Cyanobacterial Hydrogen Metabolism

Regulation and Maturation of Hydrogenases

ELLENOR DEVINE
Dissertation presented at Uppsala University to be publicly examined in Håggsalen, room 10132, Ångström laboratorium, Lägerhyddsvägen 1, Uppsala, Friday, January 14, 2011 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

In times with elevated CO₂ levels and global warming there is a need of finding alternatives to carbon based energy carriers. One such environmental friendly solution could be H₂ produced by living organisms. Cyanobacteria are good candidates since they can produce H₂ from sunlight and water through the combination of photosynthesis and H₂ producing enzymes i.e. nitrogenases and/or [NiFe]-hydrogenases. This thesis investigates the maturation and transcriptional regulation of [NiFe]-hydrogenases in cyanobacteria, with a special focus on hydrogenase specific proteases.

The core of all hydrogenases consists of the small and large subunit. The large subunit in which the catalytic site is located goes through an extensive maturation process which ends with a proteolytic cleavage performed by a hydrogenase specific protease (HupW/HoxW). This thesis shows that within the maturation process of hydrogenases, the proteolytic cleavage is probably the only step that is specific with respect to different types of hydrogenases i.e. one type of protease cleaves only one type of hydrogenase. Further in-silico analysis revealed that these proteases and the hydrogenases might have co-evolved since ancient time and that the specificity observed could be the result of a conserved amino acid sequence which differs between the two types of proteases (HupW/HoxW).

A number of different transcription factors were revealed and shown to interact with the promoter regions of several of the genes encoding maturation proteins. The results indicate that the hydrogenase specific proteases are regulated on a transcriptional level in a similar manner as the hydrogenases they cleave. This thesis contributes with knowledge concerning transcriptional regulation and protein regulation of hydrogenases which will be useful for designing genetically engineered cyanobacteria with an improved and adjustable H₂ production.

Keywords: cyanobacteria, hydrogenase, hydrogenase specific protease, maturation, regulation, transcriptional studies, hupW, hoxW, hyp, NtcA, CalA

Ellenor Devine, Department of Photochemistry and Molecular Science, Microbial Chemistry, Box 523, Uppsala University, SE-75120 Uppsala, Sweden.

© Ellenor Devine 2011

ISSN 1651-6214
ISBN 978-91-554-7968-8
urn:nbn:se:uu:diva-134246 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-134246)
In memory of Grandpa (Morfar)
Gottfrid Teodor Olsson
(1919-1995)
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
Contributions

The author of this thesis has done the following contributions to the articles;

Paper I: Planned and performed all experiments, wrote the majority of the article.
Paper II: Planned and performed all experiments, except for hydrogen measurements, and wrote the majority of the manuscript.
Paper III: Planned and performed all experiments, wrote the majority of the manuscript.
Paper IV: Performed western blot studies and did part of the writing of the manuscript.
Paper V: Planned and performed all analysis, wrote the majority of the manuscript.
Paper VI: Performed all \textit{in-silico} studies and did part of the writing of the article.

Author’s request

This work, a result of the author’s research for a doctoral degree at Uppsala University, Sweden, contains unpublished material. The author reserves her right to publish it at appropriate time. If any of the unpublished material is taken for reproduction, further study or modification it must be duly acknowledged giving references to this thesis.
# Contents

Introduction ................................................................................................... 11  
Cyanobacteria ........................................................................................... 13  
Model organisms ...................................................................................... 13  
  *Nostoc punctiforme* ATCC 29133 ....................................................... 13  
  *Nostoc* sp. strain PCC 7120 ................................................................. 14  
  Working with cyanobacteria ................................................................. 15  
Nitrogen fixation ...................................................................................... 16  
  The Nitrogenase................................................................................... 16  
  Nitrogen control and nitrogen fixation ................................................ 18  
  The Heterocyst.................................................................................... 19  
The hydrogen metabolism ........................................................................ 20  
  The uptake hydrogenase .................................................................... 21  
  The bidirectional hydrogenase ............................................................. 22  
  The hydrogenase maturation process ................................................... 23  
  The transcriptional regulation of hydrogenases ..................................... 26  
  The uptake hydrogenase .................................................................... 26  
  The bidirectional hydrogenase ............................................................. 27  
Aim of this Thesis ......................................................................................... 29  

## Result and Discussion

- Part I – The specificity of hydrogenase maturation .............................. 30  
  - The evolutionary history of hydrogen metabolism and hydrogenase 
    specific proteases (Paper I)................................................................. 30  
  - The specificity of hydrogen specific proteases (Paper I, II) .......... 32  
- Part II – Nitrogen control and hydrogen metabolism .............................. 37  
  - Transcriptional regulation of the uptake hydrogenase and its specific 
    protease (Paper I, III, V)................................................................. 37  
  - Transcriptional regulation of the bidirectional hydrogenase and its 
    specific protease (Paper I, III) ......................................................... 40  
  - The complexity of the NtcA family (Paper V) ................................. 41  
  - CalA, a repressor of the accessory genes, important for hydrogenase 
    maturation? (Paper III, VI) .............................................................. 45  

Summary and future outlook ................................................................... 48  

Svensk sammanfattning ............................................................................. 51
Abbreviations

2-OG  2-Oxoglutarate
5’RACE  5’ rapid amplification of cDNA ends
aa  Amino acids
Abr  Antibiotic resistance protein
ATCC  American type culture collection
ATP/ADP/AMP  Adenosine triphosphate/diphosphate/ monophosphate
BLAST  Basic local alignment search tool
bp  Base pair
Cal  Cyanobacterial AbrB like protein
CAP  Catabolite activator protein
CO  Carbon monoxide
CO2  Carbon dioxide
CN  Cyanide
cNMP  Cyclic nucleotide monophosphate
CRP  Cyclic AMP receptor protein
DNR  Dissimilative nitrate respiration protein
EMSA  Electrophoretic mobility shift assay
FeS  Iron sulphur
FNR  Fumarate-nitrate reduction protein
GFP  Green fluorescent protein
H2  Hydrogen gas
HGT  Horizontal gene transfer
Hox  Hydrogen oxidation protein
Hup  Hydrogen uptake protein
Hyp  Hydrogenase pleiotropic protein
IHF  Integration host factor
Lex  Lambda excision protein
LUCA  Last universally common anchestor
N2  Nitrogen gas
NADH/NAD+  Nicotinamide adenine dinucleotide (reduced/unreduced)
NADPH/NADP+  Nicotinamide adenine dinucleotide phosphate (reduced/unreduced)
NH3/NH4  Ammonia/ammonium
nt  Nucleotide
Ntc  Nitrogen control protein
$O_2$  Oxygen gas
PAS  Named after three proteins were this domain occurs; Per, Arnt, Sim
PCC  Pasteur culture collection
PCR  Polymerase chain reaction
Pi/PPi  Orthophosphate/pyrophosphate
PSI/II  Photosystem I/II
RT  Reverse transcriptase
tsp  Transcriptional start point
Introduction

Each year, the amount of carbon we humans release through the burning of fossil fuels is a staggering 400 times the amount that is fixed globally by earth’s biotopa (27). The effect this release of carbon has and will have on the earth, in the form of global warming and pollution have been the focal point of many studies and has been summarized by the Intergovernmental Panel on Climate Change (IPCC) (http://www.ipcc.ch/). The results from these studies are worrying since they suggest that the consequences of global warming will be severe i.e. raised sea levels, severe draught in several areas of the world, significant loss of biodiversity, changes in wind and water stream patterns and ultimately will be a serious threat to human health. Furthermore, with increasing energy demands and depleting coal, gas, and oil reserves it is becoming desperately clear that something needs to be done (107).

Of all the alternative renewable energy sources capturing energy directly from the sun is by far the most attractive option. Every hour more energy hits the earth from the sun then we humans use per year (22). For solar energy utilization two problems need to be solved, 1) capturing and conversion and 2) storage. The energy captured could in theory be stored either in batteries, mechanically (e.g. to pump water uphill) or as thermal energy but unfortunately several of these options are today not a cost efficient solution (61). Another option is to store the energy in the form of hydrogen. Hydrogen has the highest amount of energy per unit mass (MJ/kg) than any other conventional fuel and can be used in fuel cells to generate an electrical current. The oxidation of hydrogen only produces water which makes it one of the cleanest fuels known to man. Hydrogen combustion might be environmental friendly but today 90% of the hydrogen gas produced is from fossil fuels (39) and alternative ways of hydrogen production is therefore needed.

It has long been known that several bacterial strains have the ability to produce hydrogen under certain growth conditions (38). Cyanobacteria together with green algae are the only organisms which are able to combine oxygenic photosynthesis with the production of molecular hydrogen (116). Or more clearly speaking, they are like a 2-in-1 solution to the energy problem of the world, combining the capturing of the sun’s energy with the conversion of energy to hydrogen. Together with cyanobacteria’s minimal and inexpensive growth requirement they are a promising candidate for biohydrogen production on commercial level.
Several hurdles need to be overcome though before bio hydrogen production can be performed on commercial level. The early stages of photosynthesis is very efficient, (>95%) but a considerably amount of energy is lost in the later stages leading up to the fixation of CO₂ (22, 29). This makes direct biophotolysis the most appealing solution whereby the enzymes responsible for hydrogen production are directly coupled to photosynthesis (38). Not using any intermediates like sugar or starch, the production line within the cell gets shorter, lowering the loss of energy to bi-products and heat.

The major drawback is the oxygen sensitivity of the enzymes involved in hydrogen production which means that they will be reversible or irreversible inactivated by the oxygen produced by PSII. In cyanobacteria two sets of enzymes are used to generate hydrogen, the nitrogenase, a key enzyme of the nitrogen fixation process, and the bidirectional [NiFe]-hydrogenase. Another important enzyme involved in hydrogen metabolism is the uptake [NiFe]-hydrogenase which is strongly connected to nitrogen fixation since it recycles the hydrogen produced by the nitrogenase. In some cyanobacteria the problem with the oxygen sensitive nitrogenase has resulted in the evolution of heterocysts, specialized microaerobic cells for nitrogen fixation harbouring the nitrogenase. They have also evolved an efficient system for transferring energy from vegetative cells to heterocyst and 50% percent of the reductants produced by the photosynthetic apparatus in vegetative cells in a filament are believed to go to heterocyst development and nitrogen production (130). This makes an in-direct solution possible for hydrogen production by microorganisms whereby the energy captured by the sun in surrounding cells would be efficiently transferred to the hydrogen producing enzymes localised in heterocysts.

Both the nitrogenase and the bidirectional hydrogenase could be used for hydrogen production as well as the introduction of a foreign hydrogenase. However, to make the production of biohydrogen commercial feasible several things needs to be improved e.g. the nitrogenase or the bidirectional hydrogenase efficiency, antenna size of PSI and PSII, lowering the energy consuming pathways competing with hydrogen production, higher amount of hydrogen producing enzymes etc. The solution includes genetic engineering and several studies on both gene and protein level of hydrogen metabolism in cyanobacteria, and how it can be improved, have been performed over the years (116).

For commercial production by genetic engineered cyanobacteria, not only the enzymes and proteins need to be studied and improved. A deeper knowledge of the transcriptional control and regulation of genes involved in hydrogen metabolism is also important. The results from these studies can in the future turn out to be an important tool when designing the perfect hydrogen producing organism. It will give us the knowledge and power to turn the genes encoding these enzymes on and off at our will but also to control the amount of a specific enzyme within the cell. By knowing more about differ-
ent protein specificity we will also understand which components are crucial for a functional hydrogenase and therefore cannot be exchanged or left out. If for example a foreign hydrogenase is placed within filamentous cyanobacteria, what extra components are needed for the hydrogenase to function? Further, what regulatory controls already existing within the cyanobacteria can we use to control the expression of this foreign hydrogenase to maximize the effect?

