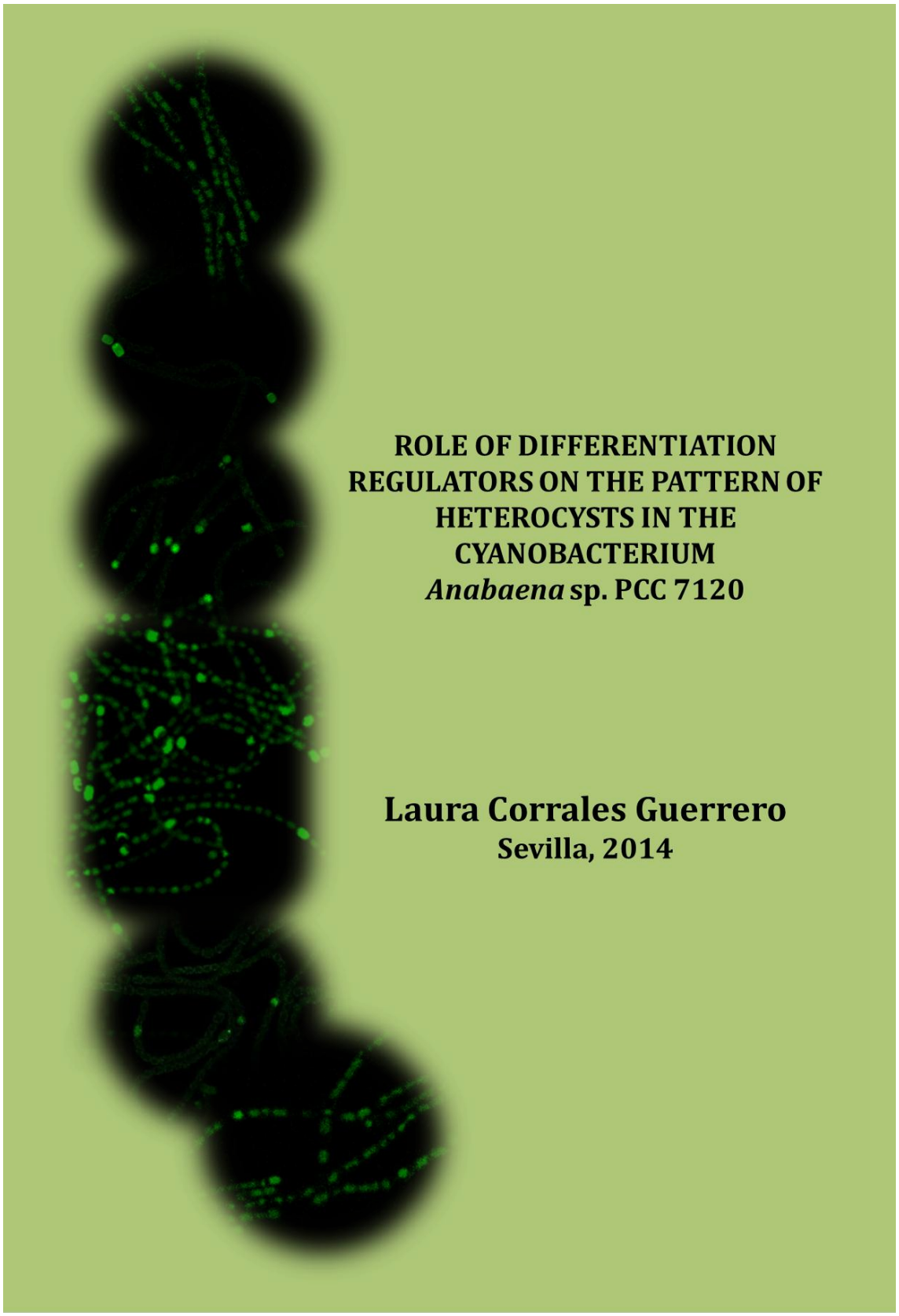


ROLE OF DIFFERENTIATION REGULATORS ON THE PATTERN OF HETEROCYSTS IN THE CYANOBACTERIUM *Anabaena* sp. PCC 7120

Laura Corrales Guerrero

Sevilla, 2014

Cover: GFP fluorescence of (up to down) HetR-GFP, HetN-sfGFP, PatS-GFP, HetR-GFP, HetP-sfGFP and HetR-GFP *patS*.

ACCIÓN DE REGULADORES DE LA DIFERENCIACIÓN EN EL ESTABLECIMIENTO DEL PATRÓN DE HETEROCISTOS EN LA CIANOBACTERIA *Anabaena* sp. PCC 7120

Trabajo presentado para optar al grado de
Doctora en Biología por la Licenciada
Laura Corrales Guerrero

Sevilla, marzo de 2014

Director

Dr. Enrique Flores García
Profesor de Investigación del CSIC

Directora

Dra. Antonia Herrero Moreno
Profesora de Investigación del CSIC

Tutor

Dr. Jose María Romero Rodríguez
Catedrático de la Universidad de Sevilla

INDEX

FIGURES INDEX

TABLES INDEX

ABREVIATIONS

A	Amperes
aa	Amino acid
Ap	Ampicilin
bp	Base pair
Chl	Chlorophyll *a*
Cm	Chloramphenicol
DCMU	Dichlorophenyl dimethylurea
dCTP	Deoxycytidine triphosphate
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNAase	Deoxyribonuclease
dNTPs	Deoxyribonucleotides
ε	Molar extinction coefficient
E	Einstein
EDTA	Ethylenediaminetetraacetic acid
F	Farads
FITC	Fluorescein isothiocyanate
GFP	Green Fluorescent Protein
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
Km	Kanamycin
LB	Luria-Bertani medium
Mch	Multiple Contiguous Heterocysts phenotype
milliQ	Ultrapure water (Trademark)
min	minute
MOPS	4-Morpholinepropanesulfonic acid
Nm	Neomycin
nt	nucleotides
Ω	Ohms
PBS	Phosphate Buffered Saline
PCC	Pasteur Culture Collection of Cyanobacteria
PCR	Polymerase Chain Reaction
PVDF	Polyvinylidene fluoride
RNA	Ribonucleic acid
RNAase	Ribonuclease
rpm	Revolutions per minute

SDS	Sodium dodecyl sulphate
S	Seconds
Sm	Streptomycin
Sp	Spectinomycin
SSC	Saline-Sodium Citrate (buffer)
SSPE	saline-sodium phosphate-EDTA (buffer)
TBE	tris-borate-EDTA (buffer)
TBS	Tris-*buffered* saline
TLC	Thin-layer chromatography
TES	N-Tris(Hydroxymethyl)methyl-2-Amino-Ethanesulfonic Acid
μ	Specific growth constant
U	Enzymatic activity units
v/v	Volume /volume ratio
V	Volts
w/v	Weight /volume ratio
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1. INTRODUCTION

1.1. CYANOBACTERIA

Cyanobacteria, also known as blue-green algae, constitute a large group of Gram-negative prokaryotes that are phylogenetically well defined and can be found in a large variety of habitats (Stanier & Cohen-Bazire, 1977; Woese, 1987). Cyanobacteria were the first organisms capable of oxygenic photosynthesis, and they are considered the ancestors of plant chloroplasts (Knoll, 2008). They are supposed to have been responsible for the change from a reductive to an oxidative atmosphere 2.2 to 2.4 billion years ago (Knoll, 2008).

Nowadays, cyanobacteria still play an important role in the carbon and nitrogen cycles, as they take CO_2 out from the air and liberate O_2 instead, and many of them are able to fix molecular nitrogen (Flores & Herrero, 2010). Additionally, they are used for many biotechnological purposes such as the generation of bioactive compounds, production of bioplastics, bioremediation, production of alternative energy sources and as a source of fertilizers and healthy food (Abed et al., 2009).

The phylum comprises groups of organisms with very different morphologies including unicellular, either single cells or aggregates, and filamentous forms, either lineal or ramified (Rippka *et al.*, 1979). Some filamentous cyanobacteria are true multicellular organisms that differentiate cells such as heterocysts, specialized in nitrogen fixation, akinetes (resting cells) or hormogonia that are small-sized motile filaments (Flores & Herrero, 2010).

The recent availability of a large number of cyanobacterial genomic sequences has allowed a complete phylogenetic analysis (Figure 1.1) (Shih *et al.*, 2013). According to this study, the heterocyst-forming cyanobacteria have a monophyletic origin. However, it has been proposed that some regulatory proteins necessary for heterocyst differentiation must have been present in a common ancestor that did not form heterocysts (Shih et al., 2013; Zhang *et al.*, 2009).

This work has focused on *Anabaena* sp. PCC 7120, hereafter *Anabaena*, which belongs to the group of lineal filamentous heterocyst-forming cyanobacteria (Figure 1.2.A,B).

