter capsulatus*: Identification of an essential glycine residue. Nucleic Acids Res. 17:5377.

Agron, P. G., Ditta, G. S., and Helinski, D. R. 1992. Mutational analysis of the *Rhizobium meliloti nifA* promoter. J. Bacteriol. 174:4120-4129.

Agron, P. G., Ditta, G. S., and Helinski, D. R. 1993. Oxygen regulation of *nifA* transcription *in vitro.* Proc. Natl. Acad. Sci. USA 90:3506-3510.

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

Andreadis, A. and Rosenthal, E. R. 1992. The nucleotide sequence of *leuB* from *Salmonella typhimurium.* Biochimica et Biophysica Acta 1129:228-230.

Anthamatten, D. and Hennecke, H. 1991. The regulatory status of the *fixL*- and *fixJ*-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti.* Mol. Gen. Genet. 225:38-48.

Anthamatten, D., Scherb, B., and Hennecke, H. 1992. Characterization of a *fixLJ*-regulated *Bradyrhizobium japonicum* gene sharing similarity with the *Escherichia coli fnr* and *Rhizobium meliloti fixK* genes. J. Bacteriol. 174:2111-2120.

Appleby, C. A. 1984. Leghemoglobin and *Rhizobium* respiration. Annu. Rev. Plant Physiol. 35:443-478.

Arnold, W., Rump, A., Klipp, W., Priefer, U., and Puhler, A. 1988. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae.* J. Mol. Biol. 203:715-738.

Arnott, M., Sidoti, C., Hill, S., and Merrick, M. 1989. Deletion analysis of the nitrogen fixation regulatory gene, *nifL* of *Klebsiella pneumoniae.* Arch. Microbiol. 151:180-182.

Austin, S. and Dixon, R. 1992. The prokaryotic enhancer binding protein NtrC has an ATPase activity which is phosphorylation and DNA dependent. EMBO J. 11:2219-2228.

Austin, S., Henderson, N., and Dixon, R. 1987. Requirements for transcriptional activation *in vitro* of the nitrogen-regulated *glnA* and *nifLA* promoters from *Klebsiella pneumoniae*: Dependence on activator concentration. Mol. Microbiol. 1:92-100.

Austin, S., Kundrot, C., and Dixon, R. 1991. Influence of a mutation in the putative nucleotide binding site of the nitrogen regulatory protein NtrC on its positive control function. Nucleic Acids Res. 19:2281-2287.

Ausubel, F. M. 1984. Regulation of nitrogen fixation genes. Cell 37:5-6.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore D. D., Seidman, J. G., Smith, J. A., and Struhl, K. 1989. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Avissar, Y. J. and Nadler, K. D. 1978. Stimulation of tetrapyrrole formation in *Rhizobium japonicum* by restricted aeration. J. Bacteriol. 135:782-789.

Bali, A., Blanco, G., Hill, S., and Kennedy, C. 1992. Excretion of ammonium by a *nifL* mutant of *Azotobacter vinelandii* fixing nitrogen. Appl. Environ. Microbiol. 58:1711-1718.

Bartkus, J. M., Tyler, B., and Calvo, J. M. 1991. Transcription attenuation-mediated control of *leu* operon expression: Influence of the number of Leu control codons. J. Bacteriol. 173:1634-1641.

Batut, J., Daveran-Mingot, M. L., David, M., Jacobs, J., Garnerone, A. M., and Kahn, D. 1989. *fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli*, regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti*. EMBO J. 8:1279-1286.

Batut, J., Santero, E., and Kustu, S. 1991. *In vitro* activity of the nitrogen fixation regulatory protein FixJ from *Rhizobium meliloti*. J. Bacteriol. 173:5914-5917.

Beltzer, J. P., Chang, L.-F. L., Hinkkanen, A. E., and Kohlhaw, G.B. 1986. Structure of Yeast LEU4. J. Biol. Chemistry 261:5160-5167.

Bender, R. A. 1991. The role of the Nac protein in the nitrogen regulation of *Klebsiella aerogenes*. Mol. Microbiol. 5:2575-2580.

Bennet, L. T., Cannon, F., and Dean, D. 1988. Nucleotide sequence and mutagenesis of the *nifA* gene from *Azotobacter vinelandii*. Mol. Microbiol 2:315-321.

Berger, D. K., Woods, D. R., and Rawlings, D. E. 1990. Complementation of *Escherichia coli* σ^{54} *(NtrA)*-dependent formate hydrogenlyase activity by a cloned *Thiobacillus ferrooxidans ntrA* gene. J. Bacteriol. 172:4399-4406.

Beringer, J. E. 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 84:188-189.

Beynon, J., Ally, A., Cannon, M., Cannon, F., Jacobson, M., Cash, V., and Dean, D. 1987. Comparative organization of nitrogen fixation-specific genes from *Azotobacter vinelandii* and *Klebsiella pneumoniae*: DNA sequence of the *nifUSV* genes. J. Bacteriol. 169:4024-4029.

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

Birkmann, A., Sawers, R. G., and Böck, A. 1987. Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. Mol. Gen. Genet. 210:535-542

Bishop, P. E. and Joerger, R. D. 1990. Genetics and molecular biology of alternative nitrogen fixation systems. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41:109-125.

Bishop, P. E. and Premakumar, R. 1992. Alternative nitrogen fixation systems. Pages 736-762 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans (eds.). Routledge, Chapman and Hall, Inc. New York.

Blanco, G., Woodley, P., Drummond, M., Bali, A., and Kennedy, C. 1993. The *nifL* gene of *Azotobacter vinelandii*: Novel features of sequence, expression and mutant phenotypes. Pages 429-434 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Boyer, H. W., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *E.coli*. J. Mol. Biol. 41:459-472.

Brendel, V., Hamm, G. H., and Trifonov, E. N. 1986. Terminators of transcription with RNA polymerase from *Escherichia coli*: What they look like and how to find them. J. Biomol. Struct. Dyn. 2:705-723.

Broughton, W. J., and Dilworth, M. Y. 1971. Control of leghaemoglobin synthesis in snake beans. Biochem. J. 125:1075-1080.

Broughton, W. J., and John, C. K. 1979. Rhizobia in tropical legumes. III experimentation and supply in Malaysia 1927-1976. Pages 113-136 in: Soil Microbiology and Plant Nutrition. W.J. Broughton, C.K. John, J.C. Rajarao and B. Lin (eds.). University of Malaysia Press. Kuala Lumpur, Malaysia.

de Bruijn, F. J. 1987. Tn*5* mutagenesis to map genes. Meth. Enzymol. 154:175-196.

de Bruijn, F. J. 1989. The unusual symbiosis between the diazotrophic stem-nodulating bacterium *Azorhizobium caulinodans* ORS571 and its host, the tropical legume *Sesbania rostrata*. Pages 457-504 in: Plant-Microbe Interactions Vol 3. E. Nester and T. Kosuge (eds.). McGraw-Hill, New York.

de Bruijn, F. J. and Ausubel, F. M. 1983. The cloning and characterization of the *glnF (ntrA)* gene of *Klebsiella pneumoniae*: Role of *glnF* (*ntrA*) in the regulation of nitrogen fixation *(nif)* and other nitrogen assimilation genes. Mol. Gen. Genet. 192:342-353.

de Bruijn, F. J. and Downie, J. A. 1991. Biochemical and molecular studies of symbiotic nitrogen fixation. Curr. Opin. in Biotechnology 2:184-192.

de Bruijn, F. J., Felix, G., Grunenberg, B., Hoffman, H. J., Metz, B., Ratet, P., Simons-Schreier, A., Szabados, L., Welters, P., and Schell, J. 1989. Regulation of plant genes specifically induced in nitrogen-fixing nodules: Role of *cis*-acting elements and *trans*-acting factors in leghemoglobin gene expression. Plant Mol. Biol. 13:319-325.

de Bruijn, F. J., Hilgert, U., Stigter, J., Schneider, M., Meyer zA, H., Klosse, U., and Pawlowski, K. 1990. Regulation of nitrogen fixation and assimilation genes in the free-living versus symbiotic state. Pages 33-44 in: Nitrogen Fixation: Achievements and Objectives. P. Gresshoff, E. Roth, G. Stacey, W. Newton (eds.). Chapman and Hall, New York-London.

de Bruijn, F. J. and Lupski, J. R. 1984. The use of transposon Tn*5* mutagenesis in the rapid generation of correlation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids - a review. Gene 27:131-149.

de Bruijn, F. J., Pawlowski, K., Ratet, P., Hilgert, U., Wong, C. H., Meyer zA, H., and Schell, J. 1988. Molecular genetics of nitrogen fixation by *Azorhizobium caulinodans* ORS571, the diazotrophic stem nodulating symbiont of *Sesbania rostrata*. Pages 351-355 in: Nitrogen Fixation: Hundert Years After. H. Bothe, F.J. de Bruijn and W.E. Newton (eds.). Gustav Fischer, Stuttgart, New York.

