Studies on genes encoding regulators of nitrogen metabolism in *Thiobacillus ferrooxidans* ATCC 33020

David Kenneth Berger

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Faculty of Science,
University of Cape Town.

CAPE TOWN
AUGUST 1990

CERTIFICATION OF SUPERVISORS

In terms of paragraph 9 of "General regulations for the degree of Ph.D." we, as supervisors of the candidate D.K. Berger, certify that we approve of the incorporation in this thesis of material that has already been published or submitted for publication.

Signed

Professor D. E. Rawlings

Thiobacillus Research Unit

University of Cape Town

Signed

Professor D. R. Woods

Deputy Vice-Chancellor (Research)

University of Cape Town

AKNOWLEDGEMENTS

My experience at the University of Cape Town has been greatly enriched on account of my supervisors, Professor D. E. Rawlings and Professor D. R. Woods; in particular, thanks, Doug, for your clear, consistent guidance and sound judgement, and your communication of a spirit of inquiry; and thanks, "Prof", for the opportunity to be swept along by your excitement and enthusiasm.

Many thanks to fellow members of the Thiobacillus Lab, past and present, for sharing the joys and sorrows of cracking this recalcitrant "nut". I am greatly indebted to Anne-Marie Clennell and Carol Kilkenny for their willing and excellent technical assistance. Thanks, also, to Greg Blatch for your kindred spirit and the relaxing lunchtime hours spent on the squash-racquets court.

A tribute is also due to the inspirational environment of UCT on the slopes of Table Mountain, where I gained relaxation and perspective on many a run through the fynbos.

My wife, Eldie, deserves special thanks for surviving another simultaneous PhD: your solid support and encouragement has made life from Eppendorf to VAX both fun and fulfilling. Your help greatly increased the efficiency of transcription of this dissertation, especially during the stationary phase of its development.

I wish to acknowledge financial assisitance from the Council for Scientific and Industrial Research and the Foundation for Research Development.

In memory of my father, John Berger, who would have been intrigued by the mysteries of the DNA molecule.....

Dedicated also to my mother, Leslie Berger, for her continued support, enthusiasm and optimism.

TABLE OF CONTENTS

ABSTRACT

The *Thiobacillus ferrooxidans* ATCC 33020 *ntrA* and *ntrC* genes were isolated from a *T. ferrooxidans* pHC79 cosmid gene bank, and cloned in *Escherichia coli*. Novel strategies were used for the isolation of each gene, since standard methods, such as hybridization with an *Azotobacter vinelandii ntrA* gene probe, or complementation of *E. coli ntrA* or *ntrC* mutants followed by selection on arginine as a sole nitrogen source, had been unsuccessful.

A method based on the inability of *E. coli ntrA* mutants to produce gas via the formate degradation pathway was developed for the isolation of the *T. ferrooxidans ntrA* gene. A *T. ferrooxidans* cosmid gene bank was transduced into the *E. coli ntrA* strain TH1, and transductants containing the *T. ferrooxidans ntrA* gene were detected using a simple agar overlay technique on the basis of NtrA-dependent expression of the gas positive phenotype. The nucleotide sequence of a 2,8-kbp *Sal*I-*Bgl*II fragment was determined. This fragment contained three open reading frames (ORF's), including the *T. ferrooxidans ntrA* gene, which encoded a predicted translation product of 475 amino acids (aa) (calculated M_r 52972). This protein showed 51, 50, 49, 40, and 28% aa sequence identity with the NtrA proteins from *Klebsiella pneumoniae*, *Pseudomonas putida*, *A. vinelandii*, *Rhizobium meliloti*, and *Rhodobacter capsulatus* (NifR4 protein), respectively. An ORF coding for a protein of 241 aa (calculated M_r 27023) was situated 12-bp upstream of the *T. ferrooxidans ntrA* gene. This protein had 57% aa identity with the product of the ORF1 located upstream of the *R. meliloti ntrA* gene. Downstream of the *T. ferrooxidans ntrA* gene the front end of an ORF, called ORF3, was identified. The predicted 78 N-terminal aa encoded by ORF3 showed 38, 38, 29, and 20% aa identity with conserved ORF's immediately downstream of the *A. vinelandii*, *P. putida*, *K. pneumoniae*, and *R. meliloti ntrA* genes, respectively.

An *E. coli ntrA* strain, containing the *T. ferrooxidans ntrA* gene without the flanking ORF's, produced gas and reduced benzylviologen, an artificial electron acceptor, when grown anaerobically on formate. These phenotypes, which are characteristic of the *E. coli* formate hydrogenlyase pathway, were repressed when formate was

replaced by nitrate in the growth media. NtrA-dependent expression from either the *E. coli fdhF*, the *T. ferrooxidans nifH*, or the *K. pneumoniae nifH* promoter in the presence of the *T. ferrooxidans ntrA* gene was demonstrated using *lacZ* fusions in *E. coli*. Expression of an *fdhF-lacZ* fusion in *E. coli* TH1 cells was increased 7-fold above basal levels in the presence of the *T. ferrooxidans ntrA* gene, and was also repressed when formate was replaced by nitrate. Expression of either a *T. ferrooxidans nifH-lacZ* or a *K. pneumoniae nifH-lacZ* fusion in *E. coli* TH1 cells (containing a constitutively expressed *K. pneumoniae nifA* gene) was increased 50-fold in the presence of the *T. ferrooxidans ntrA* gene. In the presence of either the *E. coli ntrA* gene (chromosomal copy) or the *K. pneumoniae ntrA* gene (cloned on a plasmid vector) a similar increase in expression from the *T. ferrooxidans nifH-lacZ* fusion was observed (in the presence of the *K. pneumoniae* NifA). However, levels of expression of the *K. pneumoniae nifH-lacZ* fusion (in the presence of the *K. pneumoniae* NifA) were increased 400-fold in the presence of either the *E. coli* or the *K. pneumoniae ntrA*.

NtrC-mediated expression of the *T. ferrooxidans nifH-lacZ* fusion was shown in *E. coli* cells grown under nitrogen-limiting conditions. An agar plate assay was developed for the detection of *E. coli ntrC* cells containing the *T. ferrooxidans nifH-lacZ* fusion which showed increased levels of β-galactosidase activity. The *T. ferrooxidans* cosmid bank was transduced into the *E. coli ntrC* (*T. ferrooxidans nifH-lacZ* fusion plasmid) cells, and colonies were grown on a nitrogen-limited minimal medium. Cells which showed increased β-galactosidase activity by virtue of formation of a strong yellow colour after flooding with *o*-nitrophenyl-β-D-galactopyranoside (ONPG) were selected. A 5,2-kbp *Kpn*I-*Hind*III subclone from one of the cosmids was shown to contain the *T. ferrooxidans ntrC* gene. The gene was localized by nucleotide sequence analysis of an internal 486-bp *Cla*I-*BamH*I fragment. The predicted translation product of this section of the *T. ferrooxidans ntrC* gene showed homology to the N-terminal region of NtrC proteins from other bacteria. NtrC proteins from the enteric bacteria *K. pneumoniae* (76%) and *E. coli* (72%) showed the greatest aa sequence similarity. Expression in *E. coli* of the *T. ferrooxidans nifH-lacZ* fusion in the presence of the *T. ferrooxidans ntrC* gene was regulated by nitrogen in both solid and liquid media.

ABBREVIATIONS

A	adenosine
aa	amino acids
Ap	ampicillin
ATCC	American Type Culture Collection
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
C	cytosine
C-	carboxy terminal (end of a protein)
Clm	chloramphenicol
CsCl	caesium chloride
°C	degrees Celsius
d	day(s)
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	1,4-dithio-L-threitol
EDTA	ethylene-diaminetetra-acetic acid
g	gram
G	guanine
h	hour(s)
IHF	integration host factor
IPTG	isopropyl–β-D-thiogalactopyranoside
kbp	kilobase pairs
kDal	kilodaltons
LB	Luria-Bertani medium
LBA	Luria-Bertani medium containing 1,5% agar
min	minute(s)
M_r	relative molecular mass
N-	amino terminal (end of protein)
nt	nucleotides
OD_{600}	optical density at 600 nm
ONPG	*o*-nitrophenyl-ß-D-galactopyranoside
ORF	open reading frame
R	(superscript) resistance
RNA	ribonucleic acid
s	second(s)
SDS	sodium dodecyl sulfate
sp(p)	species
T	thymine
TAE	tris-acetate EDTA buffer
Tet	tetracycline
Tn	transposon
Tris	Tris(hydroxymethyl)aminomethane
U	uracil
U	units of enzyme activity
UV	ultraviolet (light)
w/v	weight/volume
XGal	5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside
α	alpha
β	beta
Δ	delta
λ	lamda
μ	micro
σ	sigma

TABLE OF FIGURES

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1

GENERAL INTRODUCTION

1.1 *Thiobacillus ferrooxidans* - the organism

Thiobacillus ferrooxidans is a bacterium which is commonly found in rock-based environments containing a variety of metal ores (Vishniac, 1974). It is a motile, Gram-negative, rod-shaped organism with approximate cell dimensions of 0,5 x 1,5 - 2 μm. *T. ferrooxidans* is aerobic, acidophilic (pH 1,3 - 3,5), and mesophilic (20 - 37°C). This autotrophic chemolithotroph uses atmospheric CO_2 as a sole source of carbon, and obtains its energy from the oxidation of either ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}) or reduced sulphur compounds to sulphuric acid (reviewed in Ingledew, 1982). *T. ferrooxidans* grows aerobically, with oxygen serving as the final electron acceptor in the oxidation process. Sulphur oxidation may also occur through a route involving an electron acceptor other than molecular oxygen (Sugio *et al.*, 1985). A hydrogen sulphide:ferric iron oxidoreductase (SFORase) has been identified and purified (Sugio *et al.*, 1987, 1988a). This enzyme plays a crucial role in this respiratory pathway where either of the metal ions - Fe^{+3}, Mo^{+6}, or Cu^{+2} - are reduced with elemental sulfur as the electron donor (Sugio *et al.*, 1988b, 1990). *T. ferrooxidans* has been shown to grow on either ammonia or nitrate as sole nitrogen sources (Tuovinen *et al.*, 1979)(section 1.6.1). Mackintosh (1978) showed that at least one strain of *T. ferrooxidans* was able to incorporate $^{15}N_2$ label into cellular material, thus demonstrating that this bacterium was able to fix atmospheric nitrogen.

1.1.1 The importance of *T. ferrooxidans* in Industrial Biomining

Bacterial leaching is playing an increasing role in the industrial extraction of metals from low grade mineral ores (Brierley, 1982; Curtin, 1983). Metals that have been extracted using biomining include copper (20% of the world market), uranium (50% of that extracted in Canada), and gold. This is not a new technological innovation, however, as the recovery of copper from the drainage water of mines was a widespread practice in the Mediterranean basin as early as 1000 BC. The large scale leaching of copper was well established by the 18th century at the Rio Tinto mines in

Spain (Brierley, 1982). The role of bacteria, as opposed to abiotic chemical reactions, in these processes was not recognized until 1947 when the presence of bacteria in acid mine drainage was reported (Colmer and Hinkle, 1947).

Leaching is a hydrometallurgical process which consists of the dissolution of an insoluble mineral ore to produce a soluble compound. Micro-organisms that are involved in the leaching process range from the extreme thermophiles, such as members of the genus *Sulfolobus*, to the moderately thermophilic *Sulfobacillus* and acidophilic members of the genus *Thiobacillus*. The most predominant bacterium in most leaching environments is *T. ferrooxidans* and this has been attributed to its tolerance for the high acidity and high concentrations of metal ions which are associated with this habitat. Furthermore, the ability of *T. ferrooxidans* to obtain energy through the oxidation of ferrous iron and its autotrophic nature may explain its ability to out-compete other members of the genus *Thiobacillus* which are only able to oxidize sulphur compounds (eg. *Thiobacillus thiooxidans*) or exhibit heterotrophy (eg. *Thiobacillus* TH1 - Brierley and Le Roux, 1977; *Thiobacillus thermosulfidooxidans* - Golovacheva and Karavaiko, 1977). The importance of mixed cultures in improving the efficiency of industrial biomining processes has been recognized. For example, enhanced bioleaching has been shown for mixed cultures of *T. ferrooxidans* and *T. thiooxidans* (Kelly *et al.*, 1979).

The recovery of gold from arsenopyrite-pyrite ores by leaching with *T. ferrooxidans* as the dominant organism has recently been scaled up from the pilot scale (Livesey-Goldblatt *et al.*, 1983) to an industrial scale (Rawlings, 1991). The gold-bearing ore is crushed and then processed through a series of aerated tanks where the bacterial oxidation takes place. This step replaces the conventional "roasting" procedure and is followed by cyanidation to produce 92-95% yields. The industrial plant was commissioned in 1986 and is currently processing 18 tonnes of crushed ore concentrate per day (Rawlings, 1991). Although this is an "open" system (Rawlings, 1991) containing a mixed population of bacteria, *T. ferrooxidans* is regarded as the most important biotic component of the process. Mutation and selection over 5 y has produced a "super" strain of *T. ferrooxidans* which has enabled an 8-fold increase in the rate of the leaching process and a 13-fold increase in

resistance to the toxic arsenic which is solubilized during the bioleaching process (Rawlings, 1991). This success highlights the importance of *T. ferrooxidans* as a bioleaching agent not only in financial terms, but also in the desirability of replacing the polluting "roasting" stage with the less environmentally damaging bioleaching procedure.

1.1.2 The chemistry of bioleaching

An understanding of the chemistry of bioleaching processes provides an insight into the importance of bacteria in industrial biomining systems. Bioleaching is carried out by two mechanisms, namely direct and indirect leaching. Direct leaching involves bacterial enzymatic oxidation of the reduced mineral or ion without entry of this substrate into the bacterium itself. The electrons released by the oxidation reaction enter the cell and are passed along the respiratory electron transport chain to generate energy-rich adenosine triphosphate (ATP). In aerobic bioleaching organisms, such as *T. ferrooxidans*, oxygen would serve as the terminal electron acceptor. An example of a direct leaching reaction carried out by *T. ferrooxidans* on chalcopyrite ($CuFeS_2$) is shown below:

(i) $$2CuFeS_2 + 8^1/_2O_2 + H_2SO_4 \text{ ---> } 2CuSO_4 + Fe_2(SO_4)_3 + H_2O$$

(Insoluble) (Soluble)

Indirect or bacterial-assisted leaching is carried out by reaction of the products of the direct leaching process with the insoluble minerals, transforming them into the oxidized, soluble form. Chemical reactions of this nature would be carried out by the powerful oxidizing agents, such as ferric iron (Fe^{+3}) and sulphuric acid, produced during the direct leaching process. In these chemical reactions ferric iron (Fe^{+3}) is reduced to ferrous iron (Fe^{+2}) which is rapidly re-oxidized by the bacteria. Fe^{+2} is stable in acidic solution, and therefore in the absence of bacteria chemical leaching mediated by Fe^{+3} would be limited by the availability of Fe^{+3}. It has been estimated that leaching bacteria such as *T. ferrooxidans* can accelerate the oxidation process a million-fold by this recycling of the Fe^{+2} to Fe^{+3} (Brierley, 1982). This regeneration of the so-called "leach liquor" (containing oxidizing agents) by bacteria is the key to the success of industrial biomining. An example of indirect leaching is the chemical

oxidation of chalcopyrite ($CuFeS_2$) by ferric sulphate ($Fe_2(SO_4)_3$) (reaction (ii)). The ferrous iron ($FeSO_4$) produced would then be rapidly reoxidized by *T. ferrooxidans* (reaction (iii)).

(ii) $$CuFeS_2 + Fe_2(SO_4)_3 \text{ -------> } CuSO_4 + FeSO_4$$
chemical

(iii) $$2FeSO_4 + H_2SO_4 + {}^1/_2\, O_2 \text{ -------> } Fe_2(SO_4)_3 + H_2O$$
biological

1.1.3 Studies on the molecular biology of *T. ferrooxidans*

In recent years the tools of recombinant DNA technology have been used in an effort to understand aspects of the molecular biology of *T. ferrooxidans*. These studies have been undertaken with a view towards the genetic manipulation of this bacterium for improved biomining capabilities. The development of a genetic system for *T. ferrooxidans* requires three major components: (i) a suitable plasmid vector, (ii) a selectable marker for stable maintenance of the recombinant construct in *T. ferrooxidans*, and (iii) a method of introduction of DNA into *T. ferrooxidans* cells.

1.1.3.1 The development of a genetic manipulation system for *T. ferrooxidans*

Several native plasmids have been isolated from *T. ferrooxidans* strains (Holmes *et al.*, 1983; Rawlings *et al.*, 1984). One plasmid, denoted pTF-FC2, is able to replicate in a wide variety of gram-negative bacteria, and the replication (Dorrington and Rawlings, 1989) and mobilization (Rohrer, pers. comm.) functions of pTF-FC2 are currently being studied in detail.

The use of antibiotic markers for selection of *T. ferrooxidans* strains is impractical as these compounds are inactivated at the low pH conditions required for the growth of this acidophile (Rawlings *et al.*, 1983). Metal ion tolerance is an attractive alternative, and the chromosomally encoded mercury ion resistance genes recently isolated from a strain of *T. ferrooxidans* (Shiratori *et al.*, 1989) have potential in this respect.

The major stumbling block in the development of a genetic system for *T. ferrooxidans* has been the lack of a method of introduction of DNA into *T. ferrooxidans* cells.

There are four methods which have been used for this purpose in other bacteria: transduction, transformation, conjugation, and electroporation. Transduction is not feasible as no viruses which infect *T. ferrooxidans* are known. Standard transformation techniques, including the use of *T. ferrooxidans* spheroplasts, have been unsuccessful (E. Barros, PhD Thesis, University of Cape Town, 1985). Conjugation of broad host range plasmids between *Escherichia coli* and a variety of thiobacilli has been carried out (Davidson and Summers, 1983; Kulpa *et al.*, 1983). A two-step mating procedure was employed in some experiments in an attempt to use non-iron-oxidizing thiobacilli, such as *Thiobacillus novellus* and *Thiobacillus intermedius* as a bridge between *E. coli* and *T. ferrooxidans* (R. Ramesar, PhD Thesis, University of Cape Town, 1988). However, no successful transfer into *T. ferrooxidans* has been obtained. Electroporation is at present being attempted (Rohrer, pers. comm.).

1.1.3.2 Advances in the understanding of the molecular biology of *T. ferrooxidans*

Basic studies on the molecular biology of metabolic processes in *T. ferrooxidans* are complementary to the above studies. They provide information on the regulatory signals required for gene expression in *T. ferrooxidans*, and enable the choice of source DNA from organisms which have genes that are most similar to *T. ferrooxidans* genes. This includes promoter sequences, regulatory elements, codon usage, and termination signals.

The *T. ferrooxidans* genes for the enzymes glutamine synthetase (*glnA*)(Barros *et al.*, 1985; Rawlings *et al.*, 1987) and nitrogenase (*nifHDK*)(Pretorius *et al.*, 1986; Pretorius *et al.*, 1987; Rawlings, 1988), and the *recA* gene (Ramesar *et al.*, 1988, 1989) have been cloned and the nucleotide sequences have been determined. The former two, which are involved in nitrogen metabolism, will be discussed in section 1.6.2. The product of the *T. ferrooxidans recA* gene was able to complement *E. coli recA* mutants for both the recombinase and protease (DNA repair) activities. Transcription for this activity was initiated from an adjacent plasmid vector promoter, as only weak expression was obtained in *E. coli* from the *T. ferrooxidans recA* promoter (Ramesar *et al.*, 1988, 1989).

The study of genes whose products are involved in the unusual physiology of *T. ferrooxidans*, namely the ability to oxidize iron and sulphur, is particularly important. The gene for the iron-induced copper protein, rustacyanin, has been cloned (Kulpa *et al.*, 1986), however no nucleotide sequence data is available. Another protein, which is induced on the transfer of *T. ferrooxidans* cells from sulfur medium to ferrous iron medium, was shown to be membrane bound (Mjoli and Kulpa, 1988). The approach of reverse genetics has been used in an attempt to clone the gene which codes for this protein. The aa sequence of a peptide derived from the purified protein was determined. This sequence was used together with the codon usage table derived from the nucleotide sequences of five *T. ferrooxidans* genes to design an oligonucleotide probe. Characterization of a positive clone identified from screening of a *T. ferrooxidans* cosmid gene bank is in progress (Mjoli, pers. comm.).

Yates and Holmes (1987) have detected two families of repeated sequences in strains of *T. ferrooxidans*. Nucleotide sequence analysis of one of these families has revealed similarities with the structure of bacterial IS elements (Yates *et al.*, 1988). Evidence that these may be involved in transposition has come from the observation that these strains of *T. ferrooxidans* are able to undergo "phenotypic switching" of colony morphology (Schrader and Holmes, 1988). This phenomenon is accompanied by the reversible loss of the ability to oxidize iron and therefore it is tempting to speculate that transposition or a site-specific inversion in *T. ferrooxidans* is a mechanism for enhancing its ability to adapt to changing environmental conditions (Rawlings *et al.*, 1991). It is important to note that no representatives of the two families of repeated sequences mentioned above have been identified in the *T. ferrooxidans* strain, ATCC 33020, used in this PhD study (Yates and Holmes, 1987).

1.2 Assimilation of nitrogen by bacteria

Nitrogen is the basic component required for the synthesis of cellular macromolecules such as amino acids, purines, pyrimidines, amino sugars, and metabolites such as NAD and p-aminobenzoate. The most readily available source of nitrogen in most environments is ammonia, and for enteric bacteria such as *E. coli* and *Salmonella typhimurium* this is the preferred source of nitrogen (Reitzer and Magasanik, 1987).

In those environments where ammonia is not available many bacteria are able to take up nitrogen as inorganic compounds such as nitrate, nitrite, urea, or diatomic atmospheric nitrogen. Many organic compounds such as arginine, ornithine, histidine, and proline are used by bacteria as sole sources of nitrogen (Tyler, 1978). There also exist metabolic pathways within bacteria for the catabolism of nitrogen containing compounds whereby nitrogen may be recycled within the cell. The central role of ammonia appears to be paramount and most cellular nitrogen is derived either from direct incorporation from ammonia or through the metabolic intermediates glutamine and glutamate which require ammonia for their synthesis.

1.2.1 Ammonia assimilation in enteric bacteria.

The cycling of nitrogen through ammonia, glutamine, and glutamate has been studied most extensively in the Gram-negative enteric bacteria *E. coli* and *S. typhimurium* (reviewed in: Tyler, 1978; Magasanik, 1982; Magasanik and Neidhardt, 1987; Reitzer and Magasanik, 1987; Magasanik, 1988). Three key enzymes are believed to play a pivotal role in this process, namely glutamine synthetase (GS, EC 6.3.1.2), glutamate synthase (GOGAT, glutamine-oxoglutarate amido transferase, EC 1.4.1.13) and glutamate dehydrogenase (GDH, EC 1.4.1.4.). The reactions catalyzed by these enzymes are shown below:

$$\text{a) } NH_3 + \text{glutamate} + ATP \xrightarrow{GS} \text{glutamine} + ADP + P_i$$

$$\text{b) glutamine} + \alpha\text{-ketoglutarate} + NADPH \xrightarrow{GOGAT} 2\ \text{glutamate} + NADP^+$$

$$\text{c) } NH_3 + \alpha\text{-ketoglutarate} + NADPH \xrightarrow{GDH} \text{glutamate} + NADP^+$$

GS is regarded as the most important enzyme in the cycling of nitrogen in enteric bacteria as this is the only known biosynthetic route for the synthesis of glutamine. This is in contrast to members of the family *Rhizobiaceae*, such as *Rhizobium meliloti*, *Bradyrhizobium japonicum*, and *Agrobacterium tumefaciens*, which contain a second glutamine synthetase enzyme, GSII (Darrow *et al.*, 1981). This enzyme is structurally related to GS of higher plants (Carlson and Chelm, 1986). A novel third locus, *glnT*, which appears to consist of a complex operon and is clearly capable of directing the

synthesis of glutamine in an *E. coli* background, has been identified from *A. tumefaciens* (Rossbach *et al.*, 1988) and *R. meliloti* (de Bruijn *et al.*, 1989).

GS in the enteric bacteria is not only responsible for glutamine synthesis, but is also required for the assimilation of ammonia under conditions of ammonia limitation. Consequently, both the catalytic activity and synthesis of GS are highly regulated (Tyler, 1978; Magasanik, 1982). All of these regulatory systems function in response to levels of available nitrogen, which is sensed with respect to the intracellular ratio of glutamine to α-ketoglutarate. A low ratio signals nitrogen limiting conditions and results in the positive regulation of GS activity and expression. Regulation at each level is linked in a complex "bicyclic cascade" (Chock *et al.*, 1985; Rhee *et al.*, 1985; Reitzer and Magasanik, 1987; Magasanik, 1988). The P_{II} protein (encoded by the *glnB* gene) forms a common link between the regulation of the activity of GS by adenylylation and the regulation of expression of the *glnA* gene encoding GS by the products of the *ntrB* (*glnL*) and *ntrC* (*glnG*) genes (Ntr regulatory system) (Garcia and Rhee, 1983; Bueno *et al.*, 1985; Son and Rhee, 1987).

The regulation of catalytic activity by adenylylation will be discussed here, while the Ntr regulatory system will be discussed in section 1.2.2. Bacterial GS consists of a dodecamer of identical subunits. Regulation of catalytic activity is carried out by the covalent addition of an AMP group to a tyrosine residue in each subunit. This adenylylation inactivates each individual subunit so the overall activity of the enzyme is inversely related to the number of adenylylated subunits. This process involves a cascade system whereby the nitrogen status of the cell is sensed by a bifunctional polypeptide which is able to add or remove uridylyl groups from the P_{II} protein, which in turn is able to dictate the direction of adenylylation or deadenylylation of each GS subunit by an adenylyltransferase (ATase)(Garcia and Rhee, 1983). Consequently, under conditions of nitrogen excess (high glutamine/α-ketoglutarate), P_{II} is deuridylylated and promotes adenylylation of GS. Conversely, under nitrogen limiting conditions (low glutamine/α-ketoglutarate), P_{II} becomes uridylylated and directs deadenylylation of GS by ATase resulting in GS with increased catalytic activity (Stadtman *et al.*, 1980; reviewed in Reitzer and Magasanik, 1987).

A second level of regulation is through cumulative feedback inhibition by the products of glutamine metabolism. The adenylylated form of GS is inhibited partially by each of the following compounds: L-alanine, glycine, histidine, tryptophan, CTP, AMP, carbamyl phosphate, and glucosamine-6-phosphate. This inhibitory effect is cumulative and together these compounds inhibit GS completely (Stadtman and Ginsberg, 1974).

Under conditions of ammonia limitation (< 0,1 mM) the synthesis of glutamate is carried out by a two step process involving GS and GOGAT. Ammonia is incorporated into glutamine by GS followed by the synthesis of glutamate from glutamine by GOGAT. The removal of glutamine by GOGAT under nitrogen limiting conditions is important otherwise the cellular machinery would respond to the accumulation of glutamine and signal inappropriately the repression of Ntr regulated systems (Reitzer and Magasanik, 1987). The role of GDH appears to be the synthesis of glutamate under conditions of excess ammonia in the growth medium (> 1 mM). The absence of ammonia in the growth medium is considered nitrogen limiting despite the presence of an alternative organic nitrogen source such as arginine or histidine. Growth under these conditions is slower than in the presence of ammonia and the levels of GS are high. This indicates that the growth limiting factor is the rate of ammonia generation from the alternative nitrogen source and of its subsequent assimilation (Reitzer and Magasanik, 1987). The degradative enzymes and transport systems for these alternative nitrogen sources are subject to Ntr control (Shaibe *et al.*, 1985; Kustu *et al.*, 1979) and therefore show the same patterns of regulation as the *glnA* gene.

1.2.2 Ntr system of transcriptional regulation

1.2.2.1 The *glnALG* (*glnAntrBC*) operon.

The *glnALG* operon is at the centre of the global Ntr nitrogen regulation system and the expression of this operon has been studied extensively in the enteric bacteria *E. coli* (Chen *et al.*, 1982; MacNeil *et al.*, 1982b; Pahel *et al.*, 1982; Ueno-Nishio *et al.*, 1983, 1984; Reitzer and Magasanik, 1985, 1986), and *S. typhimurium* (Garcia *et al.*, 1977; Wei and Kustu, 1981; reviewed in Kustu *et al.*, 1986; Keener *et al.*, 1987), as well as in *Klebsiella pneumoniae* (de Bruijn and Ausubel, 1981; Espin

et al., 1982; Alvarez-Morales *et al.*, 1984; Macfarlane and Merrick, 1985, 1987). The pattern of regulation appears to be similar in all three organisms. *GlnA* is the structural gene for GS while the products of *glnL* (*ntrB*) and *glnG* (*ntrC*) are the regulators NR_{II} (NtrB) and NR_I (NtrC), respectively. The operon contains three promoters: *glnAp1*, *glnAp2*, and *glnLp* (Fig. 1.1).

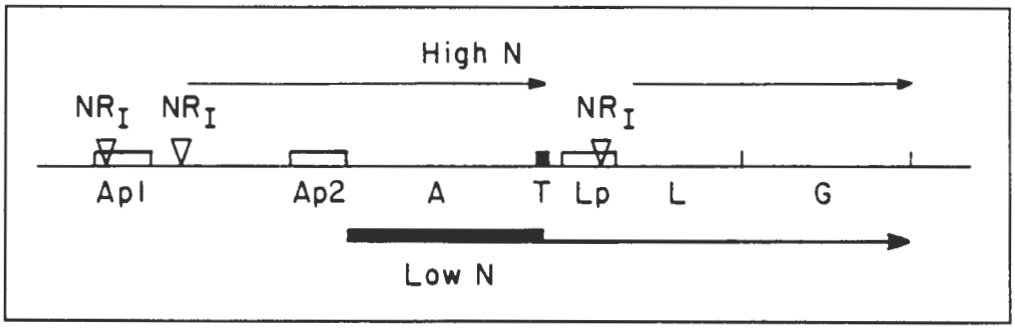


Fig. 1.1. Transcription of the gln*ALG* operon under nitrogen limiting and nitrogen excess conditions. □, promoters; ▽, NR_I-binding sites; ■, terminator (after Magasanik, 1988).

GlnAp1 and *glnLp* show sequence similarity to the canonical -35, -10 *E. coli* promoter sequences and are therefore recognized by σ70-RNA polymerase (Eσ70), while *glnAp2* has a promoter which is quite different and this is recognized by a form of RNA polymerase bound to a specialized sigma factor denoted σ54 (Eσ54). σ54 is encoded by the *ntrA* (*glnF*, *rpoN*) gene. The expression of the operon is autogenously regulated by NR_I. In cells growing in nitrogen excess, the transcription is initiated at *glnAp1* and *glnLp*. This serves to maintain a basal level of GS, NR_I, and NR_{II}. Initiation of transcription at *glnAp1* is further stimulated under conditions of carbon limitation by the catabolite gene-activator protein (CAP) and cyclic AMP (cAMP) (Magasanik and Niedhardt, 1987). Approximately three out of every four transcripts from *glnAp1* terminate at a *rho*-independent terminator upstream of *glnLp*. Concomitantly, transcription is repressed by NR_I. Binding sites for NR_I, characterized by the nucleotide sequence $GCACN_5TGGTGC$, overlap the -35 region of *glnAp1* and the -10 region of *glnLp*. There are five NR_I binding sites upstream of *glnAp2* and two of these, NR-1 and NR-2, which overlap *glnAp1* are bound with high affinity by NR_I. Thus, under nitrogen excess, NR_I limits the synthesis of both GS and itself. This results in a concentration of about five molecules of NR_I per cell.

Nitrogen deprivation results immediately in the activation of the initiation of transcription from *glnAp2*. This is possible as Eσ54 is bound to *glnAp2* in an inactive closed promoter complex irrespective of nitrogen-status (Reitzer *et al.*, 1987). Transition of this closed promoter complex to an active open promoter complex is stimulated by the activator form of NR_I. The activation of NR_I by NR_{II}-mediated phosphorylation comprises the second component of the "bicyclic cascade" regulation of GS in response to nitrogen status. The ratio of glutamine to α-ketoglutarate is sensed by the uridylylation system which controls the modification of the P_{II} protein which not only mediates the adenylylation status of GS as was discussed in section 1.2.1, but also mediates the phosphorylation status of NR_I through interaction with NR_{II}. The scheme for this control is outlined in Fig. 1.2.

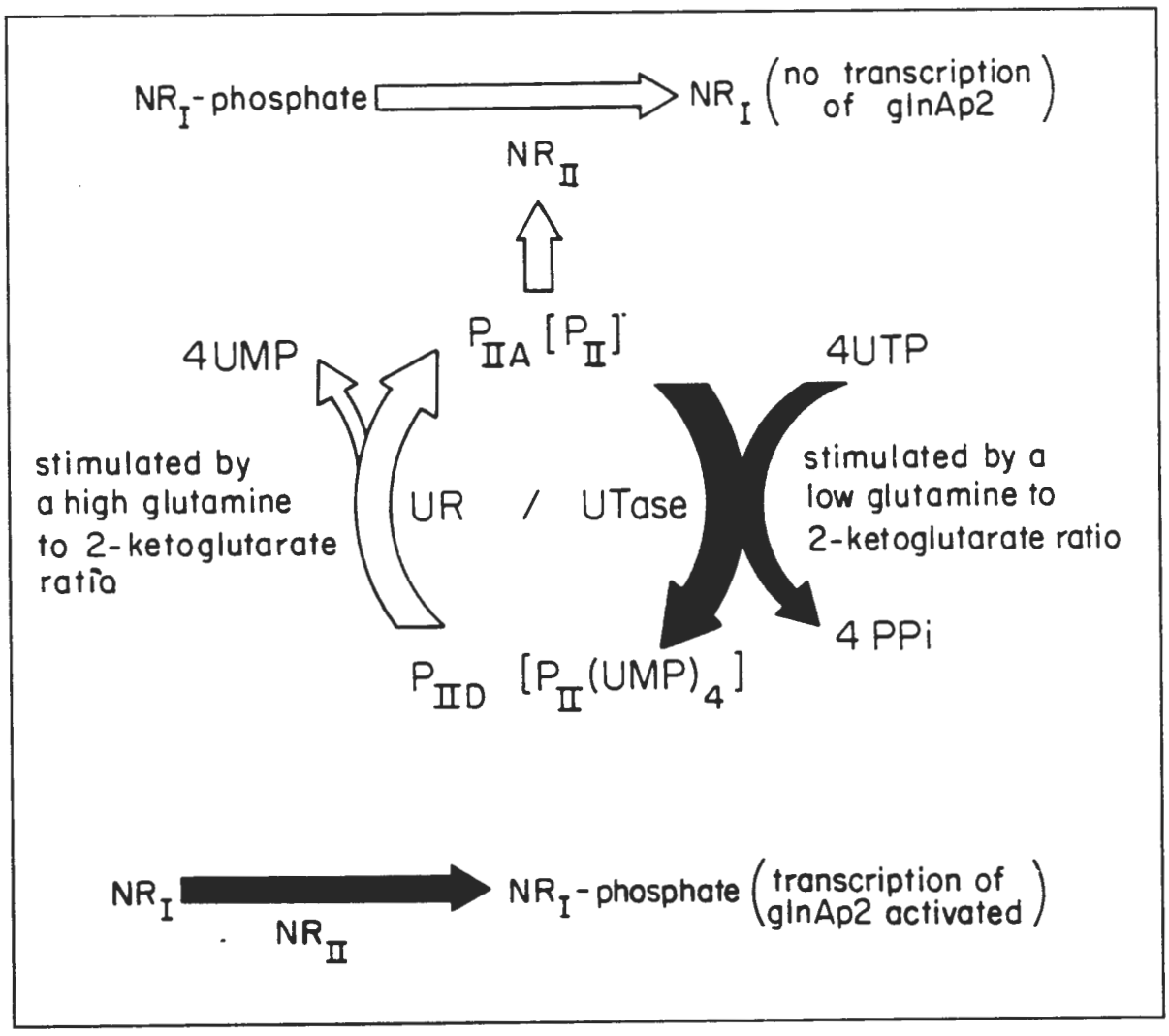


Fig. 1.2. Regulation of expression at the gln*Ap2* promoter of the gln*ALG* operon; (after Reitzer and Magasanik, 1987).

The phosphorylated form of NR_I (NR_I-PO_4) is the activator form and the role of NR_{II} is to phosphorylate NR_I. This phosphorylation is a reversible reaction (Ninfa and Magasanik, 1986; Ninfa *et al.*, 1986). The phosphorylation of NR_I is carried out by a two-step process. The first step is the reaction between NR_{II} and ATP which results in the attachment of the γ-phosphate of ATP to a histidine residue of NR_{II}. The phosphate is then transferred to an aspartate residue of NR_I by a reaction with NR_{II}. NR_I-PO_4 activates transcription from the *glnAp2* promoter, which results in increased intracellular levels of GS and NR_I-PO_4. This activation continues as long as the glutamine/α-ketoglutarate ratio remains low (nitrogen limitation) and P_{II} remains in its uridylylated form (Fig. 1.2). NR_I-PO_4 is dephosphorylated to its repressor form very rapidly when conditions of nitrogen excess occur and this reaction is signalled by the removal of the uridylyl residues from P_{II}, which in this unmodified form is required together with NR_{II} for the removal of the phosphate group from NR_I-PO_4. NR_I is thus rendered unable to activate transcription from *glnAp2* (Fig. 1.2). Enhancer qualities of the NR_I UAS's, and recent discoveries in the mechanisms of transcriptional activation by NR_I-PO_4 will be discussed in Chapter 5.

1.2.2.2 Other Ntr regulated operons.

Apart from the *glnALG* operon several other genes whose products are involved in nitrogen metabolism are regulated by the three components of the Ntr regulatory system: σ54, NR_I, and NR_{II}. This includes ammonia and glutamine uptake systems in *E. coli*, which increase the ability of the cell to scavenge a diminishing supply of these nitrogen sources (Servin-Gonzalez and Bastarrachea, 1984; Jayakumar *et al.*, 1986; Nohno *et al.*, 1986). The expression of genes which encode the uptake systems for arginine (*argT* gene expressed from the *argTr* region) and histidine (*hisJQMP* operon expressed from the *dhuA* region) in *S. typhimurium* have been shown to be under transcriptional control by the Ntr system (Kustu *et al.*, 1979; Stern *et al.*, 1984). The nucleotide sequences of the *argTr* and *dhuA* regions have revealed NR_I binding sites and promoters which may be recognized by Eσ54, and it has been shown for *argTr*, at least, that these are the regions of Ntr control (Higgins and Ames, 1982; Schmitz *et al.*, 1987, 1988). The Ntr regulatory system is responsible for the first step in the regulation of expression of nitrogen fixation genes in bacteria such as *K. pneumoniae* as will be discussed in section 1.3.2.3.

1.3 Nitrogen Fixation

Biological nitrogen fixation (N_2-fixation) is a phenomenon exhibited by members of divergent prokaryotic taxonomic groups (Burns and Hardy, 1975). These include *Azotobacteriaceae*, *Enterobacteriaceae*, *Rhodospirillaceae*, *Bacilliaceae*, *Rhizobiaceae*, *Actinomycetaceae*, and *Cyanobacteria*. The fixation of atmospheric dinitrogen (N_2) is the process whereby these organisms reduce N_2 to ammonia (NH_3), which is a source of nitrogen which may be readily incorporated into cell constituents as discussed in section 1.2. This ability, termed diazotrophy, is an important property for those organisms which inhabit environments which lack fixed nitrogen.

The current research interest in N_2-fixation is enormous and this can be attributed to the possible benefits that biological N_2-fixation poses for world agricultural yields. It is estimated that industrial nitrogen fertilizer chemically fixed from the atmosphere by the Haber process supports a third of the world's population (Postgate, 1989). The rate of human population growth indicates that over the next few decades the net input of fixed nitrogen into the world's agricultural soils will have to increase dramatically. There is much hope (as well as reason behind scientific project proposals for funding) that improvements in the current exploitation of biological N_2-fixation and the development of novel N_2-fixation systems will address this shortfall.

A rational explanation of biological N_2-fixation and the isolation of the bacteria responsible was reported a little over 100 y ago by Hellriegel and Wilfarth and by Beyerinck, respectively (discussed in Quispel, 1988); and since then many discoveries have enabled dissection of the process in great detail (Burris, 1988). The power of mutational studies and recombinant DNA technology is reflected in the increase in research into the genetics of N_2-fixation during the last two decades.

The ambit of this study does not justify a complete review of biological N_2-fixation processes, consequently the genetic regulation of N_2-fixation will be focussed upon. Two major types of N_2-fixation systems have been studied most extensively: (i) bacteria that are diazotrophic in the free living state, such as *K. pneumoniae* and *Azotobacter vinelandii*; and (ii) bacteria, such as *R. meliloti* and *B. japonicum*, that fix

nitrogen in symbiotic relationship with leguminous plants. The genetic regulation of N_2-fixation in *K. pneumoniae* will be discussed and where relevant comparisons will be made with other nitrogen fixing systems.

1.3.1 Physiological regulation of nitrogen fixation

Nitrogen fixation is energetically expensive (approximately 28 moles of ATP required per mole of nitrogen fixed; Postgate, 1982), and requires a complex metabolic machinery. Many bacteria employ N_2-fixation only under conditions when there is an absence of more easily assimilable nitrogen sources, consequently several mechanisms exist to ensure that the N_2-fixation system is active only under appropriate circumstances. This regulation consists of two components: inhibition of the activity of the nitrogenase enzyme by environmental factors, and the transcriptional regulation of the N_2-fixation genes.

The most important physiological constraint is that imposed by the oxygen sensitivity of the nitrogenase enzyme (Kelly, 1969). Various mechanisms have evolved in diazotrophs for protecting the enzyme system from O_2 denaturation. *K. pneumoniae* will only fix nitrogen under strict anaerobic conditions. Members of the genus *Azotobacter* exhibit several physiological responses which reconcile the strongly reductive process of N_2-fixation with oxidative phosphorylation (reviewed in Haaker *et al.*, 1988). These include respiratory protection, which is not a nif-specific phenomenon, but it does serve to increase the rate of respiration in response to increases of environmental oxygen (Drozd and Postgate, 1970). Other methods involve an increase in intracytoplasmic membranes and conformational protection of the nitrogenase enzyme by reversible complex formation with a Fe/S protein in response to an increase in O_2 concentration (Post *et al.*, 1983; Scherings *et al.*, 1983). Compartmentalization is a method used in the cyanobacteria, which set aside specific nitrogen fixing heterocyst cells which are separate from the photosynthesizing cells. The most sophisticated compartments for restricting access of oxygen to nitrogenase are the nodules of leguminous plants. These nodules are the result of colonization by members of the families *Rhizobiaceae* and *Bradyrhizobiaceae*, and their characteristic pink colour is due to a haem-protein, leghaemoglobin. This protein binds O_2 which enables it to supply oxygen to the bacteroids (rhizobia that have colonized the root

nodules), but at a concentration low enough not to damage the nitrogenase (reviewed in Postgate, 1982).

Control of gene expression is a second level of regulation which enables repression and derepression of nitrogenase activity in the presence and absence of oxygen and fixed nitrogen sources, respectively.

1.3.2 Nitrogen fixation genes and their products

K. pneumoniae is a facultative anaerobe, and as it has been investigated most extensively it is often regarded as the model for the study of the genetic regulation amongst free-living diazotrophs (Drummond, 1984; Ausubel, 1984), however much understanding has been obtained in recent years concerning the N_2-fixation system of members of the genus *Azotobacter* which are aerobic free-living diazotrophs (Merrick, 1988).

1.3.2.1 Organization of *nif* genes in *K. pneumoniae* and *A. vinelandii*

The *K. pneumoniae nif* regulon is confined to a single cluster of chromosomal genes adjacent to the histidine biosynthesis genes (MacNeil *et al.*, 1978). Physical mapping (Riedel *et al.*, 1979), genetic mapping involving transposon mutagenesis (Merrick *et al.*, 1980), and nucleotide sequencing (reviewed in Merrick, 1988; Arnold *et al.*, 1988; Cannon *et al.*, 1988) have revealed that this regulon comprises 20 genes in eight transcriptional units extending over 23-kbp (Fig. 1.3). The structural enzyme at the core of the N_2-fixation process is termed nitrogenase, which is a molybdenum requiring enzyme in *K. pneumoniae*. This is made up of the Fe-protein (Dinitrogenase reductase, component 2, or Kp2) which consists of two identical subunits encoded by the *nifH* gene, and the Mo-Fe-protein (Dinitrogenase, component 1, or Kp1) which is a tetramer consisting of dimers of the products of the *nifD* and *nifK* genes (Fig. 1.3). The product of *nifM* has been implicated in Fe-protein maturation. The products of *nifQ*, *nifB*, *nifE*, *nifN*, and *nifV* are thought to be involved in the synthesis of Mo-Fe-cofactor which forms part of the active site. The roles of *nifY*, *nifU*, *nifS*, *nifZ*, *nifW*, *nifC* and *nifT* products are presently not clearly understood. The *nifL*, *nifA*, and *nifX* products are regulatory proteins (Fig. 1.3)(reviewed in Merrick, 1988; Gosink *et al.*, 1990).

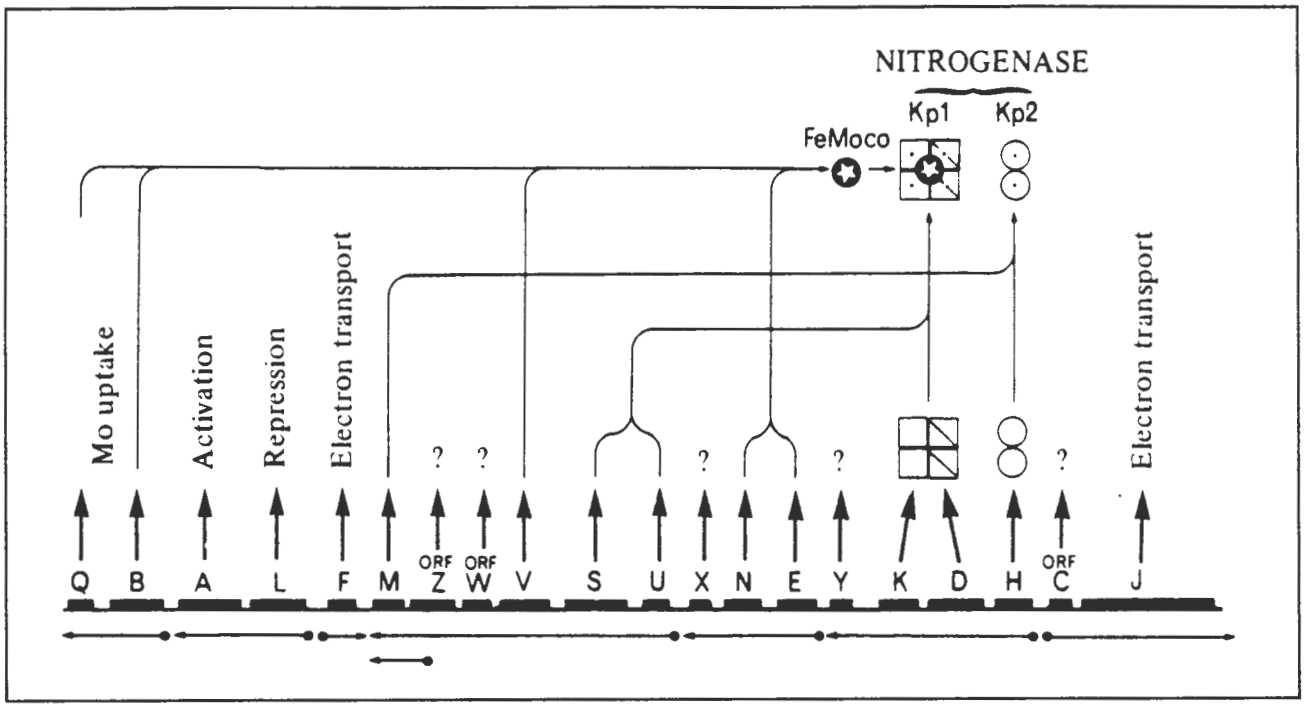


Fig. 1.3. Map of the *nif* gene cluster in *K. pneumoniae*. The roles of the gene products are indicated schematically by the vertical arrows above each gene. The horizontal arrows indicate the extent and direction of each transcript, black dots representing the location of σ54-dependent promoters (after Dixon, 1988).

A major difference between the N_2-fixing system in *K. pneumoniae* and the *Azotobacter*'s is the existence of "alternative" nitrogenase enzymes in members of the latter group (Bishop *et al.*, 1980, 1986; reviewed in Bishop and Joerger, 1990). These enzymes consist of components encoded for by different genes to those encoding the molybdenum nitrogenase. These "alternative" nitrogenase enzymes replace the Mo-nitrogenase when the cells are in an environment devoid of Mo. Both *A. vinelandii* and *Azotobacter chroococcum* possess a second (vanadium(V)-containing) nitrogenase (Hales *et al.*, 1986; Robson *et al.*, 1986a). A third nitrogenase is synthesized by *A. vinelandii* in an environment without either metal (Chisnell *et al.*, 1988). These alternative systems are encoded by reiterated genes. *A. chroococcum* has second copies of the *nifHDK* and *nifEN* genes (designated $H^*D^*K^*$ and E^*N^*). *A. vinelandii* has three copies of *nifHDK*, designated 1,2, and 3. A recent study has shown that a *nifHDK* deletion mutant of the purple non-sulphur photosynthetic bacterium *Rhodobacter capsulatus* shows slow diazotrophic growth in molybdenum-free medium, implicating the presence of an "alternative" nitrogenase system (Klipp *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). The *nifA/nifB* gene region is duplicated in this organism (Klipp *et al.*, 1988; Masepohl *et al.*, 1988).

1.3.2.2 Organization of *nif* genes in other bacteria

The organization of *nif* genes in the bacteria studied to date range from a tightly linked cluster of operons (as shown for *K. pneumoniae*) to a dispersed set of genes. These genes may be located on the chromosome or on plasmids. The reiterated *nif* genes of the *Azotobacter*'s are unlinked and situated on the chromosome. In *Azospirillum brasiliense* N_2-fixation is encoded for by *nif* and *fix* genes which are scattered on at least 65-kbp of genomic DNA (Elmerich *et al.*, 1988; Singh *et al.*, 1989). Some strains of *Enterobacter agglomerans* were found to contain a large plasmid carrying a tightly linked *nif* gene cluster (Singh *et al.*, 1988). This differed in organization from the *K. pneumoniae nif* cluster (Fig. 1.3) in that the *nifJ* and *nifF* genes were positioned and potentially co-transcribed at the opposite end of the cluster to the *nifHDK* genes (Klingmüller, pers. comm.). The *Anabaena nifH* and *nifD* genes are adjacent but are separated from *nifK* by 11-kbp of DNA in non-nitrogen fixing vegetative cells. Under conditions of nitrogen deprivation and heterocyst differentiation this organism displays a novel form of positive regulation by excision of this 11-kbp fragment by site-specific recombination to produce a contiguous active *nifHDK* operon (Golden *et al.*, 1985, 1987; reviewed in Haselkorn, 1986). A second deletion of 55-kbp reconstitutes the *nifB-fdxN-nifSU* operon in a similar manner (reviewed in Haselkorn, Eighth International Conference on Nitrogen Fixation, 1990). *B. japonicum* displays a complex organization of *nif* and *fix* genes in two major chromosomal clusters (Hennecke *et al.*, 1988). A complex organization of *nif*, *fix*, and *nod* genes are borne on large plasmids in members of the genus *Rhizobia*, such as *R. meliloti* (Kondorosi *et al.*, 1988; reviewed in Long, 1989).

This variation in organization of *nif* genes raises the question as to what came first: the clustered or dispersed organization? The answer to this is not clear. Two theories have been put forward to explain the evolution of nitrogen-fixation systems: (i) The seemingly haphazard distribution of nitrogen-fixing ability, even amongst strains of the same genus, indicates a common origin followed by divergent loss. This is supported by the congruence between the rate of divergence of eight nitrogenase Fe-protein sequences and the rate of divergence calculated for the 16s-RNA sequences of the same eight species (Hennecke *et al.*, 1985). (ii) The identification of plasmid-borne N_2-fixation systems and the ease with which *nif* transfer can occur in

the laboratory (Dixon and Postgate, 1972; Dixon *et al.*, 1976; Postgate and Kent, 1987) implies that lateral *nif* gene transfer could occur (Postgate, 1982; Postgate and Eady, 1988). Biochemical and genetic studies on the nitrogen fixing systems of the archaebacteria and their relationship to the eubacterial nitrogen fixing systems may provide some answers to these fundamental questions (Postgate, 1989).

1.3.2.3 Regulation of *nif* gene expression in *K. pneumoniae*

K. pneumoniae nif gene transcription is regulated by a two-tier cascade system. The first level of control involves the Ntr regulatory circuit which through σ54, NR_I, and NR_{II} regulates the expression of the *nifLA* operon in response to the cell's nitrogen status (Merrick, 1983; Drummond *et al.*, 1983). This regulation is similar to that described in section 1.2.2.1. The second level of control is mediated by the regulatory proteins NifL and NifA (products of the *nifLA* genes) and σ54 which regulate the other *nif* operons in response to both O_2 and N-status (Buchanan-Wollaston *et al.*, 1981b; Ow and Ausubel, 1983; reviewed in Gussin *et al.*, 1986)(Fig. 1.4).

A common feature of all *nif* operons in *K. pneumoniae* is the consensus promoter with conserved -GG- and -GC- doublets at approximately positions -24 and -12 with respect to the transcription start site, respectively. This was initially observed in *K. pneumoniae* (Beynon *et al.*, 1983) however it also appears to be a feature of *nif* genes in *Rhizobium*, *Azorhizobium*, and *Bradyrhizobium* (reviewed in Gussin *et al.*, 1986), *Azotobacter* (Brigle *et al.*, 1985; Robson *et al.*, 1986b), *T. ferrooxidans* (Pretorius *et al.*, 1987), *E. agglomerans* (Kreutzer *et al.*, 1989), and *Desulfovibrio* (Merrick, 1988). This consensus promoter has been found to be the site recognized by the form of RNA polymerase complexed to σ54 (Eσ54) and therefore has been termed a σ54-dependent promoter (an alternative name is a NtrA-dependent promoter, because σ54 is encoded by the *ntrA* gene). σ54-dependent promoters are not confined to *nif* and Ntr-regulated genes (section 1.2.2) but appear to be a feature of genes whose products have diverse physiological roles (section 1.4; reviewed in Kustu *et al.*, 1989).

The σ54-dependent promoter is a feature common to both levels of regulation of *K. pneumoniae nif* gene expression, so the question that one may ask is: how is this

regulation exercised in sequential order without a short circuit by direct activation of all the *nif* operons by NR_I when this becomes activated in response to nitrogen limitation? The fine-tuning of expression from σ54-dependent promoters relies on three factors: (i) the DNA sequence adjacent to the invariant GG,GC pairs, (ii) the recognition sequence upstream of the σ54-dependent promoter for the activator protein, and (iii) the activator and repressor proteins.

It has been shown that σ54, NR_I, and NR_{II} purified from *K. pneumoniae* are required together with core RNA polymerase to direct transcription from the *nifLA* promoter (p*nifLA*)(Austin *et al.*, 1987; Wong *et al.*, 1987). Phosphorylation of NR_I by NR_{II}, as described for the *E. coli* Ntr system, is required for activation at the *K. pneumoniae* p*nifLA* (Ninfa and Magasanik, 1986; Minchin *et al.*, 1988). However, a 5-10 fold greater concentration of NR_I was required for activation of p*nifLA* than for *glnAp2*. This reflects the absence of high affinity NR_I binding sites in p*nifLA* (two low affinity sites at -142 and -163 have been identified) as compared to the *glnA* promoter. Co-operative binding of NR_I at these two sites in p*nifLA* appears to be required to achieve maximal activation (Contreras and Drummond, 1988; Minchin *et al.*, 1988). A further difference between p*nifLA* and *glnAp2* is that transcription from p*nifLA* *in vitro* is dependent on negative supercoiling at physiological salt conditions (Dixon *et al.*, 1988) and is reduced *in vivo* by mutations in DNA gyrase and gyrase inhibitors (Dimri and Das, 1988). Minchin *et al.* (1989) have shown that NR_I mediated activation of p*nifLA* is face-of-the-helix dependent both *in vivo* and *in vitro* and that Eσ54 does not make strong contacts with the -24,-12 region of p*nifLA*. Both of these features are in contrast with expression from *glnAp2* (Magasanik, 1988). Consequently, NR_I-PO_4 is required for stabilization of the interaction of Eσ54 with p*nifLA*.

These data are consistent with the model that transcription of p*nifLA* is only initiated once the intracellular level of NR_I-PO_4 is high, which would occur once the cells were subjected to severe nitrogen starvation. This control would be carried out because Eσ54 is unable to form a stable closed promoter complex at p*nifLA* in the absence of NR_I-PO_4, in contrast to *glnAp2* (point (i) above). NR_I-PO_4 mediated transcription at p*nifLA* produces the NifL and NifA proteins and these proteins regulate expression of

not only all the other *nif* genes, but are also capable of autogenously regulating expression from p*nifLA*.

This second tier of regulation is distinguished from the regulation at p*nifLA* by differences in the activator and repressor proteins, and the activator protein recognition site (points (ii) and (iii) above). Consequently, the absence of NR_I binding sites upstream of the other *nif* genes prevents short circuiting and activation by NR_I-PO_4.

The other *nif* promoters contain a conserved upstream activator sequence (UAS) with a consensus TGT-N_{10}-AGA which is located more than 100 bp from the transcription initiation site (Buck *et al.*, 1986, 1987b). Methylation protection studies have demonstrated that this is the site of NifA binding *in vivo* (Morett and Buck, 1988). A helix-turn-helix (HTH) motif situated in the carboxy-terminus of the protein has been shown to be responsible for DNA binding at the UAS of the *nifH* promoter (Morett *et al.*, 1988). NifA therefore acts as the activator of the *nif* operons and the current model involves looping of the DNA between the UAS and the *nif* promoter to enable interaction between NifA bound at the UAS and Eσ54 at the *nif* promoter (Buck *et al.*, 1987a; Buck and Cannon, 1989; Hoover *et al.*, 1990). Integration host factor (IHF), which has been shown to bend DNA (Robertson and Nash, 1988), has been implicated in this process. Purified *E. coli* IHF has been shown to bind to a region between the UAS and σ54-dependent promoter of *nifH* genes from several bacteria, including *K. pneumoniae*, *R. meliloti*, and *A. vinelandii* (Santero *et al.*, 1989; Santero *et al.*, Eighth International Conference on Nitrogen Fixation, 1990; Hoover *et al.*, 1990). Putative IHF binding sites have been identified in the regulatory regions of *nif* genes from several bacteria, which included the *K. pneumoniae nifU* and *nifE* genes, the *E. agglomerans nifU* gene, and eight out of 12 *A. vinelandii nif* genes (Cannon *et al.*, 1990; Hoover *et al.*, 1990). Maximal expression from the *K. pneumoniae nifH* and *nifU* promoters in *E. coli* required IHF (Cannon *et al.*, 1990). Moreover, IHF was required for the establishment of a Nif^+ phenotype in *E. coli* from the plasmid pRD1, which carries the entire *K. pneumoniae nif* gene cluster (Cannon *et al.*, 1990).

