Characterization of the *nifUHD* Cluster

and a new Myoglobin-Like Gene from

*Nostoc commune* UTEX 584

by

Stephen V. Angeloni

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of
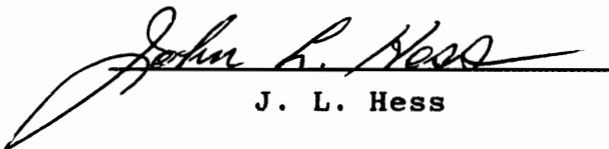
DOCTOR OF PHILOSOPHY

in

Biochemistry and Nutrition

APPROVED:

_____
M. Potts, Chairman

_____
J. L. Hess

_____
T. W. Keenan

_____
G. H. Lacy

_____
T. J. Larson

March 1992

Blacksburg, Virginia

Characterization of the *nifUHD* Cluster

and a new Myoglobin-Like Gene from

*Nostoc commune* UTEX 584

by

Stephen V. Angeloni

Committee Chairman: Malcolm Potts

Biochemistry and Nutrition

(Abstract)

Sequence analysis of the entire 3.5 kb *Hind*III genomic

DNA fragment previously isolated from *Nostoc commune* UTEX

584 (Defrancesco and Potts 1988), determined the exact

locations of the *nifU*, *nifH*, and *nifD* genes and identified

two potential stem loop structures, a direct repeat, and an

ORF that codes for a protein with a predicted amino acid

sequence similar to that of myoglobin from *Paramecium*

*caudatum*.

The *N. commune* UTEX 584 myoglobin-like protein has a

predicted length of 118 amino acids and molecular mass of

12,906 Da. A PCR copy of the gene (*glbN*) was cloned for

overexpression of the protein. The recombinant protein was

purified and used for spectral analysis and for the

production of polyclonal antisera.

Treatment of the recombinant protein with dithionite

and CO resulted in spectral shifts characteristic of

hemoproteins that bind oxygen. While some of the spectral

characteristics are unique to the protein, in general the spectra were more like those of globins than cytochromes. Based on these characteristics and the sequence similarity to the *P. caudatum* myoglobin, we proposed the name cyanoglobin, with the gene designation *glbN* and the protein designation GlbN.

Western analysis of GlbN expression was performed on *N. commune* UTEX 584 and two species of *Anabaena* (*Anabaena* sp. PCC 7120 and *Anabaena variabilis*). In *N. commune* UTEX 584 a protein with a molecular mass similar to that predicted for GlbN was detected. This protein was produced in increased amounts under the same growth conditions that resulted in increased production of nitrogenase reductase (the *nifH* gene product). No proteins of similar size to GlbN were detected in *Anabaena* sp. PCC 7120 or *A. variabilis*.

A possible function of GlbN may be for oxygen storage, transport, or protection of the nitrogenase system. These functions as well as those of the direct repeat and the potential stem loop structures and their relationship to nitrogen fixation or other physiological processes in *N. commune* UTEX 584 require further analysis.

# List of Abbreviations

bp, Base pairs

Chl *a*, Chlorophyll *a*

CIA, 24:1 chloroform:isoamyl alcohol

CO, Carbon monoxide

Da, Daltons

ddH$_2$O, Distilled deionized water

DTT, Dithiothreitol

EDTA, Ethylenediaminetetraacetic acid

*g*, Acceleration of gravity

IPTG, Isopropylthio-β-D-galactopyranoside

kb, Kilobase(s)

LBA, Luria-Bertani media containing 100 μg/ml ampicillin

ORF(s), Open reading frame(s)

PBS, Phosphate buffered saline

PCR, Polymerase chain reaction

PEG, Polyethylene glycol

PMSF, Phenylmethylsulfonyl fluoride

SDS, Sodium dodecyl sulfate

SSC, Standard saline citrate

ssRNA, Single stranded RNA

STET, Sucrose, Triton X-100, EDTA, Tris buffer

STE, Sodium chloride, Tris, EDTA buffer

TAE, Tris acetate EDTA buffer

TBE, Tris borate EDTA buffer

Tris, Tris(hydroxymethyl)aminomethane

TTBS, Tween Tris buffered saline

# Table of Contents

# List of Figures

# List of Tables

# Introduction

The utilization of nitrogen by all living organisms depends on the conversion of atmospheric nitrogen ($N_2$) into ammonia, nitrates, and eventually nitrogen-containing organic compounds. A large part of this conversion is carried out by nitrogen fixing bacteria. These bacteria can be divided into two groups: photosynthetic, such as the cyanobacteria, and non-photosynthetic. There are a number of differences between these bacteria with respect to their biochemical and morphological characteristics as well as the conditions under which they fix nitrogen. Despite these differences, the genes for their nitrogenases and accessory proteins are highly conserved and the arrangements of these genes in the different species have some similarities as well as differences (Jacobson et al. 1989, Merrick 1988). These differences represent the need to insure the proper functioning of nitrogenase during the life cycles and under the different conditions encountered by each species of bacteria.

Most of the non-photosynthetic bacteria fix nitrogen when in symbiotic association with plants. Very few can fix nitrogen under free-living conditions in nature. Most cyanobacteria are capable of fixing nitrogen under free living conditions and some during symbiotic association with plants (Söderbäck et al. 1990, Lindblad et al. 1991). This

1

ability requires specific genetic, biochemical, and
morphological characteristics which differ from the non-
photosynthetic bacteria. In addition, since cyanobacteria
are photosynthetic prokaryotes, they need to deal with
oxygen generated by the photosynthetic apparatus during
light absorption since the nitrogenase complex is oxygen
labile. Different species of cyanobacteria handle this
problem in different ways depending on their life cycle as
will be discussed below. The life cycles of the different
cyanobacterial species are quite diverse. In general, they
can be divided into two groups, filamentous and non-
filamentous. The filamentous can be further divided into
those that form specialized cells called heterocysts and
those that do not. *N. commune* UTEX 584 is considered a
heterocystous filamentous cyanobacteria.

In the absence of combined nitrogen some of the cells
along the filament of *N. commune* UTEX 584 differentiate into
heterocysts. These cells are responsible for fixing nitrogen
(Haselkorn 1978, Thomas et al. 1977, Wolk et al. 1974, Meeks
et al. 1978, Peterson and Wolk 1978). Dinitrogen in the
heterocyst is converted to ammonia by nitrogenase (Figure
1). This ammonia is used to synthesize glutamine from
glutamate which is then exported to the vegetative cells.
The heterocyst is believed to provide an anaerobic

Figure 1. Some proposed interactions between the heterocyst and vegetative cells.

Key enzymes are nitrogenase (1), glutamine synthetase (2), and glutamate synthase (3). Key amino containing or accepting substrates are glutamate (Glu), glutamine (Gln), and $\alpha$-ketoglutarate ($\alpha$-ketoglu).
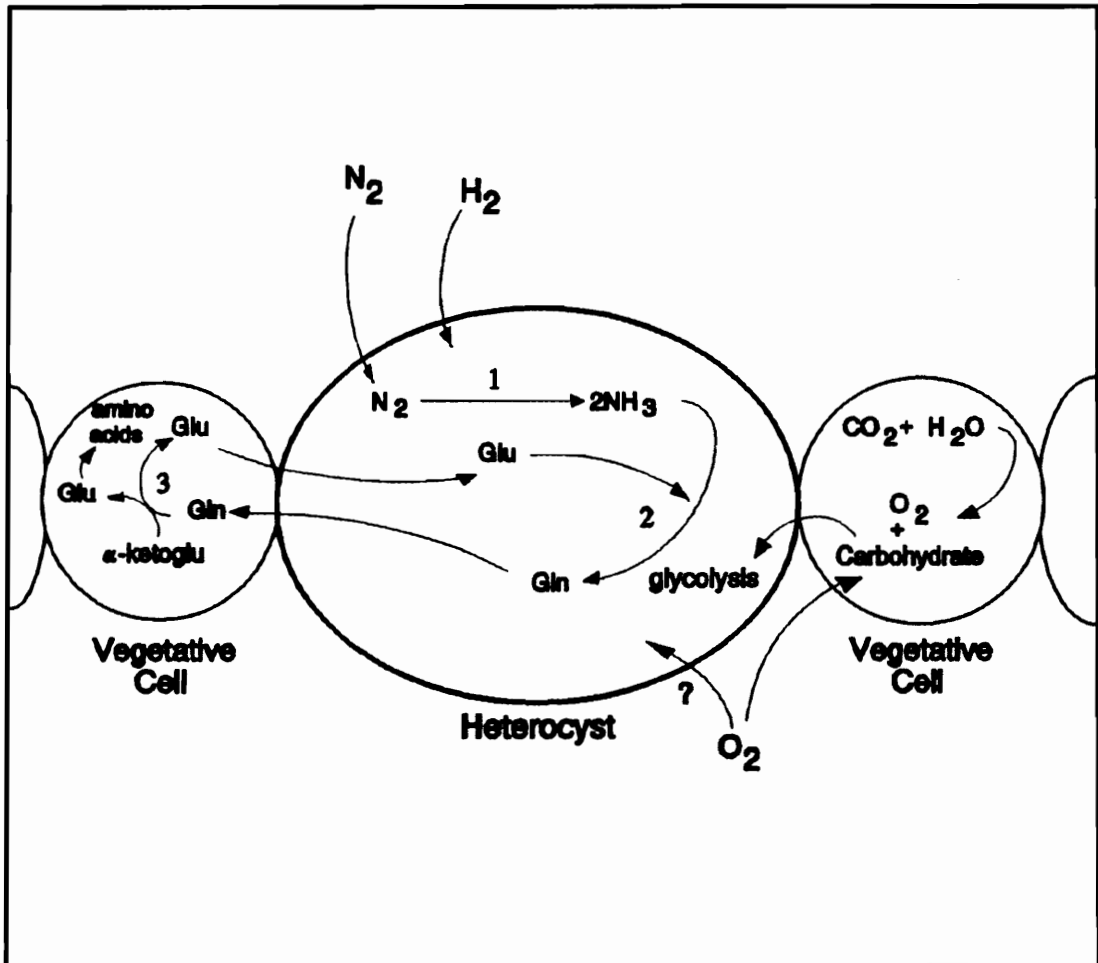
environment for the oxygen-sensitive nitrogenase complex (Haselkorn 1978, Rippka and Stanier 1978, Stanier and Cohen-Bazire 1977). This is why heterocysts lack photosystem II components which produce oxygen during non-cyclic electron flow (Haselkorn 1978, Donze et al. 1972). It has also been suggested that the thick sheath of carbohydrate around the heterocyst acts as a barrier to prevent oxygen from entering the heterocyst. This role of the sheath, however, is still a topic of debate since some studies have shown heterocysts to be able to take up or evolve oxygen, dinitrogen, and hydrogen, indicating that the sheath may not be an adequate barrier to these gases (Tabita et al. 1988, Smith et al. 1985). In general, these characteristics represent a spatial mechanism for separating nitrogen fixation from oxygen evolution which is usually employed by heterocystous cyanobacteria.

Another method of protecting nitrogenase from oxygen, which is employed by non-heterocystous cyanobacteria, is to separate temporally the periods of nitrogen fixation from oxygen evolving photosynthesis (Stal and Krumbein 1985, Mitsumi and Kumazawa 1988, Mitsumi et al. 1986, Gallon and Chaplin 1988). In these cases, nitrogen fixation usually occurs during the dark period while photosynthetic oxygen evolution occurs during the light period. Another method of oxygen protection could employ the reduction of oxygen to

water. One source of reductant for oxygen could be hydrogen, which can also serve as a source of reducing equivalents for nitrogen fixation (Bothe and Never 1988). This may be the reason for the increased rate of hydrogen uptake by some cyanobacteria during nitrogen fixation in the presence of oxygen (Scherer et al. 1980, Tabita et al. 1988).

In heterocystous cyanobacteria, the expression of nitrogen fixation genes should be coupled to the differentiation of cells into heterocysts. However, some studies have shown that induction of *nif* gene expression can be uncoupled from heterocyst differentiation by exposing the cells to a $N_2$ containing anaerobic atmosphere (Haselkorn 1978, Stal and Krumbein 1985, Helber et al. 1988a, 1988b, Kentemich et al. 1988a). Under these conditions cells did not differentiate into mature heterocysts but *nif* mRNA and nitrogenase activity was detected. These data show that in the absence of combined nitrogen and oxygen, the nitrogen fixation genes can be transcribed and nitrogenase is produced without the development of mature heterocysts, suggesting that oxygen may negatively regulate *nif* gene expression and possibly promote heterocyst development.

It has been shown under these anaerobic conditions, that the *nifH* promoter is activated in a pattern similar to the pattern of heterocyst development along the filament,

even though mature heterocysts do not form (Elhai and Wolk 1990). Hence, additional factors may be involved in stimulating heterocyst formation and its characteristic pattern. The characteristic pattern of heterocyst formation is believed to be due to chemical gradients of nitrogenous and/or photosynthetic compounds along the filament.

Another characteristic of heterocyst differentiation that affects *nif* gene regulation is a genome rearrangement that occurs in some but not all species of cyanobacteria. Two rearrangements have been reported in *Anabaena* sp. PCC 7120, both of which are deletions from the genome (Golden et al. 1988, 1987, 1985). One deletion is of 55 kb from within the 3' end of the *fdxN* gene, while the other deletion is of an 11 kb fragment from within the 3' end of the *nifD* gene (Figure 2). When these sections of the genome are deleted, the coding regions of the interrupted genes are restored. Excision of the 55 kb element is believed to produce a mRNA that contains *nifB, fdxN, nifS,* and *nifU.* An attempt to find such a transcript was undertaken, however, the results did not conclusively demonstrate if these genes are on the same transcript (Mulligan and Haselkorn 1989).

While no genes have been reported on the 55 kb element, the 11 kb element appears to code for several genes (Lammers et al. 1990). One gene called *xisA,* is believed to code for

Figure 2. Arrangement of *nif* genes in *Anabaena* sp. PCC 7120 heterocysts.

The arrangement of the genes is shown as they would occur in the heterocyst after excision of the 55 and 11 kb elements. The sites of the two excisions are shown. This figure is modified from Mulligan and Haselkorn 1989.

nifB    fdxN    nifS    nifU    nifH    nifD    nifK

428 bp

Site of 55 kb excision                    Site of 11 kb excision
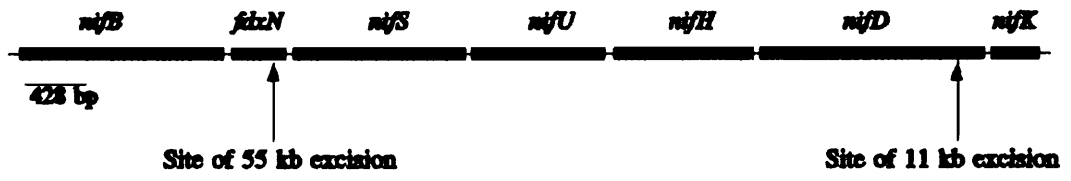
the protein responsible for excision of the 11 kb element
(Golden and Wiest 1988, Lammers et al. 1986). No mRNA for
this gene has been detected. Several other ORFs have been
discovered on this element (Lammers et al. 1990). The
predicted amino acid sequences of these ORFs are similar to
those of cytochrome type proteins. The role of these
proteins and when they are expressed is not known.

These types of genome rearrangements do not occur in
all species of cyanobacteria. While these rearrangements
have been suggested to occur in some species of *Nostoc*
(Meeks et al. 1988) it is not known if they occur in *N.
commune* UTEX 584. These rearrangements could explain why
heterocysts are not able to revert back to vegetative cells
when combined nitrogen becomes available. The loss of these
regions of the genome could prevent this reversion if the
excised regions contained genes required by vegetative
cells.

These genomic rearrangements do not occur in all
cyanobacteria and as a result the arrangement of the
nitrogen fixation genes in different species have certain
species-specific characteristics. Some differences in gene
arrangements have been observed in cyanobacteria and the
non-photosynthetic diazotrophs for the more than 20 *nif*
genes that have been found (Kallas et al. 1985, Merrick

1988, Jacobson et al. 1989). Some of the nitrogen fixation genes are usually on the same operon in all species while some genes will have different groupings.

In addition, at least three different nitrogenases have been discovered in non-photosynthetic diazotrophs (Bishop and Joerger 1990, Bishop et al. 1988, Eady et al. 1988, Pau et al. 1989). While all three nitrogenases contain iron and sulfur, they vary in their metal requirement at the reaction center. The most studied nitrogenase requires molybdenum (Mo). Another requires vanadium (V) and the third uses iron (Fe).

While these nitrogenases have different metal requirements, they have some similar characteristics. All three have an iron protein (nitrogenase reductase) (Bishop and Joerger 1990, Bergström et al. 1988, Joerger et al. 1990, 1989). The role of nitrogenase reductase is to provide electrons to nitrogenase for dinitrogen reduction. Nitrogenase reductase is a dimer of identical 30,000 Da subunits. The genes for the subunits are *nifH* for the Mo-nitrogenase, *vnfH* for the V-nitrogenase, and *anfH* for the Fe-nitrogenase. The nitrogenases themselves also have some similarities as well as differences. All three have a core composed of two α subunits and two β subunits. The α and β subunits each have a molecular mass of about 61,000 Da. The

α subunits are encoded by *nifD* for the Mo-nitrogenase, *vnfD* for the V-nitrogenase, and *anfD* for the Fe-nitrogenase. The genes for the β subunits are *nifK*, *vnfK*, and *anfK*. The V and Fe-nitrogenases have an additional subunit δ, which is not present in the Mo-nitrogenase (Bishop and Joerger 1990, Joerger et al. 1990, 1989). The genes for the δ subunits are *vnfG* and *anfG*. Thus, the subunit composition for the Mo nitrogenase is $α_2β_2$ and for the alternative nitrogenases it's $α_2β_2δ_2$. The expression of one nitrogenase instead of another depends on the availability of the required metal. This is demonstrated by the fact that in the absence of combined nitrogen, if V is available but Mo is not, the V-nitrogenase will be expressed. In addition, mutants lacking the Mo nitrogenase genes only grow under nitrogen fixing conditions if V is in the media (Bishop et al. 1986, 1980).

By growing bacteria in a medium lacking a certain combination of these metals, it was discovered that some species have all three nitrogenases while others have only two. Of the three nitrogenases, only the Fe-nitrogenase is not found in all bacteria (Bishop and Joerger 1990). The role of the alternative nitrogenases is believed to allow the bacteria to function in soils which may lack these required metals. In cyanobacteria, there is evidence that *Anabaena variabilis* may have V and Fe-nitrogenases, since it

can still grow in media lacking combined nitrogen and Mo if
V is in the media or when Mo and V are absent (Kentemich et
al. 1991, 1988a, 1988b). While the genes for these
alternative cyanobacterial nitrogenases have not been
isolated, evidence for them lies in the fact that some
cyanobacteria have more than one copy of *nifH*-like genes
(Defrancesco and Potts 1988). Among the three known
nitrogenases, the genes for the nitrogenase reductase are
more conserved than those for the α and β subunits of the
nitrogenase itself, suggesting that the other *nifH*-like
genes in some cyanobacteria are for alternative nitrogenases
(Premakumar et al. 1989). Additional evidence comes from
data showing that in *A. variabilis* one of the extra *nifH*-
like copies hybridizes with an *anfH* DNA probe (Kentemich et
al. 1991).

In nitrogen fixing bacteria, the differences in the
types of nitrogenases present and the arrangement of their
genes may reflect a need to customize the nitrogen fixation
system to suit each species' environment and life cycle.
This would insure proper functioning of the nitrogen
fixation systems under the different conditions encountered
by each species.

The organization and regulation of *nif* genes in
cyanobacteria is not as well studied as in the non-

photosynthetic nitrogen fixing bacteria. *Anabaena* sp. PCC 7120 is the most studied cyanobacterium with respect to nitrogen fixation and subsequently most of the sequence data for cyanobacterial *nif* genes are derived from this species.

This study describes the organization and novel features of the *nifUHD* region in addition to characterization of GlbN and Western analysis of GlbN expression in *N. commune* UTEX 584 and two other species of cyanobacteria. In addition, a preliminary study of the life cycle of *N. commune* UTEX 584 during growth in nitrogen-free media by air lift fermentation is presented in the appendix.

## Materials and Methods

**Growth of *N. commune* UTEX 584 and *Anabaena* sp. PCC 7120**

For protein isolation, *N. commune* UTEX 584 was usually grown in shaker flasks at 32°C on a 16-hour light/8-hour dark cycle in 100 ml of BG11 or $BG11_o$ (Rippka et al. 1979). BG11 is the same as $BG11_o$ except that $BG11_o$ does not contain $NaNO_3$. The composition of BG11 is: 1.5 g/l $NaNO_3$, 0.04 g/l $K_2HPO_4 \cdot 3H_2O$, 0.075 g/l $MgSO_4 \cdot 7H_2O$, 0.036 g/l $CaCl_2 \cdot 2H_2O$, 0.006 g/l citric acid, 0.006 g/l ferric ammonium citrate, 0.001 g/l $Na_2Mg$ EDTA, 0.02 g/l $Na_2CO_3$, and 1 ml of trace metals per liter of media. Trace metals contains, 2.86 g/l $H_3BO_3$, 1.81 g/l $MnCl_2$, 0.222 g/l $ZnSO_4 \cdot 7H_2O$, 0.39 g/l $Na_2MoO_4 \cdot 2H_2O$, 0.079 g/l $CuSO_4 \cdot 5H_2O$, and 0.0494 g/l $Co(NO_3)_2 \cdot 6H_2O$. After 7 days in BG11, the cells were shifted to $BG11_o$ by washing six times in $BG11_o$ with 50 ml per wash, then grown in 100 ml $BG11_o$ for at least 2 days before harvesting for protein isolation.