Buchanan-Wollaston, V., Cannon, M. C., Beynon, J. L., and Cannon, F. C. 1981. Role of the *nifA* gene product in the regulation of *nif* expression in *Klebsiella pneumoniae*. Nature 294:776-778.

Buck, M. 1986. Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around positions -12 and -24 for activation by the *nifA* and *ntrC (glnG)* products. J. Bacteriol. 166:545-551.

Buck, M. and Cannon, W. 1989. Mutations in the RNA polymerase recognition sequence of the *Klebsiella pneumoniae nifH* promoter permitting transcriptional activation in the absence of NifA binding to upstream activator sequences. Nucleic Acids Res. 7:2597-2612.

Buck, M. and Cannon, W. 1992a. Activator-independent formation of a closed complex between σ^{54}-holoenzyme and *nifH* and *nifU* promoters of *Klebsiella pneumoniae*. Mol. Microbiol. 6:1625-1630.

Buck, M. and Cannon, W. 1992b. Specific binding of the transcription factor sigma-54 to promoter DNA. Nature 358: 422-424.

Buck, M., Cannon, W., and Woodcock, J. 1987. Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involved DNA loop formation. Mol. Microbiol. 1:243-249.

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

Buck, M., Miller, S., Drummond, M., and Dixon, R. 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature 320:374-378.

Cannon, W. and Buck, M. 1992. Central domain of the positive control protein NifA and its role in transcriptional activation. J. Mol. Biol. 225:271-286.

Cannon, W., Claverie-Martin, F., Austin, S., and Buck, M. 1993. Core RNA polymerase assists binding of the transcription factor σ54 to promoter DNA. Mol Microbiol. 8:287-298.

Cannon, W. V., Kreutzer, R., Kent, H. M., Morett, E., and Buck, M. 1990. Activation of the *Klebsiella pneumoniae nifU* promoter: Identification of multiple and overlapping upstream NifA binding sites. Nucleic Acids Res. 18:1693-1701.

Casadaban, M. J., Martinez-Arias, A., Shapira, S. K., and Chou, J. 1983. β-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Meth. Enzymol. 100:293-308.

Chen, J.-J., Yang, J. M., Petryshyn, R., Kosower, N., and London, I. M. 1989 Disulfide bond formation in the regulation of eIF-2α kinase by heme. J. Biol. Chem. 264:9559-9564.

Chen, K. C.-K., Chen, J.-S., and Johnson, J. L. 1986. Structural features of multiple *nifH*-like sequences and very biased codon usage in nitrogenase genes of *Clostridium pasteurianum*. J. Bacteriol. 166:162-172.

Clark, J. M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res. 16:9677-9686.

Claverie-Martin, F. and Magasanik, B. 1991. Role of integration host factor in the regulation of the *glnHp2* promoter of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 88:1631-1635.

Colbeau, A. and Vignais, P. M. 1992. Use of *hupS::lacZ* gene fusion to study regulation of hydrogenase expression in *Rhodobacter capsulatus*: Stimulationby H_2. J. Bacteriol. 174:4258-4264.

Collins, J. J., Roberts, G. P., and Brill, W. J. 1986. Posttranscriptional control of *Klebsiella pneumoniae nif* mRNA stability by the *nifL* product. J. Bacteriol. 168:173-178.

Colonna-Romano, S., Arnold, W., Schluter, A., Boistard, P., Puhler, A., and Priefer, U. B. 1990. *Rhizobium leguminosarum* FnrN shows structural and functional homology to *Rhizobium meliloti* FixK. Mol. Gen. Genet. 223:138-147.

Contreras, A. and Drummond, M. 1991. Cys^{184} and Cys^{187} of NifL protein of *Klebsiella pneumoniae* are not absolutely required for inhibition of NifA activity. Gene 103:83-86.

Contreras, A. Drummond, M., Bali, A., Blanko, G., Garcia, E., Bush, G., Kennedy, C., and Merrick, M. 1991. The product of the nitrogen fixation regulatory gene *nfrX* of *Azotobacter vinelandii* is functionally and structurally homologous to the uridylyltransferase encoded by *glnD* in Enteric bacteria. J. Bacteriol.173:7741-7749.

Coomber, S. A., Jones, R. M., Jordan, P. M., and Hunter, C. N. 1992. A putative anaerobic coproporphyrinogen III oxidase in *Rhodobacter sphaeroides*. I. Molecular cloning, transposon mutagenesis and sequence analysis of the gene. Mol. Microbiol. 6:3159-3169.

Coppard, J. R. and Merrick, M. J. 1991. Cassette mutagenesis implicates a helix-turn-helix motif in promoter recognition by the novel RNA polymerase sigma factor σ^{54}. Mol. Microbiol. 5:1309-1317.

Craig, N. L. and Nash, H. A. 1984. *E. coli* integration host factor binds to specific sites in DNA. Cell 39:707-716.

de Crombrugghe, B., Busby, S., and Buc, H. 1984. Cyclic AMP receptor protein: Role in transcription activation. Science 224:831-838.

Dailey, H. A. 1990. Conversion of coproporphyrinogen to protoheme in higher eukaryotes and bacteria: terminal three enzymes. Pages 123-161 in: Biosynthesis of Heme and Chlorophylls. H. A. Dailey (ed.). McGraw Hill, New York.

David, M., Daveran, M. L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., Kahn, D. 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell 54:671-683.

David, M., Domergue, O., Pognonec, P., and Kahn, D. 1987. Transcription patterns of *Rhizobium meliloti* symbiotic plasmid pSym: Identification of *nifA*-independent *fix* genes. J. Bacteriol. 169:2239-2244.

Debarbouille, M., Martin-Verstraete, I., Kunst, F., and Rapoport, G. 1991. The *Bacillus subtilis sigL* gene encodes an equivalent of σ^{54} from gram-negative bacteria. Proc. Natl. Acad. Sci. USA 88:9092-9096.

Ding, M. and Yelton, D. B. 1993. Cloning and analysis of the *leuB* gene of *Leptospira interrogans* serovar *pomona*. J. Gen. Microbiol. 139:1093-1103.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria : Construction of a gene bank of *Rhizobium meliloti.* Proc. Natl. Acad. Sci. USA. 77:7347-7351.

Ditta, G., Virts, E., Palomares, A., and Kim, C. H. 1987. The *nifA* gene of *Rhizobium meliloti* is oxygen regulated. J. Bacteriol. 169:3217-3223.

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

Dixon, R., Eady, R. R., Espin, G., Hill, S., Iaccarino, M., Kahn, D., and Merrick, M. 1980. Analysis of regulation of *Klebsiella pneumoniae* nitrogen fixation *(nif)* gene cluster with gene fusions. Nature 286:128-132.

Dixon, R. A., Henderson, N. C., and Austin, S. 1988. DNA supercoiling and aerobic regulation of transcription from the *Klebsiella pneumoniae nifLA* promoter. Nucleic Acids Res. 16:9933-9946.

Dodd, I. B. and Egan, J. B. 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acid Res. 18:5019-5026.

Donald, R. G. K., Nees, D. W., Raymond, C. K., Loroch, A. I., and Ludwig, R. A. 1986. Characterization of three genomic loci encoding *Rhizobium* sp. strain ORS571 N_2 fixation genes. J. Bacteriol. 165:72-81.

Dreyfus, B. and Dommergues, Y. R. 1981. Nitrogen-fixing nodules induced by *Rhizobium* on the stem of the tropical legume *Sesbania rostrata.* FEMS Microbiol. Lett. 10:313-317.

Dreyfus, B., Garcia, J. L., and Gillis, M. 1988. Characterization of *Azorhizobium caulinodans* gen. nov., sp. nov., a stem-nodulating nitrogen-fixing bacterium isolated from *Sesbania rostrata.* Int. J. Syst. Bacteriol. 38:89-98.

Drummond, M., Clements, J., Merrick, M., and Dixon, R. 1983. Positive control and autogenous regulation of the *nifLA* promoter in *Klebsiella pneumoniae.* Nature 301:302-313.