The sequences adjacent to these *nif* promoters are not conducive to strong interactions with Eσ54 in the absence of an activator (as described for p*nifLA*). In this context, it has been shown for the *nifH* promoter that activation by NifA requires the correct orientation of the UAS and *nif* promoter with respect to the face of the helix (Buck *et al.*, 1987a). Furthermore, it was possible to construct a *nifH* promoter mutated at positions -17 to -15 at which activation by NifA was no longer face-of-the-helix dependent (Buck and Cannon, 1989). These results were substantiated by *in vivo* footprinting experiments, which demonstrated σ54-dependent methylation protection at this mutant *K. pneumoniae nifH* promoter, but not at the wildtype *nifH* promoter (Morett and Buck, 1989). Moreover, $KMnO_4$ footprinting revealed that this σ54-dependent protection was due to the formation of a closed promoter complex by Eσ54 at the mutant promoter (Morett and Buck, 1989). This is consistent with the model that the constraint on expression of the *K. pneumoniae nifH* gene, at least, is due to a weak NifA-independent interaction of Eσ54 with the wildtype *nifH* promoter (Morett and Buck, 1989). Fidelity of activation is ensured by the requirement for the activator, NifA, to bind to the UAS.

Repression of *nif* gene expression is controlled through the action of NifL which antagonizes the action of NifA in the presence of fixed nitrogen or oxygen (Hill *et al.*, 1981; Merrick *et al.*, 1982; Filser *et al.*, 1983). Although there is amino acid sequence homology between NifL,NifA and NR_{II},NR_I (Drummond *et al.*, 1986) the mode of action of NifL appears to differ considerably from that of NR_{II}. NifL, unlike NR_{II}, is not a bifunctional protein (ie. it is not involved in the activation function of NifA). The mechanism whereby NifL prevents activation of transcription by NifA is not yet known, however Henderson *et al.* (1989) have shown that this requires either iron or manganese ions. A possible redox sensitive site has been identified in the predicted amino acid sequence of NifL (Drummond and Wootten, 1987). Involvement of the redox environment in the activity of NifL is unlikely, however, since its repressive activity is required under aerobic conditions or microaerophilic conditions in the presence of fixed nitrogen. Immunochemical evidence suggests that NifL and NifA form a protein complex, which indicates that protein-protein interactions play a role in this repression.

A secondary level of regulation which may be carried out by NifL is the destabilization of *nif* mRNA, as studies with *nifL*⁻ mutants have shown increased stability of *nif* mRNA in these strains (Collins and Brill, 1985; Collins *et al.*, 1986). A recent report that a *nifX* insertion mutant responded more slowly to the repressive effects of O_2 and NH_4^+, and that overexpression of the *nifX* region blocked *nif* protein synthesis, protein accumulation, and *nifHDKTY* mRNA accumulation has implicated the product of the *nifX* locus in negative regulation (Gosink *et al.*, 1990). The NifA protein has been found to be thermolabile, therefore limiting N_2-fixation in *K. pneumoniae* to temperatures below 37°C (Buchanan-Wollaston *et al.*, 1981a; Brooks *et al.*, 1984).

1.3.2.4 Summary of regulation of *nif* genes in *K. pneumoniae*

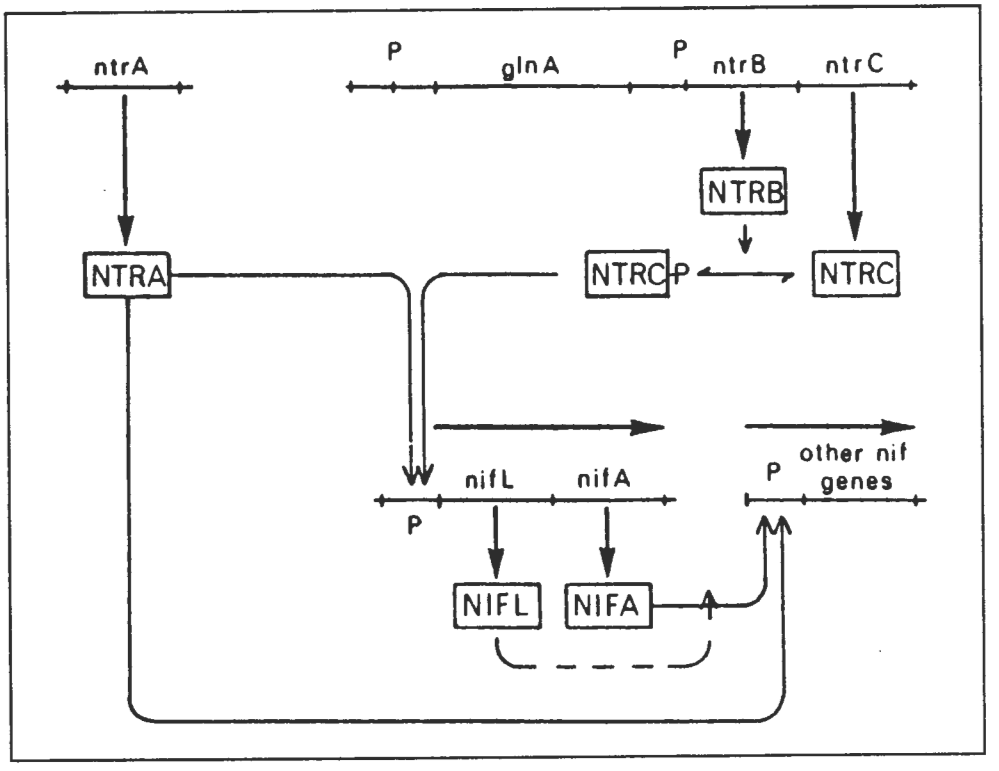


Fig. 1.4. Regulation of *K. pneumoniae nif* gene expression. NTRA = σ54; NTRB = NR_{II}; NTRC = NR_I (after Dixon, 1988).

Fig. 1.4 outlines the current model of *nif* gene expression. On transfer of *K. pneumoniae* cells from nitrogen poor to nitrogen excess conditions, the pool of NR_I proteins will be converted to the dephosphorylated form due to the action of NR_{II}. No activation of transcription from p*nifLA* will occur. Any residual NifA proteins will be

inactivated by NifL, resulting in no transcription from the second level of *nif* operons. Under nitrogen limiting aerobic conditions NR_I-PO_4 will activate p*nifLA* to produce NifL and NifA, however NifL will inactivate NifA in response to the oxygen levels - thus blocking expression of the other *nif* operons. Under anaerobic, nitrogen-limiting conditions, p*nifLA* is activated by NR_I-PO_4, NifL will be in its non-repressive form, and NifA will activate transcription of the *nif* operons, including autogenous activation of its own promoter (p*nifLA*). This will result in synthesis of the nitrogenase proteins and the capacity for N_2-fixation.

1.3.2.5 Comparison of *nif* gene regulation in *K. pneumoniae* with *nif* regulatory systems in other bacteria

The regulation of *nif* gene expression in *Azotobacter* contains elements common to the Ntr/Nif regulatory system in *K. pneumoniae*, however it is complicated by differential expression of genes for the "alternative" nitrogenases. The *ntrA*, *ntrB*, *ntrC*, and *nifA* gene equivalents in *A. vinelandii* have been cloned (Toukdarian and Kennedy, 1986; Bennet *et al.*, 1988; Merrick *et al.*, 1987). A further regulatory locus, *nfrX*, has been identified in both *A. vinelandii* and *A. chroococcum* (Santero *et al.*, 1988). Additionally, two *nifA*-like genes, denoted *anfA* and *vnfA*, have been cloned and sequenced (Joerger *et al.*, 1989). It is thought that the products of *ntrD* (another *nifA*-like gene), *nifA*, and *nfrX* are required for the expression of the Mo-dependent N_2-fixation system (Bennett *et al.*, 1988; Santero *et al.*, 1988; Walmsley *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). The *vnfA* product is required for expression of the V-dependent system in *A. vinelandii* (Joerger *et al.*, 1989). Under Mo- and V-deficient conditions, the presence of the *anfA* and *nfrX* genes is essential for N_2-fixation, while the *nifA* product is required for maximal diazotrophic growth (Joerger *et al.*, 1989). The product of *ntrD* has also been implicated in expression of this third nitrogenase system (Walmsley *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). The role of the *ntrC* product, if any, appears to be indirect (Bali *et al.*, 1988). A *nifL* gene has recently been identified upstream of the *A. vinelandii nifA* gene and the Mo-nitrogenase activity of an *A. vinelandii* strain with a mutation in this locus is not subject to repression by high concentrations of fixed nitrogen (Bali and Kennedy, Eighth International Conference on Nitrogen Fixation, 1990).

There are two major groups of N_2-fixing rhizobia which are able to form symbiotic relationships with plants. These are the fast growing species, such as *R. meliloti* and *Rhizobium leguminosarum*, which can only fix N_2 symbiotically; and the slow growing species, such as *B. japonicum*, which are capable of active N_2-fixation *ex planta*. These N_2-fixation systems are encoded by *nif* genes, which bear structural and functional homology to the *K. pneumoniae nif* genes, and *fix* genes, which have been identified as essential loci for symbiotic diazotrophy.

The *R. meliloti ntrC* gene product is believed to be involved in control of nitrogen metabolism in the free-living state (Szeto *et al.*, 1987), however it does not play a role in activation of symbiotic N_2-fixation. This difference in regulation between the *K. pneumoniae* and *R. meliloti* is consistent with the differences in N-status of the environments in which each organism fixes N_2: *K. pneumoniae* requires N-limiting conditions; *R. meliloti* must be able to fix N_2 in the nitrogen rich nodule environment (David *et al.*, 1988). The *R. meliloti* NifA, which shows homology to the *K. pneumoniae* NifA (Buikema *et al.*, 1985), plays a central role in activation of the N_2-fixation system (Batut *et al.*, 1989). The "symbiotic signal" for activation of N_2-fixation in *R. meliloti* bacteroids is not known, however the current model implicates oxygen levels as the possible environmental factor (Ditta *et al.*, 1987) which is sensed by the transmembrane *fixL* product. This protein and the *fixJ* product form a two component regulatory system (see section 1.5) which modulate the activity of the *nifA* product which is the positive activator of the structural *nif* and *fix* genes (Hertig *et al.*, 1989; Batut *et al.*, 1989; Gilles-Gonzalez *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). A further complexity is the involvement of the *fixK* product and the NifA-independent expression of the *fixN* gene (Batut *et al.*, 1989).

B. japonicum is able to fix N_2 during symbiosis or in the free-living state under microaerobic conditions. Separate regulatory elements may exist for N_2-fixation in these two environments. A two-tiered control system has been proposed with expression of a *fixR-nifA* operon at the first level (Thöny *et al.*, 1989). Homologues of the *R. meliloti fixLJ* genes have been cloned and it has been shown that the products of these genes are not required for expression of the *fixR-nifA* operon under

aerobic conditions. Their role in the induction of this operon under anaerobic conditions is under investigation (Anthamatten and Hennecke, Eighth International Conference on Nitrogen Fixation, 1990). The activation of the *nif* and *fix* genes is mediated at a second level by NifA, which requires a divalent metal ion for activation and is irreversibly oxygen sensitive (Hennecke *et al.*, 1988; Fischer and Hennecke, 1987). This highlights the difference between the *K. pneumoniae* NifA, which is not intrinsically oxygen sensitive, and the *B. japonicum* and *R. meliloti* NifA proteins, which are inactivated in the presence of oxygen. This oxygen sensitivity resides in a unique central domain of these proteins, and there is no homology between the rhizobial and *K. pneumoniae* NifA N-terminal regions (Fischer *et al.*, 1988). Furthermore, no NifL homologue has been found in these (brady)rhizobia species. NifL is thought to interact with the N-terminal region of NifA (Nixon *et al.*, 1986).

A striking feature of *nif* genes in those bacteria that have been studied to date (with the exception of *Anabaena, Clostridium,* and the Archebacteria) is the conservation of the σ54-dependent promoter. The survival of this feature throughout the evolution of many nitrogen fixation systems, irrespective of vertical or horizontal transfer, highlights the resilience and importance of the specialized sigma factor σ54, which is encoded by the *ntrA* gene.

1.4 σ54(NtrA)-dependent genes in bacteria

Eubacteria employ a number of sigma factors, which when associated with core RNA polymerase alter the specificity of promoter recognition (reviewed in Helmann and Chamberlin, 1988). The most abundant sigma factor in the Gram-negative bacterium *E. coli* is termed σ70 and this is responsible for constitutive and regulated expression of many genes whose products are responsible for a wide variety of cellular functions (Yura and Ishihama, 1979). The major vegetative sigma factor, σ43, in the Gram-positive bacterium *B. subtilis* directs recognition of a similar canonical promoter to σ70 which appears to be the principal class of promoter in this Gram-positive bacterium (Doi and Wang, 1986). In addition, alternative sigma factors have been identified in each of these groups which allow transcription of genes whose products contribute to a common physiological response. The heat shock

response in enteric bacteria is elicited through σ32 (Grossman *et al.*, 1984). Several alternative sigma factors are required during the sporulation process in *B. subtilis* (Losick *et al.*, 1986). Bacteriophages that infect these groups also encode sigma factors.

As has been described in sections 1.2.2 and 1.3.2.3, σ54 is the major sigma factor required for the expression of genes whose products are involved in nitrogen metabolism and nitrogen fixation. Recent advances in understanding have shown that σ54 differs from the other alternative sigma factors in that it is required for expression of genes whose products are involved in a wide variety of cellular functions. The role of σ54 in these systems has been reviewed comprehensively by Kustu *et al.* (1989). These and additional σ54-dependent genes, which have subsequently come to light, will be discussed in this section.

σ54 was first identified in association with nitrogen metabolism in enteric bacteria and therefore this designation was adopted to reflect the M_r of the product in enteric bacteria. Subsequently, homologues in other organisms with different M_r have been identified. The gene designation (*glnF*, *ntrA*, *rpoN*) also reflects this history, however its broader role justifies revision to *rpoE* or a *sig* designation. For convenience, the designations σ54 (or NtrA) and *ntrA* have been adopted in this study.

There are several lines of evidence which may be used to identify a gene which requires σ54 for expression:

(i) presence of a σ54-dependent promoter - conserved GC doublet (11-14 bp upstream of the transcription start site) with a GG doublet exactly 10 bp upstream.
(ii) lack of transcription in a *ntrA* mutant strain.
(iii) lack of transcription when the gene is transferred to an *E. coli ntrA* mutant.
(iv) mutational analysis of the σ54-dependent promoter
(v) *in vitro* transcription with Eσ54 (σ54-RNA polymerase holoenzyme).
(vi) dependence on a transcriptional activator which works in conjunction with Eσ54.
(vii) presence of upstream activator binding sites.

The most immediate method available is therefore sequence scanning of the upstream regulatory region of a gene for features (i) and (vii), however it is important to be aware that this evidence is speculative in the absence of biological evidence provided by the other techniques.

The assimilatory nitrite and nitrate reductase systems, which enable cells to grown on these compounds as a sole nitrogen source have been found to be inactive in *ntrA* mutant strains of *A. vinelandii*, *R. meliloti*, and *Alcaligenes eutrophicus* (Santero *et al.*, 1986; Ronson *et al.*, 1987b; Romermann *et al.*, 1988). The activator protein required for expression of these systems in *K. pneumoniae*, *A. vinelandii*, *R. meliloti*, *A. tumefaciens*, and *A. brasiliense* appears to be a NR_I homologue (Cali *et al.*, 1989; Toukdarian and Kennedy, 1986; Szeto *et al.*, 1987; Rossbach *et al.*, 1987; Pederosa and Yates, 1984).

The *dctA* gene, which is a component of the C_4-dicarboxylic acid transport system in the rhizobia, has a σ54-dependent promoter and is not expressed in *ntrA* mutant strains. The DctD protein is the activator, at least in the free-living state (Ronson *et al.*, 1987a). UAS's upstream of the *dctA* genes have been identified as the sites required for full activation by the DctD proteins in both *R. meliloti* and *R. leguminosarum* (Ledebur *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). Pleiotrophic *hno*$^-$ mutants of *A. eutrophicus*, which are thought to have a lesion in an *ntrA* gene, are defective in C_4-dicarboxylic acid transport.

The *melA* gene, which encodes the tyrosinase required for synthesis of the pigment melanin in *R. leguminosarum*, requires the activator NifA and is not expressed in *E. coli ntrA* mutants (Hawkins and Johnston, 1988). NifA-activated expression, a consensus NifA UAS, and a σ54-dependent promoter has been identified for genes located on a *R. meliloti* cryptic plasmid which may affect nodulation efficiency (Sanjuan and Olivares, 1989).

The *E. coli* anaerobic formate hydrogenlyase pathway is dependent upon σ54 for expression (*fdhF* and *hyd-17* genes). This pathway is not active in an *E. coli ntrA* mutant, and the *fdhF* gene has a σ54-dependent promoter and an UAS (Birkmann

et al., 1987a; Birkmann and Böck, 1989). The *hyd-17* locus consists of divergently transcribed operons with σ54-dependent promoters (Lutz *et al.*, 1990; Böhm *et al.*, 1990). Transcriptional activation of the hydrogenase component, at least, of this pathway is thought to require the product of the *hydG* gene (Stoker *et al.*, 1989). Other workers have identified a gene encoding a putative transcriptional activator of this pathway denoted *fhlA* (Sankar *et al.*, 1988: identified as ORFE within the *hyd-17* locus; Böhm *et al.*, 1990), but at this stage it is not clear as to whether *fhlA* is identical to *hydG* (Schlensog and Böck, in preparation).

The *hox* genes are clustered on a 450-kbp megaplasmid in *A. eutrophicus* and these provide the cells with the ability to oxidize H_2 as a source of energy (Friedrich and Friedrich, 1983). This property is lost in a *ntrA* mutant; a σ54-dependent promoter has been identified for the *hoxC* gene; and the product of *hoxA* is the activator of this pathway, which is induced in response to energy limitation (Romermann *et al.*, 1988). Expression of *hox* genes encoding a membrane bound hydrogenase in *Pseudomonas facilis* is σ54-dependent (Romermann *et al.*, 1989).

The expression of cell surface components in several bacteria has been shown to be σ54-dependent. The sequences of the regulatory regions of pilin genes in *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae*, *Bacteroides nodosus*, and *Moraxella bovis* has revealed the canonical σ54-dependent promoter (Johnson *et al.*, 1986; Meyer *et al.*, 1984). A *P. aeruginosa ntrA* mutant failed to synthesize pilin (Ishimoto and Lory, 1989). Furthermore, a putative UAS has been identified upstream of the *P. aeruginosa pilA* gene (Pasloske *et al.*, 1989). These results implicate σ54 in the increased virulence associated with pilin production in *P. aeruginosa* and *N. gonorrhoeae*.

The σ54-dependence of motility functions and flagellin synthesis has been studied intensively in *Caulobacter crescentus*. The *flbG* (hook operon) and *flaN* operons have σ54-dependent promoters and conserved UAS's, termed *ftr* (Mullin *et al.*, 1987). *E. coli* Eσ54 is able to recognize these promoters *in vitro* (Ninfa *et al.*, 1989). Furthermore, it has been demonstrated, using site-directed mutagenesis, that the *ftr* elements are required for transcription (Mullin and Newton, 1989); and the FlbD

protein has been identified as the transcriptional activator (Ramakrishnan and Newton, 1990). *NtrA* mutants of *P. aeruginosa* and *Pseudomonas putida* are non-motile and this has been shown by electron microscopy to be due to the inability to synthesize flagella (Totten *et al.*, 1990; Inouye *et al.*, 1990). These results are interesting as another type of sigma factor (sigma F in enteric bacteria, σ28 in *B. subtilis*) is required for transcription of flagellar and chemotaxis functions in these bacteria (Helmann *et al.*, 1988; Arnosti and Chamberlin, 1989). The role of this sigma factor in motility functions of *C. crescentus* or *Pseudomonas* species is unknown.

σ54 is involved in regulation of various metabolic functions in *P. putida* (Köhler *et al.*, 1989b). The enzyme carboxypeptidase G2, which hydrolyzes the C-terminal moiety from folic acid, is encoded by a gene which has the sequence of a σ54-dependent promoter (Minton and Clarke, 1985). The expression of some of the enzymes that catabolize toluene and related aromatic compounds, coded for by genes on the TOL plasmid, is σ54-dependent. The consensus promoter has been identified in the *xylCAB* and *xylS* genes (Inouye *et al.*, 1984) and the XylR protein has been shown to be the activator (Inouye *et al.*, 1987). The *xylCAB* operon was the first non-nitrogen metabolism set of genes which was shown to be σ54-dependent: expression of this operon was activated in the presence of the *E. coli ntr* genes (Dixon, 1986).

The microorganism *Myxococcus xanthus* carries out a complex lifecycle involving aggregation and fruiting body formation. The *mbhA* gene, which codes for hemagglutinin, has a consensus σ54-dependent promoter. Mutational analysis has shown this to be required for transcription (Romeo and Zusman, 1987).

It appears that each species of micro-organism has a unique mosaic of σ54-dependent genes which code for a wide variety of metabolic processes. These range from biosynthetic pathways to degradative enzymes to structural components. Each function confers upon its host organism the ability to capitalize upon a specific nutritional or environmental niche, and therefore they could be termed "luxury" genes. Undoubtedly, many more σ54-dependent genes will be identified in future. It remains

an intriguing question as to how the components required for this transcription (*ntrA* encoding σ54, the σ54 consensus promoter, and an activator and UAS specific for each function) have been acquired, and are still being acquired, during the course of evolution. The identification of more UAS sequences and the sequences of the activator proteins which recognize them may reveal some basic rules which govern this process.

1.5 Global regulation of gene expression in bacteria

Bacteria are often regarded as evolutionary "primitive" organisms in relation to eukaryotes on account of their relative simplicity of molecular organization. This, however, is a misconception as it is now evident that most species of bacteria are highly efficient organisms which have streamlined their metabolic machinery to respond rapidly to environmental conditions. Strategies employed by bacteria as an adaptive response to changing environmental conditions range from rapid changes in motility to long-term global reorganizations of gene expression and cell morphology. Global regulation of gene expression lies at the heart of these processes. The key to selective advantage, and therefore survival, for bacteria is possessing sensitive sensory mechanisms which are able to respond to a changing habitat and communicate this information to the transcription machinery.

A common mechanism whereby this "stimulus-response coupling" or "signal transduction" is carried out has in recent years been shown to be widespread in many bacterial species. This mechanism involves two types of enzymatic components: histidine protein kinases (HPK), and their associated response regulators (RR) (reviewed in Kofoid and Parkinson, 1988; Bourret *et al.*, 1989; Stock *et al.*, 1989b, 1990). In most cases, the HPK (alternatively termed the "sensor" or "modulator") acts as the sensor of the environmental status and transmits a signal to the RR which in turn regulates the activity of a set of target proteins or the expression of a specific set of genes to elicit an appropriate response. As these two components are central to the signal transduction, the term "two-component" has often been employed to describe these systems, although this may not always be the case as there are often other proteins involved in the signal transduction. Processes regulated in

this way include sporulation, transformation, competence, pathogenicity, virulence, gliding and flagellar motility, membrane transport, and nitrogen metabolism. Furthermore, this mechanism of signal transduction has been identified in both Gram-negative and Gram-positive bacterial species (Stock *et al.*, 1989b).

Stimulus-response coupling by this mechanism has been studied most extensively with respect to regulation of nitrogen metabolism, chemotaxis, osmoregulation, and phosphate balance in the enteric bacteria. Each system is controlled by a specific HPK and RR pair, and sequence analysis has revealed that these belong to two families of homologous proteins (Nixon *et al.*, 1986; Ronson *et al.*, 1987c; Stock *et al.*, 1989b; Albright *et al.*, 1989a). The pairs of proteins (HPK/RR) that are involved in signal transduction in nitrogen regulation, chemotaxis, osmoregulation, and phosphate regulation are NR_{II}/NR_{I}, CheA/CheB and CheY, EnvZ/OmpR, and PhoR/PhoB, respectively (for reviews see Magasanik, 1988; Stewart and Dahlquist, 1987; Csonka, 1989; and Wanner, 1987). Phosphorylation of the RR mediated by the HPK is an essential feature of the signal transduction mechanism and a common phosphotransfer enzymology is thought to be involved (Ninfa and Magasanik, 1986; Wylie *et al.*, 1988; Stock *et al.*, 1988b; Bourret *et al.*, 1989; Igo *et al.*, 1989). Further examples of this system have been identified on aa sequence analysis of regulatory proteins and the recognition of conserved domains which correspond to "transmitter" (HPK) and "receiver" (RR) modules (Kofoid and Parkinson, 1988).

The HPK family is defined by a region of conserved sequence situated near the C-terminus. This represents the transmitter module and contains the histidine residue, which is the site of autophosphorylation of the HPK prior to phosphotransfer to the RR (Keener and Kustu, 1988). Most HPK proteins are postulated to have an extracytoplasmic N-terminal domain which interacts with the "environment" (ie. stimulatory ligands in the periplasm) and a transmembrane domain which transmits the environmental signal to the cytoplasmic C-terminal domain which carries out the kinase function (Forst and Inouye, 1988). Exceptions to this are CheA and NR_{II} which receive signals within the cytoplasm. The nitrogen status signal for NR_{II} is related to the ratio of α-ketoglutarate to glutamine within the cell and is

mediated by the P_{II} protein (section 1.2.2.1). HPK's have also been shown to exhibit phosphatase activity as this is essential for "switching off" of the RR in appropriate circumstances. The phosphatase activity of NR_{II} requires the P_{II} protein (section 1.2.2.1).

The family of at least 25 types of RR's is characterized by a conserved N-terminal domain of about 100 aa (Stock *et al.*, 1989b). Phosphorylation of an aspartate residue within this region by the corresponding HPK is thought to stimulate the activation function of each RR (Stock *et al.*, 1988b). The functional and structural significance of conserved residues within this domain, concluded from the tertiary structure of the *S. typhimurium* CheY protein - a representative RR - (Stock *et al.*, 1989a, 1990), will be discussed in Chapter 5. The conservation in sequence at the N-terminal region of RR's is consistent with the notion that this is the site of interaction with the relevant HPK, however once the RR's have been activated by phosphorylation they direct the target response in a variety of ways. This diversity is reflected in the different sub-families of RR's which may be identified by homologies within their C-terminal regions.

CheY and SpoOF (involved in sporulation in *B. subtilis*) contain only the conserved N-terminal region. NR_I, DctD, HydG, and PgtA (phosphoglycerate transport) constitute a second sub-family and have a homologous C-terminal region which is also found in other proteins which lack the HPK-interacting N-terminal domain, namely FlbD, TyrR, and NifA. The common link between the activity of these proteins which share this conserved C-terminal region is that NR_I, DctD, and NifA have been shown to activate transcription from σ54-dependent promoters (sections 1.2.2, 1.4, and 1.3.2.3). At the extreme C-termini of NR_I, DctD, and NifA a DNA-binding region has been identified which enables these proteins to bind to UAS's (sections 1.2.2 and 1.3.2.3; Ledebur *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). Phosphorylation of NR_I (RR) within its N-terminal domain by NR_{II} (HPK) stimulates its ability to bind DNA and activate transcription (section 1.2.2.1). This example illustrates the link between the HPK activity and the response modulated by the RR, which in this case is the activation of nitrogen metabolism genes. A third sub-family based on C-terminal homology includes PhoB

and OmpR which regulate the expression of genes with promoters recognized by the major form of RNA polymerase (Eσ70 in *E. coli*). A fourth sub-family includes FixJ (section 1.3.2.4) and NarL (nitrate reductase; Nohno *et al.*, 1989) which share a homologous C-terminal domain. Finally, there are some RR's which have C-termini which show little similarity to other RR's, such as AlgR (alginate production - Deretic *et al.*, 1989) and CheB (methylesterase component in chemotaxis).

The similarity in structure and function between the different HPK transmitter C-terminal domains and between the different RR receiver N-terminal domains begs the question as to whether cross-talk is possible. The availability of purified components has enabled the demonstration that CheA can phosphorylate NR_I or OmpR to activate transcription from the *glnA* or *ompF* promoters; and EnvZ can phosphorylate NR_I and CheY (Ninfa *et al.*, 1988; Igo *et al.*, 1989). Cross-talk has also been demonstrated *in vivo*: a hyperactive variant of the nitrogen kinase, NR_{II}2302, was able to act in the place of CheA to produce tumbling/swimming behavior (Ninfa *et al.*, 1988). Moreover, the complex phenotypes exhibited by strains with HPK gene mutations may be explainable by such crosstalk (Backman *et al.*, 1983; Wanner *et al.*, 1988). This versatility indicates that although a specific HPK-RR pair may be the primary regulators of a particular response, other HPK's and RR's could exert peripheral effects. This idea introduces the concept of a global regulatory mosaic within a cell and that a cell's net response is a combination of the regulatory inputs which in part are dictated by HPK-RR interactions. It is likely that peripheral regulatory effects by other HPK-RR would be most significant when a regulatory system is poised at a threshold (eg. when a cell is depleting a rich nitrogen source, such as ammonia) (Stock *et al.*, 1989b).

1.6 Nitrogen metabolism in *T. ferrooxidans*

1.6.1 Sources of nitrogen available to *T. ferrooxidans*

Nitrogen sources that may be present in environments inhabited by *T. ferrooxidans* are likely to be inorganic nitrogen compounds such as ammonia, nitrates, nitrites, and atmospheric dinitrogen. Nitrogen is a microconstituent in igneous rocks and

sedimentary deposits, and its availability would be related to the geochemical composition and degree of weathering (Tuovinen *et al.*, 1979).

Ammonia is a preferred nitrogen source for many bacteria (discussed in section 1.2) and the concentration of this compound in an industrial bioleaching environment is low. A value of 30 mg/l has been quoted for a uranium leach liquor (Rawlings, 1981). An understanding of the chemistry of ammonia at different pH's is important when considering the availability of ammonia in this habitat. At a low pH, ammonia will exist predominantly in the protonated form, NH_4^+, while in more alkaline conditions the equilibrium will be shifted towards a predominance of the neutral form, NH_3 (Kleiner, 1981). A bioleaching environment dominated by *T. ferrooxidans* is invariably acidic. Consequently, any available ammonia will be in the NH_4^+ form. It is thought that NH_3 is able to diffuse rapidly across biomembranes; in contrast, most biomembranes are impermeable to NH_4^+ (Henderson, 1971). It is likely, therefore, that *T. ferrooxidans* would have a transport system to take up the ammonia, which would only be available in the NH_4^+ form. NH_4^+-transport systems have been demonstrated in several bacterial species (Kleiner, 1981). However, the intracellular pH of the *T. ferrooxidans* cells is neutral, which would result in conversion of the majority of the intracellular pool of ammonia to the NH_3 form. This would create a NH_3 gradient across the cell membrane driving the diffusion of NH_3 out of the cells. To prevent a futile cycle, the *T. ferrooxidans* cell membrane would have to possess a reduced permeability to NH_3. In this context, those microorganisms that are neutrophilic, such as *K. pneumoniae*, would be expected to lose NH_3 rapidly if grown at low pH. It has been shown that *K. pneumoniae* cells grown in a glucose limited chemostat at low culture pH values (pH = 4,5 - 5) but provided with excess nitrogen (80 mM NH_4^+Cl) exhibited elevated levels of GOGAT. This phenotype is characteristic of nitrogen-limited cells (see section 1.2.1), indicating that the internal nitrogen sensory system was responding to nitrogen starvation which may have been the result of rapid diffusion of NH_3 out of the cells (Buurman *et al.*, 1989).

Studies on nitrogen requirements for growth and iron-oxidation by *T. ferrooxidans* are complicated by the effect of ammonia absorbed by the acidic medium and by the contribution of nitrogen fixation. However, Tuovinen *et al.* (1979) showed that the

rate of iron-oxidation by *T. ferrooxidans* was limited by the availability of ammonia [$(NH_4^+)_2SO_4$] at concentrations below 3,6 mg/l, and was stimulated by concentrations between 3,6 and 18 mg/l. Furthermore, iron oxidation could be carried out when nitrate was provided as the sole nitrogen source at concentrations of 0,1 mM, and the rate of iron-oxidation was increased when the cells were subjected to a preculture "adaptation" phase on 0,5 mM nitrate. These results indicate the presence of an induction mechanism. Complex nitrogen sources, such as amino acids and casein, appeared to inhibit the growth of *T. ferrooxidans* (Tuovinen *et al.*, 1979). This autotrophic bacterium has also been shown to be unable to grow in the presence of many other organic molecules (Matin, 1978). It therefore appears that *T. ferrooxidans* requires low concentrations of inorganic nitrogen sources for optimal iron-oxidizing activity under aerobic conditions.

The role of nitrogen fixation by *T. ferrooxidans* is not clear, as it has been shown that the nitrogen fixing ability of *T. ferrooxidans* was oxygen sensitive (Mackintosh, 1971), and therefore incompatible with aerobic growth and iron-oxidation. Fixation of atmospheric N_2 by *T. ferrooxidans* has been demonstrated by the measurement of acetylene reduction by *T. ferrooxidans* cultures and the incorporation of ^{15}N into cell material by these cells (Mackintosh, 1978). This activity was inhibited by the addition of 18 mg/l $(NH_4^+)_2SO_4$ and occurred under microaerophilic conditions. This inhibition by O_2 and NH_4^+ is reminiscent of repression systems in other diazotrophs and suggests the existence of a regulatory circuit in *T. ferrooxidans*. The cells were grown on $FeSO_4$ medium, and N_2-fixation only occurred after an initial cycle of exponential aerobic growth, presumably to generate sufficient metabolic energy for the energetically expensive N_2-fixation process (Mackintosh, 1978). Aerobic/anoxic growth conditions are unlikely to prevail in nature, consequently it is not known under what conditions *T. ferrooxidans* is able to fix N_2. A possibility is that under anoxic conditions *T. ferrooxidans* is able to generate energy by the oxidation of inorganic sulphur compounds and the concomitant reduction of Fe^{3+}, Mo^{6+}, or Cu^{2+} (Sugio *et al.*, 1990), and these reactions may produce enough energy for N_2-fixation (Rawlings, pers. comm.).

1.6.2 Genes and enzymes involved in nitrogen metabolism from *T. ferrooxidans*

1.6.2.1 *T. ferrooxidans* glutamine synthetase

The gene encoding glutamine synthetase, *glnA*, from *T. ferrooxidans* ATCC 33020 has been cloned in *E. coli* (Barros *et al.*, 1985) and the nucleotide sequence has been determined (Rawlings *et al.*, 1987). This gene was expressed in *E. coli* and the enzyme was functional in *E. coli*. The *T. ferrooxidans* GS was purified and the subunit showed an apparent M_r of 60 000 (Barros *et al.*, 1986). Assuming the enzyme forms a typical dodecamer as shown for GS's of other Gram-negative bacteria (Almassy *et al.*, 1986), the *T. ferrooxidans* GS was predicted to have a particle M_r of 720 000. Electron microscopy of purified GS revealed characteristic disc shaped molecules with central holes, which are thought to represent the undissociated enzyme. *T. ferrooxidans* GS activity was shown to be modulated by adenylylation (Barros *et al.*, 1986).

The nucleotide sequence of the *T. ferrooxidans glnA* gene revealed some regulatory features that are common to the *E. coli glnA* gene: a *glnAp*1-like promoter, a catabolite-activating protein consensus recognition sequence, and a high affinity NR_I-binding site (as defined for *E. coli glnA* - Reitzer and Magasanik, 1986) which overlapped the *glnAp*1-like promoter (Rawlings *et al.*, 1987). However, no σ54-dependent promoter sequence could be identified (Rawlings *et al.*, 1987). The *T. ferrooxidans glnA* encoded a predicted protein of 468 aa which showed approximately 60% aa similarity with the GSI enzymes of other Gram-negative bacteria such as *E. coli*, *Vibrio alginolyticus*, *R. leguminosarum*, and *A. brasiliense* (Rawlings, 1989). Furthermore, a region thought to be the site of adenylylation in the *E. coli* GS is highly conserved in the *T. ferrooxidans* GS aa sequence, which is consistent with the biochemical results (Rawlings *et al.*, 1987). It was not known prior to this study whether the *T. ferrooxidans glnA* was linked to *ntrBC* genes.

1.6.2.2 *T. ferrooxidans nifHDK* genes

Southern hybridization studies with the *K. pneumoniae nifHDK* genes as a probe identified homologous nucleotide sequences in total DNA preparations from five different iron-oxidizing *T. ferrooxidans* strains, including ATCC 33020 (Pretorius *et al.*, 1986). These results substantiated previous experiments which demonstrated

that *T. ferrooxidans* was able to fix nitrogen (Mackintosh, 1978). The *T. ferrooxidans* ATCC 33020 *nifHDK* genes were cloned (Pretorius *et al.*, 1986) and the nucleotide sequence was determined (Pretorius *et al.*, 1987, Rawlings, 1988). The genetic organization was similar to that observed for the *K. pneumoniae nifHDK* genes with no canonical termination signals between the *nifH* and *nifD* genes or the *nifD* and *nifK* genes, and a regulatory region upstream of the *nifH* gene which contained two tandem NifA UAS's, and a consensus σ54-dependent promoter. The predicted products of the *T. ferrooxidans nifHDK* genes showed greatest aa similarity to equivalent proteins from members of the bradyrhizobia (approximately 80% similarity between NifH proteins; 68% similarity between NifD or NifK proteins), although the *T. ferrooxidans* NifH showed 75 and 71% aa similarity to the *K. pneumoniae* and *A. vinelandii* NifH proteins, respectively (Rawlings, 1988, 1989).

1.7 Aims of the project

There are many challenges associated with a study of the biology of the bacterium, *T. ferrooxidans*, which grows optimally at a pH of between 1,3 and 3,5 and is inhibited by many organic compounds, including agar. The utilization of recombinant DNA technology to isolate genes which encode specific phenotypes and the propagation of these genes in *E. coli* is therefore a desirable approach for studies on *T. ferrooxidans*. *T. ferrooxidans* is a Gram-negative bacterium and it had been shown prior to the inception of this project that *T. ferrooxidans* genes displayed regulatory signals which were recognized in *E. coli*. A key component of bacterial metabolism is the provision of nitrogen within the cell, and the genes for two important nitrogen metabolism structural enzymes, glutamine synthetase and nitrogenase, had been isolated. The aim of this project was to build upon these previous studies by isolating the genetic determinants encoding regulators of nitrogen metabolism in *T. ferrooxidans*. Clues obtained from the nucleotide sequence analysis of the regulatory regions of these genes indicated that *T. ferrooxidans* may contain a Ntr regulatory system as shown in the enteric bacteria. In this thesis I report on the isolation and sequence analysis of the *T. ferrooxidans ntrA* gene (Berger *et al.*, 1990), and the cloning of the *T. ferrooxidans ntrC* gene. In addition, I show what regulatory elements are required for transcription from the *T. ferrooxidans nifH* promoter in *E. coli*.

CHAPTER 2

ISOLATION OF THE *THIOBACILLUS FERROOXIDANS NTRA* GENE

CHAPTER 2

ISOLATION OF THE *THIOBACILLUS FERROOXIDANS NTRA* GENE

2.0 Summary

Attempts to isolate the *T. ferrooxidans ntrA* gene by complementation of an *E. coli ntrA* mutant for Ntr function or hybridization with an *A. vinelandii ntrA* gene DNA probe proved unsuccessful. A novel strategy involving complementation of an *E. coli ntrA* mutant for NtrA-dependent expression of the anaerobic gas-producing formate degradation pathway was developed. The cloned *T. ferrooxidans ntrA* gene was detected using a simple agar overlay technique on the basis of complementation of NtrA-dependent expression of the gas positive phenotype. The *T. ferrooxidans ntrA* gene was subcloned on a 2,8-kilobase-pair (kbp) *Sal*I-*Bgl*II fragment from the original recombinant cosmid isolate. Southern hybridization against *T. ferrooxidans* chromosomal DNA confirmed the origin of a 3,7-kbp *Bgl*II fragment containing the *T. ferrooxidans ntrA* gene.

2.1 Introduction

Analysis of the DNA sequence upstream of the *T. ferrooxidans nifH* gene showed the presence of a putative NtrA-dependent promoter (Pretorius *et al.*, 1987). This indicated that it was likely that *T. ferrooxidans* contained an *ntrA* gene. It was therefore decided to isolate the *T. ferrooxidans ntrA* gene as part of a study on the nitrogen metabolism of this industrially important diazotrophic chemolithoautotroph. This study would form part of the current effort to determine whether gene regulation in general and the regulation of nitrogen fixation genes in particular were similar to that of other Gram-negative diazotrophs, or whether the ability of *T. ferrooxidans* to grow in a harsh environment populated by few bacteria had resulted in genetic drift and unusual features in the structure of one of the key regulators of transcription - NtrA.

The isolation of *ntrA* genes has been greatly facilitated by the construction of *ntrA* mutants in the source organism or a closly related organism. The *ntrA* genes from *K. pneumoniae* (de Bruijn and Ausubel, 1983; Merrick and Stewart, 1985), *A. vinelandii* (Toukdarian and Kennedy, 1986), and *P. putida* (Köhler *et al.*, 1989b) were isolated by complementation of *E. coli ntrA* mutants for the ability to grow on minimal medium with arginine as a sole nitrogen source. This complementation of so-called "Ntr" function entailed derepression of *E. coli* nitrogen assimilation genes such as those for arginine utilization (*aut* operon) by the heterologous cloned *ntrA* gene products.

The *R. meliloti ntrA* gene was isolated by construction of a *R. meliloti ntrA* mutant using Tn5 mutagenesis and selection for the inability to carry out NtrA-dependent transcription of a *dctA*-lacZ fusion (Ronson *et al.*, 1987b). The *R. meliloti ntrA* gene was then isolated from a recombinant cosmid library by complementation of this *R. meliloti ntrA* mutant for growth on succinate as a sole carbon source. The rationale for this was that the cloned *ntrA* gene enabled the *ntrA* mutant to recover NtrA-dependent expression of the *dct*A gene, the structural component of C_4-dicarboxylic acid transport, providing uptake of succinate (as well as malate and fumarate). The *R. meliloti ntrA* mutant was used in a similar manner to clone the *ntrA*

gene from the broad host range *Rhizobium* sp. NGR234 by interspecies complementation (Stanley *et al.*, 1989).

The *ntrA* gene from *P. putida* was cloned by a second group (Inouye *et al.*, 1989) using hybridization with an *E. coli ntrA* gene probe, and the *P. aeruginosa ntrA* gene was cloned using an *A. vinelandii ntrA* gene probe (Ishimoto and Lory, 1989). The locus *nifR4* (*ntrA* homologue) from the purple non-sulfur photosynthetic bacterium *Rhodobacter capsulatus* was identified by sequence analysis of a recombinant clone which complemented a spontaneous Nif$^-$ regulatory mutant of this organism (Jones and Haselkorn, 1989).

The difficulty in cultivating *T. ferrooxidans* in laboratory media containing organic supplements and the absence of a genetic manipulation system for this organism precluded the design of a cloning procedure using phenotypes of the bacterium itself. The ability to identify a recombinant clone from a gene bank which could be propagated in *E. coli* was therefore fundamental to this study. Three methods were employed in attempts to clone the *T. ferrooxidans ntrA* gene:

1) Complementation of an *E. coli ntrA* mutant for Ntr function.
2) Hybridization with the *A. vinelandii ntrA* DNA probe, which was available at the beginning of the project.
3) Complementation of an *E. coli ntrA* mutant for NtrA-dependent expression of the formate degradation pathway.

E. coli cells grown under anaerobic conditions with glucose produce formate which is metabolised via the fermentative formate degradation pathway or via the respiratory nitrate-linked route (reviewed in Stewart, 1988). In the presence of a terminal electron acceptor such as nitrate the formate is respired to generate energy. However, in the absence of such an electron acceptor the fermentative formate hydrogenlyase pathway is induced which results in the evolution of H_2 and CO_2 gas (Peck and Guest, 1957). The formate hydogenlyase complex has not been fully characterized, however it is composed of a formate dehydrogenase, unknown redox compounds, and a hydrogenase (Stewart, 1988). Much debate has accompanied characterization of

E. coli hydrogenase enzymes, however it is now thought that the electrophoretically stable isoenzymes hydrogenase 1 and 2 form part of the respiratory hydrogen uptake (Hup) pathway, while the labile hydrogenase 3 is the gas evolving component of the formate degradation pathway (Sawers *et al.*, 1985).

Birkmann *et al.* (1987a) recently discovered that expression of the *fdhF* gene, which codes for the selenopolypeptide component of the formate dehydrogenase, and the *hyd-17* gene(s), which are required for expression of hydrogenase 3, is NtrA-dependent. DNA sequence analysis of the *fdhF* gene (Zinoni *et al.*, 1986) enabled identification of an NtrA-dependent promoter. The *hyd-17* locus was identified in a phage Mu *d*1 insertion mutant of *E. coli* denoted M17s which exhibited a 90% reduction in hydrogenase activity (Pecher *et al.*, 1983). M17s exhibited no reduction in the hydrogenase 1 and 2 activities as compared to its parent, indicating that the mutation was caused by integration of Mu *d*1 phage into a regulatory or structural gene specific for hydrogenase 3 (Birkmann *et al.*, 1987b).

The region of the *E. coli* chromosome (between 58/59 min) spanning the *hyd-17* locus has recently been cloned and sequenced (Lutz *et al.*, 1990; Böhm *et al.*, 1990). Two divergently transcribed operons with σ54-dependent promoters were identified (Lutz *et al.*, 1990). A DNA fragment containing one of the operons (ORF's 1 - 8) and ORF A and B from the second operon complemented M17s to wildtype levels of hydrogenase 3 activity and gas production. The predicted product of ORF5 showed sequence similarity to the large subunit from Ni/Fe hydrogenases, while the products of the other ORF's in the same operon showed homology to electron transport proteins: ORF2 and ORF6 encoded proteins with iron-sulphur clusters of the 4Fe/4S ferredoxin type; ORF7 product was homologous to protein G of the chloroplast electron transport chain; and ORF3 and ORF4 products contained putative transmembrane hydrophobic regions, and hydrophilic regions homologous to subunits 4 and 1 of mitochondrial and plastid NADH-ubiquinol oxidoreductases, respectively (Böhm *et al.*, 1990).

This operon is therefore thought to encode structural components of the formate hydrogenlyase pathway (ie. electron carriers and hydrogenase 3), but not the formate dehydrogenase (unlinked *fdh* genes)(Böhm *et al.*, 1990). The second operon, of which ORF's A - E have been characterized to date, encodes proteins involved in regulation and assembly of the system. ORFB product is required for nickel utilization (Böhm *et al.*, 1990).

The existence of NtrA-dependent expression of the unlinked *fdhF* gene and both operons at the *hyd*-17 locus lends support to the model that the formate degradation pathway is subject to co-ordinate transcriptional regulation by the binding (upstream of each promoter) of a regulator which interacts with the NtrA-RNA polymerase in a manner similar to the regulators of the *glnA* gene in enteric bacteria (reviewed in Kustu *et al.*, 1986) and the *nifH* gene of *K. pneumoniae* (reviewed in Gussin *et al.*, 1986). A possible two-component regulatory system which may fulfil this role has been identified by Stoker *et al.* (1989): the *hydHG* genes were cloned on the basis of complementation of an *E. coli* mutant defective in hydrogenase 3 activity; and the *hydH* and *hydG* predicted products were homologous to enteric NtrB and NtrC proteins, respectively. Another candidate activator is the product of *fhlA* (Sankar *et al.*, 1988; Schlensog *et al.*, 1989), which is identical to the product of ORFE within the *hyd17* locus (Böhm *et al.*, 1990). It remains to be seen whether ORFD and ORFE are identical to *hydH* and *hydG* (pers. observation). Furthermore, a *cis*-acting DNA element upstream of the *fdhF* gene that is required for regulation of expression from its NtrA-dependent promoter (Birkmann and Böck, 1989) is reminiscent of UAS's (enhancers) bound by activators, such as NtrC and NifA, which act at other NtrA-dependent promoters.

The observation that *E. coli ntrA* mutants were deficient in formate degradation and therefore unable to produce gas was employed to develop a sensitive and novel complementation screening technique for the presence of a cloned *T. ferrooxidans ntrA* gene.

2.2 Materials and Methods

2.2.1 Bacterial strains, plasmids, and media

The strains and plasmids used in this study are described in Appendix A. Complex media for growth of *E. coli* under either aerobic or anaerobic conditions were Luria-Bertani medium (LB), containing 1,5% agar (LBA), or buffered TGYEPF (pH 6,5) (Begg *et al.*, 1977), respectively (Appendix B). Nitrogen-limited minimal medium for aerobic growth of *E. coli* strains was glucose minimal medium (GMM, Appendix B) supplemented with 0,2% arginine. When required, antibiotics were added at the following concentrations : ampicillin, 100 µg/ml; tetracycline, 15 µg/ml; chloramphenicol, 20 µg/ml.

2.2.2 Preparation of DNA

The alkaline lysis method (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981) was used for both small- and large-scale plasmid preparations (Appendix C). Chromosomal DNA was prepared from *T. ferrooxidans* ATCC 33020 by the DNA extraction method as outlined in Appendix C. Construction of the gene bank of *T. ferrooxidans* ATCC 33020 DNA in the cosmid vector pHC79 has been described previously (R. Ramesar, Ph.D. thesis, University of Cape Town, South Africa, 1988).

2.2.3 DNA manipulations

Standard methods (Maniatis *et al.*, 1982) were used for restriction digests, gel electrophoresis, purification of DNA fragments from agarose gels, ligations, and the filling in of 5' sticky ends (Appendix C).

2.2.4 Southern blotting and hybridization with the *A. vinelandii ntrA* gene

T. ferrooxidans chromosomal DNA was digested in 10 µg samples with the restriction enzymes *Bam*HI, *Bgl*II, and *Bcl*I and electrophoresed on a 0,8% agarose gel, followed by transfer to a Hybond N^+ membrane by the method of Reed and Mann (1985)(Appendix C). The *A. vinelandii ntrA* gene on plasmid pNA4 was ^{32}P-labelled by nick-translation as described in Appendix C. Autoradiography was continued for up to two weeks.

2.2.5 Screening for the *T. ferrooxidans ntrA* gene by complementation of an *E. coli ntrA* mutant for Ntr function

The *T. ferrooxidans* ATCC 33020 pHC79 cosmid library was transduced into the *E. coli ntrA* mutant ET8045 as described in Appendix C. The expression mixes were washed in 0,8% NaCl before plating onto GMM containing 0,2% arginine with antibiotic selection. As a positive control the plasmid pFB71 carrying the *K. pneumoniae ntrA* gene was transformed into *E. coli* ET8045 and plated onto the test medium. A proportion of each washed transduction expression mix was plated onto LBA (+ antibiotic) plates to determine the number of transductants. Growth was scored after incubation for 24 - 48 h at 30°C.

2.2.6 Experiments using complementation of the *E. coli* formate degradation pathway to screen for the *T. ferrooxidans ntrA* gene

2.2.6.1 Testing of *E. coli* strains and controls for the gas positive phenotype. Control experiments were carried out to ensure that the *ntrA* mutants available in the laboratory were unable to produce gas and that a visual test for gas positive clones could be developed. *E. coli ntrA* mutant strains TH1, YMC22, and ET8045, as well as an *E. coli ntrA*$^+$ strain YMC10, with and without the cloned *K. pneumoniae ntrA* gene (pFB71) were inoculated into test-tubes containing either liquid TGYEPF medium (Appendix B) with Durham tubes or TGYEPF medium containing 0,8% agar, and then incubated at 35°C under anaerobic conditions for 12 - 24 h. These strains were also plated onto TGYEPF medium and overlaid with TGYEPF containing 0,8% agar (TGYEPF overlay plates) and incubated under the same conditions. This was to test whether single colonies of *ntrA*$^+$ bacteria where identifiable by the formation of a pocket of gas within the agar. Appropriate antibiotic selection was provided in all cases.

2.2.6.2 Screening for the *T. ferrooxidans ntrA* gene by complementation of an *E. coli ntrA* mutant for gas production. The *T. ferrooxidans* ATCC 33020 pHC79 cosmid library was transduced into the *E. coli ntrA* deletion mutant TH1. Transductants were plated onto TGYEPF overlay plates (as described in section 2.2.6.1) with antibiotic selection and incubated anaerobically overnight at 35°C.

Complementation of *ntrA* product activity was identified by a pocket of gas surrounding a colony in the agar.

2.2.6.3 "Shotgun" subcloning of the cosmid pT3. pT3 was digested with a range of restriction enzymes which have 6 bp recognition sites to identify enzymes which had few sites in pT3. Three suitable enzymes - *Eco*RI, *Eco*RV, and *Bam*HI - were chosen for subcloning of pT3 to increase the likelihood of using a restriction enzyme which did not cut within the *ntrA* gene. Samples of pT3 were digested with *Eco*RI, *Eco*RV, and *Bam*HI and ligated to the plasmid vectors pUC19 (*Eco*RI digested), pUC19 (*Sma*I digested), and pUC19 (*BamH*I digested) or pEcoR251 (*Bgl*II or *Bgl*II-*Bam*HI digested), respectively. The ligation mixes were transformed into *E. coli* TH1 and transformants were selected for the gas positive phenotype. Screening of large numbers of clones was carried out most effectively by direct selection on TGYEPF overlay plates, however in cases where there was a high proportion of gas positives it was found to be more efficient to screen by first plating onto LBA (+ antibiotic) plates, followed by toothpicking onto LBA (+ antibiotic) and TGYEPF overlay plates. This bypassed having to repurify a gas positive colony from within a TGYEPF overlay plate containing many other colonies.

2.2.6.4 Subcloning of pT10. All subclones are shown on Fig. 2.3. pT15 was constructed by cloning of a 3,4-kbp *Bam*HI fragment from pT10 into pUC18(*Bam*HI). pT11 and pT12 were *Bgl*II and *Sal*I deletions of pT10, respectively. A 5,5-kbp *Bam*HI fragment from pT10 was cloned into the *Bgl*II site of the vector pEcoR252 to produce pT20. Digestion of pT20 with *Eco*RI and *Sal*I, followed by filling in with Klenow enzyme and blunt-end recircularization ligation produced pT21. pT22 was a *Bgl*II-*Bam*HI deletion derivative of pT21. pT25 and pT26 were the result of insertion of a 3,7-kbp *Bgl*II fragment from pT20 in both orientations in pEcoR252. A 2,8-kbp *Sal*I-*Bgl*II fragment from pT20 was inserted into the *Sal*I-*Bam*HI sites of pUC19 to generate pT29, however loss of the pUC19 *Bam*HI site prevented use of pT29 as a convenient substrate for exonuclease III shortenings from both directions. Consequently, the same 2,8-kbp *Sal*I-*Bgl*II fragment from pT20 was recloned into the *Sal*I site of pUC19 by a two-step ligation process involving filling in with Klenow fragment of the insert *Bgl*II end and one of the vector *Sal*I ends to produce pT30.

2.2.6.5 Southern blotting and hybridization with pT10 against *T. ferrooxidans* chromosomal DNA. *T. ferrooxidans* chromosomal DNA (10 µg) and pT10 plasmid DNA (0,5 µg) were digested with *Bgl*II and electophoresed together with λDNA(*Pst*I) as a molecular weight marker (Appendix D) on a 0,8% agarose gel. This was transferred to a Hybond N^+ membrane by the method of Reed and Mann (1985)(Appendix C). pT10 and λDNA (Boeringher Mannheim) were ^{32}P-labelled by nick-translation (Appendix C). The section of the filter containing the λDNA (*Pst*I) was sliced off and hybridization as described in Appendix C was carried out seperately. The autoradiograph was exposed after 20 h.

2.3 Results

2.3.1 An attempt to clone the *T. ferrooxidans ntrA* by complementation of an *E. coli ntrA* mutant for Ntr function

The *T. ferrooxidans* cosmid bank was transduced into the *E. coli ntrA* mutant ET8045 and transductants were plated onto GMM containing arginine as a sole nitrogen source. Plating of a proportion of the transductants on LBA + Ap demonstrated that 3000 transductants were screened on the minimal medium. ET8045(pFB71) cells grew as large colonies on the GMM (+ arginine). No positive clones were identified.

2.3.2 An attempt to identify the *T. ferrooxidans ntrA* by hybridization with an *A. vinelandii ntrA* probe

No positive hybridization signal was obtained between a ^{32}-P-labelled *A. vinelandii ntrA* DNA probe and *T. ferrooxidans* chromosomal DNA digested in 10μg samples with the restriction enzymes *Bam*H1, *Bgl*II, and *Bcl*I (data not shown).

2.3.3 Testing of *E. coli* strains for the gas positive phenotype

E. coli ntrA mutant strains TH1, ET8045, and YMC22 were unable to produce gas when grown anaerobically in TGYEPF medium. These strains did not produce a bubble in the Durham tubes or cracks in the 0.8% agar (Fig. 2.1). In contrast, *E. coli ntrA*$^+$ strain YMC10 (which is the parent of the *ntrA* mutant strains) and *E. coli* TH1, ET8045, or YMC22 containing the cloned *K. pneumoniae ntrA* gene produced a gas bubble in the Durham tubes and caused dramatic cracks in the 0.8% agar (Fig. 2.1). These control strains were also used to demonstrate that plating onto a TGYEPF overlay plate was a simple and efficient manner to screen for gas positive clones (Fig. 2.2).