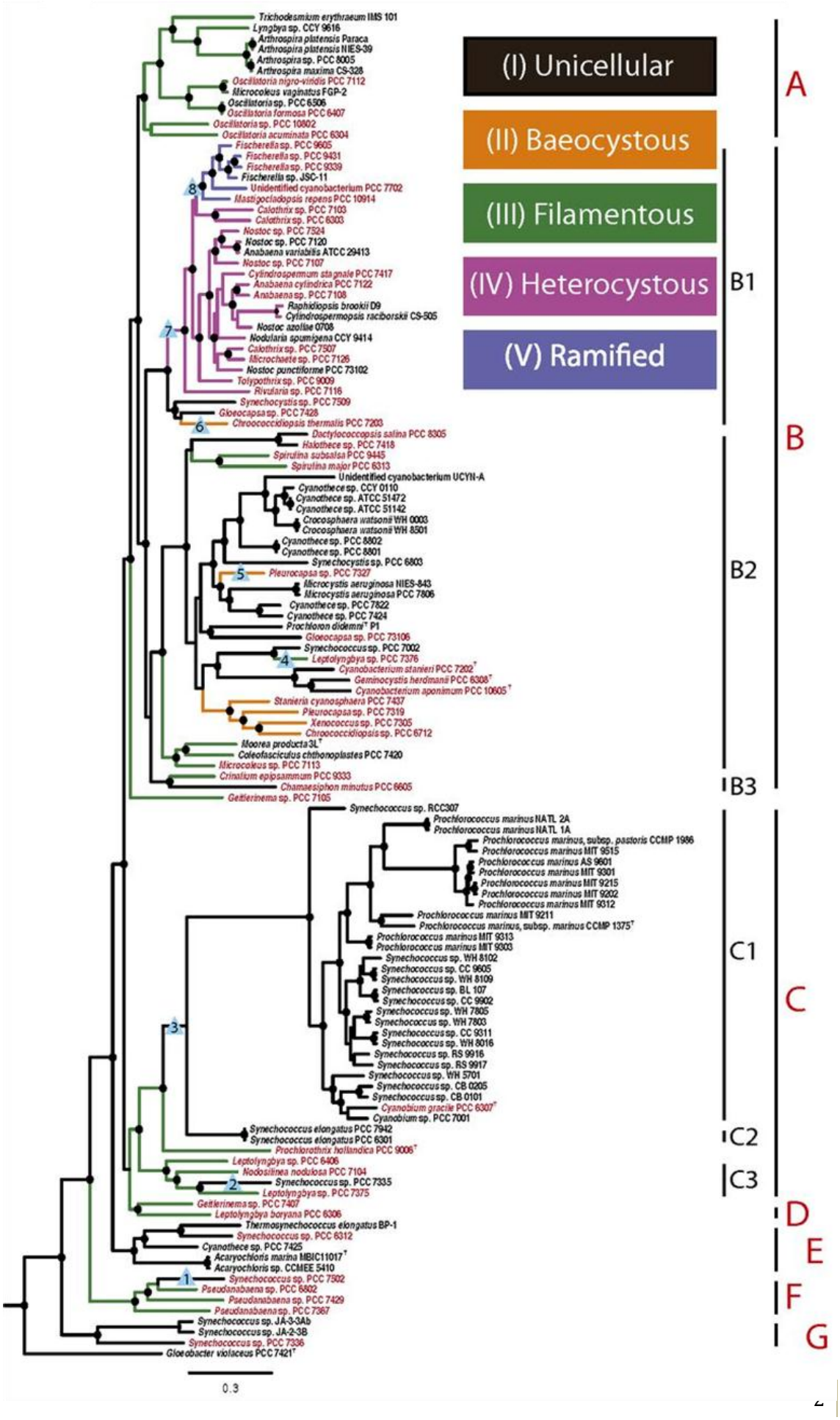

(I) Unicellular
(II) Baeocystous
(III) Filamentous
(IV) Heterocystous
(V) Ramified
A
B
B1
B2
B3
C
C1
C2
C3
D
E
F
G
Trichodesmium erythraeum IMS 101
Lyngbya sp. CCY 9616
Arthrospira platensis Paraca
Arthrospira platensis NIES-39
Arthrospira sp. PCC 8005
Arthrospira maxima CS-328
Oscillatoria nigro-viridis PCC 7112
Microcoleus vaginatus FGP-2
Oscillatoria sp. PCC 6506
Oscillatoria formosa PCC 6407
Oscillatoria sp. PCC 10802
Oscillatoria acuminata PCC 6304
Fischerella sp. PCC 9605
Fischerella sp. PCC 9431
Fischerella sp. PCC 9339
Fischerella sp. JSC-11
Unidentified cyanobacterium PCC 7702
Mastigocladopsis repens PCC 10914
Calothrix sp. PCC 7103
Calothrix sp. PCC 6303
Nostoc sp. PCC 7524
Nostoc sp. PCC 7120
Anabaena variabilis ATCC 29413
Nostoc sp. PCC 7107
Cylindrospermum stagnale PCC 7417
Anabaena cylindrica PCC 7122
Anabaena sp. PCC 7108
Raphidiopsis brookii D9
Cylindrospermopsis raciborskii CS-505
Nostoc azollae 0708
Nodularia spumigena CCY 9414
Calothrix sp. PCC 7507
Microchaete sp. PCC 7126
Nostoc punctiforme PCC 73102
Tolypothrix sp. PCC 9009
Rivularia sp. PCC 7116
Synechocystis sp. PCC 7509
Gloeocapsa sp. PCC 7428
Chroococcidiopsis thermalis PCC 7203
Dactylococcopsis salina PCC 8305
Halothece sp. PCC 7418
Spirulina subsalsa PCC 9445
Spirulina major PCC 6313
Unidentified cyanobacterium UCYN-A
Cyanothece sp. CCY 0110
Cyanothece sp. ATCC 51472
Cyanothece sp. ATCC 51142
Crocosphaera watsonii WH 0003
Crocosphaera watsonii WH 8501
Cyanothece sp. PCC 8802
Cyanothece sp. PCC 8801
Synechocystis sp. PCC 6803
Pleurocapsa sp. PCC 7327
Microcystis aeruginosa NIES-843
Microcystis aeruginosa PCC 7806
Cyanothece sp. PCC 7822
Cyanothece sp. PCC 7424
Prochloron didemni P1
Gloeocapsa sp. PCC 73106
Synechococcus sp. PCC 7002
Leptolyngbya sp. PCC 7376
Cyanobacterium stanieri PCC 7202
Geminocystis herdmanii PCC 6308
Cyanobacterium aponinum PCC 10605
Stanieria cyanosphaera PCC 7437
Pleurocapsa sp. PCC 7319
Xenococcus sp. PCC 7305
Chroococcidiopsis sp. PCC 6712
Moorea producta 3L
Coleofasciculus chthonoplastes PCC 7420
Microcoleus sp. PCC 7113
Crinalium epipsammum PCC 9333
Chamaesiphon minutus PCC 6605
Geitlerinema sp. PCC 7105
Synechococcus sp. RCC307
Prochlorococcus marinus NATL 2A
Prochlorococcus marinus NATL 1A
Prochlorococcus marinus, subsp. pastoris CCMP 1986
Prochlorococcus marinus MIT 9515
Prochlorococcus marinus AS 9601
Prochlorococcus marinus MIT 9301
Prochlorococcus marinus MIT 9215
Prochlorococcus marinus MIT 9202
Prochlorococcus marinus MIT 9312
Prochlorococcus marinus MIT 9211
Prochlorococcus marinus, subsp. marinus CCMP 1375
Prochlorococcus marinus MIT 9313
Prochlorococcus marinus MIT 9303
Synechococcus sp. WH 8102
Synechococcus sp. CC 9605
Synechococcus sp. WH 8109
Synechococcus sp. BL 107
Synechococcus sp. CC 9902
Synechococcus sp. WH 7805
Synechococcus sp. WH 7803
Synechococcus sp. CC 9311
Synechococcus sp. WH 8016
Synechococcus sp. RS 9916
Synechococcus sp. RS 9917
Synechococcus sp. WH 5701
Synechococcus sp. CB 0205
Synechococcus sp. CB 0101
Cyanobium gracile PCC 6307
Cyanobium sp. PCC 7001
Synechococcus elongatus PCC 7942
Synechococcus elongatus PCC 6301
Prochlorothrix hollandica PCC 9006
Leptolyngbya sp. PCC 6406
Nodosilinea nodulosa PCC 7104
Synechococcus sp. PCC 7335
Leptolyngbya sp. PCC 7375
Geitlerinema sp. PCC 7407
Leptolyngbya boryana PCC 6306
Thermosynechococcus elongatus BP-1
Synechococcus sp. PCC 6312
Cyanothece sp. PCC 7425
Acaryochloris marina MBIC11017
Acaryochloris sp. CCMEE 5410
Synechococcus sp. PCC 7502
Pseudanabaena sp. PCC 6802
Pseudanabaena sp. PCC 7429
Pseudanabaena sp. PCC 7367
Synechococcus sp. JA-3-3Ab
Synechococcus sp. JA-2-3B
Synechococcus sp. PCC 7336
Gloeobacter violaceus PCC 7421
1
2
3
4
5
6
7
8
0.3

Figure 1.1. Cyanobacterial maximum-likelihood phylogenetic tree. The 126 genomes sequenced to date are presented in the tree, which was generated with a sequence resulting from concatenating 31 proteins. Branches are coloured according to morphological subsection. Nodes supported with a bootstrap of at least 70% are indicated by a black dot. Morphological transitions are denoted by blue triangles. Image taken from Shih et al. (2013).

1.2. MULTICELLULARITY AND INTERCELLULAR COMMUNICATION

Many species of unicellular bacteria have developed mechanisms to communicate with each other since group behaviour confers evolutionary advantages (Dirix *et al.*, 2004; Claessen *et al.*, 2014). *Anabaena* filaments constitute a step forward in evolution as they are true multicellular organisms. That is, the cells cannot survive independently but they require cohesion, transport of regulatory substances and metabolites between cells, and cooperation between different cells types (Mariscal & Flores, 2010).

Cells in a filament possess their own cytoplasmic membrane and peptidoglycan layer(s), however a continuous outer membrane, impermeable to many substances, is continuous along the filament (Figure 1.2.C) (Nicolaisen *et al.*, 2009). As a result, the periplasm is a common space shared by all the cells of the filament (Mariscal *et al.*, 2007). Some proteinaceous structures that have been called microplasmodesmata, septosomes or septal junctions, lie in the septa between cells, and they can be observed as electron-dense thin strings perpendicular to the membranes (Figure 1.2.C) (Wilk *et al.*, 2011). These structures have been characterized as channel complexes containing the SepJ and Fra proteins (Flores *et al.*, 2007; Merino-Puerto *et al.*, 2013). They have been suggested to have a double function, binding of adjacent cells and transport between them. Therefore, two different ways of communication have been proposed in *Anabaena* filaments, directly through the cells and through the periplasm.

When growing diazotrophically, two different cellular types coexist in the *Anabaena* filament, vegetative cells that fix CO_2 photosynthetically and heterocysts that are specialised in the fixation of molecular nitrogen, which requires an anoxic environment not compatible with oxygenic photosynthesis. When the filaments are incubated in the absence of combined nitrogen, global changes in gene expression take place leading to heterocyst differentiation and, thus, the division of labour between the two different cells (Flaherty *et al.*, 2011). Heterocysts transfer combined nitrogen to vegetative cells, which in turn donate CO_2 fixation products to heterocysts (Flores & Herrero, 2010).

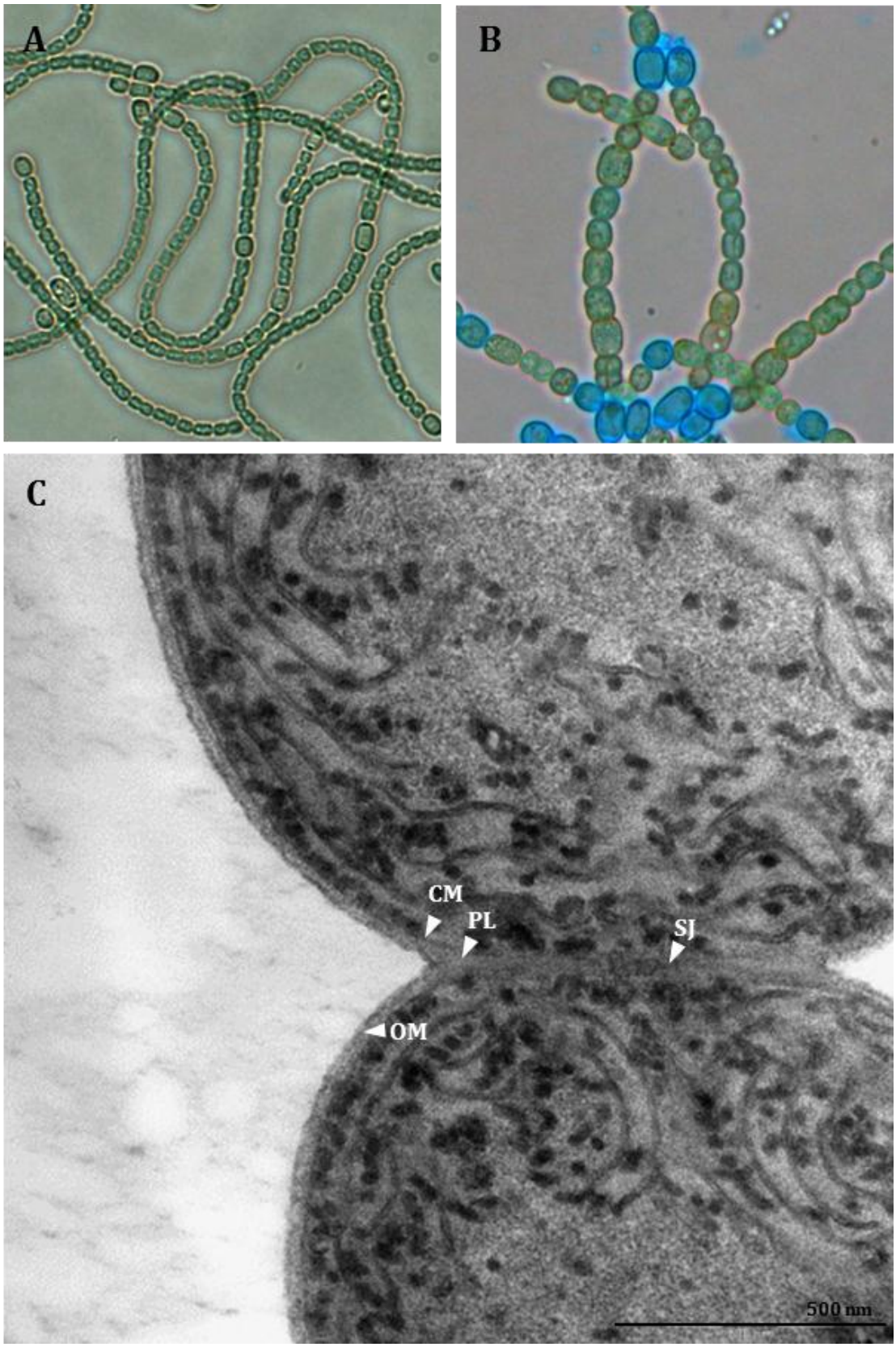


Figure 1.2. Filament and cell structure in *Anabaena*. A. Light microscopy of *Anabaena* filaments grown under diazotrophic conditions. Magnification, 40x. B. Light microscopy of heterocysts stained with Alcian Blue. Magnification, 100x. C. Electron microscopy, uranyl-acetate staining, of a septum between vegetative cells. Previously described structures are indicated: cytoplasmic membrane (CM), peptidoglycan layer (PL), septal junctions (SJ) and outer membrane (OM).