*Anabaena* sp. PCC 7120 was maintained in $BG11_o$ in shaker flasks at 32°C as described above for *N. commune* UTEX 584. The cultures were then transferred to fresh $BG11_o$, as described above, for at least 2 days, before harvesting for protein isolation.

For one set of experiments, *N. commune* UTEX 584 and *Anabaena* sp. PCC 7120 were maintained in $BG11_o$ from which

15

$Na_2MoO_4 \cdot 2H_2O$ was omitted from the trace metals solution. This medium is referred to as "Mo-free" and was supplemented with 100 nM $VSO_4$. It is assumed that the vanadium might be able to overcome the effects of trace contamination of molybdenum from the reagents or glassware used to make the media (Dr. Anneliese Ernst, personal communication). Rigorous treatments to remove trace amounts of contaminating molybdenum from the reagents and glassware were not performed.

For the life cycle studies described in the appendix, *N. commune* UTEX 584 was grown in a BRL Air Lift fermenter in $BG11_o$ at room temperature. The cultures were grown in continuous light adjusted between 1 to 100 µE $m^{-2}s^{-1}$ and were sparged continuously with air at a rate of 1 to 1.5 liters/minute while automatically maintaining the culture volume at 1.8 liters.

**Protein Extraction from *N. commune* UTEX 584 and *Anabaena***

The cells were pelleted by centrifugation at 5,000 x *g* for 15 minutes in sterile 50 ml conical centrifuge tubes. The supernatants were poured off and the cell pellets were frozen at -75°C for 30 minutes and lyophilized overnight. The lyophilized cells were frozen with liquid nitrogen and ground to a powder in a mortar. The powder was resuspended in 0.5 ml of sample buffer (80 mM Tris pH 6.8, 2% w/v SDS,

10% v/v glycerol, 1.7% v/v β-mercaptoethanol, 0.00024% w/v
bromophenol blue in methanol) and subjected to additional
grinding. The resulting lysates were placed into 1.5 ml
microcentrifuge tubes and boiled for 5 minutes. The boiled
samples were centrifuged at 12,000 x $g$ at room temperature
for 15 minutes. The supernatants were divided into 100 µl
aliquots and stored at -20˚C.

## Plasmid Isolation

Several procedures were employed, depending on the
application for the plasmid preparation. To isolate plasmids
for double stranded sequencing and other applications, a
modification of the procedure of Birnboim and Doly was used
(Birnboim and Doly 1979). For a 100 ml culture this was as
follows. The cells were pelleted by centrifuging at 3,000 x
$g$ at 4˚C for 10 minutes. After centrifuging, as much of the
supernatant as possible was removed using a pipet when
needed. The cell pellet was resuspended by vortexing in 4 ml
of GTE (50 mM glucose, 25 mM Tris pH 8.0, 10 mM EDTA pH 8.0)
with 50 mg lysozyme and 50 µl of 10 mg/ml RNase A and
incubated at room temperature for 30 minutes. After the 30
minute incubation, 8 ml of freshly made NaOH-SDS solution
(0.2 M NaOH, 1% w/v SDS) was added, mixed by inverting and
placed on ice (0˚C) for 10 minutes. After 10 minutes at 0˚C,
6 ml of 0˚C 7.5 M ammonium acetate was added, mixed by

inverting and placed at 0°C for 10 minutes, then centrifuged at 12,000 x $g$ at 4°C for 15 minutes. The supernatant was transferred to a new tube trying not to collect any of the white particles (membranes, denatured proteins and genomic DNA) floating in the suspension. The pellet was discarded and the supernatant was centrifuged at 12,000 x $g$ or higher for 15 minutes or longer to insure that all of the white particulate material was pelleted. The clear supernatant was transferred to a new tube and the pellet was discarded. To the supernatant, 0.6 volumes of isopropanol was added at room temperature, mixed by inverting, incubated at room temperature for 15 minutes and centrifuged at 12,000 x $g$ at room temperature for 15 minutes. The supernatant was discarded, the pellet was lyophilized and redissolved in 200 µl of TE buffer pH 8.0 (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) and 100 µl of 7.5 M ammonium acetate. This suspension was transferred to a sterile 1.5 ml tube, placed at 0°C for 20 minutes, and centrifuged at 12,000 x $g$ at 4°C for 15 minutes. The resulting supernatant was transferred to a new 1.5 ml tube, mixed with 600 µl of 95% v/v ethanol and incubated at room temperature for 10 minutes followed by centrifuging at room temperature at 12,000 x $g$ for 15 minutes. The supernatant was discarded and the pellet was

washed once with -20°C 80% v/v ethanol, vacuum dried and resuspended in sterile ddH$_2$O.

This procedure was also scaled down for preparing plasmids from smaller cultures. In this case, it was possible to do several preparations at the same time.

For rapid screening of many clones, the following procedure was used. From a 5 ml overnight culture of the desired colony, 3 ml was centrifuged at 12,000 x $g$ at room temperature for 5 minutes. As much of the supernatant as possible was poured off and the remainder was removed with a pipette. The cell pellet was resuspended by vortexing in 200 µl of STET buffer (8% w/v sucrose, 5% v/v Triton X-100, 50 mM EDTA pH 8.0, 50 mM Tris pH 8.0, 0.025 mg/ml RNAse A and 3 mg/ml lysozyme) and incubated at room temperature for 15 minutes followed by boiling for 40 seconds. After boiling, the sample was immediately centrifuged at 12,000 x $g$ at room temperature for 15 minutes. After centrifuging, 160 µl of the supernatant was placed into a new tube with 240 µl of NaCl-PEG (0.67 mM NaCl, 10.8% w/v PEG 8,000) and mixed by vortexing. After mixing, the tube was placed at 0°C for 45 minutes. The DNA was pelleted by centrifugation at 12,000 x $g$ at 4°C for 15 minutes. The supernatant was discarded and the DNA pellet was washed twice with -20°C 70% v/v ethanol, vacuum dried and resuspended in 50 µl of sterile ddH$_2$O.

Usually 1 µl of such a preparation was sufficient to analyze on an agarose gel or to use for restriction analysis.

Isolation of single stranded M13 DNA for sequencing was performed as described in the Promega Protocols and Applications Guide (Promega 1989/90).

## Single Stranded RNA Synthesis

Single stranded RNA was made using T7 or SP6 RNA polymerases (Promega) as described by the manufacturer. The plasmids used as templates for RNA synthesis were digested with the appropriate restriction enzyme to prevent ssRNA concatamers of the plasmid from being synthesized. Radioactive ssRNA probes and unlabeled ssRNA were purified as follows. To a 50 µl reaction, 1 µl of RQ1 DNase (1 U/µl, Promega) was added and the mixture was incubated at 37˚C for 15 minutes. After the 15 minute incubation, 5 µl of 3 M sodium acetate was added with 50 µl of equilibrated phenol and 50 µl of CIA followed by vortexing. The phases were separated by centrifugation at 12,000 x $g$ at room temperature for 2 minutes. The aqueous phase was transferred to a new tube and extracted once with 100 µl of CIA and centrifuged as above to separate the phases. The aqueous phase was transferred to a new tube, mixed with 137 µl of 95% v/v ethanol and then placed at -75˚C for 30 minutes. The sample was then centrifuged at 12,000 x $g$ for 30 minutes at

4°C. The supernatant was discarded and the pellet was washed twice with -20°C 70% v/v ethanol and vacuum dried. The dried pellet was resuspended in 25 µl of sterile 0.1x SSC (0.015 M NaCl, 0.0015 M sodium citrate pH 7.0) and stored at -75°C.

**Deletion Clone Construction**

The *Hind*III-*Eco*RI 2.4 kb fragment was ligated into pBluescript KS (Stratagene) giving the construct called pNc*nif*1 (Figure 3C). Deletion clones were made from this construct with exonuclease III (Exo III) using a modification of the procedures of Henikoff (Henikoff 1987, 1984) as follows. Approximately 5 µg (12 µl) of *Apa*I-*Hind*III digested pNc*nif*1 was mixed with 6 µl ddH$_2$O, 3 µl 10x Exo III buffer (BRL Life Technologies, Inc,: BRL), 3 µl 140 mM β-mercaptoethanol, and warmed to 37°C for 5 minutes. After warming, 6 µl of 65 Units/µl Exo III (BRL) was added. After addition of Exo III, 5 µl aliquots were removed at 30 second intervals. Each aliquot was placed into its own tube with 35 µl of 0.1x SSC and 10 µl of equilibrated phenol at room temperature. After collecting 6 aliquots, 10 µl of CIA was added to each aliquot and vortexed. After centrifuging at 12,000 x *g* for 2 minutes, each aqueous phase was extracted with 20 µl of CIA in a new tube. From each CIA extracted aliquot, 40 µl of the aqueous phase was transferred into its own new tube and mixed with 60 µl of NaCl-PEG, placed at 0°C

Figure 3. Location of genes on the 3.5 kb *Hin*dIII fragment
and on the 2.4 kb fragment cloned into several different
plasmids and on pT7ORF18.

A shows the proposed location of the *nif* genes on the
3.4 kb fragment before it was completely sequenced. B shows
the construct pNDH1 which contains the 2.4 kb *Hin*dIII-*Eco*RI
fragment from the 3.5 kb fragment cloned into pGEM-3. The
pNDH1 construct was used to make ssRNA. C represents the
construct p*Ncnif1* which was used to make deletion clones of
the 2.4 kb *Hin*dIII-*Eco*RI removed from pNDH1. D shows the
configuration of pT7ORF18. RBS is the ribosome binding site.
The T7 promoter is about 50 bases upstream of the ATG start
codon. The *Hin*dIII site is about 50 bases downstream of the
stop codon.

## A



Hind III    EcoR I    Hind III

*nifU*    *nifH*    *nifD*

220 bp

## B



Hind III    EcoR I

T7→    *nifU*    *nifH*    ←SP6

## C



Apa I    Hind III    EcoR I

T3    *nifU*    *glbN*    *nifH*    T7

220 bp    Pvu I    Pvu I    Pvu I

## D



Nde I    Hind III

T7    RBS    CATATG    *glbN*    TAATAG    AAGCTT

Met    Stop
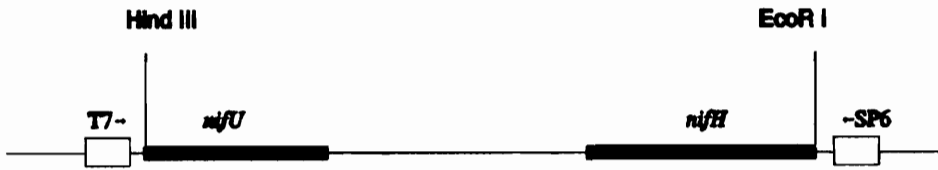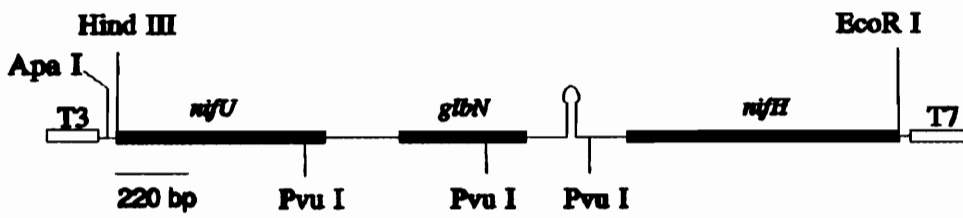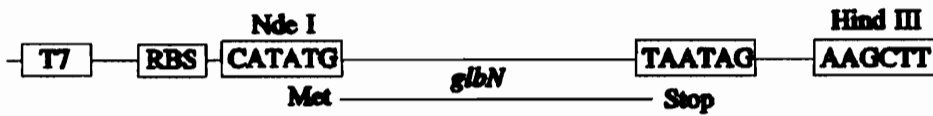
for 45 minutes followed by centrifuging at 12,000 x $g$ at 4°C
for 10 minutes. The supernatants were discarded and the
pellets were washed once with -20°C 70% v/v ethanol and
vacuum dried. Each dried pellet was resuspended in 40 µl of
1x mung-bean nuclease buffer (Promega). After resuspending,
1 µl was removed from each tube for analysis on 1% w/v
agarose gels. To each tube, 1 µl of 5 units/µl mung-bean
nuclease (Promega) was added followed by an incubation at
37°C for 30 minutes. After the 30 minute incubation, each
tube was extracted with 1:1 phenol:CIA, CIA, NaCl-PEG
precipitated, washed and vacuum dried as described above.
Each pellet was resuspended in 25 µl of ligation mix (178 µl
$H_2O$, 20 µl 10x ligation buffer (Promega), 2 µl 100 mM DTT)
from which 5 µl was removed for analysis on agarose gels. To
each tube 1 µl of T4 mix (4 µl of 1x ligation buffer
(Promega), 2 µl T4 DNA ligase (Promega)) was added then
incubated at room temperature for 4 hours followed by
storage at -20°C until needed for transformations.

*E. coli* DH5α cells were made competent for
transformation using the Hanahan procedure (Hanahan 1985).
These cells were transformed with the whole ligation
reaction and grown at 37°C on Luria-Bertani plates with 100
µg/ml ampicillin (LBA plates). Colonies were picked and
streaked onto fresh LBA plates with up to 50 colonies per

plate. These clones were screened using the rapid plasmid isolation procedure described above.

## Sequencing and Sequence Analysis

Sequencing was accomplished using the United States Biochemical Co. Sequenase kit according to the manufacturer's directions using $[\alpha\text{-}^{35}S]dATP$. The sequencing gels were run on a BRL model S2 apparatus using 5% w/v or 6% w/v polyacrylamide wedge gels with TBE buffer (Maniatis et al. 1982).

Sequence analysis was performed using the IntelliGenetics program PCGENE.

## Oligonucleotides

Oligonucleotides for sequencing and PCR were made by Dr. T. Larson on an Applied Bioystems Model 381A DNA Synthesizer. The oligonucleotides were deprotected and cleaved from the synthesis columns according to the Applied Biosystems manual. After deprotection and cleavage, the oligonucleotides were purified on Perkin Elmer Cetus Pure-Pak cartridges according to the manufacturer's directions. The oligonucleotides used are listed in Table I with their corresponding bases on the 3.5 kb sequence shown in Figure 5.

Table I. List of oligonucleotides used.

These are the oligonucleotides that were made for sequencing and PCR work. The base positions represent the locations of the oligonucleotide on the sequence in figure 5. Oligonucleotides used for PCR have only part of their sequence complementary to the 3.5 kb fragment. The complementary portion is underlined.

Table I. List of oligonucleotides used.

| serial # | oligonucleotide sequence | base positions |
|----------|--------------------------|----------------|
| 272-839 | 5'AGATTCGTCGCGTGATTC 3' | 196→213 |
| 259-445 | 5'TGATGTAGAAGGTGATCG 3' | 494→511 |
| 245-965 | 5'ATCAATGTTCGCTAGACA 3' | 299→282 |
| 250-257 | 5'GATTTACCGATACCGCCT 3' | 1,652→1,635 |
| 298-644 | 5'CACGCAGAATTGCGCCGG 3' | 2,257→2,274 |
| 208-160 | 5'GGGTCGCAACCGACG 3' | 1,724→1,710 |
| 239-427 | 5'CACGAACCATGTCGT 3' | 3,252→3,238 |
| 253-808 | 5'ATCATCGAAGAAAGA 3' | 2,661→2,675 |
| 297-392 | 5'GGTACCATGCACTTC 3' | 2,970→2,984 |
| 229-591 | 5'CTGGCA<u>TATGAGCACATTGTACG</u> 3' | <u>888→904</u> |
| 232-493 | 5'GCGAAGC<u>TTCGATAGGACTTGAT</u> 3' | <u>1,296→1,281</u> |

**RNA Isolation and Northern Transfers**

RNA was isolated from *N. commune* UTEX 584 using the
following procedure. The cells were collected from the
fermenter, at various stages of the life cycle, into about
50 ml of sterile ddH$_2$O at 0°C and partly frozen (was stored
at -20°C for 3 hours before using). About 150 ml of culture
was collected and the cells were pelleted by centrifuging at
10,000 x *g* for 10 minutes at 4°C. The supernatant was
discarded and the cell pellet was resuspended in 50 ml
sterile 50 mM Tris pH 8.1 at 0°C and centrifuged as above.
This wash was repeated two more times to remove any EDTA or
other chelating agents which could inactivate the vanadyl-
ribonucleoside complex (VRC) used in the next step. In a
chilled mortar, 5 grams of the wet cell pellet was mixed
with 10 ml of lysis buffer (50 mM Tris pH 8.1, 20 mM VRC, 30
mM MgCl$_2$, 120 mM β-mercaptoethanol). Liquid nitrogen was
added to freeze the cells and they were ground to a powder.
The frozen lysate was allowed to thaw, then mixed with 2 ml
of detergent solution (6% v/v NP-40, 6% v/v Tween 20 and 6%
w/v Brij 35 sterilized by autoclaving) and placed at 0°C for
10 minutes. To this mixture was added 15 ml of phenol
equilibrated with 0.1 M Tris (made with Trizma base (Sigma)
without adjusting the pH) which was prewarmed to 50°C. The
tubes were placed at 50°C for 10 minutes with frequent

mixing. Following the 50˚C incubation, 1 ml of 10x STE (1 M NaCl, 100 mM Tris pH 7.4, 10 mM EDTA) and 7 ml of sterile ddH$_2$O were added and mixed by vortexing. After vortexing, 15 ml of CIA was added and mixed well at 50˚C. The phases were separated by centrifuging at 1,000 x $g$ at room temperature for 5 minutes. The aqueous phase was retrieved trying not to collect any of the material at the interface. The aqueous phase was extracted two more times with 30 ml of 1:1 phenol:CIA at 50˚C, then extracted once with 20 ml of CIA at room temperature. The aqueous phase was transferred to a new tube to which 2.5 volumes of 95% w/v ethanol was added and placed at -75˚C for 30 minutes. Following centrifugation at 12,000 x $g$ at 4˚C for 30 minutes, the supernatant was discarded and the RNA pellet was washed with -20˚C 80% v/v ethanol, vacuum dried and resuspended in sterile ddH$_2$O. Additional purification on CsCl gradients or with RNase-free DNase was performed as needed to remove any DNA. RNase on all glassware and plasticware, or in solutions was inactivated by baking, diethylpyrocarbonate treatment or siliconizing as needed (siliconization with Prosil-28 from PCR Inc. Gainesville, FL, was an effective means for inactivating RNase). VRC was made as described in Maniatis et al. (1982).

RNA samples for Northern transfer were denatured and run on formaldehyde/formamide agarose gels (1% w/v agarose) as recommended by BRL (see Anonymous 1986). Native RNA was analyzed on 1% w/v agarose gels made with TAE buffer (Maniatis et al. 1982).

## Construction of the *glbN* Expression Plasmid

Copies of *glbN* were made by PCR with a Coy Tempcycler Model 50 for 30 cycles using conditions determined according to Innis and Gelfand (1990). During the PCR, an *Nde*I site was placed at the 5' ATG start codon of the gene using oligonucleotide # 229-591. Approximately 50 bases past the TAG stop signal, a *Hin*dIII site was placed using oligonucleotide # 232-493. The PCR product did not digest efficiently enough with *Nde*I and *Hin*dIII to allow cloning directly into the *Nde*I-*Hin*dIII digested expression vector pT7-7. The PCR product was therefore blunt end ligated into pBluescript KS digested with *Sma*I, using the procedure described by Scharf (1990). The PCR product was easily digested from pBluescript with *Nde*I and *Hin*dIII and isolated for ligation into pT7-7 previously digested with *Nde*I and *Hin*dIII. The ligation reaction of the *Nde*I-*Hin*dIII PCR fragment into pT7-7 was used to transform competent *E. coli* DH5α cells (Hanahan 1985) to screen for the proper construct and to isolate large quantities of the plasmid for

sequencing and transformation of the *E. coli* expression host
BL21 DE3. The resulting plasmid, pT7ORF18, is shown in
figure 3D. Sequencing of pT7ORF18 with a T7 primer and with
oligonucleotide #232-493 as a primer, showed that the
sequence of the PCR copy of the gene was identical to that
of the original gene.