Drummond, M. H., Contreras, A., and Mitchenall, L. A. 1990. The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae.* Mol. Microbiol. 4:29-37.

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

Elmerich, C., Dreyfus, B. L., Reysset, G., and Aubert, J. P. 1982. Genetic analysis of nitrogen fixation in a tropical fast-growing *Rhizobium*. EMBO. J. 1:499-503.

Essenberg, R. C. and Sharma, Y. K. 1993. Cloning of genes for proline and leucine biosynthesis from *Brucella abortus* by functional complementation in *Escherichia coli*. J. Gen. Microbiol. 139:87-93.

Evans, D. J., Jones, R., Woodley, P. R., Wilborn, J. R., and Robson, R. L. 1991. Nucleotide sequence and genetic analysis of the *Azotobacter chroococcum nifUSVWZM* gene cluster, including a new gene *(nifP)* which encodes a serine acetyltransferase. J. Bacteriol. 173:5457-5469.

Figurski, D. H. and Helinski, D. R. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA 76:1648-1652.

Fischer, H. M., Acuna, G., Anthamatten, D., Arigoni, F., Babst, M., Brouwer, P., Kaspar, T., Kullik, I., Preisig, O., Scherb, B., Weidenhaupt, M., and Hennecke, H. 1993. Two oxygen-responsive regulatory cascades control nitrogen fixation genes in *Bradyrhizobium japonicum*. Pages 411-416 in: New Horizons in Nitrogen Fixation, R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Fischer, H. M., Bruderer, T., and Hennecke, H. 1988. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: Identification of indispensible cysteine residues potentially involved in redox activity and/or metal binding. Nucleic Acids Res. 16:2207-2224.

Fischer, H. M., Fritsche, S., Herzog, B., and Hennecke, H. 1989. Critical spacing between two essential cysteine residues in the interdomain linker of the *Bradyrhizobium japonicum* NifA protein. FEBS Lett. 255:167-171.

Foster-Hartnett, D. and Kranz, R. G. 1992. Analysis of the promoters and upstream sequences of *nifA1* and *nifA2* in *Rhodobacter capsulatus*; activation requires *ntrC* but not *rpoN*. Mol. Microbiol. 6:1049-1060.

Freundlich, M., Ramani, N., Mathew, E., Sirko, A., and Tsui, P. 1992. The role of integration host factor in gene expression in *Escherichia coli*. Mol. Microbiol. 6:2557-2563.

Friedman, D. I. 1988. Integration host factor: A protein for all reasons. Cell. 55:545-554.

Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J., and Ausubel, F. M. 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium mutants*. Gene 18:289-296.

Fuhrhop, J.-H. and Smith, K. M. 1975. Laboratory methods. Pages 757-869 in: Porphyrins and Metalloporphyrins. K.M. Smith (ed.). Elsevier Sci. Pub. Comp. Amsterdam, The Netherlands.

Fytlovich, S., Gervais, M., Agrimonti, C., and Guiard, B. 1993. Evidence for an interaction between the CYP1 (HAP1) activator and a cellular factor during heme-dependent transcriptional regulation in the yeast *Saccharomyces cerevisiae*. EMBO J. 12:1209-1218.

Gemmill, R. M., Wessler, S. R., Keller, E. B., and Calvo, J. M. 1979. *Leu* operon of *Salmonella typhimurium* is controllled by an attenuation mechanism. Proc. Natl. Acad. Sci. USA. 76:4941-4945.

Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. 1991. A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. Nature 350:170-172.

Godon, J-J., Chopin, M-C., and Ehrlich, S. D. 1992. Branched-chain amino acid biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. 174:6580-6589.

Grandoni, J. A., Zahler, S. A., and Calvo, J. M. 1992. Transcriptional regulation of the *ilv-leu* operon of *Bacillus subtilis*. J. Bacteriol. 174:3212-3219.

Gribskov, M. and Burgess, R. R. 1986. Sigma factors from *Escherichia coli*, *Bacillus subtilis*, phage SPO1 and T4 are homologous proteins with conserved secondary structures. Nucleic Acids Res. 14:6745-6763.

Gussin, G. N., Ronson, C. W., and Ausubel, F. M. 1986. Regulation of nitrogen fixation genes. Ann. Rev. Genet. 20:567-591.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557-580.

Haselkorn, R. and Buikema, W. J. 1992. Nitrogen fixation in Cyanobacteria. Pages 166-190 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans (eds.). Routledge, Chapman and Hall, Inc. New York.

Helmann, J. D. and Chamberlin, M. J. 1988. Structure and function of bacterial sigma factors. Ann. Rev. Biochem. 57:839-872.

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. 1989. A simple method for site-directed mutagenesis using polymerase chain reaction. Nucleic Acids Res. 17:6545-6551.

Hennecke, H. 1990. Regulation of bacterial gene expression by metal-protein complexes. Mol. Microbiol. 4:1621-1628.

Hertig, C., Li, R. Y., Louarn, A. M., Garnerone, A. M., David, M., Batut, J., Kahn, D., and Boistard, P. 1989. *Rhizobium meliloti* regulatory gene *fixJ* activates transcription of *R. meliloti nifA* and *fixK* genes in *Escherichia coli*. J. Bacteriol. 171:1736-1738.

Heusterspreute, M., Ha-Thi, V., Emery, S., Tournis-Gamble, S., Kennedy, N., and Davison, J. 1985. Vectors with restriction site banks IV. pJRD184, a 3793-bp plasmid vector having 43 unique cloning sites. Gene 39:299-304.

Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W., and Hermodson, M. A. 1986. A family of related ATP-binding subunits coupled to many distinct processes in bacteria. Nature 323:448-450.

Hirsch, A. M. 1992. Developmental biology of legume nodulation. New Phytol. 122:211-237.

Hirsch, P. R. and Beringer, J. E. 1984. A physical map of pPH1JI and pJB41. Plasmid 12:139-141.

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 sigma factor. Proc. Natl. Acad. Sci. USA 82:7525-7529.

Holtel, A. and Merrick, M. J. 1988. Identification of the *Klebsiella pneumoniae glnB* gene: Nucleotide sequence of wild-type and mutant alleles. Mol. Gen. Genet. 215:134-138.

Holtel, A. and Merrick, M. J. 1989. The *Klebsiella pneumoniae* P_{II} protein *(glnB* gene product) is not absolutely required for nitrogen regulation and is not involved in NifL-mediated *nif* gene regulation. Mol. Gen. Genet. 217:474-480.

Hoover, T. R., Robertson, A. D., Cerny, R. L., Hayes, R. N., Imperial, J., Shah, V. K., and Ludden, P. W. 1987. Identification of the V factor needed for synthesis of the iron-molybdenum cofactor of nitrogenase as homocitrate. Nature 329:855-857.

Hoover, T. R., Santero, E., Porter, S., and Kustu, S. 1990. The integration host factor stimulates interaction of RNA polymerase with NifA, the transcriptional activator for nitrogen fixation operons. Cell 63:11-22.

Huala, E. and Ausubel, F. M. 1989. The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti nifH* promoter. J. Bacteriol. 171:3354-3365.

Hübner, P., Willison, J. C., Vignais, P. M., and Bickle, T. A. 1991. Expression of regulatory *nif* genes in *Rhodobacter capsulatus*. J. Bacteriol. 173:2993-2999.

Hunt, T. P. and Magasanik, B. 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc. Natl. Acad. Sci. USA 82:8453-8457.

Inouye, S., Yamada, M., Nakazawa, A., and Nakazawa, T. 1989. Cloning and sequence analysis of the *ntrA (rpoN)* gene of *Pseudomonas putida*. Gene 85:145-152.

Irvine, A. S. and Guest, J. R. 1993. *Lactobacillus casei* contains a member of the Crp-Fnr family. Nucleic Acids Res. 21:753.

Ishimoto, K. S. and Lory, S. 1989. Formation of pilin in *Pseudomonas aeruginosa* requires the alternative σ factor (RpoN) of RNA polymerase. Proc. Natl. Acad. Sci. USA. 86:1954-1957.

Jayaraman, P. S., Cole, J. A., and Busby, S. J. W. 1988. Mutational analysis of the nucleotide sequence at the FNR-dependent *nirB* promoter in *Escherichia coli*. Nucleic Acids Res. 17:135-145.