1.3. HETEROCYST DIFFERENTIATION

A mature heterocyst may be distinguished from vegetative cells by their larger size, diminished pigmentation due to the loss of chlorophyll and phycobilisomes, thicker cell envelopes and evident cyanophycin granules at the poles adjacent to vegetative cells (Kumar *et al.*, 2010). A mature heterocyst is recognizable 16-18 hours after nitrogen deprivation, but an intermediate state, the proheterocyst, can be noticed at around 8 hours.

1.3.2. STRUCTURAL CHANGES

Heterocysts possess an especial envelope that reduces the diffusion of gasses into the cytoplasm. This envelope consists of an inner glycolipid layer and an outer polysaccharide layer (Figure 1.3). The latter, composed of specific polysaccharides, is the first being deposited, is thick and homogenous, and is supposed to protect the inner layer from physical damage (Cardemil & Wolk, 1979; Murry & Wolk, 1989). The glycolipid layer is thin, laminated and composed of specific glycolipids that represent a permeability barrier for gases, among them O_2 (Bryce *et al.*, 1972). Many genes involved in the deposition of these layers have been characterized, and mutants of these genes are incapable of nitrogen fixation under oxic conditions (Kumar et al., 2010).

There is a reorganization of membranes during the process of heterocyst differentiation. The thylakoid membrane system, which in vegetative cells is located in the periphery of the cell, is rearranged to a reticular membrane system that is concentrated close to the poles. It contains enzymes involved in respiration, needed to eliminate traces of oxygen that could enter through the heterocyst neck. This membrane system is called "honeycomb" because of its reticular look (Lang & Fay, 1971) (Figure 1.3).

Some cellular structures typical of vegetative cells such as carboxysomes and glycogen granules disappear during differentiation. However, cyanophycin granules are formed in the poles of the heterocysts in a cup-like structure (Lang & Fay, 1971) (Figure 1.3). Cyanophycin is a polymer of arginine and aspartate that is used as nitrogen storage (Picossi *et al.*, 2004).

The septa between heterocysts and vegetative cells is narrowed to form the heterocyst neck (Lang & Fay, 1971). The narrowness of these septa and the presence of the cyanophycin granule produce a slower rate of metabolite exchange between vegetative cells and heterocysts than between vegetative cells (Mullineaux *et al.*, 2008).

1.3.3. METABOLIC CHANGES

The enzyme nitrogenase is inhibited by O_2, making an anoxic or microoxic environment essential for its activity. The enzyme is constituted by a tetramer of the

polypeptides encoded by the genes *nifD* and *nifK*, which hold the catalytic site, and a dimer of the polypeptide encoded by *nifH* that acts as electron donor (Wolk *et al.*, 1994).

The O_2-producing photosystem II is disassembled during heterocyst differentiation. This allows the recycling of the amino acids contained in antenna pigments, phycobiliproteins, that can account for up to ca. 50% of the soluble protein in cyanobacteria (Meeks & Elhai, 2002). However, photosystem I-based electron flow is operative in heterocysts and represents a source of reductant and ATP for nitrogenase (Wolk et al., 1994). In addition, the heterocysts exhibit a high rate of respiration catalysed by specific oxidases with a dual function, elimination of residual O_2 and generation of ATP (Valladares *et al.*, 2003).

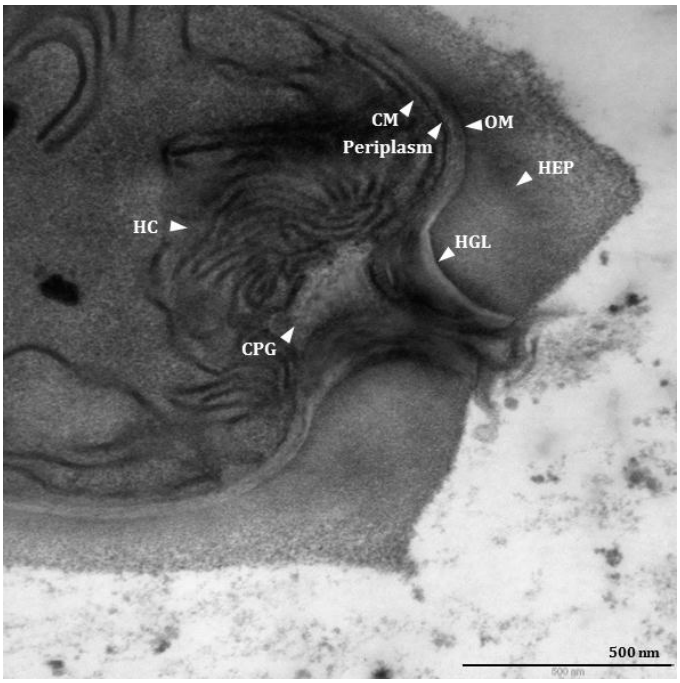


Figure 1.3. Transmission electron micrograph of a detached heterocyst. Some structures are indicated: cytoplasmic membrane (CM), periplasm, outer membrane (OM), cyanophycin granule (CPG), honeycomb membranes (HC), heterocyst glycolipid layer (HGL) and heterocyst polysaccharide layer (HEP). The OM and HGL are difficult to distinguish because they are very close to each other.

The heterocyst lacks the Calvin-Benson cycle for CO_2 fixation and depends on the vegetative cells to obtain reduced carbon. Sucrose has been proposed to be an important transferred molecule since invertases and the oxidative pentose-phosphate pathway are important in heterocysts (Flores & Herrero, 2010; López-Igual *et al.*, 2010).

The fixed nitrogen is incorporated into glutamine and other amino acids in the heterocysts. Heterocysts have high levels of glutamine synthetase and no glutamate synthase, so glutamine is donated from heterocysts to vegetative cells, whereas glutamate is transferred from the latter to the heterocysts (Flores & Herrero, 2010). In addition,

alanine is imported into heterocysts (Pernil *et al.*, 2010), and the cyanophycin-derived β–aspartyl arginine dipeptide is transferred from heterocysts to vegetative cells (Burnat *et al.*, 2014).

Heterocysts are terminally differentiated cells, which mean that they have lost the ability to divide and revert to a vegetative cell. The cell division machinery is down-regulated in these cells at both the transcriptional and translational levels (Wang & Xu, 2005; Kuhn *et al.*, 2000).

1.4. REGULATION OF HETEROCYST DIFFERENTIATION

According to microarray analysis, nitrogen deprivation significantly induces 495 genes and inhibits expression of 196 genes in *Anabaena* (Xu *et al.*, 2008; Ehira & Ohmori, 2006b). Many of these genes bear complex promoters with different transcription start sites directed by different promoter determinants (Valladares *et al.*, 2008; Picossi *et al.*, 2014; Mitschke *et al.*, 2011).

At least 30 proteins have been described to be involved in the regulation of heterocyst differentiation (see Figure 1.4 for a scheme). Two kinds of regulatory circuits take place during heterocyst differentiation: regulation of gene expression and generation of positional information (Flores & Herrero, 2010). Positive and negative elements participate in the regulation, and the balance between them controls the differentiation process to come to an end or not.

It has been proposed that during the differentiation, genes are expressed in an ordered sequence so that gene expression at one stage depends on expression in previous stages (Wolk et al., 1994). Indeed, expression of genes encoding regulatory proteins takes place within the first hours after nitrogen deprivation.

Upon N step-down, cyanobacterial cells sense the lack of combined nitrogen in the media through a rise in the intracellular 2-oxoglutarate (2-OG) concentrations (Li *et al.*, 2003). This molecule serves as a carbon skeleton for ammonium assimilation (Vázquez-Bermúdez *et al.*, 2000). Therefore, under nitrogen starvation 2-OG accumulates within the cells, this accumulation being sensed by NtcA, which is a transcriptional regulator that belongs to the CRP (cyclic AMP receptor protein) family and controls the largest bacterial regulon characterized to date (Picossi et al., 2014). However, NtcA was first described as the factor that exerts global nitrogen control in cyanobacteria (Frías *et al.*, 1994). In fact, *ntcA* mutants lacking a functional *ntcA* gene are impaired in the assimilation of nitrate and do not produce heterocysts (Frías et al., 1994). NtcA acts as a homodimer in which each subunit presents a helix-turn-helix domain for DNA binding in the C-terminal end

(Herrero *et al.*, 2001) (Figure 1.5). This regulator binds to DNA at sites with the consensus sequence GTAN$_8$TAC (Luque *et al.*, 1994). NtcA can act as an activator of transcription by a class II mechanism in promoters including a binding site centred at approx. -41.5 bp from the transcriptional start point or by a class I mechanism (Busby & Ebright, 1999), or act as a repressor (Herrero *et al.*, 2004).

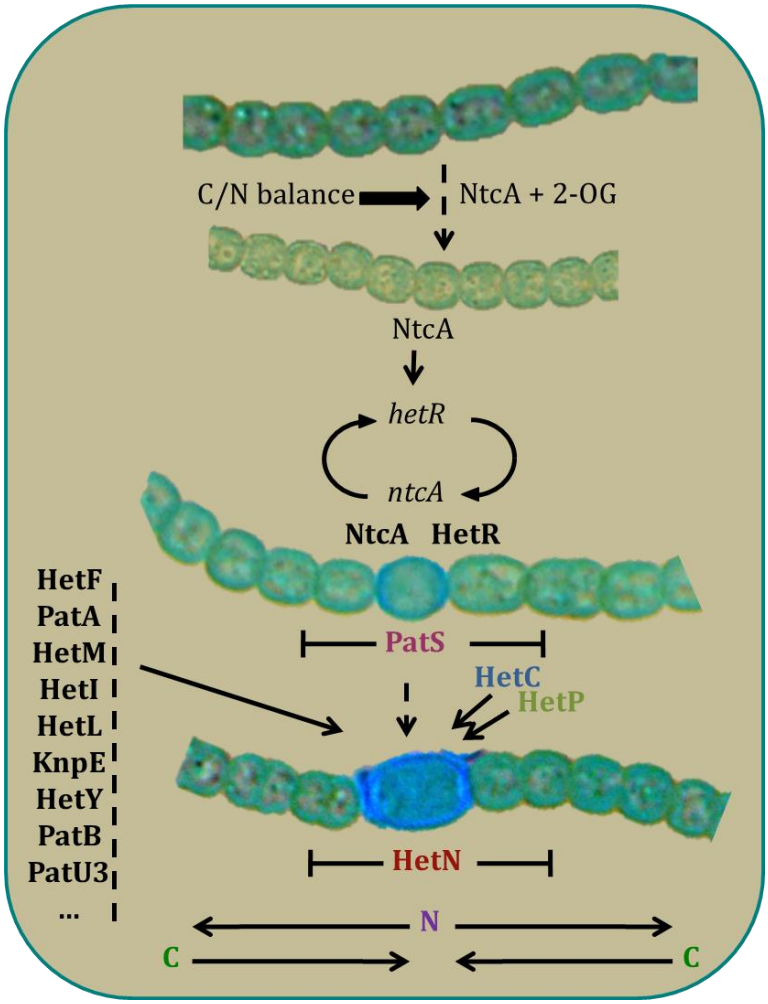


Figure 1.4. Scheme of the heterocyst differentiation process. A filament deprived of combined nitrogen experiences a complex process that results in the formation of a mature heterocyst. Some of the regulatory proteins involved in the process are indicated. Adapted from Flores & Herrero (2010).