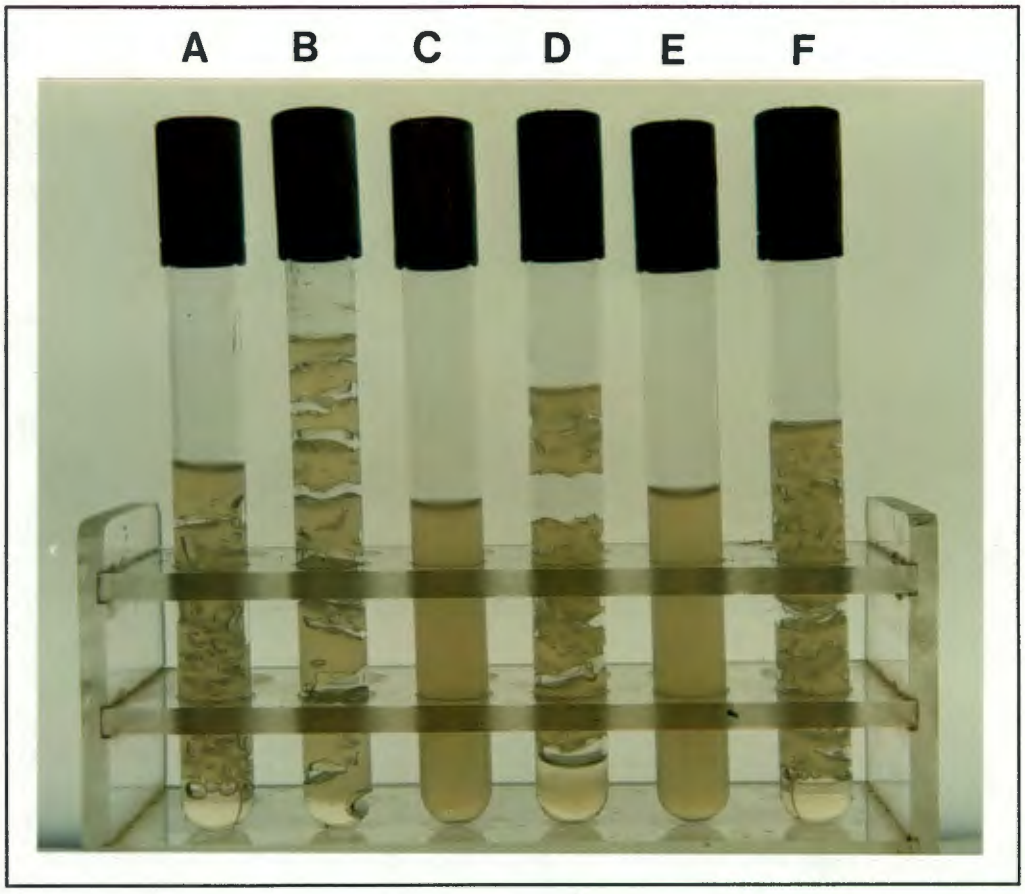


Fig. 2.1. Test of *E. coli* strains for the gas$^+$ phenotype. Strains were inoculated into TGYEPF medium containing 0,8% agar and grown anaerobically at 37°C for 15 h. Tubes: (A) YMC10; (B) YMC10(pFB71); (C) TH1; (D) TH1(pFB71); (E) YMC22; (F) YMC22(pFB71).

2.3.4 Isolation of the *T. ferrooxidans ntrA* gene by complementation of an *E. coli ntrA* mutant for gas production

The *T. ferrooxidans* cosmid bank was transduced into the gas-negative *E. coli ntrA* mutant TH1 and four gas producing colonies were isolated out of 2000 transductants screened. Transformation of cosmid DNA from each colony into *E. coli* TH1 confirmed the gas positive phenotype. Restriction enzyme fragments common to all four cosmid clones were observed and showed that the cloned inserts were greater than 40-kbp in size (results not shown). One clone, pT3, was chosen for further analysis.

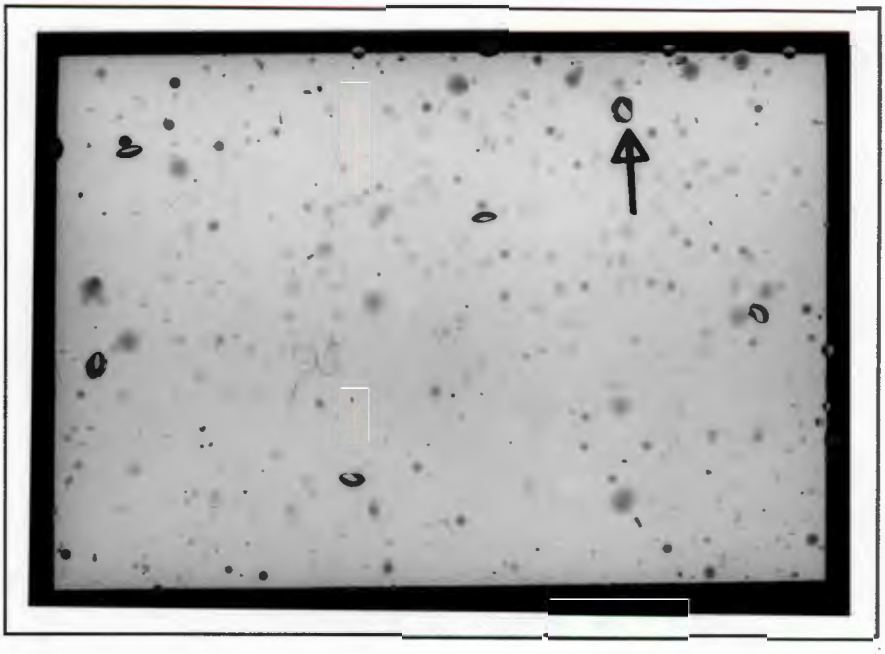


Fig. 2.2. Illustration of the screening technique for identifying gas positive colonies in a TGYEPF overlay plate. This photograph is a closeup of a plate containing a mixture of gas$^+$ and gas$^-$ *E. coli* colonies. The arrow indicates a gas pocket within the agar produced by a colony of gas$^+$ cells.

2.3.5 Subcloning of pT3 to localize the *T. ferrooxidans ntrA* gene

"Shotgun" subcloning of pT3 produced 10 subclones which conferred the gas positive phenotype on *E. coli* TH1 (see section 2.2.6.3). The smallest subclone pT10 was chosen for further work, and a restriction map of this subclone was compiled (Fig. 2.3). pT10 was constructed by ligation of a mixture of *Eco*R1 digestion products of pT3 and the vector pUC19 (*Eco*RI), however it was observed that pT10 contained the original pHC79 cosmid vector and not the pUC19 vector. As can be seen on Fig. 2.3, it consisted of two *Eco*R1 fragments of 8,8-kbp (containing the vector pHC79) and 6,8-kbp. pT10 therefore arose from either a pT3 *Eco*RI partial digest and recircularization event, or from the recloning of a non-contiguous 6,8-kbp *Eco*RI fragment by the *Eco*RI fragment spanning the cosmid vector pHC79. It is not clear as to whether these *Eco*R1 fragments were adjacent in the original cosmid clone pT3, however further work in this project was carried out only with sequences internal to the 6,8-kbp *Eco*RI fragment.

Several subclones of pT10 (Fig. 2.3 - see section 2.2.6.4 for descriptions of constructs) were tested for the gas positive phenotype in *E. coli* TH1. Plasmids pT11, pT12, and pT15 tested negative. pT20 tested positive, which enabled delineation of a 2,8-kbp *Sal*I-*Bgl*II fragment which encoded the ability to confer the gas positive phenotype on *E. coli* TH1. This fragment was subcloned into the vector pUC19 to produce pT30 in preparation for DNA sequencing.

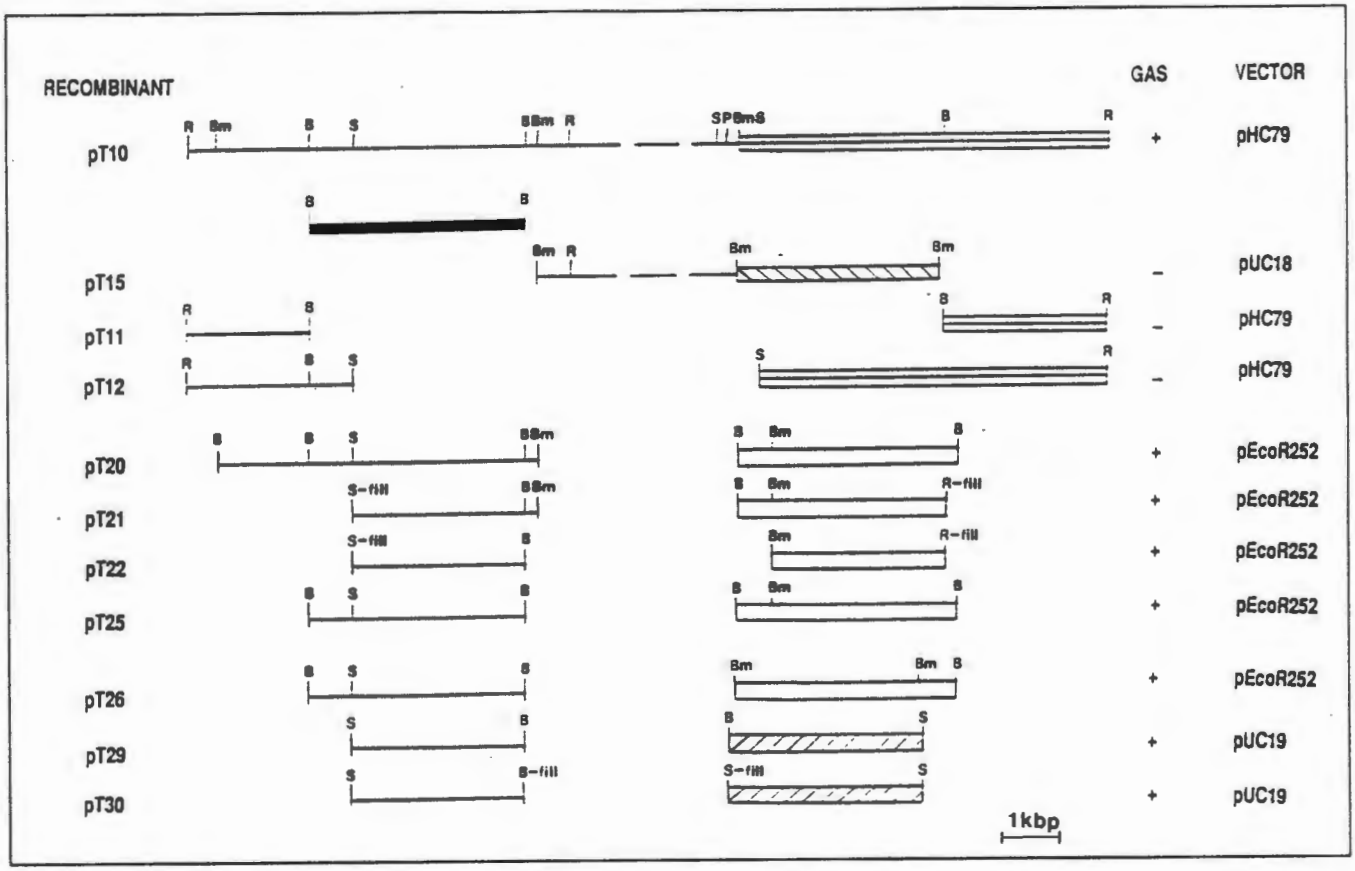


Fig. 2.3. Restriction Map and subclones of pT10. The full extent of both insert and vector sequences are shown for each plasmid subclone. *T. ferrooxidans* DNA is indicated by single lines, while vector DNA is indicated by boxes. The name given to each subclone is indicated on the left hand side, while the vector name as well as the gas positive (+) or gas negative (-) phenotype of each subclone in the *E. coli ntrA* mutant strain TH1 are indicated on the right hand side. The two *Eco*RI fragments of pT10 are indicated by a thin line (6,8-kbp fragment containing only *T. ferrooxidans* DNA and the coding sequences analyzed further in this study), and a dashed line plus a box (8,8-kbp *Eco*RI fragment containing *T. ferrooxidans* DNA + pHC79 vector sequences) to show that these 2 *Eco*RI fragments may not be contiguous in the original cosmid clone pT3. The solid box indicates the extent of the 3,7-kbp *Bgl*II fragment which gave a positive hybridization signal between pT10 and *T. ferrooxidans* chromosomal DNA (section 2.3.6). B, *Bgl*II; Bm, *Bam*HI; P, *Pst*I; R, *Eco*RI; S, *Sal*I

2.3.6 Southern hybridization of pT10 against *T. ferrooxidans* chromosomal DNA

A 3,7-kbp *Bgl*II fragment (shown on Fig. 2.3) internal to the 6,8-kbp *Eco*RI fragment of pT10 (and containing the minimal fragment encoding the gas positive phenotype) hybridized to a *T. ferrooxidans* chromosomal fragment of the same size (Lane B; Fig. 2.4), confirming the origin of this fragment. The two other hybridization signals at >20-kbp and 11-kbp in the *T. ferrooxidans Bgl*II digest (Lane B; Fig. 2.4) represent hybridization between pT10 and the two *T. ferrooxidans* chromosomal *Bgl*II fragments which flank the 3,7-kbp *Bgl*II fragment. pHC79 did not hybridize to *T. ferrooxidans* chromosomal DNA (result not shown).

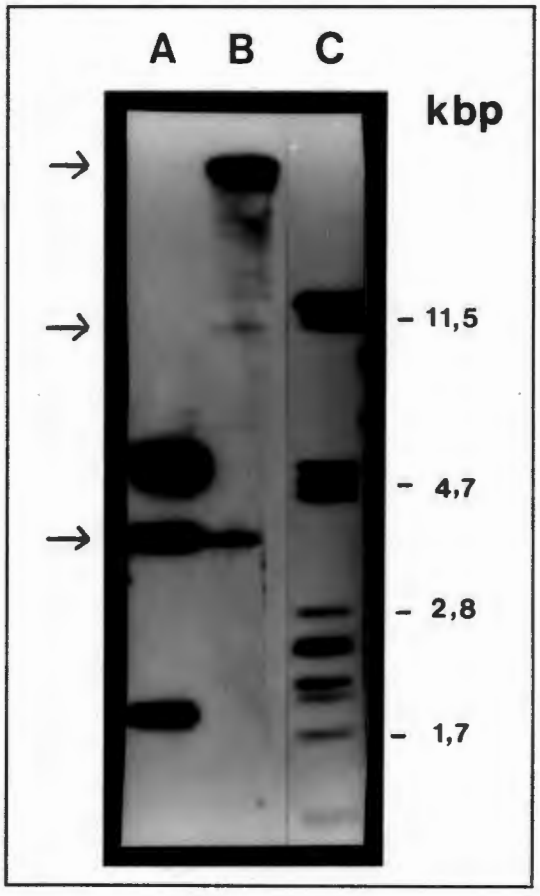


Fig. 2.4. Hybridization of pT10 against *T. ferrooxidans* chromosomal DNA. Autoradiograph of ^{32}P-labelled pT30 hybridized to pT30(*Bgl*II)(lane A) and T. ferrooxidans chromosomal DNA(*Bgl*II)(lane B). Lane (C) shows ^{32}P-labelled λDNA hybridized seperately to λDNA(*Pst*I) as a molecular weight marker. Arrows indicate the positive hybridization signals in lane B [*T. ferrooxidans* chromosomal DNA(*Bgl*II)].

2.4 Discussion

Techniques which were used for the isolation of the *ntrA* genes from *K. pneumoniae* (de Bruijn and Ausubel, 1983; Merrick and Stewart, 1985), *A. vinelandii* (Toukdarian and Kennedy, 1986), and two *Pseudomonas* strains (Köhler *et al.*, 1989; Inouye *et al.*, 1989; Ishimoto and Lory, 1989) proved to be unsuccessful when applied to *T. ferrooxidans*. This required the development of a novel strategy employing complementation of an *E. coli ntrA* mutant for expression of another NtrA-dependent phenotype: gas production under anaerobic growth conditions in the presence of formate. This method was used successfully to isolate overlapping recombinant clones from a *T. ferrooxidans* cosmid gene library which were able to confer the gas positive phenotype on an *E. coli ntrA* mutant. A selected cosmid clone was subcloned down to a minimal active fragment to enable demonstration that this complementation was due to the product of the cloned *T. ferrooxidans ntrA* gene (see Ch. 3 and Ch. 4).

The reason why no clones were isolated by complementation of an *E. coli ntrA* mutant for the ability to grow on minimal medium with arginine as a sole nitrogen source became clear when the putative *T. ferrooxidans ntrA* gene was tested for this complementation once it had been cloned. Several clones of the putative *T. ferrooxidans ntrA* gene carried on different vectors (pT3, pT20 - Fig. 2.3; pT40, pT41, pT50 - Ch. 4) were transformed into the *E. coli ntrA* mutant and plated onto minimal medium containing arginine. The amount of growth obtained was equivalent to that shown by the *E. coli ntrA* mutant transformed with vector DNA and much less than the parent *E. coli* YMC10 strain or an *E. coli ntrA* mutant carrying a cloned *K. pneumoniae ntrA* gene (results not shown).

Although expression of these cloned genes may differ, this indicates that the putative *T. ferrooxidans ntrA* gene product is not as functionally related to the *E. coli ntrA* gene product as are the *ntrA* gene products from *K. pneumoniae*, *A. vinelandii*, and *P. putida* which complemented *E. coli ntrA* mutants for Ntr function. The basis of this complementation is that these heterologous *ntrA* gene products are able to direct transcription from the NtrA-dependent promoters of the *E. coli* arginine utilization (*aut*) system. Nevertheless, the putative *T. ferrooxidans ntrA* gene product was able to replace the *E. coli ntrA* gene product activity with respect to transcription of the

NtrA-dependent promoters of the formate degradation pathway as this formed the basis of the cloning strategy. A possible explanation is that arginine utilization is tested on minimal medium and is essential for growth whereas gas production is tested on a rich medium and is not required for growth. The putative *T. ferrooxidans* NtrA may promote transcription from the *E. coli* NtrA-dependent arginine utilization promoter(s) too weakly to permit growth on arginine. However, the sensitivity of detecting gas bubble formation enabled identification of transductants in which transcription of the *E. coli fdhF* promoter was directed by the putative *T. ferrooxidans* NtrA.

The inability to detect the *T. ferrooxidans ntrA* gene by hybridization with an *A. vinelandii ntrA* gene DNA probe was also clarified once the DNA sequence of the *T. ferrooxidans ntrA* gene had been determined (Ch. 3). As the codon usage and GC bias differs between bacteria it is clear that functional similarities between two proteins which may be reflected in amino acid sequence identities may not necessarily be reflected in DNA sequence identities. Alignment of the DNA sequence of the *T. ferrooxidans ntrA* gene (Ch. 4) with that of the *A. vinelandii ntrA* gene showed 42% identity (data not shown) using the University of Wisconsin Genetics Computer Group (UWGCG) GAP program. Identity was therefore too low to enable detection under the hybridization conditions used. Indeed, it was reported by Ronson *et al.*, (1987b) that no positive signals were obtained on screening *R. meliloti* DNA with a *K. pneumoniae ntrA* gene probe where there is 38% identity at the DNA level between *ntrA* genes. The DNA sequence of the *P. aeruginosa ntrA* gene is not available, however it is likely to have very high DNA identity to the *P. putida ntrA* gene which shows 80% identity at the DNA level to the *A. vinelandii ntrA* gene (the probe used to isolate the *P. aeruginosa ntrA* gene, Ishimoto and Lory, 1989).

The simplicity and sensitivity of the TGYEPF overlay technique for screening gene libraries for heterologous *ntrA* genes in *E. coli* indicates that it has great potential for the isolation of *ntrA* genes from other bacteria. This is of importance as current interest in NtrA is increasing (note that four papers on the cloning of *ntrA* genes appeared in 1989) due to the realisation that NtrA is central to the transcription of genes involved in a wide range of physiological processes in many bacteria (see section 1.4).

CHAPTER 3

DNA SEQUENCE OF THE *T. FERROOXIDANS NTRA* GENE AND AMINO ACID COMPARISONS OF THE GENE PRODUCT WITH OTHER NTRA PROTEINS

CHAPTER 3

DNA SEQUENCE OF THE *T. FERROOXIDANS NTRA* GENE AND AMINO ACID COMPARISONS OF THE GENE PRODUCT WITH OTHER NTRA PROTEINS

3.0 Summary

The nucleotide sequence of a 2,8-kbp *T. ferrooxidans* DNA fragment from the plasmid pT30, which conferred the gas positive phenotype on the *E. coli ntrA* mutant strain TH1, was determined. Three open reading frames (ORF's) were identified. The central ORF corresponded to the *T. ferrooxidans ntrA* gene and encoded a predicted translation product of 475 amino acids (aa) (calculated M_r 52972). The *T. ferrooxidans* NtrA protein had 51, 50, 49, 40, 28% aa sequence identity with the NtrA proteins of *K. pneumoniae*, *P. putida*, *A. vinelandii*, *R. meliloti*, and *R. capsulatus*, respectively. An ORF coding for a protein of 241 aa (calculated M_r 27023) was situated 12 bp upstream of the *T. ferrooxidans ntrA* gene. This protein showed 57% aa identity with the product of the ORF1 located upstream of the *R. meliloti ntrA* gene. Downstream of the *T. ferrooxidans ntrA* gene the front end of an ORF, called ORF3, was identified. The predicted 78 N-terminal aa encoded by ORF3 showed 38, 38, 29, and 20% aa identity with conserved ORF's immediately downstream of the *A. vinelandii*, *P. putida*, *K. pneumoniae*, and *R. meliloti ntrA* genes, respectively.

3.1 Introduction

Nucleotide sequencing is a powerful tool in understanding the genetic determinants responsible for a phenotype. The nucleotide sequences of *ntrA* genes from five bacterial species have been determined to date. The DNA sequences of *ntrA* genes and flanking regions from *K. pneumoniae* (Merrick and Gibbons, 1985; Merrick and Coppard, 1989), *A. vinelandii* (Merrick *et al.*, 1987; Merrick and Coppard, 1989), *R. meliloti* (Ronson *et al.*, 1987b; Albright *et al.*, 1989b), *R. capsulatus* (Jones and Haselkorn, 1989; Alias *et al.*, 1989), and *P. putida* (Inouye *et al.*, 1989; Köhler *et al.*, 1989a) have been made available in the EMBL\Genbank\DDBJ databases under the accession numbers X103147\X16335, X05888, M16513\M24926, X12358\X15437, and M24916\X16474, respectively.

Analysis of this sequence data for ORF's has been very useful in determining the size of *ntrA* genes and identifying possible flanking genes. This has been especially useful because due to the acidic nature of NtrA proteins they have been found to migrate anomalously when analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The *K. pneumoniae* NtrA, for example, has an actual M_r of 54 kDal but it migrated at the position of a 75 kDal protein on SDS-PAGE (Merrick and Gibbons, 1985).

The predicted translation products of these five *ntrA* genes have been determined and this has enabled identification of conserved regions within NtrA proteins which may represent functional domains (Merrick *et al.*, 1987). The availability of aa databases and computer based search and alignment programs has been valuable, especially with respect to ascribing possible functions to these conserved regions and to those proteins which are encoded by the ORF's linked to *ntrA* genes.

The presence of conserved ORF's upstream and downstream of *ntrA* genes appears to be true for at least *K. pneumoniae*, *P. putida*, and *R. meliloti* (Merrick and Gibbons, 1985; Merrick and Coppard, 1989; Inouye *et al.*, 1989; Ronson *et al.*, 1987b), while in *A. vinelandii* the downstream ORF is conserved (Merrick *et al.*, 1987).

The function of the product of the upstream ORF1, which has been studied in *R. meliloti*, is not clear; however the inability to construct chromosomal mutants in this locus and the absence of transcriptional coupling between ORF1 and *ntrA* indicate that ORF1 may encode a protein involved in housekeeping functions which is not essential for NtrA function (Albright *et al.*, 1989b). The absence of transcriptional stop signals between *ntrA* and the downstream ORF in *K. pneumoniae*, *P. putida*, *R. meliloti*, and *A. vinelandii* indicates that this may form part of an operon. Merrick and Coppard (1989) have shown that at least in *K. pneumoniae* there are two ORF's downstream of the *ntrA* gene which are involved in modulation of NtrA-RNA polymerase activity. The context of the *R. capsulatus nifR4* (*ntrA*) gene which is downstream of the *nifHDK* genes and upstream of the *nifA* gene appears to be different from *ntrA* genes in the other bacteria studied (Jones and Haselkorn, 1989).

Nucleotide sequence data can also provide clues as to regulatory regions which may be involved in the expression of the *ntrA* gene itself. Previous studies had indicated that the *ntrA* gene in *E. coli* (Castano and Bastarachea, 1984), *K. pneumoniae* (de Bruijn and Ausubel, 1983; Merrick and Stewart, 1985), and *R. meliloti* (Ronson *et al.*, 1987b) was constitutively expressed at a low level and not subject to regulation by levels of available nitrogen.

In addition, the following possible regulatory elements have been recognized: -35 and -10 promoter-like regions and a weak ribosome binding site 5' to the *K. pneumoniae ntrA* gene (Merrick and Gibbons, 1985); a -10 promoter-like region and the absence of a consensus ribosome binding site 5' to the *A. vinelandii ntrA* gene (Merrick *et al.*, 1987); a -35 promoter-like region and poor consensus to the *E. coli* -10 promoter-like region 5' to the start of the mapped *R. meliloti ntrA* gene transcript (Albright *et al.*, 1989b). In contrast, expression of a *R. capsulatus nifR4*::*lacZ* fusion was increased 10-fold in the absence of ammonia and oxygen (Kranz and Haselkorn, 1985) and no consensus promoter-like sequences could be identified in the sequence 5' to the *nifR4* gene (Jones and Haselkorn, 1989).

In this study the nucleotide sequence of the 2,8-kbp *T. ferrooxidans* DNA insert in pT30 was determined to confirm that the *T. ferrooxidans ntrA* gene was responsible for complementation of the gas positive phenotype in the *E. coli ntrA* mutant TH1. The context of the *T. ferrooxidans ntrA* gene with respect to linked ORF's was elucidated. The predicted translation product of the *T. ferrooxidans ntrA* gene was compared at the level of primary aa sequence to the five NtrA protein sequences available. Conserved regions, which may have functional significance, were identified.

3.2 Materials and Methods

3.2.1 Bacterial strains and plasmids

E. coli strain LK111 (Appendix A) (Zabeau and Stanley, 1982) was used for propagation, subcloning, and exonuclease III shortening of DNA inserts in the vectors pUC18, pUC19 (Norrander *et al.*, 1983) and Bluescript SK$^+$ and KS$^+$ (Stratagene, San Diego) (Appendices A, F).

3.2.2 Media, buffers, and enzymes

All media and buffers not described in the text are given in Appendix B. Restriction endonucleases, T4 DNA ligase, and S1 nuclease were purchased from Boehringer Mannheim Biochemicals. Exonuclease III was obtained from Bethesda Research Laboratories.

3.2.3 Sequencing strategy

The plasmid pT30 (see Ch. 2) was used to generate overlapping deletions from both ends of the insert, using the exonuclease III shortening technique (Henikoff, 1984, 1987). Some templates for sequencing were generated by subcloning to confirm areas of sequence obtained from exonuclease III templates or to cover gaps (Fig. 3.1). The whole insert of pT30 was sequenced at least once on each strand.

3.2.4 Exonuclease III shortening of pT30 and cloning of shortened fragments

pT30 digested with either of the restriction endonuclease pairs *Sac*I-*BamH*I or *Sal*I-*Sph*I served as templates for unidirectional digestion using exonuclease III by an adaptation of the method of Henikoff (1984, 1987). The template created by the *Sac*I-*BamH*I enzyme pair enabled shortening from the *Bgl*II end of the *T. ferrooxidans* insert in pT30, while the template produced by the *Sal*I-*Sph*I enzyme pair was used to shorten from the other end (*Sal*I). After digestion, the linearised plasmid (10 µg) was precipitated with isopropanol, resuspended in Exo-buffer (100 µl) (Appendix B), and equilibrated at 37°C (5 min). Eleven microfuge tubes containing ice cold S1 nuclease mixture (25 µl) (Appendix B) were prepared immediately before starting the shortening reaction. A sample (9 µl) of linearised plasmid in Exo buffer was removed to a microfuge tube containing S1 nuclease mixture before the shortening reaction

was initiated by the addition of exonuclease III (300 units). Samples (9 µl) were removed at 30 s intervals and added to the 10 remaining microfuge tubes containing the S1 nuclease mixture. The microfuge tubes were incubated at room temperature for 30 min while the S1 nuclease digested any single stranded DNA present. The S1 nuclease reaction was stopped by the addition of S1 stop solution (3,4 µl) (Appendix B) to each microfuge tube and incubation at 70°C for 10 min. The extent of shortening was checked by electrophoresis of approximately 200 ng DNA (8 µl) from every second tube on an agarose gel (0,8% w/v) in Tris-acetate buffer. The exonuclease III-generated ends were filled in by the addition of DNA polymerase I (Klenow) (0,5 units per tube) in Klenow buffer (Appendix B), incubation at room temperature for 3 min, followed by a further incubation of 5 min in the presence of a mixture of dNTP's (0,125 mM each, A, C, G, and T). The shortened DNA was religated by the addition of ligation mixture (120 µl) (Appendix B) to each tube, at 15°C for 4 h. Competent *E. coli* LK111 cells were transformed with the ligation mixtures and selection was on LBA plates containing Ap (100 µg/ml).

3.2.5 Construction of subclones for nucleotide sequencing

Subclones (Fig. 3.1) were constructed using suitable restriction endonuclease sites which were identified during the course of the sequencing project. pT30-C3P was constructed by deletion of a 213 bp *Pst*I fragment (from the *Pst*I site within the pUC19 multiple-cloning-site to the *Pst*I site at position 203 on Fig. 3.1) from the exonuclease III shortened clone pT30-C3 (Fig. 3.1). The two *Cla*I restriction endonuclease sites within pT30 were found to be subject to methylation and therefore resistant to digestion if the plasmid was isolated from a *dam*$^+$ *E. coli* strain such as LK111 (Backman, 1980). This necessitated the transformation of plasmids pT30-D and pT30-C (Fig. 3.1) into the *dam*$^-$ *E. coli* strain GM41 (Appendix A) and reisolation of unmethylated plasmid DNA.

pT40 and pT41 were constructed by subcloning of a 1,69-kbp *Cla*I(at position 802 on Fig. 3.1)-*Eco*RI(within the multiple-cloning-site of pUC19) fragment from the exonuclease III shortened clone pT30-D (Fig. 3.1) into the *Cla*I, *Eco*RI sites of Bluescript KS$^+$ and Bluescript SK$^+$, respectively. Interestingly, *E. coli* LK111 (pT40) cells grew as blue colonies on YTA + Xgal plates (Appendix B). Observation of the

DNA sequence of pT40 indicated that this was most probably due to a fortuitous translational fusion between an ORF containing an ATG start codon which could be seen out of frame with and 30 bp 5' to the stop codon of the *T. ferrooxidans ntrA* gene, but in frame with the *lacZ* gene. The 552 bp *Cla*I fragment from pT30-D was inserted into the *Cla*I site of Bluescript SK$^+$ to generate pT30-DC (Fig. 3.1).

pT30-CV was the result of insertion of a 180 bp *Cla*I(position 802 on Fig. 3.1)-*Nco*I(position 982 on Fig. 3.1) fragment into the *Cla*I, *Eco*RV sites of Bluescript SK$^+$. LK111(pT30-CV) cells also grew as blue colonies on YTA + Xgal plates. DNA sequencing revealed that in the process of construction of pT30-CV the insert *Nco*I end had been filled-in with Klenow and ligated to the vector *Cla*I end, the 5'-overhang of which had been flushed off by Klenow, while the 5'-overhang of the insert *Cla*I end had also been flushed off enabling ligation to the vector *Eco*RV end. The filled-in *Nco*I site resulted in an ATG codon, which fortuitously was in frame with the *lacZ* gene, therefore probably giving rise to the *lac* positive phenotype.

3.2.6 Nucleotide sequencing

Recombinant plasmids resulting from either exonuclease III shortening, or from subcloning (Fig. 3.1), were analyzed by restriction endonuclease mapping before the preparation of CsCl purified plasmid DNA for sequencing (Appendix C). The preparation of template DNA, primer annealing, sequencing reactions, and the gel electrophoresis and autoradiography, are described in Appendix C.

3.2.7 Sequence analysis

The DNA and aa sequences were analyzed with the IBM XT computer Genepro (Version 4.1) programme and a VAX computer using the UWGCG sequence analysis package version 6.1 (Devereux, 1984) and associated databases. A table of one- and three- letter codes for aa is given in Appendix E.

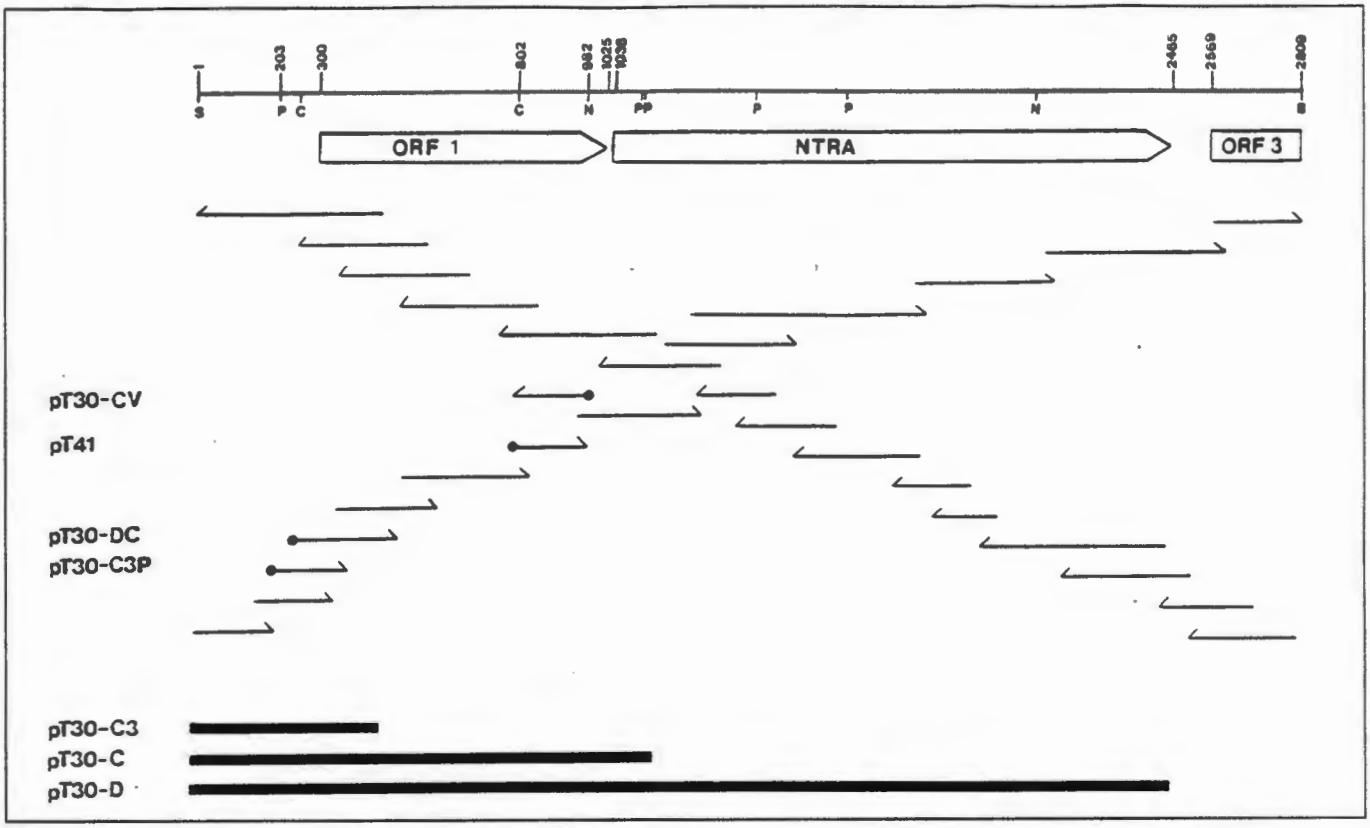


Fig. 3.1. Sequencing strategy for determination of the nucleotide sequence of the *T. ferrooxidans* DNA insert in pT30. The single line represents the 2,8-kbp *Sal*I-*Bgl*II *T. ferrooxidans* DNA insert in pT30, the complete nucleotide sequence of which was determined. Numerals refer to nucleotides from the *Sal*I site and match those in Fig. 3.2. Proposed ORF's are indicated by open boxes. The arrows indicate the extent of sequence obtained from each deletion clone produced by the method of Henikoff (1987) using exonuclease III. Those regions where sequence data was obtained from subclones are indicated by an arrow preceded by a filled circle. The extent of the *T. ferrooxidans* DNA inserts in the exonuclease III generated deletion clones pT30-C, pT30-D, and pT30-C3 used for the production of the subclones are shown by filled boxes. B, *Bgl*II; C, *Cla*I; N, *Nco*I; P, *Pst*I; S, *Sal*I.

3.3 Results

3.3.1 Nucleotide sequence of the *T. ferrooxidans ntrA* gene and flanking regions

The nucleotide sequence of the 2,8-kbp *SalI-BglII* *T. ferrooxidans* DNA insert in pT30 is shown in Fig. 3.2. Analysis of the sequence data revealed three ORFs (Fig. 3.1).

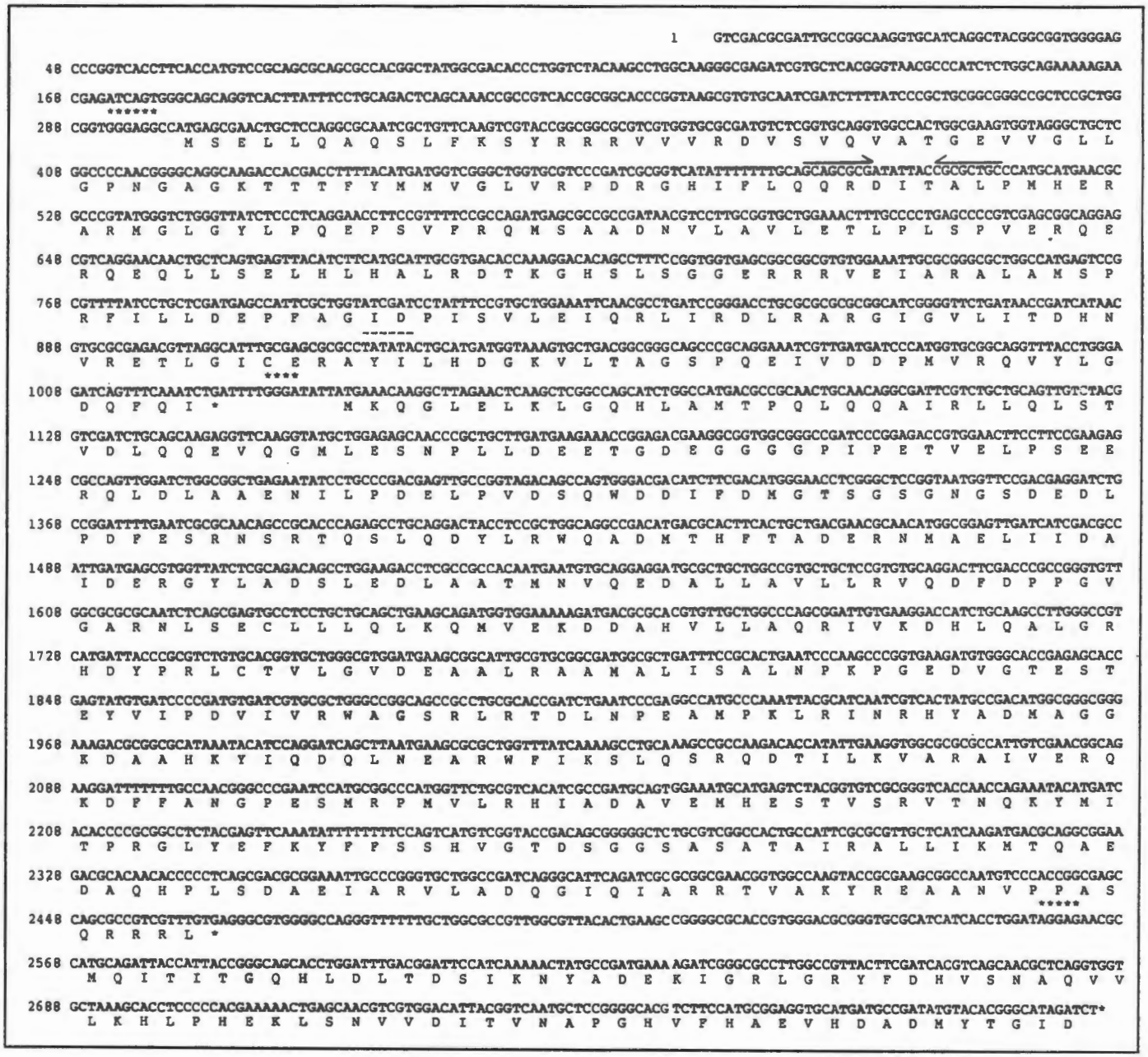


Fig. 3.2. Nucleotide sequence of the 2,8-kbp *SalI-BglII* DNA insert in pT30 containing the *T. ferrooxidans* ORF1, *ntrA*, and ORF 3, with predicted translation products. Nucleotides are numbered from the first base of the *SalI* site. Possible ribosome-binding sites are indicated with asterisks. A putative -10 promoter sequence for *ntrA* is shown by a dashed overline. The inverted repeat sequence at position 491 is indicated by arrows. These data have been assigned the Genbank accession number M33831.

Evidence for the authenticity of these ORF's as coding for functional proteins was provided by analysis of the sequence data using the UWGCG application TESTCODE which aids identification of protein-coding ORF's by plotting a measure of the non-randomness of the nucleotide composition at every third base. TESTCODE uses a statistic based on measurements of the period three compositional constraints in the entire Genbank database for regions thought to be coding and non-coding (Fickett, 1982). TESTCODE produces a graph which predicts coding regions to a 95% level of confidence, represented by any plot above a level of 9.5 TESTCODE units. As can be seen on Fig. 3.3 there is a plot above this level corresponding to each ORF indicating a bias in third base nucleotide composition throughout each ORF.

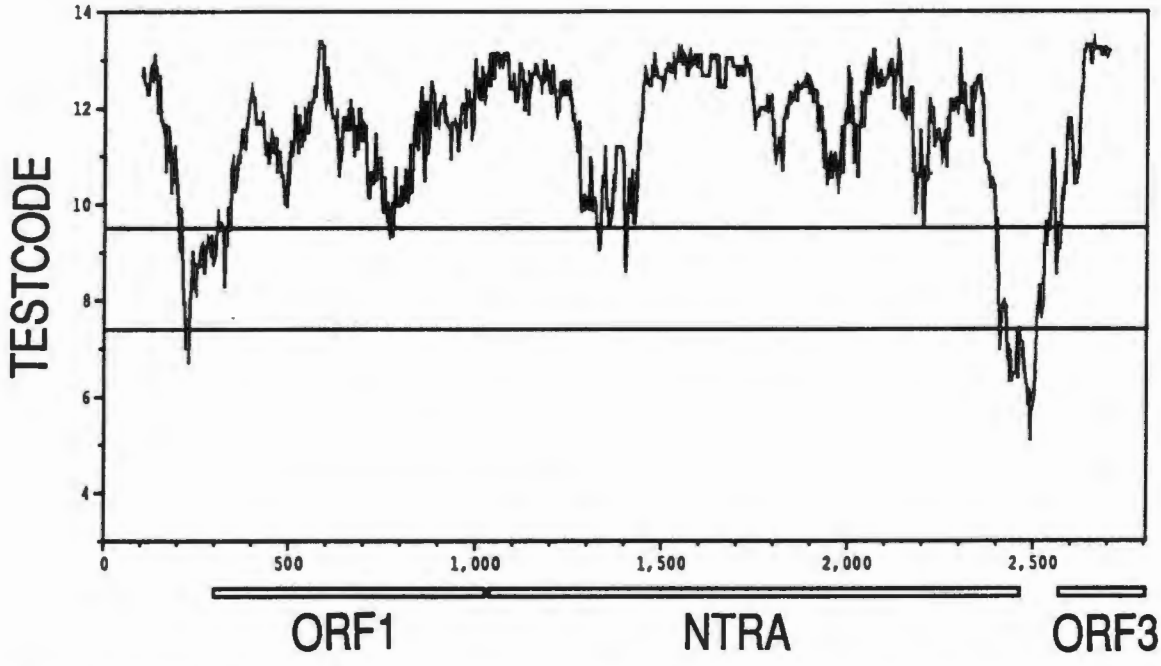


Fig. 3.3. TESTCODE analysis of the 2,8-kbp DNA insert in pT30. The horizontal axis represents the 2,8-kbp *T. ferrooxidans* DNA insert with the ORF's predicted from the nucleotide sequence (ORF1, NTRA, ORF3) shown as open boxes below. The vertical axis is in TESTCODE units, with two horizontal lines demarcating the three "windows" of probability for protein-coding sequences:- < 7.5 units : low probability; 7.5-9.5 units: intermediate probability; > 9.5 units: 95% probability.

The central ORF (1428 bp) coded for an acidic protein with a calculated M_r of 52927. The predicted aa sequence of this ORF was aligned with the *K. pneumoniae* (Merrick and Gibbons, 1985), *A. vinelandii* (Merrick *et al.*, 1987), *R. meliloti* (Ronson *et al.*, 1987b), *R. capsulatus* (Jones and Haselkorn, 1989; Alias *et al.*, 1989), and *P. putida* (Inouye *et al.*, 1989; Köhler *et al.*, 1989a) *ntrA* gene products (Fig. 3.4).

Regions of aa sequence homology, reported previously for ntrA gene products (Merrick *et al.*, 1987), were observed (see section 3.3.2) and identified this ORF as the *T. ferrooxidans ntrA* gene. The percentage similarity between the various *ntrA* gene products is shown in Table 3.1.

Table 3.1. Percent similarity of *ntrA* predicted products. Numerals refer to % identical amino acids in pairwise alignments using the UWGCG program GAP.

Bacterium[a]	T.f.	K.p.	P.p.	A.v.	R.m.
T.f.	-				
K.p.	51	-			
P.p.	50	55	-		
A.v.	49	58	83	-	
R.m.	40	39	37	38	-
R.c.	28	29	34	32	34

[a] T.f., *T. ferrooxidans*, K.p., *K. pneumoniae*, P.p., *P. putida*, A.v., *A. vinelandii*, R.m., *R. meliloti*, R.c., *R. capsulatus*.

The acidic nature of the *T. ferrooxidans ntrA* gene product, which is a feature of NtrA proteins, is demonstrated in Table 3.2. A weak ribosome binding site GGGA was present at position 1030 (Fig. 3.2). There were no clearly identifiable -35 promoter sequences although a putative -10 promoter sequence (TATATA) at position 921 was detected (Fig. 3.2).

3.3.2 Amino acid sequence similarity between the *T. ferrooxidans* NtrA and other NtrA proteins

Merrick *et al.* (1987) compared the NtrA proteins from *A. vinelandii* (AvNtrA), *K. pneumoniae* (KpNtrA), and *R. meliloti* (RmNtrA) and divided the proteins into three regions based on aa sequence conservation. This analysis has been extended to cover the more recently sequenced *P. putida* NtrA (PpNtrA), which shows 83% identity to the AvNtrA, and the least conserved *R. capsulatus* NtrA (RcNtrA) (Fig. 3.4). All three regions are represented in the *T. ferrooxidans* NtrA (TfNtrA).

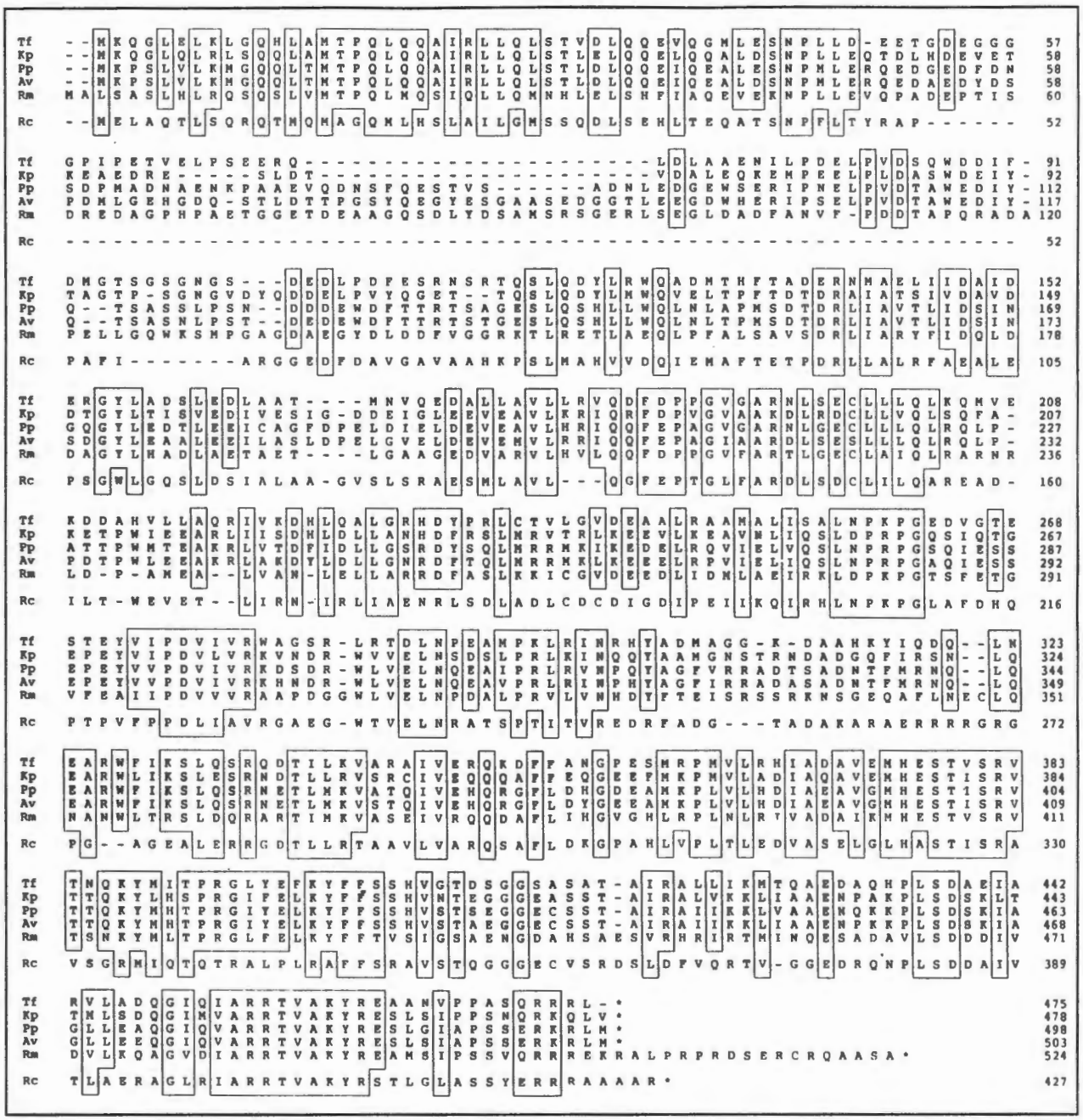


Fig. 3.4. Alignment of the deduced amino acid sequence of the *T. ferrooxidans* NtrA with other NtrA proteins. Amino acids are identified by the single letter code and regions which are conserved (see Appendix E for scheme of conservative substitutions) in at least five NtrA proteins are boxed. Numerals refer to aa positions in each NtrA protein. Tf, *T. ferrooxidans* NtrA; Kp, *K. pneumoniae* NtrA; Pp, *P. putida* NtrA; Av, *A. vinelandii* NtrA; Rm, *R. meliloti* NtrA; Rc, *R. capsulatus* NtrA.

Region 1 encompasses the first 48 aa of the NtrA proteins. The N-terminal 48 aa of the TfNtrA shows 77, 73, 75, 42 and 25% identical aa to this region of the KpNtrA, AvNtrA, PpNtrA, RmNtrA, and RcNtrA proteins, respectively (when conservative substitutions are taken into account the % similarity increases to 92, 88, 85, 60 and 40%, respectively). Region 1 of the TfNtrA is also glutamine and leucine rich, and contains a proline residue at position 45.

Region 2 refers to a variable length of non-conserved aa sequence adjacent to Region 1. This region is present in the TfNtrA from Glu49 to Gln121. Although there is little sequence conservation between the six NtrA proteins in this region there is a stretch of 15 aa from Leu81 to Phe95 of the TfNtrA which shows similarity to the KpNtrA, AvNtrA, and PpNtrA; but not to the RmNtrA and RcNtrA. A glycine residue within this region, which was identified as essential in the RcNtrA (Gly108: Alias *et al.*, 1989), is conserved in all six NtrA protein sequences (position 155 in the TfNtrA; Fig. 3.4). Region 3 corresponds to the C-terminal two thirds of the NtrA proteins which show an average of 50% similarity with sub-regions of even greater similarity. Region 3 in the TfNtrA extends approximately from Ser122 to the C-terminal end. The putative DNA-binding helix-turn-helix (HTH) motif within Region 3 is highly conserved in the TfNtrA and extends from Leu366 to Asn385. Alignment of this potential DNA-binding site in NtrA proteins with HTH motifs of other DNA-binding proteins is shown in Fig. 3.5. The stretch of 9 aa from Ala453 to Arg461 near the C-terminus of the TfNtrA is 100 % conserved in all 6 NtrA proteins.

		1								10								20		
TfNtrA	(366)	L	R	H	I	A	D	A	V	E	M	H	E	S	T	V	S	R	V	T	N
		*			*	*			*		*	*		*	*	*	*	*			
KpNtrA	(367)	L	A	D	I	A	Q	A	V	E	M	H	E	S	T	I	S	R	V	T	T
		*			*	*			*		*	*		*	*	*	*	*			
PpNtrA	(387)	L	H	D	I	A	E	A	V	G	M	H	E	S	T	I	S	R	V	T	T
		*			*	*			*		*	*		*	*	*	*	*			
AvNtrA	(392)	L	H	D	I	A	E	A	V	G	M	H	E	S	T	I	S	R	V	T	T
		*			*	*			*		*	*		*	*	*	*	*			
RmNtrA	(394)	L	T	I	V	A	D	A	I	K	M	H	E	S	T	V	S	R	V	T	S
		*			*	*			*		*	*		*	*	*	*	*			
RcNtrA	(313)	L	E	D	V	A	S	E	L	G	L	H	A	S	T	I	S	R	A	V	S
EcLacI	(6)	L	Y	D	V	A	E	Y	A	G	V	S	Y	Q	T	V	S	R	V	V	N
EcGalR	(4)	I	K	D	V	A	R	L	A	G	V	S	V	A	T	V	S	R	V	I	N
EcCytR	(12)	M	K	D	V	A	L	K	A	K	V	S	T	A	T	V	S	R	A	L	M

Fig. 3.5. Alignment of the helix-turn-helix (HTH) motif in NtrA proteins with HTH motifs in known DNA-binding proteins. Amino acids (aa) which are conserved between NtrA proteins and the helix-turn-helix (HTH) motifs at the N-terminus of the *E. coli lac* repressor (EcLacI), *gal* repressor (EcGalR), and *cytR* protein (EcCytR) are blocked, and those aa conserved between NtrA proteins are indicated by an asterisk. Numerals refer to the positions within the HTH motif. Figures in parenthesis refer to the position in the protein of the first aa in each sequence.

Table 3.2. Amino acid composition (mole%) of TfNtrA compared with KpNtrA.

Amino acids	TfNtrA	KpNtrA	Average protein[a]
Acidic (D + E)	15,5	16,7	11,5
Acid + acid amide (D + E + N + Q)	25,3	27,6	19,8
Basic (K + R + H)	12,4	11,2	13,5
Hydrophobic (L + V + I + M)	25,1	26,7	20,2
Aromatic (F + Y + W)	4,6	5,4	8,3
Charged (D + E + K + R + H)	27,9	27,9	25,1
Aliphatic (A + G)	16,6	10,9	16,9
Hydroxyl (S + T)	10,1	13,2	13,1

[a] after Dayhoff *et al.* (1978).

3.3.3 Identification of ORF1 upstream of the *T. ferrooxidans ntrA* gene

Upstream of the *T. ferrooxidans ntrA* gene a second ORF (ORF1: nts 300 - 1025, Fig. 3.1) was identified. The predicted aa sequence of the *T. ferrooxidans* ORF1 product showed 57% aa identity to the predicted product of the *R. meliloti* ORF1 (Fig. 3.6), which is located upstream of the *R. meliloti ntrA* gene (Albright *et al.*, 1989b). This was considerably greater than the number of identical residues shared between the *ntrA* gene products of the two bacteria (40%) (Table 3.1). On the basis of the aa sequence similarity to the *R. meliloti* ORF1 product and the occurrence of a strong ribosome binding sequence (GGGAGG) at position 292, the ATG codon at position 300 is the most likely start of the *T. ferrooxidans* ORF1 (Fig. 3.2). The predicted product of the *T. ferrooxidans* ORF1 was 29 aa shorter at its N-terminal end than the *R. meliloti* ORF1 product. The translation products of all three reading frames upstream of nt 300 (Fig. 3.2) shared no homology with the N-terminus of the *R. meliloti* ORF1. The *T. ferrooxidans* ORF1 terminated 12 bp from the proposed start codon of the *T. ferrooxidans ntrA* gene.

A sequence with potential to form a stem and loop RNA secondary structure was present within the coding region of the *T. ferrooxidans* ORF1 (position 491, Fig. 3.2). This stem-loop structure (ΔG= -18,6 kcal/mol; Salser, 1977) was preceded by seven U's at position 480 (Fig. 3.2). TESTCODE (Fig. 3.3) and CODONPREFERENCE (Fig. 3.7) analysis indicated that there may be a protein-coding ORF upstream of the *T. ferrooxidans* ORF1, however the predicted translation of this ORF showed no similarity to the product of the *R. meliloti* ORF0 observed upstream of the *R. meliloti* ORF1 (Albright *et al.*, 1989b). No significant similarities where found between this ORF and any sequence in the available databases from homology searches using the UWGCG TFASTA and FASTA programmes (data not shown).

```
Tf                                                MSELLQAQSLFKSYRRRVVVRDVSVQVATGE      31
                                                    * *  * **** * **  **| *  **
Rm      MQIPFLHKRKRGKKPSAAAAAAARAVDKARYDGTLIARGLTKSYRSRRVVNGVSLVVRRGE      60

Tf      VVGLLGPNGAGKTTTFYMMVGLVRPDRGHIFLQQRDITALPMHERARMGLGYLPQEPSVF      91
         ************** ***  ***  * * * ||  *|*  **  *** *|****** *|*
Rm      AVGLLGPNGAGKTTCFYMITGLVPVDEGSIEINGNDVTTMPMYRRARLGVGYLPQEASIF     120

Tf      RQMSAADNVLAVLETLPLSPVERQERQEQLLSELHLHALRDTKGHSLSGGERRRVEIARA     151
        *  |  **| ****        *  |   ** * *  ** | |  ********|*****
Rm      RGLTVEDNIRAVLEVHDENVDRRESKLNDLLGEFSITHLRKSPAIALSGGERRRLEIARA     180

Tf      LAMSPRFILLDEPFAGIDPISVLEIQRLIRDLRARGIGVLITDHNVRETLGICERAYILH     211
        **  * * ********|***** |** *|* *  ****************| |****|*
Rm      LATDPTFMLLDEPFAGVDPISVADIQALVRHLTSRGIGVLITDHNVRETLGLIDRAYIIH     240

Tf      DGKVLTAGSPQEIVDDPMVRQVYLGDQFQI                                   241
         * *** *  ||**  * **  |****|* |
Rm      AGEVLTHGRANDIVTNPDVRRLYLGDNFSL                                   270
```

Fig. 3.6. Amino acid alignment between the ORF1 proteins from *T. ferrooxidans* and *R. meliloti*. Alignment was made with the IBM XT Computer Genepro (Version 4.1) protein alignment subroutine. Identical aa are marked with asterisks and conservative substitutions are marked with vertical bars. These are based on the groups (I, L, V, M); (D, E); (K, R, H); (Q, N); (S, T); (G, A) and (F, Y). Regions with similarity to the nucleotide binding pocket are underlined. Tf, *T. ferrooxidans* ORF1; Rm, *R. meliloti* ORF1.

