Transcriptional regulation of two putative proteases in the cyanobacteria

*Nostoc* sp. PCC 7120

Daniel Vestman
Summary

There is a great interest in using microorganisms to produce molecular hydrogen, an ideal energy carrier in a time where fossil fuel reserves are starting to decrease. One possibility is to use genetically engineered cyanobacteria in large-scale bioreactors, exploiting solar energy for a clean and efficient hydrogen production without harmful waste. The background for this perspective is that many cyanobacterial species have a natural ability to produce hydrogen gas. This hydrogen is actually a bi-product in a process where the cyanobacterium utilizes atmospheric nitrogen gas to make ammonia, a nitrogen compound accessible to the organism. One problem is that the produced hydrogen is recycled to a great extent by enzymes called hydrogenases. This is a group of enzymes fairly common in microbial species and they can have somewhat different properties. Cyanobacterial hydrogenases are either hydrogen consuming (uptake hydrogenase) or both consuming and producing (bidirectional hydrogenase). Hydrogenases are very complex and require many steps of modification, so called maturation, before they are fully active. Mutants with defective uptake hydrogenase are yielding substantially more hydrogen gas than the wild type species and are showing the way for future development.

*Nostoc* PCC 7120 is a free-living filamentous cyanobacterium capable of producing hydrogen gas through the process of nitrogen fixation. It has both an uptake hydrogenase that recycles the hydrogen and a bidirectional hydrogenase which function is not clearly understood. In this project, I have studied two maturation proteins from a regulatory point of view. The proteins, HupW and HoxW, are proteases that cleave a C-terminal sequence from the uptake hydrogenase and the bidirectional hydrogenase, respectively. I performed 5'RACE experiments to find the transcriptional start point of the two corresponding genes, *hupW* and *hoxW*, and found one 234 bp upstream of *hupW*. Correctly located according to this, I found the –10 and –35 boxes. These are more or less conserved sequences that act as recognition and binding sites for the α-subunit of the RNA polymerase. This is a step towards a more complete understanding of the mechanisms that govern the genetic control of these proteins. It is also a step on the way to unlock the potentials of cyanobacteria as future energy producers.
The Cyanobacteria

Named after the pigment phycocyanin, cyanobacteria, or blue-green bacteria, are a diverse group of prokaryotic organisms present in a wide range of habitats like oceans, lakes, terrestrial and extreme environments. The cyanobacteria perform oxygenic photosynthesis based on chlorophyll $a$ and are believed to be the evolutionary ancestor of the eukaryotic chloroplast. This ancestry is reflected in the similarity of the cyanobacterial and the eukaryotic photosynthesis. Depending on species, they can be unicellular, filamentous or colonial, and many live as symbionts together with plants, fungi or mosses. The cyanobacteria are of ancient origin, at least 3500 million years old and fossilized remains are abundant, especially from the Proterozoic era (2500-570 m years ago). Oxygen evolving cyanobacteria are thought to be responsible for oxygenating the earth’s atmosphere. Remains of their early evolution and the environmental conditions of that time are for example tolerance of low oxygen levels, high tolerance of UV radiation and free sulphide. Adaptations to low levels of light and an effective system for utilizing nutrients has allowed them to dominate vast areas of the world’s oligotrophic (nutrient poor) seas. They are also important in many terrestrial environments like deserts and semi-deserts, where they live on or under the topsoil and help to stabilize it. One great advantage of many cyanobacterial species is their ability to fix atmospheric nitrogen (N$_2$) and convert it into a biologically more accessible form, ammonia (NH$_3$). This feature can give them a competitive advantage over other organisms when there is a shortage of combined nitrogen in its environment. Molecular hydrogen (H$_2$) is produced as a bi-product in this process. Many species also have special enzymes dedicated to the production and recycling of hydrogen gas. In every day life cyanobacteria are perhaps mostly known to the public for causing toxic blooms and as health food (Spirulina, figure 1).

Figure 1. Spirulina grown for commercial use in large facilities. (http://www.alibaba.com/photo/10662735/Spirulina_Plantenis_Growing_In_Greenhouse.jpg)
The heterocyst – a specialized cell

Since the mechanism for nitrogen fixation is strongly inhibited by oxygen (O₂), some filamentous species have evolved a special type of cell to facilitate this reaction where the oxygen level is kept at a minimum. These cells, called heterocysts, differentiate at a regular interval along the filaments when the need for nitrogen arises, or as a response to chemical signalling from a symbiotic organism (Figure 2). Several features help in keeping the oxygen levels low. A thicker cell wall prevents O₂ and other gases from diffusing in. There is no active photosystem II that can produce O₂ by splitting of water. A higher respiratory activity reduces the small amount of O₂ still present to almost anoxic levels. The product of the N₂ reduction, ammonia, is incorporated into glutamine and exported to the vegetative cells of the filament, which in turn exports carbohydrates to the heterocysts.