To promote heterocyst differentiation, another transcriptional regulator, HetR, is required, which is a differentiation-specific regulator. Inactivation mutants of *hetR* are also unable to produce heterocysts (this phenotype is called Het $^-$), whereas strains overexpressing this protein produce more heterocysts than the wild-type strain (Buikema & Haselkorn, 1991; Camargo *et al.*, 2012). During differentiation, the *ntcA* and *hetR* genes are both positively autoregulated and induced in a mutually-dependent manner, producing an amplification loop of gene expression necessary to reach the protein levels

needed for progression of differentiation (Muro-Pastor *et al.*, 2002; Black *et al.*, 1993) (Figure 1.4). The activation of *hetR* is also enhanced by NrrA, a response regulator whose expression is, in turn, activated by NtcA (Ehira & Ohmori, 2006a). *hetR* expression can be detected at low levels in cells growing with combined nitrogen, and it increases in clusters of cells after ca. 0.5 h after N deprivation so that there is an increase in the protein after approx. 6 hours (Black et al., 1993; Buikema & Haselkorn, 1991). After the induction of *hetR*, the cell is committed to become a heterocyst, and during the middle and late stages of differentiation all the structural and metabolic changes occur (Yoon & Golden, 2001; Cai & Wolk, 1997).

The HetR protein acts as a dimer and has autoproteolitic and DNA binding activity (Khudyakov & Golden, 2004; Kim *et al.*, 2011). A HetR protein bearing the substitution $S^{179}N$ does not allow heterocyst differentiation (Buikema & Haselkorn, 1991). Three other residues (D^{17}, G^{36} and H^{69}) are essential for maintaining normal levels of the HetR protein and therefore for a normal heterocyst differentiation (Risser & Callahan, 2007).

The crystal structure of HetR from a *Fischerella* sp. has been determined recently (Kim et al., 2011; Kim *et al.*, 2013). The protein forms a dimer comprised of a central DNA-binding unit containing the N-terminal regions of the two subunits organized with two helix-turn-helix motifs, two globular flaps extended in opposite directions and a hood over the central core formed from the C-terminal subdomains (Figure 1.5). The flap domains might interact with the target DNA in a sequence-nonspecific manner.

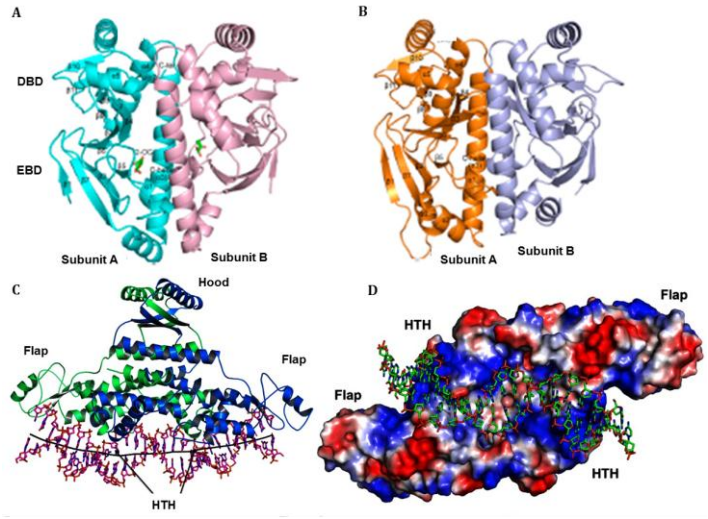


Figure 1.5. Crystal structures of NtcA and HetR. A. Ribbon representation of an NtcA dimer (represented in blue and pink) bound to two molecules of 2-OG (close to the center). The effector binding domain (EBD) and the DNA-binding domain (DBD) are indicated. B. Ribbon representation of an apo-NtcA dimer (represented in orange and light blue). C. Ribbon representation of a HetR dimer bound to DNA showing the flap, hood and helix-turn-helix domains. D. Surface representation of the HetR dimer. Adapted from (A, B) Zhao et al. (2010) and (C, D) Kim et al. (2013).

1.5. ESTABLISHMENT OF THE SPATIAL PATTERN OF HETEROCYSTS

Cyanobacterial heterocysts are distributed following a defined pattern along the filament that depends on the species. It represents the simplest and oldest evolutionary model for the study of the generation of biological patterns during development. In *Anabaena* sp. PCC 7120, clusters of 3-4 cells initiate the process of differentiation by increasing the HetR intracellular concentrations (Black et al., 1993), however only one of them complete the differentiation process, separated by approx. 10-15 vegetative cells (Figure 1.2.A,B). How these clusters are formed and resolved to a single differentiated cell is still unclear.

When nitrogen-replete filaments of *Anabaena* are deprived of nitrogen, heterocyst formation is initiated at many positions along the filament. The interval size of vegetative cells between two heterocysts rises to a maximum (normally 10-15 cells, maximum 30 cells in rare occasions), and is maintained as the filament grows by formation of a new heterocyst more or less equidistant from others. If this pattern were established randomly, the frequency distribution of the number of vegetative cells (n) between differentiating cells would be a monotonically decreasing exponential function such as $f(1-f)^{n-1}$, where f is the fraction of differentiating cells (Wolk, 1967). This is not the case and, therefore, the heterocyst distribution is not random, resulting instead from regulatory processes.

The simplest model proposed to explain the pattern was a negative regulation by products of N_2 fixation (Fay *et al.*, 1968). However, expression of the positive element, *hetR*, takes place only in defined clusters of cells just minutes after the N step-down, much before nitrogenase activity can be detected. This fact suggests the operation of regulatory circuits that at least create the *de novo* pattern independently of N_2 fixation (Wolk et al., 1994). In fact, Wolk & Quine (1975) showed, by computer simulation, that the heterocyst pattern could result from inhibition by a diffusible substance produced by these cells.

A two-stage model was later proposed to explain the differentiation process, based on the hypothesis that the stage of a cell in the cell cycle influences its competence to become a heterocyst (Meeks & Elhai, 2002). In a first stage (initiation), all the cells would sense nitrogen starvation but only a fraction on them, possibly those in a critical stage of the cell cycle, would be competent to initiate differentiation. These cells, which will be contiguous as sibling cells are normally at the same stage in the cell cycle, would increase *hetR* expression to a moderate level. In a second stage (resolution), expression of *hetR* would lead to the synthesis and release of a negative signal that would inhibit HetR in the neighbouring cells, which would end with the cluster resolved to a single cell. Which cell in

the cluster expresses before or in a higher quantity the inhibitory signal may depend on the initial levels of HetR and/or other factors.

This model is consistent with experimental observations of a phenotype called Multiple Contiguous Heterocysts (Mch) in which the heterocysts are usually found in pairs or groups, although the interval size is quite preserved. This distribution occurs naturally in species such as *Mastigocladus laminosus* or *Anabaena cylindrica* (Wolk et al., 1994), but can be also found in *Anabaena* strains overexpressing the *hetR* gene or bearing the substitution $R^{223}W$ in the HetR amino-acid sequence (Khudyakov & Golden, 2004). The strain bearing the latter mutation has been proposed to be insensitive to the inhibitory signals.

The pattern of heterocysts has been classified by Meinhardt (2008) as a reaction-diffusion-based model in which only the inhibitor and not the activator diffuses (Figure 1.6). In this way, the process must be triggered by an activator that is subject to nonlinear autocatalysis. The activator also stimulates the expression of the inhibitor that diffuses to other cells and stops the autocatalysis of the activator. This theory predicts that, if HetR were eliminated, no heterocysts would be formed. In contrast, if the inhibitor were mutated, most cells would form heterocysts. This model has been supported by fitting experimental data to mathematical simulations in which one or two different diffusible inhibitors and HetR as the activator are considered (Gerdtzen *et al.*, 2009; Zhu *et al.*, 2010).

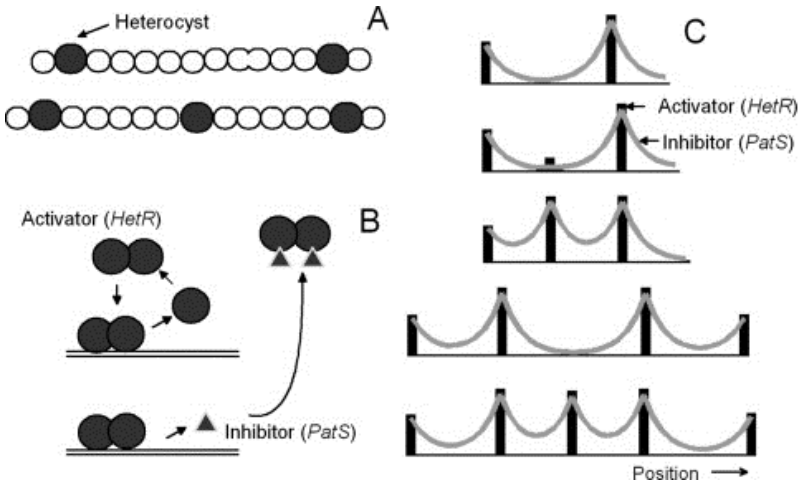


Figure 1.6. Reaction-diffusion-based model applied to heterocyst pattern formation. A. Example of the maintenance of the interval size between heterocysts (dark circles) during the growth of the filament. B. HetR dimers that directly activate *hetR* transcription and also activate the formation of a small peptide that can spread through the filament and bind to HetR. If PatS is bound to HetR, DNA-binding of HetR is no longer possible, so PatS inhibits the activator autocatalysis. C. Simulation: only the inhibitor is diffusible across the cells, so activation occurs in isolated cells. Whenever the inhibitor drops below a threshold level, a new autocatalysis of the activator is triggered from baseline activation. Since the inhibitor distribution around a minimum is shallow,

initially more than one cell can start this activation process. Due to competition, only one isolated cell eventually becomes activated. Image taken from Meinhardt (2008).

1.5.1. PatS

The PatS peptide has been demonstrated to be an inhibitor of the process of heterocyst differentiation. The *patS* gene (*asl2301*) was first identified in a cosmid that supressed heterocyst differentiation (Yoon & Golden, 1998). Subcloning experiments were used to identify a 1-kb region of the cosmid that produced a strong heterocyst inhibition phenotype. A small open reading frame was identified consisting of thirteen or seventeen codons since it contains two potential ATG translational starts. Further subcloning experiments showed that a PCR fragment containing this open reading frame suppresses heterocysts when placed downstream of a strong promoter (that of the *glnA* gene). All these constructs involved expression of the *patS* gene in a replicative plasmid, whose copy number in the cell was unpredictable, leading to overexpression of the *patS* gene in the resulting strains. A strain bearing a deletion of 381 bp including the *patS* gene substituted by the C.S3 gene cassette exhibited an Mch phenotype under combined nitrogen deficiency and formed heterocysts when growing with nitrate as a nitrogen source (Yoon & Golden, 1998). This remarkable phenotype is consistent with PatS being an inhibitor of the differentiation process.

Genes homologous to *patS*, as well as to *hetR*, are widespread among filamentous cyanobacteria, even in those that do not form heterocysts (Zhang et al., 2009). The PatS amino-acid sequence is highly conserved in *Anabaena, Anabaena variabilis* and *Nostoc punctiforme*. All these sequences share a C-terminal end containing the ERGSGR sequence and a highly hydrophobic N-terminal region.

Two transcriptional start points (tsp) were first identified in the promoter region of *patS* (Yoon & Golden, 2001). A tsp located at -39 nt with regard to the translation start of the gene is developmentally regulated to give increased transcription in differentiating cells, and the one located at -314 nt appears to be active in vegetative cells and provides a basal level of transcription under all conditions of growth. A third tsp, located at -254 nt, dependent on HetR, was proposed in a recent analysis (Mitschke et al., 2011). A strain bearing a transcriptional fusion with the GFP in a replicative plasmid shows that after differentiation of an initial patterned group of proheterocysts, the expression of *patS* reverts back to a basal level in most cells, and after 12 hours of nitrogen deprivation only proheterocysts strongly express the GFP (Yoon & Golden, 1998). Consistent results were obtained using northern analysis and *patS-lacZ* fusions (Yoon & Golden, 2001).

