UPPSALA
UNIVERSITET

UPTEC X 17 029
Examensarbete 30 hp
Januari 2018

Characterization of promoter regions for design of a synthetic cell specific promoter

Katja Annala

UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

Characterization of promoter regions for design of a cell specific promoter

Katja Annala

Global warming is one of the modern world's biggest problems and renewable fuels are needed to meet this challenge. Biohydrogen is one alternative of a renewable fuel and can be produced from genetically modified cyanobacteria. Although the production is not yet efficient enough to compete with fossil fuels on a commercial level, it has potential. By gaining knowledge regarding pathways and regulatory mechanisms in cyanobacteria, genetic tools can be built and further enhance the production of biohydrogen. A synthetic promoter, which is active only in hydrogen producing cells, heterocysts, is an example of a wanted genetic tool. Such a promoter would make it possible to overexpress proteins essential for hydrogen production in heterocysts and thereby increase the production of hydrogen. In this study, several promoters have been designed and compared quantitatively on both transcriptional and translational levels. The promoters were compared regarding their heterocyst specific expression and their strength. By comparison of the designed promoters, all with different contents, the essential promoter regions have been identified and two well characterized and heterocyst specific promoter have been proposed.

Handledare: Karin Stensjö
Ämnesgranskare: Ann Magnuson
Examinator: Jan Andersson
ISSN: 1401-2138, UPTEC X 17 029

Populärvetenskaplig sammanfattning

Den globala uppvärmningen är en av nutida människans största utmaningar. Vår dagliga, storskaliga användning av fossila bränslen har lett till höga halter koldioxid i atmosfären vilket i sin tur ger upphov till den så kallade växthuseffekten. Med stigande temperaturer följer ett opålitligt klimat och vissa platser av jorden kommer att bli rentutav svårbebodda. Transportsektorn står för en stor andel av förbrukningen av fossila bränslen och vi behöver ett nytt, miljövänligt bränsle.

Biobränslen är bränslen som är producerade av levande organismer och hit hör bland annat biodiesel, bioetanol och biogas, alla tre används i dagens samhälle. Framtagningen av dessa biobränslen kräver stora mängder råmaterial och odlingen och hanteringen av råmaterialet bidrar till utsläpp av koldioxid och konkurrerar med jordbruket om brukbar jord. För att komma ifrån användningen av råmaterial tittar forskare världen över på att genetiskt modifiera fotosytiserande organismer för produktion av biobränsle. Fördelen med dessa organismer är att de endast behöver solljus, vatten och koldioxid för att leva och föröka sig. Cyanobakterier är fotosyntetiserande organismer som kan utvinna kväve ur både mediet de lever i och genom att fixera kväve från atmosfären. Arten *N. punctiforme* är en filamentös cyanobakterie med filament som främst består av två celltyper, vegetativa celler och heterocyster. Kvävefixeringen sker i heterocysterna och när de fixerar atmosfäriskt kväve bildar de restprodukten vätgas. Vätgas är mycket reaktivt och energirikt och förbrukas därför direkt av cellen. I och med att vätgasen är mycket energirik och att förbränningen av vätgas endast resulterar i vatten som restprodukt är den ett alternativ biobränsle. Genom att genmodifiera cyanobakterier till att släppa ut vätgasen ur cellen istället för att förbruka den skulle man kunna använda cyanobakterier för att storskaligt producera vätgas.

I dagens läge finns det modifierade cellinjer som producera mätbara mängder vätgas, men processen är inte tillräckligt effektiv för att vätgasen ska vara konkurrenskraftig med de fossila bränslena. För att kunna göra processen mer effektiv behöver vi fler genetiska verktyg. Ett önskvärt verktyg skulle vara en syntetisk promotor som skulle göra det möjligt att överuttrycka vissa enzymer endast i vätgasproducerande cellerna, heterocysterna. På så sätt skulle man kunna effektivisera processen och därmed öka produktionen av vätgas utan att försämra tillväxten av filamentet genom att belasta de vegetativa cellerna med syntes av obrukbara enzym.

Detta är ett projekt inom syntetisk biologi med syftet att jämföra ett flertal designade syntetiska promotorer och dra slutsatser kring deras styrka och heterocystspecificitet. Det med förhoppningen om att kunna bygga en både stark och heterocyspecifik promotor.

Table of content

Abbreviations

cDNA	complementary deoxyribonucleic acid
DEF	deficiency-related changes
DIC	differential interference contrast
DIF	differentiation-related changes
DNA	deoxyribonucleic acid
Dps	deoxyribonucleic acid binding protein from starved cells
GFP	green fluorescent protein
ncRNA	non-coding ribonucleic acid
NsiR1	nitrogen stress induced ribonucleic acid 1
PCR	polymerase chain reaction
RBS	ribosome binding site
RNA	ribonucleic acid
RT	reverse transcriptase
RT-qPCR	real time polymerase chain reaction
TSS	transcription start site
WT	wild type
YFP	yellow fluorescent protein

1 Introduction

Due to the ongoing climate change, that is mostly caused by our usage of carbon-based fossil fuels, the need of a more environment friendly fuel is critical. Biohydrogen is such a fuel since the usage of this energy carrier does not contribute to emission of carbon dioxide and therefore has no impact on the so-called greenhouse effect sin. For future biotechnological production of biohydrogen, and other alternative fuels, cyanobacteria is an interesting production host (Khanna et al. 2016). The reason is that the energy source for cyanobacterial production is solar energy and the source for reducing power comes from water. In cyanobacteria, the enzymes involved in H_2 production are O_2 sensitive and therefore H_2 production in some cyanobacteria takes place in specific cell types, i.e. the heterocyst, with a microoxic environment (Kumar et al. 2010). Since the production of H_2 takes place in the heterocysts it is of interest to be able to control the metabolic pathways in the heterocyst without affecting the vegetative cells. At present time, is does not exist such engineering tools. There is a lack and need of a synthetic, well characterized, heterocyst specific promoter. Such a promoter is under development and this project will focus on establishing the function of the different parts of the promoter sequence.

1.1 Heterocyst

1.1.1 Cell differentiation

Cyanobacteria provide nitrogen from the media it is living in but in the absence of combined nitrogen it can fixate atmospheric nitrogen through the enzyme nitrogenase within produced heterocyst. The photosynthetic *N. punctiforme* live as filaments containing hundreds of vegetative cells and in the absence of combined nitrogen every ten to twenty cell differentiate into a heterocyst (figure 1). The differentiation will be complete within approximately 20 hours at 30 °C after nitrogen step-down (Kumar et al. 2010). Morphologically, the heterocyst are distinguishable from the vegetative cells in several ways. They are larger and have a relatively rounder shape, diminish pigmentation and thicker cell envelopes containing an inner glycolipid layer and an outer polysaccharide layer which protects the inner environment of the heterocyst from oxygen (Kumar et al. 2010).

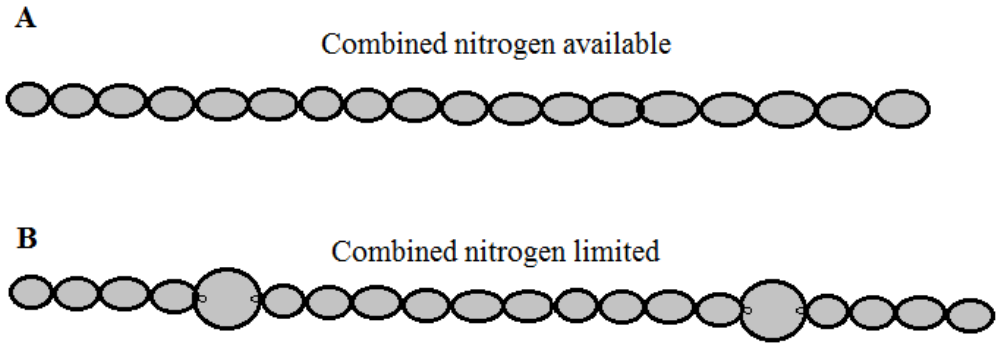

Figure 1. Heterocyst formation in *N. punctiforme*. (A) Filament growing in media supplied with combined nitrogen. (B) Filament growing in media without combined nitrogen, heterocysts are formed every ten to twentieth cell along the filament.

The thylakoid membrane, an intracellular membrane system located in the cyanobacterial cytoplasm, is the cell component where the photochemical reactions of the photosynthesis occur. Upon cell differentiation, the thylakoid membrane reorganizes and proteins not expressed in vegetative cells will be expressed. In the absence of nitrogen, the nitrogen rich and light harvesting phycobilisome antennas located in the thylakoid membrane will be degraded. This phenomenon occurs in all cells of the filament and can be detected by the decrease in auto fluorescence due to the lower amount of photosynthetic pigment. When the nitrogen providing heterocysts are formed, the vegetative cells regain their pigment level while it stays low in the heterocysts (Magnuson & Cardona 2016).

During transition to an atmospheric nitrogen fixating state, when the heterocyst are forming, the vegetative cells do not replicate. Once the heterocyst are mature, the replication of the vegetative cells continues. The division of vegetative cells will continue until the number of vegetative cells between the heterocyst is too big and next differentiating phase begins. Therefore, the pattern of vegetative cells and heterocysts of filaments living in nitrogen poor media will be conserved (Muro-Pastor et al. 2017).

1.1.2 Cell differentiation on a genetic level

Metabolites can diffuse between cells in the filament. This makes it possible for the fixed nitrogen produced by heterocysts to reach vegetative cells and nutrients produced by vegetative cells to reach heterocysts. This is essential for amino acid synthesis in vegetative cells and the maintenance of high energy consuming fixating in heterocysts. It also contributes to a certain communication between the cells since they will be affected by regulators expressed in adjacent cells and not only by themselves. This kind of communication is essential when trying to understand cell differentiation and pattern maintenance (Muñoz-García & Ares 2016).

On a genetic level, the expression of the transcriptional regulator NtcA is one of the first steps in cell differentiation (Kumar et al. 2010). *ntcA*, induced by nitrogen stress, is an important

trigger and an activator of several genes. NtcA induces *hetR* which encodes HetR, a central regulator during differentiation. HetR protein forms homodimers with DNA binding activity and is a master regulator of several genes during the development of a heterocyst (Flaherty et al. 2014). *hetR* is expressed at low levels in all cells but the expression increases during the early stages of cell differentiation. The two genes, *ntcA* and *hetR*, and their expression are mutually dependent (Muñoz-García & Ares 2016). The gene hetR is negatively regulated by the PatS protein, and more precise the caboxyle-terminal pentapeptide RGSGR of the PatS protein, which prevents the DNA binding activity of HetR. The expression of *patS* is low in vegetative cells, strongly up regulated in differentiating cell and not expressed at all in heterocyst. Due to this, the diffusing short peptide PatS will prevent nearby cells of the differentiating cell to enter differentiation (Muñoz-García & Ares 2016). HetR as a positive regulator and PatS as a negative one provides a minimal understanding of regulation of differentiation and heterocyst pattern formation, see figure 2 (Kumar et al. 2010, Muñoz-García & Ares 2016).

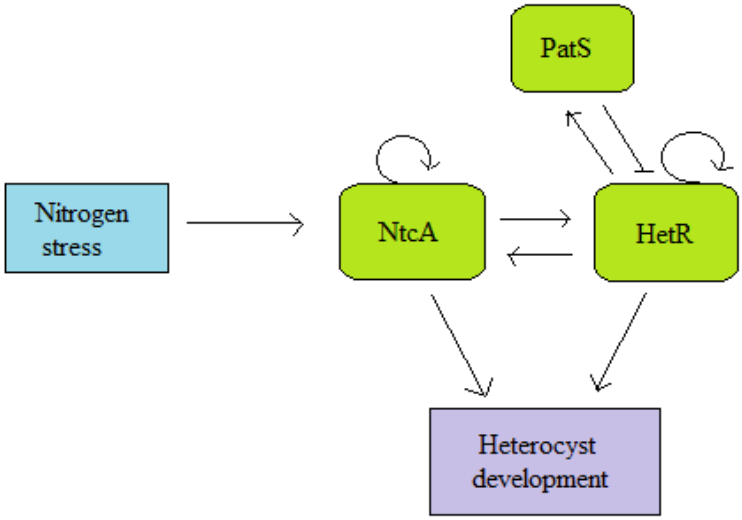

Figure 2. A schematic figure highlighting the central regulators, NtcA, HetR and PatS, involved in heterocyst development.

To be able to understand the pattern maintenance and not only the formation another key player needs to be added, the HetR inhibitory factor HetN (Muñoz-García & Ares 2016). The expression of hetN has been shown to be activated after patS and the two proteins PatS and HetN contains the same active motif, the RGSGR pentapeptide. The RGSGR pentapeptide derived from PatS can bind HetR with a stoichiometry relation of 1:1 and since HetN contains the same motif, it is assumed that HetN affects HetR in the same way (Muñoz-García & Ares 2016). In contradiction to *patS*, *hetN* is only expressed in heterocysts and not in vegetative cells. This will provide nearby cells of, first the differentiating cell, and later the mature heterocysts from entering differentiation. The cells in between two heterocysts will only be inhibited from differentiation as long as they are affected by the diffusing peptides. When the

number of vegetative cells between the heterocyst are too big, the middle cells will no longer be affected by PatS and therefore will be able to enter differentiation. Like this, the pattern with a heterocyst in every ten to twenty cell will remain (Muñoz-García & Ares 2016).

1.2 DIF motif in Nostoc spp.

1.2.1 The DIF motif

A study performed in 2011 examined the regulation of cell differentiation. They used differential RNA-seq (dRNA-seq) with the aim to map all transcription start sites (TSS) in Anabaena sp. PCC 7120 (Mitschke et al. 2011). This method is selective for the 5´ends and results in a wide-ranging determination of the localization of TSSs in the genome. Two groups of transcripts was composed. One with all the transcripts produced by the wild type (WT) during nitrogen step-down and one containing all transcripts produced by a mutant lacking the HetR gene and thus lacking heterocysts, also during nitrogen step-down. Since HetR is a major regulator for cell differentiation into heterocyst, the transcripts present in the WT group and not in the mutant group are likely to be a part of the process of heterocyst differentiation. The transcripts present in both categories were named deficiency-related changes (DEF) and the ones only found in the WT category differentiation-related changes (DIF) (Mitschke et al. 2011). The promoter of the non-coding NsiR1 was found in the DIF class and since NsiR1 is conserved in Nostocales they compared the promoter region of 51 repeats, the NsiR1 is transcribed from a tandem array, from five different strains. This resulted in the discovery of a conserved palindrome 5´TCCGGA located around the -35 position of the transcription start site. This motif is to be found in several other heterocyst specific promoters, and with a single mismatch allowed, 58 of the 209 promoters in the DIF class contained this motif. With a fusion of the promoter from NsiR1 repeat 6, containing the DIF motif, and a promoter-less GFP reporter, the relevance of the DIF motif was tested. They saw strong green fluoresces in (pro) heterocyst cells and they thereby assume that the DIF motif is essential for heterocyst-specific expression. To test this theory, the palindrome was replaced with the sequence 5´GAATTC which resulted in weak, nonheterocyst-specific fluorescence. Hence, it was concluded that the DIF motif is required for heterocyst-specific expression of the nsiR1 promoter (Mitschke et al. 2011).

1.2.2 The non-coding RNA NsiR1

In 2010, a genome wide study in Nostocales potential regulating non-coding RNAs (ncRNA) were identified (Ionescu et al. 2010). In this research they found an array of ncRNA consisting of 12 tandem repeats that was transcribed in large amounts during the entering of cell differentiation (Ionescu et al. 2010). They suggested the main accumulated transcript, which has similarities to other regulatory non-coding RNAs, to be nitrogen stress-induced RNA 1, NsiR1. In *Anabaena* sp. PCC 7120, this RNA is about 60 nt with a clear secondary structure and is expressed shortly after nitrogen decrease. In addition, the expression of NsiR1 requires NtcA and HetR and the gene of NsiR1 was located upstream of *hetF*, coding for a

product needed in heterocyst formation. NsiR1 is conserved in all completed Nostocales genomes to be found and the expression of NsiR1 is present in 19 different heterocyst-forming cyanobacteria. All this points towards NsiR being a part of the regulatory network that drives vegetative cells into heterocyst (Ionescu et al. 2010).

1.2.3 The early expression of genes regulated by the DIF

In 2014, Muro-Pastor presented a study where the expression of NsiR1was determined to be the first known step in cell differentiation seen in single cells. The expression of the NsiR1 promoter was determined before the heterocyst could be distinguished from the vegetative cells in morphology. Thereby can the expression of NsiR1 be a marker of the early steps of cell differentiation (Muro-Pastor 2014). Muro-Pastor uses fluorescence to observe differentiating cells of Anabaena sp. PCC 7120. During differentiation, the content of light-harvesting complexes called phycobilisomes is altering which can be seen as a change in fluorescence (Magnuson & Cardona 2016). The red auto fluorescence induced by the phycobilisome is connected with the thylakoid membrane and can be seen in a circular pattern around the vegetative cells, with a less fluorescent center (figure 3). In the beginning of differentiation, the phycobilisome detaches from the membrane and this will give the immature heterocyst a uniform fluorescence, covering the whole cell, compared to the vegetative cells (figure 3). The emission from mature heterocyst is very low (figure 3). Further Muro-Pastor coupled a GFP gene to the NsiR1 promoter and could thereby follow the activity of the NsiR1 promoter in different cells. She observed some cells that expressed P_{nsiR1}-GFP but still showed a circular shape of auto fluorescence close to membrane of the cell. These cells are prospective heterocysts which have entered the differentiation but do not show any morphological changes yet (figure 3). The initial induction of genes regulated by the DIF motif, such as P_{nsiR1}, is here shown to be localized to single cells, due to the GFP being expressed in single cells (Muro-Pastor 2014).

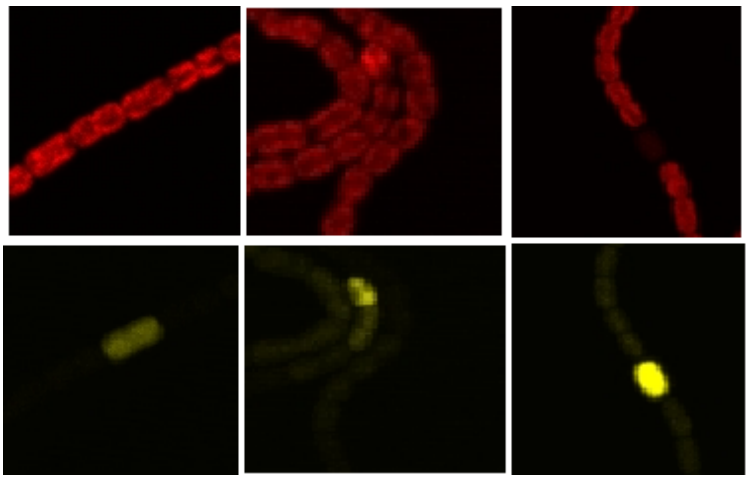

Figure 3. The three stages of heterocyst differentiation. The upper row shows the red auto fluorescence caused by the thylakoid membrane and the lower row the expression of P_{nsiR1}-YFP. Left panels images a prospective heterocyst with the same intensity in red fluorescence as the vegetative cells and low yellow fluorescence. The middle panels shows immature

heterocyst with decreased red fluorescence and increased green fluorescence. The right shows panels mature heterocyst showing almost no red fluorescence at all and clear green fluorescence.