Nitrogen fixation and hydrogen evolution

The ability to transform atmospheric nitrogen into ammonia is due to the enzyme nitrogenase. This fixation of nitrogen is accompanied by the reduction of protons (H⁺) to molecular hydrogen (H₂). The full reaction reads as follows:

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

This is a energy requiring, but very useful reaction since nitrogen is essential for growth and often in short supply in many biotopes. As can be seen in the reaction formula above, one of the products of nitrogen fixation is molecular hydrogen, H₂. This hydrogen is to a great extent recycled by an uptake hydrogenase to yield protons and electrons. The uptake hydrogenase has so far been found in all nitrogen fixing cyanobacteria examined. In addition to this, some cyanobacteria also possess a bi-directional hydrogenase that both consumes and evolves molecular hydrogen. The function of the bi-directional
complex is not really understood, but it is though to take part in the regulation of reducing potentials and electron transport (Tamagnini, et al. 2002). The cellular localization of the hydrogenases is so far somewhat uncertain. However, the uptake hydrogenase is thought to reside on the cytoplasmic side of the cytoplasmic or thylakoid membrane while the bidirectional hydrogenase is soluble or loosely associated to a membrane. There are different types of hydrogenase enzymes that are classified according to their metal content and conserved structure. The ones in Cyanobacteria are of Ni-Fe type and closely related to hydrogenases present in i.e. E.coli. Hydrogen is a good source of energy for facultative chemolithotrophes and the hydrogenases are present in almost all Bacteria and Archaea (Casalot, et al. 2001).

Hydrogenase maturation

Hydrogenases go through several steps of modification before they are fully mature and active. Though very little is known about this process in cyanobacteria, it is well studied in E.coli which possess several hydrogenases. In E.coli, one set of so called hyp-genes (hypABCDE) handle the maturation of hydrogenases. However, all hydrogenases (except one) also have a specific C-terminal endopeptidase (HyaD, HybD and HycI). It is responsible for the last part of the maturation process, the removal of an amino acid sequence at the C-terminal end of the large subunit. This leads to proper closure of its metal centre, binding to the small subunit and enzymatic activity. The data on C-terminal endopeptidases are derived mainly from the crystal structure of HybD (Fritsche, et al. 1999) and genetic and biochemical data from HycI (Rossmann, et al. 1995), both from E.coli. The key to this specific recognition seems to be the 3D structure of both the peptidase and the hydrogenase, rather than any sequence consensus. It has also been shown that the incorporation of nickel into the large subunit is essential for the binding of the peptidase. Much is still unknown about this process and the crystallization of an endopeptidase in complex with its substrate would be very useful (Theodoratou, et al. 2005). Sequence homologues of the protease genes have been found in all sequenced cyanobacteria. In Nostoc sp. PCC 7120, two peptidases can be found. One for the uptake and one for the bi-directional hydrogenase. Although the best amino acid sequence similarity to an E.coli counterpart (hybD) were as low as 42 and 30%, comparisons between the 3-D structures showed almost exact matches. Moreover, the supposed metal binding amino acids had just the right sterical conformation (Figure. 3). It is proposed that these cyanobacterial hydrogenase endopeptidases should be called HupW and HoxW according to previously used nomenclature. Transcriptional studies showed that both proteases are constantly expressed while the hyp genes only are expressed under N2 fixing conditions. The proteases also differ from the hyp genes in the respect that they are not transcribed together with or located close to the hyp operon. Based on these findings, it has been suggested that the proteases might be under some kind of posttranscriptional regulation (Wünschiers, et al. 2003).
All work in this study was performed using *Nostoc* PCC 7120, formerly known as *Anabaena* sp. PCC 7120 (PCC stands for Pasteur Culture Collection). From here on, I will reference to it as simply *Nostoc* 7120. It is a free-living heterocystous strain and the first filamentous cyanobacteria to have its genome sequenced (Kaneko, et al. 2001). *Nostoc* 7120, like all other nitrogen fixing strains, possesses an uptake hydrogenase consisting of a large and a small subunit, HupL and HupS. It also possess a bidirectional hydrogenase consisting of the five subunits Hox- E, F, H, U, and Y (Axelsson. 2003).