It has been hypothesized that the PatS peptide could act in a way similar to bacterial signalling peptides involved in quorum-sensing and regulation of cellular

differentiation (Wu *et al.*, 2004). If this were the case, the peptide would be recognised through the N-terminal end and processed to a smaller C-terminal, active form perhaps during export outside of the producing cell (Bassler, 2002).

Random chemical mutagenesis of a plasmid carrying the *patS* gene showed that all the products that failed to suppress heterocyst formation contained a missense mutation affecting one of the last five codons of the *patS* open reading frame (Yoon & Golden, 1998). Moreover, the exogenous addition of a synthetic pentapeptide, RGSGR, corresponding to the last five amino acids encoded by the *patS* gene, but not the one corresponding to the last four, also inhibits heterocyst differentiation. Thus, a processed form of PatS containing at least the last five amino acids may be the diffusible inhibitory signal that regulates heterocyst differentiation. However, the cells appear to be insensitive to the PatS-5 peptide when added after 8 hours of nitrogen deprivation, presumably because at that time the differentiating cells are already committed to differentiate (Yoon & Golden, 2001). Different versions of the *patS* gene encoding peptides with all seventeen or the last eight, seven, six, five or four amino acids were placed downstream from different promoters in a replicative plasmid in the wild type or in a *patS* deletion background (Wu et al., 2004). PatS-5, PatS-8 and PatS-6 minigenes expressed in vegetative cells, under the *petE* promoter (which is active mainly in these cells), produce the strongest inhibitory activity. Synthetic PatS-5 and PatS-6 peptides have been demonstrated to inhibit the DNA-binding activity of HetR *in vitro* (Huang *et al.*, 2004) by directly binding to the protein (Feldmann *et al.*, 2011; Feldmann *et al.*, 2012).

The peptide is likely to remain inside the individual filaments, as no inhibitory signal is detected in the medium of a strain that overexpresses the *patS* gene (Yoon & Golden, 1998). When this synthetic peptide is added to a *patS* mutant strain in the medium, the Mch phenotype reverts, but the wild-type pattern is not restored, suggesting that the inhibitory signal must be produced in specific cells (Yoon & Golden, 1998). Transport of PatS from the producing to the neighbouring cells to form an inhibitory gradient has not been demonstrated yet. However, when confined to the cytoplasm of the cell that produces it, PatS is incapable of restoring a normal pattern of heterocysts to a *patS* mutant strain, suggesting that it must diffuse from cell to cell to function properly in cell patterning (Wu et al., 2004).

1.5.2. HetN

HetN is another negative regulator of the heterocyst differentiation process. It is not necessary for proper *de novo* pattern formation, but it is needed for stabilization and maintenance of the pattern over time. Overexpression of the *hetN* gene from an inducible *petE* promoter results in a complete suppression of heterocyst differentiation, and a *hetN* mutant has been reported to show a delayed Mch phenotype that does not normalize over time (Callahan & Buikema, 2001). PatS and HetN do not require one another for activity, and in a mutant of both genes almost all cells differentiate into heterocysts after several days of nitrogen deprivation, suggesting that PatS and HetN are the two predominant inhibitors of differentiation (Borthakur *et al.*, 2005).

The *hetN* gene (*alr5358*) was discovered after substitution of a wide genomic region encompassing this ORF by a transposon. The strain bearing this insertion exhibited a Het$^-$ phenotype, but was prone to turn into a strain with Mch phenotype (Black & Wolk, 1994). The sequence substituted comprised the genes *hetM* [later renamed as *hglB*, involved in the formation of the Hgl envelope (Bauer *et al.*, 1997)], *hetN* and *hetI*. A *hetN* mutant is not impaired in Hgl deposition, therefore the gene is not involved in glycolipid synthesis (Callahan & Buikema, 2001). Surprisingly, different substitutions in the 3' region of the *hetN* coding sequence result in a Mch phenotype, whereas substitutions in the 5' region lead to inhibition of differentiation. The overexpression of *hetN* in a replicative plasmid causes inhibition of differentiation; however the overexpression of a region comprising *hetN* plus *hetI* produces a phenotype similar to the wild-type phenotype. A *hetI* mutant could not be segregated, thus this gene may be essential for the cyanobacterium (Black & Wolk, 1994).

There is a basal level of *hetN* mRNA in the presence of combined nitrogen, but its expression is induced 12 hours after of N step-down (Bauer et al., 1997; Flaherty et al., 2011), about the time at which the cells are committed to differentiation. After this time, *hetN* is expressed exclusively in differentiating and mature heterocysts (Callahan & Buikema, 2001; Li *et al.*, 2002). The HetN protein has been reported to be located in cellular membranes (Li et al., 2002; Higa et al., 2012).

The *hetN* gene encodes a 287 amino-acid protein that presents homology with ketoacyl reductases and bears an internal ERGSGR peptide segment (the same as the last 6 amino acids of the PatS peptide). It remains unclear if the inhibition of the differentiation occurs through the PatS-like sequence or the reductase activity, since contrasting results have been published in this regard in the last years (Liu & Chen, 2009; Higa *et al.*, 2012). This PatS-6 sequence is only encoded in the *hetN* gene found in some of the sequenced

heterocyst-forming cyanobacteria, but it is still possible that the inhibitory activity is determined by a degenerated consensus sequence.

1.5.3. HetC

HetC is an essential element for the differentiation of heterocysts, since its mutation leads to arrest of differentiation at an early stage. This phenotype was first described as non-differentiating and non-fragmenting (Khudyakov & Wolk, 1997). Nevertheless, after prolonged incubation under nitrogen deprivation, some patterned cells, often paired, with diminished autofluorescence and increased expression of the *hetR* gene tend to appear (Khudyakov & Wolk, 1997; Xu & Wolk, 2001). These cells, however, do not lose the ability to divide as mature heterocysts do, and they are not formed if synthetic PatS-5 is added to the culture or in a *hetR* mutant background (Xu & Wolk, 2001). The expression of late genes, such as *hepB* and *devB,* involved in the differentiation process is down-regulated in these cells of a *hetC* mutant. However, the expression of *ftsZ,* that initiates the assembly of the division ring, is up-regulated, while the expression of early genes remains unaltered (Wang & Xu, 2005). Consequently, HetC has been thought to act during commitment, perhaps being involved in the regulation of cell division that takes place during heterocyst differentiation (Xu & Wolk, 2001).

The *hetC* gene (*alr2817*) is expressed ca. 3 hours after after N step-down, reaching its peak of expression after ca. 9 hours. An NtcA-dependent tsp located at -571 nt and another NtcA- and HetR-dependent tsp located at -293 nt have been identified (Muro-Pastor *et al.*, 2009). Strains bearing transcriptional fusions with the *lux* or the *gfp* genes show an increase in expression in proheterocysts and heterocysts. In a *hetR* mutant, the positive regulation is impaired, but the basal level of expression is considerably elevated (Khudyakov & Wolk, 1997; Muro-Pastor *et al.*, 1999). The *hetC* gene has been suggested to have positive autoregulation, as an insertion inside the gene does not abolish the expression but diminishes and delays it (Khudyakov & Wolk, 1997).

The HetC protein exhibits similarity to members of the family of ATP-binding cassette (ABC) exporters. ABC transporters can act as exporters, importers or both, but prokaryotic importers need a periplasmic binding protein to direct the substrate to the cytoplasmic membrane channel (Figure 1.7) (Rees *et al.*, 2009). HetC bears in addition a putative peptidase domain, which belongs to the C39-D class. Normally, domains belonging to this family are being part of big proteins that also contain, from the N-terminus, a variable region (in this case a cAMP binding domain), the peptidase domain, the transmembrane domain and the ATPase domain (Dirix et al., 2004). The C39-D class is exclusive for cyanobacteria and clusters out separately from other Gram-negative clusters

of the C39 family. The peptidase domains of the C39 family have been reported to be bacteriocin-processing peptidases (Khudyakov & Wolk, 1997; Michiels *et al.*, 2001; Rawlings *et al.*, 2012). They recognise a target region in the N-terminal side of the pre-peptide, cut it, and export the mature peptide outside of the cell. The target region in the substrate is LSX_2ELX_2IXGG for Gram-positive bacteria and $M(R/K)ELX_3E(I/L)X_2(I/V)XG(G/A)$ for Gram-negative bacteria (Dirix et al., 2004). Bacteriocin operons are commonly formed by a gene encoding the prebacteriocin, an ABC transporter, an immunity gene and an accessory protein essential for transport (Figure 1.7). Despite the name of the family, these proteins not only can transport bacteriocins but also signalling peptides, as in the case of the symbiotic system Btr of *Rhizobium etli* (Michiels et al., 2001).

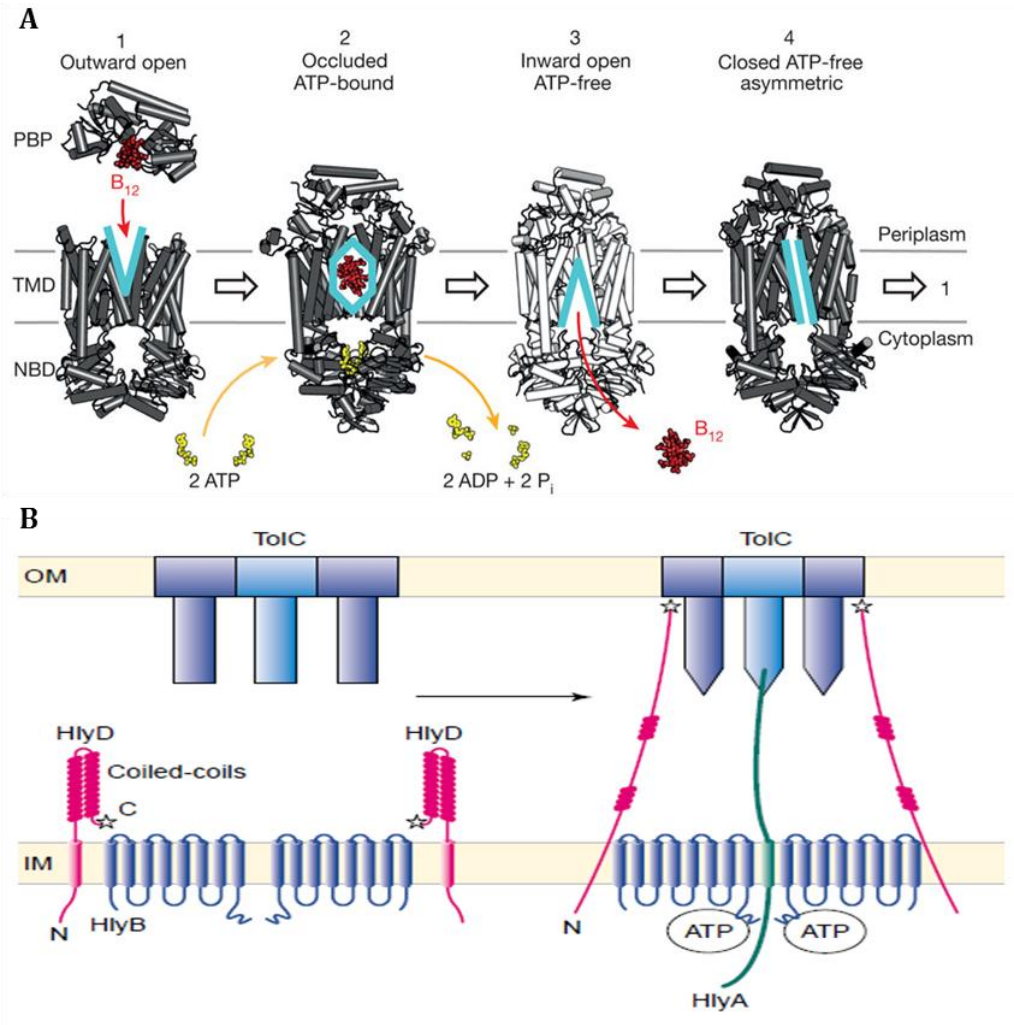


Figure 1.7. Mechanism of action of ABC importers and bacteriocin exporters. A. Example of an ABC importer mechanism, that of the Btu system for B12 vitamin import. The periplasmic binding protein (PBP), transmembrane domain (TMD) and the nucleotide-binding domain (NBD) are indicated. Image taken from Korkhov et al. (2012). B. Example of a bacteriocin secretion mechanism, that of the Hly system of hemolysin export in *E. coli*. The ABC transporter (HlyB), the accessory protein (HlyD), the outer membrane transporter (TolC) and the pre-peptide (HlyA) are indicated. Image taken from Gentschev et al. (2002).