Cyanobacteria

The phylum cyanobacteria is in many ways unique with their extraordinary diverse morphology, capacity to form symbiosis with a variety of organism and ability to survive in a wide range of environment (46). Once they revolutionized the world by being the first organism to perform oxygenic photosynthesis (~2.3 billion years ago), and with that they changed the world forever by raising the oxygen levels in the atmosphere (56). They have even made a more direct contribution to eukaryotic life since they are believed to be the origin of chloroplasts in plants (92).

While many are unicellular several strains grow in colonial or filamentous forms, some even with the ability of differencing cells to meet special needs (74, 96). Additional to the vegetative cell which is the normal state under favourable conditions, cyanobacterial cells can develop into akinetes, hormogonia and heterocysts (73). Akinetes, which are spore-like resting cells, are formed when conditions are harsh and can resist both cold and draught to eventually develop into vegetative cells again when the surrounding becomes more favourable. Hormogonia are short gliding filaments which are formed as a response to stress and are important for plant-cyanobacteria symbiosis (17). Last but not least are the heterocysts, which are specialised cells for nitrogen fixation and are further discussed under the section “Nitrogen fixation”.

Model organisms

*Nostoc punctiforme* ATCC 29133

The filamentous symbiotic *Nostoc punctiforme* ATCC 29133, also known as *Nostoc punctiforme* PCC 73102 (*N. punctiforme*) was originally isolated from the cycad *Macrozamia* (97). It is sequenced with a genome of 9.06 Mb, being one of the largest microbial genomes so far and is distributed on one chromosome and five plasmids (74). This nitrogen fixing cyanobacterium has the ability to differentiate cells from the normal vegetative state into hormogonias, akinetes or heterocysts depending on growth conditions (95).
As a symbiotic organism it has decreased growth rate, high metabolism and a higher heterocyst frequency which can reach as much as 50% of the cells (21). These abilities make N. punctiforme highly interesting from a hydrogen producing point of view since many of them (e.g. growth rate, high heterocyst frequency etc) coincide with the requirements of the ideal heterocyst forming hydrogen producing strain. Further, it can be grown in bioreactors and techniques for genetic modifications are readily available. Like many cyanobacteria they have low growth requirement and are photoautotrophic, even though N. punctiforme in particular can also grow heterotrophically (115).

*Nostoc* sp. strain PCC 7120

The origin of the free living *Nostoc* sp. strain PCC 7120 (*Nostoc* PCC 7120) is unknown but as *N. punctiforme*, it is a filamentous, nitrogen fixing strain that has the ability to form heterocysts (Fig. 1). It was previously thought to belong to the genus *Anabaena*, and is therefore also known as *Anabaena* sp. strain PCC 7120, until it was reclassified as belonging to *Nostoc*. However, contrary to *N. punctiforme*, it is not a symbiotic strain and it can form neither akinetes nor hormogonia and it is an obligate photoautotroph (95).

The genome of the size 7.21 Mb has been sequenced and contains one chromosome and six plasmids (55). Several techniques are available for genetic modifications and hence the strain has been used extensively for studies i.e. nitrogen fixation, cell differentiation and the formation of heterocyst (35).

![Figure 1](image_url). *Nostoc* sp. strain PCC 7120. Filaments of *Nostoc* PCC 7120 with vegetative cells and heterocysts (H).
Working with cyanobacteria

Cyanobacteria have been used as a model organism for many years especially for studies of the nitrogen fixation process, circadian clocks and the photosynthesis apparatus(46). Several cyanobacterial strains are easy to cultivate and only need air, water, sun and a few minerals to survive. Since a growing number of cyanobacteria are also being fully sequenced, including *N. punctiforme* and *Nostoc* PCC 7120 which have been used in this study, bioinformatics studies on a much larger scale are now possible.

One advantage from a genetic engineering perspective is that new genetic material can easily be introduced into many cyanobacteria by “natural” transformation, electrophoration or conjugation. Once inside the cell, the gene of interest can be modified or interrupted by an antibiotic cassette through homologues recombination(46). Successful modifications of the genome and introduction of foreign genes (e.g. GFP) have been performed in many strains including *N. punctiforme*, *Nostoc* PCC 7120 and *Synechocystis* sp. strain PCC 6803 (64, 70, 116).

These advantages don’t mean that there are any no drawbacks. Since cyanobacteria belong to the bacterial superkingdom it is easy to make the assumption that they are all fast growing and as easy to transform like other bacterial model organisms e.g. *Escherichia coli*. This might be true when comparing with many eukaryotes but the doubling time for cyanobacterial strains is usually around 20-24 h. This can be compared with *Escherichia coli* which has a doubling time of around 30 min in normal conditions in LB-medium. Further, since many homologues recombination events in cyanobacteria only lead to single recombination, a second selection step is usually needed after conjugation. This is usually accomplished by the *sacB* gene which encodes levan sucrose (16). The protein is harmless during ordinary cell growth but when 5% sucrose is added to the medium the enzyme produces a toxic substance. By including *sacB* gene in the vectors it is possible to select for those cells in which two recombination events have occurred (i.e. loss of vector), replacing the wild-type gene with the modified one.

Another obstacle for molecular work is their thick cell walls which make DNA and mRNA purifications more difficult and physical force like sonication and the use of strong mechanical disruption are usually needed to break the cells.
Nitrogen fixation

Nitrogen is required for the synthesis of both amino acids and nucleic acids and it is essential for life on earth (73). It might therefore seem strange that even though atmospheric nitrogen is abundant most organisms, including all eukaryotes, are incapable of using it and are limited to organic (e.g. urea) or combined nitrogen (e.g. nitrate, ammonia) as their main source. Luckily for the cyanobacterial phylum, several of its members have succeeded to overcome this problem. They are capable of producing their own ammonia by capturing atmospheric nitrogen gas and converting it to combined nitrogen by the enzyme nitrogenase, a process referred to as nitrogen fixation.

The Nitrogenase

The key enzyme for nitrogen fixation in cyanobacteria is the nitrogenase which catalyzes the reduction of nitrogen gas to NH₃ (Fig 2). There are several types of nitrogenase depending on metal content (molybdenum [Mo], vanadium and iron) but the most common one in cyanobacteria is the Mo-nitrogenase (123). It is an oxygen sensitive enzyme which consists of several subunits divided into two parts (46, 123);

I The Mo-Fe protein dinitrogenase; a heterotetramer in the shape of a αβ₂ structure, responsible for the reduction of N₂.

II The Fe-protein dinitrogenase reductase; a homodimer which donates electrons from ferredoxin or flavodoxin to the dinitrogenase.

Biological fixing of nitrogen with the nitrogenase might seem ideal since 78% of the atmosphere is made up of nitrogen gas, but it is also the most expensive source of nitrogen since at least 16 ATP is needed for the production of ammonia (Mo-nitrogenase) (46). The complete reaction of the Mo-nitrogenase is as followed:

\[ \text{N}_2 + 8\text{H}^+ + 8\text{e}^- + 16\text{ATP} \rightarrow 2\text{NH}_3 + 16\text{ADP} + 16\text{Pi} + \text{H}_2 \]

Since hydrogen is produced as a bi-product the nitrogenase is also a key player of the hydrogen metabolism in diazotrophic cyanobacteria. However, the Mo-nitrogenase might be the most efficient nitrogenase for converting nitrogen gas to NH₃ but both the V-nitrogenase and the Fe-nitrogenase have a higher hydrogen gas vs. NH₃ ration then the Mo-nitrogenase (72). Since
Figure 2. Schematic illustration showing the key enzymes involved in hydrogen metabolism in cyanobacteria; uptake hydrogenase, the bidirectional hydrogenase and the nitrogenase. *; The HupC subunit is still to be identified in cyanobacteria.

the reaction favours ammonium production, together with the high production cost in form of ATP, genetic engineering will be needed to make the nitrogenase a more feasible solution to microbiological hydrogen production (71). In *Azotobacter vinelandii* for example, a single amino acid change in the Mo-nitrogenase caused the enzyme to divert ~80% of the electrons to hydrogen production (72).
Since the nitrogenase can only function in microaerobic conditions, nitrogen fixing organisms have developed either spatial or temporal solutions. The time dependent solution is found among unicellular or filamentous cyanobacteria which separate the oxygen evolving photosystem from nitrogen fixation by performing the first during light periods and the second during dark periods (28). In heterocyst forming organisms the nitrogenase is located in the microaerobic environment of the heterocyst while vegetative cells contain the oxygen evolving photosynthesis i.e. a spatial solution (116).

Nitrogen control and nitrogen fixation
Cyanobacteria can use several organic and combined nitrogen sources (e.g. urea, nitrate and ammonium) and it is not until they are fully used up that nitrogen fixation occurs (46). Independent of source they will be metabolized into ammonium which is incorporated into cyanobacteria by glutamine synthetase (GlnA) and glutamate synthase (GlsF) using 2-oxoglutarate (2OG) as carbon backbone (47). The process takes place in two steps:

Glutamine synthetase:  Glutamate + NH₃+ ATP → Glutamine + ADP + Pₐ
Glutamate synthetase: Glutamine + 2OG + 2[H⁺] → 2Glutamate

Since cyanobacteria lack 2-oxoglutarate dehydrogenase, 2OG produced in the citric acid cycle will build up inside the cell if not used for nitrogen assimilation (47). This make 2OG a perfect indicator of the C:N balance and thereby the nitrogen status within the cell.

A metabolic signal like 2OG cannot act alone though and is sensed by the P_II signal protein(GlnB) and a transcriptional factor called NtcA which both are able to bind to 2OG (46). The homotrimeric P_II protein is one of the most widespread signalling proteins in nature and is often regulating process related to nitrogen metabolism (32). In cyanobacteria it has been connected to nitrate/nitrite transport, arginine biosynthesis and NtcA-dependent gene expression (46). The P_II protein operates in two modes of signal perception; it will either bind 2OG in the presence of ATP or get a covalent modification at an exposed T-loop (i.e. phosphorylation). Both of these result in conformation change of the protein and depending on these the protein can bind to different P_II signals (32). The connection between P_II and NtcA is a protein called PipX. This protein which only exists in cyanobacteria (46), can independently interact with both P_II and NtcA. When 2OG levels are high it forms a complex with NtcA and NtcA dependent genes are activated but when 2OG levels are intermediate or low it favours binding to P_II and the same genes are downregulated (32, 46).

The transcriptional factor homodimer NtcA can also sense the 2OG levels directly but any binding between the two has yet to be demonstrated (46).
NtcA belongs to the CRP/FNR family and usually act as an activator, like in the case of many nitrogen–source assimilation genes e.g. dev (ABC transporter), nir (nitrate assimilation), urt (urea transport), amt (ammonium assimilation) and glnA, and even to itself (31, 48). It signature sequence has been reported to be the palindromic sequence GTAN$_8$TAC (alternatively TGTN$_{10}$ACA) (46). As an activator it is usually found -41.5 nt upstreams the transcriptional start site together with a -10 box while repressor sites are usually found overlapping either the -35 or the -10 box. The arrangement as an activator makes it similar to the Class II promoter activated by CRP (46, 124).

The Heterocyst

Several filamentous cyanobacteria like *N. punctiforme* and *Nostoc* PCC 7120 have the ability to form specialized cells called heterocysts, which act as the site of nitrogen fixation in these strains. The history of heterocysts is uncertain but based on similarities in the polysaccharide layer it has been suggested that they evolved from akinetes (130). They constitute a perfect micro aerobic environment for the nitrogenase, the enzyme used to convert atmospheric nitrogen to ammonium. Like the vegetative cells they are made up of a plasmamembrane, a periplasm and peptidoglycan layer and an outer membrane but additionally they also have a heterocyst envelope consisting of glycolipids and polysaccharides (80, 130) (Fig. 1). The hydrophobic glycolipid layer is heterocyst specific and made up of two different glycolipids. It is believed to act as a barrier keeping oxygen out from the oxygen sensitive nitrogenase (80) while the polysaccharide layer mainly acts as a stabilizer adding mechanical protection and supporting glycolipid layer formation (80, 130).