The method of using confocal microscopy to visualize the cells and their status was used in this study. The filaments will be analysed as by Muro-Pastor (2014).

1.2.4 One of two promoters regulating alr3808 contains DIF

In a paper by Li et al. published in 2015, two orthologous and differentially regulated Dps (DNA-binding proteins from starved cells) genes named Npun_R5799 and alr3808 were studied (Li et al. 2015). The two genes originate from *N. punctiforme* and *Anabaena* sp. PCC 7120, respectively. By analysing the cell specific promoter regulation of these two genes, the putative importance of the encoded proteins for heterocyst specific metabolism were examined. Two different transcription start sites were found upstream of alr3080, TSS1 and TSS2, and one of them, TSS2, upstream of Npun_R5799. The DIF-motif is located 31 nucleotides upstream of TSS1 positioned at -35 (figure 4) and TSS1 is dependent of HetR. TSS1 was induced early during differentiation and active in both heterocysts and proheterocysts. TSS2 on the other hand was not dependent of neither HetR nor NtcA and TSS2 is therefore most likely to be active in all cells of the filament (Li et al. 2015).

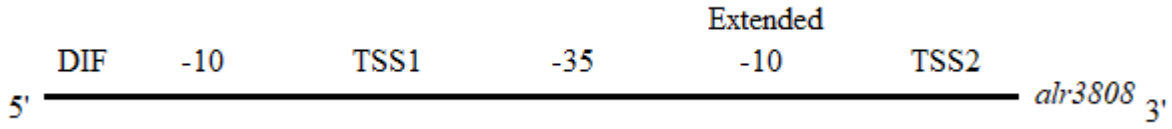

Figure 4. A figure showing the upstream regions of alr3808. The two minus ten regions, the one found together with the DIF motif and the extended minus ten upstream of TSS2, were in this study evaluated as potential minus ten regions in a synthetic promoter.

1.3 Aims and objectives

The aim of this project is to design a synthetic promoter that can be used as a genetic tool for research in the cyanobacteria *N. punctiforme*. This will be done through examination and characterization of different promoter regions.

2 The context of this study

This promoter study is part of ongoing research at the Department of Chemistry - Ångström Laboratory. The starting point for this research are previous results regarding the NsiR1 promoter, the DIF motif, the two transcription start sites TSS1 and TSS2 and an alignment of the transcripts from the DIF containing promoters found by Mitschke et al., 2011 (figure 5).

The alignment was done to identify conserved regions of the promoter sequences published by Mitschke et al. in 2011. In this alignment, a consensus region situated at -10 and a TSS was found together with the DIF.

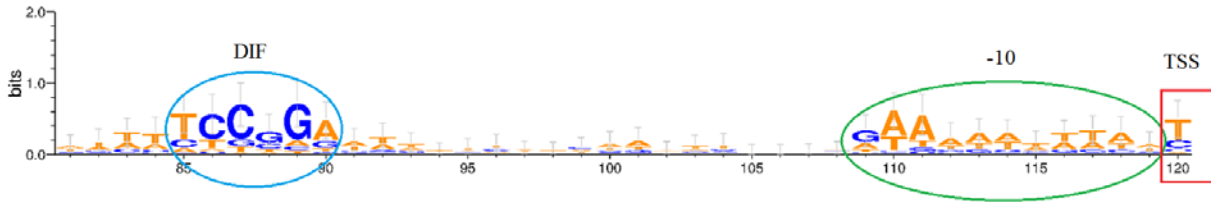

Figure 5. The result of the alignment of the promoter sequences published by Mitschke et al. in 2011. The alignment showing both the already published DIF motif and the discovered -10 consensus region situated upstream of the transcription start site.

2.1 The promoter constructs

When developing a synthetic promoter, the knowledge regarding the regulation must be thorough (Huang and Lindblad 2013). To know all possible regulations of the promoter is essential for controlled induction and regulation. It is therefore not enough to only use a native promoter containing the DIF motif since other unknown regulations can take place. Knowing that the DIF motif is essential for expression and that the NsiR1 promoter is expressed in the early steps of cell differentiation can be used as a starting-point in the development of a synthetic promoter. Based on the alignment in figure 5 and earlier published papers, several promoters were built with the aim of having a minimal, well characterized, heterocyst-specific promoter (table 1 and figure 6). The suggested promoters consisted of either three or four parts. Their location relative to the ribosome binding site are shown in figure 6.

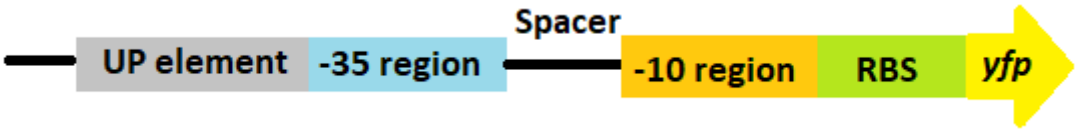

Figure 6. A schematic figure of the promoter parts being described in table 1.

Table 1. Description of the eleven different promoter designs.

Name	Upstream region	-35	Spacer	-10
Native short	no	DIF	*native	*native
Native long	yes	DIF	*native	*native
Native short -10Ptrc	no	DIF	*native	**Ptrc
Synthetic short	no	DIF	scrambled	consensus
Synthetic long	yes	DIF	scrambled	consensus
Synthetic short -35Ptrc	no	**Ptrc	scrambled	consensus
Synthetic short -10Ptrc	no	DIF	scrambled	**Ptrc
Synthetic short -10N	no	DIF	scrambled	*native
Synthetic short -10 TSS1	no	DIF	scrambled	TSS1
Synthetic short -10 TSS2	no	DIF	scrambled	TSS2
No promoter YFP	no	no	no	no

*The native sequence of the eight base pair NsiR1 promoter (Ionescu et al. 2010).

**The Ptrc promoter originates from Escherichia coli, and can be used in cyanobacteria, resulting in high expression of the gene of interest (Wang 2012).

The promoter designed in this study was divided into different regions, as a tool to investigate the importance of each region (table 1). With each construct we aimed to answer questions regarding the different promoter parts.

Native short; can the DIF motif expressed together with a native sequence be enough to give heterocyst specificity?

Synthetic short -10N; does the spacer sequence between the -35 and -10 region has an impact on the promoter activity? The only difference between the native short construct and the synthetic short -10N construct is the scrambled sequence between the -35 and the -10 region. All synthetic constructs contain this scrambled sequence. This to exclude any other potential,

unknown interaction between the native sequence and transcription factors. The scrambled sequence contains the same GC content, and therefore the same number of hydrogen bonds, as the native and is therefore as stable. The comparison between the two constructs, native short and synthetic short -10, can be done as control to ensure that there is no polymerase interaction with the non-conserved native spacer.

Synthetic short; does the consensus -10 sequence have any impact on heterocyst specificity? By comparing the synthetic short construct with the synthetic short -10 construct can the consensus -10 sequence be evaluated.

Native long; does the upstream region have any effect on heterocyst specificity? By comparing the native long mutant with the native short mutant can the upstream region be examined.

Synthetic long; by comparing all four constructs, native long, native short, synthetic long and synthetic short, the importance of the upstream region be evaluated.

Synthetic short -35Ptrc and synthetic short -10Ptrc; is both the -35 and -10 motif needed for heterocyst specificity? If only one of the motifs is needed for heterocyst specificity, will it be possible to combined the needed motif with parts of a strong promoter and thereby have strong and specific expression?

Synthetic short -10 TSS1; will the combination of a SS and the DIF motif give heterocyst specificity? Since the transcription start site called TSS1 is found together with the DIF motif in Anabaena sp. PCC 7120 and is HetR dependent could the fusion of theses motifs give heterocyst specificity.

Synthetic short -10 TSS2; can this -10 region be a substitute for the strong Ptrc promoter? Since the transcriptional start site called TSS2 is likely to be found in all cells, could a fusion together with the DIF motif result in heterocyst specific expression?

2.2 Experimental procedures

The following methods have been used during this project. Here is a general description of the usefulness of these methods.

2.2.1 Electroporation

Electroporation is used to provide transport of a wide range of molecules, there among DNA, through a biological membrane. During electroporation, which is a physical phenomenon, the cells are placed in an electrical field, which leads to an increase in the permeability of the cells. When the cell membrane is exposed to a sufficiently high electric field, pores will be

formed within the membrane leading to possible transportation of DNA into the cells. This phenomenon can be applied to all prokaryotic and eukaryotic cells (Kotnik et al. 2015).

2.2.2 Colony PCR

Colony PCR is used for detection of inserted DNA in the transformant. It is a high-throughput method with which it is possible to determine if a colony contains the wanted plasmid or not. By lysing the transformant and extracting DNA, the eventual plasmid can be analysed with a PCR reaction. Primers designed to anneal to the inserted DNA will be used to detect the inserted DNA. The PCR amplicon can then be analysed by electrophoresis with a DNA ladder for size determination and verification (New England BioLabs. 2017).

2.2.3 RNA extraction and RT-qPCR

RNA extraction is done to extract all transcripts from the different cell samples. A certain amount of RNA is then used for cDNA synthesis. Like this, all samples will contain the same amount of cDNA regardless the difference in amount of cells between the samples. A quantitative real-time PCR, qPCR, which is a method used for analysing gene expression on transcript level. With this method it is possible to estimate the quantity of a certain transcript in each sample. This makes it possible not only to detect expression but also to compare expression of the same gene in different samples (Integrated DNA technologies. 2015). SYBR® green (Quanta BioscienceTM) is a fluorescent dye that non-specifically binds to double stranded DNA and is used in the qPCR reaction. With the fluorescent dye it is possible to follow the amplification of the sequence of interest. When the dye binds to double stranded DNA it increases its emission. So for each PCR-cycle, the amount of double stranded DNA will increase and thereby will the amount of bind dye and emission increase, see figure .

Figure 7. Illustration of one amplification qPCR cycle. The primer binding sites, the cDNA amplification by DNA polymerase how the dye then can bind between the two complementary strains of cDNA, illustrated with green dots (For is the abbreviation of forward primer and Rev is the abbreviation of reversed primer).

The real-time PCR instrument detects emission after each cycle and the software can then calculate and plot a normalized value of the emission (Rn) against the number of cycles (figure 8). The cycle when the emission gets strong enough to be detected by the instrument is called quantification cycle, Cq. The Cq value of a certain gene can then be compared between samples (Integrated DNA technologies. 2015).

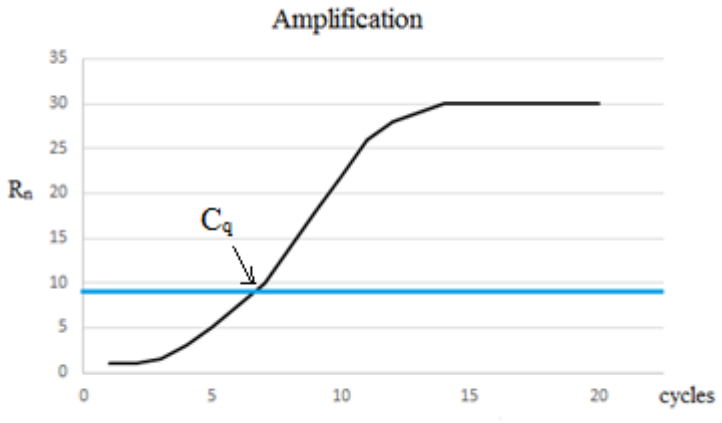

Figure 8. qPCR amplification curve showing Rn value plotted amplification cycles. The threshold line is drawn in blue. The Cq value is where the threshold line crosses the reaction curve.

When extracting RNA some DNA contamination will always be found in the RNA sample. To remove the greater proportion of DNA, a DNase treatment can be used. However, some abundant DNA will still be detectable in the sample, for example DNA encoding the 16s subunit. To ensure that the DNA contamination does not affect the qPCR result, a minus reverse transcriptase (-RT) control is included. When the cDNA synthesis is performed, two reactions for each sample will be run, one with reverse transcriptase (+RT) and one without (-RT). The –RT sample will then be used as a template in the qPCR reaction to detect potential DNA contamination. If the –RT control sample has a Cq value 10 cycles higher than the RT test sample it implies that the –RT sample contained 1000-fold less target sequence, compared to the test sample, and 100% efficiency of the RT reaction of the sample of interest is assumed (Applied Biosystems).

2.2.4 Confocal microscopy

Confocal microscopy is a laser scanning microscopy with the advantage of reducing fluorescence emission or reflection from out-of-focus signals (figure 9), which gives the capability of pictures with high optical resolution. The illumination from the point within the focal plane is the only illumination that passes the pinhole in the confocal aperture and can be detected by the photomultiplier. By scanning over the specimen and combining many point illuminations, an image of the focal plane can be provided (Confocal Laser Scanning Microscopy, C.J.R. Sheppard and D.M. Shotton, 1997).

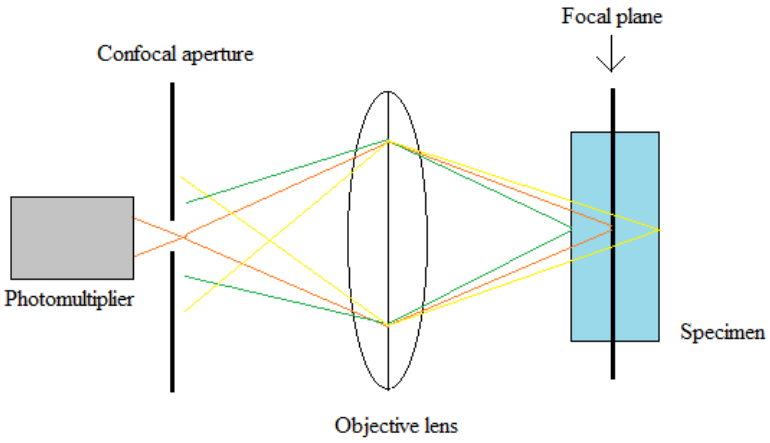

Figure 9. A simplified figure of the confocal principal showing only the emission from the specimen, and not the laser, elicited by the illuminating aperture. Only fluorescence emission from the in-focus plane, orange lines, can pass the confocal aperture unimpeded and contribute to the image. This while emission from regions above, green lines, and below, yellow lines, have other primary image plane foci and will thereby be blocked by the confocal aperture (Confocal Laser Scanning Microscopy, C.J.R. Sheppard and D.M. Shotton, 1997).

3 Materials and methods

If nothing else is mentioned, chemicals are from Sigma-Aldrich®.

3.1 Strains and growth conditions

N. Punctiforme ATCC 29133 mutant strains containing constructs described in table 1 were grown in $BG11_0$, i.e. BG11 (Rippka et al. 1979), supplemented with 5 mM HEPES (pH 7.5), 2.5 mM NH_4Cl and 25 ng/mL kanamycin. They were grown in sterile 100 mL E-flasks at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker until OD_{700} was 0.05 ± 0.02. The cultures were then inoculated in 500 mL bubbling flasks supplied with air containing 400 mL of the same media and grown at the same temperature and light intensity until OD_{700} was 0.14 ± 0.02. For nitrogen step-down, the cultures were harvested and washed in $BG11_0$. The cells were centrifuged at 2500 rpm for 6 minutes, the supernatant was discarded and the cells were suspended in $BG11_0$. The wash was repeated three times and the cultures were then resuspended in $BG11_0$. After nitrogen step-down, the cultures were divided into three biological replicates and grown in bubbling flasks under the same conditions as before.

3.2 Primers and plasmid

A pSCR plasmid was used for insertion of promoter construct into *N. punctiforme,* see supplementary 8.1.

All primers used during this study are listed in table 2.

Table 2. Primers used for colony PCR and RT-qPCR

Primer name	Primer sequence (5'-3')	Length of amplicon (base pairs)
NifH_F	CGTAGAATCTGGTGGTCCAG	342
NifH_R	GCCAAGGTTTCGACAATTCG	342
16s_F	GAATAAGCATCGGCTAACTCCG	195
16s_R	CTACACCAGGAATTCCCTCTGC	195
YFP_F	GCTACCCCGACCACATGAAG	167
YFP_R	GATGCCCTTCAGCTCGATG	167

3.3 Electroporation and colony PCR

3.3.1 Day 1:

N. Punctiforme ATCC 29133 was grown in a bubbling flask supplemented with air and in $BG11_0$ media supplemented with 5 mM HEPES (pH 7.5) and 2.5 mM NH4Cl at 25°C for 3 days. 50 mL of cell culture was spun down at 4000 rpm for 5 min in a Falcon tube and the supernatant was discarded. The cell samples were held on ice and the cells were sonicated (SONICS® Vibra cellTM), with settings amplitude 25, pulse 1p 1.0 sec, 2-3 times. The cells were resuspended in 25 mL $BG11_0$ media supplemented with 5 mM HEPES (pH 7.5) and 2.5 mM NH_4Cl and left at 30 °C, weak illumination for 4 h for recovering. After 4 h, the cells were spun down at 3500 rpm for 5 min and resuspended in 20 mL autoclaved dH_2O, this was done four times, and the fifth time the cells were spun down the pellet was resuspended in 1 mL autoclaved dH2O. The cell density was adjusted to a chlorophyll a content of 50-100 µg Chl a/mL. Precooled ShockPodTM cuvettes were held on ice and for each reaction was 1-3 µg of purified plasmid DNA from *E. coli* mixed with 40 µL cell culture in the cuvette. For electroporation BIO RAD Gene Pulse XcellTM was used with settings 2400 V, Ts 5.0, 2 mm. After electroporation the cuvettes were held on ice until the cells were transferred to sterile E-flasks with 20 mL $BG11_0$ media supplemented with mM HEPES (pH 7.5) and 2.5 mM NH_4Cl. The cultures were incubated overnight in weak illumination at 30°C.

3.3.2 Day 2:
The overnight cultures were spun down at 4000 rpm for 10 min and the supernatant was discarded. The pellet of 200 µL was plated on 1% agar $BG11_0$ plates supplemented with 5 mM HEPES (pH 7.5), 2.5 mM NH_4Cl and 25 ng/mL kanamycin with a preplaced filter, GN-6 Metricel® MCE Menbrane Disc Filters (Pall Corporation), for selection of positive colonies.

3.3.3 Day ~20:
The positive colonies were picked and restreaked on new sterile 1% agar $BG11_0$ plates supplemented with 5 mM HEPES (pH 7.5), 2.5 mM NH_4Cl and 25 ng/mL kanamycin.

3.3.4 Day ~35:
Cells from each colony were diluted in 100 µL dH_2O in 1.5 mL Eppendorf tubes and vortexed. The cells were boiled for 10 min at 100°C and then used as template in the PCR reaction. Two PCR reactions were done for each colony, one with primers annealing to the inserted plasmid and one with primers annealing to the nif gene encoding for the nitrogenases. Purified plasmid and the nif gene worked as positive controls and genomic DNA from *N. Punctiforme* as a negative control. Vice versa for the nitrogenase reaction. With these two reactions combined results regarding both presence of plasmid and genomic DNA can be drawn. The PCR products were analysed by DNA gel.