1.5.4. HetP

Downstream from the *hetC* gene, the *hetP* (*alr2818*) gene is located. The mutation of the *hetP* gene leads to inhibition of heterocyst differentiation without fragmentation, as is the case in the *hetC* mutant. When a mutant strain bearing an insertion in the *hetP* gene or the wild-type strain was supplied with a DNA fragment containing *hetC*, *hetP*, *asr2819* and a truncated *alr2820* the resultant phenotype was Mch, frequently with heterocyst pairs of different size. However, incorporation of a DNA fragment containing only the *hetP* gene is toxic in both the mutant and the wild-type backgrounds (Fernández-Piñas *et al.*, 1994). The mutation of the *hetP* gene in *Nostoc ellipsosporum* leads to inhibition of heterocyst differentiation but not of akinete formation, in contrast to the *hetR* mutation that impairs both processes (Leganés *et al.*, 1994). The expression of a transcriptional *patS-gfp* fusion in a *hetP* mutant shows a normal pattern of green fluorescent cells. The overexpression of the *hetP* gene leads to Mch in a wild-type background, and partially bypasses the lack of *hetR* (Higa & Callahan, 2010).

A transcriptional fusion with the *gfp* allowed seeing expression of the *hetP* gene mostly in heterocysts at 24 hours after nitrogen deprivation. Five transcription start sites have been suggested at -727, -545, -208, -177 and -12 nt with respect to the translation start in the *hetP* promoter, being the -727 site dependent on HetR (Higa & Callahan, 2010).

1.5.5. OTHER PROTEINS INFLUENCING THE PATTERN OF HETEROCYSTS

The HetF protein influences heterocyst differentiation exerting a positive effect on *hetR* expression, and it may be also required for a proper localization of HetR in the (pro)heterocysts (Wong & Meeks, 2001). A *hetF* mutant is impaired in heterocyst differentiation, and overexpression of *hetF* leads to an Mch phenotype in the absence of combined nitrogen (Wong & Meeks, 2001). HetF presents homology with proteases and is required specifically in differentiating cells (Risser & Callahan, 2008).

The PatA protein is a response regulator with a PATAN domain, and its C-terminal region is similar to the CheY protein, which functions as a phosphorylation-activated switch (Liang *et al.*, 1992). A *patA* mutant forms heterocysts mostly at the ends of the filaments, even in a *hetR* overexpressing background, suggesting that *patA* acts downstream of *hetR* expression (Liang et al., 1992). It is possible that PatA influences heterocyst differentiation by attenuating the negative effects of PatS and HetN on HetR or preventing HetR autoproteolysis (Orozco *et al.*, 2006).

PatL is a pentapeptide repeat protein whose mutants have the same phenotype as the *patA* mutants, with mostly terminal heterocysts. In addition, according to yeast two hybrid assays both proteins might interact (Liu & Wolk, 2011).

HetL is composed almost entirely of 40 tandem pentapeptide repeats. The *hetL* gene was isolated in a genetic screen designed to identify genes involved in PatS signalling (Liu & Golden, 2002). Overexpression of *hetL* in a *patS* overexpression strain leads to formation of heterocysts by bypassing the PatS inhibitory signal. Overexpression of *hetL* in the wild-type background produces an Mch phenotype. Moreover, *hetL* overexpression induces partial heterocyst differentiation in an *ntcA* mutant. A *hetL* mutant shows normal heterocyst differentiation and diazotrophic growth, indicating that *hetL* plays a nonessential role in heterocyst differentiation (Liu & Golden, 2002). It has been suggested that HetL might not be an element acting in the process of heterocyst differentiation but its overexpression leading to an interaction with this process. However, it is also possible that HetL could perform a fine-tuning regulation by exerting a subtle protection of HetR from degradation.

The *hetY* gene is located upstream from *patS* in the genome. Its overexpression leads to a partial suppression of heterocyst differentiation, and its inactivation delays the differentiation of heterocysts, which moreover are aberrant (Golden & Yoon, 2003).

The *asr1734* gene is expressed only in proheterocysts and heterocysts after nitrogen step-down (Wu *et al.*, 2007). An *asr1734* mutant shows elevated levels of *ntcA* expression and a weak Mch phenotype. Overexpression of *asr1734* inhibits heterocyst differentiation in the wild-type, the *patS* mutant and the *hetR* $R^{223}W$ mutant backgrounds, suggesting that Asr1734 acts downstream of PatS and HetR (Wu et al., 2007).

PatN, identified in *Nostoc punctiforme*, is a membrane protein whose deletion leads to an increase in the number of heterocysts with short intervals of vegetative cells but no heterocyst doublets (Risser *et al.*, 2012). This phenotype was called Multiple singular heterocysts (Msh). The expression level of some genes involved in differentiation, including *hetF* or *patA*, is altered in the *patN* mutant. A PatN-GFP fusion localises to the cytoplasmic membranes, showing a singular pattern due to a biased inheritance, which means that PatN is not present in all the cells of the filament. PatN concentration likely depends on distribution of the protein after cell division, and PatN has been hypothesized to decrease the competency of a vegetative cell to become a heterocyst (Risser et al., 2012).

1.6. AIMS AND OBJECTIVES

This work aims to study the regulation of the formation of the heterocyst pattern in *Anabaena* sp. PCC 7120. Despite the large amount of published works on this subject, much still remains to be learned. The general plan in this thesis was to address the study of the relationships between different elements that affect the process of differentiation.

Research objectives:

1. To study the different regions in the PatS peptide, its putative processing and transference to neighbours of the producing cells.
2. To study the role of different regions of HetN and the localization of the protein.
3. To study the role of different regions of HetC and the localization of the protein, and the relationship between HetC and the negative elements PatS and HetN.
4. To study the localization of the HetP protein and the relationship between HetP, HetC and Asr2819, which are encoded by neighbouring genes.
5. To study *hetR* expression as a reporter of heterocyst differentiation.

The results of this thesis are based on the following publications:

Corrales-Guerrero, L., Mariscal, V., Flores, E. & Herrero, A. (2013) Functional dissection and evidence for intercellular transfer of the heterocyst-differentiation PatS morphogen. *Mol Microbiol* **88**: 1093-1105.

Corrales-Guerrero, L., Flores, E. & Herrero, A. Relationships between the ABC-exporter HetC and peptides regulating the spatiotemporal pattern of heterocyst distribution in *Anabaena*. Submitted.

Corrales-Guerrero, L., Tal, A., Arbel-Goren, R., Mariscal, V., Flores, E., Herrero, A. & Stavans, J. Spatial fluctuations in expression of the heterocyst differentiation regulatory gene *hetR* in nitrogen-replete *Anabaena* filaments. In preparation.

Corrales-Guerrero, L., Mullineaux, C., Flores, E. & Herrero, A. Subcellular localization and keys for the function of the HetN factor influencing heterocyst ditribution in *Anabaena* sp. In preparation.

2. MATERIALS AND METHODS

2.1. ORGANISMS AND CULTURE CONDITIONS

2.1.1. *Escherichia coli*

2.1.1.1. Strains

The *Escherichia coli* strains used in this work are shown in Table 2.1.

Table 2.1. E. coli strains used in this work.

Strain	Genotype	Reference
BL21	F- *ompT gal dcm lon hsdS$_B$(r_{B^-} m_{B^-}) λ(DE3[lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Studier & Moffatt (1986)
DH5α	F- *supE44 hsdR17 ($r_{K^-}m_{K^+}$) recA1 girA96 (Nal^r) endA1 thi-1 relA1 Δ(lacaya-argF) (Ø80lacZΔM15) U169*	Hanahan (1983)
ED8654	*Lac-3 o LacY1 supE44 supF58 hsdR514 ($r_{K^-}m_{K^-}$) recA56mcrA1 metB1 lacY galk2 galT22 trpR55*	Murray *et al.* (1977)
HB101	F- *hsdS20(r_B-m_B) leu supE44 ara14 galK2 lacY1 proA2 rpsL20 xyl-5 mtl-1 recA13 mcrB*	Boyer & Roulland-Dussoix (1969)

2.1.1.2. Culture conditions

Escherichia coli strains were grown in Luria-Bertani medium (LB), containing 10 g·L^{-1} NaCl, 10 g·L^{-1} tryptone and 5 g·L^{-1} yeast extract (Sambrook & Russell, 2001). For the preparation of solid media, LB medium was supplemented with Bactoagar (Difco) to a final concentration of 1.5% (w/v). The solid and liquid media were sterilized in autoclave. Transformation-competent cells of strains DH5α and BL21 were prepared by José Enrique Frías (Servicio de Cultivos Biológicos del Centro de Investigaciones Científicas Isla de la Cartuja).

Liquid cultures were incubated in orbital shakers at 200 rpm and 37 °C. The liquid volume was 3-5 mL for small cultures (10-mL tubes), and 500 mL to 3 L (1 L to 5 L flasks) for large-scale cultures. Solid cultures were grown in Petri dishes (approximately 20 mL of LB medium per dish) in an incubator at 37 °C. Cultures were supplemented with antibiotics (sterilized by filtration; dissolved in water except for chloramphenicol that was dissolved in ethanol) at standard concentrations: ampicillin (Ap), 50 μg·mL^{-1}; chloramphenicol (Cm), 30 μg·mL^{-1} in ethanol; streptomycin sulfate (Sm), 25 μg·mL^{-1}; spectinomycin dihydrochloride pentahydrate (Sp), 100 μg·mL^{-1}; and kanamycin sulfate (Km), 50 μg·mL^{-1}.

2.1.2. *Anabaena* sp. PCC 7120 and derivative strains

2.1.2.1. Strains

This work has been carried out with the filamentous heterocyst-forming cyanobacterium *Anabaena* sp. (also known as Nostoc sp.) PCC 7120 (ATCC 27893) (hereafter *Anabaena*), which belongs to section IV of the taxonomic classification of Rippka *et al.* (1979), and derivative mutant strains (described in Table 2.2 and Table 2.3).

Table 2.2. Mutant strains of *Anabaena* not generated in this work.

Strain	Genotype	Reference
CSM1	*hetC-gfpmut2-pCSV3*	M. León, unpublished
CSSC2	*hetR*	Camargo *et al.* (2012)
CSVM17	P_{patS}-*gfpmut2* (α plasmid)	Mariscal *et al.* (2007)
CSVT20	*patS*	Corrales-Guerrero *et al.* (2013)

Table 2.3. Mutant strains of *Anabaena* generated in this work.