As a response to nitrogen depletion around 5-10% of the vegetative cells in heterocyst forming strains terminally differentiate into heterocyst within approximately 24 hours (1), depending on strain and growth conditions. A fully mature heterocyst contains no oxygen evolving photosystem II, while respiration is kept, resulting in a microaerobic environment. Their main purpose is to provide vegetative cells with fixed nitrogen in the form of amino acids like glutamine and in return they receive carbohydrates, most likely sucrose (35).

The differentiation of heterocysts starts with nitrogen depletion and the NtcA/2OG mediated response. This result in the expression of a large number of genes, many encoding regulatory proteins, in a way which was once described “a chorus of signals” (133). This is a complex system and the players are too many to fit within the scope of this thesis. I will therefore only give a brief description of the two major ones, *hetR* and *patS*.

The first gene to be expressed during differentiation is *hetR* whose protein product is the key regulator for heterocyst formation and depends on NtcA for expression (1, 133). It is a mutual dependence as NtcA in turns needs
HetR for full effect (78). Contrary to NtcA which is a global nitrogen response protein, HetR is heterocyst specific and over expression of the gene leads to heterocyst formation even under non-nitrogen fixing conditions (133). The direct function of HetR is not fully clear but it is a serine type protease with self degrading abilities which might also degrade other proteins during heterocyst development (1, 133).

HetR in turns up regulates the gene encoding PatS, a small peptide, only 17 aa long, that acts as an inhibitor of heterocyst differentiation (1, 53). After the protein is produced within the differencing cells/heterocysts it is believed to diffuse through the filament and suppress the development of the neighbouring cells (133). Only the last 5 aa of PatS (PatS-5) are needed for heterocyst suppression and PatS might therefore be degraded either in the vegetative cells or in the heterocyst (1). The relationship between HetR and PatS is complicated since this PatS-5 have an inhibitory effect on HetR and it has been suggested that the PatS to HetR ratio plays an important role in determine cell development (133).

The hydrogen metabolism

There are three key players directly involved in hydrogen metabolism in cyanobacteria, the nitrogenase, which have been mentioned previously, the uptake hydrogenase (HupSL) and the bidirectional hydrogenase (HoxEFUYH). The hydrogenases are not universally found among cyanobacteria and different strains might contain only one or both types of hydrogenases. An example of this are the model organisms *N. punctiforme* which only contain the uptake hydrogenase, *Nostoc* PCC 7120 which contains both and *Synechocystis* PCC 6803 which only contains the bidirectional hydrogenase (116) (Table 1).

Both these hydrogenases are [NiFe]-hydrogenases, a large group of ancient enzymes which can be found in many bacterial and archean organisms (10, 128). Based on the phylogenetic results from more than 80 microorganisms, they can been divided into four major groups, some with several subgroup (128);

I: Membrane bound H₂ uptake hydrogenases.
II: a. H₂ sensing hydrogenases *(regulatory hydrogenases).*
   b. Cyanobacterial uptake hydrogenases.
III: a. F₄₂₀-reducing hydrogenases
   b. Bifunctional hyperthermophilic hydrogenases.
   c. MV-reducing hydrogenases.
   d. Bidirectional NAD-linked hydrogenases.
IV: Membrane bound H₂ evolving hydrogenases.
The hydrogenases belonging to group I and II are mostly found among prokaryotes while group 3 is mostly archaean (except for group 3d which is only found in prokaryotes). Group 4 can be found in both prokaryotes and archaean organisms.

It was clear from these phylogenetic results that the two hydrogenases found in cyanobacteria belong to two different clades within the [NiFe]-hydrogenase family. The uptake hydrogenase is found within group 2 and resembles the sensing hydrogenases (also known as regulatory hydrogenases) found in, for example *Ralstonia eutropha* strain H16 while the bidirectional hydrogenase belongs to group 3d of the [NiFe]-hydrogenases. This makes the bidirectional hydrogenase more closely related to many archaean hydrogenases like the F_{420}-reducing hydrogenases in group 3a and the hyperthermophilic hydrogenases in group 3b (128).

Independent of type the [NiFe] hydrogenases always contains one large subunit (HupL/HoxH), containing the catalytic centre (i.e. [NiFe]) which is held in place to the aa backbone by CN- and CO-ligands, and Cys thiolates (Cys-S), which is responsible for hydrogen conversion, and one small subunit containing three [FeS] clusters important for electron transport (12, 33, 127, 128).

The uptake hydrogenase

The uptake hydrogenase is a membrane bound heterodimeric enzyme encoded by the genes *hupSL* in cyanobacteria (116). It is still not known how the hydrogenase interacts with the membrane since the enzyme lacks membrane spanning regions and it has been suggested that a third subunit, HupC, can attach the enzyme to the membrane. The small subunit of the membrane bound hydrogenases in group 1 contains a C-terminal transmembrane domain, missing in the cyanobacterial HupS. Instead the C-terminal is similar to the one observed in the small subunit of regulatory hydrogenases, like HoxBC in *Ralstonia eutropha* strain H16, and which is responsible for dimerization of the regulatory hydrogenase (i.e. the formation of [HoxBC]_{2}) and interaction with the histidine kinase HoxJ (15).

The function of the uptake hydrogenase is closely linked to nitrogen fixation and so far only one nitrogen fixing strain has been found which lacks an uptake hydrogenase (*Synechococcus* sp. BG 043511) (66). The hydrogenase recycles the hydrogen produced by the nitrogenase, and does thereby provide reducing power and ATP to different cellular functions (116).

The uptake hydrogenase is irreversible inactivated by oxygen and in heterocystious cyanobacteria, grown in air and without combined nitrogen, the enzyme is mainly or exclusively active in the heterocyst (116). The localisation of the uptake hydrogenase is not completely resolved though. In the symbiotic organism *N. punctiforme*, immunolocalisation studies indicate that
The presence of different structural genes encoding for hydrogenases and accessory maturation proteins, within selected cyanobacterial strains. Nitrogen fixing strains are marked with a tick (✓). BH; bidirectional hydrogenase, UH; uptake hydrogenase, N; nitrogen fixing abilities.

<table>
<thead>
<tr>
<th>Strain</th>
<th>N</th>
<th>BH</th>
<th>UH</th>
<th>Accessory Maturation genes</th>
<th>Protease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nostoc punctiforme ATCC 29133</td>
<td>✓</td>
<td>-</td>
<td>hupSL</td>
<td>hypABCDEF</td>
<td>hupW</td>
</tr>
<tr>
<td>Nostoc sp. strain PCC 7120</td>
<td>✓</td>
<td>hoxEFUYH</td>
<td>hupSL</td>
<td>hypABCDEF</td>
<td>hupW/hoxW</td>
</tr>
<tr>
<td>Synechocystis sp. strain PCC 6803</td>
<td>-</td>
<td>hoxEFUYH</td>
<td>-</td>
<td>hypABCDEF</td>
<td>hoxW</td>
</tr>
<tr>
<td>Anabaena variabilis ATCC 29413</td>
<td>✓</td>
<td>hoxEFUYH</td>
<td>hupSL</td>
<td>hypABCDEF</td>
<td>hupW/hoxW</td>
</tr>
<tr>
<td>Lyngbya majuscule CCAP 1446/4</td>
<td>✓</td>
<td>-</td>
<td>hupSL</td>
<td>hypABCDEF</td>
<td>hupW</td>
</tr>
</tbody>
</table>

The uptake hydrogenase is present in both vegetative cells and heterocysts (65, 90, 117). It is not known if the enzyme is fully active in both cell types though and activity studies of the uptake hydrogenase in Nostoc PCC 7120, a non-symbiotic strain, located the activity to be present in heterocyst only (89). However in Nostoc PCC 7120 the gene encoding for HupL is interrupted by an element which is removed by XisC under nitrogen fixing conditions.

The bidirectional hydrogenase

The bidirectional hydrogenase is encoded by the genes hoxEFUYH in cyanobacteria. It is a pentameric protein made up of a hydrogenase part; HoxYH (small respective large subunit) and a diaphorase part; HoxEFU and the molecular weight of the protein suggest a dimeric assembly of the protein complex, hox(EFUYH), (101, 116) (Fig 1). Even though the enzyme has the capability to both produce and consume hydrogen the preference is for the later (52). It is less sensitive, compared to the uptake hydrogenase and is only reversible inactivated by oxygen, hence its name, and stable to heating up to 70°C (51, 106).

The function of the bidirectional hydrogenase in cyanobacteria is still unknown. The protein is not essential and has been suggested to be involved in both fermentation; as a mediator in the release of excess reducing power during anaerobic growth (4), and photosynthesis; acting as an electron valve (5). It was before suggested to be part of respiratory complex I based on sequence similarity between subunits of the bidirectional hydrogenase and complex I (6). This seem unlikely though since several cyanobacterial strains which lack a bidirectional hydrogenase (i.e. N. punctiforme) show no difference in respiration rate compared to strains containing a bidirectional hydro-
The complex I subunits and the [NiFe]-hydrogenase are believed to have the same evolutionary background which could explain the sequence similarity (10, 44).

The bidirectional hydrogenase can be found both in heterocysts and vegetative cells in *Nostoc* PCC 7120 but with a higher activity within heterocyst during aerobic growth (52). It is usually considered to be soluble in *Nostoc* PCC 7120 (52) but the subcellular localisation within cyanobacteria in general is still not clear. In both *Anabaena variabilis* ATCC 29413 (105), *Synechocystis* sp. strain PCC 6803 (5) and *Synechococcus* sp. strain PCC 6301 (57) the hydrogenase might have a weak association to a cell membrane like the thylakoid- or cytoplasmic membrane (116). An N-terminal lipoprotein signal sequence has been discovered in the small subunit (HoxY) among some cyanobacteria like *Synechocystis* sp. strain PCC 6803 (82). Lipid modifications of the N-terminal are a unique bacterial post-translational modification which allows anchoring of proteins to a membrane.

Based on the dissimilar results from transcriptional, functional and subcellular localisation studies of the bidirectional hydrogenase within different strains, it cannot be ruled out that the bidirectional hydrogenase have different functions in different organisms (82). It can therefore not be assumed that the results from different strains can be directly transferred to another organism.

The hydrogenase maturation process

The core subunits of the hydrogenases, HupSL and HoxYH, need to go through a maturation process before they will be fully functioning. The small subunit (HupS/HoxY) possesses three [FeS] clusters and the incorporation and assembly of this cluster is still mostly unknown. It has been shown in *Rhizobium leguminosarum* bv. *viciae* and *Ralstonia eutropha* strain H16 that the gene clusters *hupGHIJ* and *hoxOQRT* respectively, are required for maturation but the exact function of the protein product of any of these genes is still to be explored (12, 68). No homologs have been found in cyanobacteria but the protein product of the genes alr0692 and alr0691, located upstream the *hyp*-genes in *Nostoc* PCC 7120, contain similar domains as found in HupH and HupG (3).

As a contrast extensive work has been done on the maturation of the large subunit (HupL/HoxH) in several organisms but in *Echerichia coli* especially (12, 20, 33). The large subunit contains an [NiFe]-active site and CN and CO ligands and the maturation depends on several proteins that catalyse the synthesis of the ligands and/or the insertion of metal ions [NiFe]. Unfortunately no studies have been performed on cyanobacteria but *in-silico* studies have revealed at least seven homologues genes of previously studied accessory proteins in *Echerichia coli*, i.e. the *hyp*-genes (*hypABCDEF*) and an
endopeptidase (hupW/hoxW) (116) (Table 1). The hyp-genes only exist in a single copy in most cyanobacterial genomes, independent on the number of hydrogenases, and are therefore believed to perform the maturation process on both the bidirectional and the uptake hydrogenase. On the contrast, the hydrogen specific proteases seem to be specific whereby hupW is believed to perform the cleavage of the uptake hydrogenase and hoxW is responsible for the bidirectional hydrogenase (116, 131) (Table 1).