3.3.4 Identification of ORF3 downstream of the *T. ferrooxidans ntrA* gene

A third ORF, called ORF3, was located downstream of the *T. ferrooxidans ntrA* gene (Fig. 3.1). The proposed ATG start codon of ORF3 was at position 2569 (Fig. 3.2) and was preceded by a consensus ribosome binding site (AGGAG) at position 2558 (Fig. 3.2), and ORF3 extended beyond the *Bgl*II cloning site at position 2804 (Fig. 3.2). Alignment of the predicted 78 N-terminal aa with the partial sequences of ORF3 products from *A. vinelandii* (Merrick *et al.*, 1987), *P. putida* (Inouye *et al.*, 1989), *K. pneumoniae* (Merrick and Gibbons, 1985), and *R. meliloti* (Ronson *et al.*, 1987b) showed 38, 38, 29, and 20% aa identity, respectively.

3.3.5 Nucleotide composition of the *T. ferrooxidans ntrA* gene and associated ORF's in relation to other *T. ferrooxidans* genes

The nucleotide sequences of six *T. ferrooxidans* genes have been determined, namely those of *glnA* (Rawlings *et al.*, 1987), *nifHDK* (Pretorius *et al.*, 1987; Rawlings, 1988), *recA* (Ramesar and Rawlings, 1989), and *merA* (Inoue *et al.*, 1989). A codon usage table, based on the sum of five of the genes (denoted TAB*Tf*5 for convenience) was drawn up using this data. The *merA* codon usage was not included since this was reported to be different from previously sequenced *T. ferrooxidans* chromosomal genes but similar to the Tn501 and R100 *merA* genes (Inoue *et al.*, 1989). The codon usage for the *ntrA* and ORF1 genes are shown together with the data of TAB*Tf*5 in Table 3.3.

The codon usage of the *ntrA* gene and ORF1 is very similar to that in TAB*Tf*5, with bias towards codons with a G or C in the "wobble" position. Exceptions from the trend found in TAB*Tf*5 are: greater representation of the GCG (alanine), UUU (phenylalanine), and AGC (serine) codons in *ntrA* and ORF1; preference for the AAU (asparagine) codon in *ntrA*; equal representation of the AAA and AAG (lysine) codons in *ntrA*; preference for the GAU (aspartate) and CAU (histidine) codons in ORF1; and greater representation of the GGU (glycine) and AUU (isoleucine) codons in ORF1.

Table 3.3. Codon usage of *T. ferrooxidans* genes[a]

Codon		*ntrA*	ORF1	TAB*f*5	*glnA*	*recA*	*nifHDK*
Ala	GCA	5	2	22	4	5	13
	GCC	20	6	102	29	22	51
	GCG	22	7	49	9	8	32
	GCU	2	1	12	1	5	6
Arg	AGA	0	0	2	0	0	2
	AGG	0	0	4	2	0	2
	CGA	1	0	3	0	1	2
	CGC	17	11	49	10	10	29
	CGG	6	8	20	2	6	12
	CGU	10	6	19	4	3	12
Asn	AAC	5	3	47	10	6	31
	AAU	9	0	24	7	6	11
Asp	GAC	23	2	85	19	10	56
	GAU	17	11	42	6	10	26
Cys	UGC	2	1	19	4	1	14
	UGU	0	0	7	0	1	6
Gln	CAA	9	4	4	0	0	4
	CAG	22	11	56	8	14	34
Glu	GAA	19	9	80	18	16	46
	GAG	16	7	62	9	9	44
Gly	GGA	2	2	11	3	2	6
	GGC	14	6	126	16	18	92
	GGG	4	4	17	4	2	11
	GGU	8	7	41	11	14	16
His	CAC	6	1	39	7	5	27
	CAU	5	6	23	7	3	13
Ile	AUA	1	2	4	0	2	2
	AUC	14	6	105	16	19	70
	AUU	8	6	20	4	6	10
Leu	CUA	0	0	3	1	0	2
	CUC	11	5	34	8	8	18
	CUG	29	18	91	18	20	53
	CUU	3	3	11	4	3	4
	UUA	2	2	2	1	1	0
	UUG	10	3	15	2	1	12
Lys	AAA	8	2	35	4	8	23
	AAG	8	2	96	20	12	64
Met	AUG	16	8	68	16	8	44
Phe	UUC	4	3	67	14	3	50
	UUU	6	4	21	3	6	12
Pro	CCA	1	1	8	1	2	5
	CCC	10	6	48	13	4	31
	CCG	10	2	30	8	4	18
	CCU	1	3	10	4	2	4
Ser	AGC	13	5	26	5	5	16
	AGU	1	2	19	4	5	10
	UCA	0	0	6	1	1	4
	UCC	6	3	44	11	10	23
	UCG	4	3	25	5	4	16
	UCU	3	0	11	5	1	5
Thr	ACA	2	0	4	1	0	3
	ACC	10	5	64	10	11	43
	ACG	7	3	23	6	1	16
	ACU	2	2	7	1	1	5
Trp	UGG	4	0	22	5	1	16
Tyr	UAC	6	3	46	11	8	27
	UAU	4	2	31	6	3	22
Val	GUA	1	1	10	4	2	4
	GUC	5	5	45	10	10	25
	GUG	18	12	72	16	13	43
	GUU	3	4	9	1	1	7
End	TGA	1	1	2	1	1	0
	TAG	0	0	0	0	0	0
	TAA	0	0	3	0	0	3

[a] Numerals refer to the number of times a codon occurs in each gene dataset.

The UWGCG CODONPREFERENCE application is designed to use a codon usage data file to analyze the protein translations in all six reading frames of a DNA sequence to identify ORF's that have a similar codon usage to that provided in a datafile. Fig. 3.7 shows analysis of the *T. ferrooxidans* DNA insert in pT30 using the UWGCG CODONPREFERENCE application with TAB*Tf*5 as the codon usage datafile. The *ntrA* gene and ORF1 in frame 3 and ORF3 in frame 1 show a codon preference and an absence of rare codons which correlates to TAB*Tf*5. The GC composition of the *T. ferrooxidans ntrA* gene and ORF1 were 59,5 and 58,8%, respectively, which matches the GC composition calculated for *T. ferrooxidans* chromosomal DNA (60%) (Harrison, 1984).

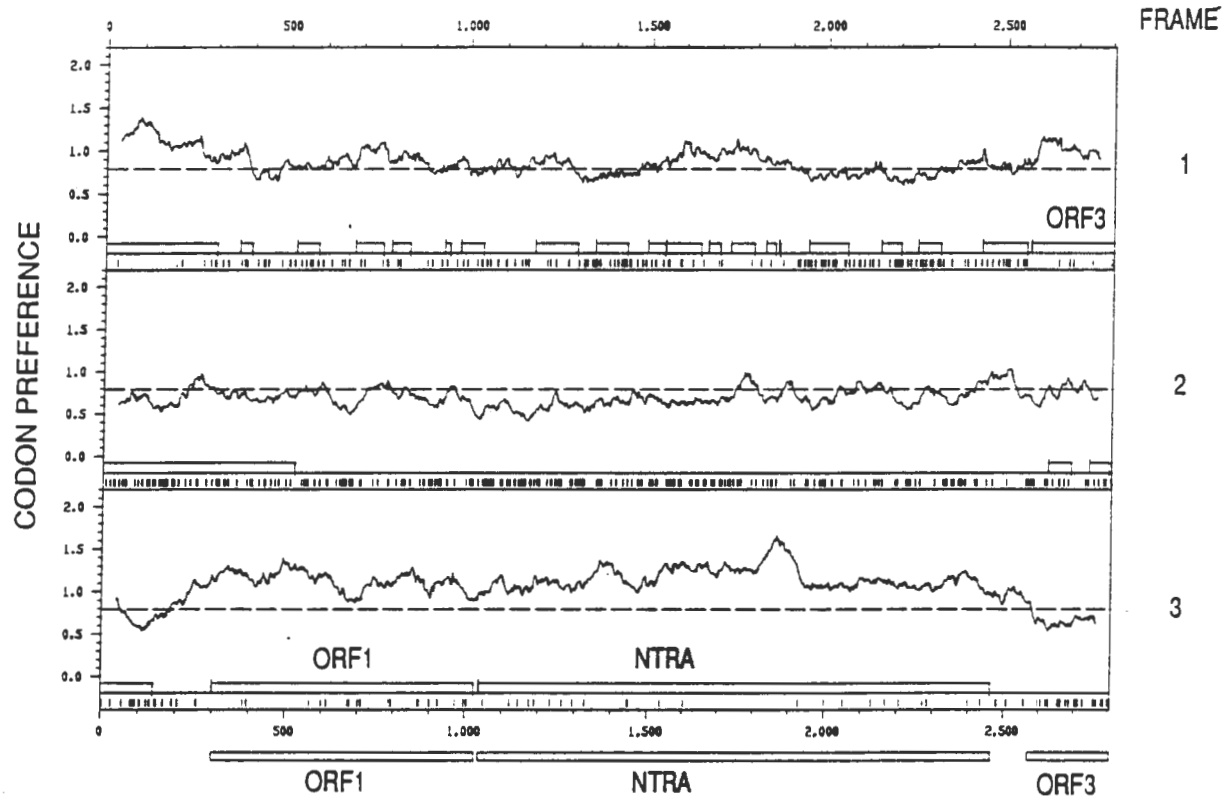


Fig. 3.7. CODONPREFERENCE analysis of the 2,8-kbp DNA insert in pT30. The horizontal axis represents the 2,8-kbp *T. ferrooxidans* DNA insert in pT30, with the predicted protein-coding ORF's (ORF1, NTRA, ORF3) shown as open boxes below. The vertical axis shows the CODONPREFERENCE plots for each of the 3 forward frames, with ORF's and the positions of rare codons (defined as those codons which appear in the TAB*Tf*5 codon usage datafile at a frequency of less than 10%) drawn below as open boxes and vertical bars, respectively. Any plot above the dotted line in any of the 3 reading frames identifies a protein-coding sequence with a similar codon usage as the TAB*Tf*5 datafile.

3.4 Discussion

The predicted protein product of the *T. ferrooxidans ntrA* gene was an acidic polypeptide of 475 aa with a calculated M_r of 52927. The term σ54 reflects the M_r of the product of the *ntrA* gene of enteric bacteria, however the M_r of *ntrA* gene products vary considerably. From previously published nucleotide sequences the M_r of other *ntrA* gene products have been calculated as 53926, 56215, 56916, 57814, and 46328 for *K. pneumoniae* (Merrick and Gibbons, 1985), *P. putida* (Inouye *et al.*, 1989), *A. vinelandii* (Merrick *et al.*, 1987), *R. meliloti* (Ronson *et al.*, 1987b), and *R. capsulatus* (Jones and Haselkorn, 1989), respectively.

The % identity between NtrA proteins can be seen in Table 3.1, however a more illustrative representation of the possible evolutionary relationships between NtrA proteins is shown in the dendrogram produced using the UWGCG program DISTANCES and the subroutine KITSCH of the program PHYLIP (J. Felsenstein, University of Washington, 1988) (Fig. 3.8). DISTANCES was used to write a matrix of the pairwise genetic distances within the group of aligned NtrA aa sequences. This matrix of distances was analyzed using KITSCH, which employs the Fitch-Margoliash and Least Squares method (Fitch and Margoliash, 1967), with the assumption that there is an evolutionary molecular clock, to produce the dendrogram shown in Fig. 3.8.

This shows that on the basis of aa sequence conservation it is possible to divide these NtrA proteins into three groups. The TfNtrA is most similar to the KpNtrA which forms a group together with the AvNtrA and PpNtrA which are very closely related. The similarities at the level of aa sequence between the *T. ferrooxidans* NtrA and this group of NtrA proteins is interesting as the autotrophic *T. ferrooxidans* inhabits a very different environment from these heterotrophic bacterial species. The RmNtrA and RcNtrA share relatively little sequence similarity with each other or the former group of NtrA proteins and appear to belong to two separate groups.

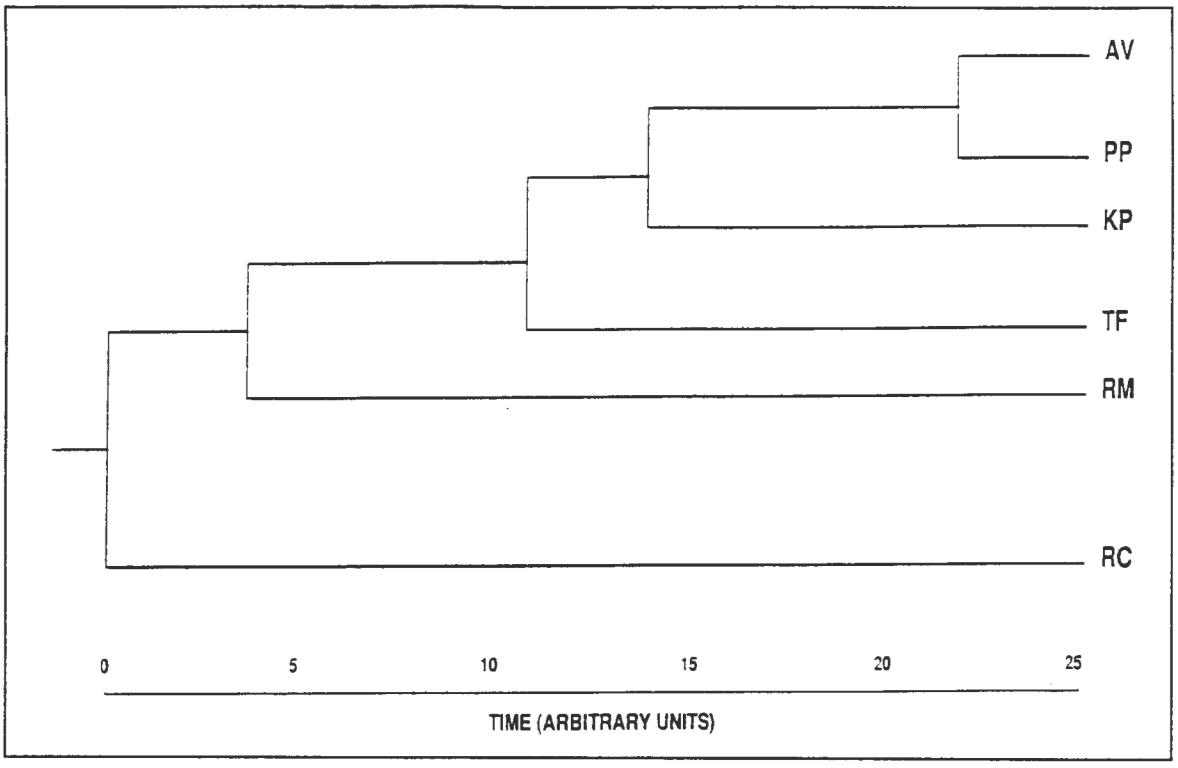


Fig. 3.8. Dendrogram of NtrA proteins based on amino acid sequence conservation. TF, *T. ferrooxidans* NtrA; KP, *K. pneumoniae* NtrA; PP, *P. putida* NtrA; AV, *A. vinelandii* NtrA; RM, *R. meliloti* NtrA; RC, *R. capsulatus* NtrA.

The codon usage of the *T. ferrooxidans ntrA* gene and ORF1 appear to be similar to that of the five *T. ferrooxidans* genes sequenced to date (TAB*Tf*5) as shown in Table 3.3, however analysis of the codon usage data using the UWGCG CORRESPOND program is interesting. The program CORRESPOND is designed to identify similar patterns of codon usage by comparing codon frequency tables. The frequency which is compared is calculated by dividing the number of incidents of the codon in question by the total number of codons specifying that aa or terminator in each table. The lower the statistic (D), the more similar the patterns of codon usage. The codon frequency tables for the *T. ferrooxidans ntrA* and ORF1 were compared to the composite (TAB*Tf*5) and individual codon frequency tables of the *T. ferrooxidans glnA*, *recA*, and *nifHDK* genes, as well as the codon frequency tables for highly expressed (UWGCG datafile "ecohigh.cod") and weakly expressed (UWGCG datafile "ecolow.cod" - Grantham *et al.*, 1981) *E. coli* genes.

The results calculated using the CORRESPOND program (Table 3.4) show that the codon usage of the *T. ferrooxidans ntrA* and ORF1 genes is more similar to the codon usage of the *T. ferrooxidans glnA* and *recA* genes and that of weakly expressed *E. coli* genes, than the codon usage of the *T. ferrooxidans nifHDK* and highly expressed *E. coli* genes. This is consistent with the hypothesis that, due to the regulatory nature of the product, *ntrA* genes are weakly expressed. In support of this view is the observation by Merrick *et al.* (1987) that the *A. vinelandii ntrA* gene utilises a number of modulator codons (Grosjean and Fiers, 1982) in contrast to the highly expressed *A. vinelandii nifHDK* genes (Brigle *et al.*, 1985).

Table 3.4. Comparison of the codon usage of the *T. ferrooxidans ntrA* and ORF1 with that of *T. ferrooxidans* and *E. coli* genes using the UWGCG program CORRESPOND[a]. Numbers represent the statistic D^2 as described in the text.

	TfrecA	*TfglnA*	*TfnifHDK*	TAB*Tf*5	*Ec*low	*Ec*high
TfntrA	1,55	1,44	3,55	2,01	1,67	4,03
TfORF1	2,61	2,70	4,66	3,16	2,07	4,78

[a] *TfntrA*, *TfORF1*, *TfrecA*, *TfglnA*, *TfnifHDK*, *Ec*low, and *Ec*high represent codon usage frequency tables of the *T. ferrooxidans ntrA*, ORF1, *recA*, *glnA*, *nifHDK*, *E. coli* weakly expressed, and *E. coli* highly expressed genes, respectively. TAB*Tf*5 is a composite codon usage frequency table of the *T. ferrooxidans recA*, *glnA*, and *nifHDK* genes.

Sigma factors carry out three major biochemical activities which include (i) binding to core RNA polymerase, (ii) promoter recognition, and (iii) facilitating the catalysis of closed to open promoter complexes which involves DNA melting (Helmann and Chamberlin, 1988). Purified NtrA protein from *E. coli*, *S. typhimurium*, and *K. pneumoniae* has been shown *in vitro* to direct transcription from NtrA-dependent promoters, thus confirming the classification of the NtrA protein as a sigma factor (Hunt and Magasanik, 1985; Hirschman *et al.*, 1985, Wong *et al.*, 1987).

Analysis of aa sequences of bacterial sigma factors has shown that many of these belong to a homologous protein superfamily (Stragier *et al.*, 1985; Gribskov and Burgess, 1986). This includes the primary σ factors in *E. coli* (Gram-negative) and *B. subtilis* (Gram-positive) which are denoted σ70 and σ43, respectively; as well as

several of the alternative σ factors in both species, such as the *E. coli* σ32 (HtpR), which is involved in the heat-shock response and many of the *B. subtilis* sigma factors involved in sporulation, such as σ30 (sigH), σ29 (sigE), and σSPOIIAC (which has not yet been shown to function as a sigma factor but shows suggestive similarities to other sequenced sigma factors) and the *B. subtilis* σ37 (SigB), the function of which is unknown (reviewed in Helmann and Chamberlin, 1988).

Merrick and Gibbons (1985) and Merrick *et al.* (1987) reported that there was no statistically significant similarity in pairwise alignments between the KpNtrA and other bacterial sigma factors, and that observation of the aa sequence of the AvNtrA confirmed their idea that NtrA sigma factors did not warrant classification in this family of bacterial sigma factors.

Helmann and Chamberlin (1988) have pointed out that although many of the phage sigma factors do not show statistically significant sequence similarity in pairwise alignments with members of the family of bacterial sigma factors they do show some similarity when viewed in multiple sequence alignments. There are several drawbacks with a computer-based statistical pairwise alignment, namely:

i) The definition of conservative substitutions may result in a biased weighting scheme. Groupings are aimed to reflect either common functional moieties, hydrophobicity, or size. The scheme adopted in Appendix E is based mainly upon common functional moieties, however in another scheme one could group Ser, Ala, and Cys together on the basis of size.
ii) Similarities between proteins may reflect repeat units common to protein secondary structural elements, such as the 3,6 aa repeat of the alpha helix. Residues conserved along one face of the helix will occur every third or fourth residue in the primary sequence.
iii) Distantly related proteins which contain short regions of similarity may be overlooked by a dilution effect of the larger regions of dissimilar sequences.
iv) Only two proteins may be compared at one time and therefore similarities amongst a family of related proteins may not be immediately evident.

The use of multiple sequence alignments rather than pairwise alignments enabled Helmann and Chamberlin (1988) to define regions that were similar amongst distantly related bacterial and phage sigma factors and thereby propose those conserved regions that may represent functional domains. As NtrA sigma factors showed little sequence similarity to the other bacterial sigma factors, Helmann and Chamberlin (1988) did not include NtrA sigma factors in their study. Nevertheless, it was thought in this study that it may be possible to identify conserved regions with functional significance by analysis of multiple alignments between the six NtrA sigma factors and the conserved regions of bacterial sigma factors.

The results of this approach are presented below. It was found that there were conserved regions within the NtrA sigma factors which showed similarity to functional domains within the bacterial sigma factors and therefore the NtrA family may represent an evolutionary related, though greatly diverged group of proteins from the other bacterial sigma factors. These and other conserved regions will be discussed within the context of identifying possible functional domains within NtrA sigma factors.

Helmann and Chamberlin (1988) observed four conserved regions amongst bacterial sigma factors and the functional attribute of each region was proposed based on biochemical and genetic evidence. This is summarized in the schematic diagram of the *E. coli* σ70 factor, a representative of this superfamily (Fig. 3.9a).

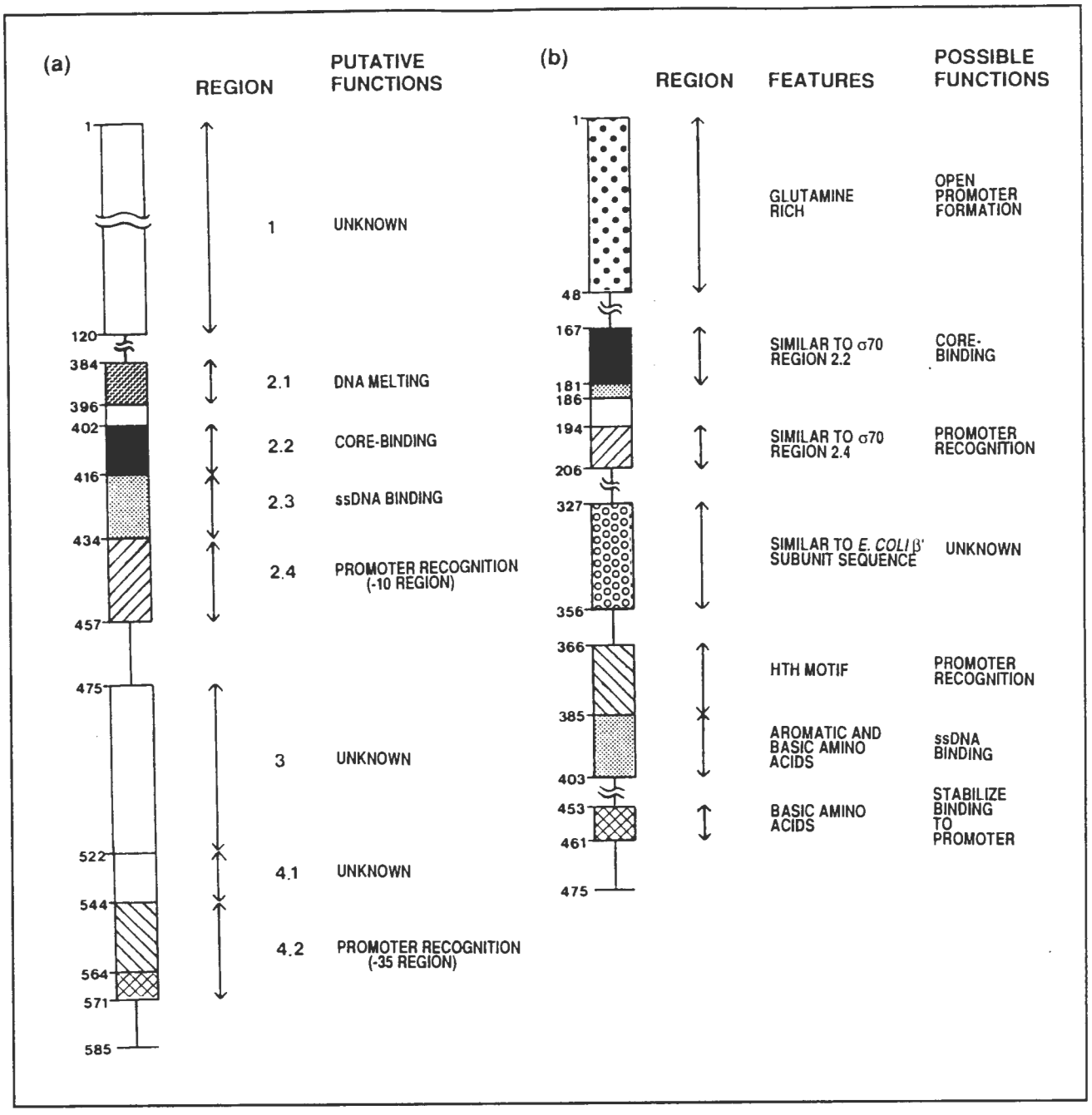


Fig. 3.9a-b. Possible functional regions within sigma factors. a) Schematic diagram of the *E. coli* σ70 factor showing regions conserved amongst members of the superfamily of bacterial sigma factors (Helmann and Chamberlin, 1988). The extent and putative functions of these regions are indicated by arrowed lines and notes on the right hand side, respectively. Boxes are shaded differently to represent different putative functional regions. Numbers on the left hand side refer to the aa positions in the primary sequence of the *E. coli* σ70 factor, starting with 1 at the N-terminus. b) Schematic díagram of the *T. ferrooxidans* NtrA showing regions conserved amongst NtrA sigma factors. Conserved regions are represented as boxes with their extent, and features and possible functions indicated by arrowed lines, and notes on the right hand side, respectively. Boxes are shaded to show regions which displayed similarity with conserved regions in the superfamily of bacterial sigma factors as drawn in Fig. 3.9a. Numerals on the left hand side refer to aa positions in the primary sequence of the *T. ferrooxidans* NtrA, starting with 1 at the N-terminus.

No similarity was found between the NtrA sigma factors and sequences corresponding to Regions 1, 2.1, 2.3, or 3 which are conserved within the superfamily of bacterial sigma factors (Fig. 3.9a), however some similarity was found between the NtrA sigma factors and sequences in Regions 2.2, 2.4, and 4.

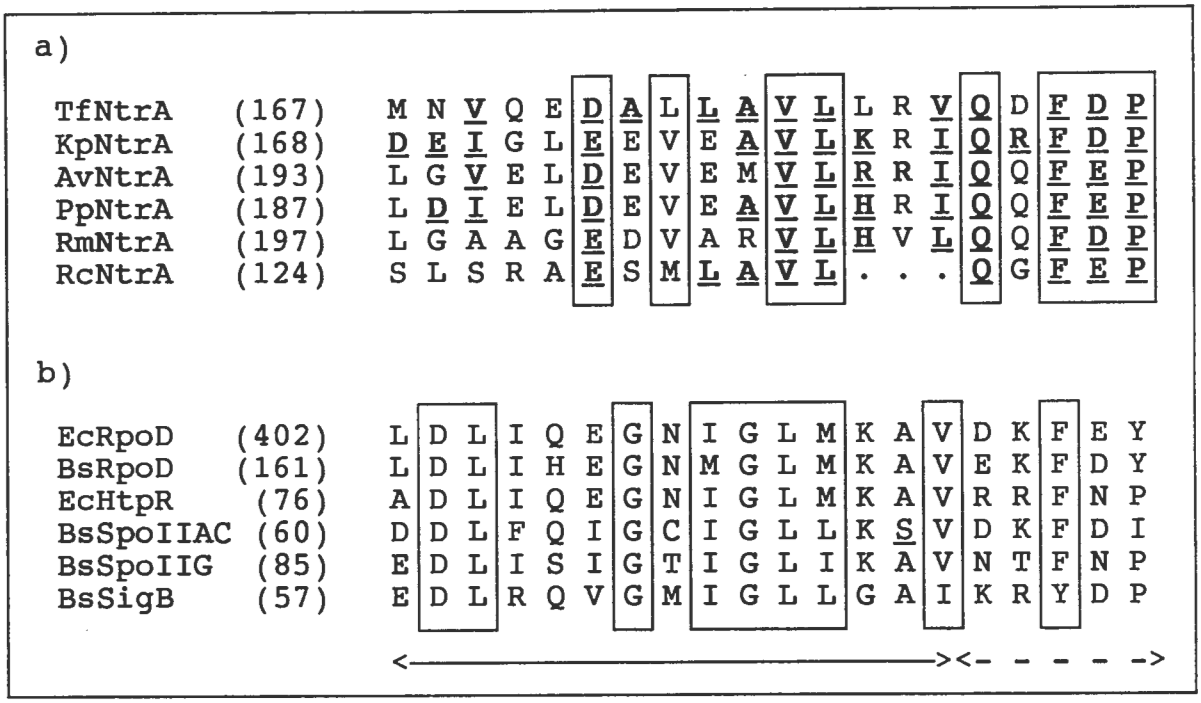


Fig. 3.10a-b. Alignment of a 20 amino acid sequence from NtrA sigma factors with conserved regions 2.2 and 2.3 identified in the superfamily of bacterial sigma factors (Gribskov and Burgess, 1986; Helmann and Chamberlin, 1988). a) Alignment of this region in NtrA sigma factors. Underlined and bold face letters identify those amino acids (aa) or conservative substitutions which are found at the corresponding position in at least three out of the six bacterial sigma factors in Fig. 3.10b. Residues conserved in all six NtrA sigma factors are blocked. b) Alignment of region 2.2 (filled arrow) and part of region 2.3 (dashed arrow) of six bacterial sigma factors according to Helmann and Chamberlin (1988). Residues that are conserved in all six sigma factors are blocked. The underlined residue is the site of a BsSpoIIAC mutation as discussed in section 3.4. Figures in parenthesis refer to the position in the protein of the first aa in each sequence.

Region 2.2 (Fig. 3.9a) contains a highly conserved group of aa amongst bacterial sigma factors and it is therefore thought that this forms part of the domain which binds the RNA polymerase (Helmann and Chamberlin, 1988). Evidence for the role of this region in core-binding is the mapping of a mutation within this region of the σSPOIIAC factor (Yudkin, 1987a). Overexpression of σSPOIIAC in *E. coli* is toxic which is attributed to the binding of σSPOIIAC to *E. coli* RNA polymerase. Mutation of Ser73 within σSPOIIAC (see Fig. 3.10b) to Phe is thought to result in a mutant

protein which is unable to bind to core polymerase therefore abolishing its toxicity when overexpressed in *E. coli*. Merrick *et al.* (1987) identified similarity between region 2.2 and the KpNtrA, AvNtrA, and RmNtrA if conservative substitutions are taken into account, and as can be seen in Fig. 3.10a,b this similarity includes the other NtrA proteins.

Region 2.3 (Fig. 3.9a) in the superfamily of bacterial sigma factors is adjacent to region 2.2, and as can be seen in Fig. 3.10a,b the similarity between the NtrA sigma factors and the bacterial sigma factors extends into part of region 2.3. There are two conserved Phe residues in region 2.3. The first of these is shown in Fig. 3.10b and this is also conserved in all 6 NtrA sigma factors (Fig. 3.10a). Region 2.3 is proposed to be involved in binding of the non-transcribed single stranded DNA (ssDNA) during open complex formation on the basis of sequence homologies with the aromatic rich domains of ssDNA binding proteins (Helmann and Chamberlin, 1988). The absence of a second conserved aromatic residue spaced 11 aa from the first in the NtrA sigma factors indicates that this region may not be responsible for ssDNA binding. A conserved region which may fulfil the criteria for a ssDNA binding domain in NtrA sigma factors is shown in Fig. 3.11. Protein structures that appear to be involved in interactions with ssDNA (Chase and Williams, 1986, McPherson *et al.*, 1979) are characterized by aromatic residues which are spaced by 5-13 aa and are flanked by basic residues. Stacking interactions between the nucleotide bases and the aromatic aa side chains appear to be important in the binding interaction, while the basic aa are thought to be involved in charge neutralization (Khamis *et al.*, 1987).

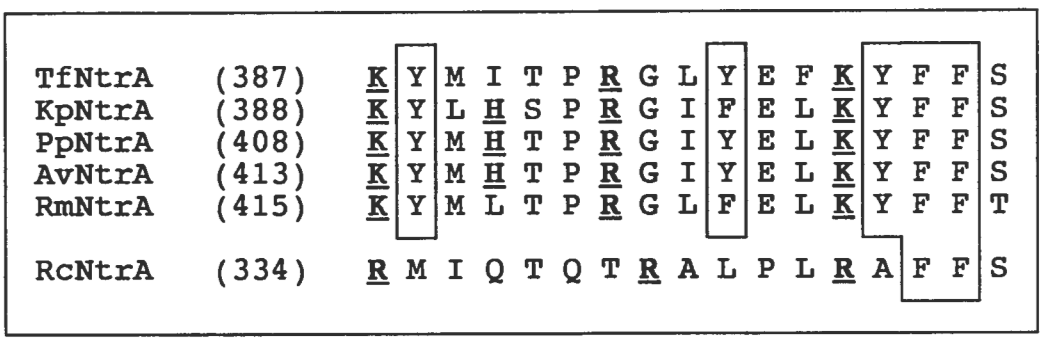


Fig. 3.11. Alignment of a 17 amino acid region conserved within NtrA proteins which may be involved in binding to single stranded DNA. Conserved aromatic residues are blocked. Underlined and bold face letters identify basic residues. Figures in parenthesis refer to the position in the protein of the first aa in each sequence.

Region 2.4 (Fig. 3.9a) contains three conserved hydrophobic residues located at every fourth aa followed by basic residues and is predicted to be alpha-helical (Chou and Fasman, 1978). The hydrophobic residues may be involved in the packing of the alpha-helix against the remainder of the protein, while mutational evidence suggests that the exposed face of the putative alpha-helix may be involved in recognition of the -10 region of promoters. Precedents for alpha helical structures which are not canonical helix-turn-helix motifs but are able to bind in a sequence specific manner to DNA may be found in the *Eco*RI endonuclease (McClarin *et al.*, 1986) and the yeast GCN4 protein (Hope and Struhl, 1986). Siegele *et al.* (1989) have shown that a mutation in this region of the *E. coli* σ70(rpoD) (Thr to Ile change at position 440 - see Fig. 3.12) stimulates transcription from P22 *ant* and *E. coli lac* promoters with mutations in the first T of the -10 hexamer (TATAAT). Similarly, a mutation in this region of the *B. subtilis* σ30(sigH)(Thr to Ile change at position 100 - see Fig. 3.12) suppresses a -13 promoter mutation in the *spo*VG gene (Zuber *et al.*, 1989). A 13 aa sequence from NtrA sigma factors (Fig. 3.12) contains three conserved hydrophobic residues spaced every fourth aa as found in region 2.4 (Fig. 3.9a), which may indicate that this conserved region may form a similar alpha-helical secondary structure.

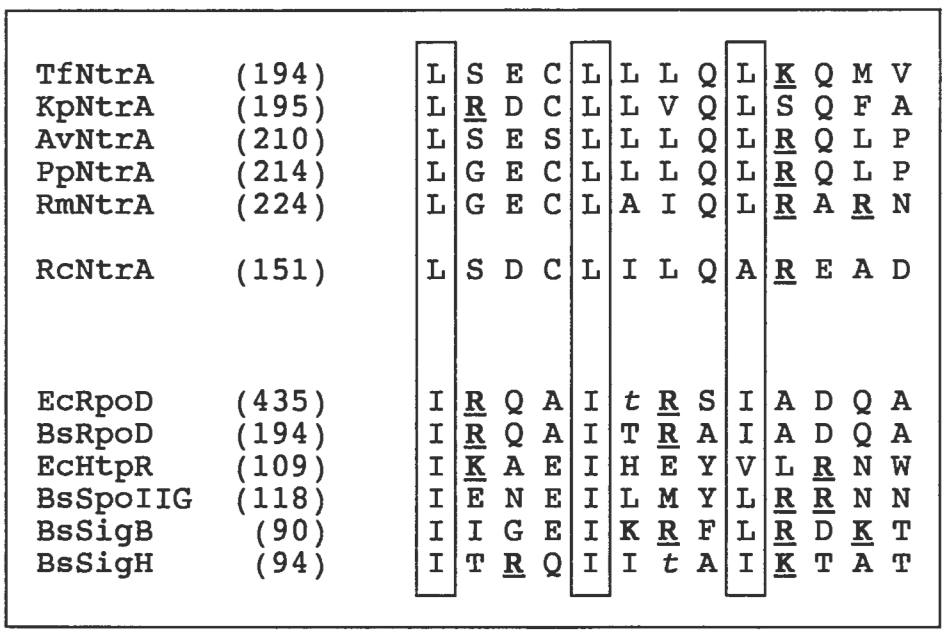


Fig. 3.12. Alignment of a 13 amino acid sequence from NtrA sigma factors with a region thought to be involved in promoter recognition in bacterial sigma factors (Region 2.4 - Helmann and Chamberlin, 1988). Conserved hydrophobic residues are boxed (Ala159 in the RcNtrA is grouped as a hydrophobic residue). Underlined and bold face letters identify basic residues. Lower case italicised letters identify sites of bacterial sigma factor mutations as discussed in the text. Figures in parenthesis refer to the position in the protein of the first aa in each sequence.

Region 4 (Fig. 3.9a) situated towards the C-terminal end of prokaryotic sigma factors contains a canonical helix-turn-helix (HTH) DNA binding motif. The strongest evidence implicating this region in promoter recognition comes from mutational studies. Two mutations in this region of the *E. coli* σ70 have been reported that specifically suppress -35 region mutations in the P22 *ant* and *E. coli lac* promoters (Siegele *et al.*, 1989; Gardella *et al.*, 1989). There is a C-terminal HTH motif which is highly conserved in all NtrA sigma factors (Fig. 3.5).

Dodd and Egan (1987) developed a quantitative method for assessing whether a given primary aa sequence contained a canonical HTH unit. The HTH motifs observed by crystallography of the prokaryotic repressor proteins formed the basis for a dataset which was used to generate an aa probability matrix. This matrix was used to calculate a table of "points" for every possible aa at each position in the HTH unit depending upon the occurrence of each aa at each position in the HTH units of the dataset proteins. A score can therefore be calculated for any test sequence by taking the sum of the "points" obtained for each of the 20 aa within the putative HTH motif. A score of >1400 is adjudged to indicate a classical HTH motif, although obviously this method is biased towards HTH motifs similar to those found within the dataset of Dodd and Egan.

Computation of the scores for the putative HTH motifs in NtrA sigma factors by this method gave scores of 1444, 1352, 1635, 1635, 1157, and 1507 for the putative HTH units of the TfNtrA, KpNtrA, AvNtrA, PpNtrA, RmNtrA, and RcNtrA, respectively. Adjustment of these scores by the method of Yudkin (1987b) - which omits the highly variable residue 7, and residues 12 and 13, which are thought to contact the DNA and therefore vary depending on the recognition sequence - resulted in further improvement of the scores (relative to the new dataset eliminated for residues 7, 12, 13) which indicates that these sequences in NtrA sigma factors are very likely to form classical HTH motifs. It is interesting to note that the second helix of a HTH motif is considered to be the primary determinant of sequence specificity (Pabo and Sauer, 1984) and this part of the proposed HTH unit of NtrA proteins (positions 10 - 20; Fig. 3.5) is entirely conserved except for three residues in the RcNtrA and position 20 of the TfNtrA.

A group of conserved basic residues near the C-terminus in subregion 4.2 (Fig. 3.9a) of members of the bacterial sigma factor superfamily is thought to be appropriately positioned to form ionic interactions with backbone phosphate residues, especially if the preceding HTH region contacts the -35 region. The basic residue rich sequence -ARRTVAKYR- is conserved in all six NtrA sigma factors close to the C-terminus and this may play a similar role in stabilization of the HTH-DNA interaction in NtrA sigma factors.

NtrA-dependent promoters are characterized by conserved GG and GC doublets at positions -24 and -12 to the transcription start site, respectively (reviewed in Kustu *et al.*, 1989). By analogy with the superfamily of bacterial sigma factors the NtrA HTH motif could interact with the -24 promoter region while another domain could interact with the -12 promoter region, however it is premature to speculate on this further.

The first 48 aa of the NtrA sigma factors are highly conserved (Fig. 3.4) and contain a high proportion of glutamine (15 - 25%) and leucine (17 - 29%) residues. This has potential to represent an N-terminal domain which terminates in a conserved proline turn motif. There is some resemblance to the glutamine-rich region required for activation of transcription by mammalian transcription factor Sp1 (Courey and Tijan, 1988). This region of the *S. typhimurium* NtrA has been shown by deletion analysis to be specifically required for NR_I-dependent isomerization of closed to open complexes at the *glnA* promoter (Kustu *et al.*, 1989).

Merrick *et al.* (1987) recognized a region conserved amongst NtrA sigma factors which showed similarity to a sequence near the N-terminus of the β' subunit of *E. coli* RNA polymerase (RpoC). Alignment of this region is showed in Fig. 3.13. The reason for this duplication of structure is not obvious, but Merrick *et al.* (1987) speculated that the binding of NtrA to core RNA polymerase alters the structure of the enzyme so that this region of NtrA assumes a role otherwise played by the homologous region of RpoC.

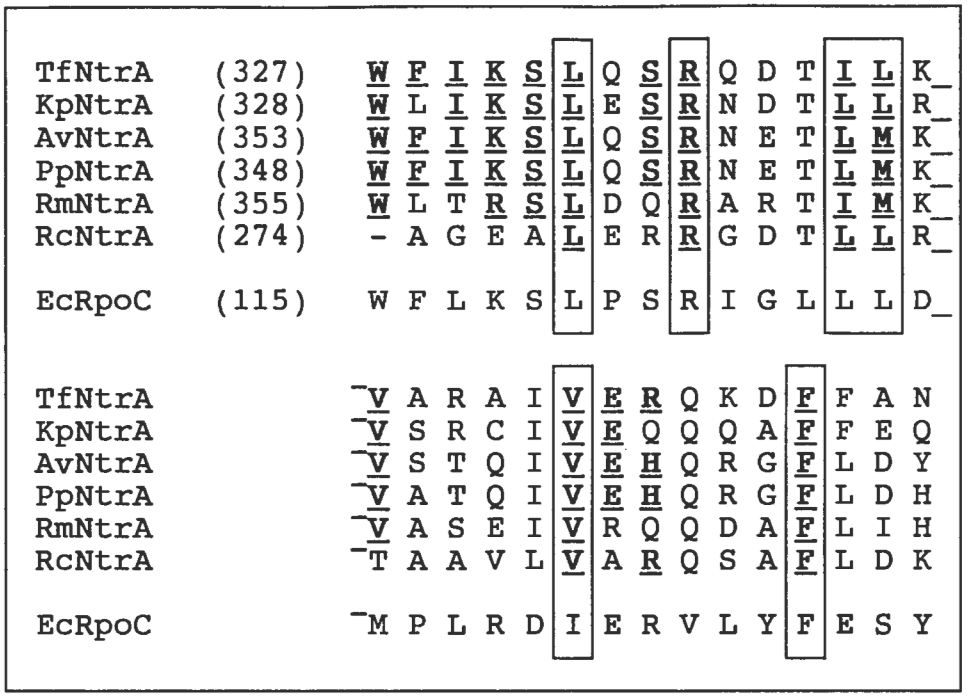


Fig. 3.13. Alignment of a 30 amino acid sequence from NtrA sigma factors with a sequence near the N-terminus of the β' subunit of *E. coli* RNA polymerase (RpoC). Amino acids conserved between all six NtrA sigma factors and RpoC are blocked. Underlined and **bold** face letters identify aa conserved in at least three NtrA sigma factors and RpoC. Figures in parenthesis refer to the position in the protein of the first aa in each sequence.

Fig. 3.9b summarizes the available data on the possible functional regions within NtrA sigma factors. These regions are shown diagrammatically using the TfNtrA as a representative of the NtrA sigma factors. The regions are drawn as shaded blocks in Fig. 3.9b to enable correlation with those regions conserved in the super family of bacterial sigma factors shown in Fig. 3.9a. The data in Fig. 3.9b is speculative and is based on primary sequence scanning with conserved regions of the six NtrA sigma factors, and therefore should be regarded as a basic guideline for the design of mutational and biochemical studies to elucidate structural/functional relationships within NtrA sigma factors. A unique functional feature of NtrA sigma factors, which is not found in other bacterial sigma factors, is the absolute requirement for an activator protein to facilitate the isomerization of closed to open promoter complexes.

Immediately upstream of the *T. ferrooxidans ntrA* gene an ORF equivalent to the ORF1 located upstream of the *R. meliloti ntrA* gene was detected. The linkage of ORF1 to *ntrA* has been reported also to occur in *K. pneumoniae* (Merrick *et al.*, 1987; Albright *et al.*, 1989b), *S. typhimurium* (Albright *et al.*, 1989b), and *P. putida* (Inouye *et al.*, 1989) and therefore appears to be a feature of bacteria of very different

physiological types. The biological function of ORF1 and the reason for its linkage to the *ntrA* gene is uncertain. Transcription of ORF1 and *ntrA* were reported to be uncoupled in *R. meliloti* (Albright *et al.*, 1989b). The observation that subclones of the *T. ferrooxidans ntrA* gene from which most of ORF1 had been deleted (pT40 and pT41 - see Ch. 4, Table 4.1) were able to complement the *E. coli ntrA* mutant is evidence that transcription of the *T. ferrooxidans ntrA* gene in *E. coli* is independent of transcription through ORF1. *In vitro* studies on NtrA-dependent promoters have indicated that NR_I-activated transcription from the *S. typhimurium glnA* (Hirschman *et al.*, 1985), *E. coli glnA* (Hunt and Magasanik, 1985), *K. pneumoniae nifLA* , or the *K. pneumoniae nifHDK* (Wong *et al.*, 1987) promoters requires only the NtrA-RNA polymerase holoenzyme. Involvement of the ORF1 product in the function of NtrA is therefore unlikely. Albright *et al.* (1989b) tried unsuccessfully to insertionally inactivate ORF1 using transposon mutagenesis, and suggested that it may code for an essential housekeeping protein.

The predicted aa sequence of the *T. ferrooxidans* ORF1 was aligned with the *R. meliloti* protein, the only other ORF1 for which a complete sequence has been published (Fig. 3.6). The two sequences have a high level of similarity over their entire length although the N-terminal end of the predicted *T. ferrooxidans* protein was 29 aa shorter. The aa sequence of the *R. meliloti* ORF1 has been compared with a family of ATP-binding proteins (Higgins *et al.*, 1986) and two regions that had homology to an ATP-binding pocket were identified (Albright *et al.*, 1989b). These are also conserved in the predicted *T. ferrooxidans* ORF1 protein (Fig. 3.6). The spacing between the *T. ferrooxidans* ORF1 and the *ntrA* gene was different from that of *R. meliloti* and *P. putida*. The *R. meliloti* and *P. putida* ORF1's terminated approximately 176 bp upstream of the respective *ntrA* genes whereas the *T. ferrooxidans* ORF1 terminated only 12 bp from the start codon of the *ntrA* gene. Since expression of the *T. ferrooxidans ntrA* gene was independent of orientation and most of ORF1 (see section 4.3.2; Ch. 4), the *ntrA* promoter recognized in *E. coli* must be located within the carboxy-terminal coding region of ORF1. The existence of a putative region of RNA stem and loop secondary structure (ΔG = -18,6 Kcal/mol; Salser, 1977) preceded by seven U's within the coding region of the *T. ferrooxidans* ORF1 (Fig. 3.2) is interesting as a similar region of potential secondary structure

(ΔG = -17.1 Kcal/mol; Salser, 1977) preceded by an A/U rich region is also present in the *R. meliloti* ORF1 nucleotide sequence (position 635 - Fig. 1 in Albright *et al.*, 1989b). These regions of potential secondary structure are situated within an area of low aa conservation on the C-terminal side of the first component of the ATP-binding pocket of the two predicted ORF1 products. A similar region of secondary structure has been shown to reduce translation efficiency in another bacterial system (Kubo and Imanaka, 1989).

The N-terminal 78 aa encoded by ORF3 situated downstream of the *T. ferrooxidans ntrA* gene showed aa similarity to equivalent ORF's downstream of the *ntrA* genes in *K. pneumoniae* (Merrick *et al.*, 1987), *A. vinelandii* (Merrick *et al.*, 1987), *R. meliloti* (Ronson *et al.*, 1987b), and *P. putida* (Inouye *et al.*, 1989). The extent of the *T. ferrooxidans* DNA insert in pT30 prevented determination of the complete sequence of ORF3. The complete sequences of this ORF from *K. pneumoniae*, *A. vinelandii* (Merrick and Coppard, 1989) and *P. putida* (Inouye *et al.*, 1989) have been determined, and are predicted to encode polypeptides of 95, 107, and 102 aa, respectively. Merrick and Coppard (1989) analyzed in *K. pneumoniae* the nucleotide sequence downstream of this ORF, which they called ORF95, and discovered two more possible ORF's, called ORF162 and ORF193. The only other sequence data available which stretches this far downstream of an *ntrA* gene is from *P. putida* (Inouye *et al.*, 1989). Merrick and Coppard (1989) observed an ORF encoding a homologue to the *K. pneumoniae* ORF162 product downstream of the *P. putida* ORF95. The only significant sequence similarity obtained with searches of the predicted products of ORF95, ORF162, and ORF193 against the available databases was between the ORF95 product and the product of URF1, an ORF adjacent to the *pheA* gene of *E. coli* (Hudson and Davidson, 1984), the function of which is unknown. Merrick and Coppard (1989) demonstrated that in an *in vitro* transcription and translation system the *K. pneumoniae* ORF95 and ORF162 genes produced polypeptides of 12 and 16 kDal, respectively. *K. pneumoniae* chromosomal mutations which were constructed in the ORF95 and ORF162 genes resulted in an increase in the level of expression from NtrA-dependent promoters. These results indicated that the products of the ORF95 and ORF162 genes function to modulate the activity of NtrA-RNA polymerase.

CHAPTER 4

BIOLOGICAL ACTIVITY OF THE *T. FERROOXIDANS NTRA* GENE PRODUCT CLONED IN *E. COLI*

CHAPTER 4

BIOLOGICAL ACTIVITY OF THE *T. FERROOXIDANS NTRA* GENE PRODUCT CLONED IN *E. COLI*

4.0 Summary

The *E. coli ntrA* mutant TH1 containing the *T. ferrooxidans ntrA* gene, cloned on a 1,69-kbp *T. ferrooxidans* DNA fragment, produced gas and reduced benzylviologen, an artificial electron acceptor, when grown anaerobically on formate. These phenotypes, which are characteristic of the *E. coli* formate hydrogenlyase pathway, were repressed when formate was replaced by nitrate in the growth media. This pattern of regulation matched that exhibited by the *E. coli ntrA*$^+$ strain YMC10 and the *E. coli* TH1 strain containing the cloned *K. pneumoniae ntrA* gene. Biological activity of the cloned *T. ferrooxidans ntrA* gene product was demonstrated using translational fusions between the NtrA-dependent promoters and N-terminal regions of either the *fdhF* gene or either of two *nifH* genes and the *lacZ* gene. The *T. ferrooxidans* NtrA in the *E. coli* TH1 cells increased expression of a *fdhF-lacZ* fusion 7-fold above the basal levels of expression observed in the absence of a *ntrA* gene. This expression was repressed by replacement of the formate in the growth medium with nitrate. The *T. ferrooxidans* NtrA in the *E. coli* TH1 cells, containing a constitutively expressed *K. pneumoniae nifA* gene, resulted in an approximately 50-fold increase in β-galactosidase activity from either a *T. ferrooxidans nifH-lacZ* fusion or a *K. pneumoniae nifH-lacZ* fusion above levels obtained in the absence of a *ntrA* gene. *E. coli* TH1 cells containing the cloned *K. pneumoniae ntrA* gene or *E. coli* YMC10 *ntrA*$^+$ cells produced a similar increase in activity from the *T. ferrooxidans nifH-lacZ* fusion in the presence of the *K. pneumoniae nifA* gene, but a 400-fold increase from the *K. pneumoniae nifH-lacZ* fusion. Under these conditions, the *T. ferrooxidans* NtrA was therefore as efficient as the *K. pneumoniae* NtrA and *E. coli* NtrA at promoting transcription from the *T. ferrooxidans nifH-lacZ* fusion, but much less efficient at promoting transcription from the *K. pneumoniae nifH-lacZ* fusion.

4.1 Introduction

The use of translational fusions between the regulatory and N-terminal regions of a gene of interest and a reporter gene is a useful stratagem in the study of the regulation of gene expression. This is particularly effective when the activity of the product of the gene of interest is difficult to measure. A favourite reporter gene amongst molecular biologists is the *E. coli lacZ* gene, which codes for the β-galactosidase enzyme (reviewed in Silhavy and Beckwith, 1985). The activity of this enzyme may be determined by a simple colourimetric reaction (Miller, 1972).

Several workers have used translational fusions between NtrA-dependent genes and the *lacZ* gene to study or confirm the activity of a cloned *ntrA* gene product. The cloned *K. pneumoniae ntrA* gene product was able to direct expression in a *K. pneumoniae ntrA* mutant of either a *K. pneumoniae* or a *R. meliloti nifH* gene (de Bruijn and Ausubel, 1983) or the *K. pneumoniae nifL* (Merrick and Stewart, 1985) gene by the increase in expression of these *nif* genes fused to *lacZ*. Ishimoto and Lory (1989) showed that the *P. aeruginosa ntrA* gene product was able to increase transcription in *E. coli* of the *R. leguminosarum dctA* gene, which has a NtrA-dependent promoter, fused to *lacZ*.

Merrick and Stewart (1985) observed a 2,5- to 3,5-fold increase in expression of a *ntrA-lacZ* fusion in *E. coli* in the presence of multiple copies of the *K. pneumoniae ntrA* gene. The reason for this possible auto-activation by the NtrA is not clear although it may not reflect the situation *in vivo* where the *ntrA* gene is present at one copy per chromosome. The expression of *ntrA* genes is thought to be at a low constitutive level irrespective of levels of available nitrogen and therefore any transcriptional regulation of *ntrA* is thought not to be a major factor in expression of NtrA-dependent genes (Merrick and Stewart, 1985; de Bruijn and Ausubel, 1983; Castano and Bastarrachea, 1984). It would be appropriate to note the limitations of regulatory studies using cloned genes on multi-copy plasmids as these are artificial conditions which may not reflect the effect of single-copy chromosomal genes. An example of this is the phenomenon of "multi-copy inhibition" which occurred on studies of *nif* genes in *K. pneumoniae*. It was found that the *K. pneumoniae*

(Riedel *et al.*, 1983) or *A. chroococcum* (Jones *et al.*, 1984) *nifH* promoters, cloned on a high copy-number plasmid, inhibited the expression of the single *nifHDK* operon situated on the *K. pneumoniae* chromosome. This was thought to be the result of titration of the limited amounts of the activator protein, NifA, by the multiple copies of *nifH* regulatory regions.

Genes which require the NtrA sigma factor for expression have been found to contain at least two types of conserved *cis*-acting DNA elements in the region upstream of the gene. These elements are the NtrA-dependent promoter itself which is conserved amongst all genes which require the NtrA sigma factor for expression, and a so-called upstream activator sequence which appears to differ between regulons and be recognized by a regulon-specific activator protein, the activity of which is mediated by a cascade of regulators in response to environmental changes (reviewed in Kustu *et al.*, 1989).

The regulatory region upstream of the *T. ferrooxidans nifH* gene showed both DNA elements required for NtrA-dependent expression. The putative *T. ferrooxidans nifH* NtrA-dependent promoter shows high similarity to the *K. pneumoniae nifH* promoter and other NtrA-dependent promoters (Ch. 5; Fig. 5.1). The *T. ferrooxidans nifH* regulatory region contains two consensus NifA binding sites (sequence -TGT-N_{10}-AGA-) in a region 119 to 73 nucleotides upstream from the NtrA-dependent promoter (Pretorius *et al.*, 1987). The presence of two of these binding sites spaced 30 bp apart (from center to center) is interesting as this is about three turns of the DNA helix, which suggests possible co-operative binding of the activator protein. This is the same distance as the spacing for the NR_I activator sites upstream of the *E. coli glnA* promoter (Reitzer and Magasanik, 1985). A current model of NifA activation in *K. pneumoniae* involves the binding of integration host factor (IHF) to a region between the NifA binding site and the NtrA-dependent promoter, which causes bending of the DNA and therefore facilitates productive interactions between NifA and NtrA-core polymerase (section 1.3.2.3; Santero *et al.*, 1989; Cannon *et al.*, 1990). Ongoing experiments have shown that the *E. coli* integration host factor is able to bind to the *T. ferrooxidans nifH* upstream regulatory region (Hoover *et al.*, 1990). Gel mobility shift experiments and hydroxyl radical

footprints with purified *E. coli* IHF showed that a region present between the NifA site and the NtrA-dependent promoter of several *nif* genes - including the *T. ferrooxidans nifH* gene -, which showed some similarity to the consensus binding site for *E. coli* IHF (derived from binding sites in lamboid phages), was the site of *E. coli* IHF binding (Hoover *et al.*, 1990; Santero *et al.*, Eighth International Conference on Nitrogen Fixation, 1990).

The similarity of the *T. ferrooxidans nifH* gene promoter-regulatory region to *K. pneumoniae nif* gene regulatory regions and the availability of the *T. ferrooxidans nifH* gene cloned in *E. coli* prompted investigation of whether the cloned *T. ferrooxidans ntrA* gene product was able to direct expression from an isogenic *nifH* promoter region. This was carried out using a translational fusion between the *T. ferrooxidans nifH* gene and the *lacZ* gene. For comparison, matching experiments were carried out using a *K. pneumoniae nifH-lacZ* fusion.

The syntheses of the formate dehydrogenase and hydrogenase components of the formate hydrogenlyase pathway in *E. coli* are known to be repressed by oxygen, repressed anaerobically by nitrate, and induced by formate (Wimpenny and Cole, 1967). Pecher *et al.* (1983) demonstrated that this regulation was at the level of transcription and it was subsequently found that the effects of oxygen, nitrate, and formate could not be separated physically on deletion analysis of the regulatory region of the *fdhF* gene (Birkmann *et al.*, 1987b). Birkmann and Böck (1989) have shown that regulation at least by formate induction is mediated by a regulatory sequence of about 25 bp in length located 110 bp 5' to the transcription start site of the *fdhF* gene. This sequence represents the second regulatory element other than the NtrA-dependent promoter which controls expression of the formate hydrogenlyase pathway. A regulatory protein(s) which acts at this site has yet to be identified. The products of *hydG* (Stoker *et al.*, 1989) and *fhlA* (Sankar *et al.*, 1988; ORFE - Böhm *et al.*, 1990), if different, are candidates.