Jensen, E. O., Marcker, K. A., and Villadsen, I. S. 1986. Heme regulates the expression in *Saccharomyces cerevisiae* of chimaeric genes containing 5'-flanking soybean leghemoglobin sequences. EMBO J. 5:843-847.

Joerger, R. D., Jacobson, M. R., and Bishop, P. E. 1989. Two *nifA*-like genes required for expression of alternative nitrogenase by *Azotobacter vinelandii*. J. Bacteriol. 171:3258-3267.

Jones, R. and Haselkorn, R. 1989. The DNA sequence of the *Rhodobacter capsulatus ntrA*, *ntrB* and *ntrC* gene analogues required for nitrogen fixation. Mol. Gen. Genet. 215:507-516.

Jordan, P. M. 1990. Biosynthesis of 5-aminolevulinic acid and its transformation into coproporphyrinogen in animals and bacteria. Pages 55-121 in: Biosynthesis of Heme and Chlorophylls. H. A. Dailey, (ed.). McGraw Hill, New York.

Kahn, D. and Ditta, G. 1991. Modular structure of FixJ: Homology of the transcriptional activator domain with the -35 domain of sigma factors. Mol. Microbiol. 5:987-997.

Kaluza, K., Alvarez-Morales, A., and Hennecke, H. 1985. Oligonucleotide-directed mutagenesis of the *Rhizobium japonicum nifH* promoter. FEBS Lett. 188:37-42.

Kaminski, P. A. and Elmerich, C. 1991. Involvement of *fixLJ* in the regulation of nitrogen fixation in *Azorhizobium caulinodans*. Mol. Microbiol. 5:665-673.

Kaminski, P. A., Mandon, K., Arigoni, F., Desnoues, N., and Elmerich, C. 1991. Regulation of nitrogen fixation in *Azorhizobium caulinodans*: identification of a *fixK*-like gene, a positive regulator of *nifA*. Mol. Microbiol. 5:1983-1991.

Kaminski, P.A., Norel, F., Desnoues, N., Kush, A., Salzano, G., and Elmerich C. 1988. Characterization of the *fixABC* region of *Azorhizobium caulinodans* ORS571 and identification of a new nitrogen fixation gene. Mol. Gen. Genet. 214:496-502.

Kammann, M., Laufs, J., Schell, J., and Gronenborn, B. 1989. Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR). Nucleic Acid Res. 17:5404.

Keithly, J. H. and Nadler, K.D. 1983. Protoporphyrin formation in *Rhizobium japonicum*. J. Bacteriol. 154:838-845.

Kent, H. M., Buck, M., and Evans, D. J. 1989. Cloning and sequencing of the *nifH* gene of *Desulfovibrio gigas*. FEMS Microbiol. Lett. 610:73-78.

Khan, H., Buck, M., and 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.

Kim, C. H., Helinski, D. R., and Ditta, G. 1986. Overlapping transcription of the *nifA* regulatory gene in *Rhizobium meliloti*. Gene 50:141-148.

Kim, H., Yu, C., and Maier, R. J. 1991. Common *cis*-acting region responsible for transcriptional regulation of *Bradyrhizobium japonicum* hydrogenase by nickel, oxygen, and hydrogen. J. Bacteriol. 173:3993-3999.

Kohler, T., Cayrol, J. M., Ramos, J. L., and Harayama, S. 1989a. Nucleotide and deduced amino acid sequence of the RpoN σ-factor of *Pseudomonas putida*. Nucleic Acids Res. 23:10125.

Kohler, T., Harayama, S., Ramos, J. L., and Timmis, K. N. 1989b. Involvement of *Pseudomonas putida* RpoN σ-factor in regulation of various metabolic functions. J. Bacteriol. 171:4326-4333.

Kolter, R. and Yanofsky, C. 1982. Attenuation in amino acid biosynthetic operons. Ann. Rev. Genet. 16:113-134.

Kong, Q. T., Wu, Q. L., Ma, Z. F., and Shen, S. C. 1986. Oxygen sensitivity of the *nifLA* promoter of *Klebsiella pneumoniae*. J. Bacteriol. 166:353-356.

Kranz, R. G. and Haselkorn, R. 1988. Ammonia-constitutive nitrogen fixation mutants of *Rhodobacter capsulatus*. Gene 71:65-74.

Kranz, R. G., Pace, V. M., and Caldicott, I. M. 1990. Inactivation, sequence, and *lacZ* fusion analysis of a regulatory locus required for repression of nitrogen fixation genes in *Rhodobacter capsulatus*. J. Bacteriol. 172:53-62.

Kullik, I. 1992. PhD thesis.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H., and Fischer, H. M. 1991. *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the σ^{54} gene *(rpoN)*. J. Bacteriol. 173:1125-1138.

Kur, J., Hasan, N., and Szybalski, W. 1989. Physical and biological consequences of interactions between integration host factor (IHF) and coliphage lambda late p'_R promoter and its mutants. Gene 81:1-15.

Kustu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. 1989. Expression of σ^{54} *(ntrA)*-dependent genes is probably united by a common mechanism. Microbiol. Revs. 53:367-376.

Labigne, A., Courcoux, P., and Tompkins, L. 1992. Cloning of *Campylobacter jejuni* genes required for leucine biosynthesis, and construction of *leu*-negative mutant of *C. jejuni* by shuttle transposon mutagenesis. Res. Microbiol. 143:15-26.

Landick, R. and Yanofsky, C. 1987. Transcription attenuation. Pages 1276-1301 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, Vol. 2. F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds.). American Society for Microbiology, Washington, D.C.

Lane, D., Prentki, P., and Chandler, M. 1992. Use of gel retardation to analyze protein-nucleic acid interactions. Microbiol. Rev 56:509-528.

Lathrop, J. T. and Timko, M. P. 1993. Regulation by heme of mitochondrial protein transport through a conserved amino acid motif. Science 259:522-525.

Layzell, D. B., Hunt, S., Moloney, A. H. M., Fernando, S. M., and Diaz del Castillo, L. 1990. Physiological, metabolic and developmental implications of O2 regulation in legume nodules. Pages 21-32 in: Nitrogen Fixation: Achievements and Objectives, P.M. Gresshoff, L.E. Roth, G. Stacey and W.E. Newton (eds.). Chapman and Hall, New York.

Lee, H.-S., Berger, D. K., and Kustu, S. 1993. Activity of purified NifA, a transcriptional activator of nitrogen fixation genes. Proc. Natl. Acad. Sci. USA. 90:2266-2270.

Leong, S. A., Ditta, G. S., and Helinski, D. H. 1982. Heme biosynthesis in Rhizobium: Identification of a cloned gene coding for δ-aminolevulinic acid synthetase from *Rhizobium meliloti*. J. Biol. Chem. 257:8724-8730.

Leong, S. A., Williams, P. H., and Ditta, G. S. 1985. Analysis of the 5'-regulatory region of the gene for δ-aminolevulinic acid synthetase from *Rhizobium meliloti*. Nucleic Acids Res. 13:5965-5976.

Liang, Y. Y., Kaminski, P. A., and Elmerich, C. 1991. Identification of a *nifA*-like regulatory gene of *Azpspirillum brasilense* Sp7 expressed under conditions of

nitrogen fixation and in the presence of air and ammonia. Mol Microbiol. 5:2735-2744.

Liang, Y. Y., de Zamaroczy, M., Arsene, F., Paquelin, A., and Elmerich, C. 1992. Regulation of nitrogen fixation in *Azospirillum brasilense* Sp7: Involvement of *nifA*, *glnA* and *glnB* gene products. FEMS Microbiol. Lett.100:113-120.

Lisser, S. and Margalit, H. 1993. Compilation of *E. coli* mRNA promoter sequences. Nucleic Acids Res. 21:1507-1516.

Lobo, A. L. and Zinder, S. H. 1992. Nitrogen fixation by methanogenic bacteria. Pages 191-211 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans (eds.). Routledge, Chapman and Hall, Inc. New York.

Lois, A. F., Ditta, G. S., and Helinski, D. R. 1993a. The oxygen sensor FixL of *Rhizobium meliloti* is a membrane protein containing four possible transmembrane segments. J. Bacteriol. 175:1103-1109.

Lois, A. F., Weinstein, M., Ditta, G. S., and Helinski, D. R. 1993b. Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. J. Biol. Chem. 268:4370-4375.