Based on the studies performed on other organisms, combined with in-silico comparison studies, a putative maturation process of the large subunit in cyanobacteria can be put together (Fig. 3). The maturation process starts with the biosynthesis of CN from carbamoyl phosphate, a two-step-process. First, HypF and HypE forms a complex and HypF transfers the carbonyl to the C-terminal of HypE, an ATP dependent process during which AMP and PPI are formed, generating a thio-carboxamide (S-CONH$_2$). During the second step another ATP is required and the tio-carboxamide is dehydrated to thio-cyanate (S-CN) and ADP and Pi is released. The resulting CN ligand is donated to the HybD-HydC complex but the exact mechanism is not known. Either the CN is donated to HypD which contain an [FeS] cluster or to a Fe atom coordinated by the N-terminal cysteiny1 of HypC. The process will be repeated subsequently adding a second CN to the HypD-HypC complex. It is not known from where the CO ligand originate but studies in *Allochromatium vinosum* indicate that they come from a different source then CN. CO is most likely incorporated after the CN addition to the Fe. When the iron active site (Fe(CN$_2$)CO) is completed it is transferred to the large subunit and a HypC-large subunit complex is formed. Nickel is then incorporated by the HypA-HypB complex. HybA is a nickel binding zinc metalloenzyme while HybB is a GFP hydrolysis protein. Since nickel can be incorporated without these proteins, when added at high concentrations in the medium, it has been suggested that the purpose of these proteins is mainly to improve the kinetics or fidelity of nickel insertion. The incorporation of nickel will result in the release of HypC and makes the active site available for the hydrogenase specific protease, HupW or HoxH. The protease will bind to the nickel atom

**Figure 3 (opposite page).** Schematic illustration of the maturation of the [NiFe] uptake hydrogenase large subunit in cyanobacteria. A) The CN ligands of the active site are biosynthesized by the HypF-HybE complex and then transferred to the HypD-HypC complex which incorporates the iron (Fe) and the ligands (CN/CO) into the large subunit (HupL). The source of CO is still to be established. After the transfer, HypC forms a complex with the apo-protein, acting as a chaperon on the large subunit by stabilizing the open formation. B) The incorporation of nickel (Ni) by the HypA/B complex. C) After nickel insertion, HypC dissociate from the apo-protein and a proteolytic cleavage take place, carried out by the protease HupW. The cleavage results in a conformational change of the large subunit which can then form a complex with the small subunit, HupS.
Uptake Hydrogenase
and around 15-30 aa, depending on type of hydrogenase, will be cleaved of from the C-terminal of the large subunit, right after the sequence recognition motif DPCXXCXXH/R (36, 67, 76, 99). This is the last step of the maturation process, believed to result in a conformational shift of the large subunit, and the two subunits HupS and HupL can thereafter go together to form a functional hydrogenase (12, 33).

The functional and structural background to how and why the endoproteases are specific is still unknown. It has previously been shown that the [Ni]-incorporation into the active site is important for any cleavage to occur and that the [Ni] acts as a substrate recognition motif for the protease (75). The protease HupD from *Echerichia coli* has been crystallised and showed that three amino acids; Glu16, Asp62 and His93, are most likely to be involved in the metal binding (34). Few *in vitro* or *in vivo* studies have been performed in other prokaryotes to study or confirm the specificity and none in cyanobacteria (69, 99, 122).

The transcriptional regulation of hydrogenases

The uptake hydrogenase

Since the uptake hydrogenase is so closely connected to the nitrogen fixing process it is not surprising that the gene transcript is often connected to nitrogen control. There are many observations of an up-regulation of the uptake hydrogenase structural genes during nitrogen depletion and NtcA binding sites have been found upstream *hupSL* in *N. punctiforme, Lyngbya majuscule CCAP 1446/4, Gloeothecae sp. strain ATCC 27152 and Anabaena variabilis ATCC 29413 and have been confirmed by EMSA in several cases (60, 63, 83, 116). NtcA binding sites have also been found upstream the accessory hyp-genes in *N. punctiforme* and *Nostoc PCC 7120* (3, 40).

The most elaborate regulation can probably be observed in *Nostoc PCC 7120* whereby the gene *hupL*, just as *nifD*, is interrupted by a 9.5 kb element containing the gene *xisC*, a site specific recombinase. When the strain is transferred to nitrogen fixing conditions, a XisC dependent rearrangement occurs in which the element is excised and transcript of *hupSL* is established (18, 19).

Additional binding factors might also be involved and the binding site of another member of the CRP/FNR family, the transcription factor FNR, has been found upstream *hupSL* in *Anabaena variabilis* ATCC 29413 (41). At least three environmental conditions, nickel, hydrogen and oxygen, can up-regulate the transcript of *hupSL* in *N. punctiforme* but how the organism senses these changes is still to be revealed (8). In other bacteria like *Ralstonia eutropha* RH16 and *Bradyrhizobium japonicum* JH hydrogen is sensed by a signal transduction system consisting of the regulatory hydro-
genase (HoxBC) together with HoxJ, a histidine protein kinase, but no such system has so far been observed in cyanobacteria (7, 15).

The bidirectional hydrogenase

Several factors have been shown to influence the expression of structural genes of the bidirectional hydrogenase, like microaerobic/anaerobic conditions, [Ni], and transfer of cultures to darkness, in an up regulatory manner (86). Furthermore the transcript has been demonstrated to be under the control of a circadian clock in *Synechocystis* sp. strain PCC 6803 and *Synechococcus* sp. PCC 7942 (86, 100).

Contrary to the uptake hydrogenase there is no clear connection between the bidirectional hydrogenase and nitrogen control. The genes of the bidirectional hydrogenase are expressed during both nitrogen and non-nitrogen fixing conditions in diazotrophic cyanobacteria and is also present in both vegetative cells and heterocysts in *Anabaena variabilis* ATCC 29413 and *Nostoc* PCC 7120 (13, 51). But the results from *hox*-gene expression studies during nitrogen depletion in cyanobacteria are conflicting. In *Synechocystis* sp. strain PCC 6803 the transcript was shown to be up-regulated during nitrogen depletion while in *Gloeocapsa alpicola* strain Fitzgerald 1051 the transcript was unaffected and instead the enzyme activity increased (4, 87, 109).

One factor that makes comparisons difficult between strains is that the *hox*-genes are often not transcribed as one single unit but sometimes interrupted by different ORFs or even separated into different clusters, like in *Nostoc* PCC 7120 (116). Several studies have shown these clusters, and sometimes even genes within the same cluster, to have different expression levels (58, 100).

So far, tree transcription factors have been identified in connection with the bidirectional hydrogenase in cyanobacteria, LexA and two AbrB-like protein (CalA/B) (84, 85). LexA bindings sites have been found upstream both *hoxE* and *hoxU* in *Nostoc* PCC 7120 and have also been confirmed by EMSA (111). It is known for its involvement in the SOS-response in *Escherichia coli*, usually acting as a repressor, but its function in cyanobacteria is still not fully understood (37, 82, 85, 111, 124). However, several studies are pointing at alternative functions of this transcription factor in cyanobacteria i.e. as a mediator of redox response and/or carbon assimilation (25, 86).

The AbrB-like transcription factors, which today also goes under the name Cal (cyanobacterial AbrB-like protein) is widely distributed among cyanobacteria (54). Several clades have been identified in the AbrB-like protein family and every cyanobacterial strains studied so far usually contain at least two of these paralogous (54). If little is known about LexA in cyanobacteria even less is known about the function of the AbrB-like proteins. However, two members, CalA and CalB of the family has been found to influence the ex-
pression of the *hox*-genes in *Synechocystis* sp. strain PCC 6803, possibly acting as both a repressor or activator depending on type (54, 84, 86)
Aim of this Thesis

In this thesis I have focused on the regulation of hydrogen metabolism on a transcriptional level in filamentous cyanobacteria with a special focus on heterocystious strains and the accessory genes (i.e. hyp and the hydrogenase specific proteases). I have also studied the hydrogen specific proteases, the only part of the maturation process which might be specific on a protein level by in-silico studies.

I have always believed that one should consider the practical applications of the research and studies one perform. For an optimal hydrogen production not just the structural hydrogen producing enzymes needs to be improved, the electron flow also needs to be directed from growth to hydrogen production. If this is to be achieved a large amount of cellular mechanisms like photosynthesis, cell division, carbon fixation will have to be altered.

Sometimes these changes can be direct, like decreasing antenna sizes of PSI and PSII, but a large part of the cells natural regulation take place on transcriptional level by altering the amount of transcript and controlling when and where a particular gene is transcribed. If we are to control hydrogen production we also need to understand these complex regulatory systems within the cell.

In this thesis, part of the aim is therefore to also identify different factors or transcriptional mechanism that might or could have an impact on hydrogen production and can be further used in genetic engineering.

The aim of this thesis can be summarised as followed;

I To study the specificity of the hydrogenase maturation proteins.

II To examine the regulation on transcriptional level of the hydrogenase maturation process.

III To identify transcriptional factors or other regulatory proteins, that might be important for cyanobacteria and/or can be used for gene and functional regulation of genetically engineered cyanobacteria.
Part I – The specificity of hydrogenase maturation

The evolutionary history of hydrogen metabolism and hydrogenase specific proteases (Paper I)

To study the evolutionary relationship between hydrogenase specific proteases a phylogenetic study was performed using the PAUP and MrBayes analysis (Paper I). Several prokaryotic and archaean proteases were selected for the construction of the tree with the criteria that the substrate of their cleavage, i.e. type of hydrogenase, should be known or possible to predict and had been classified according to Vignais et al 2001(128). The criteria was set so that the resulting non-rooted phylogenetic tree of proteases and the tree presented by Vignais et al 2001(128) of the hydrogenase large and small subunit could be compared with respect to branches and subgroups.

The different groups that could be identified in the phylogenetic tree of proteases all had an equivalent group among the hydrogenases (see Vignais et al 2001 (128)) i.e. all the proteases that cleaved the same group or type of hydrogenase were clustered together, as can be seen in the simplified tree in Fig. 4a and in full in paper I (Paper I; Fig. 1). A comprehensive list of identified groups of hydrogenase specific proteases is presented here and the groups are named after the type of hydrogenase which they cleave:

1. Proteases that cleave members of group 1; membrane bound uptake hydrogenase.
2. Proteases that cleave members of group 2; cyanobacterial uptake hydrogenase (i.e. HupSL).
3. a. Proteases that cleave members of group 3a; $F_{420}$ reducing hydrogenases.
3. d. Proteases that cleave members of group 3d; NAD/NADP-reducing hydrogenases (i.e. HoxEFUYH).
4. Proteases that cleave members of group 4; Membrane bound H$_2$ evolving hydrogenases.

The similarities between the hydrogenase specific proteases’ and the hydrogenases’ phylogenetic trees suggest they have co-evolved through history. A similar co-evolution between the large and the small subunit of the hydro-
genases have already been proposed based on a comparable similarity (128). Considering that the hydrogenases have evolved since ancient time, perhaps even pre-LUCA (10, 128), the suggested co-evolution could be of very ancient origin.