Two types of formate dehydrogenase activities have been identified in *E. coli*, and these may be distinguished by their activities on artificial electron acceptors (reviewed in Stewart, 1988). The formate dehydrogenase-N, which is found predominantly in

oxygen- and nitrate-respiring cells, is particularly active with methylene blue or phenazin methosulphate as an artificial electron acceptor. In contrast, the formate dehydrogenase-H, which is a component of the fermentative formate hydrogenlyase pathway, is active with benzylviologen as an artificial electron acceptor (Peck and Guest, 1957). Mandrand-Berthelot *et al.* (1978) developed a dye-overlay technique that could be used directly on growth plates for the detection of *E. coli* cells exhibiting the benzylviologen-linked formate dehydrogenase activity. The TGYEP overlay technique described in Ch. 2 for the detection of gas positive *E. coli* colonies and the dye-overlay technique for the detection of formate dehydrogenase-H activity provided two simple phenotypic tests to determine whether the *E. coli ntrA* mutant strain containing the *T. ferrooxidans ntrA* gene showed the same pattern of formate and nitrate regulation of formate hydrogenlyase activity as a *ntrA*$^+$ *E. coli* strain. The effect of the cloned *T. ferrooxidans ntrA* gene product on the expression of the *E. coli fdhF* gene was also studied using a *fdhF-lacZ* fusion plasmid.

This chapter describes experiments to demonstrate the biological activity of the *T. ferrooxidans ntrA* gene product cloned in *E. coli*. Phenotypes characteristic of the formate hydrogenlyase pathway and the expression of a *fdhF-lacZ* fusion were studied to show that the *T. ferrooxidans ntrA* gene complemented the *E. coli ntrA* mutant for formate hydrogenlyase activity and that this activity was subject to the same regulation pattern as a wildtype *E. coli* strain. The effect of the cloned *T. ferrooxidans ntrA* gene product on the expression of either a *T. ferrooxidans* or a *K. pneumoniae nifH-lacZ* fusion in *E. coli* was also studied.

4.2 Materials and Methods

4.2.1 Bacterial strains, plasmids, and media

The strains and plasmids used in this study are described in Appendix A. Media used for growth of *E. coli* strains is described in the text and Appendix B. When required, antibiotics were added at the following concentrations: ampicillin, 100 µg/ml; tetracycline, 15 µg/ml; chloramphenicol, 20 µg/ml.

4.2.2 Phenotypic tests for formate hydrogenlyase activity

E. coli strains were toothpicked onto duplicate TGYEP agar plates at pH 6,5 containing either 30 mM formate (TGYEPF plates) or 40 mM nitrate (TGYEPN plates), and tested for either gas production by the TGYEPF agar overlay technique as described in Ch. 2 (section 2.2.6.1) or for benzylviologen reduction by the benzylviologen overlay technique as described by Mandrand-Berthelot *et al.* (1978). This involved incubation overnight at 37°C under anaerobic conditions of the *E. coli* strains toothpicked onto the TGYEPF and TGYEPN plates. Benzylviologen reduction was tested by pouring over the growth plates a dye-overlay solution held at 45°C and containing agar (0,8% w/v), benzylviologen (1 mg/ml), and KH_2PO_4 (25 mM; pH 7,0). Colonies of *E. coli* strains which were able to reduce benzylviologen were immediately identifiable by the conversion of the colourless benzylviologen dye in the agar overlay surrounding each colony to a bright purple colour. Appropriate antibiotic selection was provided.

4.2.3 Growth conditions for β-galactosidase assays

E. coli strains containing the *fdhF-lacZ* fusion plasmid, pBN208, were grown in a pre-culture of LB for 8 h at 37°C under aerobic conditions. This pre-culture was washed in 0,8% saline and the optical density at 600 nm (OD_{600}) was measured to enable inoculation of a sample of the washed pre-culture into the assay culture media to ensure a starting OD_{600} of 0,05 units which was constant for each *E. coli* strain. The assay culture media for *E. coli* (pBN208) strains was liquid TGYEPF or TGYEPN (pH 6,5), containing 0,2% glutamine (w/v) because of the glutamine

auxotrophy of the *E. coli ntrA* mutant strain TH1. These cultures were grown under anaerobic conditions at 37°C for 15 h before the β-galactosidase assays were carried out.

E. coli strains containing the *nif-lacZ* fusion plasmids were grown by an adaptation of the method of Ow and Ausubel (1983). These strains were grown aerobically in a LB pre-pre-culture at 30°C for 8 h, washed in 0,8% saline, and inoculated into the pre-culture media at a starting $OD_{600} = 0,2$. The pre-culture medium consisted of NFDM (Appendix B) containing 500 µg/ml casamino acids, 2 mg/ml glutamine, and 2 mg/ml $(NH_4)_2SO_4$. Growth was continued in this nitrogen rich medium at 30°C in a shaking waterbath under anaerobic conditions to saturation (15 h), after which the cultures where washed and inoculated at a starting $OD_{600} = 0,5$ into nitrogen limiting medium, which consisted of NFDM supplemented with 100 µg/ml glutamine. Growth was continued for 10 h under anaerobic conditions in a shaking waterbath at 30°C before determination of the β-galactosidase activity.

4.2.4 β-Galactosidase assays

β-Galactosidase activity of the *E. coli* cultures was determined by an adaptation of the method of Miller (1972). All measurements of β-galactosidase activity reported in this thesis were the result of at least three separate determinations. A method was developed which enabled rapid processing of multiple samples with volumes convenient for use in 1,5 ml Eppendorf tubes. Cells were washed in 0,8% saline and resuspended in phosphate buffer (0,1 M K_2HPO_4; pH 7,0). Several serial dilutions of the washed cells were assayed in 250 µl amounts with an equal volume of Z buffer (Appendix B). This involved permeabilization of the cells by addition of 50 µl 0,1% (w/v) sodium dodecyl sulphate (SDS) and 50 µl of chloroform, followed by an equilibration incubation at 28°C for 5 min after which the reaction was initiated by the addition of 100 µl of the substrate *o*-nitrophenyl-β-D-galactopyranoside (ONPG)(4 mg/ml) in phosphate buffer. The mixture was incubated at 28°C until the appearance of a pale yellow colour which usually occurred after 5 - 15 min of incubation. The reaction was stopped by the

addition of 250 µl of Na_2CO_3 (14% w/v). Cell debris was sedimented by centrifugation in a bench-top microfuge for 2 min and the pale yellow supernatant was removed and diluted with water (1 ml) to provide a convenient volume for spectrophotometric analysis. The OD_{420} was measured to determine the amount of *o*-nitrophenol released and the OD_{550} was also determined to obtain a measure of the influence of the cell debris on the OD_{420}. Miller (1972) states that the factor (1,75 x OD_{550}) gives a close approximation to the absorbance at OD_{420} caused by cell debris. It was found that the inclusion of the centrifugation step in the protocol reduced this factor to zero in most cases. The β-galactosidase activity was expressed in Miller units which were calculated using the formula (Miller, 1972) where the reaction time (t) is measured in min and the culture volume assayed (v) in ml:

$$\text{Miller units} = 1000 \times \frac{OD_{420} - (1.75 \times OD_{550}) \times \text{dilution factor}}{t \times v \times OD_{600}}$$

4.2.5 Subcloning of the *T. ferrooxidans ntrA* gene

Fig. 4.1 shows the 1,69-kbp *Cla*I(position 802 on Fig. 3.2)-*Eco*RI fragment containing the *T. ferrooxidans ntrA* gene together with its putative promoter region which was subcloned into the relevant plasmid vectors to demonstrate the biological activity of the product of the *T. ferrooxidans ntrA* gene in the absence of the ORF1 and ORF3 coding sequences. This fragment was derived from pT30-D which was isolated as a result of exonuclease III shortening of pT30 from the *Bgl*II end to nucleotide 2491 (described in section 3.2.7; Fig. 3.1). The insertion of this fragment into the Bluescript KS^+ and SK^+ vectors to produce pT40 and pT41 has been described in section 3.2.7. The tetracycline resistance gene of the vector pACYC184 (Appendix F) was replaced by this 1,69-kbp *Cla*I-*Eco*RI fragment, to produce pT50. This was achieved by ligation of this 1,69-kbp *Cla*I-*Eco*RI (filled-in using Klenow at the *Eco*RI 5' overhang) fragment into *Cla*I and *Hind*II (the site 3' to the tetracycline resistance gene) digested pACYC184.

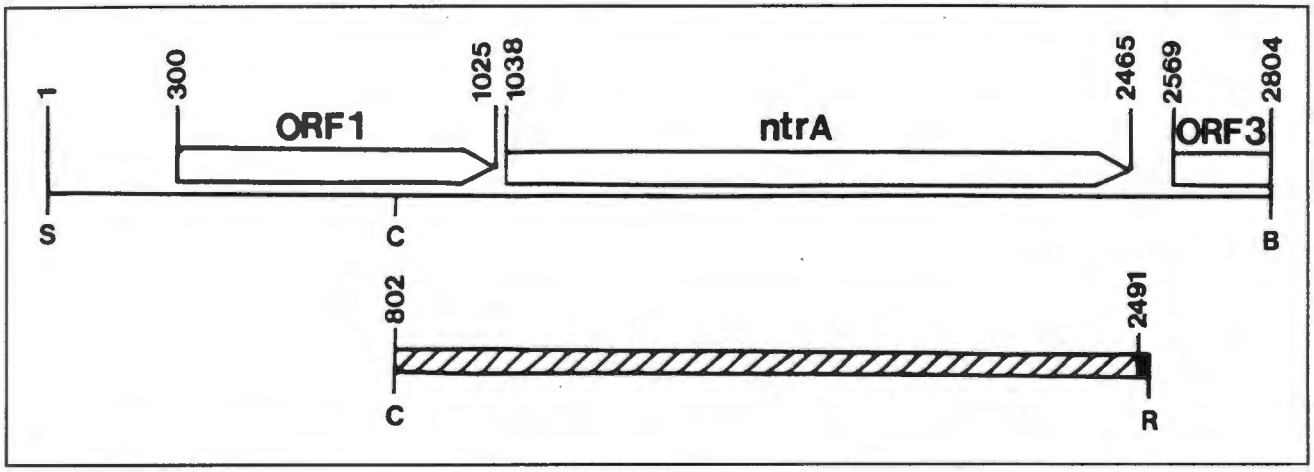


Fig. 4.1. Origin of the *T. ferrooxidans ntrA* gene subclone. The single line represents the 2,8-kbp *T. ferrooxidans* DNA insert in pT30. Numerals refer to nucleotides from the *Sal*I site and match those in Fig. 3.2. **The 1,69-kbp *Cla*I/*Eco*RI fragment from pT30-D used to construct pT40, pT41, and pT50 is shown by a hatched box (*T. ferrooxidans* DNA) and a filled box (pUC19 DNA).** ORF's are indicated by open boxes. B, *Bgl*II; C, *Cla*I; R, *Eco*RI; S, *Sal*I.

4.2.6 Construction of the plasmid vector p184t and subcloning of the *K. pneumoniae ntrA* gene

p184t was a derivative of pACYC184 deleted for the tetracycline resistance gene and a region recently implicated in destabilization of the vector pACYC184 (Kolot *et al.*, 1989). This deletion was achieved by digestion of pACYC184 with *Hind*III restriction endonuclease which resulted in a single cut 5' to the tetracycline resistance gene, fill-in of the *Hind*III 5' overhang with Klenow, followed by digestion with the *Hind*II restriction endonuclease which has a sites within and 3' to the tetracycline resistance gene (Appendix F). p184t was the result of recircularization and ligation of the plasmid at the filled-in *Hind*III end and the blunt end *Hind*II site 3' to the tetracycline resistance gene, thus deleting all of the tetracycline resistance gene coding sequences. Deletion of the tetracycline resistance gene was necessary since this was the marker carried on the plasmid pCK3 carrying the *K. pneumoniae nifA* gene. The *K. pneumoniae ntrA* gene on a 1,69-kbp *Cla*I fragment from the plasmid pFB71 was cloned into the *Cla*I site of p184t to produce pK50.

4.2.7 Construction of a translational fusion between the *T. ferrooxidans nifH* gene and the *E. coli lacZ* gene

A 1,2-kbp *Eco*RI-*Xho*I fragment from pIMP11 (Pretorius *et al.*, 1987) was cloned into the *Eco*RI site of pMC1403 (Casadaban *et al.*, 1983) to produce a translational fusion between the *T. ferrooxidans nifH* and the *E. coli lacZ* gene. The strategy employed is shown in Fig. 4.2. The translational fusion vector pMC1403 contains a multiple cloning site with unique sites for the restriction endonucleases *Eco*RI, *Sma*I, and *BamH*I 5' to codon 8 of the *E. coli lacZ* gene. Consequently, the *lacZ* gene lacks a promoter region, a ribosome binding site, and an ATG start codon. Insertion of a fragment of DNA which contains all of these elements and choice of restriction sites so that the resultant fusion is in frame with the *lacZ* gene produces a recombinant clone which is positive for β-galactosidase activity.

A 1,2-kbp *Eco*RI-*Xho*I fragment from pIMP11 which contained the promoter region and N-terminus of the *T. ferrooxidans nifH* gene was ligated to *Eco*RI digested pMC1403 to produce a linear molecule. The remaining 5' overhangs of the insert *Xho*I site and the vector *Eco*RI site were filled-in with Klenow to enable a blunt-end recircularization ligation at these ends to produce an in-frame translational fusion between codon 75 of the *T. ferrooxidans nifH* gene and the *lacZ* gene.

This junction site was checked by nucleotide sequencing using a synthetic primer (5'-CG-CCA-GGG-TTT-TCC-CAG-3') (a gift from Prof. D. Botes, Biochemistry Department, University of Cape Town, South Africa) which is complementary to the sequence situated between +42 and +64 nucleotides within the *lacZ* gene and therefore may be conveniently used to obtain sequence data at the 5' end of the *lacZ* gene.

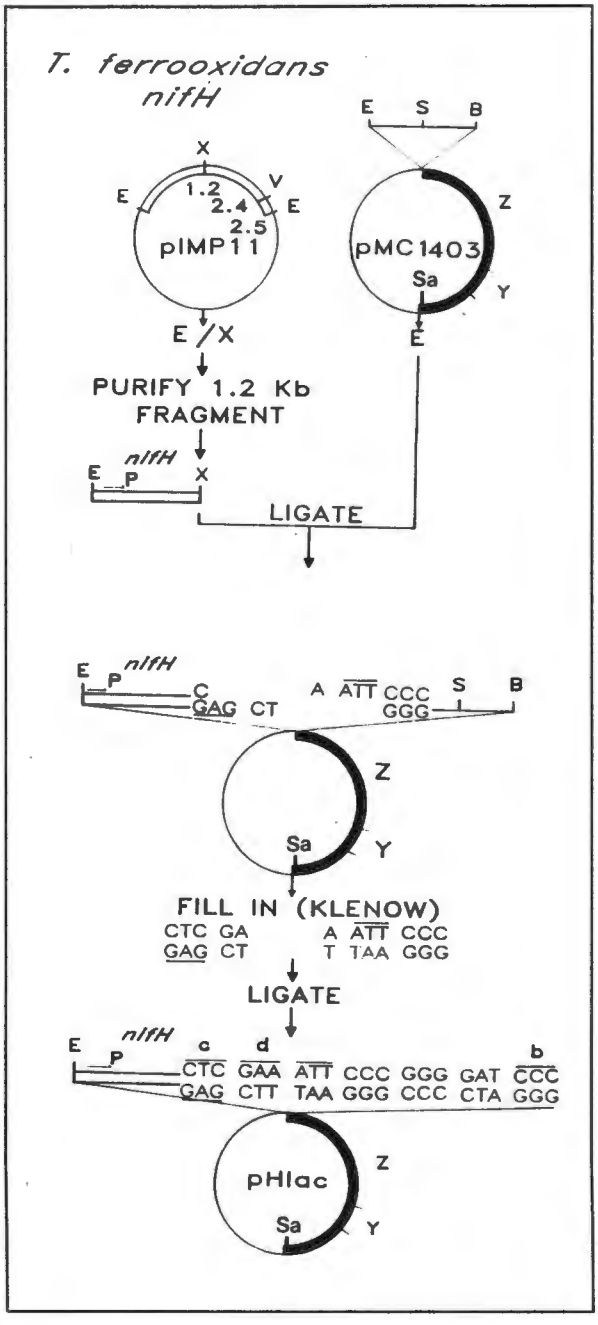


Fig. 4.2. Construction of pHlac as described in the text. *T. ferrooxidans* DNA is represented by a double line. The *E. coli lacZYA* genes are represented by a thick line. Abbreviations: b, codon 8 of *lacZ*; c, codon 75 of *T. ferrooxidans nifH*; d, chimeric codon; P, NtrA-dependent promoter; B, *Bam*HI; E, *Eco*RI; Sa, *Sal*I; S, *Sma*I; V, *Eco*RV; X, *Xho*I.

4.3 Results

4.3.1 Regulation of formate hydrogenlyase activity in *E. coli* TH1 cells containing the cloned *T. ferrooxidans ntrA* gene

Experiments were carried out to determine whether the pattern of formate and nitrate regulation of the formate hydrogenlyase pathway were the same in the mutant *E. coli* TH1 strain containing the cloned *T. ferrooxidans ntrA* gene as in the parent strain, *E. coli* YMC10. Phenotypic tests were carried out on the different *E. coli* strains with and without the cloned *T. ferrooxidans ntrA* gene to check for the gas positive phenotype and the ability to reduce benzylviologen (Table 4.1).

Table 4.1. Phenotypic tests for formate hydrogenlyase activity.

Strain[a]	Relevant genotype[b]	Gas[c]	Benzylviologen reduction[d]	Gas[c]	Benzylviologen reduction[d]
		30 mM formate		40 mM nitrate	
TH1(pT40)	*Tf ntrA*$^+$	+	+	nd	nd
TH1(pT41)	*Tf ntrA*$^+$	+	+	nd	nd
TH1(B-SK)	Δ *ntrA*	-	-	nd	nd
TH1(p184t)	Δ *ntrA*	-	-	-	-
TH1(pT50)	*Tf ntrA*$^+$	+	+	-	-
TH1(pK50)	*Kp ntrA*$^+$	+	+	-	-
YMC10(p184t)	*Ec ntrA*$^+$	+	+	-	-
FM911(p184t)	*Ec ntrA*$^+$				
	Δ *fdhF*	-	-	-	-
FM911(pT50)	*Tf ntrA*$^+$				
	Δ *fdhF*	-	-	-	-
FM911(pK50)	*Kp ntrA*$^+$				
	Δ *fdhF*	-	-	-	-

[a] *E. coli* strains were grown anaerobically on TGYEP agar (pH 6,5), containing either 30 mM formate or 40 mM nitrate.

[b] *Tf, T. ferrooxidans; Kp, K. pneumoniae; Ec, E. coli*; nd, not determined; B-SK, plasmid vector Bluescript-SK.

[c] Gas production was scored positive (+) on formation of a gas pocket around a colony growing beneath a TGYEP agar overlay.

[d] Reduction of benzylviologen was scored positive (+) on the conversion of the benzylviologen dye from colourless to purple within one minute of overlaying colonies grown anaerobically overnight on TGYEP agar.

The *E. coli* TH1 strain containing the plasmid vector p184t was unable to produce gas or reduce benzylviologen when grown with formate or nitrate (Table 4.1). In contrast, *E. coli* YMC10 and TH1 containing either the *T. ferrooxidans* or the *K. pneumoniae ntrA* gene were gas positive and were able to reduce benzylviologen when grown with formate (Table 4.1). In the presence of nitrate both gas production and benzylviologen reduction were repressed (Table 4.1). The inability of the *E. coli fdhF* mutant FM911, containing either the *T. ferrooxidans* or the *K. pneumoniae ntrA* gene, to produce gas or reduce benzylviologen indicated that a functional *fdhF* gene was required for *ntrA* complementation of formate hydrogenlyase phenotypes (Table 4.1).

4.3.2 Expression of the *T. ferrooxidans ntrA* gene in *E. coli*

E. coli TH1 cells containing the *T. ferrooxidans ntrA* gene cloned in both orientations with respect to the Bluescript vector *lacZ* gene (pT40, pT41) produced gas and reduced benzylviologen when grown with formate (Table 4.1). It was concluded that the *T. ferrooxidans ntrA* gene was expressed from its own regulatory region in *E. coli*.

4.3.3 Expression of a *fdhF-lacZ* fusion in *E. coli* in the presence of the *T. ferrooxidans ntrA* gene

The biological activity of the *T. ferrooxidans ntrA* gene product in *E. coli* was tested by investigation of the expression of a translational fusion between the N-terminus of the *E. coli fdhF* gene and the *lacZ* gene, which codes for the β-galactosidase enzyme. This *fdhF-lacZ* fusion was carried on the plasmid pBN208 (Appendix A). Positive controls were the cloned *K. pneumoniae ntrA* gene and the parent *E. coli* YMC10 strain which has a chromosomal *ntrA* gene. When *E. coli* TH1(pBN208) cells containing the *T. ferrooxidans ntrA* gene on a compatible plasmid were grown anaerobically with formate the expression of the *fdhF-lacZ* fusion was increased 7-fold above the basal level of expression obtained in the absence of a *ntrA* gene (Rows 1 and 2; Table 4.2). A similar increase in β-galactosidase activity was obtained in the presence of either the *K. pneumoniae* or the *E. coli ntrA* gene (compare rows 3 and 4 with row 2; Table 4.2). β-galactosidase activity was repressed when the cultures were grown anaerobically with nitrate (Table 4.2).

Table 4.2. Effect of the *T. ferrooxidans* NtrA on expression of β-galactosidase activity from a *fdhF-lacZ* fusion plasmid in *E. coli*.

	Strain[a]	Relevant genotype	β-galactosidase activity (Miller units)	
			30 mM formate	40 mM nitrate
1	TH1(pBN208, pT50)	*T. ferrooxidans ntrA* (*fdhF-lacZ*)	2267	180
2	TH1(pBN208, pACYC184)	Δ *ntrA* (*fdhF-lacZ*)	322	102
3	TH1(pBN208, pFB71)	*K. pneumoniae ntrA* (*fdhF-lacZ*)	2514	260
4	YMC10(pBN208, pACYC184)	*E. coli ntrA* (*fdhF-lacZ*)	3761	359

[a] *E. coli* strains were grown anaerobically in TGYEP medium (pH 6,5) + 0,2% (w/v) glutamine + 0,8% glucose, containing either 30 mM formate or 40 mM nitrate.

4.3.4 Expression of *T. ferrooxidans and K. pneumoniae nifH-lacZ* fusions in *E. coli* in the presence of the *T. ferrooxidans ntrA* gene and the *K. pneumoniae nifA* gene

The biological activity of the *T. ferrooxidans ntrA* gene product was investigated further by analysis of the expression of translational fusions between the N-terminus regions of *nifH* genes from *T. ferrooxidans* and *K. pneumoniae*, and the reporter gene, *lacZ*. Positive controls were the cloned *K. pneumoniae ntrA* gene and the parent *E. coli* YMC10 strain which has a chromosomal *ntrA* gene. NtrA-dependent expression from *nifH* promoters requires the presence of an activator, encoded by the *nifA* gene. As there is no *nifA* gene in *E. coli*, this was provided by the constitutively expressed *K. pneumoniae nifA* gene on the plasmid pCK3. Low basal levels of β-galactosidase activity were obtained from *E. coli* TH1 containing the vector p184t and compatible plasmid vectors carrying the *K. pneumoniae nifA* gene (pCK3) , and either the *K. pneumoniae nifH-lacZ* fusion (pMB1) or the *T. ferrooxidans nifH-lacZ* fusion (pHlac) when grown anaerobically in nitrogen limited medium (Rows 2 and 6; Table 4.3). Replacement of the vector p184t by the plasmid pT50 (carrying the *T. ferrooxidans ntrA* gene) increased expression of both *nifH-lacZ* fusions more than

50-fold above basal levels (compare row 1 with row 2, and row 5 with row 6; Table 4.3). The cloned *T. ferrooxidans ntrA* gene product was less efficient at promoting expression of the *K. pneumoniae nifH-lacZ* fusion than was the cloned *K. pneumoniae ntrA* gene product or the chromosomally-encoded *E. coli ntrA* gene product (compare row 1 with rows 3 and 4; Table 4.3). All three *ntrA* gene products produced similar levels of β-galactosidase activity from the *T. ferrooxidans nifH-lacZ* fusion (Rows 5, 7, and 8; Table 4.3).

Table 4.3. Effect of *T. ferrooxidans* NtrA on expression of β-galactosidase activity from *nifH-lacZ* fusion plasmids in the presence of the *K. pneumoniae nifA* gene in *E. coli*.

	Strain[a]	Relevant genotype[b]	β-galactosidase activity (Miller units)
1	TH1(pCK3, pMB1, pT50)	*T. ferrooxidans ntrA*$^+$ (*KpnifH-lacZ*) *KpnifA*^{c}	1 484
2	TH1(pCK3, pMB1, p184t)	Δ *ntrA* (*KpnifH-lacZ*) *KpnifA*^{c}	28
3	TH1(pCK3, pMB1, pK50)	*K. pneumoniae ntrA*$^+$ (*KpnifH-lacZ*) *KpnifA*^{c}	13 333
4	YMC10(pCK3, pMB1, p184t)	*E. coli ntrA*$^+$ (*KpnifH-lacZ*) *KpnifA*^{c}	10 113
5	TH1(pCK3, pHlac, pT50)	*T. ferrooxidans ntrA*$^+$ (*TfnifH-lacZ*) *KpnifA*^{c}	1 642
6	TH1(pCK3, pHlac, p184t)	Δ *ntrA* (*TfnifH-lacZ*) *KpnifA*^{c}	32
7	TH1(pCK3, pHlac, pK50)	*K. pneumoniae ntrA*$^+$ (*TfnifH-lacZ*) *KpnifA*^{c}	2 042
8	YMC10(pCK3, pHlac, p184t)	*E. coli ntrA*$^+$ (*TfnifH-lacZ*) *KpnifA*^{c}	2 044

[a] Strains were grown anaerobically in nitrogen-limiting conditions as described in the text (section 4.2.4).

[b] *Kp*, *K. pneumoniae*; *Tf*, *T. ferrooxidans*.

[c] Represents constitutive expression of the *K. pneumoniae nifA* gene.

4.4 Discussion

A recombinant cosmid clone which complemented an *E. coli ntrA* mutant for gas production was cloned and the nucleotide sequence of a 2,8-kbp fragment containing the *T. ferrooxidans ntrA* gene was determined (Chapters 2 and 3). As the *T. ferrooxidans ntrA* gene was not able to complement the *E. coli ntrA* mutant for growth on minimal medium and arginine, as had been used for the isolation of the *K. pneumoniae, A. vinelandii*, and *P. putida ntrA* genes, experiments were carried out to show that the cloned *T. ferrooxidans ntrA* gene could promote transcription from NtrA-dependent promoters. A 1,69-kbp subclone of the sequenced *T. ferrooxidans* DNA insert, which contained the coding sequence for the *T. ferrooxidans ntrA* gene, was used to confirm that this was the *T. ferrooxidans* genetic determinant responsible for the complementation of the gas positive phenotype in the *E. coli ntrA* mutant. Phenotypic tests showed that this complementation was due to NtrA-dependent expression of the *E. coli* gas producing formate hydrogenlyase pathway by the cloned *T. ferrooxidans ntrA* gene product. This complementation was tested by observing the ability of the *E. coli ntrA* mutant containing the cloned *T. ferrooxidans ntrA* gene to produce gas and reduce benzylviologen when grown anaerobically in the presence of formate, and the repression of both activities on the replacement of formate by nitrate in the growth medium.

A direct demonstration that the cloned *T. ferrooxidans* NtrA protein could promote expression from the NtrA-dependent promoters of the genes encoding components of the *E. coli* formate hydrogenlyase pathway was obtained using a translational fusion between the *E. coli fdhF* gene, which encodes the selenopolypeptide component of the formate dehydrogenase, and a reporter gene, *lacZ*. This enabled the quantitative measurement of levels of expression from the *fdhF* NtrA-dependent promoter in the presence and absence of the cloned *T. ferrooxidans ntrA* gene. The *T. ferrooxidans* NtrA was able to promote expression from the *E. coli fdhF* promoter to levels equivalent to that achieved by either the *E. coli* or the *K. pneumoniae* NtrA proteins as measured by an increase in β-galactosidase activity from the *fdhF-lacZ* fusion product (Table 4.2).

The existence of a formate hydrogenlyase pathway in *T. ferrooxidans* is unlikely because in *E. coli* it is a non-energy conserving pathway which would be incompatible with the stringent energy budget associated with *T. ferrooxidans*' chemolithoautotrophic way of life. Biological activity of the cloned *T. ferrooxidans* NtrA was therefore investigated with respect to expression of the NtrA-dependent promoter of the *nifH* gene, a component of the nitrogen fixation system, which is known to occur in *T. ferrooxidans*. The increase in expression of a *T. ferrooxidans nifH-lacZ* fusion in the *E. coli ntrA* mutant directed by the *T. ferrooxidans* NtrA was equivalent to that achieved by either the *K. pneumoniae* or *E. coli* NtrA proteins (Table 4.3). Furthermore, this activation required the presence of the *K. pneumoniae nifA* gene which indicates that the *K. pneumoniae* NifA protein is able to recognize the *T. ferrooxidans nifH* upstream activator site(s) which is a requirement for NtrA-dependent expression from *nifH* promoters. The weak expression from the *K. pneumoniae nifH* promoter directed by the *T. ferrooxidans NtrA* (Table 4.3) could be due to either poor recognition of the *nifH* promoter by the heterologous *T. ferrooxidans* NtrA complexed to the *E. coli* core RNA polymerase, or inefficient interaction between the *T. ferrooxidans* NtrA protein and the *K. pneumoniae* NifA protein. However, strict comparisons of NtrA efficiency should not be made because the *ntrA* genes are expressed at different levels and expression of each fusion is dependent on NifA, NtrA, and core RNA polymerase coded for by genes from different bacteria.

The ability of the *T. ferrooxidans* NtrA to direct expression from the *T. ferrooxidans nifH* promoter region in *E. coli* provides evidence that NtrA is involved in regulation of at least one aspect of nitrogen metabolism in *T. ferrooxidans*, namely that of nitrogen fixation. This is interesting as no NtrA-dependent promoter has been identified upstream of the *T. ferrooxidans glnA* gene (Rawlings *et al.*, 1987), which codes for a key nitrogen metabolism enzyme, glutamine synthetase.

CHAPTER 5

CLONING OF THE *T. FERROOXIDANS NTRC* GENE

CHAPTER 5

CLONING OF THE *T. FERROOXIDANS NTRC* GENE

5.0 Summary

Transduction of an *E. coli ntrC* mutant with a *T. ferrooxidans* cosmid gene bank, and selection for growth on arginine as a sole nitrogen source did not yield the *T. ferrooxidans ntrC* gene. Low levels of expression from a *T. ferrooxidans nifH-lacZ* fusion plasmid in an *E. coli ntrC* background were observed, while the presence of an *E. coli* or *K. pneumoniae ntrC* gene resulted in increased β-galactosidase activity when the cells were grown in nitrogen-limiting conditions. An agar plate assay for the detection of *E. coli ntrC* cells containing the *T. ferrooxidans nifH-lacZ* fusion which showed increased levels of β-galactosidase activity in the presence of a cloned *ntrC* gene was developed. *E. coli ntrC* (*T. ferrooxidans nifH-lacZ* fusion plasmid) cells were transduced with the *T. ferrooxidans* cosmid gene bank and colonies were grown on a low-nitrogen minimal medium. Cells which showed increased β-galactosidase activity by virtue of formation of a strong yellow colour after flooding with *o*-nitrophenyl-β-galactopyranoside (ONPG) were shown to contain the *T. ferrooxidans ntrC* gene on a recombinant cosmid. The *T. ferrooxidans ntrC* gene was subcloned on a 5,2-kbp *Kpn*I-*Hind*III fragment. Nucleotide sequence analysis was used to localize the *T. ferrooxidans ntrC* gene, part of which was situated on a 486 bp *Cla*I-*Bam*HI fragment. The predicted translation product of this section of the *T. ferrooxidans ntrC* gene showed homology to the N-terminal region of NtrC proteins from other bacteria, with the greatest similarity to NtrC proteins from enteric bacteria (76 and 72% similarity with NtrC from *K. pneumoniae* and *E. coli*, respectively). Expression in *E. coli* of the *T. ferrooxidans nifH-lacZ* fusion in the presence of the *T. ferrooxidans ntrC* gene was regulated by nitrogen in both solid and liquid media. Southern hybridization against *T. ferrooxidans* chromosomal DNA confirmed the origin of the *T. ferrooxidans ntrC* gene.

5.1 Introduction

NtrC, the product of the *ntrC* gene, is the transcriptional regulator at the heart of the Ntr regulatory system in enteric bacteria (described in detail in section 1.2.2). The importance of NtrC in the general regulation of nitrogen metabolism in enteric bacteria, the identification of *ntrC* genes in non-enteric bacteria, and the role of NtrC at the first level of regulation of expression of the *nif* genes in *K. pneumoniae* has led to intense interest in this regulatory protein. Overexpression of the *ntrC* gene in *E. coli*, and the ease with which this protein may be purified in milligram amounts (Reitzer and Magasanik, 1983; Hawkes *et al.*, 1985) has enabled dissection of the molecular mechanisms of NtrC-mediated regulation in great detail. Another nitrogen metabolism transcriptional activator, the *K. pneumoniae* NifA protein, appears to be predominantly in an insoluble form after most purification protocols (Austin *et al.*, 1990; Hoover *et al.*, 1990). In this chapter the terminology *ntrC* (as opposed to *glnG*) and NtrC (as opposed to GlnG or NR_I) has been used.

Complementation studies and phenotypic characterization of *ntrC* mutant strains are providing some answers as to the role of NtrC in nitrogen regulation in non-enteric bacteria. NtrC appears to be required for growth on nitrate as a sole nitrogen source in *A. vinelandii* (Toukdarian and Kennedy, 1986), *Azorhizobium caulinodans* (Pawlowski *et al.*, 1987), *R. meliloti* (Szeto *et al.*, 1987), *A. tumefaciens* (Rossbach *et al.*, 1987), but not *R. leguminosarum* (Moreno *et al.*, Eighth International Conference on Nitrogen Fixation, 1990). Growth on arginine, histidine, or proline as sole nitrogen sources is impaired in *ntrC* mutants of *A. caulinodans* (Pawlowski *et al.*, 1987), while the chromosomal but not the Ti-encoded arginine catabolism pathway is defective in *A. tumefaciens ntrC* strains (Rossbach *et al.*, 1987). In members of the family *Rhizobiaceae*, such as *R. meliloti*, *B. japonicum*, and *A. tumefaciens*, the *glnII* gene is subject to transcriptional regulation by NtrC in response to nitrogen levels, while the *glnA* gene is not (de Bruijn *et al.*, 1989; Martin *et al.*, 1988; Rossbach *et al.*, 1987). In *R. meliloti*, at least, the *glnT* locus appears also to be subject to NtrC control (de Bruijn *et al.*, 1989).

Current models portray NtrC as responsible for nitrogen regulation in the free-living state of rhizobia species, but its involvement in symbiotic nitrogen fixation varies between species. NtrC is essential not only for N_2-fixation in free-living *A. caulinodans* cells, but also for *nif*-derepression during symbiosis and nodule development on the legume *Sesbania rostrata* (Pawlowski *et al.*, 1987). NtrC activated transcription of the *R. meliloti nifHDK* and *fixABC* operons has been shown in free-living cells, although the significance of this is not clear as acetylene reduction by free-living *R. meliloti* cultures has not been demonstrated in the laboratory (Szeto *et al.*, 1987). However, it is clear that NtrC is not required for symbiotic N_2-fixation by *R. meliloti* (Szeto *et al.*, 1987). N_2-fixation by *A. vinelandii* was independent of NtrC (Toukdarian and Kennedy, 1986), while a *R. capsulatus nifR1* mutant was Nif$^-$ (Jones and Haselkorn, 1989).

NtrC genes have been cloned from the enteric bacteria, *E. coli*, *S. typhimurium*, and *K. pneumoniae* by virtue of their linkage to the *glnA* gene (Covarrubias and Bastarrachea, 1983; Kustu *et al.*, 1986; de Bruijn and Ausubel, 1981; Espin *et al.*, 1982). The *ntrC* gene from *A. vinelandii* was isolated by complementation of the *E. coli ntrC* strain ET8556 and selection for growth on arginine as a sole source of nitrogen (Toukdarian and Kennedy, 1986). The *glnA* gene was found to be closely linked to the *ntrC* gene in this non-enteric bacterium (Toukdarian and Kennedy, 1986). *NtrBC*-like genes (*nifR2* and *nifR1*) from the photosynthetic bacterium *R. capsulatus* were identified by complementation of a Nif$^-$ mutant (Avtges *et al.*, 1985; Jones and Haselkorn, 1989). A *ntrC*::Tn5 mutant of *R. leguminosarum* was found to be defective in GSII activity, and complementation of this phenotype yielded the *ntrC* gene (Moreno *et al.*, and Defez *et al.*, Eighth International Conference on Nitrogen Fixation, 1990).

The *R. meliloti* and *Bradyrhizobium* sp. [*Parasponia*] *ntrC* genes were isolated using Southern hybridization with the *E. coli ntrC* gene as a probe (Nixon *et al.*, 1986; Szeto *et al.*, 1987). The *ntrC* gene of *A. caulinodans* was identified and cloned using the *K. pneumoniae nifA* gene as a hybridization probe (Pawlowski *et al.*, 1987). The *A. tumefaciens ntrC* gene was isolated by two separate groups using either the *A. caulinodans* (Rossbach *et al.*, 1987) or the *R. meliloti* (Wardhan *et al.*, 1989) *ntrC*

gene as a hybridization probe. A strong hybridization signal was observed on probing of *B. japonicum* chromosomal DNA with the *Bradyrhizobium* sp. [*Parasponia*] *ntrC* gene, and this was used as a basis for the isolation of the *B. japonicum ntrC* gene (Martin *et al.*, 1988).

The use of interspecies hybridization as a technique for the isolation of *ntrC* genes indicates that there is extensive homology between *ntrC* genes from diverse prokaryotic groups. This is borne out by experiments which showed positive hybridization signals on screening of chromosomal DNA digests from *Agrobacterium radiobacter*, *Agrobacterium rubi*, *Agrobacterium rhizogenes*, and *A. eutrophicus* with the *A. tumefaciens ntrC* gene as a probe (Wardhan *et al.*, 1989).

The nucleotide sequences of *ntrC* genes from *E. coli* (Miranda-Rios *et al.*, 1987), *K. pneumoniae* (Buikema *et al.*, 1985; Drummond *et al.*, 1986), *Bradyrhizobium* sp. [*Parasponia*] (Nixon *et al.*, 1986), *R. meliloti* (Szeto *et al.*, 1987), *A. tumefaciens* (Wardhan *et al.*, 1989), and *R. capsulatus* (*nifR1* locus)(Jones and Haselkorn, 1989) have been made available in the EMBL\Genbank\DDBJ databases under the accession numbers X105173, M19277\X02617, M14227, M15810, J03678, and X12359, respectively. The deduced aa sequences of the NtrC proteins from *K. pneumoniae*, *Bradyrhizobium* sp. [*Parasponia*], *R. meliloti*, and *R. capsulatus* (NifR1) are archived in the NBRF\SWISS\PIR protein databases under the accession numbers A03564, B26499, A26934, and P09432, respectively.

It has been proposed that NtrC proteins may be comprised of three domains as analysis of primary aa sequences has revealed conserved regions which bear resemblance to functional motifs in other proteins (Drummond *et al.*, 1986; Nixon *et al.*, 1986). The N-terminal region of approximately 100 aa is highly conserved amongst NtrC proteins and this is thought to form the domain which is involved in interaction with NtrB (section 1.2.2.1). This domain has been termed the receiver module (Kofoid and Parkinson, 1988) and identifies NtrC as a member of the superfamily of response regulators (RR) which are involved in stimulus-response coupling within a bacterial cell (section 1.5; Stock *et al.*, 1989b). The central domain of NtrC proteins has been termed the "activator" domain as this region shows

homology to the central region of proteins such as NifA and DctD which are also able to activate expression from σ54-dependent promoters. It is thought that this region interacts with σ54 and/or RNA polymerase. The C-terminal region of NtrC proteins contains a DNA-binding motif, and mutational studies have shown that this region is required for the binding of the *K. pneumoniae* NtrC to the UAS's of the *glnA* gene and transcriptional activation from *glnAp2* (Contreras and Drummond, 1988).

Further evidence for the functional roles of these domains within NtrC proteins has been obtained from studies of the molecular mechanisms whereby NtrC-PO_4 activates transcription from the *glnAp2* promoter in *E. coli* and *S. typhimurium.* The isomerization of closed to open promoter complexes at the *E. coli glnAp2* by NtrC-PO_4 has been shown *in vivo* (Sasse-Dwight and Gralla, 1988). *In vitro* studies with purified *S. typhimurium* NtrC have shown that this reaction requires ATP (apart from the ATP requirement for the phosphorylation of NtrC), however once an open complex had been formed the maintenance of the open complex or the initiation of transcription could be carried out by Eσ54 alone (Popham *et al.*, 1989).

Phosphorylation by NtrB at a site within the N-terminal domain of NtrC is essential for NtrC to perform its role as an activator in enteric bacteria (Ninfa and Magasanik, 1986). In contrast, the *K. pneumoniae* NifA protein (see section 1.3.2.3) is able to activate transcription at σ54-dependent *nif* promoters without phosphorylation of its N-terminal region. This is consistent with the idea that the N-terminal regions of these two proteins are non-homologous because they are "receiver" modules which are receptive to different signal elements (ie. NtrC is modified by NtrB; NifA activity is modulated by NifL). It has been suggested, therefore, that the attachment of the phosphate group to NtrC may not play an intrinsic role in the interaction with Eσ54, but causes a conformational change that allows the central domain of NtrC to interact with the closed promoter complex (Weglenski *et al.*, 1989). Mutational studies which support this hypothesis include the identification *E. coli* strains with *ntrC* gene mutations which exhibit increased NtrC activity, which may be the result of novel NtrC proteins which more readily assume the active form. An example is a mutation in the *E. coli ntrC* gene which resulted in an NtrC protein which could activate transcription to a low level in the

unphosphorylated form, while the mutant NtrC-PO_4 was more efficient than the wild-type NtrC-PO_4 (Weglenski *et al.*, 1989). Biological evidence for the hypothesis that the central region (conserved between NtrC and NifA proteins) is involved in the interaction between the activator and Eσ54 come from studies that showed that this region of the *R. meliloti* NifA was sufficient for transcriptional activation function (Huala and Ausubel, 1989).

Maximal activation by NtrC-PO_4 at the *glnAp2* promoter requires the presence of UAS's at which the activator protein binds, however these sites are not essential when NtrC-PO_4 is present at high concentrations (Ninfa *et al.*, 1987; Reitzer and Magasanik, 1986). NtrC binding sites display characteristics similar to eukaryotic enhancers as these sites are located at least 100 bp upstream of *glnAp2*, and that relocation of these sites by more than 1000 bp does not alter the ability of NtrC-PO_4 to activate transcription from *glnAp2* (Reitzer and Magasanik, 1986; Garcairrubio and Covarrubias, 1987). This suggests that this activation could involve DNA bending or loop formation so that the NtrC-PO_4 is able to interact with Eσ54 while still bound to these sites far from *glnAp2* (Reitzer and Magasanik, 1986; Reitzer *et al.*, 1989). A recent report has provided evidence to support the DNA loop model as opposed to a tracking or topoisomerase model for interaction between NtrC-PO_4 and Eσ54 (Wedel *et al.*, 1990). The NtrC-binding sites and the *glnA* promoter were placed on different rings of a singly linked catenane, and it was shown *in vitro* that this configuration gave similar levels of NtrC-mediated transcriptional activation as obtained when these DNA elements were *in cis* (Wedel *et al.*, 1990). NtrC binding sites are therefore thought to tether NtrC-PO_4 near the *glnA* promoter, thereby increasing the frequency of collisions between NtrC-PO_4 and Eσ54. Moreover, the DNA loop predicted to exist between the NtrC binding sites and the *glnA* promoter has been visualized directly by electron microscopy (Su *et al.*, 1990).

The positions of certain bases adjacent to the conserved -GC- doublet within the σ54-dependent promoter are considered to be important in distinguishing between promoters that are activated by NtrC-PO_4 as opposed to NifA (Ow *et al.*, 1983, 1985; Ray *et al.*, 1990). σ54-dependent promoter regions of relevant genes are shown in Fig. 5.1. Examination of σ54-dependent promoter sequences has indicated that in

most cases the presence of T residues at positions -17 to -15 signalled a promoter which could be activated by NtrC-PO_4 (Ow *et al.*, 1983; Fig. 5.1). In contrast, the *K. pneumoniae nifH* promoter, which has a stringent requirement not to be activated by NtrC-PO_4, has C residues at these positions (Beynon *et al.*, 1983; Fig. 5.1).

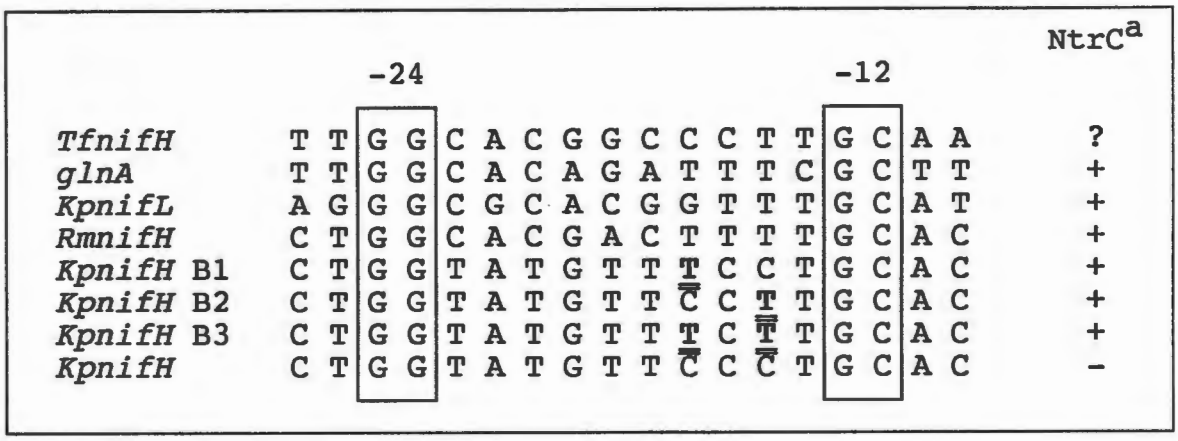


Fig. 5.1. σ54-dependent promoter regions of several nitrogen metabolism genes. The conserved -GG- and -GC- doublets at positions equivalent to -24 and -12 to the transcription start site of enteric *glnA* genes are blocked. NtrC[a] indicates whether these promoters are activated by NtrC-PO_4. Bold and double underlined letters indicate the C to T transitions constructed in the *K. pneumoniae nifH* promoter by Ow *et al.* (1985). Sequences are compiled from those identified by transcript mapping: *glnA*, sequences are identical for *E. coli* (Reitzer and Magasanik, 1985), *S. typhimurium* (Hirschman *et al.*, 1985), and *K. pneumoniae* (Dixon, 1984); *KpnifL*, *K. pneumoniae nifLA* promoter (Drummond *et al.*, 1983); *RmnifH*, *R. meliloti nifH* promoter (Sundaresan *et al.*, 1983a); *KpnifH*, *K. pneumoniae nifH* promoter (Sundaresan *et al.*, 1983a). The *T. ferrooxidans nifH* promoter (*TfnifH*) is putative.

These "rules" have been tested most extensively in *E. coli* backgrounds. Expression from the *R. meliloti nifH* promoter, which contains T residues at all three positions between -17 and -15 (Fig. 5.1), was shown to be derepressed in the presence of an *ntrC* gene (Sundaresan *et al.*, 1983b). Ow *et al.* (1985) used fusions between the *K. pneumoniae nifH* gene and *lacZ* to study the effect of point mutations in the *K. pneumoniae nifH* promoter region. Three separate mutant promoters, which were the result of C to T transitions at positions -17 to -15 (Fig. 5.1), showed activation by NtrC-PO_4 as measured by increases in β-galactosidase activity in *E. coli* (Ow *et al.*, 1985).

Moreover, Ray *et al.* (1990) have shown by transcript mapping both *in vivo* and *in vitro* that the *K. pneumoniae nifH* promoter mutant B3 (Fig. 5.1) required NtrC-PO_4 at high intracellular concentration for activation, while NtrC-PO_4 was unable to activate transcription at the *K. pneumoniae nifH* wildtype promoter. Provision of NtrC-binding sites upstream of a test promoter effectively increases the local concentration of NtrC-PO_4, and experiments where this was carried out resulted in levels of transcription from the *nifH*-B3 promoter equivalent to those obtained from the *glnAp2* promoter; in contrast, the *K. pneumoniae* wild-type promoter with NtrC-binding sites upstream was weakly activated. Further evidence for the importance of the σ54-dependent promoter (Eσ54 binding site) in transcriptional activation was demonstrated by the inability of NtrC-PO_4 to direct transcription from the σ70-dependent *lac* promoter provided with high affinity NtrC-binding sites (Ray *et al.*, 1990).

Morett and Buck (1989) have demonstrated, using *in vivo* footprinting, that Eσ54 forms a NifA-independent closed promoter complex at two σ54-dependent promoters which have T's between positions -17 and -15 (the *R. meliloti nifH*, and a mutant *K. pneumoniae nifH* promoter identical to B3 but with a further C to T transition at -16), but not at the *K. pneumoniae nifH* wildtype promoter. Taking the data of Ow *et al.* (1985), Ray *et al.* (1990), and Morett and Buck (1989) one may infer that NifA-independent NtrC-mediated activation occurs at promoters at which Eσ54 more readily forms a closed promoter complex.

It is interesting to note that the *B. japonicum nifH* and *nifDK* promoters were not activated in *E. coli* by the *K. pneumoniae* NtrC-PO_4, but were activated by NifA (Alvarez-Morales and Hennecke, 1985). Both of these promoters contain G (purine) as well as T residues between positions -17 and -15, which may have resulted in a weak interaction with *E. coli* Eσ54 which was insufficient for NtrC-PO_4 from solution to catalyze the isomerization to open complexes. In contrast, the binding of NifA to the NifA-specific UAS's of these promoters (Alvarez-Morales and Hennecke, 1986) may have increased the local concentration of NifA sufficiently to enable transcriptional activation.

The putative *T. ferrooxidans nifH* promoter has a T residue at the -15 position (Fig. 5.1), and this sequence is the same between positions -17 and -12 as the *K. pneumoniae nifH*-B2 mutant promoter, which showed some activation by NtrC-PO_4 (Ow *et al.*, 1985). The surrounding nucleotides are also likely to play a role, however it seemed possible that the *T. ferrooxidans nifH* promoter may be activated by NtrC-PO_4 in an *E. coli* background. The availability of a translational fusion between the *T. ferrooxidans nifH* gene and *lacZ* (pHlac, section 4.2.7) made this idea relatively easy to test. Experiments, aimed at the cloning of the *T. ferrooxidans ntrC* gene, which were based on NtrC-mediated activation of this fusion, are described in this chapter.

5.2 Materials and Methods

5.2.1 Bacterial strains, plasmids, and media

The strains and plasmids used in this study are described in Appendix A. *E. coli* cells were grown with LB medium (Appendix B) or with glucose minimal medium containing glucose at a concentration of 0,2% (w/v) (GMM - Appendix B) or 2% (w/v) (GGMM) with additions as described in the text. When required, antibiotics were added at the following concentrations: ampicillin, 100 µg/ml; tetracycline, 15 µg/ml; chloramphenicol, 20 µg/ml.

5.2.2 Preparation of DNA

The alkaline lysis method (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981) was used for both small- and large-scale plasmid preparations (Appendix C). Preparation of chromosomal DNA from *T. ferrooxidans* ATCC 33020 and the source of the *T. ferrooxidans* ATCC 33020 cosmid gene bank have been discussed in section 2.2.2.

5.2.3 DNA manipulations

Standard methods (Maniatis *et al.*, 1982) were used for restriction digests, gel electrophoresis, purification of DNA fragments from agarose gels, ligations, and the filling in of 5' sticky ends (Appendix C).

5.2.4 Plasmid constructions

The construction of the *T. ferrooxidans nifH-lacZ* fusion plasmid, pHlac, has been described in section 4.2.7 (Fig. 4.2). The vector, pMC1403 (Casadaban *et al.*, 1983), used to construct pHlac, and the vector, pHC79 (Hohn and Collins, 1980), used to construct the *T. ferrooxidans* cosmid bank, both contain replication functions and the ampicillin antibiotic marker derived from pBR322. It was therefore necessary to subclone the *T. ferrooxidans nifH-lacZ* fusion into a plasmid vector which was compatible with the pHC79-based cosmids. This was achieved by insertion of a 8,4-kbp *Eco*RI-*Sal*I fragment (containing the *T. ferrooxidans nifH-lacZ* fusion and the *lacYA* genes; Fig 4.2) from pHlac into the vector pACYC184 (Appendix F) digested with *Bam*HI and *Sal*I. After ligation of the *Sal*I ends of each fragment, the *Eco*RI

sticky end upstream of the *T. ferrooxidans nifH* promoter region was filled-in with Klenow and ligated to the *Bam*HI end (which was also filled-in with Klenow) within the tetracycline resistance gene of pACYC184 to produce pHlac10. pHlac10 carried the chloramphenicol resistance marker and was compatible with ColE1-based plasmids.

pHlac10 contained the promoter of the tetracycline resistance gene (*tet* promoter) from pACYC184 located upstream of the *T. ferrooxidans nifH* regulatory region and therefore transcripts could potentially originate from this -35, -10 region. Consequently, this region was deleted to check whether any spurious clones were detected as a result of increased β-galactosidase activity due to activation of expression from this vector promoter region. This deletion was achieved by digestion of pHlac10 with *Xba*I (unique site upstream of the *tet* promoter) and with *Eco*RV (unique site downstream of the *tet* promoter). Klenow enzyme was added to fill-in the *Xba*I 5' overhang, followed by a recircularization ligation to produce pHlac20. Restriction enzyme analysis of pHlac20 revealed that the *tet* promoter had been removed (loss of a unique *Hind*III site within the *tet* promoter) but that a larger deletion than expected had occurred. A *Ssp*I site immediately 5' to the *Xba*I site was intact indicating that the deletion extended in the other direction, ie. through the *Eco*RV site, removing approximately 200 bp of pACYC184 vector sequences and approximately 300 bp of *T. ferrooxidans* DNA which was at least 400 bp upstream of the *T. ferrooxidans nifH* regulatory region and therefore not thought to be necessary for expression from the *T. ferrooxidans nifH* promoter.

pK70 was constructed by insertion of a *Hind*III fragment from pFB514 which contains the *K. pneumoniae glnAntrBC* genes (de Bruijn and Ausubel, 1981) into the *Hind*III site of pHC79. The orientation of the insert placed the 5' end of the *K. pneumoniae glnA* gene closest to the *Eco*RV site within the vector.

5.2.5 Screening for the *T. ferrooxidans ntrC* gene by complementation of an *E. coli ntrC* mutant for Ntr function

The *T. ferrooxidans* ATCC 33020 pHC79 cosmid library was transduced into the *E. coli ntrC* mutant ET8556 as described in Appendix C. The expression mixes were

washed in 0,8% NaCl before plating on GMM containing 0,2% arginine with antibiotic selection. As a positive control, the plasmid p804 carrying the *E. coli ntrBC* genes was transformed into *E. coli* ET8556 and plated onto the test medium. A proportion of each washed transduction expression mix was plated onto LBA + Ap plates to determine the number of transductants. Growth was scored after incubation for 24 - 48 h at 30°C.

5.2.6 β-Galactosidase assays

β-Galactosidase activity of the *E. coli* cultures was determined as described in section 4.2.4.

5.2.6.1 Growth conditions for β-galactosidase assays to test the expression from the *T. ferrooxidans nifH* promoter in an *E. coli ntrC* mutant strain. *E. coli* cells containing the *nifH-lacZ* fusion plasmid, pHlac, were grown in a preculture of GMM containing 15 mM glutamine and 0,2% $(NH_4)_2SO_4$ (w/v) (nitrogen rich medium) for 20 h at 37°C under aerobic conditions. This preculture was washed in 0,8% saline and the optical density at 600 nm (OD_{600}) was measured to enable inoculation of a sample of the washed pre-culture into the assay culture media to ensure a starting OD_{600} of 0,05 units which was constant for each *E. coli* strain. The assay culture medium for *E. coli* (pHlac) strains was as described by Tuli *et al.* (1982): GMM supplemented with 0,15 mM glutamine and 15 mM glutamate (nitrogen poor medium). Appropriate antibiotic selection was provided at all times. These cultures were grown under aerobic conditions at 37°C, and β-galactosidase assays were carried out after 10 h and 20 h growth. No substantial differences in levels of β-galactosidase activity were obtained at these two points of the growth curve for each culture.

5.2.6.2 Growth conditions for β-galactosidase assays to measure expression from the *T. ferrooxidans nifH* promoter in the presence of the cloned *T. ferrooxidans ntrC* gene. Growth conditions were altered from those described in section 5.2.6.1 to match the rich and poor nitrogen sources provided in the solid media used to screen for the *T. ferrooxidans ntrC* gene. *E. coli* cells were grown in pre-culture as described in section 5.2.6.1, washed, and inoculated at a starting OD_{600} of 0,2 units into the assay culture media. The nitrogen rich medium consisted of GMM supplemented

with 0,2% glutamine (w/v), 0,2% $(NH_4)_2SO_4$ (w/v), 0.2% casamino acids (w/v), and 0,2% yeast extract (w/v); the nitrogen poor medium consisted of GMM supplemented with 0,02% glutamine (w/v), 0,02% casamino acids (w/v), and 0,02% yeast extract (w/v). Appropriate antibiotic selection was provided at all times. The cultures were grown in these media for 10 h before determination of the β-galactosidase activity.