**Expression of the PCR Copy of *glbN* and Purification of GlbN**

The T7 RNA polymerase-directed expression vector (pT7-
7) was used to overexpress GlbN in *E. coli* BL21 DE3 ($F^-$,
*ompT*, $r_B^- m_B^-$)(Studier et al. 1990). This strain lacks the
*lon* protease, the *ompT* outer membrane protease, and contains
the T7 RNA polymerase gene under the control of the lacUV5
promoter on a DE3 lysogenic $\lambda$-phage insert (Studier et al.
1990). This expression host was transformed with pT7ORF18
using the Hanahan procedure (Hanahan 1985). Individual
colonies were picked and grown at 37°C in LBA to log phase
($OD_{550} \approx 0.4$). At log phase, IPTG was added to 0.4 mM and the
cells were allowed to grow for another 2 hours. The cells
were pelleted by centrifugation at 5,000 x *g* for 10 minutes
and washed several times in 50 mM HEPES pH 7.5, then lysed
by passage through a chilled French pressure cell at 16,000
pounds per square inch in 50 mM HEPES pH 7.5, 1 mM PMSF
(made with 20 mM PMSF in 100% v/v methanol). The cellular
debris was removed by centrifugation at 30,000 x *g* at 4°C

for 30 minutes. The resulting supernatant had an absorbance maximum at 418 nm. The supernatant was passed through a Mono Q HR 5/5 anion exchange column (Pharmacia LKB) in 50 mM HEPES (pH 7.5) while monitoring at 280 nm. Under these conditions GlbN did not bind to the column and was eluted in the void volume. This fraction was collected and concentrated using a Centriprep 10 concentrator (Amicon, with a molecular mass exclusion limit of 10,000 Da). The concentrated fraction was applied to a Superose 12 HR 10/30 gel filtration column (Pharmacia LKB) in 50 mM HEPES at pH 7.5 while monitoring at 280 nm and 418 nm. The 418 nm peak fractions which contained GlbN had a characteristic red color. Protein concentrations were determined using the Pierce Coomassie protein assay kit according to the manufacturer's directions.

Determination of the native mobility of recombinant GlbN was performed on a Superose 12 HR 10/30 gel filtration column in 50 mM HEPES pH 7.5 while monitoring at 280 nm and 418 nm. The native molecular mass standards used for calibration of the column were ovalbumin (45,000 Da), carbonic anhydrase (29,000 Da), and cytochrome C (12,400 Da).

**Production of Antisera to GlbN**

Purified fractions of GlbN from the Superose 12 column were used to inject rabbits for antibody production using a

modification of the procedures of Harlow and Lane (Harlow
and Lane 1989). Two male New Zealand white rabbits were used
(identification numbers 762 and 763). The procedures for
maintaining, handling, inoculating, and bleeding the rabbits
were approved by the Lab Animal Resources committee of
Virginia Polytechnic Institute and State University. Before
inoculation, 15 ml of preimmune sera was collected from each
rabbit. The rabbits were inoculated with 0.5 ml of 0.5 mg/ml
partially purified protein mixed with 0.5 ml of complete
Freund's adjuvant. Both rabbits were given intramuscular
injections at four different sites (the triceps or trapezius
muscles of each front leg and near the biceps femoris of
each hind leg) with 250 µl per site. Eleven days later,
rabbit 762 received 3.5 mg more of partially purified GlbN
in incomplete adjuvant and rabbit 763 received 0.44 mg.
Eleven days later, 15 ml of blood was taken from each
rabbit, allowed to clot at room temperature and stored at
4°C overnight. The next morning the clotted blood was
centrifuged at 10,000 x $g$ for 10 minutes. The supernatants
were divided into 1 ml aliquots and stored at -20°C. These
antisera did not react well with SDS denatured GlbN, so the
rabbits were injected with purified GlbN (0.45 mg for rabbit
762 and 0.9 mg for rabbit 763) in incomplete adjuvant as
described above. Ten days later both rabbits were completely

bled. Approximately 140 ml of whole blood was obtained from each rabbit. The blood was allowed to clot at room temperature and stored at 4°C overnight. The coagulated blood was then centrifuged at 4,000 x $g$ at 4°C for 10 minutes. The supernatants were placed into new 50 ml sterile conical tubes and centrifuged again as above. These supernatants were transferred to new 50 ml sterile conical tubes and stored at -20°C.

The antisera from the terminal bleeding reacted well with native and SDS denatured GlbN and were therefore further purified by ammonium sulfate precipitation using a modification of the procedure of Harlow and Lane (1988). To 20 ml of serum at room temperature in a 50 ml sterile conical tube, 6.25 g of ammonium sulfate was gradually added (final concentration 50% w/v ammonium sulfate) while vortexing gently at room temperature until the ammonium sulfate completely dissolved. After dissolving the ammonium sulfate, the tube was stored at 4°C overnight followed by centrifugation at 3,000 x $g$ at 4°C for 30 minutes. The supernatant was discarded and the pellet was drained to remove as much supernatant as possible. The pellet was resuspended in 6 ml of PBS (1 mM $NaH_2PO_4$, 10 mM $Na_2HPO_4$, 150 mM NaCl, and 1 mM $MgCl_2$) and vortexed gently until dissolved. The dissolved pellets were desalted by using Bio-

Rad Econo-Pac 10DG columns as follows. After each column was equilibrated with 30 ml of PBS, 3 ml of the redissolved pellet solution was loaded onto each column and allowed to flow into the column without collecting the flow through. To each column, 4 ml of PBS was added and the first 4 ml of flow through, which contains the desalted antisera, was collected. Each column was washed 5 times with 20 ml of PBS per wash (to remove the ammonium sulfate retained in the column) and the remainder of each sample was desalted as described above. The desalted antisera from each rabbit was pooled with the first desalted fraction from that rabbit, diluted to 15 ml with PBS plus 30 µl of 10% w/v $NaN_3$ and stored at -20°C until ready to concentrate. The desalted antisera were then thawed and concentrated with Centriprep 30 concentrators (Amicon, with a molecular mass exclusion of 30,000 Da) by centrifuging at 1,500 x $g$ at 4°C. The antisera from rabbit 762 was concentrated to 3.6 ml and that of rabbit 763 was concentrated to 2.2 ml. These antisera preparations were divided into 200 µl aliquots and stored at -75°C.

**Protein Gels and Western Blotting**

SDS polyacrylamide gels of 15% w/v, 12% w/v, or 10% w/v were made according to a Bio-Rad Protein II procedure (after Laemmli, 1970) and were silver-stained or Coomassie-stained.

Silver staining was performed using a modification of a procedure provided by Dr. T. Keenan (personal communication) which is similar to that described by Wray et al. (1981). The gels were fixed overnight with gentle agitation in 50% v/v methanol and then washed in autoclaved ddH$_2$O with 3 changes at 5 minutes per wash. The wash water was poured off and the gels were submerged in pre-developer (described below) and gently agitated for 30 minutes at room temperature. The pre-developer was poured off and the gels were washed with 3 changes of autoclaved ddH$_2$O for a total wash time of 5 minutes. The wash water was poured off and the gels were covered with the developing solution and gently agitated at room temperature. After adding developer, the bands started to appear within 10 to 15 minutes. The reaction was stopped before the background staining became too dark, by decanting the developer and washing the gels in 20% v/v methanol.

The pre-developer was prepared in a 100 ml graduated cylinder by mixing 21 ml of 0.36% w/v NaOH with 1.4 ml of 38% v/v NH$_4$OH. To the NaOH-NH$_4$OH solution, 4 ml of freshly made silver nitrate (0.1 g AgNO$_3$/ml ddH$_2$O) was added in a drop-wise fashion while gently mixing. After all the silver nitrate was added, the solution was diluted to 100 ml with autoclaved ddH$_2$O. The pre-developer was made fresh and used within 10 minutes of being made. The developer was prepared

just before using by mixing 1 ml of 1% w/v citric acid with 0.1 ml of 37% v/v formaldehyde and diluting to 500 ml with autoclaved ddH$_2$O.

Coomassie staining was accomplished by placing the gels in Coomassie stain for 15 minutes followed by destaining overnight in a 10% v/v acetic acid, 10% v/v methanol solution. Foam packing cut into small squares was placed in the destain solution to aid the destaining processes. Coomassie stain was made according to a Bio-Rad procedure by dissolving 1 gram of Coomassie Brilliant Blue R250 in 180 ml of methanol, adding 36 ml glacial acetic acid and diluting to 400 ml with ddH$_2$O.

Blotting onto polyvinylidene fluoride microporous Immobilon-P membranes (Millipore) was done using a Hoefer TE series electrophoresis unit at 250 mA for 2 to 3 hours in millipore transfer buffer (15% v/v methanol, 25 mM Tris, and 192 mM glycine, pH 8.3). After transfer of the proteins, the membranes were blocked at room temperature for 60 minutes in 50 ml of blocking buffer (20 mM Tris pH 7.5, 100 mM NaCl, and 1% w/v BSA). The membranes were washed twice with 50 ml of TTBS (20 mM Tris pH 7.5, 100 mM NaCl, and 0.05% v/v Tween 20) at room temperature for 5 minutes each wash. The wash buffer was poured off and 50 ml of 1:10,000 antisera in antibody buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.05% v/v Tween 20, and 0.1% w/v BSA) was added and gently agitated at

room temperature overnight. After the overnight incubation, the membranes were washed with TTBS as described above and then treated for 3 hours with 50 ml of 1:3,000 goat anti-rabbit horseradish peroxidase conjugate (Bio-Rad) in antibody buffer. The membranes were washed twice with TTBS as described above, then washed another 3 times with TBS (20 mM Tris pH 7.5, 100 mM NaCl) to remove the Tween-20. The washed membranes were placed into a clean Petri dish, fresh color development solution (described below) was added and allowed to develop at room temperature for 30 minutes. Color development was stopped by washing the membranes with $ddH_2O$. While the membranes were still wet, the molecular mass marker lanes were cut from the membranes and stained for 10 minutes with amido black (Sigma: 1% w/v amido black in 35% v/v acetic acid) diluted 1:10 in $ddH_2O$, then washed with $ddH_2O$ and air dried. The remainder of the membrane was air dried and photographed.

The color development solution was made using a modification of the Bio-Rad procedure by dissolving 60 mg of the Bio-Rad horseradish peroxidase color development reagent (4-chloro-1-naphthol) in 20 ml of methanol at -20°C. To 100 ml of TBS at room temperature, 60 µl of 30% v/v $H_2O_2$ was added and mixed well. The two solutions were immediately mixed together and poured over the membranes as described above.

## Microscopy Techniques

For the life cycle study described in the appendix, the cells were observed by placing a few drops of culture on a microscope slide and placing a cover slip over the drops. The cells were not fixed or stained. The cells were photographed on Kodak Panatomic-X Film (35 mm) with an Olympus Vanox microscope using a 40x Plan Acromat objective with an Olympus OM2 camera body.

## Chlorophyll *a* Determinations

During the life cycle studies described in the appendix, chlorophyll *a* (Chl *a*) concentrations were determined by collecting 50 ml of culture in a preweighed tube and pelleting the cells by centrifugation at 10,000 x *g* at room temperature for 5 minutes. The supernatant was decanted and excess media was removed from the inside of the tube with a cotton swab. After the tube was weighed to determine the mass of the cell pellet, 10 ml of 100% v/v methanol was added along with 10 to 12 glass beads (5 mm in diameter). The sample was vortexed to disrupt the cell pellet and stored at -20°C overnight. The sample was then warmed to room temperature and filtered through a Whatman #4 paper filter into a 25 ml flask. The Chl *a* concentration was determined in a cuvette with a 1 cm light path at 665 nm where 1 $OD_{665}$= 13.42 µg Chl *a*/ml (the procedure for

determining the Chl *a* concentration was kindly provided by

Dr. K. Jäger and is similar to that of De Marsac and Houmard

1988).

## Results

### Sequence Analysis and Comparisons

A 3.5 kb *Hind*III genomic fragment from *N. commune* UTEX 584 was previously isolated and partly sequenced (Defrancesco and Potts 1988). Based on this partial sequence information, it was determined that this fragment contained part of the nitrogen fixation genes *nifU* and *nifD* in addition to all of *nifH* (Figure 3A). From this 3.5 kb fragment a 2.4 kb *Hind*III-*Eco*RI fragment containing part of *nifU* and most of *nifH* was subcloned into pGEM-4 to give the construct called pNDH1 (Figure 3B). Using this construct with T7 or SP6 RNA polymerase, it was possible to make ssRNA from the insert. The ssRNA made with SP6 is complementary to the mRNA of these genes. When pNDH1 was digested with *Pvu*I, a probe specific for mRNA transcripts containing *nifH* was made with SP6 RNA polymerase. Using this probe labeled with $^{32}$P, it was possible to detect mRNA transcripts containing *nifH*, but the results were dark smears with a few distinct bands (data not shown). Of the bands that were visible on this Northern transfer, some were larger than 2.4 kb and some were smaller than 2.4 kb. It was not possible to determine their exact size.

An additional finding was that unlabeled ssRNA, synthesized with SP6 RNA polymerase on *Hind*III linearized

41

pNDH1, produced an unexpected shorter transcript of about 1 kb in addition to the expected 2.4 kb transcript (Figures 4A lane 4 and 4B lane 5). The shorter transcript obtained with SP6 RNA polymerase, indicated that occasionally transcription terminated near the 5' end of *nifH*. With T7 RNA polymerase only the expected 2.4 kb transcript was seen with *Eco*RI linearized pNDH1 (Figure 3A lane 3).

To determine more precisely the locations of the genes on the 3.5 kb fragment, and understand why there may be several transcripts containing *nifH* in *N. commune* UTEX 584 and why SP6 RNA polymerase was producing the shorter transcript, the entire 3.5 kb fragment was sequenced using deletion clones of the 2.4 kb fragment and by sequencing the 1.1 kb *Eco*RI-*Hin*dIII fragment in M13 and with oligonucleotide primers. The nucleotide sequence of the entire 3.5 kb fragment is shown in Figure 5. From these data the sequence for all of *nifH*, and the parts of *nifU* and *nifD* present on the 3.5 kb fragment were obtained. An additional potential ORF was also found between *nifU* and *nifH*. As will be discussed, this ORF was given the name *glbN*. Other interesting features on the 3.5 kb fragment include two potential stem loop structures and a direct repeat (Figures 5 and 6). The PCGENE program HAIRPIN determined that both stem loops have ΔG values of about -15 kcal/mol and are

Figure 4. Ethidium bromide stained 1% agarose gels of SP6 and T7 ssRNA made from pNDH1.

Gel A, lanes 1 and 6 are *E. coli* rRNA markers. Lane 2 is RNA made on a control template with T7 RNA polymerase. Lane 3 is T7 ssRNA made with *Eco*RI linearized pNDH1. Lane 4 is SP6 RNA made on *Hind*III linearized pNDH1. Lane 5 is rat liver RNA. The sizes of the markers are as follows: 28s is 4.72 kb, 23s is 2.9 kb, 18s is 1.87 kb, 16s is 1.3 kb, and 4s is about 80 to 100 bp. The control template produces bands of 1.4 kb, 0.68 kb and 0.043 kb. Only the 1.4 kb band is clearly visible in this photo. On the original photo, the 0.68 kb band was also visible.

Gel B, lanes 1 and 6 are a mix of λ*Hind*III and *E. coli* rRNA markers. Lane 2 is *Pvu*I digested pNDH1. Lane 3 is SP6 RNA polymerase RNA made on *Pvu*I digested pNDH1 with biotin labeled UTP as the only source of UTP in the reaction. Lane 4 is the same as lane 3 except that only unlabeled UTP was used. Lane 5 is SP6 made RNA made with only unlabeled UTP on pNDH1 linearized with *Hind*III. The upper most band of the λ marker is an aggregate that forms when the marker is stored for long periods without a small amount of ethidium bromide in the buffer.

45
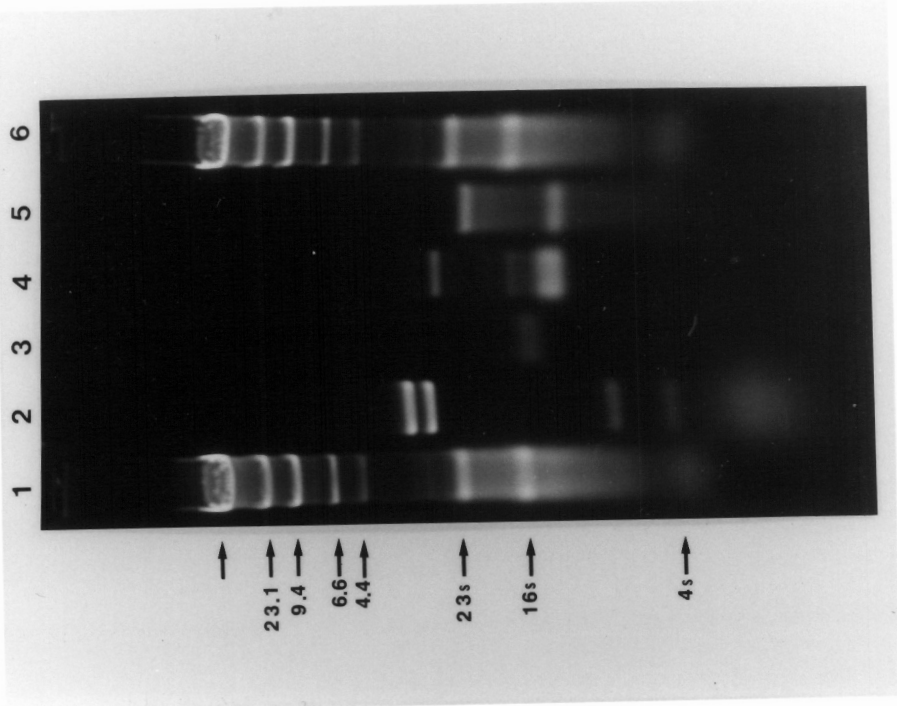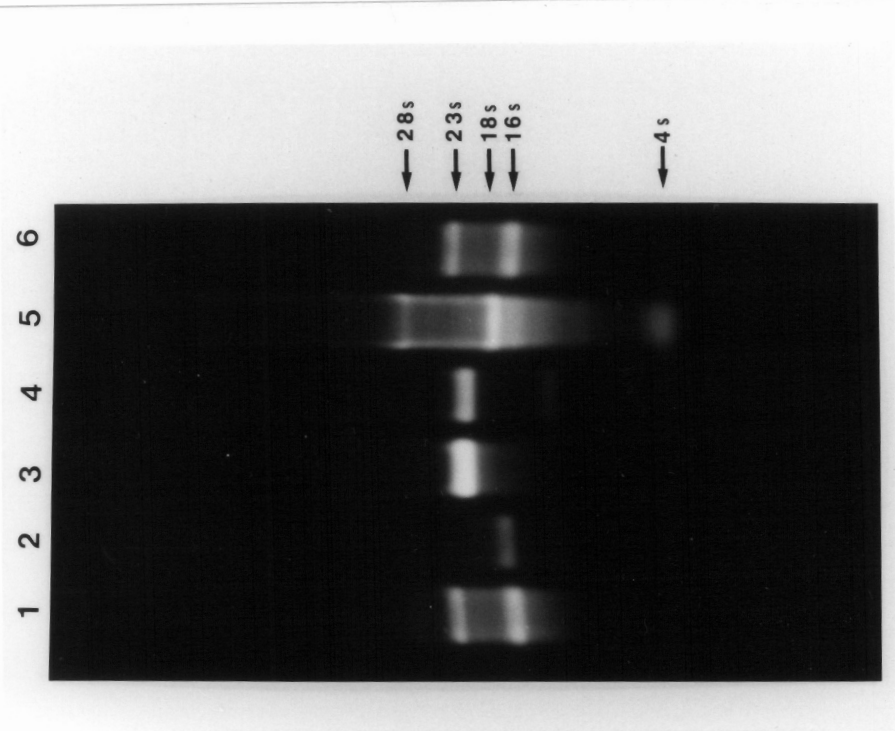
Figure 5. The complete sequence of the 3.5 kb *Hin*dIII fragment.