**Biohydrogen**

Today there is a growing concern about the future energy situation. The reserves of oil and natural gas are being depleted worldwide and rapidly developing regions in Asia and Africa are claming their part of the black gold that made the western world rich during the 20th century. At the same time we are starting to understand the environmental effects of elevated CO2 levels. Molecular hydrogen is emerging as one of the more attractive alternatives for a future energy carrier. It has a high energy-content and when combusted, the only product is water. There are several projects worldwide where fuel cells are used to drive busses, cars and trucks on hydrogen gas. However, industrial production of H2 using available methods is very energy consuming and therefore not a solution. One promising alternative is to use hydrogen-evolving microorganisms. Here, cyanobacteria are ideal candidates requiring only mineral salts, water and sunlight. The waste bodies can be used for fertilization or to feed fishes. Mutant strains with non-functional uptake hydrogenases have been engineered that can produce up to 3.7 times more H2 (167.60 µmol/mg chl a/h, *Anabaena variabilis*)
PK17R) than wild type (Lopes Pinto, et al. 2002). Outdoor experiments with bioreactors (Figure 4) show that there is still much work to be done in the future before cyanobacteria can emerge as a commercially attractive alternative (Lindblad, et al. 2002). Future research objectives include better methods of screening for strains with effective N₂ fixing and H₂ evolving properties and better performance at high light intensities (Masukawa, et al. 2002). Another exciting alternative is to introduce higher yielding H₂-evolving hydrogenases into cyanobacteria from i.e. green algae.

![Figure 4](image)

**Figure 4.** A schematic drawing showing how a bioreactor containing cyanobacteria feeds hydrogen gas into a fuel cell which generates electricity, with water as the only waste product (Lindberg. 2003).

**Materials and methods**

**Growth conditions**

*Nostoc* 7120 was grown at 25°C under constant light and aerobic conditions in BG11₀ (Rippka, et al. 1979) medium containing no combined nitrogen, in order to induce nitrogen fixation.

**RNA extraction**

Cyanobacteria posses a thick cell wall composed of polysaccharides. This makes the isolation of nucleic acids somewhat difficult. 50 ml of culture was centrifuged at 4 Krpm (Eppendorf Centrifuge 5417R with F45-30-11 rotor was used for all centrifugations) for 5 minutes, the cell-pellets
were resuspended in 1 ml TRIzol reagent. App 0.2 g of 0.5 mm glass beads was added and the cells were disrupted using Fast-prep (FP 120 BIO 101, Savant). Insoluble material was removed by spinning at 12K rpm for 10 min and collecting the clear upper phase. The samples were incubated at room temp for 5 min, 0.2 ml chloroform was added, the tubes vigorously shaken and then left for 3 min at room temperature. After spinning at 12 k rpm for 15min, the RNA remains in the upper phase. This was saved and mixed with 0.25 ml isopropanol and 0.25 ml salt precipitation solution and then incubated at room temperature for 10 min. To precipitate the RNA as a pellet, the samples are centrifuged at 12 K rpm for 10min at 4°C. The supernatant is removed and the pellet is washed by vortexing with 1 ml 70% EtOH. The sample is spun at 7,5 K rpm for 5 minutes at 4°, the supernatant is poured of and the remaining pellet left to air-dry for some minutes. To DNase treat the sample, it was resuspended with 20 μl DNAse buffer and 4 μl DNAse (Fermentas) and left to incubate for 30 min at 37°C. The RNA was then extracted from the solution by adding 200 μl ddH2O, 110 μl phenol, 110 μl chloroform, vortexing for 1 min, centrifuging at full speed for 7 min and then transferring the upper phase to a fresh tube. The final step is precipitating and washing. 1 ml 99.5% EtOH (-20°C) and 80 μl 10M LiCl (4°C) was added and the sample was left at -20°C over night to precipitate. The next day, the sample was centrifuged at full speed for 30 min, the supernatant was removed and 500μl of 70% EtOH was added. After 5 min of full speed spinning, the supernatant was poured of and the pellet dried for 15 min. Finally, the RNA is resuspended in 20μl ddH2O. The RNA concentration of the preparation was measured by spectrophotometry at a wavelength of 260 nm, and the quality was checked by electrophoresis on a 1 % agarose gel.

5' RACE

5' RACE is a technique for amplifying a specific mRNA sequence from a defined internal site to its 5' end. First, a cDNA is created using a gene specific primer. The RNA is degraded and the newly synthesized cDNA is purified. Then homopolymeric tails are added to the cDNA strands using TdT (Terminal deoxynucleotidyl Transferase). The tails serve as a binding site for a so-called Abridged Anchor Primer (AAP), which is used together with a second gene specific primer to amplify the cDNA. For even greater yield and specificity, a second PCR step with a third gene specific primer, and a so-called Abridged Universal Amplification Primer (AUAP) is done. In these experiments, I used a commercially available kit (5’RACE System for Rapid Amplification of cDNA Ends, Version 2.0; Invitrogen), according to instructions (Figure 3). All primers used are listed in Table 1. The results of the first and second PCR were checked on a 1% agarose gel.
Figure 3.
The principles of 5' RACE. Three gene specific primers are used together with anchor primers to fabricate a stretch of cDNA corresponding to the mRNA, from a given site inside the gene of interest, to its 5' starting point.
Cloning and plasmid preparation

The PCR products from the 5'RACE experiments were ligated in to a cloning vector and transformed in to chemically competent cells (TOPO® TA Cloning; Invitrogen). White and light-blue colonies were picked and the plasmids were extracted using a kit (HiYield Plasmid Mini Kit; RBC Labs) according to protocol. The final plasmid concentrations were determined by spectrophotometry at a wavelength of 260 nm.