3.4 RNA extraction and RT-qPCR

3.4.1 Harvesting of cells:
All centrifuges were kept at 4°C. 15 mL falcon tubes (Sarstedt) and 2 mL screw cap tubes (Sarstedt) were pre-cooled at -20°C and then held on ice. 13 mL cell culture was spun down at 4°C 5000 rpm for 5 min and then held on ice. The supernatant was discarded and the cell pellet was transferred to the screw cap tube which was centrifuged at 4°C 5000 rpm for 5 min. The supernatant was removed and the sample was stored at -80°C.

3.4.2 Day 1
All RNA work was done in a fume hood and on ice, if nothing else is written.

The frozen cell pellets were suspended in 1 mL TRIzol reagent. 0.2 g sterile, acid washed glass beads were added to each tube. The cells were broken with a bead beater (Precellys® 24 homogenizer, Bertin Technologies) at 6800 rpm for 30 sec three times. The samples were put on ice for 5 min between the runs. 0.2 mL chloroform was added and the tubes were mixed by hand and then incubated in room temperature for 15 min. The samples were centrifuged at 12000 rpm for 5-10 min at 4°C and the upper, aqueous phase was transferred to new 1.5 mL tubes. 1 volume of phenol-chloroform-isoamylalcohol (25:24:1) was added and the samples were mixed by pipetting and then let to incubate at room temperature for 10 min. The upper,

aqueous phase was again transferred to new 1.5 mL tubes. 1 volume of isopropanol was added and the samples were mixed by hand and the RNA was precipitated at -20°C overnight.

3.4.3 Day 2

The RNA samples were centrifuged at 12000 rpm , 4°C for 10 min, the supernatant was discarded and the pellet was washed with 1 mL 75% ethanol. The samples were centrifuged at 12000 rpm, at 4°C but for 5 min, the supernatant was removed and the pellet was left in open tube to air-dry on ice. When the pellet was dry it was dissolved in 25 µL of DEPC treated H2O. The RNA concentration was measured with NanoDropTM (Thermo Fisher Scientific). 2 µL of RNA was mixed with 2 µL DNA gel loading dye (Thermo Fisher Scientific) and separated on a 1.5% agarose gel to check the quality of the RNA. DNase treatment was done according to PerfeCta® DNase I (Quanta BioscienceTM) protocol with the addition of double U/µL DNase I. A PCR reaction with primers annealing to the reference gene and the target gen used in further qPCR reaction was set up to scan for DNA contamination. qScriptTM cDNA Synthesis Kit (Quanta BiosciencesTM) was used for cDNA synthesis with a total amount of 1 µg RNA in each sample.

3.4.4 Day 3

CFX Real-Time PCR Detection Systems (BIO RAD) with settings based on the $2^{-\Delta\Delta CT}$ method, described by Livak and Schmitting in 2001 (Livak & Schmittgen 2001), was used for analyzing the transcript levels. PerfeCTa® SYBR® Green SuperMix for iQ (Quanta BioscienceTM) was used with a total reaction volume of 25 µL and the reaction was run with 39 cycles. The qPCR reaction was set up with three biological replicates, and three technical replicates for each biological replicate, reducing pipetting errors. Non-template sample was used as a negative control for both the reference gene and the target gene. Genomic DNA from *N. Punctiforme* was used as a positive control for the 16s gene and purified plasmid was used as a positive control for the YFP gene. The cDNA samples and the positive control for the 16s gene contained 50 ng of template and the positive control for the YFP gene contained 10 ng of template.

3.5 Confocal microscopy

Leica DM 6000 CS microscope with argon laser was used for confocal microscopy. Two channels were used for detection of auto fluorescence and YFP fluorescence. Detection of auto fluorescence was done between 600-700 nm and detection of YFP was done between 527-540 nm. The smart gain was set at 612 and 560, respectively. Analyses of heterocyst formation was performed on cells cultivated on plats. After nitrogen step down, described in 3.1, 20 µL of the cell culture was plated on 60 mm TC standard dishes (Sarstedt) filled with solid 0.8% agarose BG110 media and covered with a cover glass. The plates were marked and the coordinates of the chosen filaments were noted. Like this, the same growing and differentiating filaments could be followed over time. The mutants were photographed every

24 hour until mature heterocysts were formed. The plates were stored in a moist box at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light overnight. For analyses of early promoter expression in liquid media the cultures were resuspended in 50 mL E-flask with BG110 media after nitrogen step down and held at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. 15 mL of cell culture was taken at time point zero hours, three hours, six hours and nine hours after nitrogen step down for cell harvest. Cell samples for confocal analysis were taken at the same time points.

4 Results

4.1 Electroporation and colony PCR

Electroporation was performed on the constructs named synthetic short -10TSS1, synthetic short -10TSS2, synthetic short -10N and synthetic short -10Ptrc, described in table 1. The remaining constructs were already transformed into *N. Punctiforme* strains. After electroporation the positive colonies were used as template for colony PCR and the amplified DNA was analysed, see supplementary 8.2. All constructs contained both genomic and plasmid DNA and were cultivated for further analyses.

4.2 Detection of heterocyst formation after nitrogen step down with confocal microscopy

Confocal microscopy was used to analyse the constructs after nitrogen step down. The different promoters were analysed by their expression of the *yfp* gene. Several analyses were performed for detection of promoter activity in different stages of cell differentiation. Subheadings 4.2.1 – 4.2.8 show promoter activity during heterocyst formation on plate, following the same filament. Subheadings 4.2.9 – 4.2.12 shows early promoter activity in filaments cultured in liquid media.

4.2.1 The native long construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours.

The native long promoter was expressed in all vegetative cells at the time point 0 hours after nitrogen step down (figure S5). No heterocysts were formed but an increase in YFP fluorescence appearing every tenth to twentieth cell was detected along the filament. 24 hours after nitrogen step down, YFP fluorescence was still detectable (figure S6). The native long promoter has an early expression in cells not showing any morphological changes. This can be concluded since a clear increase of YFP fluorescence is detected in some of the cells while the auto fluorescence caused by these cells has the same intensity as the rest of the filament. The DIC image do not indicate any heterocyst.

At the time point 48 hours after nitrogen step down, a heterocyst pattern is clearly seen (figure 10). Several notations can be done. The promoter expression is stronger in immature heterocysts compared to mature ones. This is indicated by the stronger YFP fluorescence in cells with increased auto fluorescence (early stage of immature heterocyst) and cells with bleaching auto fluorescence (late stage of immature heterocyst) compared to in cells with no auto fluorescence (mature heterocyst). Both vegetative cells and heterocysts express the promoter, seen as YFP fluorescence in all cells of the filament.

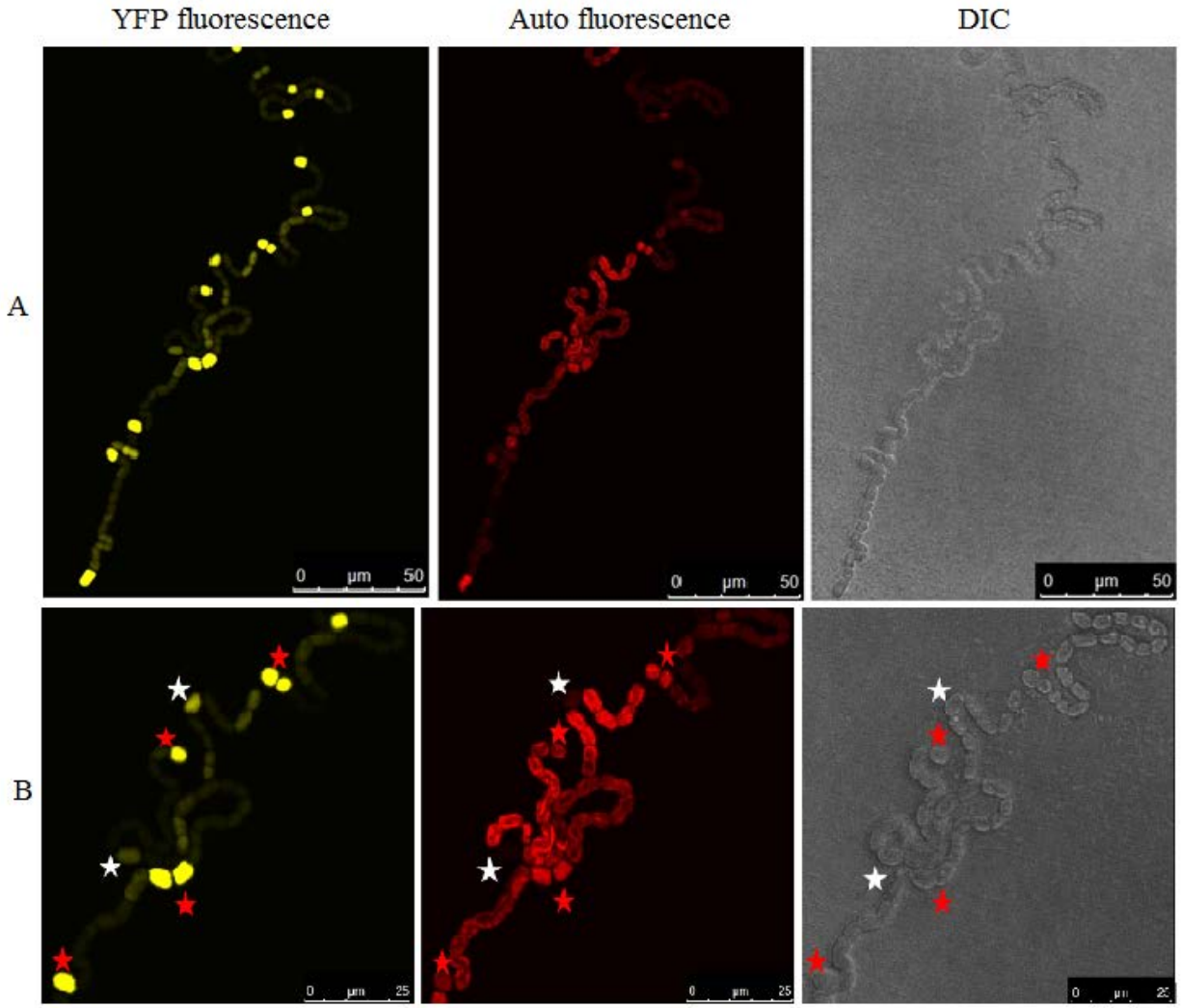

Figure 10. Confocal images of YFP fluorescence caused by the activity of the native long promoter, 48 hours after nitrogen step down. The native long promoter coupled to a YFP gene is expressed in both immature heterocysts, marked with red asterisks, and mature heterocysts, marked with white asterisks. A pattern with strong YFP fluorescence along the filament can be distinguish. This pattern corresponds with fluctuation in auto fluorescence and different morphology shown in the DIC image. The YFP fluorescence is stronger in immature heterocysts, compared to in mature heterocysts, indicating an early promoter expression. Panel A shows the whole filament and panel B are zoomed images. The left images show the

YFP fluorescence, the middle images show the auto fluorescence and the right images show the differential interference contrast (DIC) images of the same filament.

4.2.2 The native short construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours.

The native short promoter was weakly expressed in all vegetative cells of the filament at the time point 0 hours after nitrogen step down (figure S7). 24 hours after nitrogen step down, the vegetative cells still show some YFP fluorescence, either if the fluorescence is due to the promoter expression or the fluorescent protein still being stable in the cell. Immature heterocysts have appeared 24 hours after nitrogen step down and they express the promoter with a higher intensity compared to vegetative cells (figure S8).

The native short promoter is at some degree expressed in vegetative cells but not as strongly as the native long promoter is expressed in vegetative cells. It does not seem to be a difference in expression strength of the native short promoter in immature and mature heterocysts (figure 11).

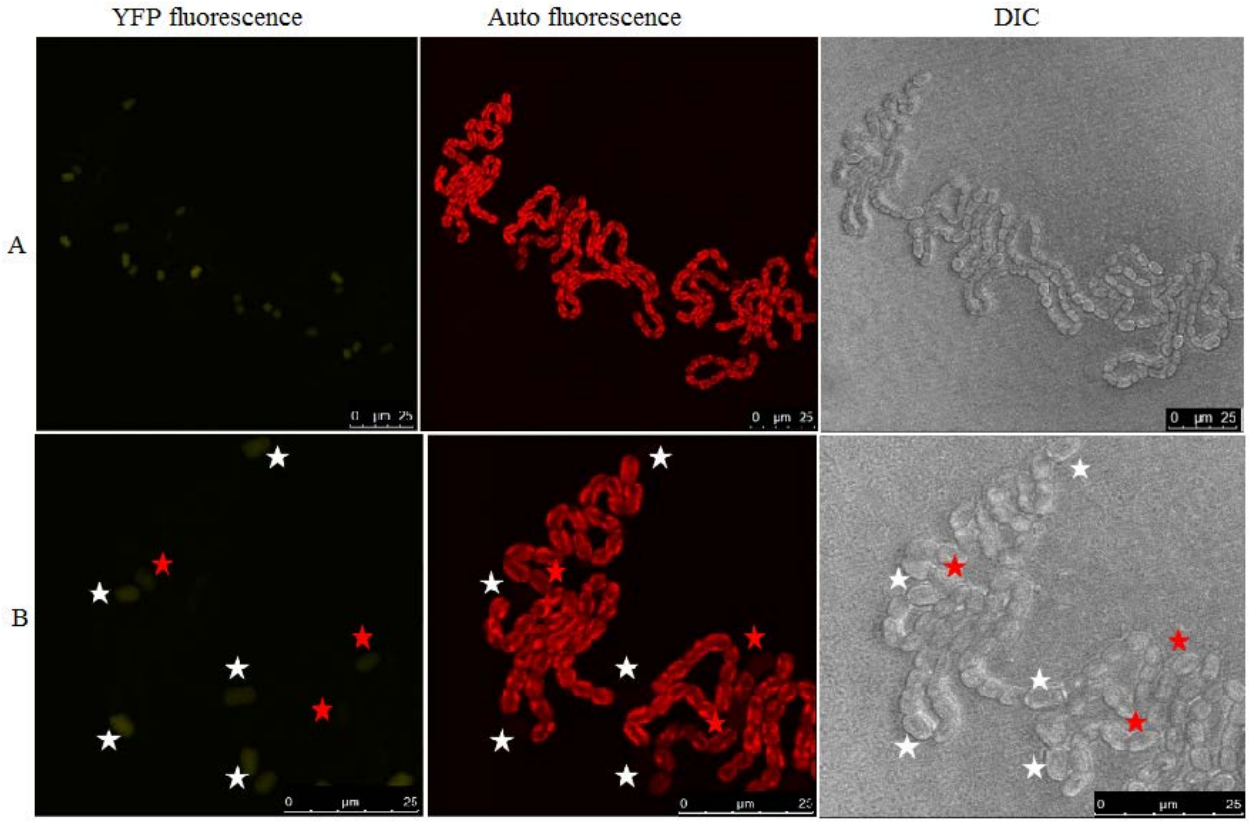

Figure 11. Confocal images of YFP fluorescence caused by the activity of the native short promoter 48 hours after nitrogen step down. The native short promoter coupled to a YFP gene is expressed in both immature heterocysts, marked with red asterisks, and mature heterocysts,

marked with white asterisks. Several heterocysts have appeared next to each other. A pattern with YFP fluorescence along the filament can be distinguish. This pattern corresponds with fluctuation in auto fluorescence and different morphology shown in the DIC image. It is noticeable that the strength in YFP fluorescence does not differs between the immature and mature heterocysts. Panel A shows the whole filament and panel B are zoomed images. The left images show the YFP fluorescence, the middle images show the auto fluorescence and the right images show the differential interference contrast (DIC) images of the same filament.

4.2.3 The synthetic long construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours.

At time point 0 hours after nitrogen, a vague YFP fluorescence forming a pattern similar to a heterocyst formation pattern can be seen (figure S9). The synthetic long promoter does not seems to be expressed in all cells. 24 hours after nitrogen step down, the stage of immature heterocysts can be seen (figure S10). The heterocyst formation pattern with YFP fluorescence in every tenth to twentieth cell is notable. In some cases do cells next to each other express YFP fluorescence, this within the pattern. This is probably because the promoter is expressed even before the immature heterocyst has stopped its division. The YFP protein is stable and therefore still detectable in the vegetative cells which sister cell is an immature heterocyst.

The synthetic long promoter has a heterocyst specific expression and is not expressed in vegetative cells. A difference in strength between the YFP fluorescence of the immature and mature heterocysts is noticeable (figure 12). The immature heterocyst, red asterisks, has a stronger expression compared to the mature heterocyst, white asterisks, indicates stronger promoter expression in the earlier stages of cell differentiation.

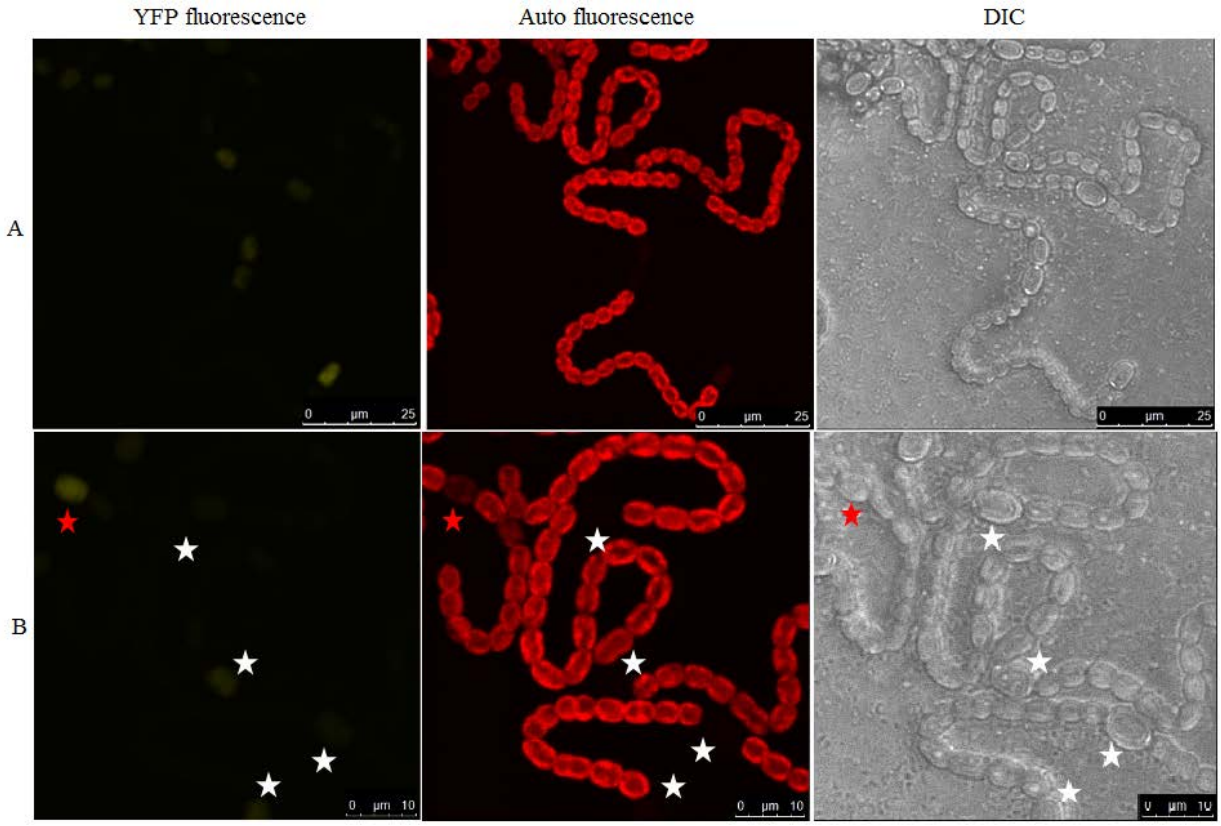

Figure 12. Confocal images of YFP fluorescence caused by the activity of the synthetic long promoter, 48 hours after nitrogen step down. The synthetic long promoter coupled to a YFP gene is weakly expressed in both immature heterocysts, marked with red asterisks, and mature heterocysts, marked with white asterisks. A pattern with YFP fluorescence along the filament can be distinguish. This pattern corresponds with fluctuation in auto fluorescence and different morphology shown in the DIC image. The YFP fluorescence is stronger in immature heterocyst compared to in mature heterocysts, showing the synthetic long promoter to be expressed early during cell differentiation. The synthetic long promoter is heterocyst specific. Panel A shows the whole filament and panel B are zoomed images. The left images show the YFP fluorescence, the middle images show the auto fluorescence and the right images show the differential interference contrast (DIC) images of the same filament.