Strain	Genotype	Resistance to antibiotics
CSL1	*patS hetC::C.S3*	SmSp
CSL3	*hetC::C.S3*	SmSp
CSL7	*hetN*	None
CSL11	*patS hetN*	None
CSL12	*hetC::C.S3 hetN*	SmSp
CSL15	*patS hetC::C.S3 hetN*	SmSp
CSL16	*hetC-p* (peptidase domain)	None
CSL17	*hetC-p patS*	None
CSL19	*patS*::pCSL30 (M·GFP·KAIMLVNFCDERGSGR)	SmSp
CSL20	*patS*::pCSL32 (MKAIM·GFP·LVNFCDERGSGR)	SmSp
CSL21	*patS*::pCSL34 (MKAIMLVNFCDERGSGR·GFP)	SmSp
CSL22	*patS*::pCSL36 (MKAIM·6xHis·LVNFCDERGSGR)	SmSp
CSL23	*patS*::pCSL38 (MKAIMLVNFCDERGSGR·6His)	SmSp
CSL30	*hetC-p hetN*	None
CSL31	*patS hetC-p hetN*	None
CSL32	*hetN* (ERGSGR)	None
CSL33	*hetC-p-gfpmut2::pCSV3*	SmSp
CSL44	*patS*::pCSL77 (AKAIMLVNFCDERGSGR)	SmSp
CSL45	*patS*::pCSL78 (MRGSGR)	SmSp
CSL46	*patS*::pCSL79 (MERGSGR)	SmSp
CSL47	*patS*::pCSL80 (MCDERGSGR)	SmSp
CSL48	*patS*::pCSL81 (MKAIMLVNFCDERGSGR)	SmSp
CSL49	*patS*::pCSL82 (MKAIALVNFCDERGSGR)	SmSp
CSL50	*patS*::pCSL83 (MKAIMQVNFCDERGSGR)	SmSp
CSL51	*patS*::pCSL84 (MKAIMLQNFCDERGSGR)	SmSp
CSL52	*patS*::pCSL85 (MKAIMLVEFCDERGSGR)	SmSp
CSL53	*patS*::pCSL86 (MKAIMLVNQCDERGSGR)	SmSp
CSL54	*patS*::pCSL87 (MKAIMLVNFADERGSGR)	SmSp
CSL55	*patS*::pCSL88 (MKAIMLVNFCAERGSGR)	SmSp
CSL56	*patS*::pCSL89 (MKAIMLVNFCDARGSGR)	SmSp
CSL57	*patS*::pCSL90 (MKAIMLVNFCDEAGSGR)	SmSp
CSL58	*patS*::pCSL92 (MKAIMLVNFCDERGAGR)	SmSp
CSL59	*patS*::pCSL91 (MKAIMLVNFCDERAGSGR)	SmSp

CSL60	*patS*::pCSL93 (MKAIMLVNFCDERGSAR)	SmSp
CSL61	*patS*::pCSL94 (MKAIMLVNFCDERGSGA)	SmSp
CSL62	*patS*::pCSL95 (MKAIVLQNFCDERGSGR)	SmSp
CSL64	*hetR-gfpmut2*::pCSV3	SmSp
CSL65	*hetR-gfpmut2*::pCSV3 *patS*	SmSp
CSL67	*hetP-sfgfp*::pCSV3	SmSp
CSL68	*hetP-sfgfp*::pCSV3 *hetC-p*	SmSp
CSL69	*hetP-sfgfp*::pCSV3 *patS*	SmSp
CSL70	*hetP-sfgfp*::pCSV3 *hetN*	SmSp
CSL71	*hetN-sfgfp*::pCSV3	SmSp
CSL72	*hetN-sfgfp*::pCSV3 *patS*	SmSp
CSL73	sp*hetN*	SmSp
CSL86	*hetR-gfpmut2*::pCSV3 *hetR*	SmSp
CSL87	*hetR-gfpmut2*::pCSV3 *patS hetC-p*	SmSp
CSL88	*hetR-gfpmut2*::pCSV3 *hetC-p*	SmSp
CSL90	*patS*::pCSL103 (MAAIMLVNFCDERGSGR)	SmSp
CSL91	*patS*::pCSL104 (MKQIMLVNFCDERGSGR)	SmSp
CSL92	*patS*::pCSL105 (MKAQMLVNFCDERGSGR)	SmSp
CSL93	*patS*::pCSL106 (MKAIMLANFCDERGSGR)	SmSp
CSL94	*patS*::pCSL110 (MDERGSGR)	SmSp
CSL95	*C.K3::asr2819* (α plasmid)	Nm
CSL96	*C.K3::asr2819* (α plasmid) *hetC-p*	Nm
CSL97	*asr2819*	None
CSL98	*asr2819 hetC-p*	None
CSL99	*asr2819::C.S3*	SmSp
CSL100	*asr2819::C.S3* p*hetC*	SmSp
CSL101	*hetC-p patS*::pCSL80	SmSp
CSL102	*hetC-p patS*::pCSL81	SmSp
CSL103	*hetN-sfgfp*::pCSV3 sp*hetN*	SmSp
CSL105	*asr2819 hetN*	None
CSL106	*patS asr2819 hetN*	None
CSL107	*hetP-gfpmut2*::pCSV3	SmSp
CSL108	*hetN-gfpmut2*::pCSV3	SmSp

2.1.2.2. Culture conditions

Anabaena strains were grown under photoautotrophic axenic conditions in $BG11_0$ medium (Rippka *et al.*, 1979). This medium contains 0.2 mM Na_2CO_3, 0.3 mM $MgSO_4$, 0.24 mM $CaCl_2$, 0.2 mM K_2HPO_4, 28.5 µM citric acid, iron(III) 6 mg/L citrate hydrate (19% Fe), 2.4 µM Na_2-EDTA, 46 µM H_3BO_3, 9.1 µM $MnCl_2$, 1.6 µM Na_2MoO_4, 0.8 µM $ZnSO_4$, 0.3 µM $CuSO_4$ and 0.2 µM $CoCl_2$. Growth media were prepared from a concentrated 100x solution containing all the components but K_2HPO_4, which was added just before sterilizing by autoclave. Ammonium chloride was used as a nitrogen source at a concentration of 4 mM for solid media and flasks or 6-10 mM for bubbled cultures. Media with NH_4Cl were buffered with TES-NaOH (pH 7.5) at a final concentration twice that of NH_4^+. NH_4Cl and TES-NaOH solutions were sterilized by filtration separately and added to the medium after its sterilization in autoclave.

Liquid cultures were incubated in 50-mL (for 25 mL of culture) or 100-mL (for 50 mL of culture) flasks in an orbital incubator set at 100 rpm in a 30 °C chamber, and illuminated from above at 30 $\mu E \cdot m^{-2} \cdot s^{-1}$.

Unless stated otherwise, the experiments were carried out with bubbled cultures after growth in flasks. Cultures were grown in 50-mL tubes (25-mL cultures), 240-mL bottles (up to 200-mL of cultures) or 1-L bottles (up to 1-L of cultures). They were illuminated laterally (40 $\mu E \cdot m^{-2} \cdot s^{-1}$) and bubbled with a sterile mixture of air/CO_2 (99:1) that additionally allowed for a correct agitation of the culture. Bubbled cultures were supplemented with $NaHCO_3$ (final concentration, 8-10 mM) (named $BG11_0C$ or $BG11_0C+NH_4^+$, depending on the nitrogen source).

Solid medium was prepared in Petri dishes by adding 1% (w/v) Bacto-Agar (Difco) sterilized in autoclave separately from the nutrients. Inoculated Petri dishes were incubated at 30 °C and illuminated from above at 15-30 $\mu E \cdot m^{-2} \cdot s^{-1}$.

When required, antibiotics were added to the media at the following final concentrations: neomycin sulfate (Nm), 5-50 $\mu g \cdot mL^{-1}$; streptomycin sulfate and spectinomycin dihydrochloride pentahydrate, 2 $\mu g \cdot mL^{-1}$ for liquid and 5 $\mu g \cdot mL^{-1}$ for solid cultures. Antibiotics were sterilized by filtration.

2.1.2.3. Methods for collecting cells

Cells were collected by filtration through 0.45 µm pore diameter filters with a vacuum system. For N step-down experiments, ammonium-grown cells were collected, washed and resuspended with $BG11_0$ or $BG11_0C$ medium.

For RNA isolation, cells corresponding to 100 mL cultures were collected by filtration, washed with 4 mL $T_{50}E_{100}$ buffer (50 mM Tris-HCl, 100 mM EDTA-Na_2, pH 8) and resuspended in 4 mL of the same buffer. Cell suspensions were transferred to two Eppendorf tubes (2 mL each), centrifuged again and the resulting pellets were frozen with liquid nitrogen and kept at -20 ºC.

2.2. MANIPULATION AND ANALYSIS OF DNA

2.2.1. Plasmids used in this work

The plasmids used in this work are described in Table 2.4 and Table 2.5.

Table 2.4. Plasmids not generated in this work.

Name	Resistance	Description	Reference
pCSAL33	Ap	*gfpmut2* in a modified pMBL-T (without NheI site)	A. López-Lozano, unpublished
pCSAL39	Km	*sfgfp* with 4xGly linker (N-terminal) and BsaI site in a modified pMBL-T (Ap resistance substituted by CK.1 cassette)	A. López-Lozano, unpublished
pCSAM135	Ap	*sepJ-gfp-mut2*, cloned in pMBL-T	Flores *et al.* (2007)
pCSAV76	Ap	*nucAnuiA* genes and C.K3 cassette cloned in a pBR322 derived	A. Valladares, unpublished

pCSBN1	Nm	pRL278 derived with polylinker of pCSV3	B. Nocea, unpublished
pCSRO	SmSp	Plasmid derived from pRL278, with Km resistance substituted by C.S3	Merino-Puerto *et al.* (2010)
pCSV3	SmSp	Plasmid derived from pRL500, with Ap resistance substituted by C.S3	Olmedo-Verd *et al.* (2006)
pCSVT51	SmSp	Flanking DNA sequences of *patS* in pCSRO to construct the deletion mutant	Corrales-Guerrero et al. (2013)
pMA-T	Ap	Commercial vector for cloning purposes	LifeTechnologies
pMBL-T	Ap	Commercial vector for cloning purposes	Dominion MBL
pRL278	Km	Vector used for the positive selection of double recombinants in *Anabaena*	Black *et al.* (1993)
pRL443	ApTc	Conjugative plasmid, mobilize ColE1 derived plasmids	Elhai & Wolk (1988)
pRL623	Cm	Derived from ColK, helper plasmid for the conjugation	Elhai *et al.* (1997)
pSPARK	Ap	Commercial vector for cloning purposes	Canvax

Table 2.5. Plasmids generated in this work.

Name	Resistance	Description
pCSL18	Ap	PCR product obtained with primers alr2817-34, alr2817-35, alr2817-36 and alr2817-37 (overlapping PCR), cloned in pMBL-T
pCSL19	ApSmSp	C.S3 inserted between the two DNA fragments of pCSL18 with EcoRV
pCSL20	KmSmSp	pCSL19 digested with BamHI and XhoI, and cloned in pRL278
pCSL22	Ap	PCR product obtained with primers alr5358-1, alr5358-2, alr5358-3 and alr5358-4 (overlapping PCR), cloned in pMBL-T
pCSL23	Km	pCSL22 digested with BamHI and XhoI and cloned in pRL278
pCSL24	Ap	PCR product obtained with primers alr2817-38, alr2817-42, alr2817-43 and alr2817-41 (overlapping PCR), cloned in pMBL-T
pCSL25	SmSp	pCSL24 digested with SpeI and XbaI and cloned in pCSRO (XbaI)
pCSL29	Ap	PCR product obtained with primers asl2301-9, asl2301-15, gfp-asl2301-1 and gfp-asl2301-2 (overlapping PCR), cloned in pMBL-T
pCSL30	SmSp	pCSL29 digested with KpnI and cloned in pCSV3
pCSL31	Ap	PCR product obtained with primers asl2301-9, asl2301-16, gfp-asl2301-3 and gfp-asl2301-4 (overlapping PCR), cloned in pMBL-T
pCSL32	SmSp	pCSL31 digested with KpnI and cloned in pCSV3
pCSL33	Ap	PCR product obtained with primers asl2301-9, asl2301-17, gfp-asl2301-5 and gfp-asl2301-6 (overlapping PCR), cloned in pMBL-T
pCSL34	SmSp	pCSL33 digested with KpnI and cloned in pCSV3
pCSL35	Ap	PCR product obtained with primers asl2301-9 and asl2301-18, cloned in pMBL-T
pCSL36	SmSp	pCSL35 digested with KpnI and cloned in pCSV3
pCSL37	Ap	PCR product obtained with primers asl2301-9 and asl2301-19, cloned in pMBL-T
pCSL38	SmSp	pCSL37 digested with KpnI and cloned in pCSV3
pCSL41	Ap	PCR product obtained with primers alr5358-11, alr5358-9, alr5358-10 and alr5358-12 (overlapping PCR), cloned in pMBL-T
pCSL42	Km	pCSL41 digested with BamHI and XhoI and cloned in pRL278
pCSL46	Ap	PCR product obtained with primers asl2301-9 and asl2301-32, cloned in pMBL-T
pCSL47	Ap	PCR product obtained with primers asl2301-9 and asl2301-33, cloned in pMBL-T
pCSL48	Ap	PCR product obtained with primers asl2301-9 and asl2301-34, cloned in pSPARK