However, the hydrogenase specific proteases of group 3d (e.g. Hox-cleaving proteases), stand out. When comparing the results in Paper I with the phylogenetic tree of the hydrogenases (128, 129) this group of hydrogenase specific proteases would have been expected to appear as a subgroup within group 3a (Fig 4b). Instead they are placed as a separate group within the tree. These results were surprising and unexpected and indicated the event of a horizontal gene transfer (HGT). HGT is today seen as a major force in evolution and has occurred numerous times between archaea and bacteria (23, 24, 59) and there are also numerous examples of HGT within cyanobacteria (93). By comparing the result from the phylogenetic tree of hydrogenase specific proteases with the tree of life (45) and results from genomic timescales of prokaryotic evolution (9, 110) and considering the proposition that methanogenesis was one of the first metabolic pathways to develop (9) the following theory were put forward in Paper I. Around 4-3 billion years ago the archean phylum started to evolve and diversify, many of them being methanogenic (9). As they evolved the hydrogenase evolved with them and group 3 of the hydrogenases started to diversify into groups e.g. 3a and 3b (both previously connected to methanogenesis and archaean organisms by Vignais et al 2004) (129). At some point between 3-3.5 billion years ago, possibly before the diversification of prokaryotes at around 3-2.5 billion years ago, a hydrogenase from this group was transferred to the prokaryotes by HGT, perhaps bringing a hydrogenase specific protease with it. Within these prokaryotes, the hydrogenase evolved to be what we today know as group 3d, bidirectional hydrogenases. Depending if a protease was transferred or not with the hydrogenase the phylogenetic tree presented in this thesis can be rooted in two different ways as seen in Fig. 4b. Either an archaean protease is the origin of all prokaryotic hydrogenase specific proteases i.e. group 1 and 2 (Fig 4b:2) or the transferred group 3d hydrogenase would have incorporated an already existing prokaryotic proteases into its maturation process (Fig 4b:1).

An extended phylogenetic study was also performed containing more hydrogenase specific proteases, some also cleaving hydrogenases of 3b type. This tree was less robust with much lower clade credibility values given for the branches, but showed proteases of 3b-type scattered in-between the group 1,2 and 3d proteases and the group 3a proteases, suggesting that an archaean protease (Fig 4b:2) is indeed the origin of the group 1 and 2 prokaryotic hydrogenase specific proteases.
Large-scale molecular genetic analysis of the DNA sequence (like studying gene order or GC content) could give a clearer picture, however, after 3 billion years of evolution, mechanisms like amelioration (e.g. the transferred gene will over time acquire the molecular characteristic of the host genome) will most likely have erased most of the evidence of a HGT event.

**The specificity of hydrogen specific proteases (Paper I, II)**

For many years it has been speculated if the hydrogenase specific proteases in cyanobacteria have a substrate preference with respect to hydrogenase type i.e. HupW would cleave the large subunit of the uptake hydrogenase
while HoxW would cleave the large subunit of the bidirectional hydrogenase (131). This type of specificity among the hydrogenase specific proteases has previously been observed in other prokaryotes like *Escherichia coli*. Another example is *Alcaligenes eutrophus* in which the HoxW protease was shown to be explicit for the cleavage of the HoxH subunit of the soluble hydrogenase in the same strain (122). In cyanobacterial though no in-vivo or in-vitro studies have been performed and the knowledge is merely based on in-silico studies (131).

To verify or dismiss the theory of protease specificity concerning HupL/HoxH cleavage in cyanobacteria a *hupW* mutant was created in *Nostoc* PCC 7120 (Paper II) (Fig. 5). The mutant strain was constructed by triparental conjugation and homologous recombination, using the vector pRL::HWNm which contained the gene *hupW* disrupted with a neomycin (Nm) cassette. Resulting clones were checked by PCR to verify full segregation of the mutant and the disruption of the *hupW* gene.

In filamentous heterocyst forming cyanobacteria hydrogen is primarily produced by the nitrogenase and subsequently recycled by the uptake hydrogenase. Any inactivation of the uptake hydrogenase, independently if it is directly or indirectly through the accessory proteins, will result in the release of hydrogen. Consequently hydrogen production measurement is an ideal method to check for uptake hydrogenase activity, as have been previously done on the *hupL* mutant in *N. punctiforme* (NHM5) (64). The mutant strain *hupW* was shown to produce hydrogen at a rate of 46 µmol H₂ mg Chl a⁻¹ h⁻¹. These results can be compared with the production observed in a *hupL* mutant in the same strain which were in the range of 45-50 nmol mg Chl a⁻¹ h⁻¹ (70). Since the production rate is the same it is clear that the uptake hydrogenase has been inactivated in the *hupW* mutant. Our results therefore suggest that the protein product of *hupW* has indeed a preference for cleavage of the uptake hydrogenase large subunit.

At this stage it cannot be excluded that the disruption of the *hupW* gene has had an effect only on the large subunit from the uptake hydrogenase (HupL) or if both hydrogenases (HupL and HoxH) have been affected, but it can be excluded that the protease will act solely on HoxH. It has previously been shown, in the same strain, that the bidirectional hydrogenase cannot replace the function of the uptake hydrogenase (70). In this study it was shown that a *hoxH* mutant in *Nostoc* PCC 7120
**hupW**

*(Nostoc PCC 7120)*

Figure 5. Illustration showing the insertion and inactivation of the gene alr1423 (*hupW*) in *Nostoc* PCC 7120 with a neomycin (Nm) cassette.

will produce even less hydrogen than the wild type (15-33%) (70) and hence the hydrogen produced by the *hupW* mutant in our study cannot be explained by an inactivation of the bidirectional hydrogenase but is most likely the result of an inactivated uptake hydrogenase. The same study also showed that it is impossible to separate a HupL− mutant from a double mutant (HupL−/HoxH+) by hydrogen measurement only and more studies are therefore needed to verify the specificity of this *hupW* mutant.

Even so, the result from these studies and previous studies in other organisms raise some questions (98, 119). How is the explicit cleavage by the hydrogenase specific protease possible and how does it distinguish between the two different substrates within the cell? The results from the crystallized hydrogenase specific proteases HydD and HycI from *Escherichia coli* have given vital clues about the substrate binding but none of them could explain the substrate recognition (34, 132). However, they did show that the amino acid residues Glu16, Asp62 and His93 in HybD in *Escherichia coli* are most likely involved in the metal binding of the nickel in the active site of the hydrogenase large subunit (the residues are represented by Asp18, Asp62 and His88 in HoxW of *Nostoc* PCC 7120) (34).

To further study the mechanism behind the specificity of hydrogenase specific proteases in cyanobacteria an *in-silico* study of HupW and HoxW from *Nostoc* PCC 7120 was performed (Paper I). By aligning the aa sequences from several cyanobacterial and other prokaryotic hydrogenase specific proteases and dividing them according to the previously defined phylogenetic group (Paper I), a group specific difference on aa level was identified. In HupW (*Nostoc* PCC 7120), a member of the phylogenetic group 2 proteases, the amino acid (aa) residues 41–44 consists of the sequence [DCGT] while among the HoxW proteases (group 3d) it is replaced by the sequence [H(Q/I)L] (aa 42–44 in HoxW of *Nostoc* PCC 7120). The sequence found among group 2 proteases (HupW) were fairly semi conserved among hydrogenase specific proteases i.e. group 1, 3a and 4 (Paper I, (34)) while
the sequence found in HoxH were unique for the proteases of Hox type, i.e. group 3d (Paper I). This is interesting since the same region had previously been suggested by others to be involved in the c-terminal cleavage of the large subunit of hydrogenases (132). The mere difference observed in this region (paper I), suggests otherwise and the HoxW specific sequence, named the HOXBOX, could instead explain part of the substrate specificity observed among hydrogenase specific proteases. Amino acid replacement, whereby Asp38 in HycI in *Escherichia coli* (equal to Asp41 in HupW in *Nostoc PCC 7120*) was changed to an asparagine also showed no effect on the cleavage process (119), which of course does not rule out that other residues within this region might be important. The same study also suggests that Asp62 is a better candidate for the proteolytic cleavage (119).

To study if and how this sequence (i.e. HOXBOX) could influence the activity or substrate binding of the protease, protein-docking experiment using the program BiGGER V2 were performed with the proteases and their respective large subunit (Paper II). Both the hydrogenase specific proteases HybD (PDB code: 1CFZ) and HycI (PDB code: 2I8L) in *Escherichia coli* have previously been crystallised and the structure resolved and could serve as template when making a 3D model of HupW and HoxW in *Nostoc PCC 7120* (34, 132)(Paper I). However, since HycI shows the protease in an open conformation (i.e. not binding the substrate) while HybD shows the protease interacting with cadmium (a substitute for [Ni] during crystallisation), the later was chosen for the modelling experiments. As template for the large subunits (HupL/HupH) the crystallized HydB (UBJ:L) from *Desulfovibrio vulgaris* Miyazaki F was chosen.

In comparison the docking experiments were also performed on HybD (1CFZ) in *Escherichia coli* and HynC in *Desulfovibrio vulgaris* Miyazaki F with their respective assigned large subunits, HybC and HybB (UBJ:L). The results from these protein docking studies showed that among the proteases that did not contain a HOXBOX (i.e. HupW from Nostoc PCC 7120, HybD from *Escherichia coli* and HynC from *Desulfovibrio vulgaris* Miyazaki F) a close interaction between the protease and their respective hydrogenase large subunit could be observed. However the result from the docking experiments with HoxW showed the protease leaning away from the large subunit, resulting in that only a small part of the protease were involved in substrate-protease interaction (Paper I: Fig. 7).

The result from the protein-docking experiments suggest that the HOXBOX could have an effect on the substrate binding whereby the protease will have a different organization toward the large subunit depending on the presence or absence of a HOXBOX.

These results matches the results from another *in-silico* study performed in the same article (Paper I) whereby conserved amino acids residues on the surface of closely related homologs of the different proteases HybD, HupW and HoxW were observed and marked on the 3D structure of the different
proteases (Paper I: Fig. 7). These studies showed that one part in particular
of the surface area of the HybD/HupW type proteases were conserved espe-
cially around alpha helix 1, beta sheet 2 and alpha helix 4 (34) of the prote-
ases. This is the same area that in the protein-docking studies showed a close
interaction with respective large subunit of the hydrogenase. However,
among the HoxH homologues the same degree of conserved residues could
not be observed, again matching the results from the protein-docking studies
which suggested a lower degree of interaction between substrate and prote-
ase.

It has previously been suggested that a conformational recognition takes
place between the protease and the large subunit and the results from the
protein-docking experiments support this theory (120, 121). Further close
interaction could through the years have enhanced substrate specificity i.e.
as the hydrogenase large subunit evolved the surface structure changes and the
protease had to adopt for proteolytic cleavage to occur. This could explain
the suggested co-evolution observed between the hydrogenase and the hy-
drogenase specific protease in the phylogenetic studies.

If the proteases are specific with respect to the large subunit of the hydro-
genase, which the result in this thesis suggest (Paper I, II), then the photo-
lytic cleavage would be the only directly known part of the maturation proc-
ess which is specific in cyanobacteria. Hypothetically the protease could
then have an effect on the hydrogenase activity on a protein level, by turning
the hydrogenases “on” by the proteolytic cleavage of the large subunit at a
specific moment. There are several other [NiFe]-hydrogenases, i.e. the hy-
drogen sensing hydrogenases of group 2 and some Achaean hydrogenases
(11, 128) which are not cleaved by a protease suggesting that the proteolytic
cleavage is not essential for the maturation of a functional hydrogenase. One
might therefore ask why this elaborated and specific system of hydrogenase
C-terminal cleavage has ever evolved, a question that is still waiting to be
answered.
Part II – Nitrogen control and hydrogen metabolism

As mentioned before the hydrogenase specific protease is the only step of the maturation process of hydrogenases which might be specific in cyanobacteria. Theoretically they could therefore have a large impact on protein level on the amount of active hydrogenase in the cell and could also be a useful tool for controlled hydrogenase activation in a genetically engineered cyanobacterium. But as a famous 90’s cartoon novel once said “Who watches the watchmen?” (77).

Transcriptional regulation of the uptake hydrogenase and its specific protease (Paper I, III, V)

At the beginning of my PhD not much was known about the gene expression and regulation of the hydrogenase specific proteases in cyanobacteria. A transcriptional analysis was therefore performed, primary on the \( hupW \) operon but also to some degree on the large subunit of the uptake hydrogenase in both strains.