Lonetto, M., Gribskov, M., and Gross, C. A. 1992. The σ^{70} family: Sequence conservation and evolutionary relationships. J. Bacteriol. 174:3843-3849.

de Lorenzo, V., Herrero, M., Metzke, M., and Timmis, K. N. 1991. An upstream XylR- and IHF-induced nucleoprotein complex regulates the σ54-dependent Pu promoter of TOL plasmid. EMBO J. 10:1159-1167.

Lutz, S., Jacobi, A., Schlensog, V., Böhm, R., Sawers, G., and Böck, A. 1991. Molecular characterization of an operon *(hyp)* necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*. Mol. Microbiol. 5:123-135.

Maas, R. 1983. An improved colony hybridization method with significantly increased sensitivity for detection of single genes. Plasmid 10:296-298.

Magasanik, B. 1982. Genetic control of nitrogen assimilation in bacteria. Ann. Rev. Genet. 16:135-168.

Magasanik, B. and Neidhardt, F. C. 1987. Regulation of carbon and nitrogen utilization. Pages 1318-1325 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, Vol. 2. F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds.). American Society for Microbiology, Washington, D.C.

Mandon, K., Hillebrand, H., Mougel, C., Desnoues, N., Dreyfus, B., Kaminski, P. A., and Elmerich, C. 1993. Characterization of *fixK*-regulated *Azorhizobium caulinodans*

genes. Page 478 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Masepohl, B., Keuntje, B., Krey, R., and Klipp, W. 1993. Regulation of nitrogen fixation in *Rhodobacter capsulatus:* Identification of a novel mechanism of negative regulation by ammonia. Page 479 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Masepohl, B., Klipp, W., and Puhler, A. 1988. Genetic characterization and sequence analysis of the duplicated *nifA/nifB* gene region of *Rhodobacter capsulatus.* Mol. Gen. Genet. 212:27-37.

McClung, C. R., Somerville, J. E., Guerinot, M. L., and Chelm, B. K. 1987. Structure of the *Bradyrhizobium japonicum* gene *hemA* encoding 5-aminolevulinic acid synthase. Gene 54:133-139.

Meade, H. M., Long, S. R., Ruvkun, G. E., Brown, S. E., and Ausubel, F. M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. 149:114-122.

Meijer, W. G. and Tabita, F. R. 1992. Isolation and characterization of the *nifUSVW-rpoN* gene cluster from *Rhodobacter sphaeroides.* J. Bacteriol. 174:3855-3866.

Merrick, M. J. 1992. Regulation of nitrogen fixation genes in free-living and symbiotic bacteria. Pages 835-876 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans (eds.). Routledge, Chapman and Hall, Inc. New York.

Merrick, M. J. and Chambers, S. 1992. The helix-turn-helix motif of σ^{54} is involved in recognition of the -13 promoter region. J. Bacteriol. 174: 7221-7226.

Merrick, M. J. and Coppard, J. R. 1989. Mutations in genes downstream of the *rpoN* gene (encoding σ^{54}) of *Klebsiella pneumoniae* affect expression from σ^{54}-dependent promoters. Mol. Microbiol. 3:1765-1775.

Merrick, M. J. and 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. Nucleic Acids Res. 13:7607-7620.

Merrick, M., Gibbins, J., and Toukdarian, A. 1987. The nucleotide sequence of the sigma factor gene *ntrA (rpoN)* of *Azotobacter vinelandii* : Analysis of conserved sequences in NtrA proteins. Mol. Gen. Genet. 210:323-330.

Miflin, B. J. and Cullimore, J. V. 1984. Nitrogen assimilation in the legume-Rhizobium symbiosis: A joint endeavour. Pages129-178 in: Genes Involved in

Microbe-Plant Interactions. D.P.S. Verma and T. Hohn (eds.). Springer-Verlag. New York, NY.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Minchin, S. D., Austin, S., and Dixon, R. A. 1988. The role of activator binding sites in transcriptional control of the divergently transcribed *nifF* and *nifLA* promoters from *Klebsiella pneumoniae*. Mol. Microbiol. 2:433-442.

Minchin, S. D., Austin, S., and Dixon, R. A. 1989. Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NtrC is face-of-the-helix dependent and the activator stabilizes the interaction of sigma 54 -RNA polymerase with the promoter. EMBO J. 8:3491-3499.

Minton, N. P. 1984. Improved plasmid vectors for the isolation of translational *lac* gene fusions. Gene 31:269-273.

Monson, E. K., Weinstein, M., Ditta, G. S., and Helinski, D. R. 1992. The FixL protein of *Rhizobium meliloti* can be separated into a heme-binding oxygen-sensing domain and a functional C-terminal kinase domain. Proc. Natl. Acad. Sci. USA. 89:4280-4284.

Morett, E. and Buck, M. 1988. NifA-dependent *in vivo* protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. Proc. Natl. Acad. Sci. USA. 85:9401-9405.

Morett, E. and Buck, M. 1989. *In vivo* studies on the interaction of RNA polymerase-σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti nifH* promoters: The role of NifA in the formation of an open promoter complex. J. Mol. Biol. 210:65-77.

Morett, E., Cannon, W., and Buck, M. 1988. The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of *nif* promoters and can be separated from the positive control function of NifA. Nucleic Acids Res. 16:11469-11488.

Morrett, E., Olvera, L., and Hennecke, H. 1993. Overlapping promoters for two polymerase holoenzymes regulate transcription of *nifA* in *Bradyrhizobium japonicum*. Page 483 in: New Horizons in Nitrogen Fixation, R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Mullin, D. A., and Newton, A. 1989. *Ntr*-like promoters and upstream regulatory sequence *ftr* are required for transcription of a developmentally regulated *Caulobacter crescentus* flagellar gene. J. Bacteriol. 171:3218-3227.

Nadler, K. D. and Avissar, Y. J. 1977. Heme synthesis in soybean root nodules. Plant Physiol. 60:433-436.

Nees, D. W., Stein, P. A., and Ludwig, R. A. 1988. The *Azorhizobium caulinodans nifA* gene: identification of upstream-activating sequences including a new element, the 'anaerobox'. Nucleic Acids Res. 16:9839-9853.

Newton, W. E. 1993. Nitrogenases: distribution, composition, structure and function. Pages 5-18 in: New Horizons in Nitrogen Fixation, R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Ninfa, A. J., Gottlin Ninfa, E., Lupas, A. N., Stock, A., Magasanik, B., and Stock, J. 1988. Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the *ntr* regulon: Evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proc. Natl. Acad. Sci. USA. 85:5492-5496.

Ninfa, A. J. and 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. Natl. Acad. Sci. USA 83:5909-5913.

Ninfa, A. J., Reitzer, L. J., and Magasanik, B. 1987. Initiation of transcription at the bacterial *glnAp2* promoter by purified *E.coli* components is facilitated by enhancers. Cell 50:1039-1046.

Nixon, B. T., Ronson, C. W., and Ausubel, F. M. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. Proc. Natl. Acad. Sci. USA. 83:7850-7854

Norel, F. and Elmerich, C. 1987. Nucleotide sequence and functional analysis of the two *nifH* copies of *Rhizobium* ORS571. J. Gen. Microbiol. 133:1563-1576.

North, A. K., Klose, K. E., Stedman, K. M., and Kustu, S. 1993. Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: The puzzle of nitrogen regulatory protein C. J. Bacteriol. 175:4267-4273.

O'Brian, M. R., Kirshbom, P. M., and Maier, R. J. 1987. Bacterial heme synthesis is required for expression of the leghemoglobin holoprotein but not the apoprotein in soybean root nodules. Proc. Natl. Acad. Sci. USA. 84:8390-8393.

O'Brian, M. R. and Maier, R. J. 1989. Molecular aspects of the energetics of nitrogen fixation in *Rhizobium*-legume symbioses. Biochim. Biophys. Acta 974:229-246.

Ow, D. W. and Ausubel, F. M. 1983. Regulation of nitrogen metabolism by *nifA* gene product in *Klebsiella pneumoniae*. Nature 301:307-313.

Ow, D. W., Xiong, Y., Gu, Q., and Shen, S. C. 1985. Mutational analysis of the *Klebsiella pneumoniae* nitrogenase promoter: sequences essential for positive control by *nifA* and *ntrC (glnG)* products. J. Bacteriol. 161:868-874.

Parkinson, J. S. and Kofoid, E. C. 1992. Communication modules in bacterial signaling proteins. Annu. Rev. Genet. 26:71-112.