The coding regions (in italicized capitals) are as follows: 1 to 620 represents most of *nifU*, 889 to 1,247 is *glbN*, 1,594 to 2,487 is *nifH*, and 2,647 to 3,525 is most of *nifD*. The potential stem loop structures are indicated by the double underlined regions. The direct repeat (AATTACG) is indicated by the underlined region. Stop codons are in bold lower case letters.
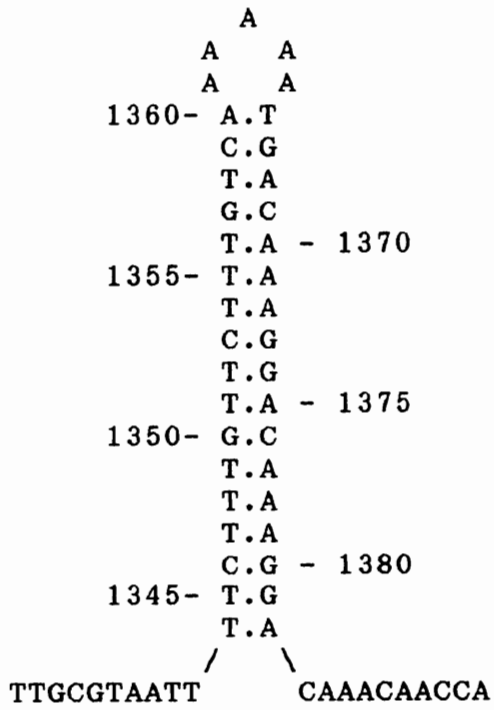
```
                10          20          30          40          50
                 |           |           |           |           |
    1 AAGCTTTGAA AGTGTCCAAC AAAGACATTG CAGATTACCT GGGTGGTTTG
   51 CCAGAAGCAA AGATGCATTG CTCTGTCATG GGACAAGAAG CCCTAGAAGC
  101 CGCTATCTAC AATTATCGTG GCATACCTCT AGCTACCCAC GATGATGATG
  151 ATGAAGGTGC ATTAGTTTGC AGTTGCTTTG GCATCAGCGA ATCTAAGATT
  201 CGTCGCGTGA TTCTAGAAAA TCATCTCACG GATGCCGAGC AAGTAACAAA
  251 TTATGTAAAA GCTGGTGGCG GATGCGGTTC CTGTCTAGCG AACATTGATG
  301 ATATTATTAG ATCAGTTCAA CAGGAATACG CCTCACCTGC TCTCAACAAT
  351 TACGGCGTAA AAGTTGCCAC AGAAATTGTT ACATCTAAGG AACGAGCGCT
  401 AACAAACGTC CAAAAGATTG CTCTAATTCA AAAGGTTCTT GATGAAGAAG
  451 TCAGACCCGT ACTGATTGCA GACGGGGGAG ATGTAGAACT CTATGATGTA
  501 GAAGGTGATC GCTTAAAGTT GTACTGCAAG GTGCTTGTGG TTCCTGTTCT
  551 AGTAGTACAG CAACTCTCAA AATTGCGATC GAAGCCAGAT TACAAGATCG
  601 TGTCAGCAAG AGCCTTGtag tcgaagcagt ttaggaattg ctgtaggact
  651 tatcgtactc tacaaatacg ccctaatatg cattcagcca ggaatagaaa
  701 gtaggatttg aaccgaatca ttgtctccac tcaatgcttc ataccaattc
  751 tttcttAATT ACGAATTACG AATTACGAAT TACGaattgg tatcagtcct
  801 accacAATTA CGgctaacag catataggcc aagagattag cgcctacttg
  851 caaatttcat cattcaactc gcttcaaagg agacaagtAT GAGCACATTG
  901 TACGACAACA TTGGTGGACA ACCGGCTATT GAACAAGTAG TAGATGAACT
  951 CCACAAACGC ATTGCTACAG ACAGCCTCCT CGCTCCCGTT TTTGCTGGTA
 1001 CAGACATGGT GAAGCAACGC AATCATCTAG TTGCTTTCTT AGCTCAAATA
 1051 TTTGAAGGGC CAAAGCAATA CGGCGGTCGT CCAATGGACA AAACCCACGC
 1101 AGGATTGAAC CTACAGCAAC CACACTTTGA TGCGATCGCC AAGCATCTCG
 1151 GCGAACGAAT GGCTGTGCGT GGAGTGTCAG CAGAAAACAC CAAAGCTGCA
 1201 CTAGATCGCG TCACAAATAT GAAGGGCGCT ATTTTGAACA AGtaatagga
 1251 cttatgcatt gattgttgac aaataacaac atcaagtcct atcgaatctt
 1301 gcaagctgaa acacattatg acgtatttgt aaaTTGCGTA ATTTTCTTTG
 1351 TTCTTTGTCA AAAAATGACA AAGGACAAAG GACAAACAAC CAaaagaaaa
 1401 aaatttttaa cgaggaagat cgatgtactt actgtcaata acttgggctg
 1451 tggagcgatc gcccccagca ggcatttatg ccaacgctac caacgatgct
 1501 ctacacgcgc tcactactgt cgctcaactg gcgtgaacgc aactgacaaa
 1551 cgcaaagacc caccaaccaa ccaattgcag ggaaacacga acaATGACAG
 1601 AAGAAACAT TAGACAGATA GCTTTCTACG GTAAAGGCGG TATCGGTAAA
 1651 TCTACCACCT CTCAAAATAC CTTGGCAGCT ATGGCAGAAA TGGGTCAACG
 1701 CATCCTCATC GTCGGTTGCG ACCCTAAAGC TGACTCCACC CGTTTGATGC
 1751 TACACAGCAA AGCTCAAACA ACCGTTCTTC ACCTCGCCGC AGAACGTGGT
 1801 GCAGTAGAAG ATATCGAAAT CGAAGAAGTA ATGTTAACCG GTTTCCGTAA
 1851 CGTTCGTTGC GTAGAATCTG GTGGGCCAGA ACCCGGTGTA GGTTGCGCCG
 1901 GTCGTGGTAT TATCACCGCC ATCAACTTCT TAGAAGAAAA TGGTGCTTAC
 1951 CAAGACCTAG ACTTCGTATC TTACGACGTA TTAGGTGACG TTGTGTGTGG
 2001 TGGTTTCGCA ATGCCTATCC GCGAAGGTAA GGCACAAGAA ATCTACATCG
 2051 TTACATCAGG TGAAATGATG GCGATGTATG CTGCTAACAA CATCGCTCGT
 2101 GGTGTTCTTA AGTATGCTCA CACTGGTGGC GTGCGTTTGG GTGGTTTGAT
 2151 TTGTAACAGC CGTAACACAG ACCGAGAAAT CGAATTAATC GAAACCTTGG
 2201 CAAAACGGTT GAACACCCAA ATGATTCACT ACGTACCCCG TGACAACATT
 2251 GTTCAACACG CAGAATTGCG CCGGATGACT GTTAACGAGT ACGCACCAGA
```

**Fig. 5 continued**

```
2301 AAGCAACCAA GCTAACGAAT ATCGGATTTT GGCTCAAAAA ATTATCGACA
2351 ACAAAAATCT GGCTATTCCT ACACCCATCG AAATGGAAGA ACTAGAAGAA
2401 TTGTTGATTG AATTCGGTAT CCTCGAAAGC GACGAAAATA CTGCCATGCT
2451 GGTTGGTAAG ACTGCTACTG AAGCTCCCGT AGTCtaagtc gccgctaaga
2501 aataacgctt tagcacaagg aaaaaggaag agttccatat cttacctttt
2551 cctctctctc ttgcatctcc cccCTAACCT TAGAGGGGAC TCCTAGTCCC
2601 CTATGCCCCG ATccttacga gtcacagagg ctaagtATGA CACCTCCAGA
2651 AAATCAAAAC ATCATCGAAG AAAGAAAAGA ACTAATTAAA GAAGTTCTCA
2701 GTGCTTACCC AGAAAAAGCT GCTAAAAAGC GCGAAAAGCA CTTAAGTGTA
2751 TACGAAGAAG GTAAGTCCGA TTGCGGCGTT AAGTCTAACA TCAAATCCCT
2801 ACCTGGGGTA ATGACCGCTC GTGGTTGTGC TTACGCCGGT TCTAAAGGTG
2851 TGGTTTGGGG CCCTATTAAG GACATGATCC ACATCAGCCA CGGGCCTGTA
2901 GGTTGCGGTT ACTGGTCTTG GTCTGGTCGT CGTAACTACT ACATCGGCAC
2951 CACAGGTATT GATACCTTTG GTACCATGCA CTTCACCTCT GACTTCCAAG
3001 AAAGAGATAT CGTGTTTGGT GGTGACAAGA AACTTGTTAA GCTGATCCAA
3051 GAACTTGATG TACTTTTCCC CCTCAACCGT GGTGTTTCCA TTCAATCTGA
3101 ATGCCCCATC GGTCTAATTG GGGATGACAT CGAAGCAGTA GCTCGGAAGA
3151 CATCGAAAGA AATTGGCAAG CCAGTAGTAC CTGTGCGTTG CGAAGGCTTC
3201 CGGGGTGTTT CCCAATCTTT AGGACACCAC ATCGCTAACG ACATGGTTCG
3251 TGACTGGGTA TTCACCAGAT CCGACCAAGC TAAGAAAGAC GGTACACTCA
3301 AGTTTGAAGG TACTCCTTAT GATGTAGCCA TCATTGGTGA CTACAACATC
3351 GGTGGAGATG CTTGGGCTAG CCGCATCCTG TTAGAAGAAA TCGCGTTGCG
3401 CGTAGTCGCT CAGTGGTCAG GTGATGGCAC CATCAACGAA ATGTTGATGA
3451 CCCCCAATGT GAAGATGAAC CTCATCCACT GTTACCGGTC GATGAACTAC
3501 ATCAGCCGTC ACATGGAAGA AGCTT
```

Figure 6. The structure of the two potential stem loops on the 3.5 kb *Hind*III fragment.

A is the potential stem loop in the AT rich region between *glbN* and *nifH*. This loop has a $\Delta$G of -15.4 kcal/mol. B is the potential stem loop between *nifH* and *nifD* which has a $\Delta$G of -15.8 kcal/mol. The numbers represent the locations of the bases on the sequence in Figure 5. The prediction of loop A was done allowing the formation of GT (or GU) base pairs. Both loops are predicted to form in the mRNA for these genes with the same $\Delta$G values as predicted above from the DNA sequence.

**A**                                        **B**

```
              A
          A       A
          A       A
  1360-  A.T
          C.G
          T.A
          G.C
          T.A - 1370
  1355-  T.A
          T.A
          C.G
          T.G
          T.A - 1375                              C
  1350-  G.C                                   C     T
          T.A                                  T.A - 2595
          T.A                          2590-  C.G
          T.A                                  A.T
          C.G - 1380                           G.C
  1345-  T.G                                   G.C
          T.A                                  G.C - 2600
         /    \                        2585-  G.C
TTGCGTAATT    CAAACAACCA                       A.T
                                              /    \
                                     CTAACCTTAG    ATGCCCCGAT
```
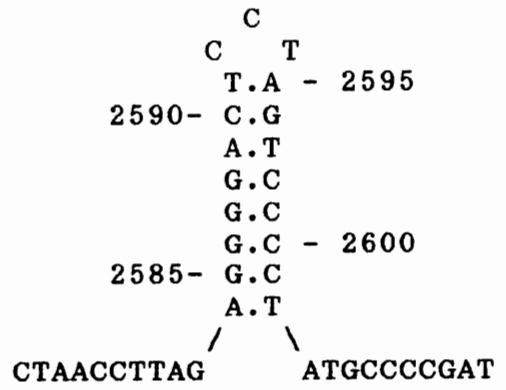
located between genes, as is the direct repeat. The loops
are predicted to form in the DNA as well as the RNA with the
same ΔG values for both nucleic acids.

The predicted amino acid sequences of the *nifU*, *nifH*
and *nifD* genes are very similar to their counterparts in
*Anabaena* sp. PCC 7120 (Figure 7). Alignment of the *N.
commune* UTEX 584 and *Anabaena* proteins show that their
sequence identities are higher towards the N-terminal
portions and become progressively less towards the C-
terminal portions (Figure 7). The predicted amino acid
sequence for the portion of NifU coded for on the 3.5 kb
fragment, has an exact identity of 68% and a conserved
similarity of 9% with NifU from *Anabaena* sp. PCC 7120
(Figure 7A). The predicted amino acid sequence of NifH from
*N. commune* UTEX 584 has an exact identity of 84% and a
conserved similarity of 7% with its counterpart in *Anabaena*
sp. PCC 7120 (Figure 7B). The predicted amino acid sequence
for the portion of *N. commune* UTEX 584 NifD coded for on the
3.5 kb fragment has an exact identity of 82% and a conserved
similarity of 7% with *Anabaena* sp. PCC 7120 NifD (Figure
7C).

Neither the nucleotide sequence nor the predicted amino
acid sequence of *glbN* had any significant homology to any

Figure 7. Comparison of the predicted Nif and globin protein
sequences.

    A, B, and C show the alignments of NifU, NifH, and NifD
respectively, of *N. commune* UTEX 584 (Nc) and *Anabaena* sp.
PCC 7120 (AB). D is the alignment of GlbN (Nc) with the
*Paramecium caudatum* myoglobin (Pc). The bases coding for the
N-terminal portion of NifU and the C-terminal portion of
NifD for *N. commune* UTEX 584 are not on the 3.5 kb fragment.
Identical amino acids are indicated by *, while · represents
similar residues. In figure D, the conserved histidine
residues which may be involved in heme binding, are
indicated by †. Similar residues were determined by the
PCGENE alignment program using the program's criteria that
the following were similar: A,S,T; D,E; N,Q; R,K; I,L,M,V;
and F,Y,W.

**A**

```
AB   MWDYTDKVLELFYDPKNQGVIEENGEPGVKVATGEVGSIACGDALRLHIK  -50

Nc                                    ALKVSNKDIAD  -11
                                     **********
AB   VEVESDKIVDSRFQTFGCTSAIASSSALTEMIKGLTLDEALKVSNKDIAD -100

Nc   YLGGLPEAKMHCSVMGQEALEAAIYNYRGIPLATHDDDDEGALVCSCFGI  -61
     ***********************************·**·********·***·
AB   YLGGLPEAKMHCSVMGQEALEAAIYNYRGIPLAAHDEDDEGALVCTCFGV -150

Nc   SESKIRRVILENHLTDAEQVTNYVKAGGGCGSCLANIDDIIRSVQQEYAS -111
     ** *·**···** ***********·*********** *****·  *   *·
AB   SENKVRRIVIENDLTDAEQVTNYIKAGGGCGSCLAKIDDIIKDVKENKAA -200

Nc   PALNNYGVKVATEIVTSK-ERALTNVQKIALIQKVLDEEVRPVLIADGGD -160
      *** *      * *   *   * ***************************
AB   TNLNNKGGSKPTNIPNSGQKRPLTNVQKIALIQKVLDEEVRPVLIADGGD -250

Nc   VELYDVEGDRLKLYCKVLVVPVLVVQQLSKLRS--KPDYKIV---SARAL -205
     ******·** ·*·          *·    ·  ·*       *·
AB   VELYDVDGDIVKVVLQGACGSCSSSTATLKIAIESRLRDRINPSLVVEAV -300
```

**B**

```
Nc   MTEENIRQIAFYGKGGIGKSTTSQNTLAAMAEMGQRILIVGCDPKADSTR  -50
     **·*******************************·*************
AB   MTDENIRQIAFYGKGGIGKSTTSQNTLAAMAEMGQRIMIVGCDPKADSTR  -50

Nc   LMLHSKAQTTVLHLAAERGAVEDIEIEEVMLTGFRNVRCVESGGPEPGVG -100
     **************************·*·  ******* *·***********
AB   LMLHSKAQTTVLHLAAERGAVEDLELHEVMLTGFRGVKCVESGGPEPGVG -100

Nc   CAGRGIITAINFLEENGAYQDLDFVSYDVLGDVVCGGFAMPIREGKAQEI -150
     **************************************************
AB   CAGRGIITAINFLEENGAYQDLDFVSYDVLGDVVCGGFAMPIREGKAQEI -150

Nc   YIVTSGEMMAMYAANNIARGVLKYAHTGGVRLGGLICNSRNTDREIELIE -200
     ************************·*****·************** *** ***
AB   YIVTSGEMMAMYAANNIARGILKYAHSGGVRLGGLICNSRKVDREDELIM -200

Nc   TLAKRLNTQMIHYVPRDNIVQHAELRRMTVNEYAPESNQANEYRILAQKI -250
     ** ********·**************** ·*** ·*** ** **
AB   NLAERLNTQMIHFVPRDNIVQHAELRRMTVNEYAPDSNQGQEYRALAKKI -250

Nc   IDNKNLAIPTPIEMEELEELLIEFGILESDENTAMLVGKTATEAPVV    -297
     * *·****·**·*** * **·*·*· *    ·  ··** *   ·
AB   NNDK-LTIPTPMEMDELEALKIEYGLLDDDTKHSEIIGKPAEATNRSCRN -299
```

Fig. 7 continued

## C

```
Nc  MTPPENQNIIEERKELIKEVLSAYPEKAAKKREKHLSVYEEGKSDCGVKS  -50
    ****** *··· * **** *** ***** · ******* * ** ********
AB  MTPPENKNLVDENKELIQEVLKAYPEKSRKKREKHLNVHEENKSDCGVKS  -50

Nc  NIKSLPGVMTARGCAYAGSKGVVWGPIKDMIHISHGPVGCGYWSWSGRRN  -100
    ****·*********************************************
AB  NIKSVPGVMTARGCAYAGSKGVVWGPIKDMIHISHGPVGCGYWSWSGRRN  -100

Nc  YYIGTTGIDTFGTMHFTSDFQERDIVFGGDKKLVKLIQELDVLFPLNRGV  -150
    **·* *** ·************************* *** ************
AB  YYVGVTGINSFGTMHFTSDFQERDIVFGGDKKLTKLIEELDVLFPLNRGV  -150

Nc  SIQSECPIGLIGDDIEAVARKTSKEIGKPVVPVRCEGFRGVSQSLGHHIA  -200
    ********* ********** ·**** ******* ·***************
AB  SIQSECPIGSIGDDIEAVAKKTSKQIGKPVVPLRCEGFRGVSQSLGHHIA  -200

Nc  NDMVRDWVFTRSDQAKKDGTLKFEGTPYDVAIIGDYNIGGDAWASRILLE  -250
    ** ·***·* * **· * ** ·*****·************** ·***
AB  NDAIRDWIFPEYDKLKKETRLDFEPSPYDVALIGDYNIGGDAWASRMLLE  -250

Nc  EIALRVVAQWSGDGTINEMLMTPNVKMNLIHCYRSMNYISRHMEEA      -295
    *· ************··**·· *  *· ********** * ·**
AB  EMGLRVVAQWSGDGTLNELIQGPAAKLVLIHCYRSMNYICRSLEEQYGMP  -300

AB  WMEFNFFGPTKIAASLREIAAKFDSKIQENAEKVIAKYTPVMNAVLDKYR  -350

AB  PRLEGNTVMLYVGGLRPRHVVPAFEDLGIKVVGTGYEFAHNDDYKRTTHY  -400

AB  IDNATIIYDDVTAYEFEEFVKAKKPDLIASGIKEKYVFQKMGLPFRQMHS  -450

AB  WDYSELGDGVQMSDEVRFFCEGRKKSLFLA  -480
```

## D

```
Nc  MSTLYDNIGGQPAIEQVVDELHKRIATDSLLAPVFAGTDMVKQRNHLVAF  -50
     ·*····*** *·  *        * ·*· ·* * * ** * *    **
Pc   SLFEQLGGQAAVQAVTAQFYANIQADATVATFFNGIDMPNQTNKTAAF   -48

Nc  LAQIFEGPKQYGGRPMDKTHAGLNLQQPHFDAIAKHLGERMAVRGVSAEN  -100
    *   ** · ** ·   †*· · · *·· †*   ·· **·*
Pc  LCAALGGPNAWTGRNLKEVHANMGVSNAQFTTVIGHLRSALTGAGVAAAL  -98

Nc  TKAALDRVTNMKGAILNK  -118
     ··     ··* ··
Pc  VEQTVAVAETVRGDVVTV  -116
```

cyanobacterial or nitrogen fixation related sequences (such as *nif, vnf, anf, fix*, etc.) in the GenBank or SWISS-PROT databases. However, the predicted amino acid sequence of GlbN has an exact identity of 27% and a conserved similarity of 22% with the myoglobin from *Paramecium caudatum* (Iwaasa et al. 1989). The GlbN protein (118 amino acids) is two amino acids longer than the *P. caudatum* myoglobin (116 amino acids) and the histidine residues presumed to be the sites of heme binding in the *P. caudatum* myoglobin, appear to be conserved (Figure 7D). These are histidines 69 and 86 in the GlbN amino acid sequence and residues 68 and 84 in the *P. caudatum* myoglobin sequence.

**Characterization of GlbN**

By expressing the PCR copy of *glbN* in *E. coli* BL21 DE3, large amounts of the protein were purified. After 2 hours of IPTG induced expression, the cells turned red as shown in Figure 8 and the 30,000 x *g* supernatant of the lysed cells was dark red. Based on the predicted amino acid sequence, the protein has a molecular mass of 12,906 Da excluding the iron porphyrin. On SDS polyacrylamide gels the protein migrated with a molecular mass between 11,900 Da and 12,900 Da (about 12,400 Da, Figures 9 and 10). The sequence of the PCR copy of the gene was identical to the original sequence,

Figure 8. The appearance of IPTG-induced *E. coli* BL21 DE3 containing pT7-7 or pT7ORF18.

A, IPTG-induced *E. coli* BL21 DE3 containing the expression plasmid pT7-7 with no inserts. B, *E. coli* BL21 DE3 containing pT7ORF18 after induction with IPTG. Both were induced for 2 hours.
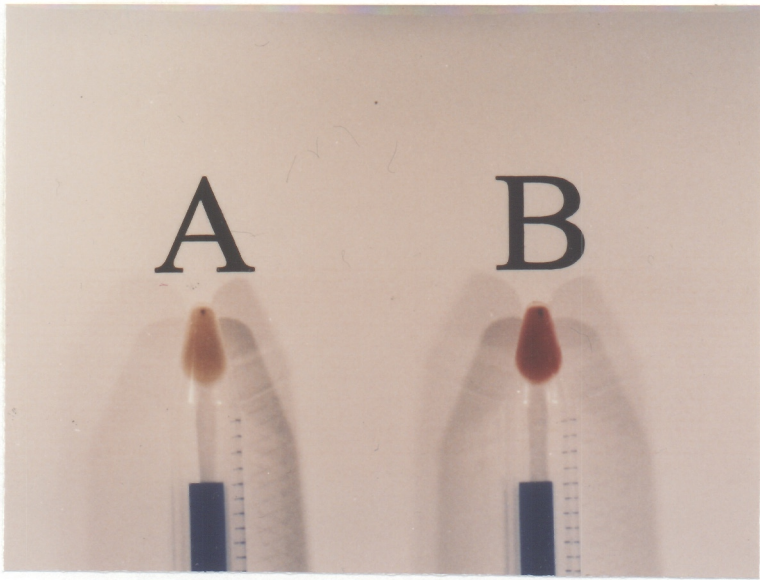
Figure 9. Fractions of GlbN from the Mono Q column.