As a first step to understanding the transcriptional regulation of \( hupW \), the transcriptional start point (tsp) was established by 5'RACE and any putative co-transcription with upstream genes was studied by RT-PCR (Paper I). Northern blots were also performed, on RNA extracted from both nitrogen and non-nitrogen fixing conditions, to detect any co-transcription with other genes as well as revealing any nitrogen related response (Paper I). The results were unexpected since they showed that \( hupW \) in both \textit{N. punctiforme} and \textit{Nostoc} PCC 7120 was co-transcribed with an upstream gene of unknown function (Fig 6).

In \textit{N. punctiforme} the gene \( hupW \) were co-transcribed with the upstream gene Npun_F0373. The protein product of this gene is only 140 aa long, contains no known domains or motifs except for a transmembrane region between amino acids 84–105. The function is unknown but homologous genes could be found in several heterocyst forming cyanobacteria like \textit{Nostoc} PCC 7120, \textit{Anabaena variabilis} ATCC 29413, \textit{Nostoc} sp. strain PCC 7422 and \textit{Nodularia spumigena} CC9414, sometimes in connection to the structural genes of the uptake hydrogenase. These homologs also have a highly conserved promoter region containing a putative NtcA binding sites, \( \sigma^{70-10} \) box, putative Shine-Dalgarno sequence and even suggests a putative tsp for some of them, indicating that these genes are under the same regulatory control in all strains.

The result from the Northern blot studies revealed a single transcript, expressed during nitrogen fixing conditions only, of about 1300 nt. The size correlates with the co-transcription of \( hupW \) and Npun_F0373 in \textit{N. puncti-
Figure 6. A schematic figure of *hupW* in *N. punctiforme* (A) and *hupW* and *hoxW* in *Nostoc* PCC 7120 (B) together with their respective promoter region (as described in paper I) and surrounding or co-transcribed genes. Marked on the picture are the transcriptional start points (arrows) and the NtcA or LexA binding sites as well as -35 and -10 boxes (when present). The sequence of respective binding site can be seen on the right.

The tsp was identified 607 bp upstream of *hupW*, together with the already mentioned $\sigma^{70}$-10 box sequence and a putative NtcA binding site (GTAN$_8$CAC), located 40 bp upstream from the tsp.
In *Nostoc* PCC 7120 the gene encoding the hydrogenase specific protease were transcribed together with the upstream gene alr1422. Homologs of this gene can be found only in a few strains e.g. *Anabaena variabilis* ATCC 29413 and *Trichodesmium erythraeum* IMS101 and the encoded protein contains no known domains or motif. Three tsp were found in total, one 234 bp upstream of *hupW* and two inside the upstream gene alr1422, 4 bp and 14 bp downstream of the putative translation start site. It is possible that alr1422 have a second translation start site since a new site was found within the same frame 115 bp downstream from the previously suggested start site. The possibility formation of a single transcript was also confirmed by the Northern blot, which revealed one long transcript (~1600 bp), only expressed during nitrogen fixing conditions, and one short constitutively expressed transcript (~500 bp). Again putative NtcA binding sites were identified, two upstream alr1422 (GTAN₈AAC and GTN₁₀AC) and two imperfect sites upstream *hupW* (TGAN₈CAC and GTAN₁₂TAC).

All the identified putative NtcA binding sites upstream *hupW* in both strains were later confirmed by non radioactive EMSA, creating a shift in all three promoter regions (PNpun_F0373, Palr1422 and *HupW-Nostoc* PCC7120) when co-incubated with His-tagged purified NtcA (Paper III). As mentioned in the introduction, putative NtcA binding sites have previously been identified upstream *hupSL* in several cyanobacterial strains (116) and the direct effect of NtcA has also been confirmed in *Anabaena variabilis* ATCC 29413 (126). Considering the strong connection between the uptake hydrogenase and nitrogen fixation it is therefore not surprising if the hydrogenase specific proteases are also under direct or indirect nitrogen control, and this is also supported by the findings of the NtcA binding sites. This would indicate that the uptake hydrogenase and the hydrogenase specific protease are under the same or similar regulatory control on transcriptional level. Therefore the effect of nitrogen depletion on the transcriptional expression of the genes of the hydrogenase specific protease was further studied in a time course study by semi-quantitative RT-PCR (Paper III) to see if the onset and increase of the gene expression would be the same. The cells were shifted from non-nitrogen to nitrogen fixing conditions and samples were taken at time point 0, 12, 24 and 48 h after nitrogen depletion. The results from the transcriptional analyses showed a clear synchronized increase of transcripts during nitrogen depletion of the genes encoding the large subunit of the uptake hydrogenase and the *hupW* genes in both strains supporting the theory of a similar regulatory control (Paper III). In *Nostoc* PCC 7120 it is also apparent that the transcriptional pattern of *hupW* mirror *hupL* and not *hoxH* which again supports the idea that these proteases are specific with respect to their substrate, the large subunit of the hydrogenase.

The cellular localisation of the uptake hydrogenase in *N. punctiforme* has been under some debate. In *Nostoc* PCC 7120, the uptake hydrogenase has
been shown to be localised to heterocysts, both by activity measurement and immunolocalisation studies of HupL (52, 104). However, similar immunolocalisation studies performed in N. punctiforme revealed that the HupL protein in this strain is present in both vegetative cells and heterocysts (65, 104). These results were puzzling especially since the promoter of the uptake hydrogenase operon (hupSL) has been shown to only be active in heterocysts (50). There is no known transport system for the uptake hydrogenase in cyanobacteria and the N-terminal TAT-signal found in the small subunit among many bacterial hydrogenases are missing in the cyanobacterial uptake hydrogenase (116). So how could this be?

To study any possible regulation or post-maturation transport of the uptake hydrogenase, which could explain these results, a HupS-GFP reporter construct was constructed to investigate the cellular and subcellular localisation of the uptake hydrogenase in N. punctiforme (Paper IV). The study revealed that the small subunit of the uptake hydrogenase, HupS, in Nostoc punctiforme is solely localised to the heterocysts and no sign of cellular transport to vegetative cells could be observed. In this study the intracellular fluorescence from the HupS-GFP fusion protein was observed during heterocyst differentiation, caused by nitrogen depletion, for up to 7 days. The weak GFP fluorescence could be observed from proheterocysts around 24 h after nitrogen depletion, and later, approximately at 30-34 h, several small and a few larger clusters were observed. An interesting observation was that the large clusters were frequently observed close to the polar regions of the heterocysts. In the related cyanobacterial strains Anabaena cylindrica and Nostoc PCC 7120 this particular area contains “honeycomb” structures, a membrane system suggested to be devoted especially to respiration (125). Since one targets for the electrons released during H₂ oxidation is the respiratory electron transport chain (51), the potential localisation of the oxygen-sensitive uptake hydrogenase to the honeycomb membrane structures could be of relevance for the function of this enzyme.

Transcriptional regulation of the bidirectional hydrogenase and its specific protease (Paper I, III)

The hoxW gene in Nostoc PCC 7120 was shown by Northern blot studies to have multiple transcripts (Paper I). This could be the result of both multiple tsp as well as multiple transcriptional stops. So far two tsp have been found, 44 bp and 70 bp upstream of hoxW. When analysing the promoter region a sequence consisting of four repeats were observed, situated 204–231 bp upstream of hoxW. The repeats were suggested to form a hairpin loop and could potentially be a transcriptional stop for the gene upstream hoxW (e.g. all0771; 4-hydroxyphenylpyruvate dioxygenase). The promoter also contained a putative LexA binding site (TAGTAGTTATGTAAT) and the inter-
action of LexA with the promoter region were later confirmed by EMSA (Paper III) (Fig. 6).

The Northern blot study also showed that \textit{hoxW} is constitutively expressed in \textit{Nostoc} PCC 7120 with a slight increase of transcript level during a shift from non-nitrogen fixing to nitrogen fixing conditions. This increase of transcript was later confirmed by semi-quantitative RT-PCR time course study (Paper III) whereby the transcript level between 0-48 h after during nitrogen depletion (i.e. inducing nitrogen fixing conditions) were studied, as described for \textit{hupW} in the previous section. The study brought some interesting results since they also showed a clear increase of the \textit{hupH} transcript levels, the large subunit of bidirectional hydrogenase. Further it could be shown that the increase of transcript was located within the vegetative cells and not the heterocysts. This was unexpected since it has previously been shown that the bidirectional hydrogenase is equally active in heterocysts and vegetative cells in \textit{Nostoc} PCC 7120 (52). It is known from before, that vegetative cells have a higher level of oxygen compared with heterocysts (30). A plausible explanation is therefore that the oxygen sensitive bidirectional hydrogenase has a higher turn-over rate in these cells and that a higher transcript is needed for maintaining enzyme activity.

In addition, during nitrogen depletion and the differentiation of heterocysts the vegetative cells need to be metabolically reprogrammed to be able to produce and transport reducing carbohydrates from the vegetative cell to the heterocyst. The bidirectional hydrogenase has previously been suggested to act as an electron valve in cyanobacteria and it could therefore be part of keeping the redox homeostasis within the vegetative cell during nitrogen depletion (103). It is possible that the bidirectional hydrogenase has a special function during a nitrogen shift from non-nitrogen fixing to nitrogen fixing conditions in vegetative cells in these strains.

The complexity of the NtcA family (Paper V)
An interesting observation in the transcriptional analysis of the uptake hydrogenase and its respective protease was that even though the level of transcript of the large subunit of the uptake hydrogenase and the hydrogenase specific protease were synchronized in time a similar expression pattern could not be observed for \textit{nifD}, one of the nitrogenase subunits (Paper III). This could be the result of several parallel signal pathways involved in nitrogen response in heterocyst forming cyanobacteria resulting in a different onset of the \textit{nifD} transcript compared with \textit{hupL/hupW}. Considering that NtcA is a key regulator during nitrogen response and heterocyst differentiation and have been connected to both the uptake hydrogenase and the \textit{hupW} gene expression it might be an important and useful transcriptional factor in future genetic engineering projects concerning hydrogen production.
An *in-silico* study was therefore performed on the NtcA family to give more detailed knowledge about the different NtcA family members and their possible role regarding nitrogen response in cyanobacteria (Paper V).

In this study an un-rooted phylogenetic tree was constructed with MrBayes analysis with paralogs to NtcA from different cyanobacterial strains i.e. *N. punctiforme*, *Nostoc* PCC 7120, *Anabaena variabilis* ATCC 29413, *Gloeobacter violaceus* PCC 7421, 'Nostoc azollae' 0708, *Synechocystis* sp. strain PCC 6803 and *Trichodesmium erythraeum* IMS101 (Fig 7b). The tree revealed several new uncharacterised members of the NtcA family which were clustered into different groups with different characteristic, as seen in the list below (Fig 7a):

1. Traditional NtcA, as defined by Herrero et al 2008 (46). Containing one cNMP binding domain and one (helix-turn-helix) HTH motif
2. Putative transcriptional factors containing one PAS domain, one cNMP binding domain and one HTH motif.
3. DevH transcriptional factors, containing no domains but one HTH motif.
4. A putative NtcA family related transcriptional factors containing one cNMP binding domain and one HTH motif.

Since group X is only a putative NtcA family member at this stage, it was not given a definite number but is merely called X in this thesis.

As mentioned earlier NtcA is a member of the CRP/FNR family and is today a well characterized and studied transcriptional factor in cyanobacteria (46). Group 1 was therefore used as a reference point when examining the other groups within the NtcA family.

Except for NtcA, group 3 (containing DevH and homologs) are the only transcriptional factors within this family which have previously been studied (43). Studies performed on DevH in *Nostoc* PCC 7120 suggest that DevH acts downstream of NtcA, at an early stage of heterocyst development, but not much is known about its function. A devH- strain develop heterocyst but perform no nitrogen fixation during aerobic growth and no nitrogenase genes (*nifHDK*) were expressed but so far DevH has only been shown to be important for the production of the heterocyst glycolipid layer (91). The same study suggested that the control of *nifHDK* might be indirect as the lack of glycolipids could result in elevated oxygen levels within the heterocyst (91). If the cells are grown under anoxic conditions the nitrogenase activity is restored (91).