Restriction analysis

To visualize the length of the inserts into the plasmids, samples of the plasmid preparations were digested with an appropriate restriction enzyme. In this case, it was BamH I. The reaction set-up was as follows: 2 μl of undiluted plasmid DNA, 1 μl Reaction Buffer 10×, 0.2 μl Restriction enzyme and 6.8 μl ddH₂O. Samples were left to incubate at 37°C for 1 h.

RT-PCR

An RT(reverse transcriptase) -PCR was performed to visualize the length of the transcript upstream of hoxW. The procedure was performed using a kit (RevertAid™ First Strand cDNA Synthesis Kit; Fermentas) according to protocol.

Table 1. Oligonucleotide primers used in this work. Optimal annealing temperatures were determined experimentally.
using a gene-specific primer, \textit{hoxW\_NB\_Reverse}. Primers directed at the two orfs \textit{all0771} and \textit{all0772}, preceding the \textit{hoxW} gene, was used in a PCR to see if it was possible to amplify internal sequences from those. A product would indicate the presence of that orf in the \textit{hoxW}-transcript. All primers used are listed in table1.

Results

Sequencing of the 5'RACE products revealed a single putative transcriptional start point (Tsp) located 234 bp upstream of the startsite of \textit{hupW}. With this in mind I identified a –10 box (tataat) centred 10,5 bp upstream of the Tsp. A less conserved –35 box (ttaaa) centred 34,5 bp upstream was also found (figure 4). The 5'RACE analysis of the \textit{hoxW} gene did not lead to any conclusive result, and nothing resembling a –10 or –35 region was found upstream of the ORF. However, the RT-PCR indicates that the ORFs proceeding \textit{hoxW} are not part of the same transcript. Only the primers for \textit{hoxW} gave any detectable product (Figure 5).

\begin{center}
\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Intergenic region ahead of \textit{hupW}. The transcriptional starting point for \textit{hupW} is marked with blue. –10 and –35 regions are underlined and bp that conform to the \textit{E.coli} consensus are in bold. Start codon of \textit{hupW} is in bold lower-case.}
\end{figure}
\end{center}
Discussion

–35 sites in cyanobacteria are often hardly conserved and only present in about half of the promoter regions that have been studied. The sequence upstream of hupW show some conformation to the E.coli consensus and Gel electrophoresis of the 5’ RACE products show three bands at different lengths for the hupW. These might be due to unspecific priming, but I believe that it is to early to rule out the possibility of additional TSPs at this point. A –35 sequence identical to the one from hupW has been found upstream one of four TSPs for the glutamine synthetase gene (glnA) in Nostoc 7120 (Tumer, et al. 1984). That transcript (III) was only found in cells that were grown with combined nitrogen, while others (I &IV) were found consequently or only under nitrogen fixing conditions (II). For the hoxW gene, there is no obvious explanation to why the 5’RACE failed to provide any relevant data. As shown below (Figure 6), a distinct band is present
in the negative control, lacking TdT tailing, for hoxW. This product showed up consequently in several experiments (data not shown), seemingly without explanation. For further 5’RACE studies, I suggest that modifying the present protocol, changing kit or primers might provide additional data. To position the primers further upstream can sometimes improve results when dealing with long transcripts. It would also be very interesting to do Northern hybridisation for both of the endopeptidase genes to get a better idea of the length of the transcripts and possibly verify the putative transcriptional starting point for the hupW gene.

There are often one or several operator sites located in the promoter region, even though no such sequences could be identified here. More research is required to get a full picture of the way in which the hydrogenase maturing enzymes affect the H₂ production in cyanobacteria. Further sequencing of cyanobacterial genomes and bioinformatic methods might also provide relevant clues since comparative genomics is proving to be a powerful tool.

![Figure 6](image)

*Figure 6.* Products from the 5’RACE nested PCR. hupW (1). hupW – TdT tailing, negative control (2). hoxW (3). hoxW – TdT tailing, negative control. hupW – PCR product, negative control (5). hoxW – PCR product, negative control (6).

**Acknowledgements**

I would like to thank Ellenor Olsson for supervising my work, giving me invaluable insight and advice and for designing primers used in the 5’RACE procedures. I would also like to thank Peter Lindblad for supervising and advising. And big thanks to the entire cyano group, and the rest of PhysBot for being so nice to me!
References


**Internet sources**