5.2.7 Methods used for the identification of the *T. ferrooxidans ntrC* gene

5.2.7.1 Development of a screening technique for NtrC-mediated activation of the *T. ferrooxidans nifH-lacZ* fusion in *E. coli*. An agar plate medium was developed whereby it was possible to distinguish between single colonies of *E. coli* (pHlac10) cells which exhibited a low level of expression of the *T. ferrooxidans nifH-lacZ* fusion (ie. in the absence of an *ntrC* gene) and those that produced a high level of expression of the *T. ferrooxidans nifH-lacZ* fusion (ie. in the presence of an *ntrC* gene). This medium had to fulfil two main criteria: (i) it had to provide enough nitrogen for growth, but at a low enough level to ensure that the NtrC protein was predominantly in its activator form; and (ii) it had to provide a visual means whereby it was possible to detect those colonies of *E. coli* cells which displayed an increase in β-galactosidase activity above the basal levels produced in the absence of an *ntrC* gene. *E. coli ntrC* mutant ET8556 (pHlac10) cells were used as a control for cells with low β-galactosidase activity, while *E. coli ntrC*$^+$ YMC10 (pHlac10) and *E. coli* ET8556 (pHlac10, pK70) cells were used as controls which had increased β-galactosidase activity for the testing of a variety of media.

Several media were tried before a successful medium was developed. Mackonkey/lactose medium (Difco) contained too high a concentration of undefined nitrogen sources to be workable, and the phosphate buffer within lactose minimal medium containing pH indicator dyes such as bromophenol purple or neutral red had too great a buffering capacity for the detection of differences in organic acid production. The inclusion of the chromogenic dye 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) into GMM proved to be too sensitive for detecting differences in β-galactosidase activities. The optimal method, which was chosen for subsequent experiments, was growth of the *E. coli*

colonies on GGMM supplemented with 0,02% casamino acids (w/v), and 0,02% yeast extract (w/v) (GGMMLN medium) containing 16 g/l agar (GGMMLN16 plates) for 48 h at 37°C after which each plate was flooded with 2 ml of ONPG (4 mg/ml; w/v). The high concentration of glucose was required for the growth of *E. coli* strains into visible colonies from a single cell. *E. coli* YMC10 (pHlac10) or ET8556 (pHlac10, pK70) cells (showing high levels of β-galactosidase activity) were detected as colonies which turned a bright yellow colour within 5 - 15 min. *E. coli* ET8556 (pHlac10) cells were visible as dull brown colonies after flooding with ONPG. Increase of the agar concentration from the usual 15 g/l to 16 g/l ensured that the colonies were more tightly attached to the medium surface and did not float off during flooding with ONPG. It was also found that incubation of the plates at 4°C for a few hours prior to ONPG flooding decreased the likelihood of colonies becoming dislodged. Differences in yellow staining intensity of colonies were discerned more clearly if plates were viewed on a light box after ONPG flooding (see Fig. 5.2 for an illustration of the technique).

5.2.7.2 Screening for the *T. ferrooxidans ntrC* gene by testing for NtrC-mediated activation of the *T. ferrooxidans nifH-lacZ* fusion in *E. coli*. The *T. ferrooxidans* ATCC 33020 pHC79 cosmid library was transduced into the *E. coli ntrC* mutant ET8556 cells containing the *T. ferrooxidans nifH-lacZ* fusion on the plasmid pHlac10. Transductants were plated onto GGMMLN16 plates with antibiotic selection and incubated aerobically for 48 h at 37°C. Colonies which exhibited an increased level of β-galactosidase activity were selected as those that formed yellow colonies 5 - 15 min after flooding with ONPG as described in section 5.2.7.1. Positive colonies were picked off with a sterile toothpick and purified by repeated streaking onto fresh LBA + Ap + Cm plates.

Each putative positive clone was tested further by toothpicking several individual colonies onto GMM containing 16 g/l agar, supplemented with either (i) 0,02% casamino acids (w/v) and 0,02% yeast extract (w/v) (GMMLN16 plates); or (ii) 0,2% $(NH_4)_2SO_4$ (w/v), 0,2% glutamine (w/v), 0,2% casamino acids (w/v), and 0,2% yeast extract (w/v)(GMMHN16 plates). Since several thousand cells are transferred at the point of a toothpick, a glucose concentration of 0,2% (w/v) was

sufficient for *E. coli* strains to form visible colonies after toothpicking onto these media. GMMLN16 plates were incubated for 48 h at 37°C before ONPG flooding, while GMMHN16 plates were incubated for 10 h at 37°C before ONPG flooding. An additional control *E. coli* strain used on the GMMHN16 plates was the *lac*$^+$ strain CSH36 (pHC79, pACYC184), which produced β-galactosidase activity which was not repressed by high levels of nitrogen. Cosmid clone DNA was purified from pHlac10 plasmid DNA by transformation into *E. coli* ET8556 cells and selection on LBA + Ap plates. Plasmid DNA was prepared from these transformants, and restriction enzyme analysis enabled choice of strains that had lost the plasmid pHlac10.

5.2.8 Subcloning of the cosmid pT101

Samples of pT101 were digested with either of the restriction enzymes *Hind*III or *Bgl*II and these were ligated to samples of the vector pEcoR252 digested with *Hind*III or *Bgl*II, respectively. The ligation mixes were transformed into ET8556 (pHlac10) cells, and transformants were screened for β-galactosidase activity above basal levels as described in section 5.2.7.1. The smallest active subclone from these experiments, pT110, was chosen for further studies (Fig. 5.3). A 5,2-kbp *Hind*III-*Kpn*I fragment from pT110 was subcloned into the *Hind*III, *Kpn*I sites of the vector Bluescript SK to produce pT120. pT125 was constructed by deletion of a 2-kbp *Bam*HI fragment from pT120 (from the insert *Bam*HI site to the *Bam*HI within the vector multiple-cloning-site). A 2,4-kbp *Hind*III-*Cla*I fragment from pT120 was cloned into the *Hind*III, *Cla*I sites of Bluescript SK to produce pT127.

5.2.9 Nucleotide sequencing

Nucleotide sequencing of the subclones pT125 and pT127 was carried out as described in section 3.2.8 and Appendix C. Nucleotide sequence data was analyzed with a VAX computer using the UWGCG sequence analysis package version 6.1 (Devereux, 1984) and associated databases.

5.3 Results

5.3.1 Attempt to clone the *T. ferrooxidans ntrC* by complementation of an *E. coli ntrC* mutant for Ntr function

The *T. ferrooxidans* cosmid bank was transduced into the *E. coli ntrC* mutant ET8556 and transductants were plated onto GMM containing arginine as a sole nitrogen source. Plating of a proportion of the transductants on LBA + Ap demonstrated that approximately 2500 transductants were screened on the minimal medium. ET8556 (p804) cells grew as large colonies on the GMM (+ arginine). No positive clones were identified out of the ET8556 cells transduced with the *T. ferrooxidans* cosmid bank.

5.3.2 Expression of the *T. ferrooxidans nifH-lacZ* fusion in the presence and absence of the *E. coli* and *K. pneumoniae ntrC* genes

Experiments were carried out to test whether there was an increase in expression from the *T. ferrooxidans nifH* promoter in the presence of a cloned *ntrC* gene in *E. coli*. *E. coli* ET8556 *ntrC* mutant cells containing the *T. ferrooxidans nifH-lacZ* fusion plasmid, pHlac, produced low basal levels of β-galactosidase activity when grown in nitrogen-limiting conditions (Table 5.1). In contrast, *E. coli* YMC10 *ntrC*$^+$ (pHlac) cells displayed 150-fold greater levels of β-galactosidase activity (Table 5.1). The presence of the *K. pneumoniae ntrC* gene was able to partially suppress the mutation in the *E. coli ntrC* gene: ET8556 (pHlac, pFB514) cells produced levels of β-galactosidase activity 60-fold greater than obtained from ET8556 (pHlac) cells (Table 5.1). These results formed the rationale behind the strategy to isolate the *T. ferrooxidans ntrC* gene by activation of expression of the *T. ferrooxidans nifH-lacZ* fusion.

Table 5.1. Expression of the *T. ferrooxidans nifH-lacZ* fusion in *E. coli* in the presence and absence of *ntrC* genes.

Strain[a]	Relevant genotype[b]	β-galactosidase activity (Miller units)
ET8556(pHlac)	*ntrC* *TfnifH-lacZ*	214
YMC10(pHlac)	*EcntrC*$^+$ *TfnifH-lacZ*	32 024
ET8556(pHlac) (pFB514)	*KpntrC*$^+$ *TfnifH-lacZ*	13 326

[a] Strains were grown aerobically in nitrogen-limiting conditions as described in the text (section 5.2.6.1).
[b] *Tf, T. ferrooxidans*; *Ec, E. coli*; *Kp, K. pneumoniae*.

5.3.3 Cloning of the *T. ferrooxidans ntrC* gene by activation of expression of a *T. ferrooxidans nifH-lacZ* fusion in *E. coli*

The *T. ferrooxidans* cosmid bank was transduced into the *E. coli ntrC* mutant ET8556 containing the *T. ferrooxidans nifH-lacZ* fusion plasmid pHlac10. Transductants (1800) were screened for an increase in β-galactosidase activity by the ONPG flooding technique after growth on GGMMLN16 plates. Six positive clones were identified and purified. Cosmid DNA was extracted from cultures of each clone, purified from pHlac10 DNA, and transformed into *E. coli* ET8556 (pHlac10) cells. These transformants were retested for increased levels of β-galactosidase activity by toothpicking onto GMMLN16 plates. It was shown that this activity was repressed when the transformants were toothpicked onto GMMHN16 plates, which contained a high concentration of nitrogen. Restriction enzyme fragments common to the purified cosmid clones were observed. One clone, pT101, was chosen for further analysis. The difference in yellow staining intensity after growth on GMMLN16 plates and ONPG flooding between ET8556(pT101, pHlac10) and ET8556(pHC79, pHlac10) cells can be seen on Fig. 5.2.

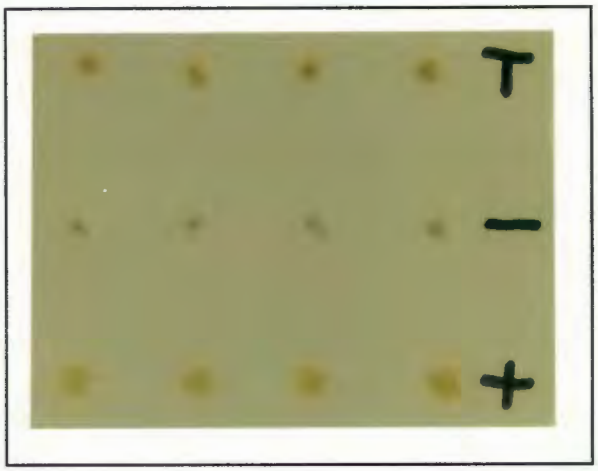


Fig. 5.2. Illustration of the screening technique for the *T. ferrooxidans ntrC* gene. Cells were toothpicked onto a GMMLN16 plate and grown for 48 h at 37°C before flooding with ONPG. Strains: (T), ET8556(pT101, pHlac10)
(-), ET8556(pHC79, pHlac10)
(+), YMC10(pHC79, pHlac10).

5.3.4 Subcloning of the cosmid clone pT101 to localize the *T. ferrooxidans ntrC* gene

Subcloning of *Bgl*II digestion fragments from pT101 into the positive selection vector pEcoR252 yielded no clones which produced an increase in expression from the *T. ferrooxidans nifH-lacZ* fusion, however some positive clones were identified from the *Hind*III subcloning (section 5.2.8). All of these subclones contained a common 13,5-kbp *Hind*III fragment, and one subclone, pT110, contained only this fragment and the vector pEcoR252 3,4-kbp *Hind*III fragment. pT110 was used to localize the *T. ferrooxidans ntrC* gene. A restriction map of pT110 was constructed (Fig. 5.3). A 5,2-kbp *Hind*III-*Kpn*I fragment from pT110 was cloned into Bluescript SK to produce pT120 (Fig. 5.3), and this construct was shown to contain the coding sequences necessary for increased β-galactosidase activity from the *T. ferrooxidans nifH-lacZ* fusion. ET8556(pHlac10) cells containing either pT125 or pT127 (Fig. 5.3) tested negative after growth on GMMLN16 plates and ONPG flooding. pT125 and pT127 were therefore used to obtain nucleotide sequence data for the *T. ferrooxidans ntrC* gene.

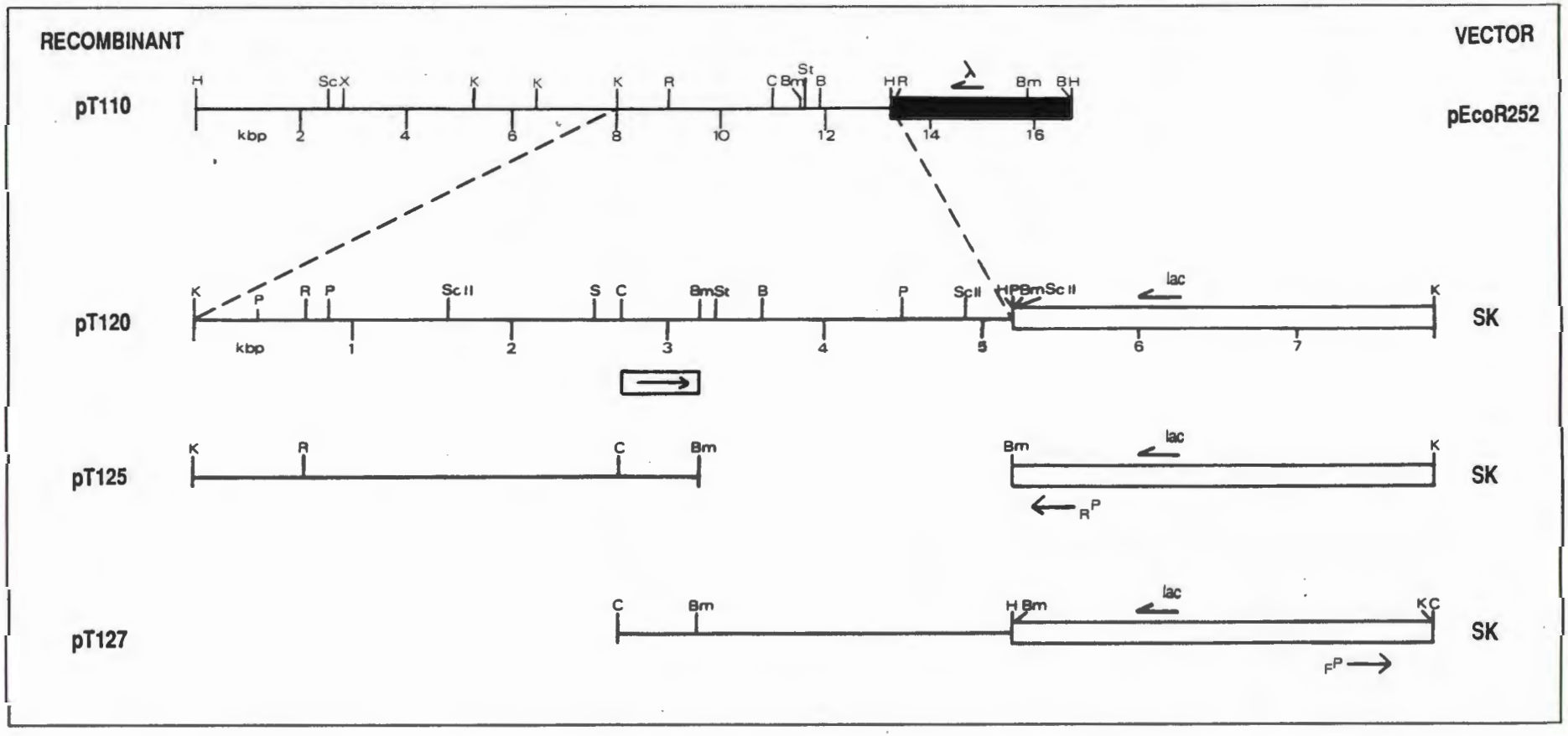


Fig. 5.3. Restriction map of pT110 and subclones. The full extent of both vector and insert sequences are shown for each subclone. *T. ferrooxidans* DNA is indicated by single lines, while vector DNA is indicated by boxes. The name given to each recombinant subclone is indicated on the left hand side, while the vector name is given on the right. The scale used for pT110 has been enlarged 1,5 x for pT120, pT125, and pT127. The 486 bp *Cla*I-*Bam*HI fragment within the *T. ferrooxidans ntrC* gene is shown by an open box; the arrow within indicates the orientation of the gene. SK, Bluescript SK vector; λ, λP_R promoter; lac, *lac* promoter; R^P, reverse primer; F^P, forward primer; B, *Bgl*II; Bm, *Bam*HI; C, *Cla*I; H, *Hind*III; K, *Kpn*I; P, *Pst*I; R, *Eco*RI; Sc, *Sac*I; ScII, *Sac*II; St, *Stu*I; X, *Xho*I.

5.3.5 Nucleotide sequence of a section of the *T. ferrooxidans ntrC* gene.

The nucleotide sequence of the 486 bp *Cla*I-*Bam*HI *T. ferrooxidans* DNA fragment (Fig. 5.3) is shown in Fig. 5.4.

```
ATCGATGACGACCCTTCCATTCGCTGGGTTTTGGAGAAGGCGTTGACTCAGGCCGATATCCTGGTG   66
I  D  D  D  P  S  I  R  W  V  L  E  K  A  L  T  Q  A  D  I  L  V

CGCACGTTTGCAGACGCCGACAGCGCGCTGCGGGCCTTGGGGCGTGACAGGCCAGGGGCCGTTGTC  132
 R  T  F  A  D  A  D  S  A  L  R  A  L  G  R  D  R  P  G  A  V  V

ACCGATCTGCGTATGCCCGGCCTGGACGGATTGGCGTTTCTGCGCGAAGTTCAGTCGCGCTGGCCA  198
T  D  L  R  M  P  G  L  D  G  L  A  F  L  R  E  V  Q  S  R  W  P

AAACTGCCGGTGATCGTGATGACGGCGCACTCTGATCTCGACAATGCCGTCGCCGCTTTTCAAAGT  264
K  L  P  V  I  V  M  T  A  H  S  D  L  D  N  A  V  A  A  F  Q  S

GGGGCCTTTGAGTATCTGCCCAAGCCGTTTGACATGGACGAGGCGGTTACGCTGGTGCAGCGCGCC  330
G  A  F  E  Y  L  P  K  P  F  D  M  D  E  A  V  T  L  V  Q  R  A

CTCGGCAGCCGGGCTGGACCGGGCGCGGCGGGTGTGGAGGAGAGCGATCCGGCGGAAATGATCGGT  396
L  G  S  R  A  G  P  G  A  A  G  V  E  E  S  D  P  A  E  M  I  G

GAGGCTCCGGCCATGCAAGAGGTGTTTCGTGCCATCGGGCGTCTGTCGCGCTCACAGATCAATGTG  462
E  A  P  A  M  Q  E  V  F  R  A  I  G  R  L  S  R  S  Q  I  N  V

CTGATTACCGGCGAATCGGGATCC  486
L  I  T  G  E  S  G  S
```

Fig. 5.4. Nucleotide sequence of the 486 bp *Cla*I-*Bam*HI DNA fragment which forms part of the *T. ferrooxidans ntrC* gene, with the predicted amino acid sequence of the corresponding region of the *T. ferrooxidans* NtrC protein. Nucleotides are numbered from the first base of the *Cla*I site.

These data were obtained from nucleotide sequencing of (i) pT125 with the reverse primer across the *Bam*HI site into the *T. ferrooxidans* insert; and (ii) pT127 with the forward primer across the *Cla*I site into the *T. ferrooxidans* insert (Fig. 5.3). Some sequence overlap was obtained in the middle, however these data represent sequence obtained from a single strand only. Consequently, a comprehensive sequencing project is necessary to obtain the sequence data with 100% accuracy. Nevertheless, the data obtained confirmed that the strategy for cloning the *T. ferrooxidans ntrC* gene had been successful. A search of the Genbank (Release 62) database using the UWGCG FASTA program revealed that this *T. ferrooxidans* DNA fragment showed sequence identity to a region situated within *ntrC* genes from several bacteria, including *K. pneumoniae*, *E. coli*, *A. tumefaciens*, and *R. meliloti*.

DNA-DNA alignments using the UWGCG GAP program showed that this 486 bp *Cla*I-*Bam*HI *T. ferrooxidans* DNA fragment had 62, 60, 54, 51, 51, and 51% identity with the 5' end region of *ntrC* genes from *E. coli*, *K. pneumoniae*, *Bradyrhizobium* sp. [*Parasponia*], *A. tumefaciens*, *R. meliloti*, and *R. capsulatus* (*nifR1*), respectively. These data indicated that this fragment was situated within the *T. ferrooxidans ntrC* gene. A reading frame (RF) without any stop codons was identified in frame 1, and the predicted translation product of this RF is shown in Fig. 5.4.

5.3.6 Amino acid sequence alignment studies of the predicted product of a section of the *T. ferrooxidans ntrC* gene and NtrC proteins from other bacteria

The deduced aa sequence of the *T. ferrooxidans ntrC* RF was aligned with the *K. pneumoniae* (Buikema *et al.* 1985) and *E. coli* (Miranda-Rios *et al.*, 1987) *ntrC* gene products using the UWGCG GAP program (Fig. 5.5).

```
Tf                    <I D D D P S I R W V L E K A L T Q A D I L V R T F A D   27
Kp  M Q R G I A W I V D D D S S I R W V L E R A L T G A G L S C T T F E S   35
Ec  M Q R G I V W V V D D D S S I R W V L E R A L A G A G L T C T T F E N   35

Tf  A D S A L R A L G R D R P G A V V T D L R M P G L D G L A F L R E V Q   63
Kp  G N E V L D A L T T K T P D V L L S D I R M P G M D G L A L L K Q I K   70
Ec  G A E V L E A L A S K T P D V L L S D I R M P G M D G L A L L K Q I K   70

Tf  S R W P K L P V I V M T A H S D L D N A V A A F Q S G A F E Y L P K P   98
Kp  Q R H P M L P V I I M T A H S D L D A A V S A Y Q Q G A F D Y L P K P  105
Ec  Q R H P M L P V I I M T A H S D L D A A V S A Y Q Q G A F D Y L P K P  105

Tf  F D M D E A V T L V Q R A L G S - - R A G P G A A G V E E S D P A E M  130
Kp  F D I D E A V A L V D R A I S H Y Q E Q Q Q P R N A P I N S P T A D I  140
Ec  F D I D E A V A L V E R A I S H Y Q E Q Q Q P R N V Q L N G P T T D I  140

Tf  I G E A P A M Q E V F R A I G R L S R S Q I N V L I T G E S G S >      162
Kp  I G E A P A M Q D V F R I I G R L S R S S I S V L I N G E S G T >      173
Ec  I A K - P A M Q D V F R I I G R L S R S S I S V L I N G E S G T >      172
```

Fig. 5.5. Alignment of the deduced amino acid sequence of a region of the *T. ferrooxidans* NtrC with the N-terminal region of NtrC proteins from *K. pneumoniae* and *E. coli*. Amino acids are identified by the single letter code, and conserved regions (see Appendix E for scheme of conservative substitutions) are boxed. Numerals refer to aa positions in each protein. Tf, *T. ferrooxidans* NtrC; Kp, *K. pneumoniae* NtrC; Ec, *E. coli* NtrC.

There was extensive aa similarity throughout the entire *T. ferrooxidans ntrC* RF predicted product and the region towards the N-terminus of both the *K. pneumoniae* and *E. coli* NtrC proteins. These results provided evidence that the *T. ferrooxidans ntrC* RF encoded 162 aa of the N-terminal region of the *T. ferrooxidans* NtrC protein (starting about 10 aa from the N-terminus of the primary aa sequence, assuming that the *T. ferrooxidans* NtrC is similar to enteric NtrC proteins at the N-terminal end). Amino acid alignments using the UWGCG GAP program, showed that this region of the *T. ferrooxidans* NtrC protein had 76, 72, 65, 64, 62, and 61% similarity with the N-terminal 200 aa of NtrC proteins from *K. pneumoniae*, *E. coli*, *Bradyrhizobium* sp. [*Parasponia*], *A. tumefaciens*, *R. meliloti*, and the NifR1 protein from *R. capsulatus*, respectively.

5.3.7 Expression of the *T. ferrooxidans nifH-lacZ fusion* in the presence of the cloned *T. ferrooxidans ntrC* gene

β-galactosidase assays of *E. coli* cells grown in liquid culture were carried out to provide quantitative measurements of the differences in yellow staining intensity of cells previously obtained by the ONPG flooding technique. Cells were assayed after growth in nitrogen-limiting and nitrogen-rich media to investigate whether expression of the *T. ferrooxidans nifH-lacZ* fusion was affected by levels of available nitrogen, a characteristic of genes that are regulated by the Ntr system in enteric bacteria.

Low basal levels of β-galactosidase activity were obtained from ET8556(pHlac20, pHC79) cells grown with high or low levels of nitrogen (Row 1; Table 5.2). ET8556(pHlac20, pT101) cells, which contained the *T. ferrooxidans ntrC* gene on the cosmid insert in the vector pHC79, produced levels of β-galactosidase activity which were 3-fold greater than the basal levels obtained from ET8556(pHlac20, pHC79) cells when both strains were grown under nitrogen-limiting conditions (Rows 1 and 2; Table 5.2). ET8556(pHlac20, pT101) cells showed a 2-fold repression of β-galactosidase activity when grown in the nitrogen-rich medium (Row 2; Table 5.2). ET8556(pHlac20, pT120) and ET8556(pHlac10, pT120) cells, which contained the cloned *T. ferrooxidans ntrC* gene on the vector Bluescript SK, produced levels of β-galactosidase activity 3- to 5-fold greater than the basal levels obtained from ET8556(pHlac20, Bluescript SK) cells when grown in nitrogen-limiting medium

(compare rows 3 and 4 with row 5; Table 5.2). However, these high levels of expression were repressed when the cells were grown in nitrogen-rich medium (Rows 3 and 4; Table 5.2). These results also showed that removal of the vector *tet* promoter from pHlac10 to produce pHlac20 did not affect the pattern of expression of the *T. ferrooxidans nifH-lacZ* fusion in the presence of the *T. ferrooxidans ntrC* gene. *NtrC*+ strains YMC10(pHlac20, pHC79) and ET8556(pHlac20, pK70) showed similar patterns of β-galactosidase activity in response to nitrogen sources in the media (Rows 6 and 7; Table 5.2). *E. coli* CSH36(pACYC184, pHC79) cells, which contained a constitutively expressed chromosomal *lacZ* gene, displayed high levels of β-galactosidase activity in both nitrogen-rich and -poor media (Row 8; Table 5.2).

Table 5.2. Expression of the *T. ferrooxidans nifH-lacZ* fusion in the presence of the cloned *T. ferrooxidans ntrC* gene in *E. coli*.

	Strain	Relevant genotype	β-galactosidase activity (Miller units)	
			LN	HN
1	ET8556(pHlac20, pHC79)	*ntrC* *TfnifH-lacZ*	312	266
2	ET8556(pHlac20, pT101)	*TfntrC*+ *TfnifH-lacZ*	1 048	553
3	ET8556(pHlac20, pT120)	*TfntrC*+ *TfnifH-lacZ*	2 035	483
4	ET8556(pHlac10, pT120)	*TfntrC*+ *TfnifH-lacZ*	3 337	224
5	ET8556(pHlac20, B-SK)	*ntrC* *TfnifH-lacZ*	654	413
6	YMC10(pHlac20, pHC79)	*EcntrC*+ *TfnifH-lacZ*	7 634	464
7	ET8556(pHlac20, pK70)	*KpntrC*+ *TfnifH-lacZ*	2 182	452
8	CSH36(pACYC184, pHC79)	*EclacZ*+	2 620	14 568

[a] Strains were grown in GMM supplemented with either a low (LN) or high (HN) concentration of nitrogen sources as described in section 5.2.6.2.

[b] *Tf, T. ferrooxidans*; *Ec, E. coli*; *Kp, K. pneumoniae*; B-SK, Bluescript SK.

5.3.8 Southern hybridization of pT120 against *T. ferrooxidans* chromosomal DNA

*Pst*I fragments (3,6- and 0,5-kbp), and a 3,2-kbp *Sac*II fragment - all of which are internal to the insert on pT120 - hybridized to *T. ferrooxidans* chromosomal DNA restriction digest fragments of the same size (results not shown). These results correlated with the restriction map of pT120 shown on Fig. 5.3, and confirmed that the fragment containing the *T. ferrooxidans ntrC* gene was cloned without rearrangement.

5.4 Discussion

The *K. pneumoniae nifH* promoter is not activated by NtrC-PO_4 in *K. pneumoniae* (Merrick, 1983). This is consistent with the two-tiered model of regulation of the *nif* genes in *K. pneumoniae*, in which NtrC is responsible for regulation at the first level only, ie. expression from *pnifLA* (section 1.3.2.3). The unresponsiveness of the *nifH* promoter (as well as other *nif* promoters) to NtrC-PO_4 prevents expression of this system under inappropriate circumstances, such as in aerobic nitrogen-limiting conditions. Although the *K. pneumoniae nifH* promoter is also not activated by NtrC-PO_4 (even at high concentration) in an *E. coli* background, the expression of *nifH* genes from other bacteria cloned in *E. coli* may not necessarily mimic that in the original host organism.

The *T. ferrooxidans nifH* regulatory region contains two canonical NifA UAS's, which indicates that a NifA-like protein is likely to function as the activator of this gene in *T. ferrooxidans* (Pretorius *et al.*, 1987). However, a *T. ferrooxidans nifH-lacZ* fusion was activated in the presence of a *ntrC* gene in *E. coli*. This phenomenon formed the basis of a method which was used successfully to isolate the *T. ferrooxidans ntrC* gene. Evidence that the *T. ferrooxidans ntrC* gene had been cloned was obtained from nucleotide sequence data, and from β-galactosidase assays which showed that expression from the *T. ferrooxidans nifH-lacZ* fusion was regulated in response to nitrogen supply in the presence of the *T. ferrooxidans ntrC* gene. Activation of the *T. ferrooxidans nifH* promoter by NtrC in *E. coli* may have been the result of (i) the high concentration of NtrC expressed from an *ntrC* gene carried on a high copy-number plasmid, and (ii) the positions of certain bases within the Eσ54-binding site of the *T. ferrooxidans nifH* promoter which enabled activation by NtrC. Alternatively, sequences which are similar to NtrC-binding sites in the region upstream of the *T. ferrooxidans nifH* promoter may have resulted in some increase in the local concentration of NtrC. Examination of the nucleotide sequence data upstream of the *T. ferrooxidans nifH* putative promoter (300 bp; Pretorius *et al.*, 1987; unpublished data) revealed sequences which showed identity to the consensus half-sites of NtrC-binding sites (results not shown). Reitzer *et al.* (1989) have provided evidence that NtrC-PO_4 binds weakly to a partial NtrC-binding site.

These factors may be artifactual and limited to the *E. coli* system used. However, they provided a convenient method to screen for the *T. ferrooxidans ntrC* gene.

This method involved a plate assay whereby large numbers of recombinant clones could be screened for increased expression of β-galactosidase activity from the *T. ferrooxidans nifH-lacZ* fusion plasmid in the *E. coli ntrC* strain ET8556. The medium used contained a low concentration of nitrogen, and positive clones which exhibited an elevated level of β-galactosidase activity where detected by flooding of the colonies with ONPG. This simple colourimetric assay has potential for broader use in the isolation of genes encoding transcriptional regulators - both positive and negative - provided a translational fusion between the target gene and *lacZ* is available (reviewed in Silhavy and Beckwith, 1985). Screening of recombinant genomic genebanks by colony hybridization with a heterologous DNA probe often requires careful fine-tuning to achieve a workable signal-noise ratio (Maniatis *et al.*, 1982). The use of a biological technique, such as complementation, which also does not require the use of radioactivity, is therefore preferable.

The screening technique developed in this study was tried since another selection method - growth on arginine as a sole nitrogen source - for identifying clones that were able to complement an *E. coli ntrC* mutant was unsuccessful when applied to the *T. ferrooxidans* gene bank. The reason for this became apparent once the *T. ferrooxidans ntrC* gene had been cloned.

E. coli ntrC ET8556 cells, containing several subclones of the *T. ferrooxidans ntrC* gene carried on different plasmid vectors, were tested for growth on agar plates containing GMM + 0,2% arginine (w/v) (results not shown). YMC10(pHC79), YMC10(Bluescript SK), ET8556(pK70), or ET8556(p804) cells (which contained either the *E. coli* or *K. pneumoniae ntrC* gene) formed large colonies on this medium. ET8556(pHC79), ET8556(Bluescript SK), ET8556(pT125), ET8556(pT127), ET8556(pT101), or ET8556(pT110) cells did not grow. ET8556(pT120) cells did, however, form small colonies after 72 h of incubation at 30°C.

The *T. ferrooxidans* NtrC protein may have been present at a high enough concentration to activate the *E. coli* arginine utilization operon in ET8556(pT120) cells because the *T. ferrooxidans ntrC* gene was expressed from the high copy-number Bluescript SK vector. pT101 and pT110 were constructed from the plasmid vectors pHC79 and pEcoR252, respectively, which replicate at a lower copy-number. The plasmid copy number may have been too low to raise the concentration of the active *T. ferrooxidans* NtrC protein within the ET8556(pT101) or ET8556(pT110) cells to enable the cells to grow on arginine as a sole nitrogen source. Adjacent vector promoters are oriented in the opposite direction to the *T. ferrooxidans ntrC* gene in pT110 and pT120 (Fig. 5.2), so transcriptional readthrough from these promoters is not responsible for differential levels of the *T. ferrooxidans* NtrC. The orientation of the *T. ferrooxidans ntrC* gene in the cosmid clone pT101 is not known. These results are consistent with the inability to clone the *T. ferrooxidans ntrC* gene by screening the *T. ferrooxidans* cosmid gene bank (vector pHC79) in *E. coli* ET8556 cells using growth on arginine as a selection method (section 5.3.1). *E. coli* ET8556 cells containing the *A. tumefaciens ntrC* gene (pBR325 vector) grew well when provided with arginine as a sole nitrogen source (Wardhan *et al.*, 1989), however a separate isolate of the same gene in a different (lower copy-number) vector, pACYC184, appeared unable to confer upon *E. coli* YMC17 (*ntrC*::Tn5) cells the ability to grow on minimal medium containing arginine (Rossbach *et al.*, 1987).

The predicted aa sequence of the section of the *T. ferrooxidans* NtrC obtained in this study covers the N-terminal domain which is highly conserved in NtrC proteins. These results suggested that the regulation of activity of the *T. ferrooxidans* NtrC may match that shown for other NtrC proteins, ie. the involvement of phosphorylation by an NtrB protein in response to levels of available nitrogen. *E. coli* ET8556 contains a point mutation in the *ntrC* gene, but the *ntrB* gene is intact. An explanation for the regulation of expression of the *T. ferrooxidans nifH-lacZ* in ET8556(pHlac20, pT120) cells in response to nitrogen levels may be modulation of the activity of the *T. ferrooxidans* NtrC by the *E. coli* NtrB. Alternatively, the *T. ferrooxidans ntrC* gene may be linked to an *ntrB* gene, the product of which may have fulfilled this function. Nucleotide sequence analysis of the region adjacent to the *T. ferrooxidans ntrC* on the plasmid pT120 should resolve this question.

Further evidence that the *T. ferrooxidans* NtrC comprises the RR element of a Ntr regulatory system from this organism is revealed by aa sequence comparisons with the N-terminal 100 aa of RR's which have been identified as involved in signal transduction in various processes in prokaryotes (reviewed in Stock *et al.*, 1989b; section 1.5). The NtrC and CheY proteins from enteric bacteria are two representative RR's which have been studied in most detail. Both proteins have been shown to be converted to an activator form by phosphorylation within the N-terminal domain mediated by a specific HPK (NtrB for NtrC; CheA for CheY)(Stock *et al.*, 1988a; Weiss and Magasanik, 1988).

S. typhimurium CheY is 129 aa in size, and therefore the whole protein is the equivalent of the N-terminal domain of other RR's, including NtrC. The three-dimensional structure of the *S. typhimurium* CheY has been solved to 0.27 nm resolution (Stock *et al.*, 1989a, 1990). A structural model is particularly useful in assigning functional significance to aa residues which may be conserved in members of a superfamily of proteins (eg. 25 RR's compared by Stock *et al.*, 1989b). The positions of aa residues, conserved within the N-terminal domain of these 25 RR's, were identified within the 3-dimensional structure of the CheY protein (Stock *et al.*, 1989b, 1990). These conserved residues were also found in the predicted aa sequence of the front end of the *T. ferrooxidans* NtrC (Fig. 5.6). This discussion has been included to provide further evidence that this sequence represents part of the *T. ferrooxidans* NtrC.

CheY is composed of a central core of five parallel β-strands surrounded by five α-helices (Fig. 5.7). Part of the hydrophobic core consists of three internal β-strands (β1, β3, and β4 - Fig. 5.6), which are characterized by clusters of hydrophobic residues. The sequence corresponding to the β1 strand is N-terminal to the aa sequence obtained for the *T. ferrooxidans* NtrC, however the clusters of hydrophobic residues corresponding to β3 and β4 are conserved in the *T. ferrooxidans* and *Bradyrhizobium* sp. [*Parasponia*] NtrC proteins (Fig. 5.6). Moreover, characteristically spaced hydrophobic residues, which correspond to the internal faces of amphipathic α-helices that flank the β-sheet, are conserved in the NtrC proteins (Fig. 5.6).

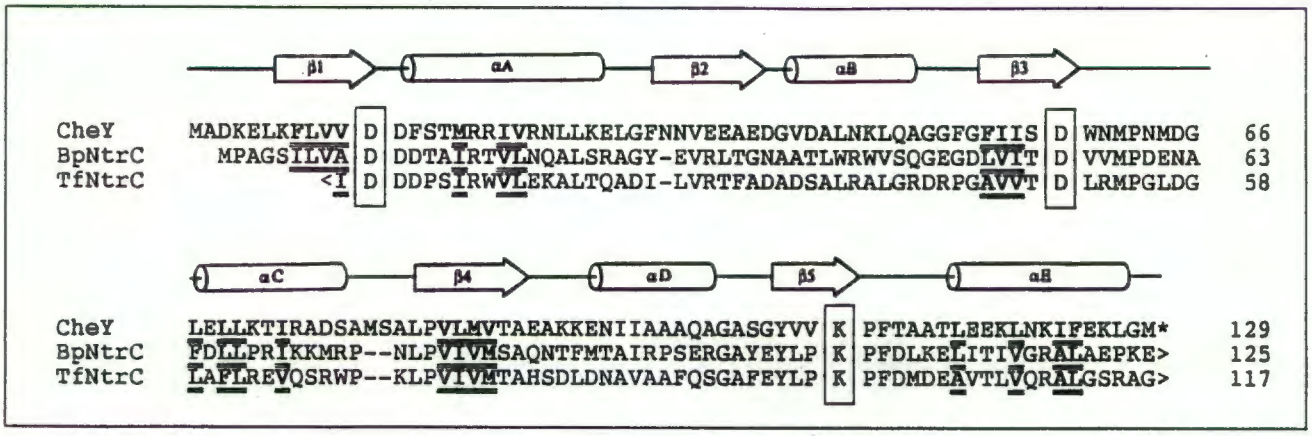


Fig. 5.6. Amino acid sequence alignment of part of the N-terminal region of the *T. ferrooxidans* NtrC with the *S. typhimurium* CheY and the N-terminal domain of the *Bradyrhizobium* sp. [*Parasponia*] NtrC. The secondary structure of CheY is represented with symbols which correlate to those structural elements shown in Fig. 5.7: arrows represent β-strands, cylinders represent α-helices, and lines represent loop regions (taken from Stock *et al.*, 1989b). Conserved hydrophobic residues are shown as double underlined bold-face letters; the three residues which are highly conserved amongst RR's are boxed.

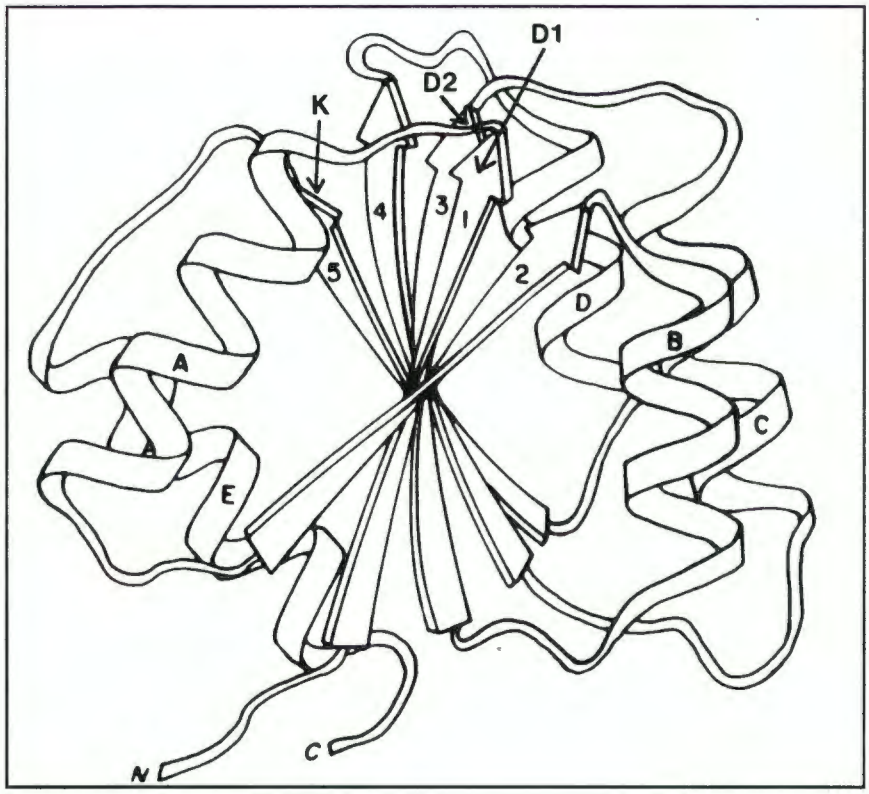


Fig. 5.7. Tertiary structure of the *S. typhimurium* CheY protein. Filled arrows represent the five-stranded parallel β-sheet (β-strands 1-5 on Fig. 5.6). The five outer α-helices are labelled A-E (α-helices A-E on Fig. 5.6). The three residues situated near the C-terminal edge of the β-sheet, which are highly conserved in corresponding positions of RR's, are indicated: Asp-13 (D1), Asp-57 (D2), Lys-109 (K) (taken from Stock *et al.*, 1989b).

Of functional significance is the 100% conservation of two Asp and one Lys residue amongst at least 25 RR's (including the *T. ferrooxidans* NtrC - Fig. 5.6). These residues are clustered at the edge of the β-sheet in the CheY protein - a position which is likely to be the active site for interaction with the HPK (Fig. 5.6, 5.7). Asp-57 of CheY has been shown to be a site of phosphorylation, and mutations at this position abolish chemotaxis (Sanders *et al.*, 1989). A common phosphotransfer enzymology amongst RR's has been suggested since NtrC has been shown to be phosphorylated at an Asp residue (Weiss and Magasanik, 1988).

It is interesting to note that the last few residues of the 162 aa section of the *T. ferrooxidans* NtrC determined in this study -VLITGESGS- (Fig. 5.4) are homologous to the residues at the start of the central region which is conserved between NtrC proteins and NifA proteins (Buikema *et al.*, 1985; Drummond *et al.*, 1986).

The linkage of the *T. ferrooxidans ntrC* gene to the *T. ferrooxidans glnA* gene is unlikely. The *T. ferrooxidans glnA* gene is located within 2,25- and 2,35-kbp *Pst*I fragments and these have been shown to hybridize to *Pst*I fragments of the same size from *T. ferrooxidans* chromosomal DNA (Barros *et al.*, 1985). Experiments were carried out to determine whether *Pst*I fragments of the same size were present on any of the cosmid clones which carried the *T. ferrooxidans ntrC* gene. pT101 and another positive cosmid clone, pT106, were chosen as the most likely clones which contained DNA inserts representative of a large section of the *T. ferrooxidans* chromosome surrounding the *T. ferrooxidans ntrC* gene. pT106 shared common restriction fragments with pT101 and therefore was assumed to carry the *T. ferrooxidans ntrC* gene but appeared most dissimilar to pT101 with respect to restriction fragments out of the six positive cosmid clones isolated (section 5.2.1). *Pst*I digests of pT101 and pT106 shared only 15-kbp of common fragments out of the 40-kbp of each cosmid clone (data not shown). However, neither cosmid clone contained a 2,25- or 2,35-kbp *Pst*I fragment which would be present if they contained the *T. ferrooxidans glnA* gene (data not shown).

Furthermore, *T. ferrooxidans* does not appear to have a *glnAntrBntrC* genetic organization since nucleotide sequence analysis of the region immediately downstream of the *T. ferrooxidans glnA* has revealed an ORF (encoding a protein with a putative N-terminal signal sequence) which bears no resemblance to any of the *ntr* genes sequenced to date or to any sequence in the available Genbank (Release 62) database (data not shown). The *ntrC* gene is not linked to the *glnA* gene in *R. meliloti* (Szeto *et al.*, 1987), *R. leguminosarum* (Filser *et al.*, 1986), or *A. tumefaciens* (Rossbach *et al.*, 1987).

CHAPTER 6

GENERAL DISCUSSION

The ecological niche inhabited by the autotrophic chemolithotroph *Thiobacillus ferrooxidans* is populated by relatively few other bacterial species, thus placing a limit on the potential for genetic exchange with other soil bacteria. This isolation raises the question as to whether enzyme structures and genetic regulatory systems that differ from those in the more extensively studied heterotrophic Gram-negative bacteria have evolved in *T. ferrooxidans*.

Prior to this study, it had been shown that two key components of nitrogen metabolism in *T. ferrooxidans* - glutamine synthetase and nitrogenase - were homologous to corresponding enzymes from heterotrophic bacteria (Rawlings *et al.*, 1987; Pretorius *et al.*, 1987). In this study, the *ntrA* and *ntrC* genes from *T. ferrooxidans* were identified and cloned in *Escherichia coli*. These components of an Ntr regulatory system from *T. ferrooxidans* were shown to act together with regulatory proteins from the enteric bacteria *E. coli* or *K. pneumoniae* to control expression from NtrA-dependent promoters in *E. coli*. In the context of the current understanding of the molecular mechanisms of transcriptional initiation, these results indicated that the *T. ferrooxidans* NtrA was able to combine in a functional complex with *E. coli* RNA polymerase which is believed to form a closed promoter complex at each NtrA-dependent promoter. The isomerization to open promoter complexes catalyzed by the relevant activator protein (uncharacterized for *fdhF*, NifA for *nifH*) could be inferred from the successful expression obtained for each system. It is not known whether there is a direct protein-protein interaction between the activator protein and the NtrA component of the RNA polymerase holoenzyme, although the correlation between one or two base changes in the NtrA recognition site and the activation specificity of either NifA or NtrC-PO_4 supports this hypothesis (Ow *et al.*, 1985; Buck and Cannon, 1989; Ray *et al.*, 1990).

These results extended the view that the major components of nitrogen metabolism in *T. ferrooxidans* - at the level of regulation of gene expression, as well as protein

structure - were similar to those in other Gram-negative bacteria. This similarity is not sufficient for full complementation of some functions in *E. coli*, such as NtrA-dependent growth on arginine as a sole nitrogen source; and the *T. ferrooxidans* NtrA appeared to be not as potent as the *E. coli* or *K. pneumoniae* NtrA proteins at mediating expression from the *K. pneumoniae nifH* promoter. Furthermore, the *T. ferrooxidans glnA* does not appear to have all the components of an Ntr regulatory system as found in enteric bacteria (see later). Elements of *T. ferrooxidans* physiology other than the systems associated with nitrogen metabolism may show a greater divergence from those in other bacteria. An example is transcriptional regulation of the *T. ferrooxidans recA* gene, which has no identifiable consensus DNA control elements, such as an SOS box or *lexA* binding site (Ramesar *et al.*, 1988).

The functional significance of the DNA control elements (NifA UAS and NtrA-dependent promoter) recognized upstream of the *T. ferrooxidans nifH* gene (Pretorius *et al.*, 1987) was supported by the results obtained in this study. Expression of the *T. ferrooxidans nifH-lacZ* fusion in *E. coli* was NtrA-dependent and activated by the *K. pneumoniae* NifA. NtrC-activated expression from the *T. ferrooxidans nifH* promoter in the presence of an *E. coli ntrA* gene was also observed. NtrC encoded by either *E. coli*, *K. pneumoniae*, or *T. ferrooxidans ntrC* genes could fulfil this function. It would be interesting to test expression of the *T. ferrooxidans nifH-lacZ* fusion in the presence of both the cloned *T. ferrooxidans ntrA* and *ntrC* genes in an *E. coli ntrAntrC* mutant. Activation by the *T. ferrooxidans* NtrC in *E. coli* raises the question as to whether NtrC is involved in expression of *nif* genes in *T. ferrooxidans* itself. The construction of *T. ferrooxidans ntrC* mutants, and the complementation of these mutants would provide an answer to this question. However, the difficulty of growing *T. ferrooxidans* on solid laboratory media, the absence of defined mutants, and the current rudimentary state of a genetic manipulation system for this organism precludes such an investigation in the immediate future.

This highlights a major challenge for studies on nitrogen fixation in *T. ferrooxidans*. A fundamental study is required to determine the environmental conditions under which *T. ferrooxidans* is able to carry out this process. Is it possible to grow

T. ferrooxidans anaerobically on a nitrogen-free defined solid medium? A major question is how *T. ferrooxidans* obtains enough energy under anaerobic conditions for this energetically expensive process, since *T. ferrooxidans* grows most vigorously when producing energy from the oxidation of reduced iron and sulphur compounds under aerobic conditions. Aerobic/anaerobic cycles, which were used to demonstrate N_2-fixation by *T. ferrooxidans* (Mackintosh, 1978) are unlikely to prevail in nature.

The study of *T. ferrooxidans* genes and their products in *E. coli*, where the tools of recombinant DNA technology can be applied, provides an insight into the molecular biology of this organism. Nucleotide sequence analysis of the *T. ferrooxidans ntrA* gene region revealed the presence of an upstream ORF1 which is also conserved in sequence and position in *R. meliloti*. The role of the ORF1 product is not known, although its conservation in a chemolithoautotroph and a symbiotic bacterium lends support to the hypothesis that it fulfils a general "housekeeping" role (Albright *et al.*, 1989b). The conservation of ORF3 in *T. ferrooxidans*, *P. putida*, *A. vinelandii*, *K. pneumoniae*, and *R. meliloti* indicates that the modulation of NtrA activity by the products of genes downstream of the *ntrA* gene observed in *K. pneumoniae* (Merrick and Coppard, 1989) may be a general phenomenon.

The predicted aa sequence of the *T. ferrooxidans* NtrA protein is of value as it provides further information as to the structural/functional significance of specific aa residues in NtrA proteins. NtrA has been shown to be a specialized sigma factor which directs expression of genes encoding components of diverse physiological processes, not only nitrogen metabolism (section 1.4; reviewed in Kustu *et al.*, 1989). Unravelling the molecular mechanism whereby NtrA is able to confer promoter specificity and the role of core RNA polymerase in these protein-DNA interactions is therefore of high priority. Construction of mutations in the putative C-terminal HTH DNA-binding motif, which is highly conserved in all six NtrA protein sequences available, and study of expression from mutant NtrA-dependent promoters would reveal which doublet, -GG- (-24) or -GC- (-12), this motif recognizes. This approach would also pinpoint which region of NtrA is responsible for the presumed second DNA-recognition function. Molecular modelling of the amino-acid-side-chain/base interactions would follow, and these data would aid interpretation of the

3-dimensional structure of an NtrA protein using X-ray crystallography. This would be particularly useful if the structure of the NtrA-RNA polymerase complex was to be determined. Two-dimensional crystals of *E. coli* σ70-RNA polymerase have been successfully grown on positively charged lipid bilayers, and a low-resolution 3-dimensional structure has been determined by electron crystallography (Darst *et al.*, 1989).

It is important to note that the cloning strategies for regulatory genes developed in this study were based on expression of genes whose products were not essential for growth. Positive recombinant *E. coli* transductants were identified by amplification of clearly visible phenotypes (gas formation or increased β-galactosidase activity). In contrast, the cloned *T. ferrooxidans ntrA* or *ntrC* genes carried on the cosmid gene bank vector pHC79 were unable to confer on the respective *E. coli ntrA* or *ntrC* mutants the ability to grow on arginine as a sole nitrogen source. This illustrates the difficulty of identifying heterologous genes, whose products show some functional similarity to the original *E. coli* product, by selection for growth. The techniques described in this thesis therefore provide simple methods for the screening of gene banks from other bacteria for *ntrA* and *ntrC* genes provided these genes are expressed in *E. coli*.

In this vein, experiments are currently underway to isolate the *T. ferrooxidans nifA* gene by an adaption of the method used to clone the *T. ferrooxidans ntrC* gene. The presence of consensus NifA UAS's (Pretorius *et al.*, 1987), and the demonstration of NifA-activated expression from the *T. ferrooxidans nifH* promoter (this study) point to the existence of such a gene in *T. ferrooxidans*. A *T. ferrooxidans* cosmid clone carrying a *nifA* gene may contain other linked *nif* genes. It would be particularly interesting to identify a gene encoding a negative regulator of *nif* expression from *T. ferrooxidans*. The homology observed between NifA proteins from several bacteria (such as *K. pneumoniae*, *A. vinelandii*, *R. meliloti*, and *B. japonicum* - Buikema *et al.*, 1985; Bennett *et al.*, 1988; Fischer *et al.*, 1988) does not appear to extend to the negative regulatory proteins; ie. NifL in *K. pneumoniae* (Drummond and Wootten, 1987) and *A. vinelandii* (Bali and Kennedy, Eighth International Conference on Nitrogen Fixation, 1990) is replaced in *R. meliloti* by a complex cascade of

regulators, including FixK, a homologue of the *E. coli* regulator Fnr (Batut *et al.*, 1989). This is consistent with the idea that N_2-fixation is carried out in each organism under different conditions; for example, the symbiotic bacteria fix nitrogen in the relatively nitrogen rich nodule environment, while *K. pneumoniae* requires nitrogen-limiting conditions for this process. The primary structure of a *T. ferrooxidans* negative regulator may provide some clues as to what signals mediate repression of N_2-fixation in *T. ferrooxidans*.

The strategy aimed at cloning the *T. ferrooxidans nifA* gene requires the development of a nitrogen-limited solid medium on which *E. coli ntrC* mutant ET8556 strains containing the *T. ferrooxidans nifH-lacZ* fusion plasmid pHlac20 are able to grow under anaerobic conditions. The use of NFDM (Cannon *et al.*, 1974), supplemented with low concentrations of casamino acids and yeast extract, is being investigated. Once it has been established that ET8556(pHlac20) cells exhibit low levels of expression of β-galactosidase activity under these conditions (ie. colonies stain a dull brown colour after flooding with ONPG), and that β-galactosidase activity is increased in ET8556(pHlac20) cells containing, in addition, the constitutively expressed *K. pneumoniae nifA* gene, it will be possible to screen the *T. ferrooxidans* cosmid gene bank. Positive clones should carry a NtrC-independent transcriptional activator of the *T. ferrooxidans nifH* gene which is active under anaerobic nitrogen-limiting conditions. The possibility of re-cloning the *T. ferrooxidans ntrC* gene will be easy to check. Isolation of a NtrC-dependent activator could be attempted using a *K. pneumoniae nifH-lacZ* fusion in *E. coli* YMC10 cells.

The effect of the cloned *T. ferrooxidans ntrA* and *ntrC* gene products on the expression of the cloned *T. ferrooxidans glnA* gene in *E. coli* would constitute an important set of experiments to be carried out. Barros *et al.* (1985) showed that glutamine synthetase activity in *E. coli glnAntrBntrC* deletion mutant cells containing the *T. ferrooxidans glnA* gene was regulated in response to nitrogen levels. Subsequent studies on the expression of *E. coli* glutamine synthetase activity (Magasanik, pers. comm.) have indicated that the use of glutamate in the media (which is taken up very poorly by *E. coli*) and the quantity of glutamine (0,15 mM) used for the "low" nitrogen medium in these experiments may not have provided sufficient growth for a true reflection of *T. ferrooxidans* GS activity in response to

available nitrogen. The *T. ferrooxidans glnA* gene shows a consensus -35, -10 promoter sequence which was overlapped by the only detectable consensus NtrC-binding site (repressor site?)(Rawlings *et al.*, 1987). No downstream *ntrBntrC* genes (this study) or any detectable NtrA-dependent promoter sequence (Rawlings *et al.*, 1987) were identified, which argues against a classical Ntr system of transcriptional regulation as shown for *glnA* genes from enteric bacteria (section 1.2.2.1). If the results of Barros *et al.* (1985) are correct, other regulatory factors encoded upstream of the *T. ferrooxidans glnA* are responsible for this regulation. Alternatively, crosstalk by other *E. coli* regulatory factors in the *E. coli glnAntrBntrC* strain may have played a role (discussed in section 1.5).

Studies on which factors are required for the expression of the *T. ferrooxidans glnA* gene in *E. coli* would be greatly facilitated by the construction of a translational fusion between the *T. ferrooxidans glnA* and the *lacZ* gene. The transcriptional regulation of expression of the *T. ferrooxidans glnA* gene in various *E. coli* backgrounds could then be studied by experiments in parallel: glutamine synthetase assays to measure expression in strains carrying the intact *T. ferrooxidans glnA* gene; and β-galactosidase assays to measure expression from the *T. ferrooxidans glnA* regulatory region without the influence of *T. ferrooxidans* GS activity on cell metabolite balances under different growth conditions. The *glnA-lacZ* fusion would also be useful in *E. coli glnA*$^+$ strains.

The functional significance of the putative NtrC-binding site upstream of the *T. ferrooxidans glnA* could be tested directly with the *T. ferrooxidans* NtrC in *E. coli*. These experiments, however, must await the completion of the nucleotide sequence analysis of the *T. ferrooxidans ntrC* gene and its flanking regions to determine whether it is linked to an *ntrB*-like gene. Evidence for such linkage on the recombinant clone pT120 (Fig. 5.3) could be obtained from parallel β-galactosidase assays of *E. coli ntrB*$^+$*ntrC* ET8556(pHlac20, pT120) cells and *E. coli ntrBntrC* YMC11(pHlac20, pT120) cells [Either of the *glnAntrBntrC* strains YMC11 (Backman *et al.*, 1981) or ET8051 (Tuli *et al.*, 1982) could be used]. If expression of the *T. ferrooxidans nifH-lacZ* fusion was also regulated in response to levels of available nitrogen in the *E. coli ntrB* background one could infer that the product of a cloned *T. ferrooxidans ntrB*-like gene was responsible.

NtrA-dependent expression of the *T. ferrooxidans glnA-lacZ* fusion in the presence of the cloned *T. ferrooxidans ntrA* could be tested in an *E. coli ntrA* strain, such as TH1. The effect of the cloned *T. ferrooxidans ntrC* (and *ntrB*, if it is linked) gene product could be checked in the *E. coli* YMC11 background. *E. coli* ET6362 (*ntrA*::Tn10, *ntrC*::Tn5)(Backman *et al.*, 1981) or, more preferably, a strain deleted for the *glnAntrB* genes as well (*E. coli* YMC23 or YMC27, Backman *et al.*, 1983) would be a useful host strain for complementation studies with the cloned *T. ferrooxidans ntrA* and *ntrC* genes together, if necessary.