The members of group 2 are interesting since they were the only group within the phylogenetic tree for which a putative function could be suggested. The members of this group all contain a PAS domain on the N-terminal side of the protein. PAS domains are usually involved in sensing
oxygen, redox potential and light within the cell (118) and studies in *Nostoc* PCC 7120 and *N. punctiforme* have shown that PAS domains are extraordinarily well represented in these strains compared to other bacteria (79, 81). Considering the oxygen sensitivity of both the nitrogenase and the uptake hydrogenase it is important for the cell to adapt to oxygen level within the cell and all known oxygen sensors so far that control nitrogen fixing, contains a PAS domain. It is therefore possible that the members of group 2 act as direct or indirect sensors of oxygen.

Additionally it should be mentioned that members of both group 2 and 3 might be essential for filamentous cyanobacteria. In *Raphidiopsis brookii* D9, a filamentous non-heterocyst forming cyanobacteria and *Cylindrosper-
mopsis raciborskii CS-505, a filamentous heterocyst forming cyanobacteria, genome searches revealed that homologous genes to both group 2 and 3 could be identified. These two cyanobacterial strains are fully sequenced and contain the smallest cyanobacterial genomes known so far and it has been suggested that these strains should be compared with other nitrogen fixing and heterocyst forming cyanobacteria in the search of key characteristics for both nitrogen fixing and heterocyst formation (114).

A key feature of the NtcA family is there homology within the HTH motif. It has previously been suggested, based on this homology that DevH and NtcA might interact with the same consensus sequence (43). The findings in Paper IV suggest that also other members of the NtcA family, like group 2, might have this ability.

This idea is strengthened by the fact that the first two amino acids (aa) (Arg193, Val194 in NtcA of Nostoc PCC 7120) in the DNA binding region of the HTH motif are conserved among members of the NtcA family within filamentous cyanobacteria. The first two aa in the DNA binding region have previously been shown to interact with the DNA helix in the CRP protein from Escherichia coli (102). Further it has recently been shown in DNR, another member of the CRP/FNR family, a correlation between the sequence of these two aa and the reported DNA binding consensus sequence of different DNR family members (94).

The differences observed on the N-terminal side of the NtcA family members, e.g. the presence or absence of cNMP binding domains and PAS domains, together with the preserved HTH motif suggest that the different members could act on different stimuli even though they bind very similar or the same consensus sequence. This would give the cell a possibility to fine tune the cellular metabolism and act on very different conditions while having only one binding site in its promoter region.

As pointed out in paper IV there are also several promoter regions of different genes connected to nitrogen or hydrogen metabolism in cyanobacteria (i.e. N. punctiforme) that contains non-perfect NtcA binding sites with respect to the consensus sequence GTAN₈TAC. These differences could reflect the binding affinity within one single group of the NtcA family to the DNA sequence. In Paper I for example a competitive EMSA was performed with purified NtcA and the promoter regions of both alr1422 and hupW from Nostoc PCC 7120. As was expected NtcA had a preference for the promoter region of alr1422, which contains a more conserved NtcA binding site. But it could also be that the differences reflect the preferred consensus sequence of each individual group within the NtcA family (i.e. 1, 2, 3 and X).

Since some of these non-perfect binding sites can also be observed in the promoter region of representatives of each individual NtcA group in N. punctiforme, it cannot be ruled out that the members of different groups within the NtcA family might interact with each other’s promoter region. If that is true the NtcA depended nitrogen control within cyanobacteria might
take the shape of a complicated spider web of interactions, even more complicated then previously believed. If the different members of the NtcA family can bind to the same consensus sequence, perhaps even with different affinity to different versions of the sequence, and further as a response to different stimuli it might turn into an interesting challenge to solve the puzzle of nitrogen related transcriptional regulation.

CalA, a repressor of the accessory genes, important for hydrogenase maturation? (Paper III, VI)

Independent on the number of hydrogenases found in a cyanobacterial strain they will have only one set of hyp-genes. These genes encode for accessory proteins important for the maturation of the large subunit. Any regulation of these genes could therefore have an effect on the amount of active hydrogenase on both types of hydrogenases in the same strain i.e. Nostoc PCC 7120. A DNA affinity assay was therefore performed to establish any possible interactions with the promoter region of the hyp-genes and previously unknown or known transcriptional factors in Nostoc PCC 7120. The result revealed the interaction of a new putative transcription factor, CalA, with the promoter region of hypC, encoding one of the accessory proteins important for hydrogenase maturation (Paper VI). This interaction was later confirmed in the same paper by using purified His-tagged CalA in EMSA experiments.

CalA is similar to a protein family of transcriptional factors called AbrB, which have been extensively studied in Bacillus subtilis. AbrB have been directly or indirectly connected to over 60 genes in different physiological and metabolic pathways. In cyanobacteria CalA has been found to act as an activator of the hox operon in Synechocystis sp. strain PCC 6803 (84, 86) and it was also recently discovered to interact with the promoter region of hupSL in both N. punctiforme and Nostoc PCC 7120 (49). Further it was shown by the same method (DNA affinity assay) in Paper III that CalA also interacts with the promoter region of the hydrogenase specific proteases in Nostoc PCC 7120 and N. punctiforme Additional to hydrogen metabolism CalA has also been associated with the regulation of genes from very diverse metabolic pathways like photosynthesis (62), nitrogen fixation (CalB) (54), oxidative stress (2) and production of the toxin cylindrospermopsin (108). This indicates that CalA act on a global level in the cell.

As mentioned in the introduction the genes encoding cyanobacterial AbrB like transcriptional factors usually exist in two copies in the genome and are classified into two distinct clades, CalA and CalB (arl0946 and all2080 in Nostoc PCC 7120). CalB have been shown to also interact with the promoter region of hupSL in both N. punctiforme and Nostoc PCC 7120 (49) and was also shown to interact with the hupW promoter regions in the same strains in this thesis (Paper III).
These results presented in Paper VI and III suggest a direct function of CalA as a regulator of hydrogen metabolism and its function was therefore further studied in Paper VI. By using a strain containing an NO\textsubscript{3}\textsuperscript{-} induced overexpression vector (pNir) carrying the gene *calA*, the expression of both *hypC*, *hoxE* and the gene encoding CalA itself were studied. It was clearly seen in these experiments that *hypC* expression was downregulated in the CalA-overexpression strain suggesting a negative regulatory role for CalA under these conditions. Unfortunately the expression levels of *hoxE* was impossible to study since the shift from NH\textsubscript{4}\textsuperscript{+} to NO\textsubscript{3}\textsuperscript{-} (leading to induction of *calA*) resulted in such a strong increase of *hoxE* in both wild type and the overexpression strain. Activity measurement of HoxE showed a decrease in activity which indicates a negative effect of CalA. However these results could also be indirectly caused by the down regulation of the gene encoding the maturation protein *hypC*.

The transcriptional study of *calA* itself was established through the down-stream gene alr0947, which were shown to be co-transcribed with *calA* in the same study by RT-PCR (Paper VI). The results from these studies indicates that CalA act as a repressor on its own promoter, which is not uncommon among transcriptional factors (124). The expression level of *calA* also seem to be generally high in the cell and the Northern blot results also showed a very strong band of approximately 2100 nt, which correspond to the co-transcription of both genes (Paper VI). The band was stronger in nitrogen fixing cultures which is in agreement with two quantitative proteomic iTRAQ studies in *N. punctiforme* and *Nostoc* PCC 7120 (88, 112).

The observation of CalA as repressor of the *hypC*-operon (Paper VI) suggests that it might also act as a repressor of *hupW*. A recent study has also shown that *calA* is down regulated in heterocyst cells relative to vegetative cells in a nitrogen fixing culture (42). A putative role for CalA might therefore be to act as a repressor of the hydrogenase maturation genes *hypC* and *hupW* in vegetative cells in particular.

Since not much is known of CalA in cyanobacteria an *in-silico* study was also performed whereby a 3D model was created of the DNA binding region of CalA (Paper VI). Contrary to AbrB, CalA and its homologs have the DNA binding region located in the C-terminal part of the protein ((113), Paper VI) and only one of four arginines, which are believed to be important for DNA binding in AbrB, might have a counterpart in CalA. It should be mentioned though that when comparing the aa sequence with other members of the AbrB family it was apparent that two of the substitutes (Gln15 and Lys24) are actually common substitute in several AbrB proteins (Paper VI). The *in-silico* studies also showed that CalA, just as AbrB, act as dimer and that the loop-hinge regions, which enables a flexible binding to the DNA and is typical for AbrB proteins, seem to be conserved. No consensus sequence has yet been established for AbrB proteins and instead it have been suggested that the transcription factor recognise the 3D shape of the DNA spiral.
The loop-hinge region has been suggested to be part of the explanation imposing a flexible DNA binding among AbrB protein (26). The observation of a loop-hinge region in CalA therefore suggests that CalA might interact with the DNA in a similar manner.
Summary and future outlook

It has always been in my interest to try to connect the result of my research with a practical usage. I also believe that many times the most interesting things are found where we least expect them. In this thesis I have studied both the maturation process of hydrogenases in cyanobacteria and the transcriptional regulation of the hyp-genes, the genes encoding the hydrogenase specific proteases (i.e. hupW and hoxW) and hydrogenases in two heterocyst forming filamentous cyanobacteria e.g. Nostoc sp. Strain PCC 7120 and Nostoc punctiforme ATCC 29133. My findings support the theory of specificity among hydrogenases specific proteases with respect to their substrate, the large subunit of the hydrogenase. A hupW mutant in Nostoc PCC 7120 was shown to produce hydrogen as an effect of an inactive uptake hydrogenase. Cyanobacterial protease specificity could be explained by my observation of group specific differences on amino acids level (e.g. the HOX-BOX) among these proteases and my evolutionary findings which suggest that the protease has been part of the maturation process since ancient time and most likely co-evolved with the hydrogenase. These finding can be used when planning future genetic engineering projects involving the introduction of a foreign [NiFe]-hydrogenases into cyanobacterial strains. If each type of protease can only cleave one type of hydrogenase large subunit it cannot be assumed that the original hydrogenase specific protease in cyanobacteria will be able to perform a proteolytic cleavage on the introduced hydrogenase. The phylogenetic tree can help by revealing the relationship between different hydrogenase specific proteases and thereby the likeliness of such a cleavage to occur. Further it is important when aa modifications of the hydrogenase are planned to take in mind the possible effect on the surface area since proteylotic cleavage might bind not only the nickel situated in the active site but recognize the 3D structure of the large subunit.

Further I have shown that these proteases seem to be under the same regulatory control on transcriptional level as the hydrogenases they cleave in both N. punctiforme and Nostoc PCC 7120. An elaborative system of transcriptional factors and regulatory control on gene level is emerging showing the involvement of several transcriptional factors like NtcA, CalA and CalB for hupW and LexA and CalA for hoxW. The transcriptional studies also indicate that the bidirectional hydrogenases have a new not yet identified role in vegetative cells during nitrogen depletion. If cyanobacteria are to be used as a host for either a foreign hydrogenase or a modified cyanobacterial
hydrogenase, to produce hydrogen on a commercial level we need to understand how the expression of the structural genes should be controlled for maximum hydrogen production. Further it might not be only the structural genes that need to be modified for optimal hydrogen production but also the regulatory system within the cell. The energy usage of cyanobacterial strains today are not optimized for hydrogen production but are aimed towards growth and survival. The energy will therefore need to be redirected and a large part of how genes and functions are regulated is on transcriptional level. The results from my studies show the involvement of several transcription factors in the regulation of hydrogenase maturation in *Nostoc* PCC 7120 and *N. punctiforme* and what stimulus they act on, which signal pathways they belong to and if and how they influence each other will need further studies, especially considering the findings of new uncharacterized transcription factors within the NtcA family.