　　　This is a 15% w/v polyacrylamide Coomassie stained gel
showing various fractions collected from the Mono Q column
(conditions described in Materials and Methods). Lanes 1 and
10 are the molecular mass standards. Lane 2 is the flow
through that contains GlbN. Lanes 3 to 9 are major peaks
that eluted from the column after the salt gradient (from 0
to 1 M NaCl in 50 mM HEPES pH 7.5) was started. Lanes 2
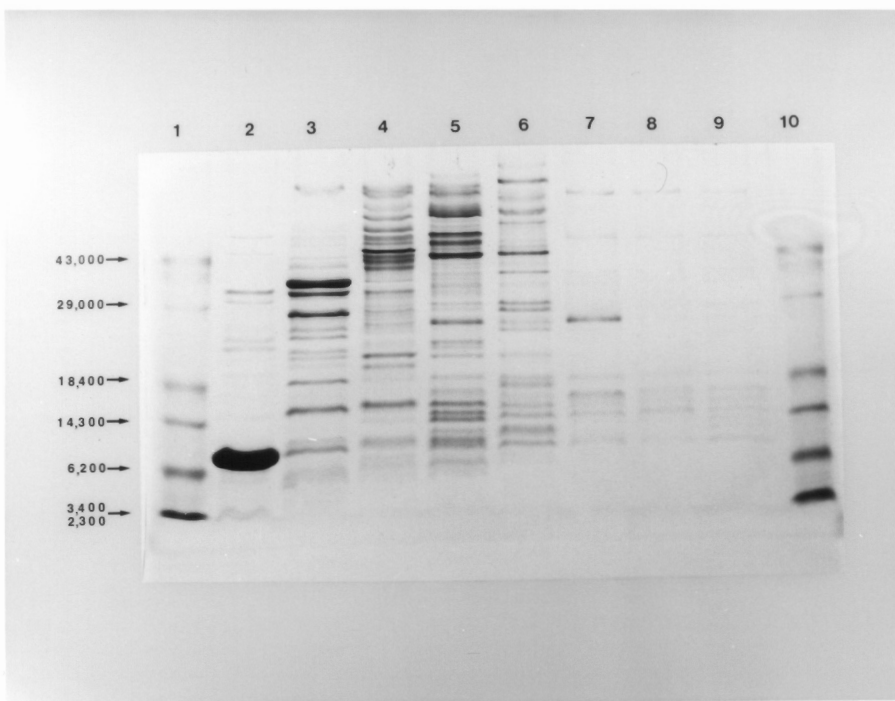through 9 represent 10 µl from 500 µl fractions.

Figure 10. Fractions of purified GlbN from the Superose 12 column.

This is a silver stained 15% w/v polyacrylamide gel showing purified and partly purified fractions of GlbN from one Superose 12 fractionation (conditions described in Materials and Methods). Lane 1 is the molecular mass standards (1 µg). Lane 2 is a peak that elutes from the column a few fractions before the GlbN peak. Lanes 3, 4, and 5 are fractions from the beginning, middle and end respectively, of the GlbN peak monitored at 418 nm. Lanes 2, 3, 4, and 5 represent 10 µl from 500 µl fractions.
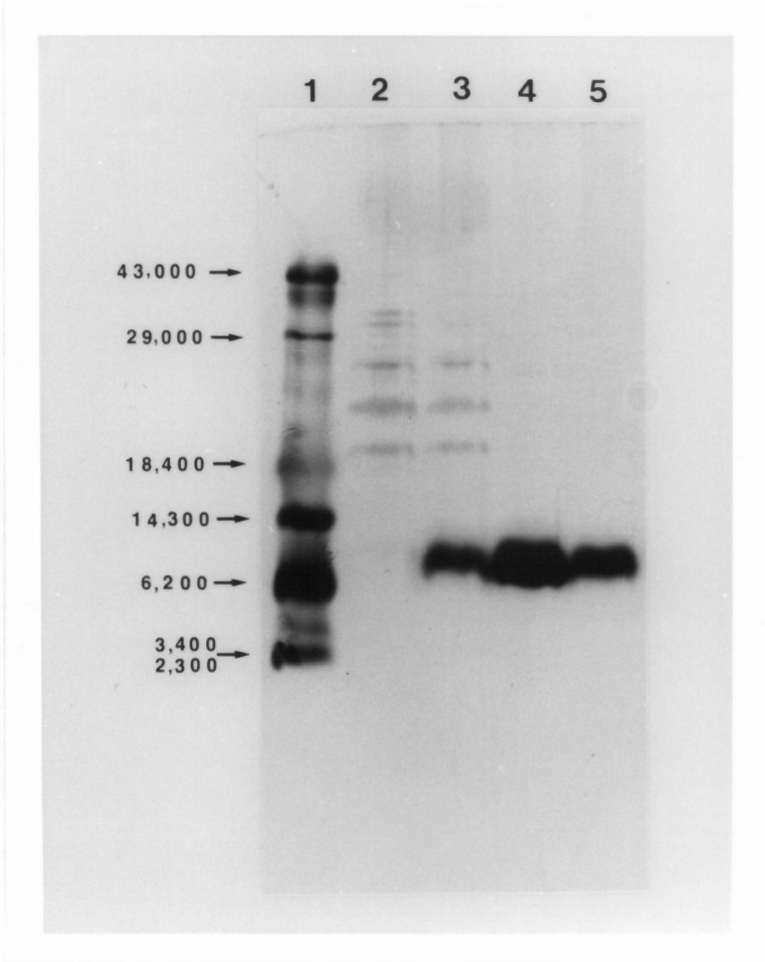
suggesting that any differences in the molecular mass from
that predicted from the DNA sequence, are not due to the
introduction of stop codons or other changes in the sequence
during PCR.

Based on the predicted amino acid sequence, GlbN has a
theoretical pI of 9.73 without the iron porphyrin. However,
in 50 mM HEPES at pH 7.5 the native protein did not bind to
Mono S (a cation exchange resin) or Mono Q (an anion
exchange resin), while most of the other *E. coli* proteins
did. This characteristic was the basis for using the Mono Q
column as the first step in purification of the protein
(Figure 9). The next purification step was by gel filtration
on a Superose 12 column in 50 mM HEPES pH 7.5, where the
native protein migrated with a molecular mass of
approximately 20,000 Da and fractions containing relatively
pure GlbN, as determined on silver-stained gels, could be
obtained (Figure 10). In all of the column fractionation
steps, the fraction containing GlbN was characterized by a
reddish color. These fractions also had a maximum absorbance
peak at about 416 nm.

An absorption spectrum of purified GlbN from 700 nm to
250 nm showed a major peak at about 413 nm and additional
peaks near 541 nm and 567 nm (Figure 11). Treatment with
sodium dithionite resulted in a decrease in absorption of

Figure 11. Spectra of GlbN treated with dithionite and CO.

Absorption peaks for untreated GlbN (-------) were determined to be at about 413 nm, 541 nm, and 567 nm. For dithionite treated GlbN (············) the peaks were at 420 nm, 535 nm, 559 nm, and 626 nm. For CO treated GlbN (———————) the peaks are 420 nm, 535 nm, 559 nm, and 626 nm. The absorbance scale for 250 nm to 500 nm is 0.0 to 0.6 while the scale for 500 nm to 700 nm is 0.00 to 0.06.
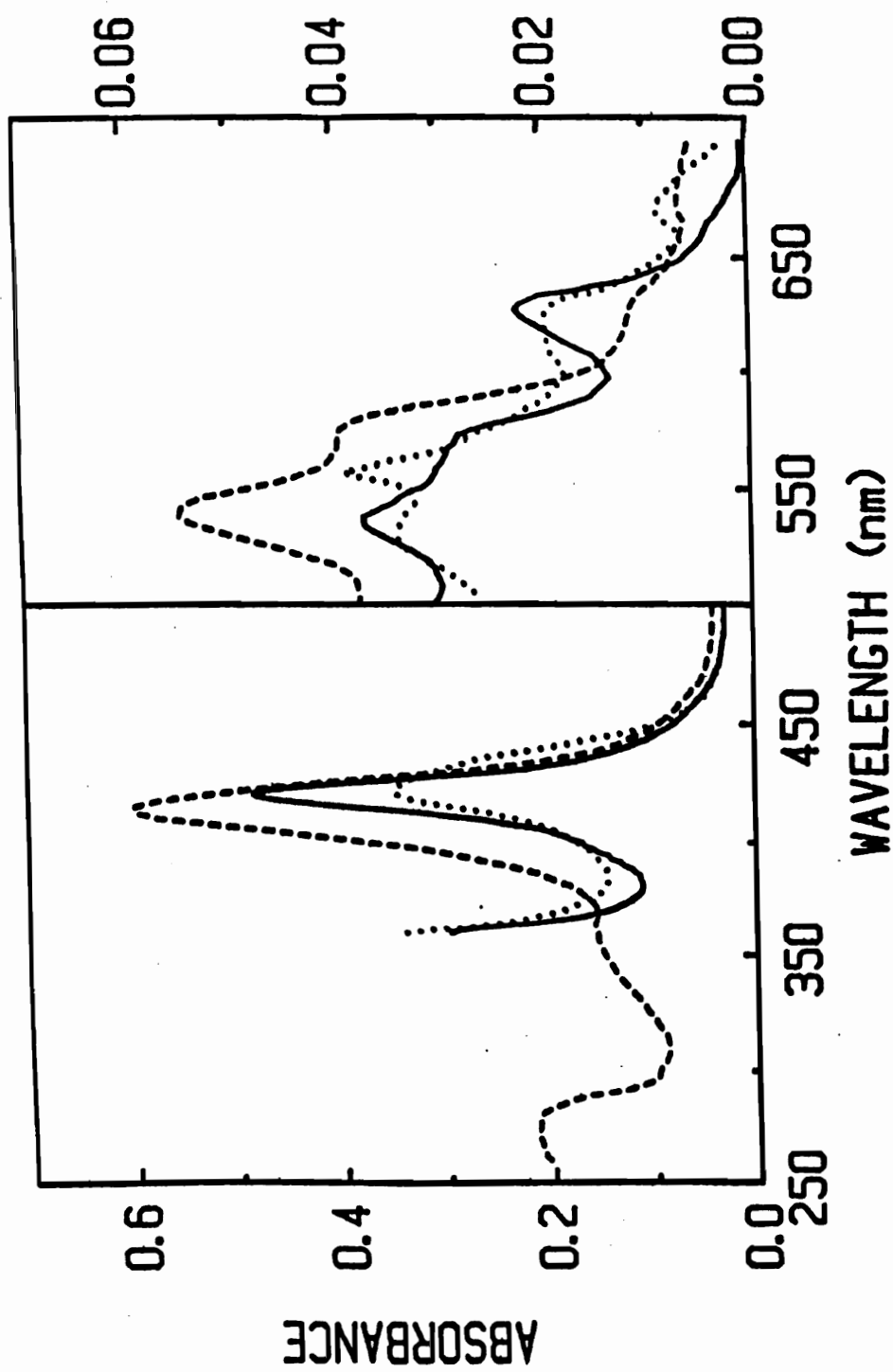
the major peak and a shift to about 420 nm (Figure 11). A drop in the absorbance of the peak at 541 nm also occurred along with a shift to 529 nm. The peak at 567 nm decreased in absorbance, but not as much as the 529 nm peak, and there was a shift from 567 nm to 559 nm. An additional peak at 626 nm was observed after dithionite treatment but was not present in the untreated protein spectra.

Treatment of dithionite-reduced GlbN with saturating amounts of carbon monoxide (CO), resulted in changes in the absorbance of most peaks and shifts for some (Figure 11). The major peak at 420 nm remained at this wavelength but the amount of absorption increased. The peak at 529 nm shifted to 535 nm and increased in absorbance. The absorbance of the peak at 559 nm decreased even more with CO treatment but its absorption wavelength did not change. The peak at 626 nm remained at this wavelength but its absorption increased slightly with CO treatment. In general, these shifts in the spectral characteristics with dithionite and CO treatment confirm that GlbN is a heme protein (Dickerson and Geis 1983, Paulon 1988, Lemberg and Barrett 1973).

Western blots using antisera made against recombinant GlbN, showed two major bands that were reactive with the antisera in *N. commune* UTEX 584, a large protein with a molecular mass of approximately 15,300 Da and a smaller

protein of about 13,000 Da. In *N. commune* UTEX 584 the amount of the larger protein being detected did not change significantly under the different growth conditions. However, there were significant changes in the amount of the smaller protein being detected (Figure 12). The smaller protein was never observed in *Anabaena* sp. PCC 7120 under the growth conditions used in this study (Figures 12 and 13). When *N. commune* UTEX 584 was shifted from BG11 to BG11$_o$ with aeration, the large protein was present with only trace amounts of the smaller protein (Figure 12 lane 1). When *N. commune* UTEX 584 was grown on BG11$_o$ in a shaker flask (less dissolved oxygen than with aeration), as more nitrogenase reductase was synthesized, more of the smaller protein was produced (Figures 13 and 14 lanes 2 and 3). When *N. commune* UTEX 584 was grown in Mo-free media supplemented with 100 nM VSO$_4$, there was also an increase in the amount of the smaller protein but not as much as seen in BG11$_o$ without aeration (Figure 12). Under all of these growth conditions the amount of the larger protein did not appear to change significantly. In *Anabaena* sp. PCC 7120, under the same growth conditions as described above for *N. commune* UTEX 584, only the larger protein was detected (Figures 12 and 13). In *A. variabilis* a protein of approximately 17,000 Da

Figure 12. Western blot of *N. commune* UTEX 584 and *Anabaena* sp. PCC 7120 protein extracts probed with antisera to GlbN made in BL21 DE3.

This blot was from a 12% w/v polyacrylamide gel. Lanes 1 to 4 are *N. commune* UTEX 584 grown on $BG11_o$ with aeration (lane 1), $BG11_o$ on a shaker without aeration (lane 2), $BG11_o$ minus Mo plus 100 nM $VSO_4$ on a shaker (lane 3), and $BG11_o$ on a shaker (lane 4). Lane 5 contains purified GlbN from BL21 DE3. Lanes 6 to 8 are *Anabaena* sp. PCC 7120 grown on $BG11_o$ with aeration (lane 6), $BG11_o$ on a shaker without aeration (lane 7), and $BG11_o$ minus Mo plus 100 nM $VSO_4$ on a shaker (lane 8). The amount of protein in each lane is about 25 µg for the *N. commune* UTEX 584 samples, 50 µg for the *Anabaena* sp. PCC 7120 samples, 3 µg for *A. variabilis*, 0.7 µg of recombinent GlbN, and about 1 µg for the molecular mass standards.
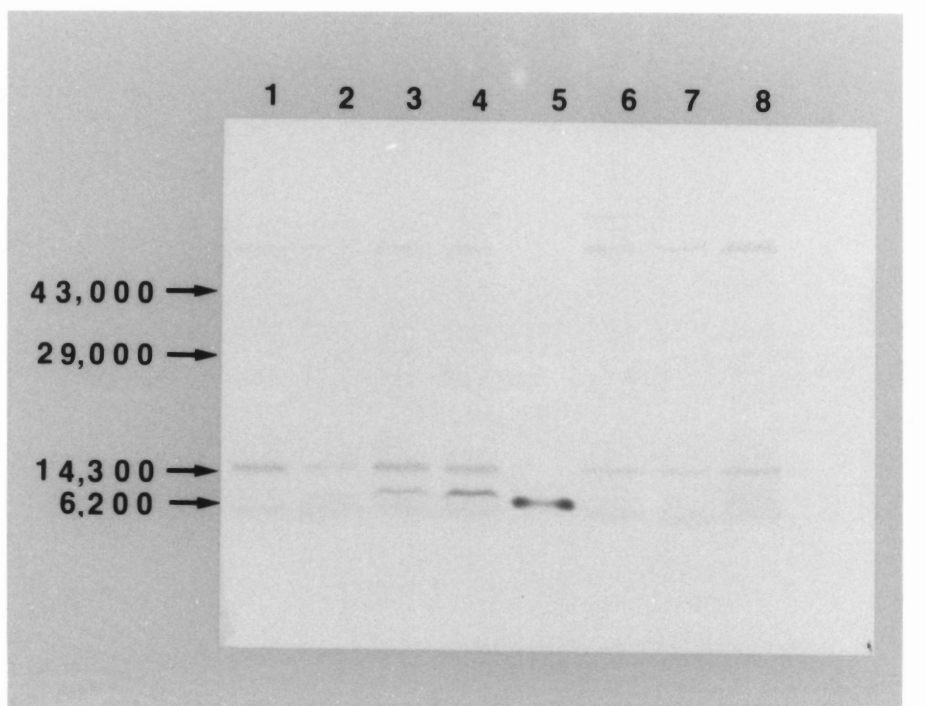
Figure 13. Western blot of *N. commune* UTEX 584 and *Anabaena* protein extracts from a 15% w/v polyacrylamide gel probed with antisera to GblN made in BL21 DE3.

Lanes 1 to 3 are *N. commune* UTEX 584 grown in $BG11_o$ with aeration (lane 1), $BG11_o$ without aeration on a shaker for 7 days (lane 2), and in $BG11_o$ on a shaker for 2 days (lane 3). Lane 4 is purified GlbN from BL21 DE3. Lanes 5 to 7 are *Anabaena* sp. PCC 7120 grown on $BG11_o$ with aeration (lane 5), $BG11_o$ without aeration on a shaker (lane 6), and on BG11 without aeration on a shaker (lane 7). Lane 8 is an extract of *A. variabilis*. The amount of protein in each lane is about 15 µg for the *N. commune* UTEX 584 samples, 30 µg for the *Anabaena* sp. PCC 7120 samples, 3 µg for *A. variabilis*, 0.7 µg of recombinant GlbN, and 1 µg for the molecular mass standards.
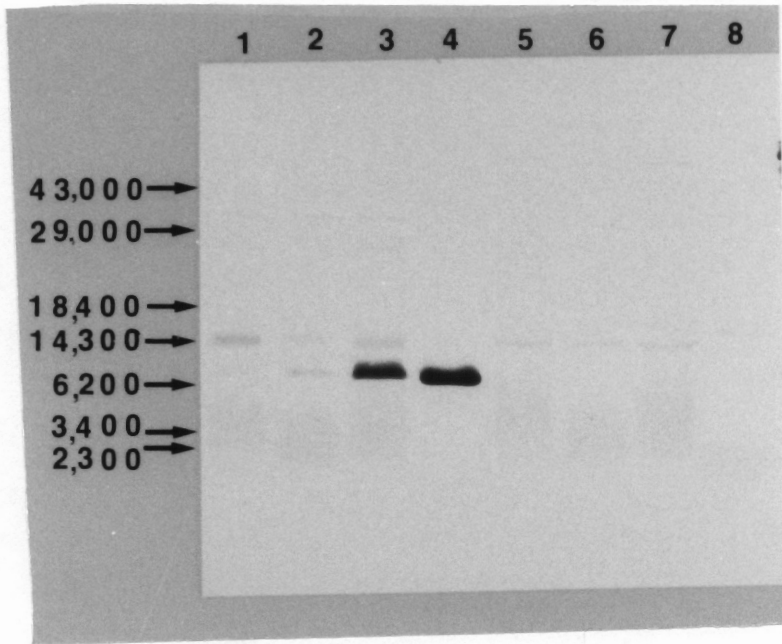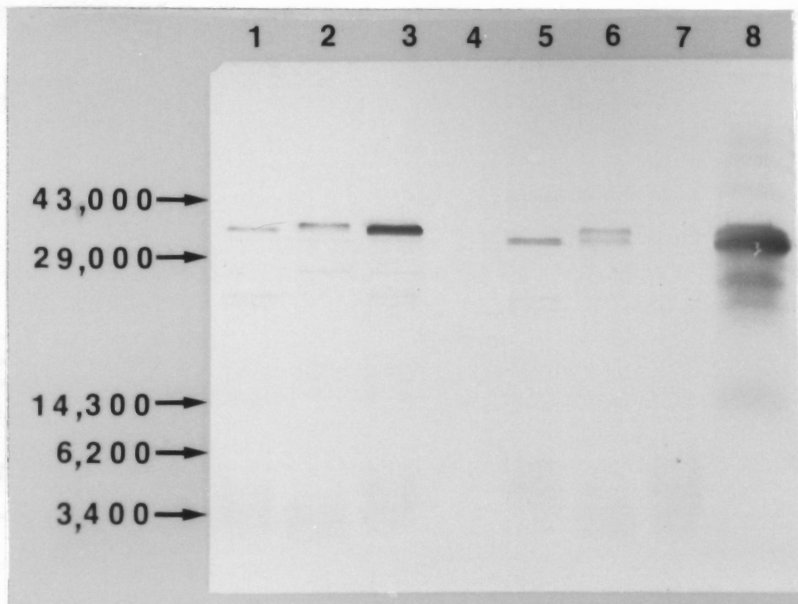
Figure 14. Western blot of *N. commune* UTEX 584 and *Anabaena* protein extracts from a 15% w/v polyacrylamide gel probed with antisera to the nitrogenase reductase of *Rhodospirillum rubrum*.