Pau, R. N. 1993. Metal regulation of nitrogenases and molybdenum transport in *Azotobacter vinelandii*. Pages 117-122 in: New Horizons in Nitrogen Fixation, R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

Pawlowski, K., Gough, S. P., Kannangara, C. G., and de Bruijn, F. J. 1993. Characterization of a 5-aminolevulinic acid synthase mutant of *Azorhizobium caulinodans* ORS571. Mol. Plant-microbe Int. 6:35-44.

Pawlowski, K., Klosse, U., and de Bruijn, F. J. 1991. Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism. Mol. Gen. Genet 231:124-138.

Pawlowski, K., Ratet, P., Schell, J., and de Bruijn, F. J. 1987. Cloning and characterization of *nifA* and *ntrC* genes of the stem nodulating bacterium ORS571, the nitrogen fixing symbiont of *Sesbania rostrata* : Regulation of nitrogen fixation *(nif)* genes in the free living versus symbiotic state. Mol. Gen. Genet. 206:207-219.

Pfeifer, K., Kim, K.-S., Kogan, S., and Guarente, L. 1989. Functional dissection and sequence of yeast HAP1 activator. Cell 56:291-301.

de Philip, P., Batut, J., and Boistard, P. 1990. *Rhizobium meliloti* FixL is an oxygen sensor and regulates *R. meliloti nifA* and *fixK* genes differently in *Escherichia coli*. J. Bacteriol. 172:4255-4262.

Popham, D. L., Szeto, D., Keener, J., and Kustu, S. 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science 243:629-635.

Popham, D., Keener, J., and Kustu, S. 1991. Purification of the alternative σ factor, σ^{54}, from *Salmonella typhimurium* and characterization of σ^{54}-holoenzyme. J. Biol. Chem. 266:19510-19518.

Preisig, O., Anthamatten, D., and Hennecke, H. 1993. Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis. Proc. Natl. Acad. Sci. USA. 90:3309-3313.

Preker, P., Hübner, P., Schmehl, M., Klipp, W., and Bickle, T. A. 1992. Mapping and characterization of the promoter elements of the regulatory *nif* genes *rpoN*, *nifA1* and *nifA2* in *Rhodobacter capsulatus*. Mol. Microbiol. 6:1035-1047.

Prentki, P., Chandler, M., and Galas, D. J. 1987. *Escherichia coli* integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. EMBO J. 6:2479-2487.

Rabin, R. S., Collins, L. A., and Stewart, V. 1992. *In vivo* requirement of integration host factor for *nar* (nitrate reductase) operon expression in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA. 89:8701-8705.

Raina, R., Bageshwar, U. K., and Das, H. K. 1993. The *Azotobacter vinelandii nifL*-like gene: Nucleotide sequence analysis and regulation of expression. Mol Gen. Genet. 237:400-406.

Ratet, P., Pawlowski, K., Schell, J., and de Bruijn, F. J. 1989. The *Azorhizobium caulinodans* nitrogen-fixation regulatory gene, *nifA*, is controlled by the cellular nitrogen and oxygen status. Mol. Microbiol. 3:825-838.

Reitzer, L. J. and Magasanik, B. 1983. Isolation of the nitrogen assimilation regulator NR_I, the product of *glnG* gene of *Escherichia coli.* Proc. Natl. Acad. Sci. USA. 80:5554-5558.

Reitzer, L. J. and Magasanik, B. 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell 45:785-792.

Reitzer, L. J., Movsas, B., and Magasanik, B. 1989. Activation of *glnA* transcription by nitrogen regulator I (NR_I)-phosphate in *Escherichia coli*: Evidence for a long-range physical interaction between NR_I-phosphate and RNA polymerase. J. Bacteriol. 171:5512-5522.

Renalier, M.H., Batut, J., Ghai, J., Terzaghi, B., Gherardi, M., David, M., Garnerone, A. M., Vasse, J., Truchet, G., Huguet, T., and Boistard, P. 1987. A new symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional *fix* gene repeat and a *nod* locus. J. Bacteriol. 169:2231-2238.

Ricca, E. and Calvo, J. M. 1990. The nucleotide sequence of *leuA* from *Salmonella typhimurium*. Nucleic Acids Res. 18:1290

Richaud, P., Colbeau, A., Toussaint, B., and Vignais, P. M. 1991. Identification and sequence analysis of the $hupR_1$ gene, which encodes a response regulator of the NtrC family required for hydrogenase expression in *Rhodobacter capsulatus*. J. Bacteriol 173:5928-5932.

Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G. D., and Gold, L. 1992. Translation initiation in *Escherichia coli*: Sequences within the ribosome-binding site. Mol. Microbiol. 6:1219-1229.

Robidoux, S., Eskild, W., Kroepelin, C. F., Hansson, V., and Guerin, S. L. 1992. Salt-dependent formation of DNA/protein complexes *in vitro*, as viewed by the gel mobility shift assay. Biotechniques 13:354-358.

Romermann, D., Warrelmann, J., Bender, R. A., and Friedrich, B. 1989. An *rpoN*-like gene of *Alcaligenes eutrophus* and *Pseudomonas facilis* controls expression of diverse metabolic pathways, including hydrogen oxidation. J. Bacteriol. 171:1083-1099.

Ronson, C. W., Nixon, B. T., Albright, L. M., and Ausubel, F. M. 1987a. *Rhizobium meliloti ntrA (rpoN)* gene is required for diverse metabolic functions. J. Bacteriol. 169:2424-2431.

Ronson, C. W., Nixon, B. T., and Ausubel, F. M. 1987b. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell 49:579-581.

Ruvkun, G. B. and Ausubel, F. M. 1981. A general method for site-directed mutagenesis in prokaryotes. Nature 289:85-88.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sangwan, I. and O'Brian, M. R. 1991. Evidence for an inter-organismic heme biosynthetic pathway in symbiotic soybean root nodules. Science 251:1220-1222.

Santero, E., Hoover, T., Keener, J., and Kustu, S. 1989. *In vitro* activity of the nitrogen fixation regulatory protein NifA. Proc. Natl. Acad. Sci. USA. 86:7346-7350.

Santero, E., Hoover, T. R., North, A. K., Berger, D. K., Porter, S. C., and Kustu, S. 1992. Role of integration host factor in stimulating transcription from the σ54-dependent *nifH* promoter. J. Mol. Biol. 227:602-620.

Sasse-Dwight, S. and Gralla, J. D. 1988. Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. Proc. Natl. Acad. Sci. USA. 85:8934-8938.

Sasse-Dwight, S. and Gralla, J. D. 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ^{54}. Cell 62:945-954.

Schwacha, A., and Bender, R. A. 1993. The *nac* (nitrogen assimilation control) gene from *Klebsiella aerogenes*. J. Bacteriol. 175:2107-2115.

Schmitz, G., Nikaido, K., and Ames, G. F. L. 1988. Regulation of a transport operon promoter in *Salmonella typhimurium*: Identification of sites essential for nitrogen regulation. Mol. Gen. Genet. 215:107-117.

Schroeder, W. A. 1968. Page 174 in: The Primary Structure of Proteins. H. O. Halvorson, H. L. Roman and E. Bell (eds.). Harper & Row, New York.

Shatters, R. G., Somerville, J. E., and Kahn, M. L. 1989. Regulation of glutamine synthetase II activity in *Rhizobium meliloti* 104A14. J. Bacteriol. 171:5087-5094.

Shaw, D. J. and Guest, J. R. 1982. Nucleotide sequence of the *fnr* gene and primary structure of the Fnr protein of *Escherichia coli*. Nucleic Acids Res. 10:6119-6130.

Shimotsu, H., Kuroda, M. I., Yanofsky, C., and Henner, D. J. 1986. Novel form of transcription attenuation regulates expression of the *Bacillus subtilis* tryptophan operon. J. Bacteriol. 166:461-471.

van Slooten, J. C. and Stanley, J. 1991. Molecular analysis of an essential gene upstream of *rpoN* in *Rhizobium* NGR234. FEMS Microbiol Lett. 82:195-202.

van Slooten, J. C., Cervantes, E., Broughton, W. J., Wong, C. H., and Stanley, J. 1990. Sequence and analysis of the *rpoN* sigma factor gene of *Rhizobium* sp. Strain NGR234, a primary coregulator of symbiosis. J. Bacteriol. 172:5563-5574.

Smith, K. M. 1975. Appendix: Electronic Absorption Spectra. Pages 871-889 in: Porphyrins and Metalloporphyrins. K.M. Smith (ed.). Elsevier Sci. Pub. Comp. Amsterdam, The Netherlands.