Personally one question stills puzzles me though... so why do the proteases exist at all? It could be that since the maturation of the large subunit involves several steps and such a complex maturation might be costly for the cell to perform, any dysfunctional version of the hydrogenase might need to be abolished at an early stage. It could also be that the dysfunctional hydrogenases interfere with any downstream reactions within the cell. The endoproteolytic cleavage of the C-terminal could therefore be a way for the cell to verify that the large subunit is correctly assembled. It could also be that since the *hyp*-genes are important for the maturation of both types of hydrogenases, a pattern which can also be seen among bacteria which contains several other [NiFe]-hydrogenases (12, 20), they need to be expressed on such a level that they will not restrict the maturation of either hydrogenase (uptake or bidirectional). To avoid any activation of each individual hydrogenase under the wrong growth conditions the protease will therefore act as a second check point, making sure that no active hydrogenase is present at the wrong time at the wrong place. Further, if the protease might need a close interaction with the large subunit for the proteolytic cleavage to occur, which the *in-silico* studies suggest, this could have indirectly resulted in an evolutionary pressure for specific proteases for each type of hydrogenase. Over time individual cyanobacterial strains might have lost one (or both) hydrogenase(s) but by this time the system with highly specific proteases was already established.

So what will come? Will there in the end be commercial hydrogen production by cyanobacteria? Well, I cannot guarantee you that the future spell “cyanobacteria” but I firmly believe that it is a good candidate and have great potentials. There are today several genetically modified strains (not just *hupW*) that produce more hydrogen then the wild type. But just as there is no single technical solution for the world’s future energy problems, which will need multiple solution like solar panels, wind power, wave power etc
there will probably be no single bacterial solution. We will most likely see several different strains being used; sometimes even simultaneous (72). Some of them might be cyanobacterial but, in the world of ‘Craig Venter’ and synthetic biology, it might not even be anything we ever seen before. An organism might one day be designed from scratch, containing only the parts that we decide, bringing genetic engineering to a new level, and there is nothing saying that at least some of those parts might be of cyanobacterial origin. “Playing God” will definitely have a new meaning in the future bringing new challenges to the human race, not just from a scientific perspective but from an ethical and moral perspective too. Creating organism that will solve our energy crisis will hopefully be on the positive side on that list of pro’s and con’s.

To design a completely new organisms for future hydrogen production will need more research and a much higher understanding of every part of the cell i.e. gene regulation, metabolism, protein function, protein regulation etc. We need to know which part to use, how they will behave within the cell, which effect they will have on their surrounding and how they should be regulated. And there we are again, the importance of transcriptional regulation...

And with those words I will end this thesis. I might not have answered all the questions about transcriptional regulation and control of hydrogenases in cyanobacteria, you could even argue that I have instead created more questions. But as the saying goes, “even the longest journey begins with a single step”.

50
Svensk sammanfattning

För över 2 miljarder år sedan utvecklades de första cyanobakterierna (även kallad blågrönalger). Dessa bakterier var de första organismerna som kunde utföra fotosyntes där solens energi tillsammans med vatten och koldioxid omvandlas till druvsocker. Det var genom dessa cyanobakterier som jordens atmosfär syresattes vilket var förutsättningen för den snabba utvecklingen av flercelliga organismer som djur och växter. Cyanobakterier är även grunden till de kloroplasters som finns i alla växter idag och där växters fotosyntes sker. De är också en av få organismer som kan binda luftens kväve och omvandla det till ammonium, en process som kallas kvävefixering, och än idag står cyanobakterierna för en stor del av världens biomassan, och är en av de primära producenterna i havet.

Deras enkla krav för överlevnad och deras förmåga att fånga solljusets energi för att skapa socker, vilket utgör bränsle för cellerna, gör dem intressanta för genetisk modifiering och odlings i kommersiellt skala. Det finns planer att idag använda dem till produktion av både olja, metangas och vätgas och deras förmåga att fånga solljus gör att de skulle kunna betecknas som ”Gröna fabriker”. Grön i dubbel bemärkelse då tillverkningen av t.ex. vätgas skulle kunna ske med endast solljus som energikälla och vatten som basmaterial, vilket såklart ger en extremt miljövänlig process. Denna typ av forskning blir extra intressant med vetenkapen om att jordens fossila resurser inte är oändliga utan beräknas ta slut inom de närmaste 100-200 åren. Användandet av fossilt bränsle har också fått oanade följder där global uppvärmning måste anses vara den mest allvarliga. Varje år släpper vi ut 400 gånger mer kol, i form av koldioxid och kolmonoxid, än vad jordens alla växter och bakterier lyckas fånga in via fotosyntesen. Effekten av alla dessa utsläpp har studerats länge och summerats av ”the Intergovernmental Panel on Climate Change” (IPCC). Slutsatsen var att den globala uppvärmningen kan leda till att isen på jordens poler smälter, förhöjd vattennivå, torka i stora delar av världen, ändrade vatten och vind strömmar, att ett flertal arter utrotas och även utgöra ett stort hot mot mänskligheten.

Att använda cyanobakterier för produktion av t.ex. vätgas är ingen enkel process. Bland cyanobakterierna finns det framförallt två enzymer som kan bilda vätgas (nitrogenaser samt ett reversibelt hydrogenaser) men det finns också enzymer som tar upp vätgas (upptags hydrogenaser). De båda hydrogenaserna är besläktade och båda går igenom en lång process där olika delar av enzymet sätts samman, innan de är funktionella. Ett steg i denna process

Sammanfattningsviss visar denna avhandling att:

I …de proteaser som klyver hydrogenaserna inte bara är specifika för hydrogenaser i cyanobakterier utan att även de enskilda proteasererna är specifika för en viss typ av hydrogenaser. Det betyder att de skulle kunna användas i framtiden för en specifik aktivering av en viss typ av proteaser.

II …dessa proteaser regleras i cellerna på ungefär samma sätt som de hydrogenaser de klyver och att de påverkas av samma regulatoriska proteiner som hydrogenaset, dvs NtcA, LexA och CalA. Detta ger oss ökad kunskap om vilka regulatoriska mekanismer som finns naturligt i cellen på gennivå och som är viktig för vätgasmetabolism.

III …de regulatoriska proteiner NtcA består utav en hel familj av närbesläktade tidigare ej studerade proteiner som kan tänkas reglera både hydrogenaser och nitrogenaser på gennivå. Detta ger oss en ökad förståelse för det komplexa reglerings system som naturligt finns i cellen och som kan behöva ytterligare kartläggning ifall detta regulatoriska system skall användas i en genetisk modifierad organism.

IV …hydrogenaserna har en lång historia och antagligen utvecklades hos de ursprungliga arkebakterierna för över 3 miljarder år sedan, och då framförallt hos de så kallade metanogener. Dessa metanogener har metan som slutprodukt i sina energigivande processer anses ha haft sin storhetstid innan cyanobakterierna utvecklades och syrehalten i atmosfären höjdes.
Welcome to the most popular page of this thesis. I think it is impossible to spend so many years doing a PhD without having a lot of people to thank, and believe me, I am so thankful! I therefore hope that I have not forgotten anyone, and if I have I am truly sorry.

I want to start with thanking my supervisor Professor Peter Lindblad who gave me this great opportunity to start and complete a PhD within his research group. Throughout the years I have received invaluable knowledge in manuscript writing and research thinking and for that I will always be truly grateful for.

I also want to thank my supervisor Dr. Karin Stensjö, who has been such a support through the years. You started off as an “ordinary” supervisor but today I would rather call you a friend and most of all a mentor, and I guess that is how it should be. I will always miss our scientific discussions and to be able to pop into your room at any time with even the smallest and silliest question (which I also did several times).

Marie and Åsa, gosh, were should I start! We started off in the same little office room so many years ago and it definitely felt awkward to finish off without you here. Even though I guess it was for the best, drinking tea and gossiping with you girls was just too nice. We have so many (strange) memories to share and I will always count you as some of my best friends. I would also like to give a special thanks to the “Marie and Daniel day-care service” whose opening hours and service is extraordinary!

To the rest of the team in the Cyanogroup, past and present, I want to give a big thank to all of you, it has been an honour working with you all. If I should mention any names... :-)

Well, Paulo, I know I had a tendency to use you as a walking, talking encyclopaedia, but it was just too easy. I guess that being bothered with questions all the time is the drawback of being such a brilliant and kind researcher as you are!!! I have a feeling you can live with that.

Pia, I got to know you when you were my exam project supervisor, so many years ago, and I truly missed you when you left the group for France and then the US. You are a great person, supervisor and researcher and I think getting you back into the group was one of the best things that could have ever happened to all of us.
To Martin and Petra, a special thank for great discussions and being entertaining laboratory companions, especially during these last 6 months. Daniel, never stop being who you are, you are so inspiring and always looks ahead into the future. You might have some mad ideas once in a while, but I think you need that ability to be able to come up with the truly great ones. I’m looking forward to seeing what you come up with next. To Fernando, Åsa S., Hsing-ho, Thorsten, Wantanee and Pond, many, many thanks for making my time in the cyanogroup so special. I also want to thank exam-workers, Daniel and Peter and project student, Linnea for working so hard within my projects and for being such great students.

I would also like to thank everyone else at Fotomol for being great co-workers, you are all fun and entertaining people during coffee breaks and lunches and all your scientific input through discussion and general comments has been brilliant.

I also want to thank the people that make everything go round. Leif Hammarsröm for being such a great prefect, Stenbjörn Styring for running the Consortium of Artificial Photosynthesis and Sven for being the best janitor we could ever have. A BIG thank you (!) to all the administrative staff, past and present, especially to Åsa and Susanne who have been great through the years. I really don’t know what I should thank you the most for, your fantastic skills as administration personal or your fantastic skills as great companions during “fika” breaks?

Gunilla... I have certainly not forgotten you. You are such a positive, funny and witty person and I have truly missed you since you retired and left the department. A bit of sunshine disappeared from Fotomol when you retired.

Elin, what a friend! You have always been there for me and I hope I will always be able to be there for you.

When it comes to my family I want to start with giving a special thank to my mum and dad, who always have supported and been there for me. Your love and care have meant a lot for me and made me to who I am today! I also want to thank all the strong-minded women in my family which have been such an inspiration for me through the years even though you might not have known it, e.g. (my mum), my sister Helena, my “farmor” (granny) Goldis, my late “mormor” (grandma) Inga (1920-2004) and my aunt Anita. You are (were) all fantastic people and role models and are all part of me.

I also want to thank my parents-in-law, Caroline and William, for being such a wonderful support, not just for my husband but to our whole family. We are not always that good at expressing our feelings but we have both really appreciated all the help you have given us over the years.

Last, but not least I want to give a big thank and hug to the most special people in my life today, my husband and two sons.

Stewart I think you know that your support during these last couple of months (years!) have been fantastic and meant so much to me. You have
always been there for me and supported me in a way that is hard to understand sometimes. The simple fact that you moved from your own home country to be with me here in Sweden says it all. You have had to put up with a lot through the years, mind-numbing science talk, weekend jobs, late hours and a wife with a completely absent mind. Hopefully those first signs of dementia will, at least, only be temporary. You make me so happy and give me so much energy and what can I say... you complete me!

Teodor and George, you are still too young to understand what a PhD and a thesis is. But when you are old, and read this, I want you both to know that the hugs and kisses you have overloaded me with since you were born, all the smiles and the love, have given me so much strength and energy. I might have started off doing this PhD for myself but now, several years down the line, I feel more and more that I am doing it all for you (and no, that doesn’t mean I will do your laundry, dishes or clean your rooms in the future!!).

Stewart, Teodor and George, I love you so much!


42. **He, D. L., and X. D. Xu.** 2010. CalA, a cyAbrB protein, binds to the upstream region of *ftsZ* and is down-regulated in heterocysts in *Anabaena* sp PCC 7120. Archives of Microbiology 192:461-469.


120. Theodoratou, E., A. Paschos, A. Magalon, E. Fritsche, R. Huber, and A. Böck. 2000. Nickel serves as a substrate recognition motif for the


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 792

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)