4.2.4 The synthetic short construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours.

At time point 0 hours after nitrogen step down, a weak YFP fluorescence can be detected at different sites along the filament (figure S11). The auto fluorescence is slightly degraded. The synthetic short promoter does not seem to be expressed in all vegetative cells. At time point 24 hours after nitrogen step down, no change in auto fluorescence can be detected but YFP fluorescence can still be detected along the filament (figure S12).

The synthetic short promoter is heterocyst specific with expression only in the heterocysts (figure 13). Compare to the images of the synthetic long promoter where the expression in the immature heterocysts seemed to be stronger than the expression in the mature heterocysts, there is no significant difference between the expression of the synthetic short promoter in immature and mature heterocysts. The auto fluorescence of the filament is relatively weak at all time points (figures S11, S12 and 13).

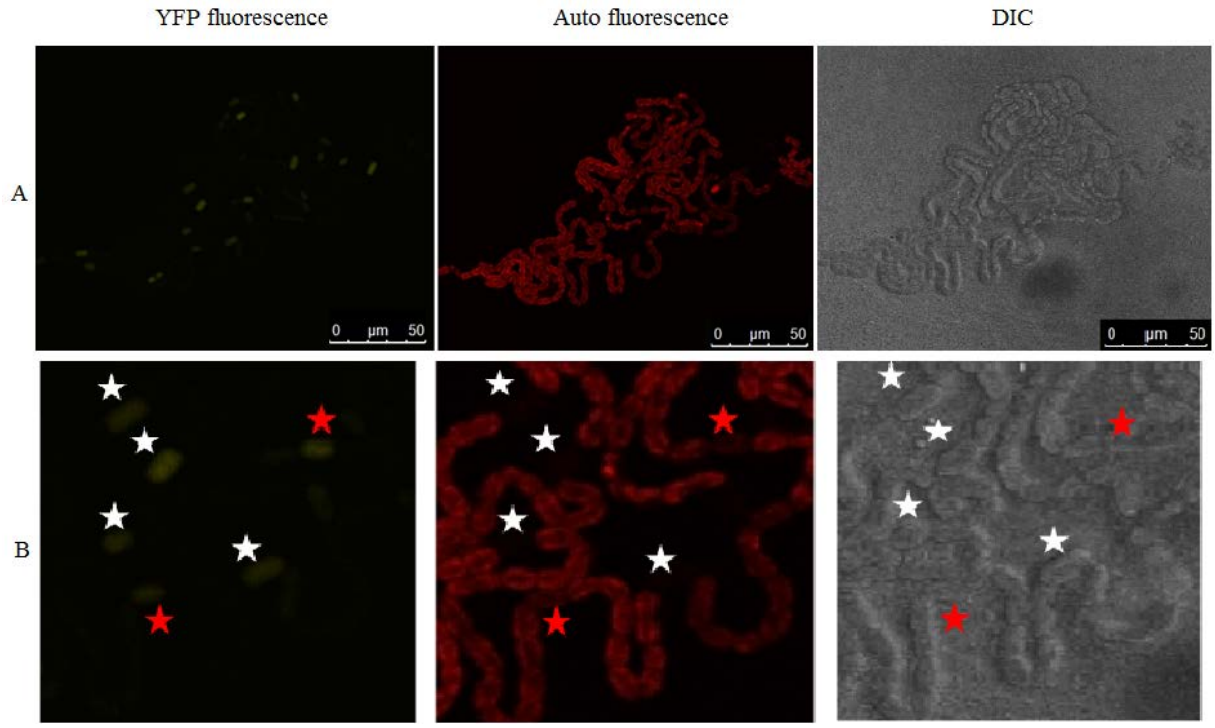

Figure 13. Confocal images of YFP fluorescence caused by the activity of the synthetic short promoter, 48 hours after nitrogen step down. The synthetic short promoter coupled to a YFP gene is weakly expressed in both immature heterocysts, marked with red asterisks, and mature heterocysts, marked with white asterisks. A pattern with YFP fluorescence along the filament can be distinguish. This pattern corresponds with fluctuation in auto fluorescence and different morphology shown in the DIC image. It is noticeable that the strength in YFP fluorescence does not differs between the immature and mature heterocysts. The synthetic short promoter is heterocyst specific. Panel A shows the whole filament and panel B are zoomed images. The left images show the YFP fluorescence, the middle images show the auto fluorescence and the right images show the differential interference contrast (DIC) images of the same filament.

4.2.5 The native short -10Ptrc construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours. Pictures from time point 0 hours and when mature heterocysts were visible are shown.

At time point 0 hours after nitrogen step down, all vegetative cells express the native short -10Ptrc promoter (figure 13). At the stage when mature heterocysts are visible, the promoter is still equally expressed in all cells and no pattern of increase in YFP fluorescence can be seen (figure 14).

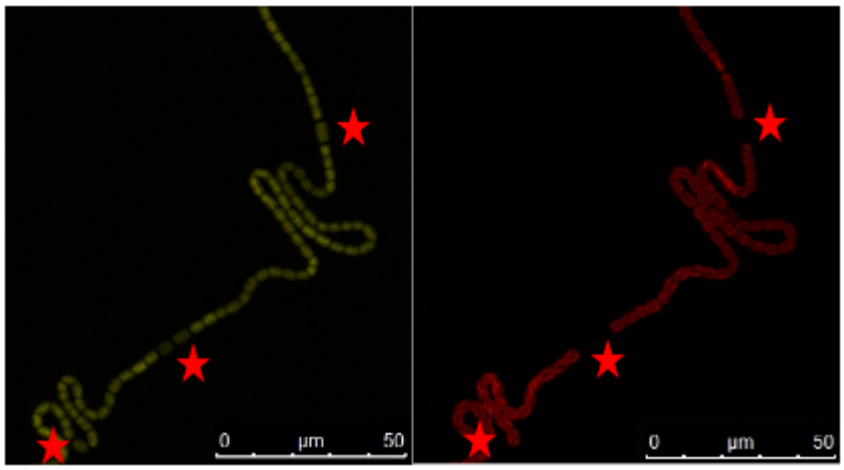

Figure 14. Confocal images of YFP fluorescence caused by the activity of the native short -10Ptrc promoter. Mature heterocyst are marked white asterisks. All cells of the filament express the native short -10Ptrc promoter indicating that the promoter is not heterocyst specific. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

4.2.6 The synthetic short -10TSS1 construct on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate and pictures of the same filament were taken every 24 hours.

The -10TSS1 promoter is not expressed in any cells along the filament, neither at time point 0, 24 nor 48 hours after nitrogen step down nor at time point (figure S14, S15 and 15).

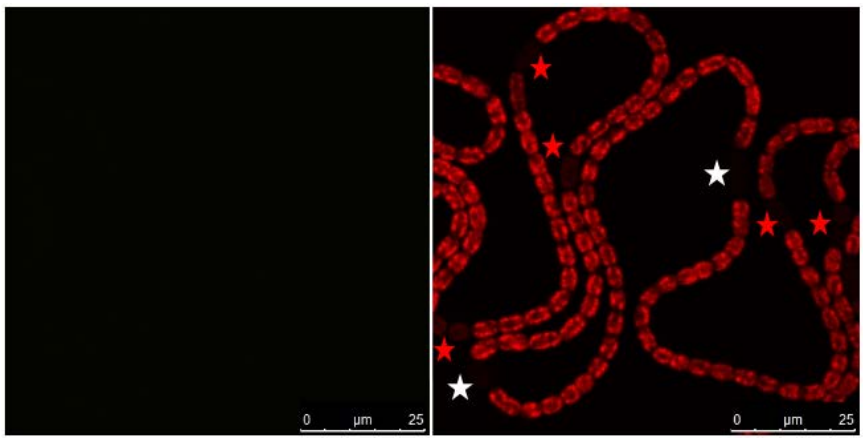

Figure 15. Confocal images of YFP fluorescence caused by the activity of the synhetic short -10TSS1 promoter, 48 hours after nitrogen step down. None of the cells along the filament expresses the synthetic short -10TSS1 promoter. The left image shows the YFP fluorescence and the right image shows the auto fluorescence. Immature heterocysts are marked with red asterisks and mature heterocysts are marked with white asterisks.

4.2.7 The synthetic short -10TSS2 on $BG11_0$ plate

After nitrogen step down, the cells were grown on a $BG11_0$ plate. Pictures of filament at time point 48 hours after nitrogen step down are shown.

The difference in YFP fluorescence among the three filaments in figure 16 indicates some stress related regulation of the synthetic short -10TSS2 promoter. Expression of the synthetic short -10TSS2 promoter was detectable in filaments with degraded auto fluorescence (fig 16B and C) while there was no promoter expression in filament showing a healthy thylakoid membrane (figure 16A).

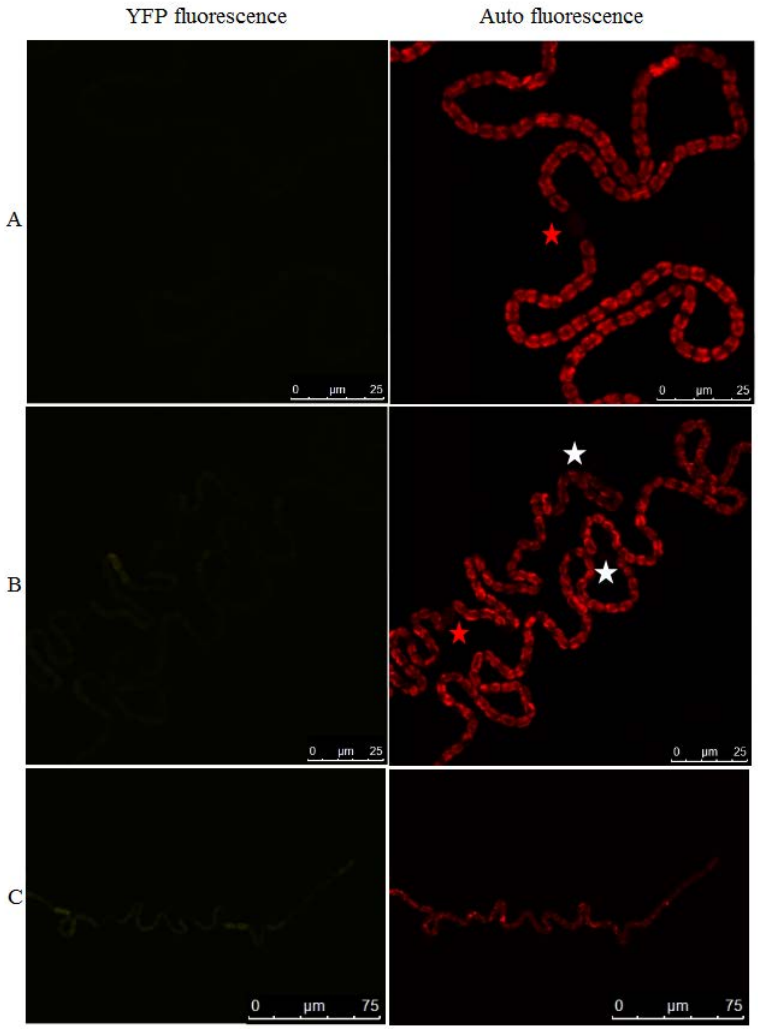

Figure 16. Confocal images of YFP fluorescence caused by the activity of the native short -10PTSS2 promoter, 48 hours after nitrogen step down. The filament shown in panel A do not express the synthetic short -10TSS2 promoter in neither heterocysts nor vegetative cells. The auto fluorescence, with its clear circular pattern, indicates a healthy filament. The filament shown in panel B consists of both immature and mature heterocysts. The auto fluorescence indicates some degradation of the thylakoid membrane and the synthetic short -10TSS2 promoter in weekly expressed in the majority of the cells. The auto fluorescence from the

filament in panel C does not form the circular pattern seen in filaments in panels A and B. The synthetic short -10TSS2 promoter is expressed in all cells of the filament. The left images shows YFP fluorescence and the right images shows auto fluorescence. Immature heterocysts are marked with red asterisks and mature heterocysts are marked with white asterisks.

4.2.8 The synthetic short -10N construct on BG11$_0$ plate

After nitrogen step down, the cells were grown on a BG11$_0$ plate and pictures of the same filament were taken every 24 hours.

At both time point 0 hours after nitrogen step down and 24 hours after nitrogen step down the synthetic short -10N promoter was expressed in all vegetative cells (figure S16 and S17).

When mature heterocysts are formed, the synthetic short -10N promoter is expressed in all cells of the filament (figure 17). Over the filament, it is possible to see some changes in YFP fluorescence but the fluctuation does not correspond with the heterocyst pattern seen in the auto fluorescence. It is therefore no correlation between YFP fluorescence and heterocyst formation. The YFP fluorescence is evenly spread within the cell, compared to the auto fluorescence that originates from the thylakoid membrane and not seen in middle of the cell (figure 17). Hence, the YFP fluorescence originated form the promoter activity and not the thylakoid membrane.

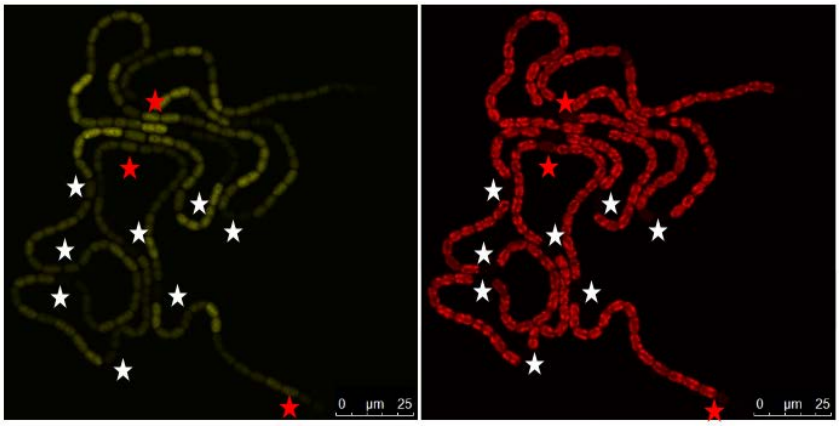

Figure 17. Confocal images of YFP fluorescence caused by the activity of the native short -10N promoter, 48 hours after nitrogen step down. The synthetic short -10N promoter is expressed in all cells of the filament. The left image shows the YFP fluorescence and the right image shows the auto fluorescence. Immature heterocysts are marked with red asterisks and mature heterocysts are marked with white asterisks.

4.2.9 The native long construct in liquid media

After nitrogen step down the cells were resuspended in BG110 and grown in a sterile 25 mL E-flask at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. Samples for confocal microscopy were taken every third hour.

The native long promoter is expressed in all cells of the filament, both in vegetative cells and in heterocysts (figure 18). This corresponds with earlier results, see figure 10. The same pattern with an increase in YFP fluorescence in every tenth to twentieth cell 0 hours after

nitrogen step down, as seen in figure 10, is noticeable in figure 18. There is no significant difference between the images taken 3 hours after nitrogen step down and 0 hours after nitrogen step down. At the time point 6 hours after nitrogen step down, an increase in YFP in certain cells along the filament can be detected. At the time point 9 hours after nitrogen step down, morphological change, seen as degradation of auto fluorescence, is visible (figure 19). Increase of YFP fluorescence can be seen along the filament. There is a difference in YFP strength between the time points 6 and 9 hours after nitrogen step down, the fluorescence is stronger 9 hours after nitrogen step down.

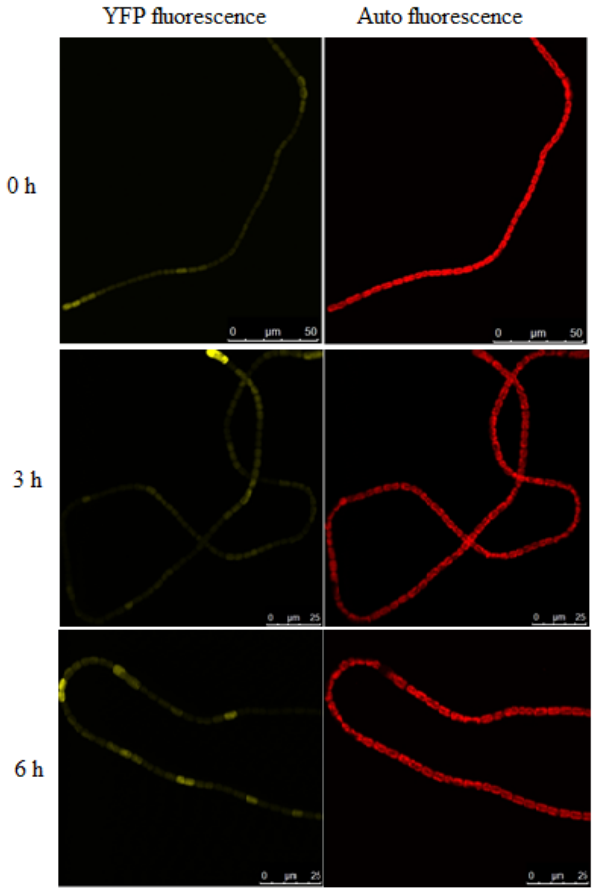

Figure 18. Confocal images of YFP fluorescence caused by the activity of the native long promoter. Cell samples from a culture living in media lacking combined nitrogen were taken every third hour for analyse of promoter expression in early stages of cell differentiation. The left images shows YFP fluorescence and the right images shows auto fluorescence. (0h) All vegetative cells of the filament express the synthetic long promoter. (3h) There is no significant difference between the images taken 3 hours after nitrogen step down and 0 hours after nitrogen step down. The promoter is till expressed in all cells of the filament. The left parts of the filament is not in focal plane and therefor seems to have a weaker fluorescence. (6h) A distinguish increase of YFP fluorescence in certain vegetative cells is detected. There is still no change in auto fluorescence and the increase of the YFP fluorescence appears in two cells next to each other along the filament. The promoter is still weakly expressed in all cells.