Somers, J. M., Amzallag, A., and Middleton, R. B. 1973. Genetic fine structure of the leucine operon of *Escherichia coli* K-12. J. Bacteriol. 113:1268-1272.

Somerville, J. E. and Chelm, B. K. 1988. Pages 111-112 in: Molecular Genetics of Plant-Microbe Interactions. R. Palacios and D.P.S. Verma (eds.). American Phytopathological Society Press, St. Paul, Minnesota.

Souillard, N. and Sibold, L. 1989. Primary structure, functional organization and expression of nitrogenase structural genes of the thermophilic archaebacterium *Methanococcus thermolithotrophicus.* Mol. Microbiol. 3:541-551.

Spiro, S. and Guest, J. R. 1990. Fnr and its role in oxygen-regulated gene expression in *Escherichia coli.* FEMS Microbiol. Rev. 75:399-428.

Stanley, J., van Slooten, J., Dowling, D. N., Finan, T., and Broughton, W. J. 1989. Molecular cloning of the *ntrA* gene of the broad host-range *Rhizobium* sp. NGR234 and phenotypes of site-directed mutagenesis. Mol. Gen. Genet. 217:528-532.

Stenzel, T. T., Patel, P., and Bastia, D. 1987. The integration host factor of *Escherichia coli* binds to bent DNA at the origen of replication of the plasmid pSC101. Cell 49:709-717.

Szeto, W. W., Nixon, T., Ronson, C. W., and Ausubel, F. M. 1987. Identification and characterization of the *Rhizobium meliloti ntrC* gene: *R. meliloti* has seperate regulatory pathways for activating nitrogen fixation genes in free-living and symbiotic cells. J. Bacteriol. 169:1423-1432.

Thöny, B., Anthamatten, D., and Hennecke, H. 1989. Dual control of the *Bradyrhizobium japonicum*symbiotic nitrogen fixation regulatory operon *fixRnifA*: Analysis of *cis*- and *trans*-acting elements. J. Bacteriol. 171:4162-4169.

Thöny, B., Fischer, H. M., Anthamatten, D., Bruderer, T., and Hennecke, H. 1987. The symbiotic nitrogen fixation regulatory operon *(fixRnifA)* of *Bradyrhizobium japonicum* is expressed aerobically and is subject to a novel, *nifA*-independent type of activation. Nucleic Acid Res. 15:8479-8499.

Thöny, B. and Hennecke, H. 1989. The -24/-12 promoter comes of age. FEMS Microbiol. Rev. 63:341-358.

Tjepkema, J. D. and Yocum, C. S. 1974. Measurement of oxygen partial pressure within soybean nodules by oxygen microelectrodes. Planta 119:351-360.

Totten, P. A., Lara, J. C., and Lory, S. 1990. The *rpoN* gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. J. Bacteriol. 172:389-396.

Toukdarian, A. and Kennedy, C. 1986. Regulation of nitrogen metabolism in *Azotobacter vinelandii*: isolation of *ntr* and *glnA* genes and construction of *ntr* mutants. EMBO J. 5:399-407.

Turner, N. E., Robinson, S. J., and Haselkorn, R. 1983. Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature 306:337-342.

Umbarger, H. E. 1987. Biosynthesis of the branched-chain amino acids. Pages 352-367 in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, Vol. 1. F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds.). American Society for Microbiology, Washington, D.C.

Vieira, J. and Messing, J. 1987. Production of single-stranded plasmid DNA. Meth. Enzymol. 153:3-11.

Virts, E. L., Stanfield, S. W., Helinski, D. R., and Ditta, G. S. 1988. Common regulatory elements control symbiotic and microaerobic induction of *nifA* in *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA. 85:3062-3065.

Vujaklija, D., Horinouchi, S., and Beppu, T. 1993. Detection of an A-factor-responsive protein that binds to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. J. Bacteriol. 175:2652-2661.

Waelkens, F., Foglia, A., Morel, J.-B., Fourment, J., Batut, J., and Boistard, P. 1992. Molecular genetic analysis of the *Rhizobium meliloti fixK* promoter: Identification of sequences involved in positive and negative regulation. Mol. Microbiol. 6:1447-1456.

Walker, J. E., Saraste, M., Runswick, M. J., and Gray, N. J. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO. J. 8:945-951.

Wang, S-Z., Chen, J-S., and Johnson, J. L. 1988. The presence of five *nifH*-like sequences in *Clostridium pasteurianum*: Sequence divergence and transcription properties. Nucleic Acids Res. 16:439-454.

Wang, S-Z., Dean, D. R., Chen, J-S., and Johnson, J. L. 1991. The N-terminal and C-terminal portions of NifV are encoded by two different genes in *Clostridium pasteurianum*. J. Bacteriol. 173:3041-3046.

Ward, J. B. and Zahler, S. A. 1973. Genetic studies of leucine biosynthesis in *Bacillus subtilis*. J. Bacteriol. 116:719-726.

Warrelmann, J., Eitinger, M., Schwartz, E., Romermann, D., and Friedrich, B. 1992. Nucleotide sequence of the *rpoN (hno)* gene region of *Alcaligenes eutrophus*: Evidence for a conserved gene cluster. Arch. Microbiol. 158:107-114.

Wedel, A., Weiss, D. S., Popham, D., Droge, P., and Kustu, S. 1990. A bacterial enhancer functions to tether a transcriptional activator near a promoter. Science 248:486-490.

Weiss, D. S., Batut, J., Klose, K. E., Keener, J., and Kustu, S. 1991. The phosphorylated form of the enhancer-binding protein NtrC has an ATPase activity that is essential for activation of transcription. Cell 67:155-167.

Wessler, S. R. and Calvo, J. M. 1981. Control of *leu* operon expression in *Escherichia coli* by a transcription attenuation mechanism. J. Mol. Biol. 149:579-597.

Whitehall, S., Austin, S., and Dixon, R. 1992. DNA supercoiling response of the σ^{54}-dependent *Klebsiella pneumoniae nifL* promoter *in vitro*. J. Mol. Biol. 225:591-607.

van de Wiel, C., Scheres, B., Franssen, H., van Lierop, M.-J., van Lammeren, A., van Kammen, A., and Bisseling, T. 1990. The early nodulin transcript ENOD2 is located in the nodule specific parenchyma (inner cortex) of pea and soybena root nodules. EMBO J. 9:1-7.

Witty, J. F., Minchin, F. R., Skot, L., and Sheehy, J. E. 1986. Nitrogen fixation and oxygen in legume nodules. Oxford Surv. Plant Mol. Cell. Biol. 3:275-314.

Wong, P. K., Popham, D., Keener, J., and Kustu, S. 1987. *In vitro* transcription of the nitrogen fixation regulatory operon *nifLA* of *Klebsiella pneumoniae*. J. Bacteriol. 169:2876-2880.

Wu, Z.-L., Charles, T. C., Wang, H., and Nester, E. W. 1992. The *ntrA* gene of *Agrobacterium tumefaciens*: Identification, cloning, and phenotype of a site-directed mutant. J. Bacteriol. 174:2720-2723.

Xu, K., Delling, J., and Elliott, T. 1992. The genes required for heme synthesis in *Salmonella typhimurium* include those encoding alternative functions for aerobic and anaerobic coproporphyrinogen oxidation. J. Bacteriol. 174:3953-3963.

Young, J. P. W. 1992. Phylogenetic classification of nitrogen-fixing organisms. Pages 43-86 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans (eds.). Routledge, Chapman and Hall, Inc. New York.

Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K., Mizobuchi, K., and Nakata, A. 1992. Systematic sequencing of the *Escherichia coli* genome: analysis of the 0 - 2.4 min region. Nucleic Acids Res. 20:3305-3308.

de Zamaroczy, M., Liang, Y. Y., Kaminski, A., Arsene, F., and Elmerich, C. 1993a. Regulation of *nifA* synthesis and activity in *Azospirillum brasilense* Sp7. Pages 423-427 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora and W.E. Newton (eds.). Kluwer Academic Publishers, The Netherlands.

de Zamaroczy, M., Paquelin, A., and Elmerich, C. 1993b. Functional organisation of the *glnB-glnA* cluster of *Azospirillum brasilense*. J. Bacteriol. 175:2507-2515.

Zhang, L., Bermingham-McDonogh, O., Turcotte, B., and Guarente, L. Antibody-promoted dimerization bypasses the regulation of DNA binding by the heme domain of the yeast transcriptional activator HAP1. Proc. Natl. Acad. Sci. USA. 90:2851-2855.