The lanes are the same as in figure 13.

was detected but no protein of about 13,000 Da was seen

(Figure 13 and Dr. Malcolm Potts personal communication).

## Discussion

By obtaining the entire sequence of the 3.5 kb *Hind*III
fragment, it was shown that the predicted amino acid
sequences of the *N. commune* UTEX 584 *nifU, nifH* and *nifD*
genes are homologous to their counterparts in *Anabaena* sp.
PCC 7120 (Figures 7 A, B, and C). The portion of NifU coded
for on the 3.5 kb fragment, has an exact identity of 68% and
a conserved similarity of 9% with NifU of *Anabaena* sp. PCC
7120. The nitrogenase reductases (NifH) of *N. commune* UTEX
584 and *Anabaena* sp. PCC 7120 have an exact identity of 84%
and a conserved similarity of 7%. Both NifH proteins have an
arginine at position 104 which may function as a site for
reversible ADP-ribosylation, as occurs at arginine 101 in
*Rhodospirillum rubrum* (Kanemoto and Ludden 1984, Ludden et
al. 1989, 1988). Western blots in this study showed two
forms of nitrogenase reductase in *N. commune* UTEX 584 and
*Anabaena* sp. PCC 7120, suggesting that they may be modified
(Figure 14). Covalent modification of nitrogenase reductase
from *A. variabilis* has been described but it is not known if
the modification is by ADP-ribosylation (Reich and Böger
1989). The NifH proteins of *N. commune* UTEX 584 and *Anabaena*
sp. PCC 7120 have a cysteine at position 101, which in the
*Azotobacter vinelandii* NifH protein is believed to be

involved in FeS ligation of the two subunits to the shared FeS cluster (Hausinger and Howard 1983).

The predicted amino acid sequence of *N. commune* UTEX 584 *nifD* has an exact identity of 82% and a conserved similarity of 7% with its counterpart in *Anabaena* sp. PCC 7120. Both proteins also have features like those of their counterparts in other nitrogen fixing bacteria. For example, they have the five conserved domains around their cysteine residues that occur in the α subunits of other species (Figure 15). The *N. commune* UTEX 584 protein also has the conserved sequence H-G-X-X-G-C at its second conserved cysteine. Recent work has shown that the cysteine residues analogous to *N. commune* UTEX 584 cysteine-64 and cysteine-156 in other species are essential for nitrogenase function, while the residues analogous to histidine-85 and cysteine-90 are not (Dean et al. 1990). As to whether or not there is a genomic rearrangement at a site in the 3' end of the *N. commune* UTEX 584 *nifD* gene, as occurs in *Anabaena* sp. PCC 7120, is not known since that part of *nifD* is not on the 3.5 kb fragment. All of these sequence data show that the nitrogenase proteins of *N. commune* UTEX 584 and *Anabaena* sp. PCC 7120 are similar to each other and those of other nitrogen fixing bacteria.

Figure 15. Comparison of the five conserved domains of the nitrogenase α-subunits from several nitrogen fixing bacteria.

The different amino acid sequences are for *N. commune* UTEX 584 (Nc), *Anabaena* sp. PCC 7120 (AB), *Azotobacter vinelandii* (Av), *Klebsiella pneumoniae* (Kp), and *Clostridium pasteurianum* (Cp). Residues that are different from the *N. commune* UTEX 584 sequence are indicated with an * above the residue. The numbers at the beginning and within each sequence represent the number of residues from the beginning of the protein and between the domains respectively. The sequences for Av, Kp and Cp are from Dean et al. 1990.

```
              I                        II                    III                      IV                    V
NC  61-RGCAYAGSKGVV-8-HISHGPVGCGYWSW-60-CPIGLIGDDI-19-CEGFRGVSQSLGHHIAN-79-HCYRSMNY

AB  61-RGCAYAGSKGVV-8-HISHGPVGCGYWSW-60-CPIGSIGDDI-19-CEGFRGVSQSLGHHIAN-79-HCYRSMNY
                              **                 *

Av  59-RGCAYAGSKGVV-8-HISHGPVGCGQYSR-60-CPIGLIGDDI-19-CEGFRGVSQSLGHHIAN-74-HCYRSMNY
                             ** *

Kp  59-RGCAYAGSKGVV-8-HISHGPVGCGQYSR-60-CPVGLIGDDI-19-CEGFRGVSQSLGHHIAN-73-HCYRSMNY
            *            * * ****         *                          *          *  *

Cp  59-RGCAYAGCKGVV-8-HITHGPIGCCFYTW-60-CPVGLIGDDI-19-CEGYKGVSQSAGHHIAN-70-QCHRSINY
```

While the amino acid sequences of the *N. commune* UTEX 584 nitrogen fixation proteins are well conserved, neither the nucleotide sequence nor the predicted amino acid sequence of *glbN* had any significant homology or identity to any nitrogen fixation related sequences (such as *nif, vnf, anf, fix,* etc.) in the GenBank or SWISS-PROT databases. The amino acid sequence of the *P. caudatum* myoglobin was the only sequence with a significant exact identity and conserved similarity. The two protein sequences differ by only two amino acids in length and have a similar hydropathic index (Figures 7D and 16).

The recombinant protein made in *E. coli* BL21 DE3 from the PCR copy of the gene has several characteristics that are typical for a myoglobin-like protein. When it is made in the expression host, the cells become red as would be expected for a heme containing protein that can bind oxygen.

Changes in the spectral characteristics of GlbN upon treatment with dithionite and CO are also similar to those that occur in heme proteins such as the globins and cytochromes (Dickerson and Geis 1983, Lemberg and Barrett 1973). The major peak near 413 nm is similar to the Soret peak of other heme proteins. This peak also shifts upon dithionite and CO treatment in a manner similar to that of heme proteins. Changes in the peaks at 451 nm and 567 nm

Figure 16. Comparison of the hydropathic index of GlbN with that of *P. caudatum* myoglobin.

The plots were generated by the PCGENE program SOAP with an amino acid interval of 9. The Y axis is the hydropathic index where a positive value indicates a more hydrophobic nature. The X axis is the amino acid position.

*Nostoc commune*

*Paramecium caudatum*
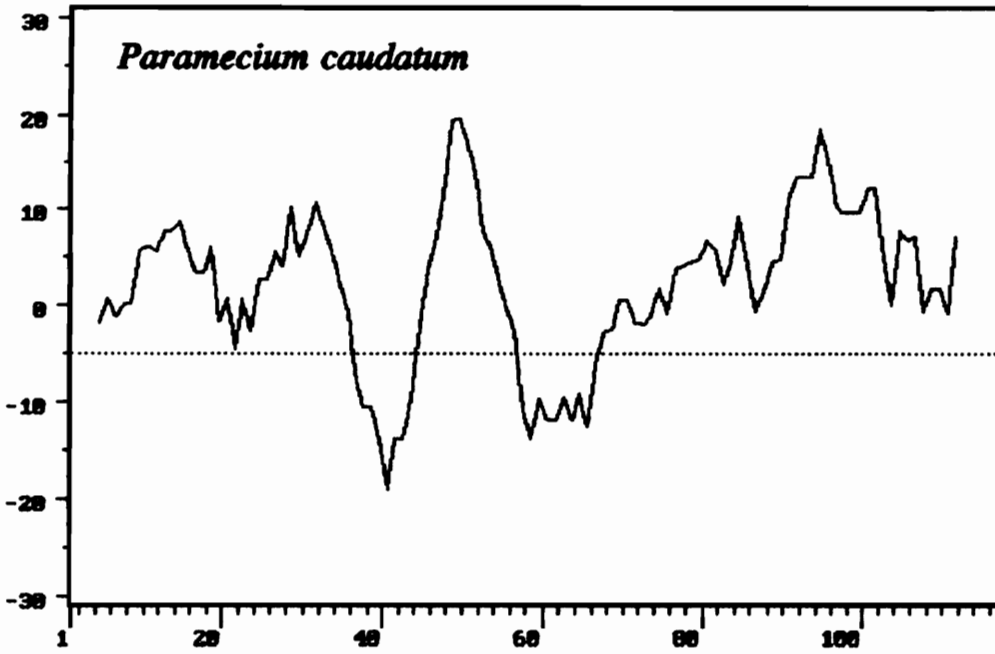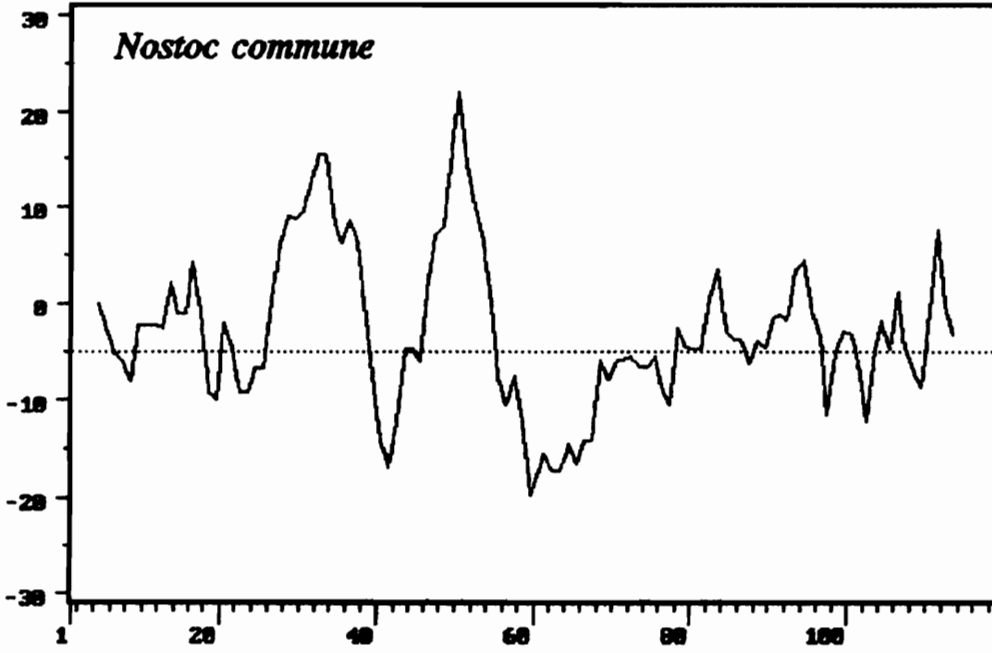
with dithionite and CO treatment have some unique
characteristics but in general are more like the changes
that occur with the globins than the changes seen with
cytochromes. Spectral characteristics for the *P. caudatum*
myoglobin with dithionite and CO treatment have not been
reported, however, the spectral characteristics of the
oxymyoglobin and acidic metmyoglobin forms show similar
Soret peaks around the 420 nm region, but the peaks from 430
nm to 600 nm are not like those seen with GlbN (Tsubamoto
1990).

The spectral characteristics of the GlbN peaks are
similar to those of another bacterial heme protein, the
*Vitreoscilla* sp. strain C1 hemoglobin (*VHb*) (Webster 1988).
The peaks of *VHb* in the oxygenated form are at 414 nm, 543
nm and 576 nm (compare to Figure 11). With CO treatment the
peaks shift to 419 nm, 535 nm and 566 nm which is similar to
the CO treated GlbN spectra. When *VHb* is treated with CN, a
peak at 625 nm is produced (Webster 1988). This peak is
similar to the 626 nm peak that appears in the GlbN spectra
after dithionite and CO treatment. Despite the spectral
similarities, the two proteins are not alike in other
respects. GlbN and *VHb* do not have high amino acid sequence
identities or conserved similarities (11% and 15%
respectively, data not shown). *VHb* is a dimer of identical

subunits of 15,770 Da per subunit (Wakabayashi et al. 1986).
It is not clear if GlbN is a monomer or dimer. The size of
native GlbN made from the PCR copy of the gene determined by
gel filtration is approximately 20,000 Da. On SDS gels, GlbN
appears to be about 12,400 Da. With the iron porphyrin, GlbN
could be expected to have a native molecular mass of about
15,000 Da (if monomeric) or about 30,000 Da (if dimeric),
indicating that further analysis of the protein's mobility
under native and denaturing conditions is needed.

Even still, these data confirm that GlbN is a heme
protein and that the porphyrin bound Fe is accessible to
molecular oxygen ($O_2$) and CO. This accessibility is typical
of myoglobin and hemoglobin due to the positioning of the
distal histidine residue (possibly histidine 86 in GlbN)
such that it does not prevent the porphyrin Fe from reacting
with $O_2$ and CO. In most cytochromes, the distal histidine
blocks entry of $O_2$ and CO into the pocket where the Fe is
located (Dickerson and Geis 1983, Poulos 1988). While GlbN
is a heme protein, it is still not known if it functions as
a globin (ie. binds $O_2$ for storage, transport and/or
protection of the nitrogen fixation system), as a
cytochrome, in which case it is capable of transferring
electrons, or if it has some other function. Since its
spectral characteristics are more like those of the globins,

we proposed the name of cyanoglobin (GlbN) for the protein
with the gene designation of *glbN*.

Western analysis showed that there are slight
differences between the size of GlbN made in BL21 DE3 and
the protein detected in *N. commune* UTEX 584. Under
denaturing conditions, the GlbN made in BL21 DE3 is slightly
smaller than the *N. commune* UTEX 584 protein (Figures 12 and
13). In addition, a larger protein was detected in *N.
commune* UTEX 584 with a molecular mass of about 15,300 Da. A
protein of similar size was detected in *Anabaena* sp. PCC
7120 while the smaller protein was not detected with the
growth conditions used in this study. In *A. variabilis* a
protein of about 17,000 Da was detected and no smaller
protein was seen (Figure 13 lane 8).

Initially it was thought that the large protein in *N.
commune* UTEX 584 was a modified form of GlbN. However, the
large protein was present under nitrogen fixing and non-
nitrogen fixing growth conditions, suggesting that the
larger protein may be another protein which is cross
reactive with the antisera. If it is another protein, since
it is about the same size as *VHb*, and GlbN has an exact
sequence identity of 11% with *VHb*, it is possible that some
antigenic sites on GlbN could be cross reactive with sites
on a *VHb*-like protein or some other protein in *N. commune*

UTEX 584. However, it is also possible that the larger
proteins are simply cross reactive with antibodies in the
antisera to *E. coli* like proteins. Whether or not the larger
protein is related to GlbN could be determined using
antisera preabsorbed with *E. coli* proteins or by analysis of
the amino acid sequences or compositions of the large and
small proteins.

If the larger protein is a modified form of GlbN, there
are several ways by which it could be modified. One way is
by phosphorylation as occurs in non-photosynthetic
diazotrophs by FixL. FixL is an oxygen-sensing heme protein
that has kinase activity, with a subunit molecular mass of
42,803 Da and can undergo autophosphorylation (Gilles-
Gonzalez et al. 1991, De Philip et al. 1990). Whether or not
GlbN has kinase activity or can be phosphorylated has not
been determined.

GlbN could also be modified by reversible ADP-
ribosylation as occurs with the nitrogenase reductase of the
purple-photosynthetic bacterium *R. rubrum* at a specific
arginine residue (Kanemoto and Ludden 1984, Ludden et al.
1988, 1989). GlbN has several arginine residues, any of
which may be a site for ADP-ribosylation.

An additional method of modification may be to have the
larger form of the protein made by two different genes as

occurs with *nifV* in *Clostridium pasteurianum* (Wang et al. 1991). Here the *nifV* protein has its N-terminal portion and C-terminal portion made on different genes.

If a mechanism of modification is used, it probably has a regulatory function. However, Western blots detecting nitrogenase reductase, showed that as the synthesis of nitrogenase reductase increased there was also an increase in the amount of GlbN (Figures 13 and 14 lanes 1 to 3). Under these conditions there was little change in the amount of the larger protein, suggesting that it may not be a modified form of GlbN. Further, in *Anabaena* sp. PCC 7120 under similar growth conditions, no protein near the predicted size of GlbN was detected but the larger protein was always present. In *A. variabilis* the large protein of 17,000 Da was seen under all growth conditions and was present in purified heterocyst fractions but no protein the size of GlbN was seen (Dr. Malcolm Potts personal communication). These data suggest that *glbN* may not be in *Anabaena* sp. PCC 7120 (and possibly not in *A. variabilis*). Further, Southern hybridization of genomic DNA from *Anabaena* sp. PCC 7120 with a *glbN* probe indicates that the gene is not in this species (Dr. Anneliese Ernst personal communication). If *glbN* is not present, this suggests that the larger proteins could be cross reactive proteins which

may not be related to *glbN*. However, further analysis to determine the significance of the larger protein is needed.

*N. commune* UTEX 584 is known to have 3 copies of *nifH* like genes and one copy of *nifD* (Defrancesco and Potts 1988). The extra copies of the *nifH* like genes suggests that this species may have alternative nitrogen fixation systems. While the sequence data showed that the *nifU*, *nifH*, and *nifD* genes are homologous to the *nif* genes of other species, difficulties with detecting GlbN in the early stages of Western analysis, suggested the possibility that these genes may be for an alternative nitrogen fixation system. Since it was not known during the early stages of Western analysis if these difficulties were due to technical problems or if they were of physiological significance, an attempt to determine if *glbN* and the *nif* genes around it encode for an alternative nitrogen fixation system, was undertaken by growing *N. commune* UTEX 584 and *Anabaena* sp. PCC 1720 in Mo-free medium. Once the technical difficulties were overcome, Western blotting of proteins from these cultures showed lower levels of GlbN in the Mo-free cultures than in cultures grown in $BG11_o$ (Figure 12). If GlbN and possibly the *nif* genes around it, were for an alternative nitrogen fixing system, it might be expected that more GlbN would be detected in the Mo-free medium than in $BG11_o$, however, more

GlbN was detected in BG11$_o$. The low levels of GlbN in the Mo-free medium may reflect low levels of expression of the *glbN* gene due to trace amounts of molybdenum contamination in the growth medium (see Materials and Methods), suggesting that GlbN and the *nif* genes around it are not for an alternative nitrogen fixing system (further supporting the sequence data).

Experiments to determine the function and conditions controlling expression of GlbN were performed by comparing the amounts of GlbN in aerated and unaerated cultures of *N. commune* UTEX 584 grown in BG11$_o$. The results showed that more GlbN (and nitrogenase reductase) was detected in unaerated cultures than in aerated cultures (Figures 12 and 13). These results suggest that expression of GlbN and nitrogenase reductase may be related to oxygen concentration. In aerated cultures, the increased oxygen content appeared to suppress the expression of GlbN and to some extent that of nitrogenase reductase (Figures 12, 13 and 14). If the function of GlbN is for oxygen protection of nitrogenase, it might be expected that there would be an increase in the amount of GlbN in aerated cultures *vs.* unaerated cultures. Yet, as noted above, the opposite was true. These results could be interpreted as indicating that the increased oxygen concentration in the aerated medium may

have inhibited expression of GlbN and of the nitrogen
fixation related genes, suggesting that GlbN may not be
involved in protecting nitrogenase from oxygen. This
interesting response suggests that oxygen is capable of
inhibiting expression of GlbN and the *nif* genes of *N.
commune* UTEX 584 as has also been observed in *Anabaena* sp.
PCC 7120 (Haselkorn 1978, Helber et al. 1988a, 1988b).
Similarly, increased oxygen concentrations were shown to
inhibit transcription of the *VHb* gene in *Vitreoscilla* sp.
strain C1 (Dikshit et al. 1989). However, another
explanation is that if GlbN expression is controlled by
internal cellular oxygen concentrations, its expression may
be dependent on the light/dark cycle. In this case, assuming
the role of GlbN is for oxygen protection, this could be the
reason why more GlbN (and nitrogenase reductase) was seen in
the unaerated cultures than in the aerated cultures (all
samples were collected during the light period). In the
aerated cultures, the increased oxygen concentration may
have been too high to be controlled by a GlbN protection
system, resulting in the inhibition of gene expression. To
understand more precisely the function of GlbN, the role of
oxygen in regulating these genes, and the effects of the
light/dark cycle on GlbN synthesis, more detailed studies
are needed. Whatever the function of GlbN may be and

whatever the role of oxygen is in regulating expression of these genes, the results showed that in *N. commune* UTEX 584 increased amounts of GlbN were seen under the same growth conditions that resulted in the detection of increased amounts of nitrogenase reductase.

Based on the above results, one possible function of GlbN that can probably be ruled out is its functioning as an oxygen sensor to regulate gene expression. Such a function might require constitutive expression (possibly at lower levels then observed in this study) which was not seen with GlbN. However, since there is still not enough data to determine the function of GlbN, it may be prudent not to totally disregard this last possibility.

Whatever the function of GlbN, since it is present between *nifU* and *nifH* in *N. commune* UTEX 584 but is not in this region in *Anabaena* sp. PCC 7120 (see Figure 2), *glbN* expression and that of the *nif* genes around it, may be controlled in a fashion slightly different from that of *Anabaena* sp. PCC 7120. This is suggested by the fact that the two species have a number of different features in their DNA and RNA sequences. *Anabaena* sp. PCC 7120 not only lacks *glbN* between *nifU* and *nifH*, it also lacks the potential stem loop structures as those found in the *N. commune* UTEX 584 sequence. The potential stem loop structures in *N. commune*

UTEX 584 are predicted to form on the DNA strand that codes for the mRNA of these genes and are also predicted to form on the mRNA when the RNA sequence is analyzed. In this respect, the loops could function to regulate gene expression by acting as rho-dependent or rho-independent terminators or possibly by attenuation, as occurs with the tryptophan operon of *E. coli* and the histidine operon of *Salmonella typhimurium* (see reviews by Platt 1981, 1986 and Yanofsky 1981). Since the loops are not predicted in *Anabaena* sp. PCC 7120, this mode of regulation, if it occurs, may not be used by *Anabaena* sp. PCC 7120.