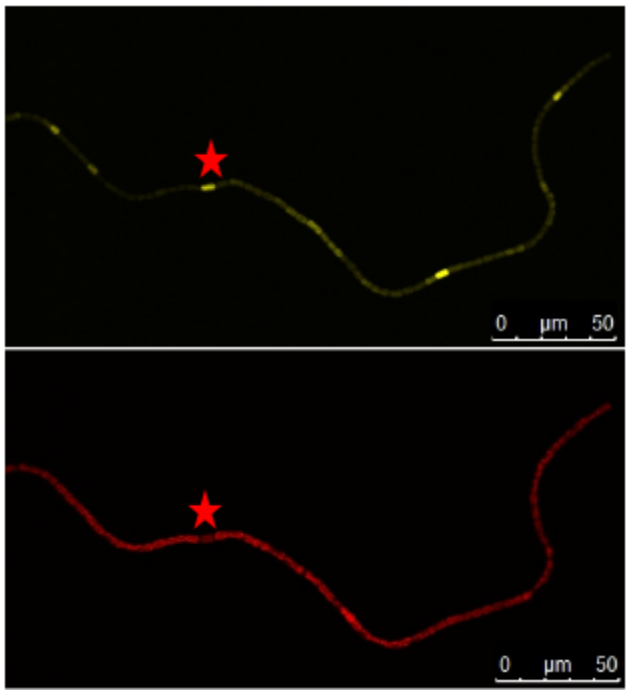

Figure 19. Confocal images of YFP fluorescence caused by the activity of the native long promoter. The culture had grown in media lacking combined nitrogen for 9 hours after nitrogen step down. Immature heterocyst with an increase in YFP fluorescence and a decrease in auto fluorescence are visible. The upper image shows the YFP fluorescence and the lower image shows the auto fluorescence. Immature heterocyst are market with red asterisks.

4.2.10 The synthetic long construct in liquid media

After nitrogen step down the cells were resuspended in $BG11_0$ and grown in a sterile 25 mL E-flask at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. Samples for confocal microscopy were taken every third hour.

At the time points 0 and 3 hours after nitrogen step down, some vegetative cells along the filament express the synthetic long promoter (figure 20). There is no significant difference between the time points 0 and 3 hours after nitrogen step down. At the time point 6 hours after nitrogen step down, single cells fluoresce in yellow. The appearance of YFP fluorescence corresponds well with a heterocyst pattern. The rest of the vegetative cells show no fluorescence at all. At the time point 9 hours after nitrogen step down, immature heterocysts with deceased auto fluorescence are visible (figure 21). These cells fluoresce in yellow with a stronger intensity compared to 6 hours after nitrogen step down. No YFP fluorescence from the vegetative cells can be detected.

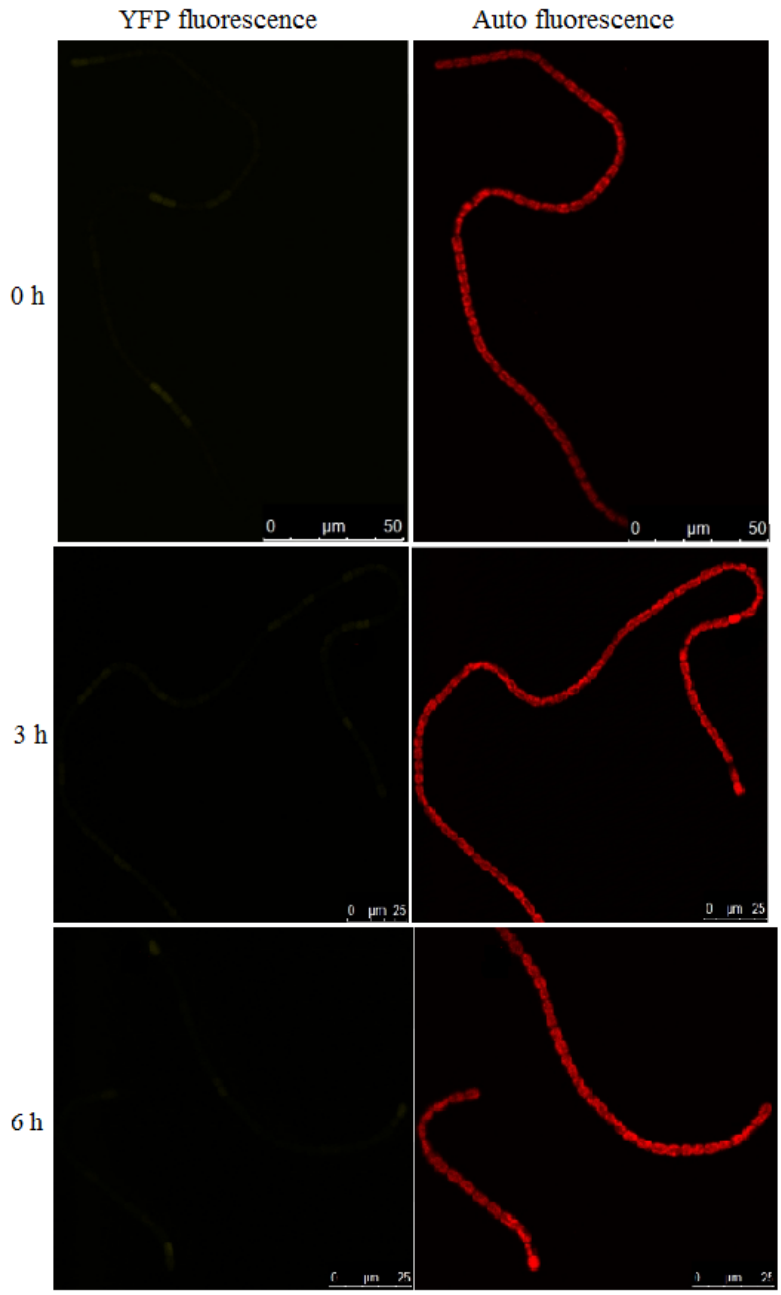

Figure 20. Confocal images of YFP fluorescence caused by the activity of the synthetic long promoter. Cell samples from a culture living in media lacking combined nitrogen were taken every third hour for analyse of promoter expression in early stages of cell differentiation. The left images shows YFP fluorescence and the right images shows auto fluorescence. (0h) Some detection of YFP fluorescence is seen along the filament. (3h) There is no significant difference between the images taken 3 hours after nitrogen step down and 0 hours after nitrogen step down. (6h) A weak YFP fluorescence is detectable in the ends of the filament

and in the middle of the right filament. YFP fluorescence in cells forming a heterocyst pattern and is not in remaining vegetative cells.

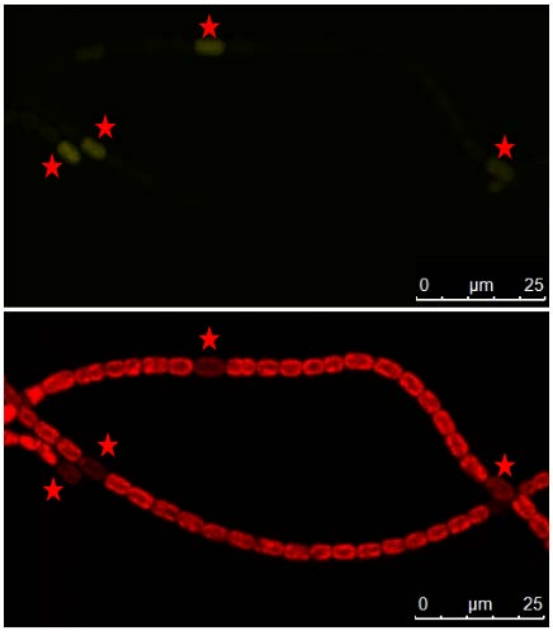

Figure 21. Confocal images of YFP fluorescence caused by the activity of the synthetic long promoter. The culture had grown in media lacking combined nitrogen for 9 hours after nitrogen step down. Immature heterocysts are notable in and marked with red asterisks. A clear decrease in auto fluorescence can be seen in the immature heterocyst as well as an increase in YFP fluorescence. The synthetic long promoter is only expressed in the heterocyst with no detectable signal in the vegetative cells.

4.2.11 The native short construct in liquid media

After nitrogen step down the cells were resuspended in $BG11_0$ and grown in a sterile 25 mL E-flask at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. Samples for confocal microscopy were taken every third hour.

At the time points 0 and 3 hours after nitrogen step down, the native short promoter is evenly and weakly expressed in all vegetative cells of the filament (figure 22). An increase in YFP fluorescence can be detected in certain cells of the filament at time point 6 hours after nitrogen step down. A distinguish increase in YFP fluorescence and a decrease in auto fluorescence can be seen in the immature heterocyst marked with red asterisks at time point 9 hours after nitrogen step down (figure 23). The vegetative cells also indicate promoter expression even if the difference in YFP fluorescence between the immature heterocyst and the vegetative cells is significant.

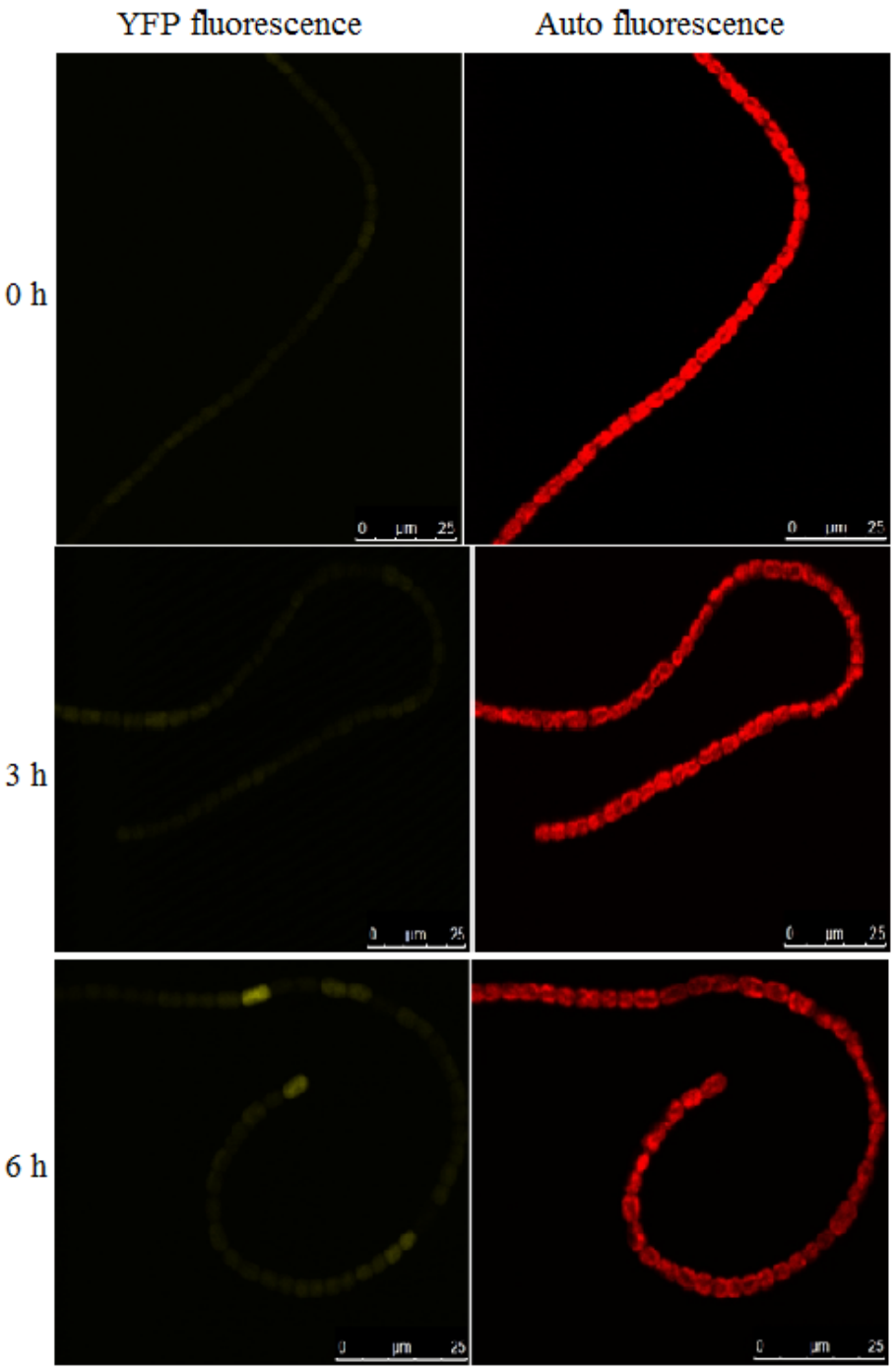

Figure 22. Confocal images of YFP fluorescence caused by the activity of the native short promoter. Cell samples from a culture living in media lacking combined nitrogen were taken every third hour for analyse of promoter expression in early stages of cell differentiation. The left images show YFP fluorescence and the right images shows auto fluorescence. (0h) All vegetative cells express the native short promoter to some degree. (3h) There is no significant difference between the images taken 3 hours after nitrogen step down and 0 hours after nitrogen step down. (6h) All vegetative cells still fluoresce in yellow and some cells along the filament show a weak increase in YFP fluorescence in a heterocyst like pattern.

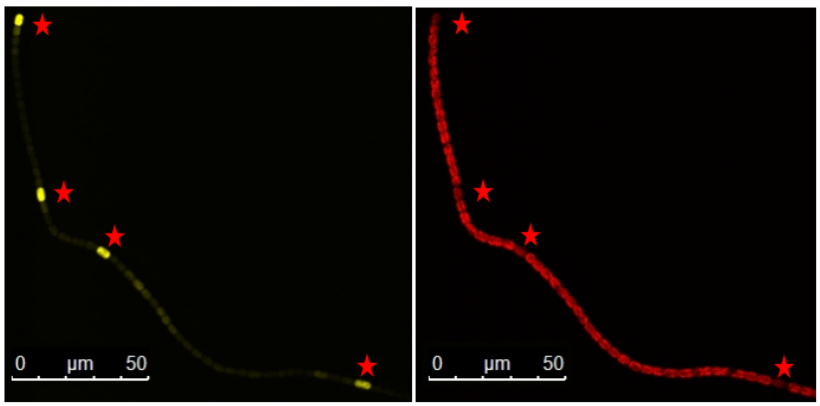

Figure 23. Confocal images of YFP fluorescence caused by the activity of the native short promoter. The culture had grown in media lacking combined nitrogen for 9 hours after nitrogen step down. Immature heterocysts are clearly seen along the filament. All cells express the native short promoter. The left image shows the YFP fluorescence and the right image shows the auto fluorescence. Immature heterocyst are marked with red asterisks.

4.2.12 The synthetic short construct in liquid media

After nitrogen step down the cells were resuspended in $BG11_0$ and grown in a sterile 25 mL E-flask at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. Samples for confocal microscopy were taken every third hour.

The synthetic short promoter is not expressed in any of the cells at the time point 0 hours after nitrogen step down (figure 24). There is no significant difference between the time points 3 and 0 hours after nitrogen step down. Some weak fluorescence can be seen 6 hours after nitrogen step down. At the time point 9 hours after nitrogen step down, no immature heterocysts are visible (figure 25). Due to the lack of immature heterocysts, the synthetic short promoter cannot fully be examined. No clear YFP fluorescence can be detected and the auto fluorescence do not differ between vegetative cells and heterocysts. The synthetic short promoter is not expressed before any morphological changes can be distinguished.

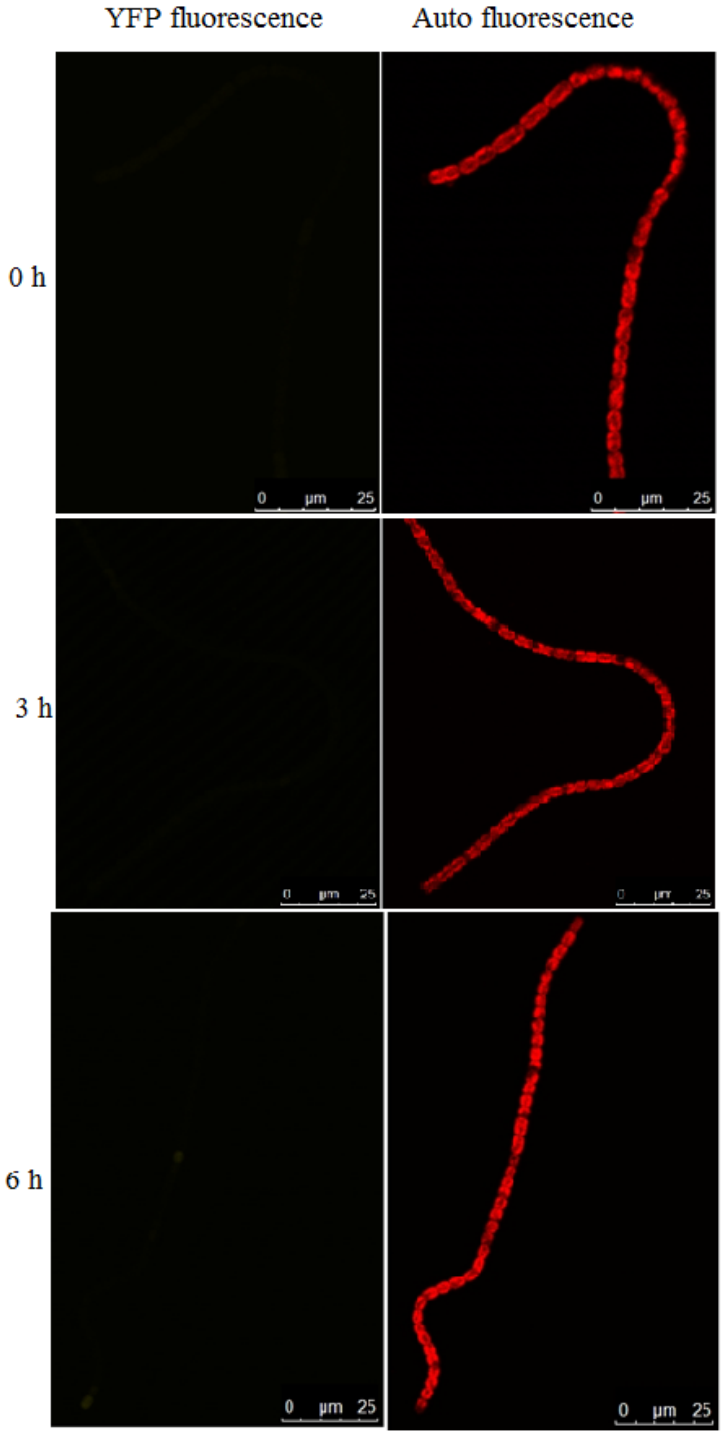

Figure 24. Confocal images of YFP fluorescence caused by the activity of the synthetic short promoter. Cell samples from a culture living in media lacking combined nitrogen were taken every third hour for analyse of promoter expression in early stages of cell differentiation. The left images show YFP fluorescence and the right images shows auto fluorescence. (0h) The synthetic short promoter is not expressed in any cells. (3h) There is no significant difference between the images taken 3 hours after nitrogen step down and 0 hours after nitrogen step down. (6h) Some weak YFP fluorescence can be detected.