Alternatively, the results of Barros *et al.* (1985) could be confirmed by identifying the genetic determinants upstream of the *T. ferrooxidans glnA* by providing them *in trans* and showing nitrogen regulation. A *glnB*-like gene, which encodes the P_{II} protein, has been identified (Holtel and Merrick, 1988) immediately upstream of the *glnA* gene in *R. leguminosarum* (Colonna-Romano *et al.*, 1987), *B. japonicum* (Carlson *et al.*, 1985; Martin *et al.*, 1989), *A. brasiliense* (Bozouklian and Elmerich, 1986), and *R. capsulatus* (Kranz *et al.*, 1990), but *glnB* and *glnA* are unlinked in the enteric bacteria *K. pneumoniae* (Holtel and Merrick, 1988) and *E. coli* (Son and Rhee, 1987). The *T. ferrooxidans* GS is regulated by adenylylation, which implicates the presence of a cyclic cascade involving a P_{II} protein in *T. ferrooxidans*, however no *glnB*-like gene is present within 470-bp upstream of the *T. ferrooxidans glnA* (Rawlings *et al.*, 1987).

APPENDIX A

ESCHERICHIA COLI STRAINS AND PLASMIDS USED.

E. coli strain	Genotype/description	Reference/origin
CSH36	F'*lacI proA*$^+$*B*$^+$ Δ(*lacpro*)*supE thi*	Miller (1972)
ET8045	*ntrA*208::Tn10 Δ*lacU*169 *thi strA rhaD* previously called ET6059	MacNeil *et al.* (1982a) Pahel and Tyler (1979)
ET8556	*ntrC*1488 *lacZ*::IS*I gyrA rbs hut*C_K[c]	Merrick (1983)[a]
FM911	Δ*fdhF recA*56 Δ(*argF-lac*)*U*169	Zinoni *et al.* (1986)[b]
GM41	*dam thi*-1 *rel*-1 Hfr H	Marinus (1973)
LK111	lacI^q lacZΔM15 lacY$^+$ *thi*-1	Zabeau and Stanley (1982)
TH1	Δ*ntrA* Δ*lacU*169 *thi*-1 *endA supE*44 *hsdR*17	de Bruijn and Ausubel (1983)[c]
YMC10	Δ*lacU*169 *thi endA hsr hut*C_K[c] Pro$^+$	Backman *et al.* (1981)
YMC22	*ntrA*208::Tn10 Δ*lacU*169 *thi endA hsr* previously called YMC18	Reitzer and Magasanik (1986) Ueno-Nishio *et al.* (1983)
Plasmid	**Description**	**Reference/origin**
Bluescript SK$^+$	Ap^R	Stratagene, San Diego
Bluescript KS$^+$	Ap^R	Stratagene, San Diego
pACYC184	Clm^R Tet^R	Chang and Cohen (1978)
pBN208	Ap^R *fdhF*::*lacZ*	Birkmann *et al.* (1987b)[b]
pCK3	Tet^R *K. pneumoniae nifA*^{constitutive} IncP	Kennedy and Drummond (1985)[a]
pEcoR251	Ap^R *EcoRI*	Zabeau and Stanley (1982)
pEcoR252	pEcoR251 derivative (*Pst*I site in *amp* mutated)	P. Janssen
pFB71	Clm^R Tet^R *K. pneumoniae ntrA*$^+$	de Bruijn and Ausubel (1983)[c]
pFB514	Clm^R *K. pneumoniae glnA*$^+$*ntrB*$^+$*C*$^+$	de Bruijn and Ausubel (1981)[c]
pHC79	Ap^R cos (pBR322 derivative)	Hohn and Collins (1980)
pIMP11	Ap^R *T. ferrooxidans nifH*$^+$	Pretorius *et al.* (1986)
pMB1	Ap^R *K. pneumoniae nifH*::*lacZ*	Buck *et al.* (1985)[a]
pMC1403	Ap^R *lac*'*ZYA*	Casadaban *et al.* (1983)
pUC18	Ap^R	Norrander *et al.* (1983)
pUC19	Ap^R	Norrander *et al.* (1983)
p804	Ap^R *E. coli ntrB*$^+$*ntrC*$^+$	Tuli *et al.* (1982)[d]

[a] Gift from C. Kennedy
[b] Gift from A. Böck
[c] Gift from F. Ausubel
[d] Gift from R. Haselkorn

APPENDIX B

MEDIA, BUFFERS, AND SOLUTIONS

APPENDIX B

MEDIA, BUFFERS, AND SOLUTIONS

All media, buffers, and solutions were sterilized by autoclaving at 121°C for 20 min unless otherwise indicated. Heat labile substances were sterilized by filtration through 0,22 µm membrane filters (Millipore).

B.1 Media

B.1.1 Acidified TK salts (500 ml)(Tuovinen and Kelly, 1974)

$(NH_4)_2SO_4$	3 g
KCl	0,1 g
K_2HPO_4	0,5 g
$MgSO_4.7H_2O$	0,5 g
$Ca(NO_3)_2$	0,01 g
Distilled H_2O	500 ml

pH was adjusted to 1,6 and the solution was autoclaved.

B.1.2 Glucose minimal medium (GMM) (1 000 ml)

Salts solution:	K_2HPO_4	10,5 g
	KH_2PO_4	4,5 g
	$Na_3Citrate$	2,5 g
	Distilled H_2O	200 ml
	$MgSO_4.7H_2O$	0,2 g
	Distilled H_2O	10 ml
	Glucose	2 g
	Distilled H_2O	10 ml
	Agar (Oxoid No. 1)	15 g
	Distilled H_2O	780 ml

Each solution was autoclaved seperately, and combined together with the appropriate nitrogen source (amino acid or $(NH_4)_2SO_4$) and 1 ml of a filter sterilized solution of thiamine (0,005g/ml stock). **GGMM** contained glucose at a concentration of 20 g/l.

B.1.3 Luria-Bertani medium (LB) (1 000 ml)

Bacto tryptone	10 g
Yeast extract	5 g
NaCl	5 g
Distilled water	1 000 ml

Solid media (**LBA**) contained 1,5% (w/v) agar.

B.1.4 NFDM liquid medium (1 000 ml)(Cannon *et al.*, 1974)

Solution A:	$MgSO_4.7H_2O$	0,1 g
	$Na_2MoO_4.2H_2O$	25 mg
	$FeSO_4.7H_2O$	25 mg
	Distilled H_2O	400 ml
Solution B:	K_2HPO_4	12,06 g
	KH_2PO_4	3,4 g
	Distilled H_2O	500 ml
Solution C:	Glucose	20 g
	Distilled H_2O	100 ml

Solutions A, B, and C were autoclaved seperately, and then combined with 1 ml of a filter sterilized solution of thiamine (0,005g/ml stock).

B.1.5 TGYEP medium (1 000 ml)(Begg *et al.*, 1977)

Tryptone	10 g
Yeast extract	5 g
K_2HPO_4	5,49 g
KH_2PO_4	9,32 g
Distilled H_2O	900 ml
Glucose	8 g
Distilled H_2O	100 ml

Solutions were autoclaved seperately, combined, and the pH adjusted to 6,5. For solid medium 15 g agar was added to the first solution. **TGYEPF** medium consisted of 1 l **TGYEP** supplemented with 30 ml NaFormate (1 M stock); **TGYEPN** consisted of 1 l **TGYEP** supplemented with 40 ml KNO_3 (1 M stock). **TGYEPF** overlay plates contained 0,8% agar (w/v).

B.1.6 TK liquid medium (1 000 ml)(Tuovinen and Kelly, 1974)

Solution A:	$(NH_4)_2SO_4$	3 g
	KCl	0,1 g
	K_2HPO_4	0,5 g
	$MgSO_4.7H_2O$	0,5 g
	$Ca(NO_3)_2$	0,01 g
	Distilled H_2O	500 ml
Solution B:	$FeSO_4.7H_2O$	45 g
	Distilled H_2O	500 ml
	Adjust to pH 1,8 with H_2SO_4	

Solution A was autoclaved, allowed to cool, and mixed aseptically to filter sterilized solution B. After mixing, the pH was adjusted to 1,7 with H_2SO_4.

B.1.7 YT medium (1 000 ml)

Bacto tryptone	16 g
Yeast extract	10 g
NaCl	5 g
Distilled water	1 000 ml

Solid media (YTA) contained 1,5% (w/v) agar. For pUC or Bluescript recombinant selection, IPTG (0,1 ml) and XGal (0,8 ml) were added to 250 ml YTA before pouring the plates.

B.2 Media additives

Media were cooled to 50°C before addition of antibiotics, XGal, or IPTG. Plates containing these additives were stored for no longer than one week at 4°C.

B.2.1 Antibiotics

Antibiotic stock solutions were as follows:

Ampicillin (Ap)	100 mg/ml water
Chloramphenicol (Clm)	20 mg/ml ethanol (96%)
Tetracycline (Tet)	15 mg/ml ethanol (50%)

All antibiotics were filter sterilized and stored at -20°C, except for Tet which was always made fresh.

B.2.2 Amino acids

Solutions were made fresh before use and filter sterilized.

L-glutamine (Sodium salt; $M_r = 146,1$) dissolved completely in distilled H_2O at concentrations below 0,219 g/10 ml (0,15 M).

L-arginine (Sodium salt; $M_r = 174,2$) dissolved completely in distilled H_2O at concentrations below 0,2 g/10 ml (2%).

B.2.3 IPTG (isopropyl-β-D-thio-galactopyranoside)

IPTG (100mM)	23,4 mg
Distilled water	1 ml

The solution was stored in aliquots at -70°C.

B.2.4 XGal (5-bromo-4-chloro-3-indolyl-β-galactoside)

XGal (2% w/v)	0,2 g
Dimethylformamide	10 ml

The solution was stored in aliquots at -70°C.

B.3 Buffers and solutions

B.3.1 ATP (10x) (Maniatis *et al.*, 1982)

Adenosine triphosphate	30 mg
Distilled water	5 ml

pH was adjusted to 7,0 with 0,1 N NaOH before making up to 5 ml. Stored in 100 µl aliquots at -70°C. Remainder was discarded once defrosted.

B.3.2 Denhardt's solution (50x) (Maniatis *et al.*, 1982)

Ficoll	1 g
Polyvinylpyrrolidone	1 g
BSA (Fraction V)	1 g
Distilled water	to 100 ml

Filter sterilized and stored in aliquots at -20°C.

B.3.3 DNA loading solution (6x)

Bromophenol blue	0,25 g
Sucrose	40 g
Distilled water	to 100 ml

DNAase-free RNAase A was routinely added at a final concentration of 500 µg/ml. The solution was stored at 4°C.

B.3.4 DTT (1M)

DTT	0,618 g
Sodium acetate (0,01 M, pH 5,2)	4 ml

Filter sterilized.

B.3.5 EDTA (0,5 M, pH 8,0) (Maniatis *et al.*, 1982)

EDTA.$2H_2O$	168,1 g
Distilled water	to 1 000 ml

EDTA will only dissolve when pH has been adjusted to 8,0. (Used approximately 20 g NaOH pellets for this purpose).

B.3.6 Ethidium bromide solution

A solution of 10 mg/ml (2,7-diamino-10-ethyl-9-phenyl-phenanthridinium bromide) was made in distilled water and stored in a dark bottle.

B.3.7 Exo-nuclease III shortening solutions (Henikoff, 1987)

B.3.7.1 Exo buffer

Tris/HCl (1 M, pH 8,0)	660 µl
$MgCl_2$ (0,1 M)	66,4 µl
Distilled water	9,27 ml

B.3.7.2 Klenow mixture

Tris/HCl buffer (0,1 M, pH 8,0)	3 µl
$MgCl_2$ (1 M)	6 µl
Distilled water	20 µl

B.3.7.3 Ligase mixture

Ligation buffer (10x)	144 µl
Distilled water	1440 ml

B.3.7.4 S_1 buffer (10x)

KOAc (3 M)	1,1 ml
NaCl (5 M)	5 ml
Glycerol	5 ml
$ZnSO_4$	30 mg

B.3.7.5 S_1 mixture

S_1 buffer (10x)	41 µl
Distilled water	259 µl
S_1 nuclease (60 U)	1,5 µl

B.3.7.6 S_1 stop

Trisma Base (no HCl)	0,3 M
EDTA (pH 8,0)	0,05 M

B.3.8 Isopropanol (salt saturated)

Isopropanol and 5 M NaCl (dissolved in 10mM Tris-HCl and 1 mM EDTA, pH 8,5) were mixed at a ratio of 2:1. The precipitated NaCl was allowed to settle out of solution, and the upper solvent phase was used.

B.3.9 Klenow (DNA polymerase I) buffer

The buffer was made according to the following table and stored at -20°C.

Stock solution	Final conc.	/10 ml
Tris-Cl (1 M, pH 7,6)	0,1 M	1 ml
$MgCl_2$ (1 M)	0,1 M	1 ml
NaCl (5 M)	0,5 M	1 ml
β-mercaptoethanol (14 M)	0,7 M	500 µl
Distilled water		6,5 ml

B.3.10 Ligase dilution buffer

The buffer was made according to the following table and stored at -20°C.

Stock solution	Final conc.	/10 ml
Tris-Cl (1 M, pH 7,6)	20 mM	0,2 ml
EDTA (0,5 M, pH 8,0)	1 mM	2 µl
DTT (0,5 M)	5 mM	10 µl
KCl (1 M)	60 mM	0,6 ml
Glycerol	44% (v/v)	4,4 ml
Distilled water		4,788 ml

B.3.11 Ligation buffer (10x)

The buffer was made according to the following table and stored in aliquots at -70°C.

Stock solution	Final conc.	/ml
Tris-Cl (1 M, pH 7,6)	66 mM	660 µl
$MgCl_2$ (1 M)	6 mM	66 µl
ATP (0,1 M)	1 mM	100 µl
DTT	0,1 M	15,4 mg
Distilled water		174 µl

B.3.12 Phenol (TE saturated)

Phenol (200 g, Merck) was melted at 65°C and 0,3 g of 8-hydroxyquinoline was added. The phenol was extracted three times with TE buffer (10x) or until the pH of the aqueous phase was approximately 7,6. The phenol was stored under TE buffer (1x) at -20°C.

B.3.13 Prehybridization solution

Stock solution	Final conc.	/100 ml
SSC buffer (20x)	(6x)	30 ml
SDS (10%)	0,5%	5 ml
Denat. SS-DNA (10 mg\ml)	100 µg/ml	1 ml
Denhardt's solution (50x)	5x	10 ml
EDTA (0,5 M, pH 8,0)	10 mM	2 ml
Distilled water		52 ml

B.3.14 Restriction enzyme buffers (10x)

Stock solution	Final conc.
Tris-Cl (1 M, pH 7,9)	0,1 M
$MgCl_2$ (1 M)	0,1 M
DTT (0,5 M)	10 mM
BSA (10 mg/ml)	1 mg/ml
Glycerol	44% (v/v)
NaCl (5M)	0, 50, 100, or 150 mM

The buffers were made using the following table and stored at -20°C.

	Salt concentration (mM)			
Stock solution	0	50	100	150
Tris-Cl	1 ml	1 ml	1 ml	1 ml
$MgCl_2$	1 ml	1 ml	1 ml	1 ml
DTT	0,2 ml	0,2 ml	0,2 ml	0,2 ml
BSA	1 ml	1 ml	1 ml	1 ml
Glycerol	4,4 ml	4,4 ml	4,4 ml	4,4 ml
NaCl	-	1 ml	2 ml	87,7 mg
Distilled H_2O	2,4 ml	1,4 ml	0,4 ml	2,4 ml

B.3.15 Restriction enzyme dilution buffer

The buffer was made according to the following table and stored at -20°C.

Stock solution	Final conc.	/10 ml
Tris-HCl (1 M, pH 7,5)	10 mM	0,1 ml
NaCl (5 M)	50 mM	0,1 ml
Distilled water		5,3 ml
Filter sterilize this solution and then add the following constituents:		
β-mercaptoethanol (14 M)	10 mM	7 µl
Gelatin (10 mg/ml)	100 µg/ml	0,1 ml
Glycerol	44% (v/v)	4,4 ml

B.3.16 RNAase A (DNAase-free)

A stock of pancreatic RNAase (RNAase A) at 10 mg/ml in 10 mM Tris-Cl (pH 7,5), 15 mM NaCl was heated to 100°C for 15 min. This was allowed to cool slowly to room temperature before being aliquotted for storage at -20°C.

B.3.17 *Sma*I restriction enzyme buffer (10x)

Stock solution	Final conc.	/10 ml
Tris-HCl (1 M, pH 8,0)	0,1 M	1 ml
KCl (1 M)	0,2 M	2 ml
$MgCl_2$ (1 M)	0,1 M	1,ml
DTT (0,5 M)	10 mM	0,2 ml
Glycerol	44% (v/v)	4,4 ml
Distilled water		1,4 ml

B.3.18 Salmon sperm DNA (SS-DNA)

A 10mg/ml solution was made in TE buffer. The DNA solution was sheared by repeated passage (12x) through a 18-gauge hypodermic needle. The solution was aliquotted and stored at -20°C. Immediately before use the DNA was denatured by boiling for 10 min followed by cooling on ice.

B.3.19 SM buffer (1x) (phage λ storage, dilution)

NaCl	5,8 g
$MgSO_4.7H_20$	2 g
Tris-Cl (1 M, pH 7,5)	50 ml
2% gelatin solution	5 ml
Distilled water	to 1 000 ml

Autoclaved.

B.3.20 Sodium acetate (3 M, pH 5,2)

Sodium acetate.$3H_2O$	4,08 g
Distilled water	to 10 ml

Adjusted pH with glacial acetic acid. Autoclaved.

B.3.21 Sodium phosphate buffer (0,1 M, pH 7,0)

Solution A:	$NaH_2PO_4.H_2O$ (0,2 M)	27,6 g/l
Solution B:	$Na_2PO_4.12H_2O$ (0,2 M)	71,6 g/l

Solution A (195 ml), Solution B (305 ml), and distilled water (500 ml) were combined, and the pH adjusted to 7,0.

B.3.22 SSC (20x)

NaCl (3 M)	175,3 g
Sodium citrate (0,3 M)	88,2 g
Distilled water	to 1 000 ml

Adjusted pH to 7,0 with NaOH (10 N). Autoclaved.

B.3.23 Tris acetate buffer (50x)

Tris base	242 g
Acetic acid	57,1 ml
EDTA (0,5 M, pH 8,0)	100 ml
Distilled water	to 1 000 ml

B.3.24 TE (Tris-EDTA) buffer (100x)

Tris base	121 g
EDTA (0,5 M, pH 8,0)	200 ml
Distilled water	to 1 000 ml

Adjusted pH to 8,0, autoclaved, and diluted with sterile water before use.

B.3.25 TSB solution

LB 150 ml

pH to 6,1 with approximately 0,5 ml 0,1 M HCl.

PEG 4000 15 g

Dispensed in 20 ml aliquots and autoclaved. Added DMSO (5% final), $MgSO_4$ (10 mM final), $MgCl_2$ (10 mM final), and, when necessary, glucose (10 mM final) immediately before use.

B.3.26 Z-buffer (pH 7,0)

Stock solution	Final conc.	/1 000 ml
$Na_2HPO_4.12H_2O$	60 mM	21,49 g
$NaH_2PO_4.H_2O$	40 mM	5,52 g
KCl	10 mM	0,75 g
$MgSO_4.7H_2O$	1 mM	0,246 g
β-mercaptoethanol (14 M)	50 mM	3,6 ml
Distilled water		to 1 000 ml

pH was adjusted to 7,0. Buffer was not autoclaved, and stored at 4°C

APPENDIX C

GENERAL TECHNIQUES

APPENDIX C

GENERAL TECHNIQUES

C.1 DNA preparation

C.1.1 Extraction of chromosomal DNA from *T. ferrooxidans*

Cells were inoculated into two litres of TK medium (Appendix B.1.6), and incubated at 30°C with constant aeration until the iron was completely oxidized (approximately 10 d). The cells were harvested by centrifugation at 10 000 rpm for 10 min, and washed twice in acidified TK salts (Appendix B.1.1). The pellets were resuspended in 4 ml of a 25% sucrose, 2 mM EDTA, 50mM Tris-HCl (pH 8,0) solution, and incubated at -20°C for 1 h. Proteinase K (1 mg/ml final concentration) was added to the frozen suspension, and the samples were shaken gently at room temperature until they thawed. SDS was added to a final concentration of 1%, and the samples were kept on ice for 15 min, followed by RNase (50 µg/ml) digestion for 20 min at 37°C. The viscous, opaque and often brown-coloured samples were dialysed at room temperature against numerous changes of TE buffer (pH 8,0) until they became translucent (24 - 60 h). The samples were purified of proteins by three phenol-chloroform-isoamyl alcohol (25:24:1) extractions followed by two diethyl-ether extractions. The DNA purification procedure was completed by an overnight dialysis against TE buffer (pH 8,0), at room temperature.

C.1.2 Plasmid preparation: small scale (miniprep)

Plasmid was isolated from a 5 ml overnight culture (LB + Ap, 100 µg/ml) using an adaption of the method of Birnboim and Doly (1979) as described by Ish-Horowicz and Burke (1981). Cells from a 1,5 ml sample of the culture were harvested by centrifugation in an Eppendorf microfuge tube for 1 min. The pellet was resuspended in 200 µl Solution I (50 mM glucose; 25 mM Tris-HCl, pH 8,0), incubated for 5 min at room temperature, and then 400 µl of Solution II (0,2 M NaOH, 1% (w/v) SDS) was added. The sample was vortexed briefly and placed on ice for 5 min, before the addition of 300 µl ice-cold Solution III (5 M KOAc, pH 4,8). The sample was vortexed briefly, and, after 5 min on ice, cellular debris and denatured chromosomal

DNA were pelleted by centrifugation for 10 min. The supernatant (700 µl) was removed to a fresh tube and sedimented with an equal volume of isopropanol by centrifugation for 5 min. The pellet was resuspended in 540 µl TE buffer (pH 8,0), and 60 µl 5 M $NaClO_4$ was added. The DNA was resedimented with an equal volume of isopropanol by centrifugation for 5 min, washed with 70% ethanol, air dried, and resuspended in 50 µl TE buffer (pH 8,0).

C.1.3 Plasmid preparation: Rapid miniprep for sizing exonuclease III shortened clones

This method enabled processing of at least 48 clones from bacterial colony to sized plasmid within 12 h. Best results were obtained for preparation of plasmids which replicated at a high copy-number, such as those derived from the pUC or Bluescript vectors. *E. coli* cells were inoculated from a colony into 500 µl YT medium (+ antibiotic) in an Eppendorf tube and incubated with agitation for 6 h at 37°C. After incubation, the cells were pelleted by centrifugation for 2 min. The pellet was resuspended in 60 µl STE buffer (100 mM NaCl; 20 mM Tris pH7,5; 10 mM EDTA). An equal volume of phenol-chloroform (1:1) was added, the sample was vortexed for 15 sec, and then centrifuged for 5 mins. Approximately 40 µl of the aqueous phase was transferred to a fresh tube. A 10 µl sample + 5 µl loading buffer (containing 1mg/ml DNase-free RNAase A)(Appendix B.3.16) was loaded on a 0,8% agarose gel and electrophoresed for 5 h at 70 volts. The distance moved by the fastest-migrating band (covalently closed circular DNA) was inversly proportional to plasmid size. Preparation by the same method of control plasmids of known size (eg. parental plasmid used for exonuclease III shortening [5,6-kbp] and vector [2,8-kbp]) and electrophoresis of these on the same gel provided a rapid method to identify exonuclease III shortened clones of a specific size. Plasmid prepared by this method was not suitable for restriction endonuclease digestion.

C.1.4 Plasmid preparation: large scale (maxiprep)

A 200 ml culture was grown overnight at 37°C in the presence of the appropriate antibiotic. The cells were harvested by centrifugation at 6 000 *g* for 5 min and then resuspended in 4 ml Solution I. After 5 min at room temperature 8 ml Solution II was added, and the mixture was kept on ice for 5 min, before the addition of 6 ml ice cold Solution III. After a further 5 min on ice the cellular debris was removed by centrifugation at 27 000 *g* for 15 min. An equal volume of isopropanol was added to the supernatant and the DNA was precipitated by centrifugation at 27 000 *g* for 10 min. The pellet was washed with 70% ethanol, resuspended in 4,2 ml TE buffer (pH 8,0), and purified by isopycnic CsCl-EtBr ultracentrifugation (Maniatis *et al.*, 1982). The plasmid preparation was prepared for ultracentrifugation by the addition of CsCl (1,3 mg/ml final conc.) and EtBr (0,5 ml of a 10 mg/ml stock). The solution was centrifuged at 25 000 *g* for 15 min to precipitate any remaining protein debris. The refractive index of the supernatant was adjusted to 1,394, the sample sealed in Beckman Quickseal ultracentrifuge tubes and centrifuged for 12 h at 55 000 rpm at 15°C in a Beckman Vti 65.2 rotor. The plasmid DNA band was visualized by long wave UV light (350 nm), and removed in the smallest volume possible. The EtBr was removed by extraction (3 times) with equal volumes of NaCl-saturated isopropanol (Appendix B.3.8). The DNA was precipitated from the CsCl solution by the addition of two volumes of water followed by an equal volume of isopropanol, and centrifugation in an Eppendorf microfuge for 15 min. The pellet was washed in 70% ethanol, resuspended in 200 µl TE buffer (pH 8,0), and the concentration was determined spectrophotometrically by measuring the absorbance of 10 µl (diluted in TE) between 220 and 310 nm. The concentration was determined by using the relationship A_{260} = 1 for 50 µg/ml double-stranded DNA.

C.2 Restriction endonuclease digestion

Restriction digests were carried out using one of the four restriction buffers (Appendix B.3.14) according to the salt requirements of the particular enzyme. The enzyme *Sma*I required a unique buffer (Appendix B.3.17). Digestion volumes were routinely 20 µl containing 300 - 500 ng DNA and one unit of restriction enzyme. Digestions were done at 37°C (most enzymes) for 1 h. Concentrated enzyme stocks were diluted to 1 or 2 units using universal restriction enzyme dilution buffer

(Appendix B.3.15). For electrophoretic analysis, the digestions were terminated by the addition of 5 µl DNA loading solution (Appendix B.3.3) to the 20 µl digestions. If the sample was to be used for ligation the digestion was terminated by a phenol-chloroform extraction. The DNA solution was extracted with the addition of phenol (1/10 volume, TE-saturated) and an equal volume of chloroform:isoamyl alchohol (24:1). The mixture was vortexed briefly, and the two phases were separated by centrifugation. The aqueous phase was extracted twice with chloroform:isoamyl alchohol (24:1). The DNA was precipitated by the addition of 5M $NaClO_4$ (1/10 volume), an equal volume of isopropanol, and 15 min centrifugation. If the DNA concentration was less than 2 µg/100 µl, *E. coli* tRNA was added (2 µg) before precipitation. After centrifugation the pellet was washed with 70% ethanol and resuspended in TE buffer (pH 8,0). When necessary, 5'-protruding termini produced from restriction endonuclease digestion were "filled-in" using the Klenow fragment from *E. coli* DNA Polymerase I (Amersham). Routinely, 0,5 µg DNA was incubated with 0,5 U Klenow (Amersham) in Klenow buffer (Appendix B.3.9) with 0,25 mM of each dNTP in a volume of 25 µl at 37°C for 10 min. The reaction was terminated by a phenol-chloroform extraction.

C.3 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out using a horizontal submerged gel system. Tris-acetate buffer (Appendix B.3.23) was used routinely. Sigma type II agarose was used at varying concentrations. The amount of DNA loaded/lane also varied with the sizes and number of fragments but under normal circumstances about 300 ng of plasmid DNA was used. The gels were electrophoresed at 2 V/cm for 16 h. Gels were stained in electrophoresis buffer containing EtBr (0,5 µg/ml) for 15 - 30 min. DNA bands were visualised using a 254 nm transilluminator. A 310 nm transilluminator was used if the DNA was to be recovered from the gel. Gels were photographed using a Polaroid CU-5 Land camera fitted with a red filter and a fixed focal length attachment. Polaroid type 667 film (ASA 3 000) was used with an exposure time of 1-2 sec at f4.7. If a negative was required then a Polaroid type 665 film (ASA 64) with an exposure of 120-140 sec at f4.7 was used.

C.4 DNA ligation reactions

DNA ligation reactions were of two basic types: recircularization of plasmids for the isolation of deletion clones (low DNA concentrations, 1 pmole DNA/ml) and recombination reactions, for example in subcloning (5 pmole DNA/ml). DNA concentration was calculated using the formula 1 pmole = (0,662 x kbp)µg. Ligation reactions containing DNA, ligation buffer (Appendix B.3.11) and water to the required volume, were performed in sterile microfuge tubes. Sticky-end ligations were performed at room temperature for 3 h or at 15°C overnight using 0,1-0,25 U of ligase, whereas blunt-end ligations were performed at room temperature for 3-20 h using 20-100 x more ligase.

C.5 Subcloning protocol

The rapid subcloning protocol of Struhl (1985) was used. The DNA fragments were separated by electrophoresis through low melting point (LMP) agarose (1%) (Seaplaque^R) in Tris-acetate buffer (50 mM, pH 8,2, no EDTA, no EtBr). The gel was stained with EtBr after electrophoresis and the DNA bands were viewed under UV light (310 nm), as briefly as possible. The desired bands were excised using sterile scalpel blades in as small a volume as possible. The gel slices were melted at 70°C for 5 min in a microfuge tube and the required amounts (2 µl vector DNA, 8 µl insert DNA) were added hot to the prepared ligation mixture containing ligation buffer, ligase and water (10 µl). The ligation was incubated at room temperature for 3 h. Before transformation of *E. coli* competent cells, the gelled ligation reaction was melted at 70°C for 5 min, and then diluted with 4 volumes of TSB solution (Appendix B.3.25).

C.6 The preparation and transformation of competent *E. coli* cells

E. coli cells were made competent for DNA uptake according to the method of Chung and Miller (1988). A 1/100 dilution of an overnight *E. coli* culture in LB was inoculated into 25 ml prewarmed LB and incubated at 37°C, with shaking, until the culture had reached early exponential phase (OD_{600} = 0,3-0,6)(2-4 h). The cell culture was poured into a pre-cooled sterile SS34 tube and the cells were harvested at 8 000 *g* for 5 min at 4°C. The cell pellet was resuspended in 2,5 ml (1/10 volume) ice-cold transformation and storage buffer (TSB) and held on ice for 10 min. The

E. coli cells (100 µl) were then mixed with DNA (routinely 50 ng) and held on ice for a further 30 min. TSB solution (0,9 ml) containing glucose (20 mM) was added to each transformation mixture and incubated at 37°C for 60 min, to allow expression of the plasmid borne antibiotic marker. Unused cells could be stored at -70°C after rapid freezing in a dry ice/ethanol bath or liquid nitrogen, and retained viability provided that the cells were thawed slowly on ice when needed.

C.7 Transduction of *E. coli* cells with the *T. ferrooxidans* pHC79 cosmid gene bank

The *T. ferrooxidans* pHC79 cosmid gene bank was stored as a phage lysate above chloroform at 4°C (R. Ramesar, PhD thesis, University of Cape Town, South Africa, 1988). Recipient *E. coli* cells were grown overnight in 25 mls LB medium containing 0,25% (w/v) maltose. A sample of 1 ml was sedimented by centrifugation in an Eppendorf microfuge, resuspended in 10 mM $MgSO_4$, and placed at room temperature for 1 h. A 10 µl sample of an appropriate dilution (routinely 10^{-7} or 10^{-6} in SM buffer; Appendix B.3.19) of the cosmid gene bank phage lysate was mixed with 200 µl of cells. Incubation was continued at 37°C for 30 min, and followed by the addition of 1 ml of LB and a further 60 min incubation at 37°C. Samples of 100 µl were plated onto the appropriate agar plates.

C.8 Nucleotide sequencing

C.8.1 Primer annealing reaction

The supercoiled DNA (6-10 µg, in TE buffer, pH 8,0) was diluted to a final volume of 20 µl in distilled water. Alkaline denaturation in 0,2 N NaOH (5 min at room temperature) was followed by the addition of 5 µl of 3 M sodium acetate (pH 5,2), 25 µl of distilled water and 150 µl of ethanol. This mixture was chilled to -70°C, centrifuged at 4°C for 20 min in a microfuge and washed with 200 µl of ethanol (70%). The DNA pellet was dried and resuspended in a final volume of 10 µl of sequencing buffer (40 mM Tris-HCl, pH 7,5; 20 mM $MgCl_2$; 50 mM NaCl) and 12 ng of primer. This mixture was annealed for 30 min at 40°C immediately prior to sequencing. The primers used were the forward sequencing primer as supplied in the "Sequenase" DNA sequencing kit (US Biochemical Corp., Cleveland, Ohio) and the M13 reverse sequencing primer (Amersham).

C.8.2 Sequencing reactions

DNA sequencing was done by the dideoxy chain termination method of Sanger *et al.* (1977) according to the protocol of Tabor and Richardson (1987), using T7 DNA polymerase and a "Sequenase" sequencing kit supplied by the US Biochemical Corporation, Cleveland, Ohio. The DNA chain was radiolabelled with [α-^{35}S]dATP (1200 Ci/mmol; Amersham).

C.8.3 Gel electrophoresis and autoradiography

The sequencing reactions were analyzed on standard 6% denaturing acrylamide urea sequencing gels. The composition and running conditions of the gels were as described in the Amersham M13 Sequencing Handbook. After electrophoresis the gels (0,2mm thick) were dried onto Whatman No. 3 filter paper using a Dual Temperature Slab Gel Dryer (Model 1125B; Hoefer Scientific Instruments, San Francisco). Gels containing ^{35}S-labelled DNA were placed under XAR-5 autoradiographic film and exposed for 1-2 d. The autoradiographs were developed using Kodak GBX X-ray developer and fixer.

C.9 Radioactive labelling of DNA probes

DNA probes were labelled with [α-^{32}P]dCTP to high specific activity by nick-translation (Rigby *et al.*, 1977). The reagents were obtained in kit form (Amersham) and used according to the suppliers specifications. Contaminating nucleotides were removed from the radioactively labelled probe preparation using a Sephadex G50 spin column as described by Maniatis *et al.* (1982). Radioactively labelled probes were stored in lead containers at -20°C. Probes were denatured by boiling (5 min) in a fume hood just before use.

C.10 DNA hybridization

DNA fragments resolved by agarose gel electophoresis were transferred to a Hybond-N+ hybridization membrane (Amersham) essentially by the protocol of Reed and Mann (1985). The use of a nylon transfer membrane allows the capillary transfer of DNA restriction fragments in alkali rather than in neutral, high ionic strength solvents (used in conventional Southern transfer), and eliminates the need for post-transfer fixation (Reed and Mann, 1985). After electrophoresis the gel was rinsed in two volumes of HCl (0,25 M) for 20 min at room temperature with gentle agitation, followed by a brief rinse in distilled water. The gel was then placed on top of two sheets of Whatman 3 MM filter paper (wetted with 0,4 N NaOH, and placed on top of an inverted gel-casting tray in a plastic box, such that the filter paper touched the base of the box, forming a wick), and was flooded with 50-100 ml of 0,4 N NaOH. A sheet of Hybond-N+, wetted by floating onto, and then immershed in, distilled water was placed on top of the gel, and any air bubbles were removed. Three sheets of Whatman 3 MM filter paper, wetted in 0,4 N NaOH, were laid onto the membrane, followed by a 4 cm thick layer of absorbent paper. A light weight was placed on top of this, and transfer left to continue overnight. After transfer, the membrane was removed and rinsed for 20 min in 2 x SSC (Appendix B.3.22).

Hybridization and washing conditions were essentially according to Maniatis *et al.* (1982). The membrane was gently shaken in pre-hybridization solution (Appendix B.3.13) for 4 h at 65°C, while the probe was being prepared. The radioactively-labelled probe to be used was denatured by boiling for 5 min and was added to the pre-hybridization fluid. Hybridization was carried out at 65°C overnight. The membrane was washed in decreasing concentrations of SSC (1x - 0,1x) at 65°C for 20 min each, and, after checking the radioactivity by means of a Geiger-counter, the washing was terminated and the membrane sealed in a plastic bag. The membrane was exposed to autoradiographic film (XAR-5) at -70°C.

APPENDIX D

MOLECULAR WEIGHT MARKER

Agarose gel electrophoresis

Phage Lambda DNA: *Pst*I restriction fragments

Fragment No.		size (base pairs)
1	(2 + 7, "cos" ends)	14057
2		11497
3		5077
4		4749
5		4507
6		2838
7		2560
8	(7, 8 = doublet)	2459
9		2443
10		2140
11		1986
12		1700
13		1159
14		1093
15		805
16		514
17	(16, 17 = doublet)	468
18		448
19		339
20		264
21	gel limit	247
22		216
23		211
24		200
25		164
26		150
27		94
28		87
29		72
30		15

APPENDIX E

ONE- AND THREE-LETTER CODES USED FOR AMINO ACIDS

Amino acid	Codes		Code	Amino acid
Alanine	Ala	A	A	Alanine
Arginine	Arg	R	C	Cysteine
Asparagine	Asn	N	D	Aspartic acid
Aspartic acid	Asp	D	E	Glutamic acid
Cysteine	Cys	C	F	Phenylalanine
Glutamine	Gln	Q	G	Glycine
Glutamic acid	Glu	E	H	Histidine
Glycine	Gly	G	I	Isoleucine
Histidine	His	H	K	Lysine
Isoleucine	Ile	I	L	Leucine
Leucine	Leu	L	M	Methionine
Lysine	Lys	K	N	Asparagine
Methionine	Met	M	P	Proline
Phenylalanine	Phe	F	Q	Glutamine
Proline	Pro	P	R	Arginine
Serine	Ser	S	S	Serine
Threonine	Thr	T	T	Threonine
Tryptophan	Trp	W	V	Valine
Tyrosine	Tyr	Y	W	Tryptophan
Valine	Val	V	Y	Tyrosine

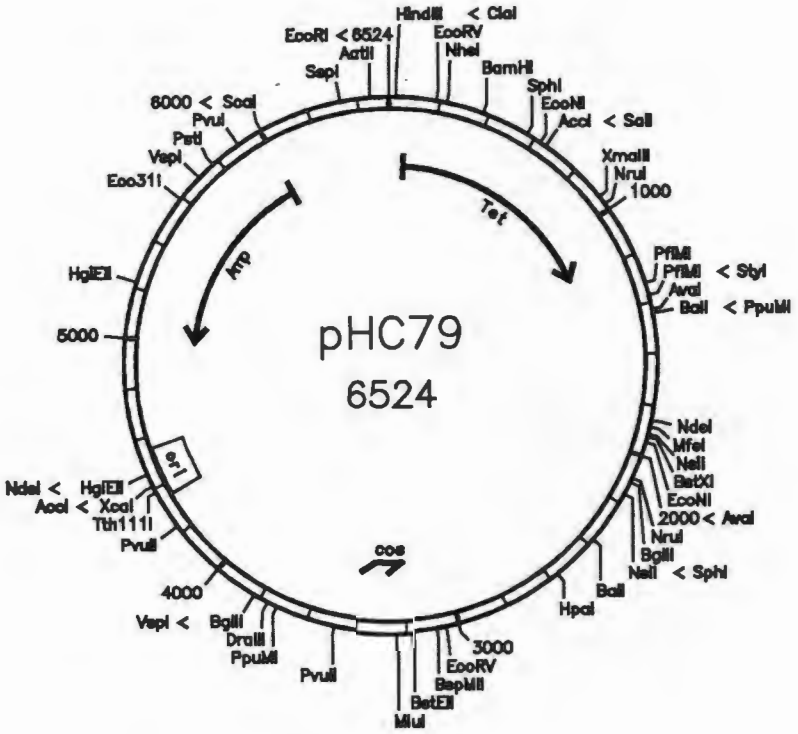
Groups of amino acids based on conservative substitutions
(used for amino acid sequence alignments)

i)	I, L, M, V	vi)	F, Y
ii)	D, E, N, Q	vii)	C
iii)	K, R, H	vi)	P
iv)	S, T	viii)	W
v)	G, A		

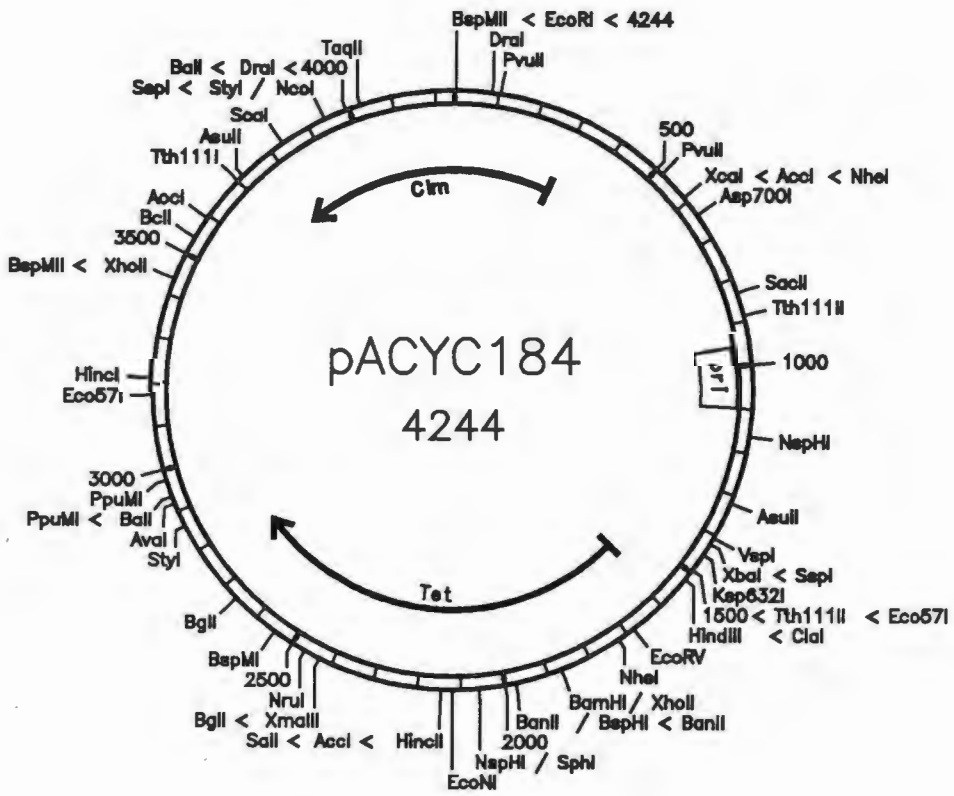
APPENDIX F

PLASMID VECTOR MAPS

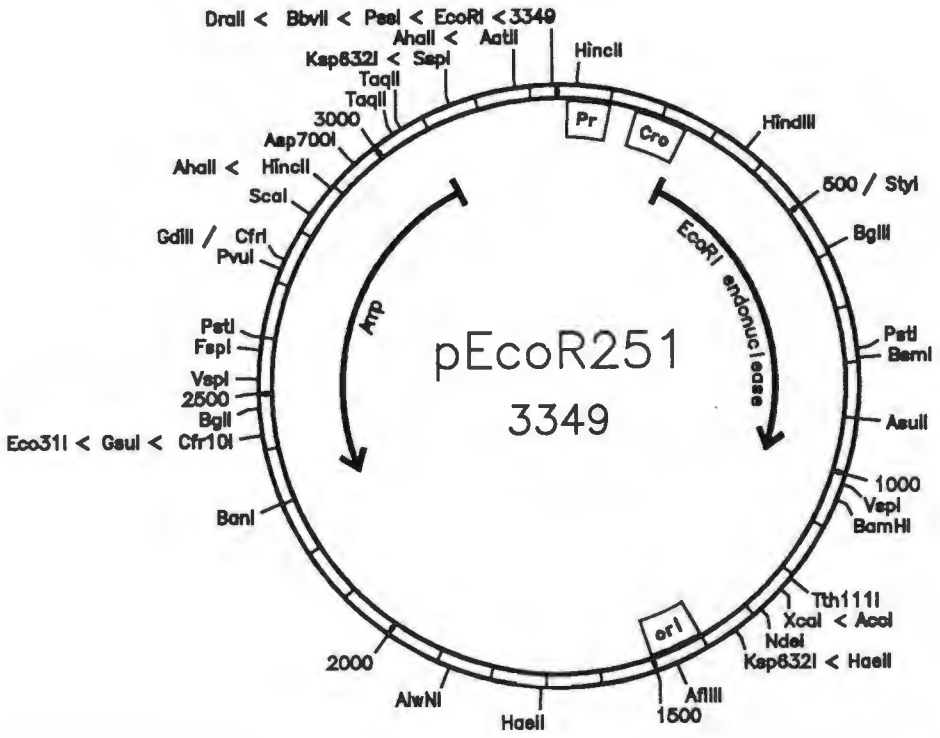
Maps show restriction enzymes (6 bp recognition) with a maximum of two sites.



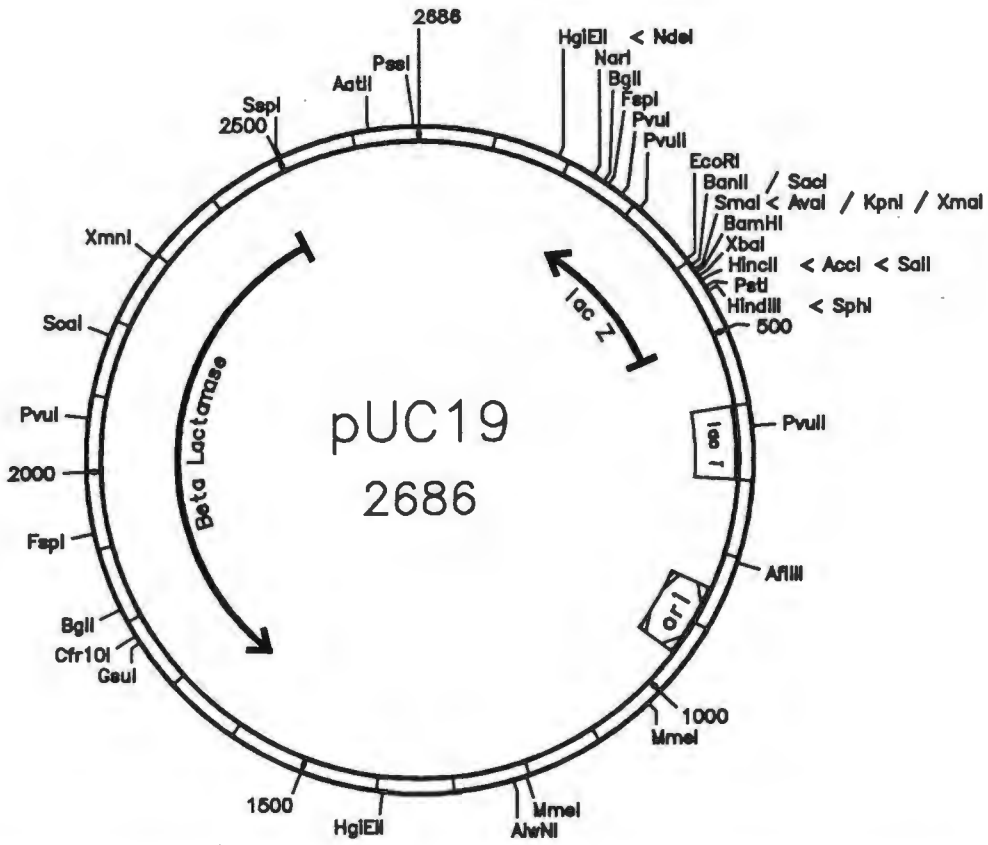
Hohn and Collins (1980)



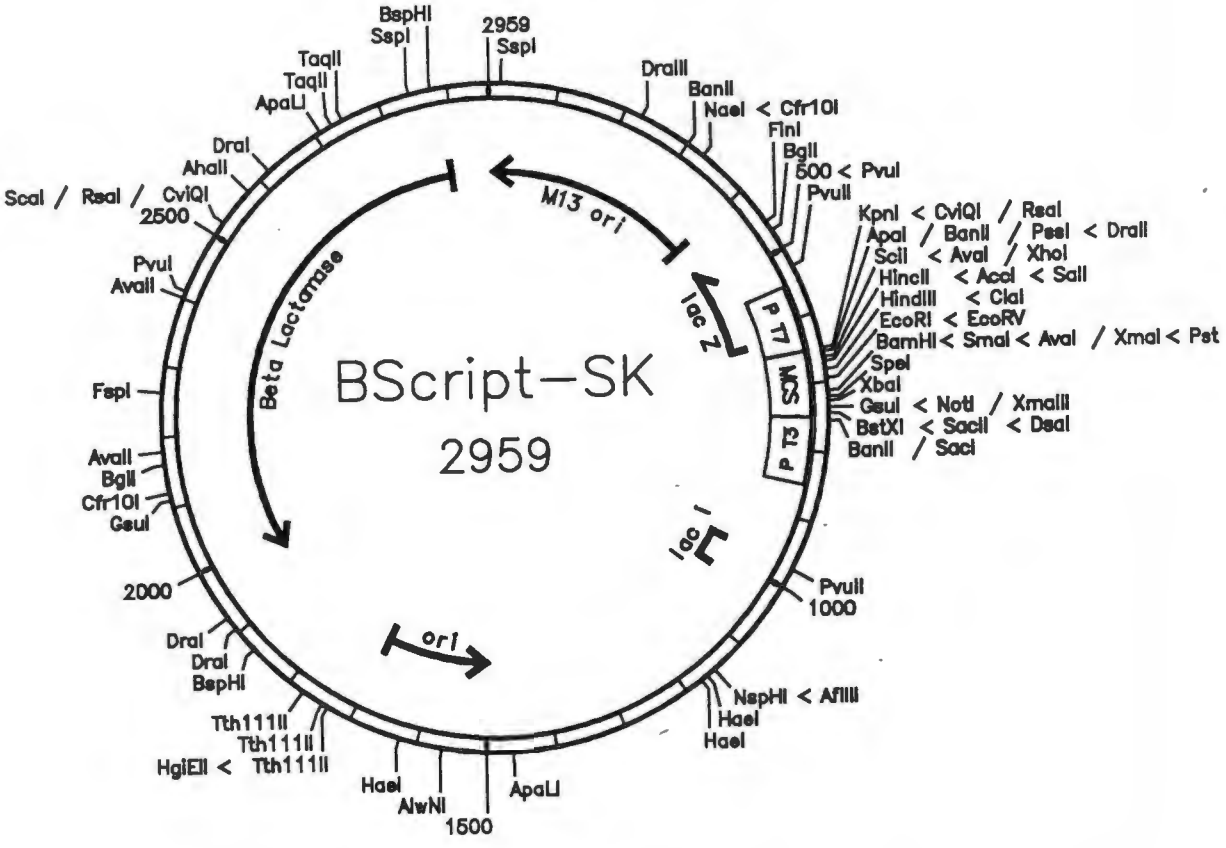
Chang and Cohen (1978)



pEcoR251 is a positive selection vector used routinely by insertional activation of the *Eco*RI endonuclease gene at the *Bgl*II or *Hind*III site.
Zabeau and Stanley (1982).



pUC18 differs from **pUC19** in the orientation of the multiple cloning site (MCS) polylinker between the *Eco*RI and *Hind*III sites.
The forward primer (-40) (5'-GTT TTC CCA GTC ACG AC-3') is situated on the *Eco*RI side of the MCS in **pUC19**. Norrander *et al.* (1983).



Map of **Bluescript SK+**. **Bluescript KS+** differs in the orientation of the MCS polylinker between the *Kpn*I and *Sac*I sites. The forward primer (-40) is situated on the *Kpn*I side of the MCS of **SK+**. Stratagene, San Diego.

REFERENCES

Albright LM, Huala E and Ausubel FM (1989a) Prokaryotic signal transduction mediated by sensor and regulator protein pairs. ***Annu. Rev. Genet.*** **23**:311-336

Albright LM, Ronson CW, Nixon BT and Ausubel FM (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 FJ, Chabert J, Willison JC and Vignais PM (1989) Nucleotide sequence of the wild-type and mutant *nifr4* (*ntrA*) genes of *Rhodobacter capsulatus* - Identification of an essential glycine residue. ***Nucleic Acids Res.*** **13**:5377

Almassy RJ, Janson CA, Hamlin R, Xuong N-H and Eisenberg D (1986) Novel subunit-subunit interactions in the structure of glutamine synthetase. ***Nature*** **323**:304-309

Alvarez-Morales A and Hennecke H (1985) Expression of *Rhizobium japonicum nifH* and *nifD* operons can be activated by the *Klebsiella pneumoniae nifA* protein but not by the product of *ntrC*. ***Mol. Gen. Genet.*** **199**:306-314

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

Alvarez-Morales A, Dixon R and Merrick M (1984) Positive and negative control of the *glnAntrBC* regulon in *Klebsiella pneumoniae*. ***EMBO J.*** **3**:501-507

Arnold W, Rump A, Klipp W, Priefer UB and Puhler A (1988) Nucleotide sequence of a 24,206-basepair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*, p. 303-304. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Arnosti DN and Chamberlin MJ (1989) Secondary σ factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*. ***Proc. Natl. Acad. Sci. USA.*** **86**:830-834

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*. ***Mol. Microbiol.*** **1**:92-100

Austin S, Henderson N and Dixon R (1990) Characterization of the *Klebsiella pneumoniae* nitrogen fixation regulatory proteins NifA and NifL *in vitro*. ***Eur. J. Biochem.*** **187**:353-360

Ausubel FM (1984) Regulation of nitrogen fixation genes. ***Cell*** **37**:5-6

Avtges P, Kranz RG and Haselkorn R (1985) Isolation and organization of genes for nitrogen fixation in *Rhodopseudomonas capsulata*. ***Mol. Gen. Genet.*** **201**:363-369

Backman K (1980) A cautionary note on the use of certain restriction endonucleases with methylated substrates. ***Gene*** **11**:169-171

Backman K, Chen YM and Magasanik B (1981) Physical and genetic characterization of the *glnA-glnG* region of the *Escherichia coli* chromosome. ***Proc. Natl. Acad. Sci. USA*** **78**:3743-3747

Backman KC, Chen Y-M, Ueno-Nishio S and Magasanik B (1983) The product of *glnL* is not essential for regulation of bacterial nitrogen fixation. ***J. Bacteriol.*** **154**:516-519

Bali A, Hill S, Santero E, Toukdarian A, Walmsley J and Kennedy C (1988) NifA and NtrC activate seperate nitrogen fixation pathways in *Azotobacter*, p. 316. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Barros ME, Rawlings DE and Woods DR (1985) Cloning and expression of the *Thiobacillus ferrooxidans* glutamine synthetase gene in *Escherichia coli*. ***J. Bacteriol.*** **164**:1386-1389

Barros ME, Rawlings DE and Woods DR (1986) Purification and regulation of a cloned *Thiobacillus ferrooxidans* glutamine synthetase. ***J. Gen. Microbiol.*** **132**:1989-1995

Batut J, Daveranmingot ML, Jacobs MDJ, Garnerone AM 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

Begg IA, Whyte JN and Haddock BA (1977) The identification of mutants of *Escherichia coli* deficient in formate dehydrogenase and nitrate reductase activities using dye indicator plates. ***FEMS Microbiol. Lett.*** **2**:47-50

Bennett LT, Cannon F and Dean DR (1988) Nucleotide sequence and mutagenesis of the *nifA* gene from *Azotobacter vinelandii*. ***Mol. Microbiol.*** **2**:315-321

Berger DK, Woods DR and Rawlings DE (1990) Complementation of *Escherichia coli* σ54(NtrA)-dependent formate hydrogenlyase activity by a cloned *Thiobacillus ferrooxidans ntrA* gene. ***J. Bacteriol.*** **172**(8):4399-4406

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 and Böck A (1989) Characterization of a *cis* regulatory DNA element necessary for formate induction of the formate dehydrogenase gene (*fdhF*) of *Escherichia coli*. ***Mol. Microbiol.*** **3**:187-195

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

Birkmann A, Zinoni F, Sawers G and Böck A (1987b) Factors affecting transcriptional regulation of the formate-hydrogen-lyase pathway of *Escherichia coli*. ***Arch. Microbiol.*** **148**:44-51

Birnboim HC and Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. ***Nucleic Acids Res.*** **7**:1513-1523

Bishop PE, Jarlenski DML and Hetherington DR (1980) Evidence for an alternative nitrogen fixation system in *Azotobacter vinelandii*. ***Proc. Natl. Acad. Sci. USA.*** **77**:7342-7346

Bishop PE and Joerger RD (1990) Genetics and molecular biology of alternative nitrogen fixation systems. ***Annu. Rev. Plant Phys. Plant Mol. Biol.*** **41**:109-125

Bishop PE, Premakumar R, Dean DR, Jacobson MR, Chisnell JR, Rizzo TM and Kopczynski J (1986) Nitrogen fixation by *Azotobacter vinelandii* strains having deletions in structural genes for nitrogenase. ***Science*** **232**:92-94

Bourret RB, Hess JF, Borkovich KA, Pakula AA and Simon MI (1989) Protein phosphorylation in chemotaxis and two-component regulatory systems of bacteria. ***J. Biol. Chem.*** **264**:7085-7088

Bozouklian H and Elmerich C (1986) Nucleotide sequence of the *Azospirillum brasiliense* Sp7 glutamine synthetase structural gene. ***Biochimie*** **68**:1181-1187

Brierley CL (1982) Microbiological mining. ***Sci. Am.*** **247**:42-51

Brierley JA and Le Roux NW (1977) A facultative thermophilic *Thiobacillus*-like bacterium: oxidation of iron and pyrite, p. 55-66. In: *Conference on bacterial leaching*, W Scwartz (ed.). Verlag Chemie, New York

Brigle KE, Newton WE and Dean DR (1985) Complete nucleotide sequence of the *Azotobacter vinelandii* nitrogenase structural gene cluster. ***Gene*** **37**:37-44

Brooks SJ, Collins JJ and Brill WJ (1984) Repression of nitrogen fixation in *Klebsiella pneumoniae* at high temperature. ***J. Bacteriol.*** **157**:460-464

Buchanan-Wollaston V, Cannon FC, Beynon JL and Cannon FC (1981a) Role of the *nifA* gene product in the regulation of *nif* expression in *Klebsiella pneumoniae*. ***Nature*** **294**:776-778

Buchanan-Wollaston V, Cannon MC and Cannon FC (1981b) The use of cloned *nif*(nitrogen fixation) DNA to investigate transcriptional regulation of *nif* expression in *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **184**:102-106

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.*** **17**:2597-2612

Buck M, Cannon W and Woodcock J (1987a) Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve 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

Buck M, Woodcock F, Cannon W, Mitchenall L and Drummond M (1987b) Positional requirements for the function of *nif*-specific upstream activator sequences. ***Mol. Gen. Genet.*** **210**:140-144

Bueno R, Pahel G and Magasanik B (1985) Role of *glnB* and *glnD* gene products in regulation of the *glnALG* operon of *Escherichia coli*. ***J. Bacteriol.*** **164**:816-822

Buikema WJ, Szeto WW, Lenley PV, Orme-Johnson WH and Ausubel FM (1985) Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulation regulatory gene *ntrC* of *Klebsiella pneumoniae*. ***Nucleic Acids Res.*** **13**:4539-4555

Burns RC and Hardy RWF (1975) *Nitrogen fixation in bacteria and higher plants*, p. 14-38. Springer Verlag, New York

Burris RH (1988) 100 years of discoveries in biological N_2 fixation, p. 21-30. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Buurman ET, Teixeira de Mattos MJ and Neijssel OM (1989) Nitrogen-limited behaviour of micro-organisms growing in the presence of large concentrations of ammonium ions. ***FEMS Microbiol. Lett.*** **58**:229-232

Böhm R, Sauter M and Böck A (1990) Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. ***Mol. Microbiol.*** 4:231-243

Cali BM, Micca JL and Stewart V (1989) Genetic regulation of nitrate assimilation in *Klebsiella pneumoniae* M5A1. ***J. Bacteriol.*** **171**:2666-2672

Cannon FC, Dixon RA, Postgate JR and Primrose SB (1974) Chromosomal integration of *Klebsiella* nitrogen fixation genes in *Escherichia coli*. ***J. Gen. Microbiol.*** **80**:227-239

Cannon M, Cannon F, Buchanan-Wollaston V, Ally D, Ally A and Beynon J (1988) The nucleotide sequence of the *nifJ* gene of *Klebsiella pneumoniae*. ***Nucleic Acids Res.*** **16**:11379

Cannon WV, Kreutzer R, Kent HM, 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

Carlson TA and Chelm BK (1986) Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. ***Nature*** **322**:568-570

Carlson TA, Guerinot ML and Chelm BK (1985) Characterization of the gene encoding glutamine synthetase I (*glnA*) from *Bradyrhizobium japonicum*. ***J. Bacteriol.*** **162**:698-703

Casadaban MJ, Martinez-Arias A, Shapira SK and Chou J (1983) β-galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. ***Meth. Enzym.*** **100**:293-308

Castano I and Bastarachea F (1984) *glnF-lacZ* fusions in *Escherichia coli*: studies on *glnF* expression and its chromosomal orientation. ***Mol. Gen. Genet.*** **195**:228-233

Chang ACY and Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. ***J. Bacteriol.*** **134**:1141-1166

Chung CT and Miller RH (1988) A rapid and convenient method for preparation and storage of competent bacterial cells. ***Nucleic Acids Res.*** **16**:3580

Chase JW and Williams KR (1986) Single-stranded DNA binding proteins required for DNA replication. ***Annu. Rev. Biochem.*** **55**:103-136

Chen Y -M, Backman K and Magasanik B (1982) Characterization of a gene, *glnL*, the product of which is involved in the regulation of nitrogen utilization in *Escherichia coli*. ***J. Bacteriol.*** **150**:214-220

Chisnell JR, Premakumar R and Bishop PE (1988) Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii*. ***J. Bacteriol.*** **170**:27-33

Chock PB, Schacter E, Jurgensen SR and Rhee SG (1985) Cyclic cascade systems in metabolic regulation. ***Curr. Top. Cell. Regul.*** **27**:3-11

Chou PY and Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. ***Adv. Enzymol.*** **47**:45-147

Collins JJ and Brill WJ (1985) Control of *Klebsiella pneumoniae nif* mRNA synthesis. ***J. Bacteriol.*** **162**:1186-1190

Collins JJ, Roberts GP and Brill WJ (1986) Post-transcriptional control of *Klebsiella pneumoniae nif* mRNA stability by the *nifL* product. ***J. Bacteriol.*** **168**:173-178

Colmer AR and Hinkle ME (1947) The role of microorganisms in acid mine drainage: a preliminary report. ***Science*** **106**:253

Colonna-Romano S, Riccio A, Guida M, Defez R, Lamberti A, Iaccarino M, Arnold W, Preifer U and Puhler A (1987) Tight linkage of *glnA* and a putative regulatory gene in *Rhizobium leguminosarum*. ***Nucleic Acids Res.*** **15**:1951-1964

Contreras A and Drummond M (1988) The effect on the function of the transcriptional activator NtrC from *Klebsiella pneumoniae* of mutations in the DNA-recognition helix. ***Nucleic Acids Res.*** **16**:4025-4039

Courey AJ and Tjian R (1988) Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel, glutamine-rich activation motif. ***Cell*** **55**:887-898

Covarrubias AA and Bastarrachea F (1983) Nucleotide sequence of the *glnA* control region of *Escherichia coli*. ***Mol. Gen. Genet.*** **190**:171-175

Csonka LN (1989) Physiological and genetic responses of bacteria to osmotic stress. ***Microbiol. Rev.*** **53**:121-147

Curtin ME (1983) Microbial mining and metal recovery corporations take the long and cautious path. ***Biotechnol.*** **1**:229-238

Darrow RA, Crist D, Evans WR, Jones BL, Keister DL and Knotts RR (1981) Biochemical and physiological studies on the two glutamine synthetases of *Rhizobium*, p. 182-185. In: *Current perspectives in nitrogen fixation*, AH Gibson and WE Newton (ed.). Australian Acad. Sci., Canberra

Darst SA, Kubalek EW and Kornberg RD (1989) Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. ***Nature*** **340**:730-732

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

Davidson MS and Summers AO (1983) Wide-host-range plasmids function in the genus *Thiobacillus*. ***Appl. Environ. Microbiol.*** **46**:565-572

Dayhoff MO, Hunt LT and Hurst-Calderone S (1978) In: *Atlas of Protein Sequence and Structure*, Vol 5, p. 363-369, MO Dayhoff (ed.).