The first stem loop between *glbN* and *nifH* appears to be located in the DNA sequence at the position where RNA synthesis by SP6 occasionally terminates, producing the unexpected short transcript (Figures 4 A and B). While this loop is not predicted to form in the RNA being made by SP6, it has not been determined if the occasional termination of transcription by SP6 can be attributed to the transient formation of the loop in the DNA or some other feature of the template or RNA being made. Irrespective of this, since the loop is predicted on the mRNA as discussed above, this loop could function in *N. commune* UTEX 584 to separate the expression of *glbN* from *nifH* and/or to regulate *glbN* expression (whether or not *glbN* is expressed from its own

promoter or another promoter like *nifB*, depending on where *nifB* is in *N. commune* UTEX 584).

The second potential stem loop may regulate the expression of *nifH* relative to *nifD* to insure that excess nitrogenase reductase is made for the nitrogenase complex. In *Klebsiella pneumoniae* a small stem loop between *nifY* and *nifE* is believed to function as a transcription termination signal (Beynon et al. 1983). The results of the preliminary Northern transfer suggest that the loop between *nifH* and *nifD* may also be a transcription termination signal. However, with respect to both of the predicted loops, more information about the genes downstream of *nifD* and upstream of *nifU* is needed to investigate further this mode of regulation.

Another potential regulatory element in the DNA sequence is a direct repeat upstream of *glbN* in *N. commune* UTEX 584. This repeat is similar with regard to the order and content of purine and pyrimidine residues, to direct repeats (of unknown function: Mulligan and Haselkorn 1989) upstream of *fdxN* and *nifU* in *Anabaena* sp. PCC 7120 (Figure 17). As discussed above, the presence of the direct repeat upstream of *glbN* and the stem loop downstream of the gene may indicate that these features are involved in regulation of *glbN* expression relative to the *nif* genes around it.

Figure 17. Comparison of the direct repeats in N. commune
UTEX 584 and *Anabaena* sp. PCC 7120.

   The direct repeat in the *N. commune* UTEX 584 sequence
just upstream of *glbN* is labeled *Nc glbN*. The 21 in the *N.
commune* UTEX 584 sequence represents 21 bases between the
repeats. The sequences of the *Anabaena* sp. PCC 7120 direct
repeats that are upstream of *fdxN* and *nifU* are labeled AB
*fdxN* and AB *nifU* respectively. The *Anabaena* sp. PCC 7120
sequences are from Mulligan and Haselkorn 1989.

*Nc glbN*   AATTACGAATTACGAATTACGAATTACG-21-AATTACG

AB *nifU*   TCCCAATTCCCAGTACCCAGCCCCAGTACCCAATCCCCCAT

AB *fdxN*   ACCCAATCCCCAATCCCGATACCTAATCCCCAATCTCCAATCCCCAAT

While the direct repeat of *N. commune* UTEX 584 and those of *Anabaena* sp. PCC 7120 are slightly similar, it is not known if they function as promoters. If they are promoters, the similarity of the repeats may indicate that initiation of expression for these genes is controlled by similar factors or mechanisms. However, determination of the function of the *N. commune* UTEX 584 direct repeat, like that of the stem loops, requires additional information about the genes upstream and downstream of the *nifUHD* region.

In summary, while it is clear that GlbN is a novel heme protein and is associated with a molybdenum nitrogen fixation system, the function of GlbN and the mechanisms for regulating its expression and that of the genes around it, still need to be determined. Other interesting facets of this system to study include determining which cell types express GlbN, and when this occurs during the different stages of the life cycle. Growth conditions effecting the expression of GlbN relative to nitrogenase reductase expression could be studied in more detail. Whether or not GlbN is capable of transporting electrons could also be studied. Analysis of the transcripts produced from the *nifUHD* region should be done to determine what roles, if any, the direct repeat and the stem loop structures have in regulating expression of these genes. The information gained

from this additional work and other studies, will aid in determining the function and evolutionary significance of the *glbN* gene and this novel *nif* gene arrangement.

# References

Allen, M.M. (1988) Inclusions: Cyanophycin. In *Methods Enzymol.* 167:207-213.

Anonymous (1986) Formaldehyde gel conditions. In *Focus* Vol. 8 #2 p14. Edited by N. Sasavage.

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

Bergström, J., Eady, R.R., Lowe, D.J., and Thorneley, T.N.F. (1988) The Fe proteins of the Mo- and V-containing nitrogenases of *Azotobacter chroococcum*. In *Nitrogen fixation: Hundred years after.* p87. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

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

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

Bishop, P.E., Premakumar, R., Joerger, R.D., Jacobson, M.R., Dalton, D.A., Chisnell, J.R., and Wolfinger, E.D. (1988) Alternative nitrogen fixation systems in *Azotobacter vinelandii*. In *Nitrogen fixation: Hundred years after.* pp71-79. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Bishop, P.E., Premakumar, R., Dean, D.R., Jacobson, M.R., Chisnell, J.R., Rizzo, T.M., and Kopczynski, J. (1986) Nitrogen fixation by *Azotobacter vinelandii* strains having deletions in the structural genes for nitrogenase. *Science* 232:92-94.

Bishop, P.E., Jarlenski, D.M.L., and Hetherington, D.R. (1980) Evidence for an alternative nitrogen fixation system in *Azotobacter vinelandii*. *Proc. Natl. Acad. Sci. U.S.A.* 77:7342-7346.

Bothe, H. and Never, G. (1988) Electron donation to nitrogenase in heterocysts. In *Methods Enzymol.* 167:496-501.

Damerval, T., Gugliemi, G., Houmard, J., and Tandeau de Marsac, N. (1991) Hormogonia differentiation in the cyanobacterium *Calothrix*: A photoregulated developmental process. *The Plant Cell.* 3:191-201.

Dean, D.R., Setterquist, R.A., Brigle, K.E., Scott, D.J., Laird, N.F., and Newton, W.E. (1990) Evidence that conserved residues Cys-62 and Cys-154 within the *Azotobacter vinelandii* nitrogenase MoFe protein α-subunit are essential for nitrogenase activity but conserved residues His-83 and Cys-88 are not. *Molec. Microbiol.* 4:1505-1512.

Defrancesco, N. and Potts, M. (1988) Cloning of *nifHD* from *Nostoc commune* UTEX 584 and of a flanking region homologous to part of the *Azotobacter vinelandii nifU* gene. *J. Bacteriol..* 170:3297-3300.

De Marsac, N.T. and Houmard, J. (1983) Complementry chromatic adaptation: Physiological conditions and action spectra. In *Methods Enzymol.* 167:318-328.

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

Dickerson, R.E. and Geis, I. (1983) In *Hemoglobin; structure, function, evolution, and pathology.* Reading, Massachusetts:Benjamin/Cummings.

Dikshit, K.L., Spaulding, D., Braun, A., and Webster, D.A. (1989) Oxygen inhibition of globin gene transcription and bacterial haemoglobin synthesis in *Vitreoscilla. J. of Gen. Microbiol.* 135:2601-2609.

Donze, M., Haveman, J., and Schiereck, P. (1972) Absence of photosystem II in heterocysts of the blue-green alga *Anabaena. Biochem. Biophys. Acta* 256:157-161.

Eady, R.R., Robson, R.L., Pau, R.N., Woodley, P., Lowe, D.J., Miller, R.W., Thorneley, R.N.F., Smith, B.E., Gormal, C., Fisher, K., Eldridge, M., and Bergström, J. (1988) The vanadium nitrogenase of *Azotobacter chroococcum*. In *Nitrogen fixation: Hundred years after.* pp81-86. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Elhai, J. and Wolk, C.P. (1990) Developmental regulation and spatial pattern of expression of the

structural genes for nitrogenase in the cyanobacteria *Anabaena*. *EMBO Journal* 9:3379-3388.

Gallon, J.R. and Chaplin, A.E (1988) Recent studies on $N_2$ fixation by non-heterocystous cyanobacteria. In *Nitrogen fixation: Hundred years after.* pp183-188. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

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

Golden, J.W. and Wiest, D.R. (1988) Genome rearrangement and nitrogen fixation in *Anabaena* blocked by inactivation of *xisA* gene. *Science* 242:1421:1423.

Golden, J.W., Carrasco, C.D., Mulligan, M.E., Schneider, G.J., and Haselkorn, R. (1988) Deletion of a 55 kilobase-pair DNA element from the chromosome during heterocyst differentiation of *Anabaena* sp. strain PCC 7120. *J. Bacteriol.* 170:5034-5041.

Golden, J.W., Mulligan, M.E., and Haselkorn, R. (1987) Different recombination site specificity of two developmental regulated genome rearrangements. *Nature* (London) 327:526-529.

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

Hanahan, D. (1985) Techniques for transformation of *E. coli*. In *DNA cloning volume I: A practical approach.* pp.109-135. Edited by D.M. Glover. Washington D.C.:IRL Press.

Harlow, E. and Lane, D. (1988) Antibodies: A laboratory manual. Cold Spring Harbor Press.

Haselkorn, R. (1978) Heterocysts. In *Annu. Rev. Plant Physiology* Vol. 29:319-344.

Hausinger, R.P. and Howard, J.B. (1983) Thiol reactivity of the nitrogenase Fe-protein from *Azotobacter vinelandii*. *J. Biol. Chem.* 258:13486-13492.

Helber, J., Johnson, T.R., Yarbough, L.R., and Hirschberg, R. (1988a) Regulation of nitrogenase gene expression in anaerobic cultures of *Anabaena variabilis*. *J. Bacteriol*. 170:552-557.

Helber, J., Johnson, T.R., Yarbough, L.R., and Hirschberg, R. (1988b) Effects of nitrogenous compounds on nitrogenase gene expression in anaerobic cultures of *Anabaena variabilis*. *J. Bacteriol*. 170:558-563.

Henikoff, S. (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. In *Methods Enzymol*. 155:156-165.

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

Herdman, M. and Rippka, R. (1988a) Cellular differentiation: Hormogonia and baeocytes. In *Methods Enzymol*. 167:232-242.

Herdman, M. and Rippka, R. (1988b) Addendum to artical [22]. In *Methods Enzymol*. 167:851-853.

Innis, M.A. and Gelfand, D.H. (1990) Optimization of PCRs. In *PCR Protocols a guide to methods and applications*. pp3-12. Edited by M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White. New York:Academic Press. Inc.

Iwaasa, H, Takagi, T., and Shikama, K. (1989) Protozoan myoglobin from *Paramecium caudatum*. It's unusual amino acid sequence. *J. Mol. Biol*. 208:355-358.

Jacobson, M.R, Brigle, K.E., Bennett, L.T., Setterquist, R.A., Wilson, M.S., Cash, V.L., Beynon, J., Newton, W.E., and Dean, D.R. (1989) Physical and genetic map of the major *nif* gene cluster from *Azotobacter vinelandii*. *J. Bacteriol*. 171:1017-1027.

Joerger, R.D., Loveless, T.M., Pau, R.N., Mitchenall, L.A., Simon, B.H., and Bishop, P.E. (1990) Nucleotide sequence and mutational analysis of the structural genes for nitrogenase 2 of *Azotobacter vinelandii*. *J. Bacteriol*. 172:3400-3408.

Joerger, R.D., Jacobson, M.R., Premakumar, R. Wolfinger, E.D., and Bishop, P.E. (1989) Nucleotide sequence

and mutational analysis of the structural genes (*anfHDGK*) for the second alternative nitrogenase from *Azotobacter vinelandii*. *J. Bacteriol.* 171:1075-1086.

Kallas, T., Coursin, T., and Rippka, R. (1985) Different organization of *nif* genes in non-heterocystous and heterocystous cyanobacteria. *Plant Mol. Biol.* 5:321-329.

Kanemoto, R.H. and Ludden, P.W. (1984) Effect of ammonia, darkness and phenzine methosulfate on whole-cell nitrogenase activity and Fe protein modification in *Rhodospirillum rubrum*. *J. Bacteriol.* 158:713-720.

Kentemich, T., Haverkamp, G., and Bothe, H. (1991) The expression of a third nitrogenase in the cyanobacterium *Anabaena variabilis*. *Z. Naturforsch.* 46c:217-222.

Kentemich, T., Danneberg, G., Hundeshagen, B., and Bothe, H. (1988a) Evidence for the occurrence of the alternative, vanadium-containing nitrogenase in the cyanobacterium *Anabaena variabilis*. *FEMS Microbiol. Letters* 51:19-24.

Kentemich, T., Danneberg, G., Hundeshagen, B., and Bothe, H. (1988b) The occurrence of the alternative nitrogenase in the cyanobacteria *Anabaena variabilis*. In *Nitrogen fixation: Hundred years after.* p89. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. *Nature* (London) 227:680-685.

Lammers, P.J, McLaughlin, S., Papin, S., Trujillo-Provencio, C., and Ryncarz II, A.J. (1990) Developmental rearrangement of cyanobacterial *nif* genes: Nucleotide sequence, open reading frames, and cytochrome P-450 homology of the *Anabaena* sp. strain PCC 7120 *nifD* element. *J. Bacteriol.* 172:6981-6990.

Lammers, P.J, Golden, J.W., and Haselkorn, R. (1986) Identification and sequence of a gene required for a developmentally regulated DNA excision in *Anabaena*. *Cell* 44:905-911.

Lawry, N.H. and Simon, R.D. (1982) The normal and induced occurance of cyanophycin inclusion bodies in several blue-green algae. *J. Phycol.* 18:391-399.

Lazaroff, N. (1973) Photomorphogenesis of *Nostocacean* development. In *The Biology of Blue-Green Algae.* p270-319. Edited by N. G. Carr and B. A. Whitton. Oxford:Blackwell Scientific Publications.

Lazaroff, N. (1966) Photoinduction and photoreversal of the *Nostocacean* developmental cycle. *J. Phycol.* 2:7-17.

Lemberg, R. and Barrett, J. (1973) In *Cytochromes.* New York:Academic Press.

Lindblad, P., Atkins, C.A., and Pate, J.S. (1991) $N_2$-fixation by freshly isolated *Nostoc* from coralloid roots of the Cycad *Macrozamia riedlei* (Fisch. ex Gaud) Gardn. *Plant Physiol.* 95:753-759.

Ludden, P.W., Lehman, L., and Roberts, G.P. (1989) Reversible ADP-ribosylation of dinitrogenase reductase in a *NifD⁻* mutant of *Rhodospirillum rubrum*. *J. Bacteriol.* 171:5210-5211.

Ludden, P.W., Roberts, G.P., Lowery, R.G., Fitzmaurice, W.P., Sarri, L.L., Lehman, L., Lies, D., Woehle, D., Wirt, H., Murrell, S.A., Pope, M.R., and Kanemoto, R.H. (1988) Regulation of nitrogenase activity by reversible ADP-ribosylation of dinitrogenase reductase. In *Nitrogen fixation: Hundred years after*. pp157-162. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Maniatis, T., Fritsh, E.F., and Sambrook, J. (1982) Molecular cloning a laboratory manual. Cold Spring Harbor Ney York:Cold Spring Harbor press.

Meeks, J.C. (1988) Symbiotic associations. In *Methods Enzymol.* 167:113-121.

Meeks, J.C., Joseph, C.M., and Haselkorn, R. (1988) Organization of the *nif* genes in cyanobacteria in symbiotic association with *Azolla* and *Anthoceros*. *Arch. Microbiol.* 150:61-71.

Meeks, J.C., Wolk, C.P., Lockau, W., Schilling, N., and Shaffer, P. (1978) Pathways of assimilation of $[^{13}N]N_2$ and

$^{13}NH_4^+$ by cyanobacteria with and with out heterocysts. *J. Bacteriol.* 134:125-130.

Merrick, M.J. (1988) Organization and regulation of nitrogen fixation genes in *Klebsiella* and *Azotobacter*. In *Nitrogen fixation: Hundred years after.* pp293-302. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Mitsumi, A. and Kumazawa, S. (1988) Nitrogen fixation by synchronously growing unicellular aerobic nitrogen-fixing bacteria. In *Methods Enzymol.* 167:484-490.

Mitsumi, A., Kumazawa, S., Takahashi, A., Ikemato, H., Cao, S., and Arai, T. (1986) Strategy by which nitrogen fixing unicellular cyanobacteria grow photoautotrophically. *Nature* (London) 323:720-722.

Mulligan, M.E. and Haselkorn, R. (1989) Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. *J. Biol. Chem.* 264:19200-19207.

Pau, R., Mitchenall, L.A., and Robson, R.L. (1989) Genetic evidence for an *Azotobacter vinelandii* nitrogenase lacking molybdenum and vanadium. *J. Bacteriol.* 171:124-129.

Peterson, R.B. and Wolk, C.P. (1978) High recovery of nitrogenase activity and of $^{55}$Fe-labeled nitrogenase in heterocysts isolated from *Anabaena variabilis*. *Proc. Natl. Acad. Sci. U.S.A.* 75:6271-6275.

Platt, T. (1986) Transcription termination and regulation of gene expression. In *Annu. Rev. Biochem.* 55:339-372.

Platt, T. (1981) Termination of transcription and its regulation in tryptophan operon of *E. coli. Cell* 24:10-23.

Potts, M. and Bowman, M.A. (1985) Sensitivity of *Nostoc commune* UTEX 584 (Cyanobacteria) to water stress. *Arch. of Microbiol.* 141:51-56.

Poulos, T.L. (1988) Heme enzyme crystal structure. In *Heme Proteins 7.* p. 1-36. Edited by G. L. Eichhorn and L. G. Marzilli. New York:Elsevier.

Premakumar, R., Chisnell, J.R., and Bishop, P.E. (1989) A comparison of the three dinitrogenase reductases expressed by *Azotobacter vinelandii*. *Can. J. Microbiol.* 35:344-348.

Promega 1989/90 protocols and applications guide.

Reich, S. and Böger, P. (1989) Regulation of nitrogenase activity in *Anabaena variabilis* by modification of the Fe protein. *FEMS Microbiol. Letters* 58:81-86.

Rippka, R., Dervelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* 111:1-61.

Rippka, R. and Stanier, R.Y. (1978) The effects of anaerobiosis on nitrogen synthesis and heterocyst development by *Nostocacean* cyanobacteria. *J. Gen. Microbiol.* 105:83-94.

Scharf, S.J. (1990) Cloning with PCR. In *PCR Protocols a guide to methods and applications.* pp84-91. Edited by M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White. New York:Academic Press. Inc.

Scherer, S., Kerfin, W., and Böger, P. (1980) Increase in nitrogenase activity in the blue-green alga *Nostoc commune* (cyanobacteria). *J. Bacteriol.* 144:1017-1023.

Shively, J.M. (1988) Inclusions: Granules of polyglucose, polyphosphate and poly-$\beta$-hydroxybutyrate. In *Methods Enzymol.* 167:195-203.

Smith, R.L., Kumar, D., Xiankong, Z., Tabita, F.R., and Van Baalen, C. (1985) $H_2$, $N_2$ and $O_2$ metabolism by isolated heterocysts from *Anabaena* sp. strain CA. *J. Bacteriol.* 162:565-570.

Simon, R.D. (1971) Cyanophycin granules from the blue-green alga *Anabaena cylindrica*: A reserve material consisting of copolymers of aspartic acid and arginine. *Proc. Natl. Acad. Sci. U.S.A.* 68:265-267.

Simon, R.D. (1973) The effects of chloramphenicol on the production of cyanophycin granule polypeptide in the blue-green alga *Anabaena cylindrica. Arch. Microbiol.* 92:115-122.

Söderbäck, E., Lindblad, P., and Bergman, B. (1990) Developmental patterns related to nitrogen fixation in the *Nostoc-Gunnera magellanica* Lam. symbiosis. *Planta* 182:355-362.

Stal, L.J. and Krumbein, W.E. (1985) Nitrogen activity in the non-heterocystous cyanobacterium *Oscillatoria* Sp. grown under alternating light-dark cycles. *Arch. Microbiol.* 143:67-71.

Stanier, R.Y and Cohen-Bazire, G. (1977) Photosynthetic prokaryotes: The cyanobacteria. *Annu. Rev. Microbiol.* 31:225-274.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1991) Use of T7 RNA polymerase to direct expression of cloned genes. In *Methods Enzymol.* 185:60-89.

Tabita, F.R., Van Baalen, C., Smith, R.L., and Kumar, D. (1988) *In vivo* protection and recovery of nitrogenase from oxygen in heterocystous cyanobacteria. In *Nitrogen fixation: Hundred years after.* pp189-194. Edited by H. Bothe, F.J. de Bruijn and W.E. Newton. Stuttgart:Gustav Fischer.

Thomas, J., Meeks, J.C., Wolk, C.P, Shaffer, P.W., Austin, S.M., and Chien, W.-S. (1977) Formation of glutamine from [$^{13}$N]ammonia, [$^{13}$N]dinitrogen, and [$^{14}$C]glutamate by heterocysts isolated from *Anabaena cylindrica. J. Bacteriol.* 129:1545-1555.