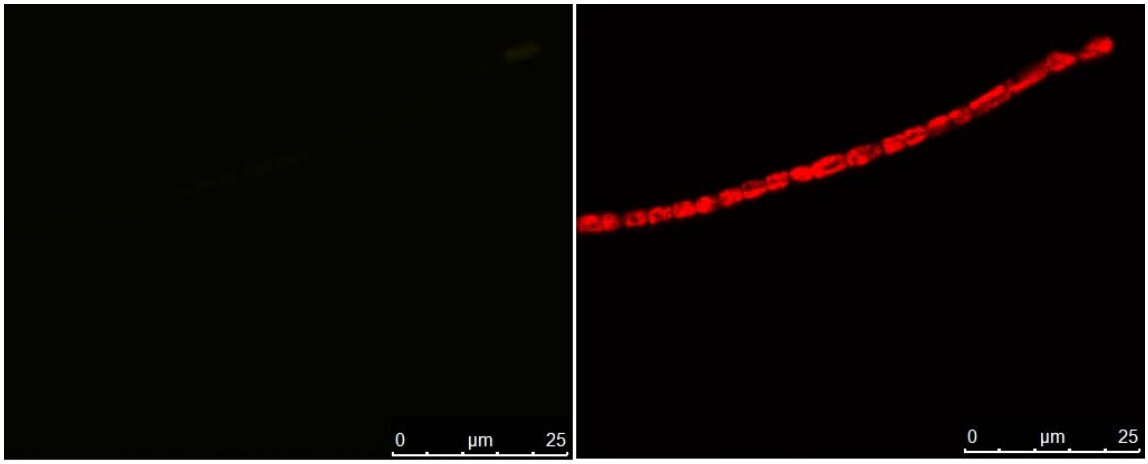

Figure 25. Confocal images of YFP fluorescence caused by the activity of the native short promoter. The culture had grown in media lacking combined nitrogen for 9 hours after nitrogen step down. A weak YFP fluorescence can be seen at the right end of the filament. No immature heterocysts are visible. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

4.3 Transcription level analysis by RT-qPCR

To complement the semi quantitative analysis method of confocal microscopy, RT-qPCR was applied. RT-qPCR provides results regarding transcription level, and therefore changes in gene expression between samples can be studied. Since the two promoters synthetic long and synthetic short have been proven to be heterocyst specific, it is of interest to examine the promoter activity on transcription level for these two promoters.

RT-qPCR was done on the samples imaged in sub heading 4.2.10 and 4.2.12. Nitrogen step down was performed on the two constructs synthetic long and synthetic short. After nitrogen step down the cells were resuspended in BG110 and grown in a sterile 25 mL E-flask at 30°C under 45 µmol photons s^{-1} m^{-2} continuous light on a shaker. Samples were taken at time points 0, 3, 6 and 9 hours after nitrogen step down. Three biological replicates were used. Total RNA was extracted and the quality of the RNA was analysed by gel electroporation, see S18 and S20. Further, possible DNA contaminations were evaluated by PCR on the RNA extracts. No contamination of DNA was identified, see S19 and S21.

4.3.1 Activity analysis of the synthetic long and the synthetic short promoters

The activity of the synthetic long promoter has been evaluated (figure 26). A significant increase in activity of the synthetic long promoter can be seen. Already 3 hours after nitrogen step down, an increase in level of the yfp transcript and hence a higher activity of the promoter is noticeable. The level of yfp transcript continues to increase during the first 9 hours after nitrogen step down (figurer 26). This indicates an early expression of the

promoter. The results presented in figure 26 correspond well with the confocal images taken of the same samples presented in figure 20 and 21. Since the transcript level of *yfp* does not reaches its peak during this time period, several time points should be included and the experiment redone.

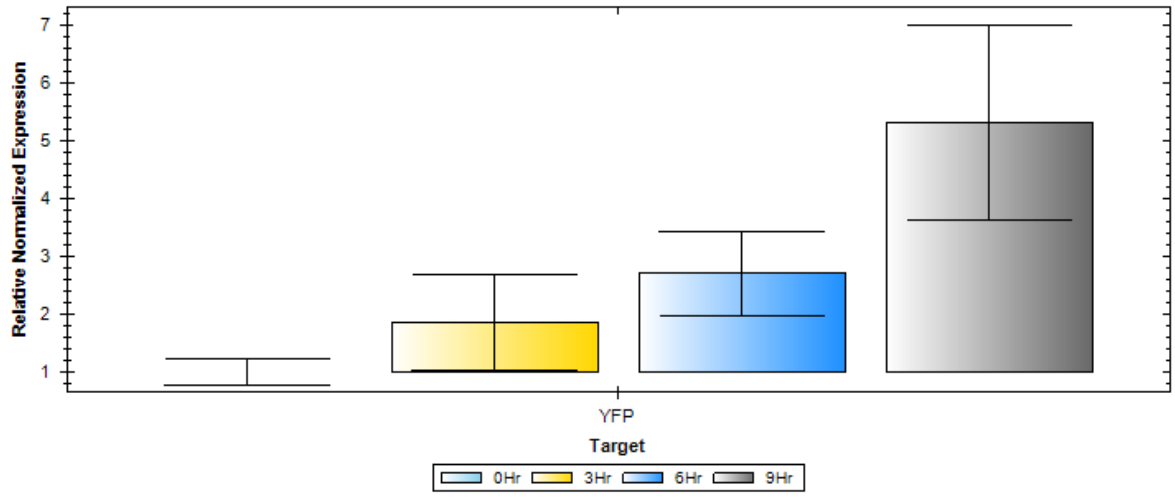

Figure 26. Analyse of early activity of the synthetic long promoter by measuring *yfp* transcription level at different time points with RT-qPCR. After nitrogen step down, the culture was gown in media lacking combined nitrogen. Samples were taken 0, 3, 6 and 9 hours after nitrogen down. A clear increase of the synthetic long promoter activity can be seen, indicating early activity of the promoter. Error bars is standard deviation.

The activity of the synthetic short promoter has been examined (figure 27 and 28). Normalized to the time point 0 hours after nitrogen step down, which has an expression value of 1, the expression of *yfp* varies between 1,22 and 1,42 during the first 9 hours after nitrogen step down (figure 27). The synthetic short promoter is not significantly activated during the first 9 hours after nitrogen step down. These results correspond well with the results presented in figures 24 and 25. In figures 24 and 25, where the activity of the synthetic short promoter at time points 0, 3, 6 and 9 hours after nitrogen step down was studied, almost no increase in promoter activity can be seen. 12 hours after nitrogen step down, the activity of the synthetic short promoter has increased significantly (figure 28). The result presented in figure 28 are not derived from the same experimental setting as the results presented in figure 27. Due to this, the expression values are not directly comparable. Despite this, the activity tendency of the synthetic short promoter can be analysed. The synthetic short promoter is expressed at a later stage of the cell differentiation, compared to the synthetic long promoter. To be able to compare the strength of the two promoters, the activity must be followed until the cultures show mature heterocysts. To be comparable, the samples from the same time points for the two promoters must be analysed in the same run of RT-qPCR.

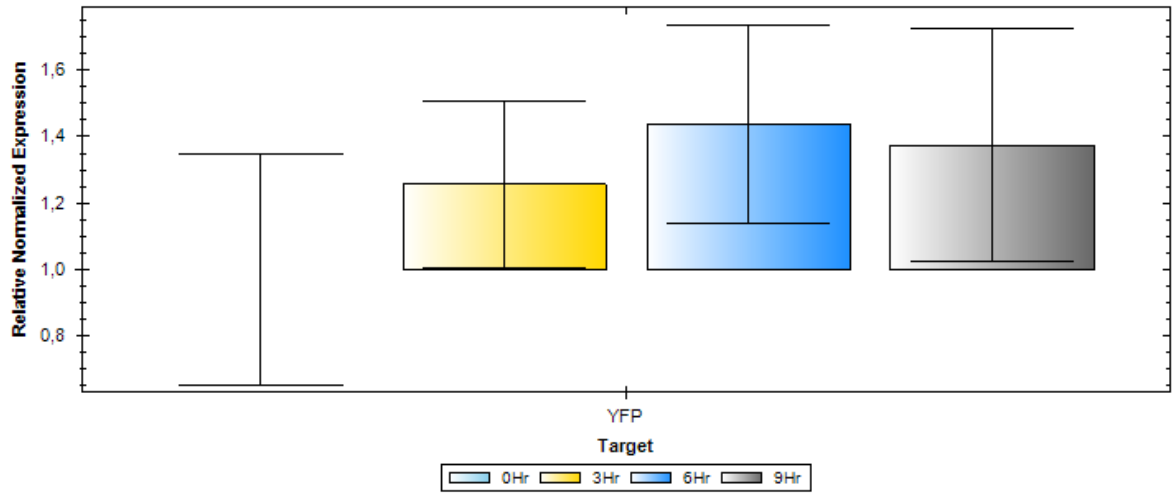

Figure 27. Analyse of early activity of the synthetic short promoter by measuring *yfp* transcription level at different time points with RT-qPCR. After nitrogen step down, the culture was gown in media lacking combined nitrogen. Samples were taken 0, 3, 6 and 9 hours after nitrogen down. No significant increase of promoter activity can be seen. The three biological replicates vary (table S2), resulting in big error bars. Error bars is standard deviation.

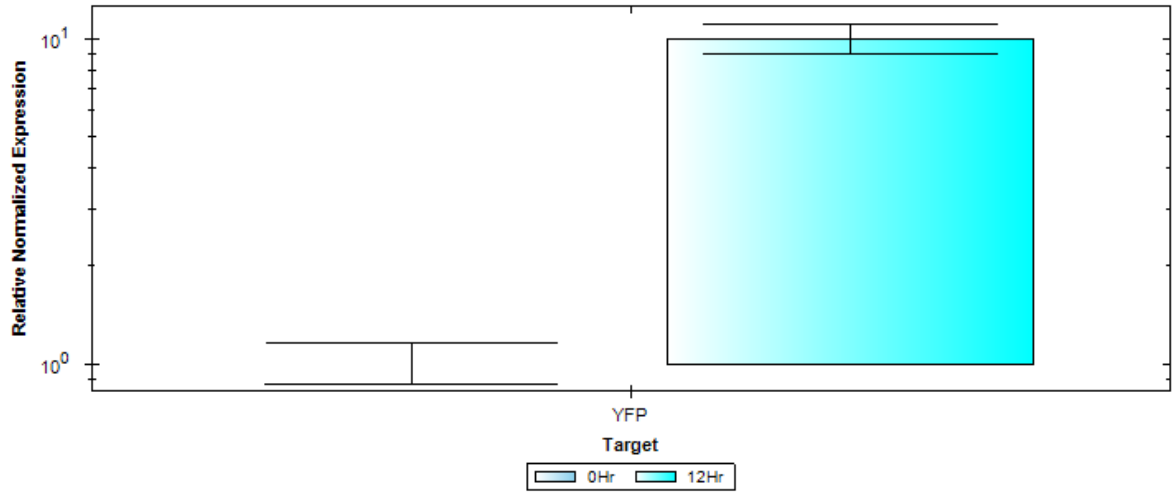

Figure 28. Analyse of early activity of the synthetic short promoter by measuring *yfp* transcription level at different time points with RT-qPCR. After nitrogen step down, the culture was gown in media lacking combined nitrogen. Samples were taken 0 and 12 hours after nitrogen down. A significant difference in expression of *yfp* between the time points is noticeable. Error bars is standard deviation.

5 Discussion

All promoter constructs consist of either three or four promoter regions, as described in table 1. The long promoter sequences consist of four parts; the upstream region, the -35 region, the spacer sequence and the -10 region. The short promoter sequences consist of three parts; the -35 region, the spacer sequence and the -10 region. By comparing the results of this study, I can draw conclusions regarding the different promoter parts. It is desirable of a synthetic promoter to only be induced by the wanted inducer, in this case the cell differentiation. Therefore, it should not interact with any other native regulator system. This is one of the main criteria when discussing the results. Other wanted characteristics are strength and consistency, the promoter should be similarly activated at all time. By using our experimental strategy, I was able to identify the importance of the different promoter regions of our synthetic promoters.

Since the long promoter constructs, the synthetic long and the native long containing the upstream region, gave earlier expression compared to the same constructs without the upstream region, the synthetic short and the native short, I conclude that the upstream region has an impact on the expression of the promoter. The upstream region might contribute to stronger binding of the RNA polymerase. This can be seen in the results, in the figures following the cultures growing in liquid media and in the graphs presenting the qPCR results. The expression of the synthetic long promoter has a notable increase during the first nine hours after nitrogen step down while the expression of the synthetic short promoter does not change significantly during these nine hours. This also corresponds with the images of the filament.

The overall interpretation of the results showing the expression of native long, native short, synthetic long and the synthetic short promoter is that the long constructs have a stronger expression compared to the short ones. The results from the images are only semi quantitative and there is no qPCR data from later time points for all constructs, therefore the conclusion is not fully supported. This especially since the qPCR data for the synthetic short promoter comparing time point zero hours and twelve hours after nitrogen step down showed a tenfold increase in promoter expression. To be able to draw any conclusions regarding the increase in promoter strength provided by the upstream region, further analyses needs to be carried out. However, the upstream region contributes to an earlier expression of the promoter.

The DIF motif present in all four constructs, native long, native short, synthetic long and synthetic short, was found by comparison of transcripts from wild type and a ΔhetR mutant after nitrogen step down. It was concluded that the DIF motif is dependent of HetR (Mitschke et al. 2011). In the same study, it was found that promoters dependent of NtcA, and directly regulated by NtcA, contain an NtcA binding site at the position -41.5, relative to the TSS (Mitschke et al. 2011). Since HetR and NtcA are dependent of each other and the DIF motif is dependent of HetR. The DIF motif might also in some way be dependent of NtcA and thereby

regulated by it. If so, it could explain way the upstream region, which originally is expressed together with the DIF motif, results in earlier expression and potential enhancement of the promoter.

Another perspective would be a potential upstream open reading frame (uORF) encoded in the upstream region of the long promoter constructs. It has been shown that several near-cognate uORFs that do not start with an AUG codon can regulate the translation of the downstream main ORF. The occupancy of the uORF by the ribosome enhances the translation of the main ORF. Several mRNAs needed for sporulation in yeast contains these uORF (Xue & Barna 2012).

The synthetic short -35Ptrc promoter did not show any expression in any of the filaments. It is possible that the construct does not function, that some mutation has occurred or that the fusion of a strong *E. coli* promoter does not cooperate with the other parts of the promoter. Since the mutated cell line has been grown in the presence of antibiotics at all time a loss of plasmid seems unlikely. Looking at the promoter expression from the other mutated cell lines, where they all express some level of YFP, a total lack of YFP expression derived from the synthetic short -35Ptrc promoter also seems unlikely. This mutant, and the role of the -35 region, needs to be further investigated.

The Ptrc promoter system has successfully been inserted in several cyanobacteria (Huang et al. 2010 and Ma et al. 2014) which implies that the synthetic short -35Ptrc promoter should function in *N. Punctiforme* as well. On the contrary, in previous studies the promoter system consisted of both the -35 and the -10 region and the system has not been tested in *N. Punctiforme*. As mentioned earlier, this mutant and the region itself needs to be further investigated.

Concerning the spacer sequence between the -35 region and the -10 region, this sequence seems to have some impact on the expression. The only difference between the native short promoter and the synthetic short -10N promoter is the spacer sequence. The native short promoter contains the native spacer and the synthetic short -10N contains the scrambled spacer. By comparing these two promoters, one can see that the native short promoter has an increased expression in heterocyst compared to vegetative cells while the synthetic short -10N is equally expressed in all cells of the filament.

The spacing between the -35 and the -10 region has been proven to have an impact on the promoter strength. The tac promoter of *E. coli* has a spacing of 16 base pairs between the -35 and -10 region. By inserting either one or two base pairs into the spacing sequence, the expanded promoters are named *trc* promoter and *tic* promoter, the activity of the promoter decreased to 90 respectively 65 percentage of the activity of the *tac* promoter (Brosius et al. 1985). In *Lactococcus lactis*, where the consensus sequences were kept intact and the spacing sequence was randomized, a 400-fold change in activity of the promoter was obtained (Jensen & Hammer 1998).

There is a difference in expression between the native and the synthetic promoters. It is noticeable that the native promoters have a stronger expression. The native promoters are also in some extension, expressed earlier. They are on the other hand, with different strength, expressed in all cells whilst the synthetic promoters are heterocyst specific. These results advocate the -10 region to be of importance for heterocyst specificity, which is one of the desired characterisations of the promoter we aim for. Additionally, the comparison of the synthetic short promoter and the synthetic short -10N promoter reinforces the conclusion regarding the -10 region. Whilst the synthetic short promoter is relative weak but heterocyst specific, the synthetic short -10N is expressed in all cells with no significant stronger expression in heterocysts. The only difference between these two promoters is the -10 region, the synthetic short -10N promoter contains the native -10 region and the synthetic short promoter contains the consensus -10 region. Hence, I conclude that the -10 region is of importance regarding the heterocyst specificity of the promoter.

The mutated cell line named synthetic short -10TSS1 did not express YFP in any of the cells. Either the -10 region does not function together with the DIF motif or if some changes has occurred in the sequence preventing the expression.

The fluorescence from the synthetic short -10TSS2 promoter were very diverse, and therefore no conclusions regarding the promoter expression could be drawn. The promoter activity did seem to be higher in filaments with decrease in auto fluorescence, indicating the synthetic short -10TSS2 promoter to be stress induced. Therefore, it is not possible to evaluate this promoter construct and further analysis needs to be implemented.

Change of the -10 region has earlier been proven to have an effect on the promoter strength (Haung and Lindblad 2013). The R40 promoter is a strong promoter in *E. coli* but does not has the same strength in *Synechocystis* sp. strain PCC 6803. A study performed in *Synechocystis* sp. strain PCC 6803 showed that when the -10 region in the R40 promoter was replaced with a consensus -10 sequence of *Synechocystis*, and transferred into Synechocystis, the promoter strength increase significant (Haung and Lindblad 2013). These results also emphasize the importance of the -10 region of the promoter.

6 Conclusions

The aim with this study was to design a synthetic promoter that can be used as a genetic tool for research in the cyanobacteria *N. punctiforme*. The promoter should be well characterized and heterocyst specific. Among the suggested promoters, the two promoters synthetic long and synthetic short were shown to be heterocyst specific. The consensus -10 region was proven to be of importance regarding the heterocyst specificity. The other promoter regions of these two promoters, the upstream region, the -35 region and the spacer region, have been evaluated but to fully understand the regulation, further investigation needs to be done.

Between the heterocyst specific promoters, the synthetic long promoter has an earlier expression and to some degree, a stronger expression. The degree of expression has only been evaluated with a semi quantitative method and should therefore be further examined. With this genetic tool it is possible to develop the production of biohydrogen in cyanobacteria.