De Bruijn FJ and Ausubel FM (1981) The cloning and transposon Tn5 mutagenesis of the *glnA* region of *Klebsiella pneumoniae*: identification of *glnR*, a gene involved in the regulation of the *nif* and *hut* operons. ***Mol. Gen. Genet.*** **183**:289-297

De Bruijn FJ and Ausubel FM (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 FJ, Rossbach S, Schneider M, Ratet P, Messmer S, Szeto WW, Ausubel FM and Schell J (1989) *Rhizobium meliloti* 1021 has 3 differentially regulated loci involved in glutamine biosynthesis, none of which is essential for symbiotic nitrogen fixation. ***J. Bacteriol.*** **171**:1673-1682

Deretic V, Dikshit R, Konyecsni WM, Chakrabarty AM and Misra TK (1989) The *algR* gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. ***J. Bacteriol.*** **171**:1278-1283

Devereux J (1984) A comprehensive set of sequence analysis programs for the VAX. ***Nucleic Acids Res.*** **12**:387-395

Dimri GP and Das HK (1988) Transcriptional regulation of nitrogen fixation genes by DNA supercoiling. ***Mol. Gen. Genet.*** **212**:360-363

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 (1986) The *xylABC* promoter from the *Pseudomonas putida* TOL plasmid is activated by nitrogen regulatory genes in *Escherichia coli*. ***Mol. Gen. Genet.*** **203**:129-136

Dixon R (1988) Genetic regulation of nitrogen fixation, p. 417-438. In: *The Nitrogen and Sulphur Cycles*, JA Cole and SJ Ferguson (ed.). Cambridge University Press, Cambridge, New York

Dixon R, Cannon F and Kondorosi A (1976) Construction of a P plasmid carrying nitrogen fixation genes from *Klebsiella pneumoniae*. ***Nature*** **260**:268-271

Dixon RA and Postgate JR (1972) Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. ***Nature*** **237**:102-103

Dixon RA, Henderson NC and Austin S (1988) DNA supercoiling and aerobic regulation of transcription from the *Klebsiella pneumoniae nifLA* promoter. ***Nucleic Acids Res.*** **16**:9933-9946

Dodd IB and Egan JB (1987) Systematic method for the detection of potential λ Cro-like DNA-binding regions in proteins. ***J. Mol. Biol.*** **194**:557-564

Doi RH and Wang L-H (1986) Multiple procaryotic ribonucleic acid polymerase sigma factors. ***Microbiol. Rev.*** **50**:227-243

Dorrington RA and Rawlings DE (1989) Identification and sequence of the basic replication region of a broad-host-range plasmid isolated from *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **171**:2735-2739

Drozd J and Postgate JR (1970) Interference by oxygen in the acetylene-reduction test for aerobic nitrogen-fixing bacteria. ***J. Gen. Microbiol.*** **60**:427-429

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

Drummond M and Wootton J (1987) Sequence of *nifL* from *Klebsiella pneumoniae*: mode of action and relationship to two families of regulatory proteins. ***Mol. Microbiol.*** **1**:37-44

Drummond MH (1984) The nitrogen fixation genes of *Klebsiella pneumoniae*: a model system. ***Microbiol. Sci.*** **1**:29-32

Drummond MH, Clements J, Merrick M and Dixon R (1983) Positive control and autogenous regulation of the *nifLA* promoter in *Klebsiella pneumoniae*. ***Nature*** **301**:302-307

Elmerich C, Galimand M, Vieille C, Delorme F and De Zamaroczy M (1988) Nitrogen fixation genes of *Azospirillum*, p. 327-331. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Espin G, Alvarez-Morales A, Cannon F, Dixon R and Merrick M (1982) Cloning of the *glnA*, *ntrB*, and *ntrC* genes of *Klebsiella pneumoniae* and studies of their role in the regulation of the nitrogen fixation (*nif*) gene cluster. ***Mol. Gen. Genet.*** **186**:518-524

Ficket JW (1982) Recognition of protein coding regions in DNA sequences. ***Nucleic Acids Res.*** **10**:5303-5318

Filser M, Merrick M and Cannon F (1983) Cloning and characterization of *nifLA* regulatory mutations from *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **191**:485-491

Filser MMK, Moscatelli CLA, Vineze E, Guida M, Salzano G and Iaccarino M (1986) Characterization and cloning of two *Rhizobium leguminosarum* genes coding for glutamine synthetase activities. ***J. Gen. Microbiol.*** **132**:2561-2569

Fischer H-M and Hennecke H (1987) Direct response of *Bradyrhizobium japonicum* NifA-mediated *nif* gene regulation to cellular oxygen status. ***Mol. Gen. Genet.*** **209**:621-626

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 reactivity and/or metal binding. ***Nucleic Acids Res.*** **16**:2207-2224

Fitch WM and Margoliash E (1967) Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome C sequences is of general applicability. ***Science*** **155**:279-284

Forst S and Inouye M (1988) Environmentally regulated gene expression for membrane proteins in *Escherichia coli*. ***Annu. Rev. Cell Biol.*** **4**:21-42

Friedrich CG and Friedrich B (1983) Regulation of hydrogenase formation is temperature sensitive and plasmid coded in *Alcaligenes eutrophus*. ***J. Bacteriol.*** **153**:176-181

Garcia E, Bancroft S, Rhee SG and Kustu S (1977) The product of a newly identified gene *glnF*, is required for synthesis of glutamine synthetase in *Salmonella*. ***Proc. Natl. Acad. Sci. USA.*** **74**:1662-1666

Garcia E and Rhee SG (1983) Cascade control of *Escherichia coli* glutamine synthetase: purification and properties of P_{II} uridylyltransferase and uridylyl-removing enzyme. ***J. Biol. Chem.*** **258**:2246-2253

Garciarrubio AA and Covarrubias AA (1987) Promoter selection by a bacterial enhancer-like activator element (BELE) in *Escherichia coli*. ***Gene*** **54**:275-280

Gardella T, Moyle H and Susskind MM (1989) A mutant *Escherichia coli* σ70 subunit of RNA polymerase with altered promoter specificity. ***J. Mol. Biol.*** **206**:579-590

Golden JW, Mulligan ME and Haselkorn R (1987) Different recombination site specificity of two developmentally regulated genome rearrangements. ***Nature*** **327**:526-529

Golden JW, Robinson SJ and Haselkorn R (1985) Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. ***Nature*** **314**:419-423

Golovacheva RS and Karavaiko GI (1977) A new facultative thermophilic *Thiobacillus* isolated from sulphide ore, p. 108-109. In: *Microbial growth on C1-compounds*. Pushchino, USSR Acad. Sci.

Gosink MM, Franklin NM and Roberts GP (1990) The product of the *Klebsiella pneumoniae nifX* gene is a negative regulator of the nitrogen fixation (Nif) regulon. ***J. Bacteriol.*** **172**:1441-1447

Grantham R, Gautier C, Guoy M, Jacobzone M and Mercier R (1981) Codon catalogue usage is a genome strategy modulated for gene expressivity. ***Nucleic Acids Res.*** **9**:43-74

Gribskov M and Burgess RR (1986) Sigma factors from *Escherichia coli*, *Bacillus subtilis*, phage *SPO1*, and phage *T4* are homologous proteins. ***Nucleic Acids Res.*** **14**:6745-6763

Grosjean H and Fiers W (1982) Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. ***Gene*** **18**:199-209

Grossman AD, Erickson JE and Gross CA (1984) The *hptR* gene product of *Escherichia coli* is a sigma factor for heat shock promoters. ***Cell*** **38**:383-390

Gussin GN, Ronson CW and Ausubel FM (1986) Regulation of nitrogen fixation genes. ***Annu. Rev. Genet.*** **20**:567-591

Haaker H, Wassink H, Mensink R and Veeger C (1988) Regulation of whole cell nitrogenase activity in *Azotobacter vinelandii*, p. 243-248. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Hales BJ, Case EE, Morningstar JE, Dzeda MF and Mauterer LA (1986) Isolation of a new V-containing nitrogenase from *Azotobacter vinelandii*. ***Biochem.*** **25**:7251-7255

Harrison AP (1984) The acidophilic thiobacilli and other acidophilic bacteria that share their habitat. ***Annu. Rev. Microbiol.*** **38**:265-292

Haselkorn R (1986) Organization of the genes for nitrogen fixation in photosynthetic bacteria and cyanobacteria. ***Ann. Rev. Microbiol.*** **40**:525-547

Hawkes T, Merrick M and Dixon R (1985) Interaction of purified NtrC protein with nitrogen regulated promoters from *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **201**:492-498

Hawkins FK and Johnston AW (1988) Transcription of a *Rhizobium leguminosarum* biovar *phaseoli* gene needed for melanin synthesis is activated by NifA of *Rhizobium* and *Klebsiella pneumoniae*. ***Mol. Microbiol.*** **2**:331-337

Helmann J and Chamberlin MJ (1988) Structure and function of bacterial sigma factors. ***Annu. Rev. Biochem.*** **57**:839-872

Helmann JD, Marquez LM and Chamberlin MJ (1988) Cloning, sequencing, and disruption of the *Bacillus subtilis* σ28 gene. ***J. Bacteriol.*** **170**:1568-1574

Henderson N, Austin S and Dixon RA (1989) Role of metal ions in negative regulation of nitrogen fixation by the *nifL* gene product from *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **216**:484-491

Henderson PJF (1971) Ion transport by energy-conserving biological membranes. ***Annu. Rev. Microbiol.*** **25**:393-428

Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. ***Gene*** **28**:351-359

Henikoff S (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis, p. 156-165. In: *Methods in Enzymology*, Vol. 155, R Wu (ed.). Academic Press, New York, London

Hennecke H, Fischer H-M, Gubler M, Thöny B, Anthamatten D, Kullik I, Ebeling S, Fritsche S and Zurchler T (1988) Regulation of *nif* and *fix* genes in *Bradyrhizobium japonicum* occurs by a cascade of two consecutive gene activation steps of which the second one is oxygen, p. 339-344. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Hennecke H, Kaluza K, Thöny B, Fuhrmann M, Ludwig W and Stackebrandt E (1985) Concurrent evolution of nitrogenase genes and 16s rRNA in *Rhizobium* species and other nitrogen fixing bacteria. ***Arch. Microbiol.*** **142**:342-348

Hertig C, Li RY, Louarn AM, Garnerone AM, 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

Higgins CF and Ferro-Luzzi Ames G (1982) Regulatory regions of two transport operons under nitrogen control: nucleotide sequence. ***Proc. Natl. Acad. Sci. USA.*** **79**:1083-1087

Higgins CF, Hiles ID, Salmond GPC, Gill DR, Downie JA, Evans IJ, Holland IB, Gray L, Buckel SD, Bell AW and Hermodson MA (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. ***Nature*** **323**:448-450

Hill S, Kennedy C, Kavanagh E, Goldberg RB and Hanau R (1981) Nitrogen fixation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *Klebsiella pneumoniae*. ***Nature*** **290**:424-426

Hirschman J, Wong P-K, Sei K, Keener J and 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

Hohn B and Collins J (1980) A small cosmid for efficient cloning of large DNA fragments. ***Gene*** **11**:291-298

Holmes DS, Lobos JH, Bopp LH and Welch GC (1983) Cloning of a *Thiobacillus ferrooxidans* plasmid in *Escherichia coli*. ***J. Bacteriol.*** **157**:324-326

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

Hoover TR, Santero E, Porter S and Kustu S (1990) The integration host factor (IHF) stimulates interaction of RNA polymerase with NifA, the transcriptional activator for nitrogen fixation operons. ***Cell*** **6**: in press

Hope IA and Struhl K (1986) Functional dissection of a eukaryotic transcriptional activator protein GCN4 of yeast. ***Cell*** **46**:885-894

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

Hudson GS and Davidson BE (1984) Nucleotide sequence and transcription of the phenylalanine and tyrosine operons of *Escherichia coli* K12. ***J. Mol. Biol.*** **180**:1023-1051

Hunt TP 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

Igo MM, Ninfa AJ, Stock JB and Silhavy TJ (1989) Phosphorylation and dephosphorylation of a bacterial activator by a transmembrane receptor. ***Genes Dev.*** **3**:1725-1734

Ingledew WJ (1982) *Thiobacillus ferrooxidans* The bioenergetics of an acidophilic chemolithotroph. ***Biochim. Biophys. Acta.*** **683**:89-117

Inoue C, Sugawara K, Shiratori T, Kusano T and Kitagawa Y (1989) Nucleotide sequence of the *Thiobacillus ferrooxidans* chromosomal gene encoding mercuric reductase. ***Gene*** **84**:47-54

Inouye S, Ebina Y, Nakazawa A and Nakazawa T (1984) Nucleotide sequence surrounding transcription initiation site of *xylABC* operon on TOL plasmid of *Pseudomonas putida*. ***Proc. Natl. Acad. Sci. USA.*** **81**:1688-1691

Inouye S, Kimoto M, Nakazawa A and Nakazawa T (1990) Presence of flagella in *Pseudomonas putida* is dependent on the *ntrA* (*rpoN*) gene. ***Mol. Gen. Genet.*** **221**:295-298

Inouye S, Nakazawa A and Nakazawa T (1987) Expression of the regulatory gene *xylS* on the TOL plasmid is positively controlled by the *xylR* gene product.
Proc. Natl. Acad. Sci. USA. **84**:5182-5186

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

Ish-Horowicz D and Burke JF (1981) Rapid and efficient cosmid cloning.
Nucleic Acids Res. **9**:2989-2998

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

Jayakumar A, Schulman I, MacNeil D and Barnes EMJr (1986) Role of the *Escherichia coli glnALG* operon in regulation of ammonium transport. ***J. Bacteriol.*** **166**:281-284

Joerger RD, Jacobson MR and Bishop PE (1989) Two *nifA*-like genes required for expression of alternative nitrogenases by *Azotobacter vinelandii*. ***J. Bacteriol.*** **171**:3258-3267

Johnson K, Parker ML and Lory S (1986) Nucleotide sequence and transcriptional initiation site of two *Pseudomonas aeruginosa* pilin genes. ***J. Biol. Chem.*** **261**:15703-15708

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

Jones R, Woodley P and Robson R (1984) Cloning and organization of some genes for nitrogen fixation from *Azotobacter chroococcum* and their expression in *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **197**:318-327

Keener J and Kustu S (1988) Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. ***Proc. Natl. Acad. Sci. USA.*** **85**:4976-4980

Keener J, Wong P, Popham D, Wallis J and Kustu S (1987) A sigma factor and auxiliary proteins required for nitrogen-regulated transcription in enteric bacteria, p. 159-175. In: *RNA polymerase and the regulation of transcription*, WS Reznikoff *et al.* (ed.). Elsevier Science Publishing Co., New York

Kelly DP, Morris PR and Brierley CL (1979) Microbial Technology: Current state, future prospects, p. 263-308. In: *Microbiological methods for the extraction and recovery of metals*, AT Bull *et al.* (ed.). Cambridge University Press, Cambridge

Kelly M (1969) Comparisons and cross reactions of nitrogenase from *Klebsiella pneumoniae*, *Azotobacter chroococcum* and *Bacillus polymyxa*. ***Biochim. Biophys. Acta.*** **191**:527-540

Kennedy C and Drummond MH (1985) The use of cloned *nif* regulatory elements from *Klebsiella pneumoniae* to examine *nif* regulation in *Azotobacter vinelandii*. ***J. Gen. Microbiol.*** **131**:1787-1795

Khamis MI, Casas-Finet JR, Maki AH, Murphy JB and Chase J (1987) Investigation of the role of individual tryptophan residues in the binding of *Escherichia coli* single-stranded DNA binding protein to single-stranded polynucleotides. ***J. Biol. Chem.*** **262**:10938-10945

Kleiner D (1981) The transport of NH_3 and NH_4^+ across biological membranes. ***Biochim. Biophys. Acta*** **639**:41-52

Klipp W, Masepohl B and Puhler A (1988) Identification and mapping of nitrogen fixation genes of *Rhodobacter capsulatus*: duplication of a *nifA-nifB* region. ***J. Bacteriol.*** **170**:693-699

Kofoid EC and Parkinson JS (1988) Transmitter and receiver modules in bacterial signalling proteins. ***Proc. Natl. Acad. Sci. USA.*** **85**:4981-4985

Köhler T, Cayrol JM, Ramos JL and Harayama S (1989a) Nucleotide and deduced amino acid sequence of the *rpoN* sigma factor of *Pseudomonas putida*. ***Nucleic Acids Res.*** **17**:10125

Köhler T, Harayama S, Ramos JL and Timmis KN (1989b) Involvement of *Pseudomonas putida RpoN* sigma factor in regulation of various metabolic functions. ***J. Bacteriol.*** **171**:4326-4333

Kolot MN, Kashlev MV, Gragerov AI and Khmel IA (1989) Stability of the pBR322 plasmid is affected by the promoter region of the tetracycline-resistance gene. ***Gene*** **75**:335-339

Kondorosi A, Kondorosi E, Gyorgypal Z, Banfalvi Z, Gyuris J, Putnoky P, Grosskopf E, John M, Schmidt J, Cam Ha DT, Lados M, Slaska-Kiss K and Schell J (1988) *Rhizobium meliloti nod* and *fix* genes controlling the initiation and development of root nodules, p. 399-403. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Kranz RG and Haselkorn R (1985) Characterization of the *nif* regulatory genes in *Rhodopseudomonas capsulata* using *lac* gene fusions. ***Gene* 40**:203-215

Kranz RG and Haselkorn R (1986) Anaerobic regulation of nitrogen fixation genes in *Rhodopseudomonas capsulata*. ***Proc. Natl. Acad. Sci. USA*. 83**:6805-6809

Kranz RG, Pace VM and Caldicott IM (1990) Inactivation sequence and *lacZ* fusion analysis of a regulatory locus required for repression of nitrogen fixation genes. ***J. Bacteriol.* 172**:53-62

Kreutzer R, Singh M and Klingmüller W (1989) Identification and characterization of the *nifH* and *nifJ* promoter regions located on the Nif plasmid Pea3 of *Enterobacter Agglomerans* 333. ***Gene* 78**:101-109

Kubo M and Imanaka T (1989) mRNA secondary structure in an open reading frame reduces translation efficiency in *Bacillus subtilis*. ***J. Bacteriol.* 171**:4080-4082

Kulpa CF, Roskey MT and Mjoli N (1986) Construction of chromosomal gene banks of *Thiobacillus ferrooxidans* and induction of the iron oxidation pathway. ***Biotech. Appl. Biochem.* 8**:330-341

Kulpa CF, Roskey MT and Travis MT (1983) Transfer of plasmid RP1 into chemolithotrophic *Thiobacillus neapolitanus*. ***J. Bacteriol.* 156**:434-436

Kustu S, Santero E, Keener J, Popham D and Weiss D (1989) Expression of σ54(*ntrA*)-dependent genes is probably united by a common mechanism. ***Microbiol. Rev.* 53**:367-376

Kustu S, Sei K and Keener J (1986) Nitrogen regulation in enteric bacteria, p. 139-154. In: *Regulation of gene expression*, I Booth and C Higgins (ed.). Symp. Soc. Gen. Microbiol., Cambridge University Press, Cambridge

Kustu SG, McFarland NC, Hui SP, Esmon B and Ferro-Luzzi Ames G (1979) Nitrogen control in *Salmonella typhimurium*: co-regulation of synthesis of glutamine synthetase and amino acid transport systems. ***J. Bacteriol.* 138**:218-234

Livesey-Goldblatt E, Norman P and Livesey-Goldblatt D (1983) Gold recovery from arsenopyrite ore by bacterial leaching and cyanidation, p. 627-641. In: *Recent progress in biohydrometallurgy*, G Rossi and AE Torma (ed.). Associazione Mineraria Sarda, Cagliari

Long SR (1989) *Rhizobium* genetics. ***Annu. Rev. Genet.* 23**:483-506

Losick R, Youngman P and Piggot PJ (1986) Genetics of endospore formation in *Bacillus subtilis*. ***Annu. Rev. Genet.* 20**:625-670

Lutz S, Böhm R, Beier A and Böck A (1990) Characterization of divergent σ54-dependent promoters in the anaerobically expressed *hyd-17* gene cluster of *Escherichia coli*. ***Mol. Microbiol.* 4**:13-20

MacNeil T, MacNeil D and Tyler B (1982a) Fine-structure deletion map and complementation analysis of the *glnA-glnL-glnG* region in *Escherichia coli*. ***J. Bacteriol.* 150**:1302-1313

MacNeil T, MacNeil D, Roberts GP, Supiano MA and Brill WJ (1978) Fine-structure mapping and complementation analysis of *nif* (nitrogen fixation) genes in *Klebsiella pneumoniae*. ***J. Bacteriol.*** **136**:253-266

MacNeil T, Roberts GP, MacNeil D and Tyler B (1982b) The products of *glnL* and *glnG* are bifunctional regulatory proteins. ***Mol. Gen. Genet.*** **188**:325-333

Macfarlane SA and Merrick M (1985) The nucleotide sequence of the nitrogen regulation gene *ntrB* and the *glnA-ntrBC* intergenic region of *Klebsiella pneumoniae*. ***Nucleic Acids Res.*** **13**:7591-7606

Macfarlane SA and Merrick M (1987) Analysis of the *Klebsiella pneumoniae ntrB* gene by site-directed *in vitro* mutagenesis. ***Mol. Microbiol.*** **1**:133-142

Mackintosh ME (1971) Nitrogen fixation by *Thiobacillus ferrooxidans* species. ***J. Gen. Microbiol.*** **66**:i-ii

Mackintosh ME (1978) Nitrogen fixation by *Thiobacillus ferrooxidans*. ***J. Gen. Microbiol.*** **105**:215-218

Magasanik B (1982) Genetic control of nitrogen assimilation in bacteria. ***Annu. Rev. Genet.*** **16**:135-168

Magasanik B (1988) Reversible phosphorylation of an enhancer binding protein regulates the transcription of bacterial nitrogen utilization genes. ***TIBS*** **13**:475-479

Magasanik B and Neidhardt FC (1987) Regulation of carbon and nitrogen utilization, p 1318-1325. In: Escherichia coli *and* Salmonella typhimurium *cellular and molecular biology*, FC Neidhardt (ed.). Amer. Soc. Microbiol., Washington DC

Mandrand-Berthelot MA, Wee MYK and Haddock BA (1978) An improved method for the identification and characterization of mutants of *Escherichia coli* deficient in formate dehydrogenase activity. ***FEMS Microbiol. Lett.*** **4**:37-40

Maniatis T, Fritsch EF and Sambrook J (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Marinus MG (1973) Location of DNA methylation genes on the *Escherichia coli* K-12 genetic map. ***Mol. Gen. Genet.*** **12**:47-55

Martin GB, Chapman KA and Chelm BK (1988) Role of the *Bradyrhizobium japonicum ntrC* gene product in differential regulation of the glutamine synthetase II gene (*glnII*). ***J. Bacteriol.*** **170**:5452-5459

Martin GB, Thomashow MF and Chelm BK (1989) *Bradyrhizobium japonicum glnB*, a putative nitrogen regulatory gene, is regulated by NtrC at tandem promoters. ***J. Bacteriol.*** **171**:5638-5645

Masephol 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

Matin A (1978) Organic nutrition of chemolithotrophic bacteria. ***Annu. Rev. Microbiol.*** **32**:433-468

McClarin JA, Frederick CA, Wang BC, Greene P and Boyer HW (1986) Structure of the DNA-*Eco*RI endonuclease recognition complex at 3 Å resolution. ***Science*** **234**:1526-1541

McPherson A, Jurnak FA, Wang AHJ, Molineux I and Rich A (1979) Structure at 2.3 Å resolution of the gene 5 product of bacteriophage *fd*: a DNA unwinding protein. ***J. Mol. Biol.*** **134**:379-400

Merrick M (1983) Nitrogen control of the *nif* regulon in *Klebsiella pneumoniae*: involvement of the *ntrA* gene and analogies between *ntrC* and *nifA*. ***EMBO J.*** **2**:39-44

Merrick M and Coppard JR (1989) Mutations in genes downstream of the *rpoN* gene (encoding σ54) of *Klebsiella pneumoniae* affect expression from σ54-dependent promoters. ***Mol. Microbiol.*** **3**:1765-1775

Merrick M, Filser M, Dixon R, Elmerich C, Sibold L and Houmard J (1980) Use of translocatable genetic elements to construct a fine-structure map of the *Klebsiella pneumoniae* nitrogen fixation (*nif*) gene cluster. ***J. Gen. Microbiol.*** **117**:509-520

Merrick M, Gibbons 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

Merrick M, Hill S, Hennecke H, Hahn M, Dixon R and Kennedy C (1982) Repressor properties of the *nifL* gene product in *Klebsiella pneumoniae*. ***Mol. Gen. Genet.*** **185**:75-81

Merrick MJ (1988) Organization and regulation of nitrogen fixation genes in *Klebsiella* and *Azotobacter*, p. 293-302. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Merrick MJ and Gibbons JR (1985) The nucleotide sequence of the nitrogen regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial sigma factors. ***Nucleic Acids Res.*** **13**:7607-7620

Merrick MJ and Stewart WDP (1985) Studies on the regulation and function of the *Klebsiella pneumoniae ntrA* gene. ***Gene*** **35**:297-303

Meyer TF, Billyard E, Haas R, Storzbach S and So M (1984) Pilus genes of *Neisseria gonorrhoeae*: chromosomal organization and DNA sequence. ***Proc. Natl. Acad. Sci. USA.*** **81**:6110-6114

Miller JH (1972) Experiments in molecular genetics, p. 352-355. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Minchin SD, Austin S and Dixon RA (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 SD, Austin S and Dixon RA (1989) Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NtrC is face of the helix dependent and the activator stabilizes the interaction of σ54 RNA polymerase with the promoter. ***EMBO J.*** **8**:3491-3499

Minton NP and Clarke LE (1985) Identification of the promoter of the *Pseudomonas* gene coding for carboxypeptidase G2. ***J. Mol. Appl. Genet.*** **3**:26-35

Miranda-Rios J, Sanchez-Pescador R, Urdea M and Covarrubias AA (1987) The complete nucleotide sequence of the *glnALG* operon of *Escherichia coli* K12. ***Nucleic Acids Res.*** **15**:2757-2770

Mjoli N and Kulpa CF (1988) The identification of a unique outer membrane protein required for iron oxidation in *Thiobacillus ferrooxidans*, p. 89-102. In: *Biohydrometallurgy 88*, PR Norris and DP Kelly (ed.). Science and Technology Letters, Kew, Surrey

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 σ54 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 seperated from the positive control function of NifA. ***Nucleic Acids Res.*** **16**:11469-11487

Mullin D, Minnich S, Chen L-S and Newton A (1987) A set of positively regulated flagellar gene promoters in *Caulobacter crescentus* with sequence homology to the *nif* gene promoters of *Klebsiella pneumoniae*. ***J. Mol. Biol.*** **195**:939-943

Mullin DA 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

Ninfa AJ 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 AJ, Mullin DA, Ramakrishnan G and Newton A (1989) *Escherichia coli* σ54 RNA polymerase recognises *Caulobacter crescentus flaG* and *flaN* flagellar gene products *in vitro*. ***J. Bacteriol.*** **171**:383-391

Ninfa AJ, Ninfa EG, Lupas A, Stock A, Magasanik B and Stock J (1988) Crosstalk between bacterial chemotaxis signal transduction proteins and the 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 AJ, Reitzer LJ and Magansanik B (1987) Initiation of transcription at the bacterial *glnAp2* promoter by purified *Escherichia coli* components is facilitated by enhancers. ***Cell*** **50**:1039-1046

Ninfa AJ, Ueno-Nishio S, Hunt TP, Robustell B and Magasanik B (1986) Purification of nitrogen regulator II, the product of the *glnL*(*ntrB*) gene of *Escherichia coli*. ***J. Bacteriol.*** **168**:1002-1004

Nixon TB, Ronson CW and Ausubel FM (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

Nohno T, Noji S, Taniguchi S and Saito T (1989) The *narX* and *narL* genes encoding the nitrate-sensing regulators of *Escherichia coli* are homologous to a family of prokaryotic two-component regulatory genes. ***Nucleic Acids Res.*** **17**:2947-2957

Nohno T, Saito T and Hong J-S (1986) Cloning and complete nucleotide sequence of the *Escherichia coli* glutamine permease operon (*glnHPQ*). ***Mol. Gen. Genet.*** **205**:260-269

Norrander J, Kempe T and Messing J (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. ***Gene*** **261**:101-106

Ow DW and Ausubel FM (1983) Regulation of nitrogen metabolism genes by the *nifA* gene product in *Klebsiella pneumoniae*. ***Nature*** **301**:307-313

Ow DW, Sundaresan B, Rothstein DM, Brown SE and Ausubel FM (1983) Promoters regulated by the *glnG*(*ntrC*) and *nifA* gene products share a heptameric consensus sequence in the -15 region. ***Proc. Natl. Acad. Sci. USA.*** **80**:2524-2528

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

Pabo C and Sauer RT (1984) Protein - DNA recognition. ***Annu. Rev. Biochem.*** **53**:293-321

Pahel G, Rothstein DM and Magasanik B (1982) Complex *glnA-glnL-glnG* operon of *Escherichia coli*. ***J. Bacteriol.*** **150**:202-213

Pahel G and Tyler B (1979) A new *glnA*-linked regulatory gene for glutamine synthetase in *Escherichia coli*. ***Proc. Natl. Acad. Sci. USA.*** **76**:4544-4548

Pasloske BL, Drummond DS, Frost LS and Paranchych W (1989) The activity of the *Pseudomonas aeruginosa* pilin promoter is enhanced by an upstream regulatory site. ***Gene*** **81**:25-34

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

Pecher A, Zinoni F, Jatisatienr C, Wirth R, Hennecke H and Böck A (1983) On the redox control of synthesis of anaerobically induced enzymes in enterobacteriaceae. ***Arch. Microbiol.*** **136**:131-136

Peck HD Jr and Gest H (1957) Formic dehydrogenase and the hydrogen-lyase enzyme complex in the *coli-aerogenes* group. ***J. Bacteriol.*** **73**:706-721

Pederosa FO and Yates MG (1984) Regulation of nitrogen fixation (*nif*) genes of *Azospirillum brasiliense* by *nifA* and *ntr* (*gln*) type gene products. ***FEMS Microbiol. Lett.*** **23**:95-101

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

Post E, Kleiner D and Oelze J (1983) Whole cell respiration and nitrogenase activity in *Azotobacter vinelandii* grown in oxygen controlled continuous culture. ***Arch. Microbiol.*** **134**:68-72

Postgate J (1989) Trends and perspectives in nitrogen fixation research. ***Adv. in Microbial Phys.*** **30**:1-22

Postgate JR (1982) *The fundamentals of nitrogen fixation.* Cambridge University Press, Cambridge

Postgate JR and Eady RR (1988) The evolution of biological nitrogen fixation, p. 31-40. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart

Postgate JR and Kent HM (1987) Qualitative evidence for expression of *Klebsiella pneumoniae nif* genes in *Pseudomonas putida*. ***J. Gen. Microbiol.*** **133**:2563-2566

Pretorius I-M, Rawlings DE and Woods DR (1986) Identification and cloning of *Thiobacillus ferrooxidans* structural *nif* genes in *Escherichia coli*. ***Gene*** **45**:59-65

Pretorius I-M, Rawlings DE, O'Neill EG, Jones WA, Kirby R and Woods DR (1987) Nucleotide sequence of the gene encoding the nitrogenase iron protein of *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **169**:367-370

Quispel A (1988) Hellriegel and Wilfarth's discovery of (symbiotic) nitrogen fixation hundred years ago, p. 3-12. In: *Nitrogen fixation: hundred years after*, H Bothe *et al.* (ed.). Gustav Fischer, Stuttgart, New York

Ramakrishnan G and Newton A (1990) FlbD of *Caulobacter crescentus* is a homologue of the NtrC (NR_I) protein and activates σ54 dependent flagellar gene promoters. ***Proc. Nat. Acad. Sci. USA.*** **87**:2369-2373

Ramesar RS, Abratt V, Woods DR and Rawlings DE (1989) Nucleotide sequence and expression of a cloned *Thiobacillus ferrooxidans recA* gene in *Escherichia coli*. ***Gene*** **78**:1-8

Ramesar RS, Woods DR and Rawlings DE (1988) Cloning and expression in *Escherichia coli* of a *recA*-like gene from the acidophilic autotroph *Thiobacillus ferrooxidans*. ***J. Gen. Microbiol.*** **134**:1141-1146

Rawlings DE (1981) Nutritional requirements of the microorganisms active in the oxidation of ferrous iron in acid mine leach liquors. ***Appl. Bacteriol.*** **51**:267-275

Rawlings DE (1988) Sequence and structural analysis of the α- and β-dinitrogenase subunits of *Thiobacillus ferrooxidans*. ***Gene*** **69**:337-343

Rawlings DE (1989) A comparison of the structure and expression of several genes from *Thiobacillus ferrooxidans* with those of other bacteria, p. 3-8. In: *Biotechnology in Minerals and Metal Processing*, S Scheiner *et al.* (ed.). SME Press, Littleton, Colorado

Rawlings DE, Gawith C, Petersen A and Woods DR (1983) Characterization of plasmids and potential genetic markers in *Thiobacillus ferrooxidans*, p. 555-570. In: *Recent progress in biohydrometallurgy*, G Rossi and AE Torma (ed.). Associazione Mineraria Sarda, Cagliari

Rawlings DE, Jones WA, O'Neill EG and Woods DR (1987) Nucleotide sequence of the glutamine synthetase gene and its controlling region from the acidophilic autotroph *Thiobacillus ferrooxidans*. ***Gene*** **53**:211-217

Rawlings DE, Pretorius I-M and Woods DR (1984) Expression of a *Thiobacillus ferrooxidans* origin of replication in *Escherichia coli*. ***J. Bacteriol.*** **158**:737-738

Rawlings DE, Pretorius I-M and Woods DR (1988) A comparative sequence analysis of *Thiobacillus ferrooxidans* nitrogen metabolism gene structure. In: *Biohydrometallurgy 87*, PR Norris and DP Kelly (ed.). Science and Technology letters, London

Rawlings DE (1991) Engineering *Thiobacillus* for mining. In: *Biotechnology of Open Systems*, S Silver (ed.). Academic Press, Florida, in press

Rawlings DE, Woods DR and Mjoli NP (1991) The cloning and structure of genes from the autotrophic biomining bacterium, *Thiobacillus ferrooxidans*. In: *Advances in Gene Technology*, PJ Greenway (ed.). JAI Press, London, in press

Ray L, Claveriemartin F, Weglenski P and Magasanik B (1990) Role of the promoter in activation of transcription by nitrogen regulator I phosphate in *Escherichia coli*. ***J. Bacteriol.*** **172**:818-823

Reed KC and Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. ***Nucleic Acids Res.*** **13**:7207-7221

Reitzer LJ, Bueno R, Cheng WD, Abrams SA, Rothstein DM, Hunt TP, Tyler B and Magasanik B (1987) Mutations that create new promoters suppress the σ54 dependence of *glnA* transcription in *Escherichia coli*. ***J. Bacteriol.*** **169**:4279-4284

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

Reitzer LJ and Magasanik B (1985) Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. ***Proc. Natl. Acad. Sci. USA.*** **82**:1979-1983

Reitzer LJ and Magasanik B (1986) Transcription of *glnA* in *Escherichia coli* is stimulated by an activator bound to sites far from the promoter. ***Cell*** **45**:785-792

Reitzer LJ and Magasanik B (1987) Ammonia assimilation and the biosynthesis of glutamine, glutamate, aspartate, asparagine, L-alanine and D-alanine, p 302-320. In: Escherichia coli *and* Salmonella typhimurium *cellular and molecular biology*, FC Niedhardt (ed.). Amer. Soc. Microbiol., Washington DC

Reitzer LJ, 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

Rhee SG, Chock PB and Stadtman ER (1985) Glutamine synthetase from *Escherichia coli*. ***Meth. Enzymol.*** **113**:213-241

Riedel GE, Ausubel FM and Cannon FC (1979) Physical map of chromosomal nitrogen fixation (*nif*) genes of *Klebsiella pneumoniae*. ***Proc. Natl. Acad. Sci. USA.*** **76**:2866-2870

Riedel GE, Brown SE and Ausubel FM (1983) Nitrogen fixation by *Klebsiella pneumoniae* is inhibited by certain multicopy hybrid *nif* plasmids. ***J. Bacteriol.*** **153**:45-56

Rigby PWJ, Dieckmann M, Rhodes C and Berg P (1977) Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase. ***J. Mol. Biol.*** **113**:237-251

Robertson CA and Nash HA (1988) Bending of the bacteriophage λ attachment site by *Escherichia coli* integration host factor. ***J. Biol. Chem.*** **263**:3554-3557

Robson RL, Eady RR, Richardson TH, Miller RW, Hawkins M and Postgate JR (1986a) The alternative nitrogenase of *Azotobacter chroococcum* is a vanadium enzyme. ***Nature*** **322**:388-390

Robson RL, Woodley PR and Jones R (1986b) Second gene (*nifH*) coding for a nitrogenase iron-protein in *Azotobacter chroococcum* is adjacent to a gene coding for a ferredoxin-like protein. ***EMBO J.*** **5**:1159-1163

Romeo JM and Zusman DR (1987) Cloning of the gene for myxobacterial hemagglutinin and isolation and analysis of structural gene mutations. ***J. Bacteriol.*** **169**:3801-3808

Romermann D, Lohmeyer M, Friedrich CG and Friedrich B (1988) Pleiotrophic mutants from *Alcaligenes eutrophus* defective in the metabolism of hydrogen, nitrate, urea, and fumarate. ***Arch. Microbiol.*** **149**:471-475

Romermann D, Warrelmann J, Bender RA and Friedrich B (1989) A *rpoN*-like gene of *Alcaligenes eutrophus* and *Pseudomonas facilis* controls expression of diverse metabolic pathways, including hydrogen oxidation. ***J. Bacteriol.*** **171**:1093-1099

Ronson CW, Astwood PM, Nixon BT and Ausubel FM (1987a) Deduced products of C4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products. ***Nucleic Acids Res.*** **15**:7921-7934

Ronson CW, Nixon BT, Albright LM and Ausubel FM (1987b) *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. ***J. Bacteriol.*** **169**:2424-2430

Ronson CW, Nixon BT and Ausubel FM (1987c) Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. ***Cell*** **49**:579-581

Rossbach S, Schell J and De Bruijn FJ (1987) The *ntrC* gene of *Agrobacterium tumefaciens* C58 controls glutamine synthetase (GSII) activity, growth on nitrate and chromosomal but not Ti-encoded arginine catabolism pathways. ***Mol. Gen. Genet.*** **209**:419-426

Rossbach S, Schell J and De Bruijn FJ (1988) Cloning and analysis of *Agrobacterium tumefaciens* C58 loci involved in glutamine biosynthesis: Neither the *glnA* (GSI) nor the *glnII* (GSII) gene plays a special role in virulence. ***Mol. Gen. Genet.*** **212**:38-47

Salser W (1977) Globin mRNA sequences: analysis of base-pairing and evolutionary implications. ***Cold Spring Harbor Symp. Quant. Biol.*** **42**:985-1002

Sanders DA, Gillece-Castro BL, Stock AM, Burlingame AL and Koshland DE (1989) Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. ***J. Biol. Chem.*** **264**:21770-21778

Sanger F, Nicklen S and Coulson AR (1977) DNA sequencing with chain terminating inhibitors. ***Proc. Natl. Acad. Sci. USA*** **74**:5463-5467

Sanjuan J and Olivares J (1989) Implication of NifA in regulation of genes located on a *Rhizobium meliloti* cryptic plasmid that affect nodulation efficiency. ***J. Bacteriol.*** **171**:4154-4161

Sankar P, Lee JH and Shanmugam KT (1988) Gene-product relationship of *fhlA* and *fdv* genes of *Escherichia coli*. ***J. Bacteriol.*** **170**:5440-5445

Santero E, Hoover T, Keener J and Kustu S (1989) *In vitro* activity of the nitrogen fixation regulatory protein NifA. ***Proc. Nat. Acad. Sci. USA.*** **86**:7346-7350

Santero E, Luque F, Medina JR and Tortolero M (1986) Isolation of *ntrA*-like mutants of *Azotobacter vinelandii*. ***J. Bacteriol.*** **166**:541-544

Santero E, Toukdarian A, Humphrey R and Kennedy C (1988) Identification and characterization of two nitrogen fixation regulatory regions, *nifA* and *nfrX*, in *Azotobacter vinelandii* and *Azotobacter chroococcum*. ***Mol. Microbiol.*** **2**:303-314

Sasse-Dwight S and Gralla JD (1988) Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. ***Proc. Natl. Acad. Sci. USA.*** **85**:8934-8938

Sawers GR, Ballantine SP and Boxer DH (1985) Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12: Evidence for a third isoenzyme. ***J. Bacteriol.*** **164**:1324-1331

Scherings GS, Haaker H, Wassink H and Veeger C (1983) On the formation of an oxygen-tolerant three-component nitrogenase complex from *Azotobacter vinelandii*. ***Eur. J. Biochem.*** **135**:591-599

Schlensog V, Birkmann A and Böck A (1989) Mutations in trans which affect the anaerobic expression of a formate dehydrogenase (*fdhF*) structural gene. ***Arch. Microbiol.*** **152**:83-89

Schmitz G, Durre P, Mullenbach G and Ferro-Luzzi Ames G (1987) Nitrogen regulation of transport operons: analysis of promoters *argTr* and *dhuA*. ***Mol. Gen. Genet.*** **209**:403-407

Schmitz G, Nikaido K and Ferro-Luzzi Ames G (1988) Regulation of a transport operon promoter in *Salmonella typhimurium*: identification of sites essential for nitrogen regulation. ***Mol. Gen. Genet.*** **215**:107-117

Schrader JA and Holmes DS (1988) Phenotypic switching of *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **170**:3915-3023

Servin-Gonzalez L and Bastarrachea F (1984) Nitrogen regulation of synthesis of the high affinity methylammonium transport system of *Escherichia coli*. ***J. Gen. Microbiol.*** **130**:3071-3077

Shaibe E, Metzer E and Halpern YS (1985) Control of utilization of L-arginine, L-ornithine, agmatine, and putrescine as nitrogen sources in *Escherichia coli* K-12. ***J. Bacteriol.*** **163**:938-942

Shiratori T, Inoue C, Sugawara K, Kusano T and Kitagawa Y (1989) Cloning and expression of *Thiobacillus ferrooxidans* mercury ion resistance genes in *Escherichia coli*. ***J. Bacteriol.*** **171**:3458-3464

Siegele DA, Hu JC, Walter WA and Gross CA (1989) Altered promoter recognition by mutant forms of the σ70 subunit of *Escherichia coli* RNA polymerase. ***J. Mol. Biol.*** **206**:591-603

Silhavy TJ and Beckwith JR (1985) Uses of *lac* fusions for the study of biological problems. ***Microbiol. Rev.*** **49**:398-418

Singh M, Kreutzer R, Acker G and Klingmüller W (1988) Localization and physical mapping of a plasmid-borne 23-kb *nif* gene cluster from *Enterobacter agglomerans* showing homology to the entire *nif* gene cluster of *Klebsiella pneumoniae* M5a1. ***Plasmid*** **19**:1-12

Singh M, Tripathi AK and Klingmüller W (1989) Identification of a regulatory *nifA* type gene and physical mapping of cloned new *nif* regions of *Azospirillum brasilense*. ***Mol. Gen. Genet.*** **219**:235-240

Son HS and Rhee SG (1987) Cascade control of *Escherichia coli* glutamine synthetase. Purification and properties of P_{II} protein and nucleotide sequence of its structural gene. ***J. Biol. Chem.*** **262**:8690-8695

Stadtman ER and Ginsburg A (1974) The glutamine synthetase of *Escherichia coli*: structure and control, p. 755-807. In: *The enzymes*, Vol. 10, PD Boyer (ed.). Academic Press, New York

Stadtman ER, Mura E, Chock PB and Rhee SG (1980) The interconvertible enzyme cascade that regulates glutamine synthetase activity, p. 59-123. In: *Glutamine: metabolism, enzymology and regulation*, Vol. 41, J Mora and R Palacios (ed.). Academic Press, New York

Stanley J, Vanslooten J, Dowling DN, Finan T and Broughton WJ (1989) Molecular cloning of the *ntrA* gene of the broad host-range Rhizobium Sp Ngr234 and phenotypes of a site-directed mutant. ***Mol. Gen. Genet.*** **2**:528-532

Stern MJ, Higgins CF and Ferro-Luzzi Ames G (1984) Isolation and characterization of *lac* fusions to two nitrogen-regulated promoters. ***Mol. Gen. Genet.*** **195**:219-227

Stewart RC and Dahlquist FW (1987) Molecular components of bacterial chemotaxis. ***Chem. Rev.*** **87**:997-1025

Stewart V (1988) Nitrate respiration in relation to facultative metabolism in enterobacteria. ***Microbiol. Rev.*** **52**:190-232

Stock AM, Chen T, Welsh D and Stock J (1988a) CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. ***Proc. Natl. Acad. Sci. USA.*** **85**:1403-1407

Stock AM, Mottonen JM, Stock JB and Schutt CE (1989a) Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. ***Nature*** **337**:745-749

Stock AM, Wylie DC, Mottonen JM, Lupas AN, Ninfa EG, Ninfa AJ, Schutt CE and Stock JB (1988b) Phosphoproteins involved in bacterial signal transduction. ***Cold Spring Harbor Symp. Quant. Biol.*** **53**:49-57

Stock JB, Ninfa AJ and Stock AM (1989b) Protein phosphorylation and regulation of adaptive responses in bacteria. ***Microbiol. Rev.*** **53**:450-490

Stock JB, Stock AM and Mottonen JM (1990) Signal transduction in bacteria. ***Nature*** **344**:395-400

Stoker K, Reijnders WNM, Oltmann LF and Stouthamer AH (1989) Initial cloning and sequencing of *hydHG*, an operon homologous to *ntrBC* and regulating the labile hydrogenase activity in *Escherichia coli* K-12. ***J. Bacteriol.*** **171**:4448-4456

Stragier P, Parsot C and Bouvier J (1985) Two functional domains conserved in major and alternative bacterial sigma factors. ***FEBS Lett.*** **187**:11-15

Struhl K (1985) A rapid method for creating recombinant DNA molecules. ***Biotechniques*** **3**:452-453

Su W, Porter S, Kustu S and Echols H (1990) DNA-looping and enhancer activity: Association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. ***Proc. Natl. Acad. Sci. USA.*** **87**:5504-5508

Sugio T, Domatsu C, Munakata O, Tano T and Imai K (1985) Role of a ferric ion-reducing system in sulfur oxidation of *Thiobacillus ferrooxidans*. ***Appl. Environ. Microbiol.*** **49**:1401-1406

Sugio T, Katagiri T, Moriyama M, Li Zhen Y, Inagaki K and Tano T (1988a) Existence of a new type of sulfite oxidase which utilizes ferric ions as an electron acceptor in *Thiobacillus ferrooxidans*. ***Appl. Environ. Microbiol.*** **54**:153-157

Sugio T, Mizunashi W, Inagaki K and Tano T (1987) Purification and some properties of sulfur:ferric ion oxidoreductase from *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **169**:4916-4922

Sugio T, Tsujita Y, Inagaki K and Tano T (1990) Reduction of cupric ions with elemental sulfur by *Thiobacillus ferrooxidans*. ***Appl. Environ. Microbiol.*** **56**:693-696

Sugio T, Tsujita Y, Katagiri T, Inagaki K and Tano T (1988b) Reduction of Mo^{+6} with elemental sulfur by *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **170**:5956-5959

Sundaresan V, Jones JDG, Ow DW and Ausubel FM (1983a) *Klebsiella pneumoniae nifA* product activates the *Rhizobium meliloti* nitrogenase promoter. ***Nature*** **301**:728-732

Sundaresan V, Ow DW and Ausubel FM (1983b) Activation of *Klebsiella pneumoniae* and *Rhizobium meliloti* nitrogenase promoters by *gln*(*ntr*) regulatory proteins. ***Proc. Natl. Acad. Sci. USA.*** **80**:4030-4034

Szeto WW, Nixon BT, Ronson CW and Ausubel FM (1987) Identification and characterization of the *Rhizobium meliloti ntrC* gene: *R. meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. ***J. Bacteriol.*** **169**:1423-1432

Tabor S and Richarson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase ***Proc. Natl. Acad. Sci. USA.*** **84**:4767-4771

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

Totten PA, Lara JC 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

Tuli R, Fisher R and Haselkorn R (1982) The *ntr* genes of *Escherichia coli* activate the *hut* and *nif* operons of *Klebsiella pneumoniae*. ***Gene*** **19**:109-116

Tuovinen OH and Kelly DP (1974) Studies on the growth of *T. ferrooxidans*. V. Factors affecting growth in liquid culture and development of colonies on solid media containing inorganic sulphur compounds. ***Arch. Microbiol.*** **98**:351-364

Tuovinen OH, Panda FA and Tsuchiya HM (1979) Nitrogen requirement of iron-oxidizing thiobacilli for acidic ferric sulfate regeneration. ***Appl. Environ. Microbiol.*** **37**:954-958

Tyler B (1978) Regulation of the assimilation of nitrogen compounds. ***Annu. Rev. Biochem.*** **47**:1126-1162

Ueno-Nishio S, Backman KC, and Magasanik B (1983) Regulation at the *glnL*-operator-promoter of the complex *glnALG* operon of *Escherichia coli*. ***J. Bacteriol.*** **153**:1247-1251

Ueno-Nishio S, Mango S, Reitzer LJ and Magasanik B (1984) Identification and regulation of the *glnL* operator-promoter of the complex *glnALG* operon of *Escherichia coli*. ***J. Bacteriol.*** **160**:379-384

Vishniac WV (1974) Genus 1. *Thiobacillus* Beiernick 1904, p. 458-460. In: *Bergey's manual of determinative bacteriology*, Vol. 597, 8th Ed., RE Buchanan and NE Gibbons (ed.). The Williams and Wilkins Co., Baltimore

Wanner B (1987) Phosphate regulation of gene expression in *Escherichia coli*, p. 1326-1333. In: Escherichia coli *and* Salmonella typhimurium*: cellular and molecular biology*, FC Neidhardt *et al.* (ed.). Amer. Soc. Microbiol., Washington DC

Wanner BL, Wilmes MR and Young DC (1988) Control of bacterial alkaline phosphatase synthesis and variation in an *Escherichia coli* K12 *phoR* mutant by adenyl cyclase, the cyclic AMP receptor protein, and the *phoM* operon. ***J. Bacteriol.*** **170**:1092-1102

Wardhan H, McPherson MJ and Sastry GRK (1989) Identification, cloning, and sequence analysis of the nitrogen regulation gene *ntrC* of *Agrobacterium tumefaciens* C58. ***Mol. Plant-Microbe Inter.*** 2:241-248

Wedel A, Weiss DS, Popham D, Droge P and Kustu S (1990) A bacterial enhancer functions to tether a transcriptional activator near a promoter. ***Science*** **248**:486-490

Weglenski P, Ninfa AJ, Ueno-Nishio S and Magasanik B (1989) Mutations in the *glnG* gene of *Escherichia coli* that result in increased activity of nitrogen regulator I. ***J. Bacteriol.*** **171**:4479-4485

Wei GR and Kustu S (1981) Glutamine auxotrophs with mutations in a nitrogen regulatory gene, *ntrC*, that is near *glnA*. ***Mol. Gen. Genet.*** **183**:392-399

Weiss V and Magasanik B (1988) Phosphorylation of nitrogen regulator I (NR_I) of *Escherichia coli*. ***Proc. Natl. Acad. Sci. USA.*** **85**:8919-8923

Wimpenny JWT and Cole JA (1967) The regulation of metabolism in facultative bacteria. III. The effect of nitrate. ***Biochim. Biophys. Acta.*** **148**:133-242

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

Wylie D, Stock A, Wong C-Y and Stock J (1988) Sensory transduction in bacterial chemotaxis involves phosphotransfer between Che proteins. ***Biochem. Biophys. Res. Commun.*** **151**:891-896

Yates JR, Cunningham RP and Holmes DS (1988) IST2: An insertion sequence from *Thiobacillus ferrooxidans*. ***Proc. Nat. Acad. Sci. USA.*** **85**:7284-7287

Yates JR and Holmes DS (1987) Two families of repeated sequences in *Thiobacillus ferrooxidans*. ***J. Bacteriol.*** **169**:1861-1870

Yudkin MD (1987a) Structure and function in a *Bacillus subtilis* sporulation-specific sigma factor: Molecular nature of mutations in *spoIIAC*. ***J. Gen. Microbiol.*** **133**:475-481

Yudkin MD (1987b) The prediction of helix-turn-helix DNA-binding regions in proteins. ***Prot. Eng.*** 1:371-372

Yura T and Ishihama A (1979) Genetics of bacterial RNA polymerases. ***Annu. Rev. Genet.*** **13**:59-97

Zabeau M and Stanley KK (1982) Enhanced expression of cro-β-galactosidase fusion proteins under the control of the P_R promoter of the bacteriophage lambda. ***EMBO J.*** 1:1217-1224

Zinoni F, Birkmann A, Stadtman TC and Bock A (1986) Nucleotide sequence and expression of the selenocysteine containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase linked) from *Escherichia coli*. ***Proc. Natl. Acad. Sci. USA.*** **83**:4650-4654

Zuber P, Healy J, Carter HL, Cutting S, Moran CP and Losick R (1989) Mutation changing the specificity of an RNA polymerase sigma factor. ***J. Mol. Biol.*** **206**:605-614