Tsubamoto, T., Matsuoka, A., Yusa, K., and Shikama, K. (1990) Protozoan myoglobin from *Paramecium caudatum.* Its autoxidation reaction and hemicrome formation. *Eur. J. Biochem.* 193:55-59.

Wakabayashi, S., Matsubara, H., and Webster, D.A. (1986) Primary sequence of a dimeric bacterial haemoglobin from *Vitreoscilla. Nature* (London) 322:481-483.

Wang, S-Z., Dean, D.R., Chen, J-S., and Johnson, J.L. (1991) The N-terminal and C-terminal portions of nifV are encoded by two defferent genes in *Clostridium pasteurianum. J. Bacteriol.* 173:3041-3046.

Webster, D.A. (1988) Structure and function of bacterial hemoglobin and related proteins. In *Heme Proteins*

7. p.245-265. Edited by G. L. Eichhorn and L. G. Marzilli. New York:Elsevier.

Whitton, B.A. (1973) Interactions with other organisms. In *The Biology of Blue-Green Algae.* p415-433. Edited by N. G. Carr and B. A. Whitton. Oxford:Blackwell Scientific Publications.

Wolk, C.P., Austin, S.M., Bortins, J., and Galonsky, A. (1974) Autoradiographic localization of $^{13}$N after fixation of $^{13}$N-labeled nitrogen gas by a heterocyst forming blue-green alga. *J. Cell Biol.* 61:440-453.

Wray, W., Boulikas, T., Wray, V.P., and Hancock, R. (1981) Silver staining of proteins in polyacrylamide gels. *Anal. Biochem.* 118:197-203.

Yanofsky C. (1981) Attenuation in the control of expression of bacterial operons. *Nature* (London). 289:751-758.

## Appendix

**Preliminary Studies of the Life Cycle of *N. commune* UTEX 584**

During the initial stages of this project, culturing of
*N. commune* UTEX 584 was done in a BRL Air Lift fermenter in
the nitrogen-free medium $BG11_o$ (see Materials and Methods
for the composition of $BG11_o$) since we were interested in
studying the expression of nitrogen fixation genes. What
prompted the study of the life cycle changes were two
observations. The first was that upon inoculation of the
fermenter with starter cultures from shaker flasks, the
cells changed from a dark green color, in the shaker
culture, to a pale yellow-green after about 16 to 24 hours
in the fermenter. The second reason was that microscopic
observation showed that along with this color change there
was also a dramatic change in the cells' morphology.

The implications of these observations raised two
questions. The first was related to the fact that growth of
the air lift cultures was monitored by the concentration of
chlorophyll *a* (Chl *a*) in µg/ml of culture. Therefore, is
this parameter measuring the increase in the culture density
or would it be a measure of the shift between cell types?
Since the µg Chl *a*/ml of culture can be expressed in terms
of cell mass as shown in equation 1 below (<u>Chl *a*</u> represents
µg Chl *a*/ml of culture and g of cells represents grams wet

105

weight), if the different cell types do not have the same Chl *a* content, the Chl *a* concentration of the culture will be related to the Chl *a* content of the different cell types as expressed in equation 2 (grams wet weight of the different cell types is g A or g B cells, etc.).

Equation 1

$$\underline{Chl\ a} = \left( \frac{g\ of\ cells}{ml\ of\ culture} \right) \times \left( \frac{\mu g\ Chl\ a}{g\ of\ cells} \right)$$

Equation 2

$$\underline{Chl\ a} = \left( \frac{g\ A\ cells}{ml\ culture} \times \frac{\mu g\ Chl\ a}{g\ A\ cells} \right) + \left( \frac{g\ B\ cells}{ml\ culture} \times \frac{\mu g\ Chl\ a}{g\ B\ cells} \right) etc.$$

The second question was: at what stage in the life cycle did heterocysts form? Since we were interested in studying the expression of nitrogen fixation genes, it would be useful to determine the conditions producing this shift and the stage at which heterocysts were most active and abundant.

To determine why these changes were occurring, observation of when the shifts occurred during the normal

usage of the cultures was undertaken. The following
observations were made.

Once the inoculum in the fermenter recovered as
indicated by its becoming a dark green, it was noticed that
the predominant cell type was a packet of cells referred to
as the aseriate form. Typical aseriates are shown in Figure
18. This form consists of packets of vegetative cells that
can be up to 5 μ in diameter and are incased in a thick
sheath. Aseriate packets have heterocysts at their ends and
often several packets may be connected by heterocysts.
Heterocysts at this stage may not be viable, as indicated by
their inability to form Formazan crystals when treated with
2,3,5,-triphenyl-2-tetrazolium chloride (Potts and Bowman
1985).

Upon dilution with fresh $BG11_o$ medium, within 16 to 24
hours, the aseriate packets ruptured and released their
vegetative cells as filaments (Figure 19). These filaments
soon lost their heterocysts and started to divide rapidly
(Figure 20). The rapid rate of division resulted in a slight
decrease in the cell diameter and after the heterocysts were
lost the filaments became motile. The motile filaments are
referred to as hormogonia. Hormogonia are the predominant
cell type which forms after inoculation when the culture
turns pale yellow-green. This stage also appears to be very
light sensitive since the light intensity needs to be kept

Figure 18. Typical *N. commune* UTEX 584 aseriate packets in a fermenter culture grown on BG11$_o$.

Mature aseriate packets are indicated as As. Some filaments in early stages of aseriate formation are also seen. Bar, 10 µm.
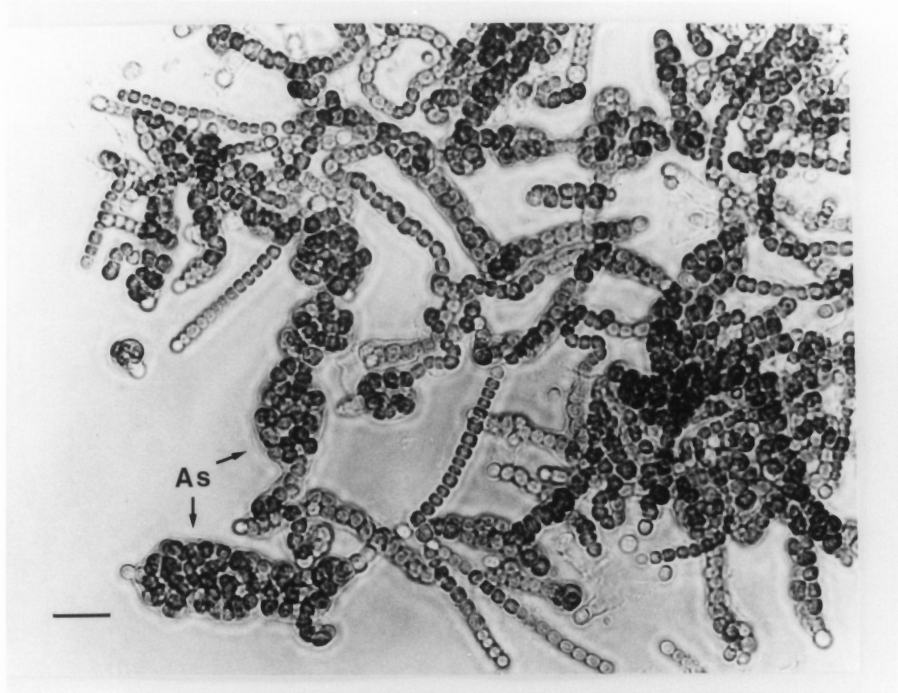
Figure 19. Release of a filament from an aseriate packet.

    A mature aseriate packet that has ruptured after introduction into fresh media is releasing a filament of vegetative cells. The emerging filament is indicated as Nf. Also shown are young rapidly dividing hormogonial filaments (YHo) that have not lost their heterocysts and are not yet motile. Bar, 10 μm.
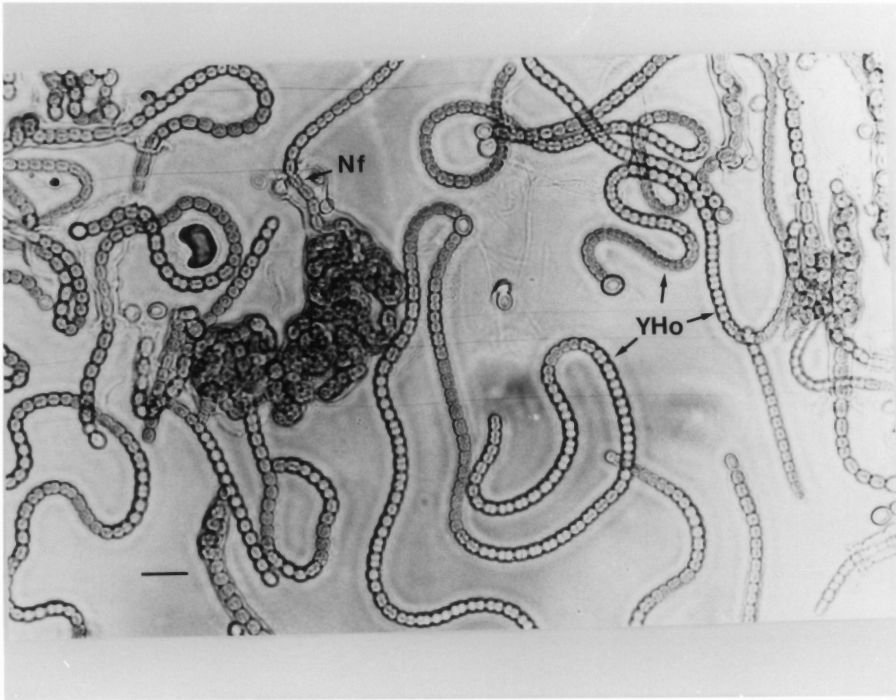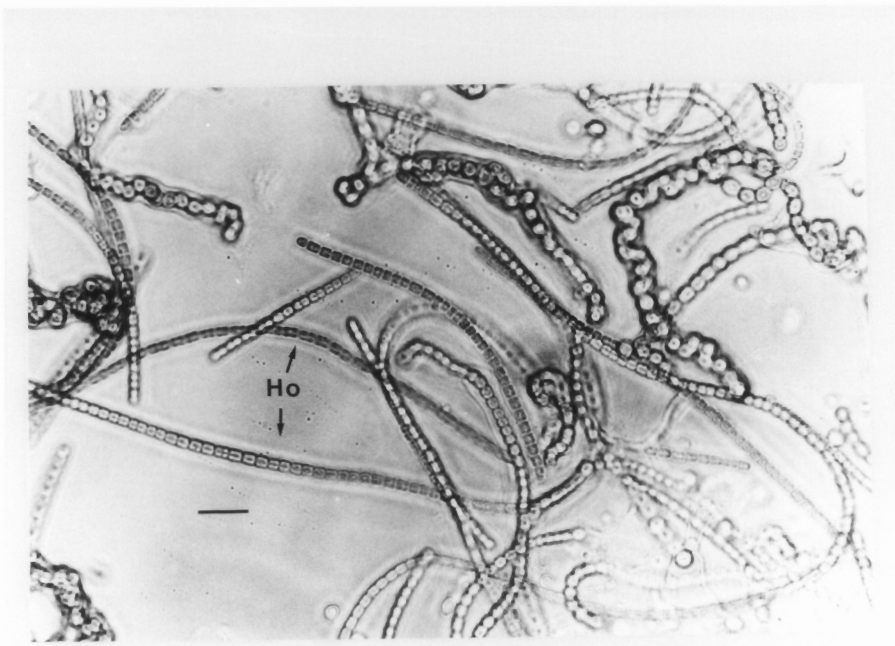
Figure 20. Motile hormogonia.

Rapidly dividing motile hormogonia lacking heterocysts are labeled Ho. Bar, 10 μm.

as low as possible or the culture will not survive. Concomitant with this increased light sensitivity and change in color is a drop in the cells' Chl *a* content. For example, the culture pictured in Figures 19 and 20 had a drop in its Chl *a* content from 653 µg/g wet weight of cells at the aseriate stage to 173 µg/g wet weight of cells after shifting to hormogonia.

Within 24 to 48 hours after shifting to hormogonia, the filaments lost their motility and started to divide more slowly. The cells began to increase in diameter and heterocysts started to form as shown in Figure 21. At the same time, a sheath started to form around the filament. The heterocysts appeared to degenerate but still remained attached to the filament. As these changes occurred, the cells' Chl *a* content increased and the cells continued to divide at a slower rate, and eventually developed into mature aseriates as those shown in Figure 18, thereby completing the cycle.

As fermenter cultures grew older, diluting them with fresh media did not stimulate the shift from aseriate to filaments as strongly as with younger cultures, suggesting that there may be factors released by the cells which are inhibitors of hormogonia formation. In young as well as old cultures, decreasing the light intensity did not stimulate

Figure 21. The shift back to the aseriate form.

The filaments pictured have lost their motility and are dividing more slowly than the motile hormogonia. Heterocysts are starting to form and are labeled He. Bar, 10 μm.
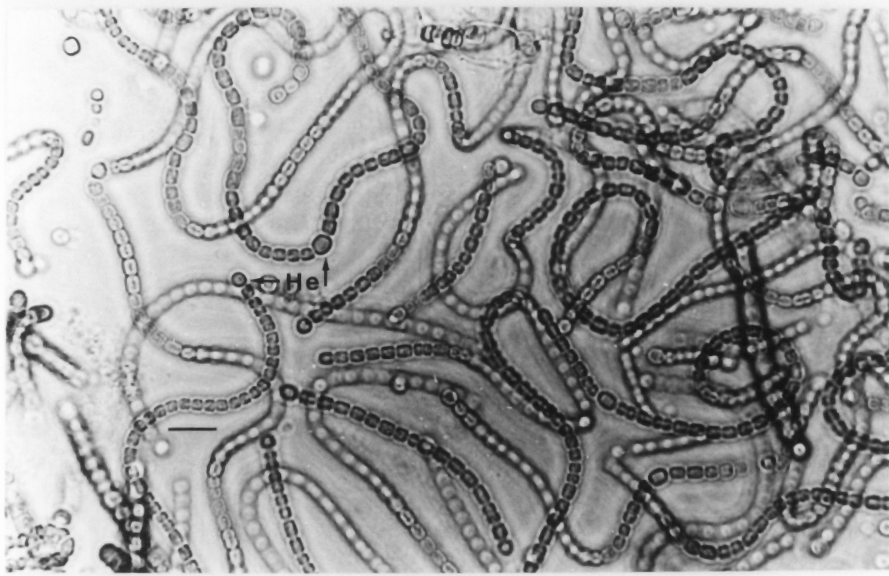
the shift to hormogonia and increasing the light intensity appeared to be detrimental to the hormogonia stage. As discussed above, hormogonial light sensitivity was especially noticeable after inoculating the fermenter. In most cases, upon inoculation of 1.8 liters of fresh media with a starter culture, virtually all of the inoculum formed hormogonia. Depending on the cell density of the inoculum, several days to a week or more were needed before enough aseriates formed after which time the light intensity could be increased.

In summary, these results suggest that the amount of nutrients and/or the accumulation of cellular products, are among the factors capable of controlling shifts between the different stages in the *N. commune* UTEX 584 life cycle. The possible physiological implications of these changes is best illustrated by following the shifts in the life cycle along with the type of culture conditions that may be present at the time these changes take place.

For example, when a young *N. commune* UTEX 584 fermenter culture in the aseriate stage is diluted with fresh $BG11_o$ medium or when the fermenter is inoculated with a washed starter culture, within 16 to 24 hours the aseriate packets rupture and release filaments which start to divide rapidly, eventually loose their heterocysts and become motile

hormogonia. The lack of heterocysts indicates that hormogonia do not need to fix nitrogen. However, the rapid division, motility, and growth occurring during this stage will require nitrogen and nitrogen-containing compounds in addition to other nutrients to support this activity. Since the only nitrogen source in $BG11_o$ is $N_2$, and hormogonia lack heterocysts, nitrogen must be coming from internal stores or else the vegetative cells of hormogonia can fix nitrogen.

Vegetative cells are known to have a rich store of nitrogen in the form of cyanophycin (Simon 1971). Cyanophycin is a copolymer consisting of an aspartic acid backbone with arginine side chains and is not made by translation (Simon 1973). Cyanophycin and a number of other stored metabolites in vegetative cells could be used to support the activity of hormogonia (Allen 1988, Shively 1988, Lawry and Simon 1982). Utilization of these internal stores during this stage of the life cycle may decrease the need for photosynthesis which may be one reason why the Chl *a* content of cells in the hormogonia stage decreases. Whether this decrease is from the breakdown of Chl *a* or from the cells dividing faster than Chl *a* can be synthesized, has not been determined.

The loss of motility and the slowing down of cell division may indicate a depletion of these internal stores.

At this point heterocysts form, indicating the need to resume nitrogen fixation. This will allow the filaments to replenish their cyanophycin reserve and support continued growth at a slower division rate.

As the filaments start to revert back to the aseriate stage they begin to produce their thick sheath. Sheath synthesis will require carbohydrate which may be another reason for the increase in the cells' Chl *a* content as the aseriate stage is formed. This demonstrates that the different cell types have different Chl *a* contents and that the culture's Chl *a* concentration is related to the cells' Chl *a* content as shown in equation 2. This indicates that the parameter of µg Chl *a*/ml of culture at some phases of culture growth will be more related to the shift between the different stages of the life cycle than to the actual growth rate of the culture.

Once the aseriate stage has formed, cultures can continue to grow at a slower rate in the same media for long periods in the fermenter as well as in shaker flasks. However, as discussed above, dilution of these older cultures with fresh media does not produce near as large a shift to hormogonia as when the cells are washed with fresh media before inoculation or subculturing. This supports the idea that in addition to the lack of nutrients in older

cultures, some factor or factors released by the cells may be able to inhibit hormogonia formation. Similar observations suggesting that cellular factors may affect hormogonia development in other species of cyanobacteria have been reported (Herdman and Rippka 1988A, 1988B).

While the data of this study indicate the possible role of nutrients and cell products in regulating these developmental changes, other studies have shown that light may be involved with some species such as *Calothrix* sp. PCC 7601 and *Nostoc muscorum* (Damerval et al. 1991, Lazaroff 1973). In *N. muscorum*, the shift from aseriate to hormogonia can be induced or inhibited by white light (Lazaroff 1973, 1966). It was found that specific wavelengths around 650 nm stimulated hormogonia formation while wavelengths from 480-580 nm appeared to inhibit differentiation. This photoinduction or inhibition has also been described for *N. commune* 584 (Lazeroff 1973). This study did not specify if this strain was UTEX 584. Even still, the stimulation by light seems contradictory to the findings in this study which showed light to be detrimental to the hormogonia stage. The different results between these studies may in part be due to the different culture conditions used. In this study the cells were grown in a well aerated nitrogen-free liquid media in continuous white light, while in the

earlier studies of Lazaroff the cells were grown on agar plates in the dark with sugars as the carbon source before being subjected to the light stimulus. However, whether or not specific wavelengths of light have the ability to induce or inhibit hormogonia formation of fermenter grown cells has not been studied.
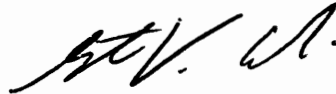
In nature, in addition to light and the availability of nutrients, other factors such as interactions with bacteria, fungi and other organisms may also be involved in regulating these changes (Meeks 1988, Whitton 1973). Whatever the stimulus for the change, as the changes occur, there should be differences in the patterns of gene expression which to some degree may be species specific. For example, when *Anabaena* sp. PCC 7120 is grown in nitrogen-free media, it always contains filaments with heterocysts which should be expressing its nitrogenase genes. When *Anabaena* sp. PCC 7120 was shifted to fresh nitrogen-free media, there was no change in the filaments' morphology (data not shown). *N. commune* UTEX 584 under the same conditions goes through several morphological changes involving the formation of heterocyst free hormogonia. Since hormogonia do not possess heterocysts, this suggests that the nitrogen fixation genes in *N. commune* UTEX 584 are not expressed during this stage of the life cycle even though the cells are in nitrogen-free

media (with $N_2$ as the only source of nitrogen). As discussed above, the utilization of cyanophycin by hormogonia as a source of nitrogen (which may have been built up during the aseriate stage), may negate the need of hormogonia to fix nitrogen.

In conclusion, while the complex life cycle of *N. commune* UTEX 584 may be a reflection of this species' having different genes and different patterns of gene expression than *Anabaena* sp. PCC 7120, how these differences will effect gene expression, organization and other physiological processes between the two species still requires more detailed study.

Vita

Stephen V. Angeloni was born in Washington D.C. on January 10, 1955 to Frederick J. and Dorothy E. Angeloni. After graduating from Florida State University in 1980 he worked as a research assistant at the University of Miami Medical School in the Department of Physiology and Biophysics before returning to Florida State to work on a Masters degree in Molecular Biology. His masters research involved determining several factors that controlled amoeboid motility in the soil amoeba, *Naegleria gruberi*. After completing his Masters Degree in 1984, he started to work for Dr. Potts as a lab technician at Florida State and moved with the lab to the Department of Biochemistry and Nutrition at Virginia Polytechnic Institute and State University in 1985. He remained a technician until he was accepted into the Ph.D. program in the Department of Biochemistry and Nutrition in 1986.

Stephen V. Angeloni