7 Acknowledgement

Firstly, I would like to thank my supervisor Karin Stensjö for giving me the opportunity to work in her science group with this project. I thank you for many rewarding discussions and for always giving me the freedom to explore my ideas. Thank you Federico Turco for introducing me to the project and thanks to Xin Li for providing me with useful knowledge regarding the project. A special thanks to Adam Wegelius for guiding me in the lab and always answering my questions. Thank you Inés Varela for being my opponent and moral support in the lab. I would also like to thank Ann Magnusson for being my subject reviewer, thank you for useful feedback. Last but not least, thanks to all members of the cyano group! Thank you for the warm and open working environment, for all the help you have given me and all the laughs.

8 Supplementary

8.1 Plasmid

As seen in figure S1 do the plasmid contain all essential parts. Replication ORI for both E.coli and *N. punctiforme*, kanamycin resistance, a ccd operon as a control and the YFP gene.

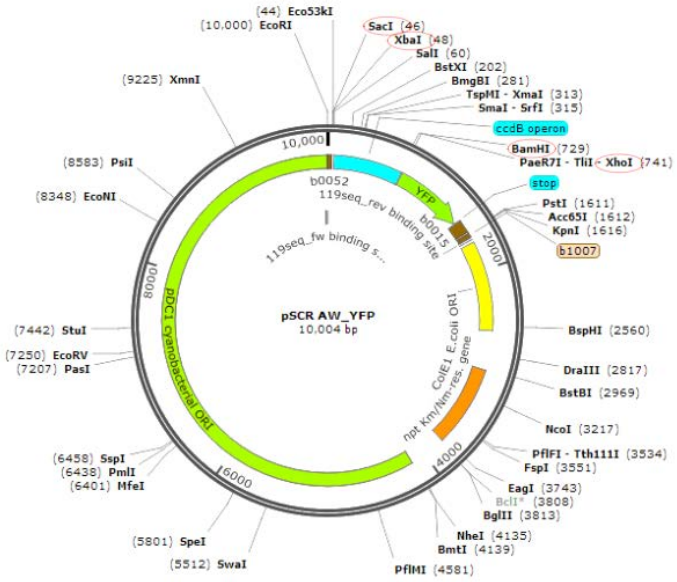

Figure S1. A schematic figure of the pSCR plasmid used in this study. The plasmid is designed by Adam Wegelius and contains replication ORI for both *E.coli* and *N. punctiforme*, kanamycin resistance, a ccd operon as a control and the YFP gene.

8.2 Analyse of amplified DNA from colony PCR

Detection of genomic DNA from constructs synthetic short -10TSS1 and synthetic short -10Ptrc is shown in figure S2 and detection of plasmid DNA from the same constructs are shown in figure S3. Detection of both genomic and plasmid DNA from constructs synthetic short -10N and synthetic short -10TSS2 is shown in figure S2. As seen in figure S2 do all colonies contain genomic DNA. As seen in figure S3 three colonies containing the -10TSS1 construct are positive, well 1-3. All colonies containing the -10Ptrc construct are positive. The colonies containing both genomic and plasmid DNA were cultivated for further analyses. As seen in figure S4 all colonies of the two constructs contained both genomic and plasmid DNA and were therefor cultivated for further analyses.

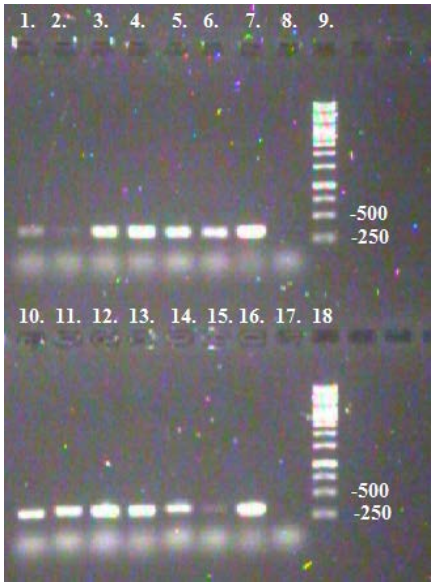

Figure S2. Detection of genomic DNA by gel electrophoresis. Well 1-6 corresponds to six colonies containing the -10TSS1 construct. Well 10-15 corresponds to six colonies containing the construct -10Ptrc. Well 7 and 16 shows a positive control and well 8 and 17 shows a negative control. Well 9 and 18 contains a 1 kb DNA ladder with the two smallest bands indicated.

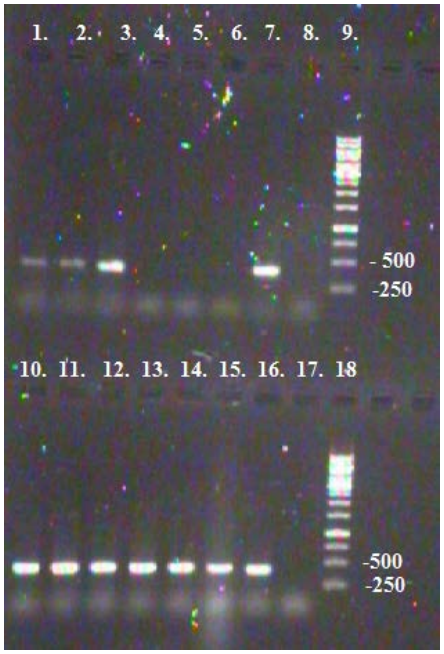

Figure S3. Detection of plasmid DNA by gel electrophoresis. Well 1-6 are the same colonies as in figure S2, corresponding to six colonies containing the -10TSS1 construct. Well 10-15 are the same colonies as in figure S2, corresponds to six colonies containing the construct -10Ptrc. Well 7 and 16 shows a positive control and well 8 and 17 shows a negative control. Well 9 and 18 contains a 1 kb DNA ladder with the two smallest bands indicated.

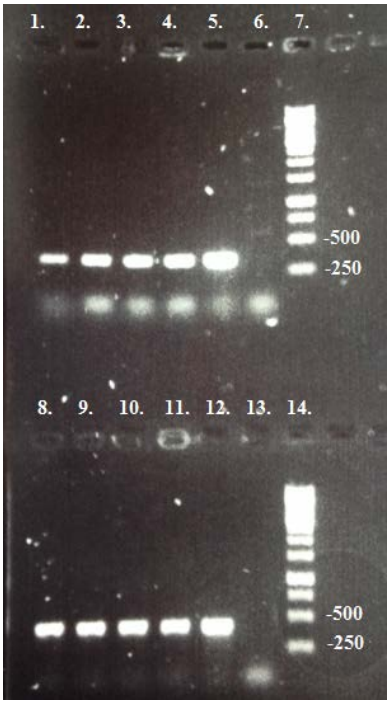

Figure S4. Detection of both genomic and plasmid DNA by gel electrophoresis. The upper row, well 1-6, shows detection of genomic DNA and the lower row, well 8-13, shows detection of plasmid DNA. Well 1-3 and 8-10 corresponds to three colonies containing the -10N construct and well 4 and 11 corresponds to a colony containing the -10TSS2 construct. Well 5 and 12 shows a positive control and well 6 and 13 shows a negative control. Well 7 and 14 contains a 1 kb DNA ladder with the two smallest bands indicated.

8.3 Confocal pictures

Pictures of all time points presented in the result.

8.3.1 Native long

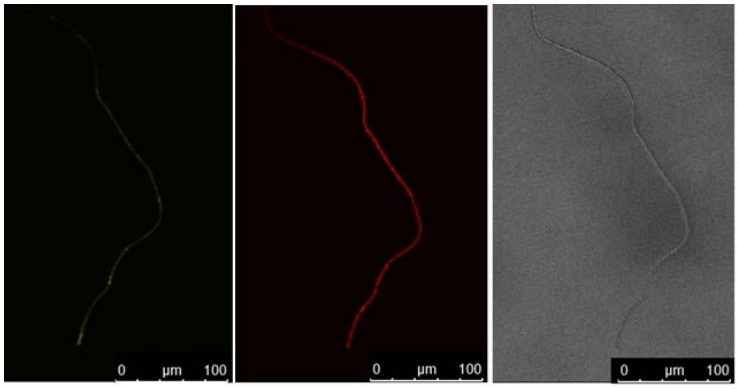

Figure S5. The native long construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

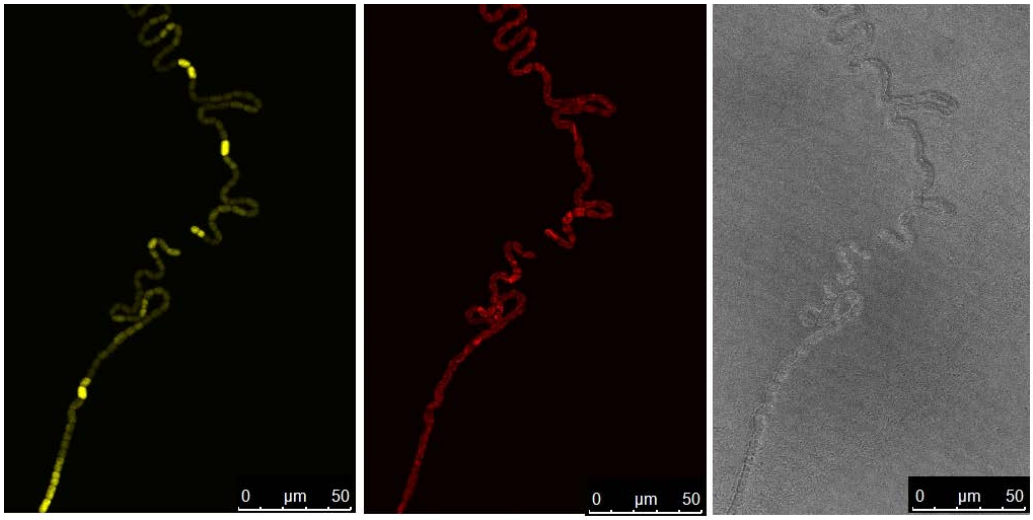

Figure S6. The native long construct 24 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

8.3.2 Native short

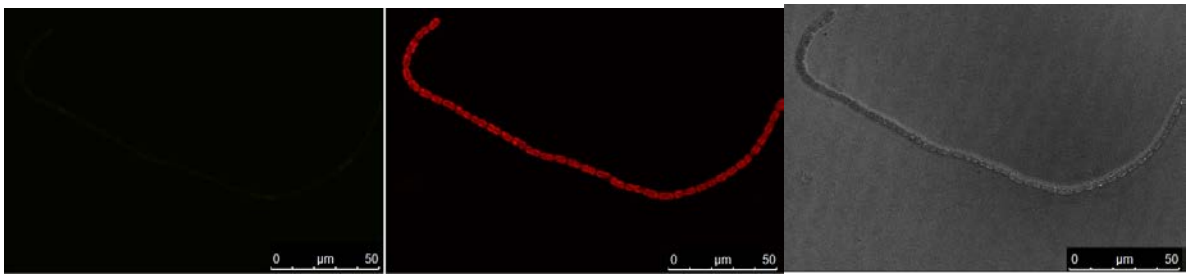

Figure S7. The native short construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

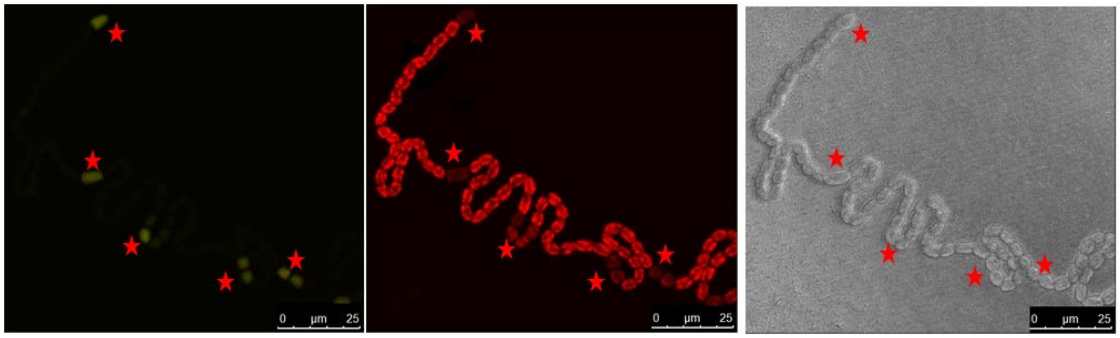

Figure S8. The native short construct 24 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament. Immature heterocysts are marked with white heterocysts.

8.3.3 Synthetic long

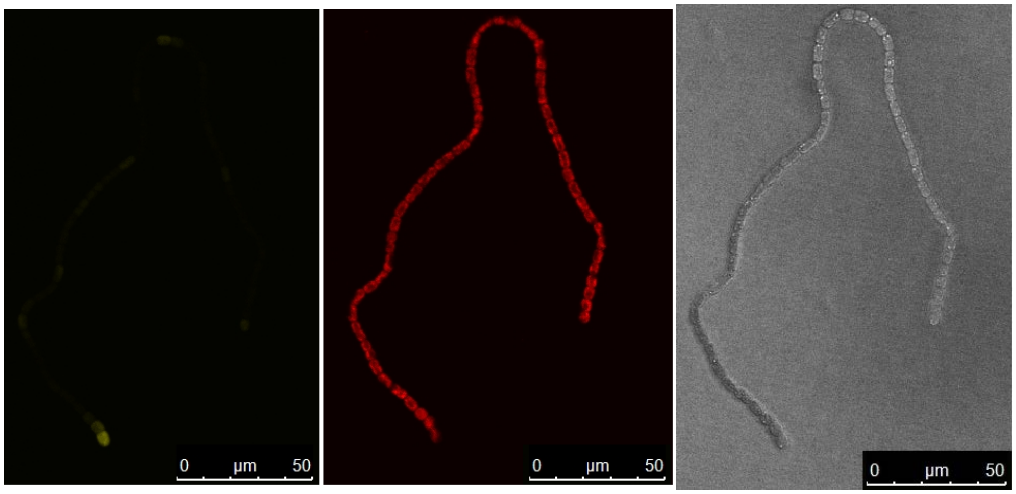

Figure S9. The synthetic long construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

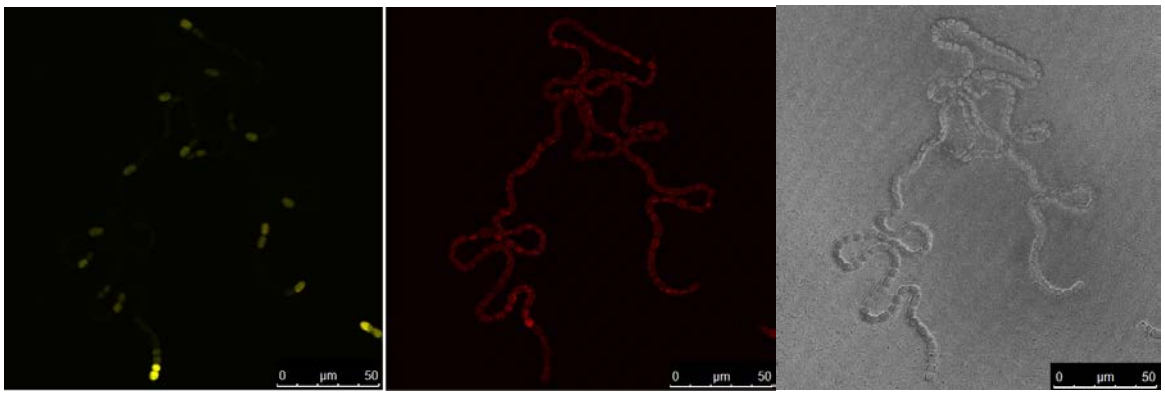

Figure S10. The synthetic long 24 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

8.3.4 Synthetic short

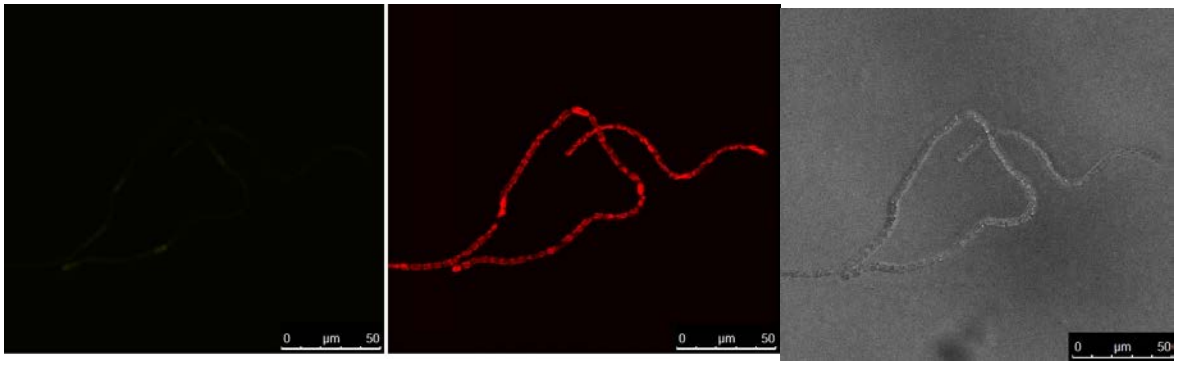

Figure S11. The synthetic short construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

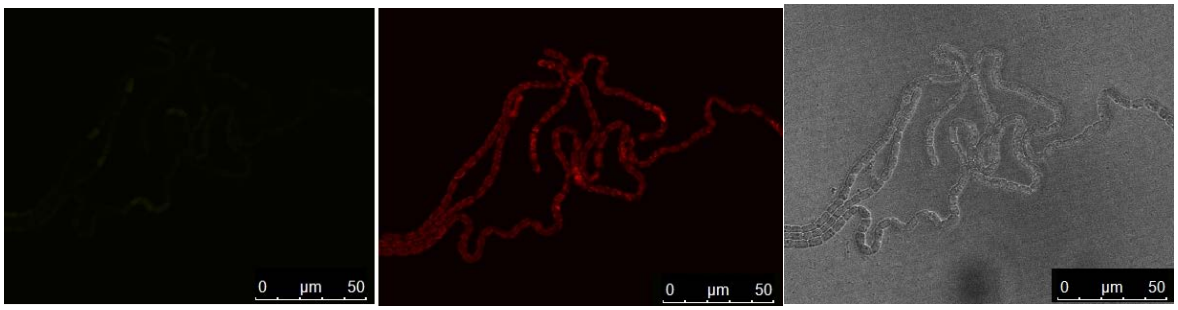

Figure S12. The synthetic short construct 24 hours after nitrogen step down. The left image shows the YFP fluorescence, the middle image shows the auto fluorescence and the right image shows the differential interference contrast (DIC) image of the same filament.