SAMENVATTING

Regulatie van de expressie van stikstof bindende genen in de bacterie *Azorhizobium caulinodans* ORS571

De tropische bodembacterie, *Azorhizobium caulinodans* ORS571, heeft een unike plaats in de groep van stikstof bindende bacterien. *A. caulinodans* heeft de uitzonderlijke eigenschap dat het zowel vrij levend als in samenwerkingsverband (symbiose) met een tropische vlinderbloemige plant, Brem (*Sesbania rostrata*) atmosferische stikstof kan omzetten in ammonium.

De stikstof wordt omgezet met behulp van het enzym nitrogenase. Deze reactie kost zeer veel energie en het enzym is zeer zuurstof gevoelig. Daarom binden vrij levende *A. caulinodans* cellen alleen atmosferische stikstof indien zuurstof- en gebonden stikstof concentraties laag zijn. De gevormde ammonium wordt gemaakt voor eigen gebruik, vergelijkbaar met de diazotrofe (diazo=stikstof; troof=eten) groei bij *Klebsiella pneumoniae*.

A. caulinodans kan in symbiose met brem, zowel wortel- als stengelknollen induceren. In deze gespecialiseerde structuren op de plant, differentieert de bacterie in een stikstof fixerende machine (bacteroide). In dit geval geven de bacteroiden de ammonia aan de plant, terwijl de plant de bacteroiden van een photosynthetisch verkregen koolstof bron (voedsel) voorziet.

De moleculaire regulatie van de genen, die betrokken zijn bij de biologische stikstof binding, is gedurende deze twee levensvormen duidelijk verschillend. Voor vrij levende stikstof bindende bacterien, zijn zuurstof en stikstof limitatie de belangrijkste signalen voor de activering van het stikstof bindings apparaat. Voor symbiotische bacterien schijnt lage zuurstof concentratie de balangrijkste factor te zijn voor de activering van de stikstof bindende genen (het nitrogenase, het electronen-transport systeem, enzymen voor de cofactor synthese, enz). Deze lage zuurstof concentratie wordt door de plant gecreert door middel van de productie van een speciale cellaag rondom de knol. Omdat de bacteroiden niet totaal zonder zuurstof kunnen leven, produceert de plant een zuurstof-bindend-eiwit het zogenaamde leg(=legume)-hemoglobine, dat de zuurstof concentratie in de cel precies reguleert.

Uit onderzoekingen aan verschillende stikstof binders is gebleken dat het stikstof bindings apparaat wordt geactiveerd door een specifiek eiwit, NifA. NifA is een transcriptionele activator dat samen met een speciale sigmafactor, NtrA (RpoN,

sigma54) de promoters (aan- en uitschakel eenheid) van de stikstof bindings genen activeert.

Hoofdstuk 2 geeft een algemeen overzicht hoe NifA en NtrA dit doen, en vervolgens hoe de expressie en activiteit van NifA zelf is geregeld in diverse micro-organismen.

In hoofdstuk 3 wordt door middel van base-paar veranderingen, de functie van een specifiek DNA motief in de *A. caulinodans nifA* promoter onderzocht. Dit motief is homoloog aan de bindingsplaats voor de speciale sigma factor (NtrA) dat met behulp van een regulator (activator enzym, zoals NifA) de expressie kan aanschakelen. Base-paar veranderingen wijzen erop dat dit motief belangrijk is voor de activering van *nifA*, en inderdaad de binding plaats voor deze sigma factor kan zijn.

Hoofdstuk 4 beschrijft de isolatie en analyse van deze sigma factor. De base-paar- en daarvan afgeleide aminozuur volgorde van NtrA werd bepaald. Een *A. caulinodans* mutant, waarin dit eiwit niet meer wordt gemaakt, werd onderzocht op expressie van NifA en het enzym nitrogenase. Zowel nitrogenase als NifA hebben een NtrA bindings-motief in de promoter. Als verwacht werd er geen nitrogenase meer gemaakt en konden de mutanten niet meer stikstof binden. De expressie van NifA zelf was verrasend niet aangetast. Dit leidt tot de suggestie dat er waarschijnlijk meer van deze sigma factoren aanwezig zijn in *A. caulinodans*, een die specifiek het motief in het *nifA* gen herkend.

In Hoofdstuk 5 wordt de binding van eiwitten in de cel aan de DNA regulator region van *nifA* bestudeerd. Door middel van verschil in electroforese snelheid van ongebonden en gebonden DNA moleculen werd een eiwit (complex) geidentificeerd dat specifiek in het *nifA* promoter gebied bind. Welk gen voor dit of deze eiwiten coderen, en welk motief op het DNA wordt herkend is (nog) niet bekend.

Hoofdstuk 6 beschrijft de isolatie en karacterisatie van een rood, fluorescent pigment dat uitgescheden werd door *A. caulinodans* mutanten met een defect in zuurstof regulerende genen *(fixLJ* en *fixK)*, waneer deze groeide onder lage zuurstof concentratie.

Hoofdstuk 7 beschrijft de isolatie en base-paar volgorde bepaling van een *A. caulinodans* gen dat per ongeluk geisoleerd werd. Met behulp van een *E. coli* gen voor een anaerobe activator, Fnr, werd een *A. caulinodans leuA* gen gevonden, dat een rol speelt in aminozuur biosynthese. Interessant is dat de organisatie van dit gen verschilt van wat men heeft gevonden in andere bacterien en dat de regulatie van de expressie van dit gen via een speciaal mechanisme kan plaats vinden.

Tenslotte wordt in Hoofdstuk 8 het laatste bijgewerkte model gepresenteerd van de regulatie van de stikstof bindende genen in *A. caulinodans*.

Curriculum vitae

John Stigter werd op 22 juli 1961 in Vreeswijk geboren. Na het behalen van het HAVO diploma in 1979 begon hij een studie voor biochemisch analist aan de Hogere en Middelbare Laboratoriumschool (HBO-B) te Oss.
Zijn stage en afstudeeropdracht werd uitgevoerd aan de Katholieke Universiteit te Nijmegen, op de afdeling Biochemie (Prof. dr. J.J. de Pont en Dr. W. Peters). Onderzoek werd verricht naar de rol van phospholipiden in het membraan-gebonden enzym-complex Na^+K^+-ATPase. In juni 1982 werd het HBO-B diploma behaald.
Hierna was hij tot september 1983 in militaire dienst.
In september 1983 begon hij met zijn studie Moleculaire Wetenschappen (T33; chemisch-biologische orientatie) aan de toenmalige Landbouwhogeschool te Wageningen.
Gedurende zijn doctoraalonderzoek (van januari tot september 1987) bij de vakgroep Moleculaire Biologie (Prof. dr. A. van Kammen, Dr. R.C. van den Bos en Dr. P. Roelvink) kwam hij voor het eerst in contact met biologische stikstofbinding en NifA.
Van september 1987 tot juni 1988 werkte hij in het kader van zijn praktijktijd bij de vakgroep Moleculaire Biologie van het Massachusetts General Hospital en de vakgroep Genetica van de Harvard Medical School in Boston, USA. Onder leiding van Prof. dr. F.M. Ausubel en Dr. L. Albright werkte hij aan de regulatie van *nifA* in *R. meliloti* en *dctA* in *R. leguminosarum*. In september 1988 werd het doctoraalexamen Moleculaire Wetenschappen behaald.
Gefacineerd door de moleculaire en genetische regulatie van stikstofbinding werd besloten om een promotieonderzoek te beginnen in de Rhizobium-groep, aan het Max-Planck Institut in Keulen, Duitsland, onder leiding van Dr. F.J. de Bruijn. Toen in mei 1990 Dr. F.J. de Bruijn verhuisde naar het DOE-Plant Research Laboratory, Michigan State University in East-Lansing, Michigan, USA, werd het onderzoek voortgezet in het laboratorium in East-Lansing.
Het onderzoek tijdens deze periode, beschreven in dit proefschrift, omvatte de analyse van de genetische regulatie van stikstof bindingsgenen in *Azorhizobium caulinodans* ORS571.
Vanaf 1 september 1993 is hij als post-doctoraal onderzoeker werkzaam aan de University of Michigan in Ann Arbor, Michigan, USA, bij de vakgroep Biologie (Prof. dr. J. Adams) waar hij de moleculaire evolutie en genetische veranderingen in *Escherichia coli* in een constante omgeving en selectiedruk, bestudeerd.