8.3.5 Native short -10Ptrc

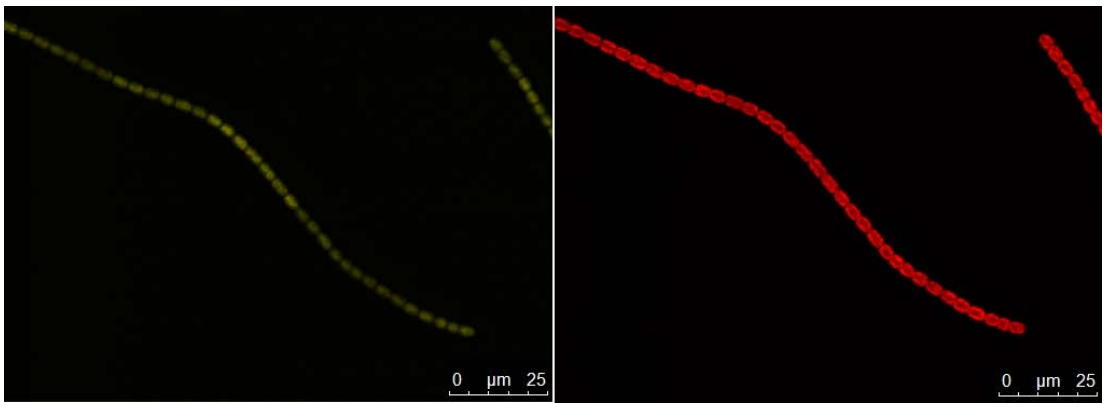

Figure S13. The native short -10Ptrc construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

8.3.6 Synthetic short -10TSS1

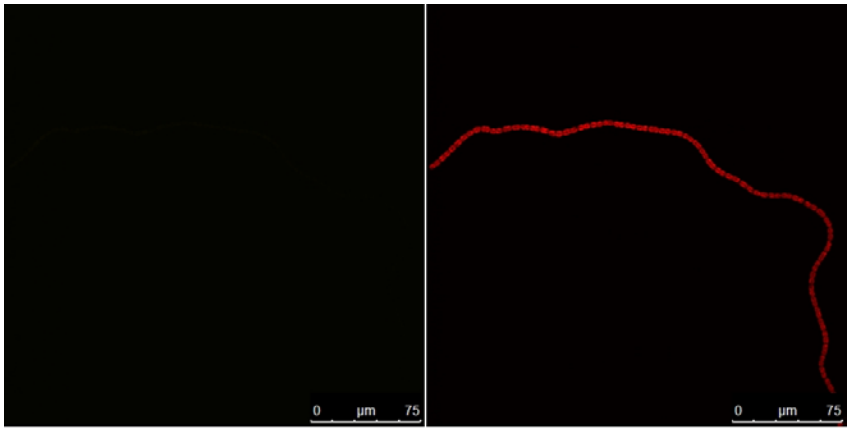

Figure S14. The synthetic short -10TSS1 construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

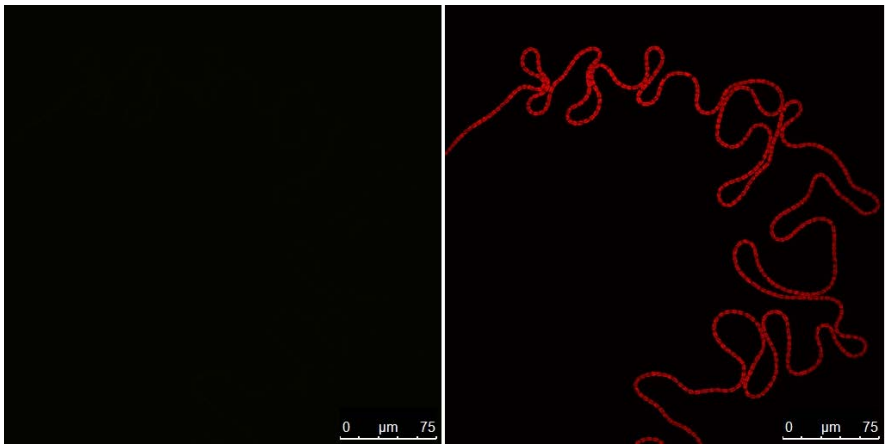

Figure S15. The synthetic short -10TSS1 construct 24 hours after nitrogen step down. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

8.3.7 Synthetic short -10N

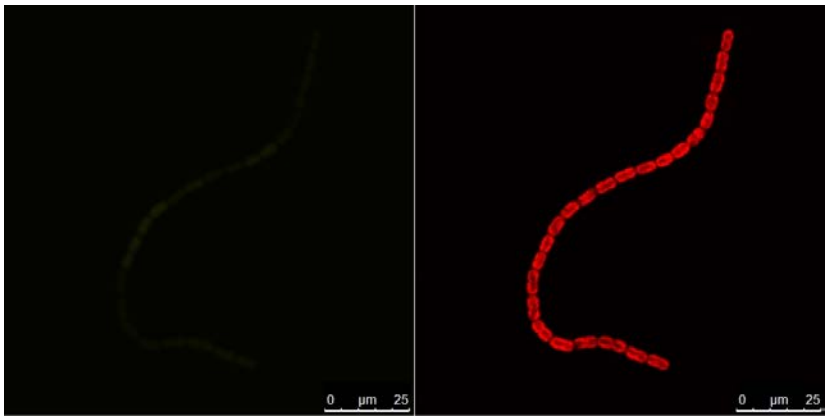

Figure S16. The synthetic short -10N construct 0 hours after nitrogen step down. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

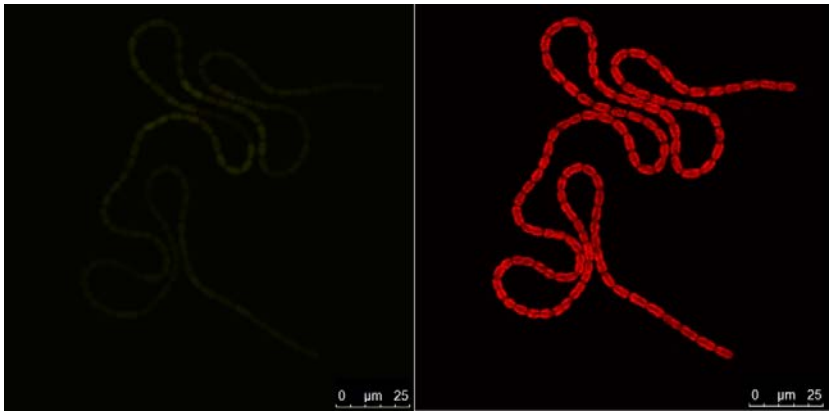

Figure S17. The synthetic short -10N construct 24 hours after nitrogen step down. The left image shows the YFP fluorescence and the right image shows the auto fluorescence.

8.4 RNA quality and RT- qPCR data

8.4.1 RNA extracted from the synthetic long and synthetic short construct

All biological replicates for the synthetic long construct contained RNA of acceptable quality for qPCR, figure S18.

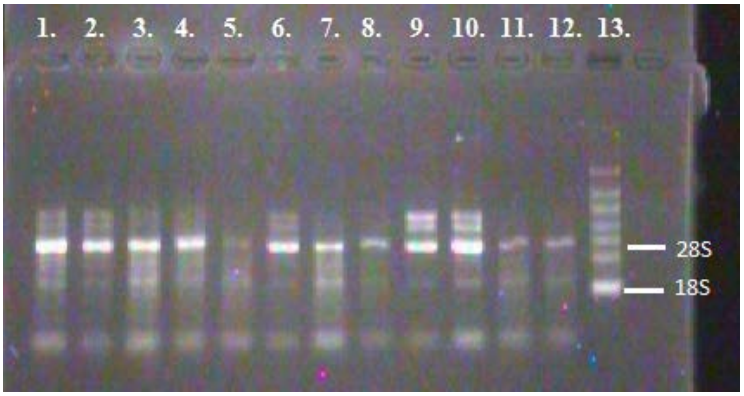

Figure S18. Well 1-3 contain RNA of the three biological replicates from the time point zero hours after nitrogen step down. Well 4-6 contain RNA of the three biological replicates from the time point three hours after nitrogen step down. Well 7-9 contain RNA of the three biological replicates from the time point six hours after nitrogen step down. Well 10-12 contain RNA of the three biological replicates from the time point nine hours after nitrogen step down. Well 13 contains a RNA ladder used to indicate the two subunits 28S and 18S.

As seen in figure S19, none of the biological replicates contain DNA contamination coding for the yfp gene whilst all samples to some extent contain DNA contamination of the 16s gene. Hence, -RT controls was implemented for all biological replicates and all control samples had a Cq value 10 cycles higher than the corresponding RT sample, see table S1.

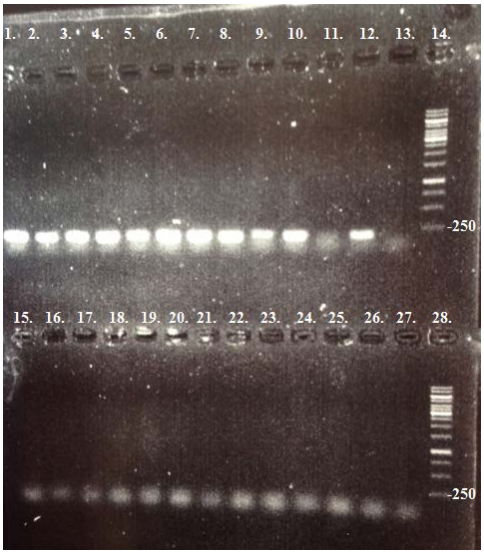

Figure S19. The upper row are the PCR products from the run with primers annealing to 16s gene and the lower row are the PCR products from the run with primers annealing to the yfp gene. Wells 1-3 and 15-17 contain the three biological replicates of time point 0h. Wells 4-6 and 18-20 contain the three biological replicates of time point 3h. Wells 7-9 and 21-23 contain the three biological replicates of time point 6h. Wells 10-12 and 24-26 contain the three biological replicates of time point 9h. Wells 13 and 27 are negative controls for each primer pair.

All biological replicates for the synthetic short construct contained RNA of acceptable quality for qPCR, figure S20.

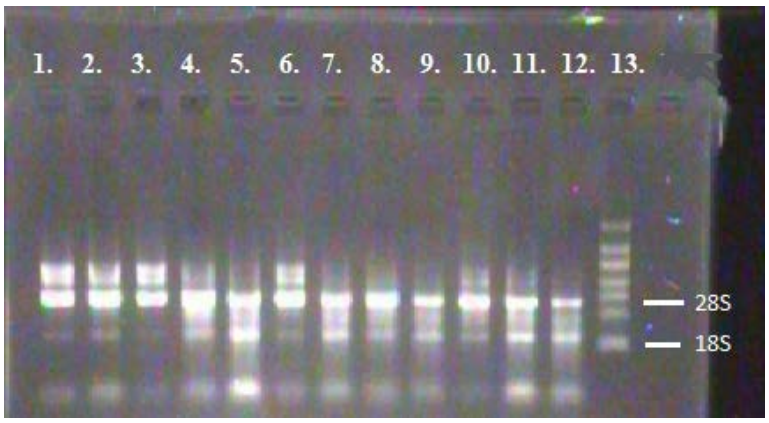

Figure S20. Well 1-3 contain RNA of the three biological replicates from the time point zero hours after nitrogen step down. Well 4-6 contain RNA of the three biological replicates from the time point three hours after nitrogen step down. Well 7-9 contain RNA of the three biological replicates from the time point six hours after nitrogen step down. Well 10-12 contain RNA of the three biological replicates from the time point nine hours after nitrogen step down. Well 13 contains a RNA ladder used to indicate the two subunits 28S and 18S.

As seen in figure S21 do not any of the biological replicates contain DNA contamination coding for the yfp gene whilst all samples to some extent contain DNA contamination of the 16s gene. Hence, -RT controls was implemented for all biological replicates and all control samples had a Cq value 10 cycles higher than the corresponding RT sample, see table S1.

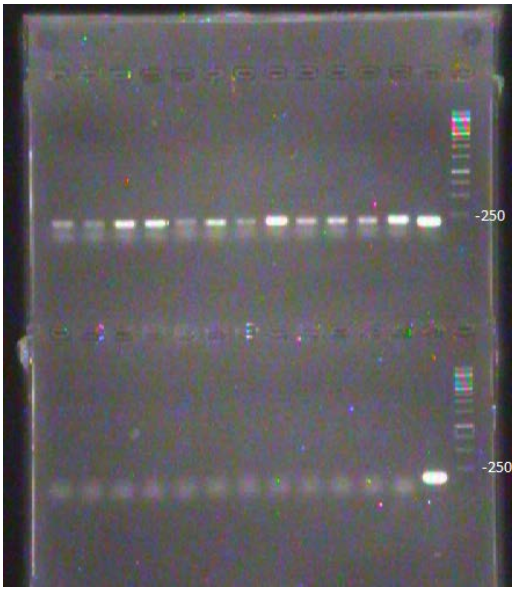

Figure S21. The upper row are the PCR products from the run with primers annealing to 16s gene and the lower row are the PCR products from the run with primers annealing to the yfp gene. Wells 1-3 and 15-17 contain the three biological replicates of time point 0h. Wells 4-6 and 18-20 contain the three biological replicates of time point 3h. Wells 7-9 and 21-23 contain the three biological replicates of time point 6h. Wells 10-12 and 24-26 contain the three biological replicates of time point 9h. Wells 13 and 27 are positive controls for each primer pair.

Table S1. RT-qPCR data of the activity of the synthetic long and synthetic short promoters. The mean Cq value for all three biological replicates for the four time point listed for both the RT reaction and the –RT reaction. SS means synthetic short and SL means synthetic long.

	Bio. rep.	SS, RT	SS, -RT	SL, RT	SL, -RT
0h	#1	7,11	27,08	9,01	31,68
	#2	5,5	23,47	9,3	31,6
	#3	4,72	28,34	8,21	31,59
3h	#1	6,21	31,07	7,84	31,63
	#2	7,3	31,6	11,47	31,74
	#3	6,45	31,56	8,8	31,55
6h	#1	7,63	31,73	9,29	31,57
	#2	8,64	31,53	9,4	31,57
	#3	7,19	31,56	7,65	31,59

9h	#1	6,21	31,69	7,99	31,54
	#2	6,63	31,68	9,44	31,63
	#3	7,57	31,68	9,18	31,66

8.4.2 RT- qPCR data for the synthetic short construct

All C_q values for the biological and technical replicates from the synthetic short construct are listed in table S2.

Table S2. The Cq values for the technical replicates for the synthetic short construct and the mean Cq value for each biological replicate, both for the 16s and the YFP gene.

0h	16s	YFP	3h	16s	YFP	6h	16s	YFP	9h	16s	YFP
Replicate #1	7,16	17,87		6,14	18,71		7,53	19,06		5,87	18,03
	6,99	17,62		6,54	17,5		8,15	18,38		6,27	17,77
	7,17	17,75		5,96	18,04		7,32	19,54		6,48	18,02
Mean #1	**7,10666667**	**17,74667**		**6,213333**	**18,08333**		**7,666667**	**18,99333**		**6,206667**	**17,94**
Replicate #2	5,21	17,41		7,29	18,73		8,99	19,73		6,74	18,16
	5,9	17,16		7	19,08		8,72	19,62		6,74	18,38
	5,39	18,09		7,62	19,2		8,19	20,36		6,43	18,33
Mean#2	**5,5**	**17,55333**		**7,303333**	**19,00333**		**8,633333**	**19,90333**		**6,636667**	**18,29**
Replicate #3	4,09	16,82		6,16	17,94		6,97	18,35		6,37	18,48
	3,84	17,57		7,15	19,09		7,42	19,44		7,22	19,24
	4,86	17,13			18,16		7,18	19,08		9,12	19,5
Mean #3	**4,26333333**	**17,17333**		**6,655**	**18,39667**		**7,19**	**18,95667**		**7,57**	**19,07333**

9 References

Brosius, Erfle M, Storella J. 1985. Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity. Journal of Biological Chemistry 260: 3539–3541.

Flaherty, Johnson DB, Golden JW. 2014. Deep sequencing of HetR-bound DNA reveals novel HetR targets in Anabaena sp. strain PCC7120. BMC Microbiology, doi 10.1186/s12866-014-0255-x.

Ionescu, Voss B, Oren A, Hess WR, Muro-Pastor AM. 2010. Heterocyst-Specific Transcription of NsiR1, a Non-Coding RNA Encoded in a Tandem Array of Direct Repeats in Cyanobacteria. Journal of Molecular Biology 398: 177–188.

Jensen, Hammer K. 1998. The Sequence of Spacers between the Consensus Sequences Modulates the Strength of Prokaryotic Promoters. Applied and Environmental Microbiology 64: 82–87.

Khanna, Raleiras P, Lindblad P. 2016. Fundamentals and Recent Advances in Hydrogen Production and Nitrogen Fixation in Cyanobacteria. In: Borowitzka, Beardall J, Raven

JA (ed.). The Physiology of Microalgae, pp. 101–127. Springer International Publishing,

Kotnik, Frey W, Sack M, Haberl Meglič S, Peterka M, Miklavčič D. 2015. Electroporation-based applications in biotechnology. Trends in Biotechnology 33: 480–488.

Kumar, Mella-Herrera RA, Golden JW. 2010. Cyanobacterial Heterocysts. Cold Spring Harbor Perspectives in Biology 2: a000315.

Li, Sandh G, Nenninger A, Muro-Pastor AM, Stensjö K. 2015. Differential transcriptional regulation of orthologous dps genes from two closely related heterocyst-forming cyanobacteria. FEMS Microbiology Letters, doi 10.1093/femsle/fnv017.

Livak, Schmittgen TD. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. Methods 25: 402–408.

Ma, Schmidt CM, Golden JW. 2014. Regulation of Gene Expression in Diverse Cyanobacterial Species by Using Theophylline-Responsive Riboswitches. Applied and Environmental Microbiology 80: 6704–6713.

Magnuson, Cardona T. 2016. Thylakoid membrane function in heterocysts. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1857: 309–319.

Mitschke, Vioque A, Haas F, Hess WR, Muro-Pastor AM. 2011. Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in Anabaena sp. PCC7120. Proceedings of the National Academy of Sciences 108: 20130–20135.

Muñoz-García, Ares S. 2016. Formation and maintenance of nitrogen-fixing cell patterns in filamentous cyanobacteria. Proceedings of the National Academy of Sciences of the United States of America 113: 6218–6223.

Muro-Pastor. 2014. The Heterocyst-Specific NsiR1 Small RNA Is an Early Marker of Cell Differentiation in Cyanobacterial Filaments. mBio 5: e01079-14.

Muro-Pastor, Brenes-Álvarez M, Vioque A. 2017. A combinatorial strategy of alternative promoter use during differentiation of a heterocystous cyanobacterium. Environmental Microbiology Reports 9: 449–458.

Wang. 2012. Application of synthetic biology in cyanobacteria and algae. Frontiers in microbiology 3: 344.

Xue, Barna M. 2012. Specialized ribosomes: a new frontier in gene regulation and organismal biology. Nature reviews Molecular cell biology 13: 355–369.

Integrated DNA technologies. 2015. Prime Time qPCR Application Guide. Experimental Overview, Protocol and Troubleshooting 4th Edition.

New England BioLabs. Colony PCR. WWW-document 2017: https://www.neb.com/applications/cloning-and-synthetic-biology/dna-analysis/colony-pcr. Gathered 2017-11-15.

Applied Biosystems. Determining genomic DNA contamination. WWW-document: http://www3.appliedbiosystems.com/WebTroubleshooting/NTCPositiveAmplification/index.htm#Amplification_of_the_No_Template_Control_NTC.htm. Gathered 2017-11-15.