pCSL49	Ap	PCR product obtained with primers asl2301-9 and asl2301-35, cloned in pMBL-T
pCSL50	Ap	PCR product obtained with primers asl2301-9 and asl2301-36, cloned in pMBL-T
pCSL51	Ap	PCR product obtained with primers asl2301-9 and asl2301-37, cloned in pMBL-T
pCSL52	Ap	PCR product obtained with primers asl2301-9 and asl2301-38, cloned in pMBL-T
pCSL53	Ap	PCR product obtained with primers asl2301-9 and asl2301-39, cloned in pSPARK
pCSL54	Ap	PCR product obtained with primers asl2301-9 and asl2301-40, cloned in pMBL-T
pCSL55	Ap	PCR product obtained with primers asl2301-9 and asl2301-41, cloned in pMBL-T
pCSL56	Ap	PCR product obtained with primers asl2301-9 and asl2301-42, cloned in pSPARK
pCSL57	Ap	PCR product obtained with primers asl2301-9 and asl2301-43, cloned in pMBL-T
pCSL58	Ap	PCR product obtained with primers asl2301-9 and asl2301-44, cloned in pMBL-T
pCSL59	Ap	PCR product obtained with primers asl2301-9 and asl2301-45, cloned in pMBL-T
pCSL60	Ap	PCR product obtained with primers asl2301-9 and asl2301-46, cloned in pMBL-T
pCSL61	Ap	PCR product obtained with primers asl2301-9 and asl2301-47, cloned in pMBL-T
pCSL62	Ap	PCR product obtained with primers asl2301-9 and asl2301-48, cloned in pMBL-T
pCSL63	Ap	PCR product obtained with primers asl2301-9 and asl2301-49, cloned in pMBL-T
pCSL64	Ap	PCR product obtained with primers asl2301-9 and asl2301-40, cloned in pMBL-T. An additional spontaneous mutation occurred during the PCR process.
pCSL66	Ap	PCR product obtained with primers alr2339-12 and alr2339-36, cloned in pSPARK
pCSL67	Ap	pCSL66 digested with XhoI and NheI and cloned in pCSAL33 digested with the same enzymes
pCSL68	SmSp	pCSL67 digested with KpnI and cloned in pCSV3 digested with the same enzyme
pCSL69	Km	PCR product obtained with primers alr2818-8 and alr2818-9, cloned in pCSAL39 digested with BsaI and HindIII
pCSL70	SmSp	pCSL69 digested with KpnI and cloned in pCSV3 digested with the same enzyme
pCSL71	Km	PCR product obtained with primers alr5358-18 and alr5358-19, cloned in pCSAL39 digested with BsaI and HindIII
pCSL72	SmSp	pCSL71 digested with KpnI and cloned in pCSV3 digested with the same enzyme
pCSL73	Km	PCR product obtained with primers alr5358-20, alr5358-13, alr5358-14 and alr5358-21 (overlapping PCR), cloned in pRL278 digested with BamHI and XhoI
pCSL77	SmSp	pCSL46 digested with KpnI and cloned in pCSV3
pCSL78	SmSp	pCSL47 digested with KpnI and cloned in pCSV3
pCSL79	SmSp	pCSL48 digested with StuI and HincII and cloned in pCSV3 (HincII)
pCSL80	SmSp	pCSL49 digested with KpnI and cloned in pCSV3
pCSL81	SmSp	pCSL50 digested with KpnI and cloned in pCSV3
pCSL82	SmSp	pCSL51 digested with KpnI and cloned in pCSV3
pCSL83	SmSp	pCSL52 digested with KpnI and cloned in pCSV3
pCSL84	SmSp	pCSL53 digested with StuI and HincII and cloned in pCSV3 (HincII)
pCSL85	SmSp	pCSL54 digested with KpnI and cloned in pCSV3
pCSL86	SmSp	pCSL55 digested with KpnI and cloned in pCSV3
pCSL87	SmSp	pCSL56 digested with StuI and HincII and cloned in pCSV3 (HincII)
pCSL88	SmSp	pCSL57 digested with KpnI and cloned in pCSV3

pCSL89	SmSp	pCSL58 digested with KpnI and cloned in pCSV3
pCSL90	SmSp	pCSL59 digested with KpnI and cloned in pCSV3
pCSL91	SmSp	pCSL60 digested with KpnI and cloned in pCSV3
pCSL92	SmSp	pCSL61 digested with KpnI and cloned in pCSV3
pCSL93	SmSp	pCSL62 digested with KpnI and cloned in pCSV3
pCSL94	SmSp	pCSL63 digested with KpnI and cloned in pCSV3
pCSL95	SmSp	pCSL64 digested with KpnI and cloned in pCSV3
pCSL103	SmSp	pCSL107 digested with PstI and cloned in pCSV3 digested with the same enzymes
pCSL104	SmSp	PCR product obtained with primers asl2301-69 and asl2301-71, cloned in pCSV3 digested with PstI
pCSL105	SmSp	PCR product obtained with primers asl2301-69 and asl2301-72, cloned in pCSV3 digested with PstI
pCSL106	SmSp	pCSL108 digested with PstI and cloned in pCSV3 digested with the same enzymes
pCSL107	Ap	Synthetic DNA fragment from 2771446 to 2771734 of the *Anabaena* genome comprising substitution K^2A and PstI restriction site (GeneArt, LifeTechnologies) cloned in pMA-T
pCSL108	Ap	Synthetic DNA fragment from 2771446 to 2771734 of the *Anabaena* genome comprising substitution V^7A and PstI restriction site (GeneArt, LifeTechnologies) cloned in pMA-T
pCSL109	Ap	PCR product obtained with primers asl2301-9 and asl2301-77, cloned in pMBL-T
pCSL110	SmSp	pCSL109 digested with PstI and cloned in pCSV3 digested with the same enzymes
pCSL111	SmSp	PCR product obtained with primers asr2818-13, asr2819-1, asr2819-2 and alr2820-1 (overlapping PCR), cloned in pCSRO digested with SacI
pCSL112	KmSmSp	pCSL113 with C.S3 inserted between the PCR fragments (EcoRV)
pCSL113	Km	PCR product obtained with primers asr2818-13, asr2819-1, asr2819-2 and alr2820-1 (overlapping PCR), cloned in pCSBN1 digested with SacI
pCSL114	Km	PCR product obtained with primers asr2819-3 and asr2819-4, cloned in pCSAV76 digested with BglII
pCSL123	SmSp	PCR product obtained with primers alr2818-19, GFP-13, GFP-14 and GFP-11 (overlapping PCR), cloned in pCSV3 digested with KpnI
pCSL124	SmSp	PCR product obtained with primers alr5358-26, GFP-13, GFP-14 and GFP-11 (overlapping PCR), cloned in pCSV3 digested with KpnI

2.2.2. DNA isolation

2.2.2.1. Plasmidic DNA from *E. coli*

Plasmidic DNA was isolated using the alkaline lysis method (Sambrook *et al.*, 1989) following the general protocol, with some exceptions: RNAse was added to Solution I and the purified DNA was resuspended in 30 µL of milliQ water. When highly purified DNA was required, the *Nucleospin Plasmid* kit (Macherey Nagel) was used.

2.2.2.2. Total DNA from *Anabaena*

Cells of *Anabaena* were harvested (10 to 50 mL of liquid culture) and resuspended in a final volume of 500 µL with $T_{1/10}$ E (10 mM Tris-HCl, 0.1 mM EDTA-Na_2, pH 8) in an Eppendorf tube. Then, 150 µL of sterile glass beads (0.25-0.30 mm diameter, previously washed with acid and milliQ water, and sterilized at 180 °C), 20 µL of 10% sodium dodecyl

sulfate and 450 µL of phenol-chloroform (1:1, v/v) were added. The mixture was shaken in a vortex for 1 min and incubated on ice for 1 min. This process was repeated 3 times. The resulting lysate was centrifuged at 16,100 x *g* for 15 min, and the resulting clear supernatant solution was transferred to a new Eppendorf tube for successive extractions with 1 volume of phenol, phenol:chloroform, and twice with chloroform. Then DNA was precipitated with 2.5 volumes of ethanol and 0.1 volumes of 3 M potassium acetate (pH 5.2) incubating for 1 h at -20 °C. After a 15-min centrifugation (as above), the pelleted DNA was washed with 70% ethanol (v/v), dried and resuspended with 30 µL of milliQ water.

2.2.3. Estimation of the concentration of DNA

The concentration of DNA was determined using a *NanoDrop Spectrophotometer* ND-1000. The absorbance was measured at 260 nm (extinction coefficient ε = 0.020 $mg^{-1}{\cdot}mL{\cdot}cm^{-1}$). The DNA concentration was also estimated by comparison with samples of known concentration, after running in agarose gel electrophoresis.

2.2.4. DNA electrophoresis in agarose gels

DNA fragments were separated by agarose gel electrophoresis (Sambrook & Russell, 2001). Gels were prepared with 0.8% agarose (w/v) in TBE buffer (90 mM Tris-Borate, 2 mM EDTA-Na_2, pH 8), supplemented with 3 µL (for 100 mL gels) of *GelRed Nucleic Acid* Gel Stain (Biotium). 0.1 volumes of 10x sample buffer (50% glycerol [v/v], 0.4% bromophenol blue [w/v] and 0.4% xylene cyanol FF [w/v]) were added to the sample before loading it into the gel. The size standard used was λ phage DNA digested with ClaI or the commercial *1kb ladder* (Labotaq). The electrophoresis was carried out using a *Mini-Sub Cell GT* or *Wide Mini-Sub Cell GT* (Bio-Rad), and the DNA was visualized with ultraviolet light on the *Bio-Rad Gel-Doc XR* system using the *Bio-Rad Quantity One* 4.6.2 software.

2.2.5. Purification of DNA fragments

DNA isolation from gels or from liquid solutions was carried out using the commercial kit *Gel Band Purification Kit* (GE Healthcare), using the protocol provided.

2.2.6. PCR

Polymerase Chain Reaction (PCR) was carried out in a *Tpersonal* (Whatman-Biometra) thermal cycler using *Biotaq DNA Polymerase* (for routine purposes) and *iProof Polymerase* (high fidelity polymerase for cloning purposes) following the manufacturer's instructions.

For cloning purposes, the Taq9-10 Polymerase was also used. The reaction mixture contained in 50 µL of final volume, 1-5 ng of DNA, 0.2 mM dNTPs, 20 pmol of primers and 2.25 U of DNA polymerase in the appropriate buffer.

2.2.7. Oligodeoxynucleotides used in this work

The oligodeoxynucleotide primers used in this work are described in Table 2.6.

Table 2.6. Oligodeoxynucleotides used in this work.

Name	Sequence 5´-3´
alr2339-12	AACTCTGGACTTCTGGCT
alr2339-36	GCTAGCACCTCCACCGCCCTTGATCAGATCGATGT
alr2339-37	CGTCACTTTTGGGGCAATTGCTT
alr2339-4	GGCACACTTAGCCGCCGTTGCTGC
alr2817-13	GTTTCAAGAACGCAACG
alr2817-14	GATATCAGCTAAGTGGTGATAAAG
alr2817-26	CGACTCAATGATGTTTATGCG
alr2817-27	TGACGACTGACAACTGACTA
alr2817-29	GACACAAGCAGTAGTAGAGC
alr2817-34	TCTCTCTTGGGTGGGATTCT
alr2817-35	ATCGCCATTAGTTCCGATATCTTTGCAGCCCTCCCT
alr2817-36	AGGGAGGGCTGCAAAGATATCGGAACTAATGGCGAT
alr2817-37	CTCGCCTACTGTGCATTTGAGACT
alr2817-38	TGGTATCGGCTGAACAAA
alr2817-41	TGTCGCCACCCCAGTCATAAT
alr2817-42	CTTGAGGGCTTGAAAACTTTCTGGGTAGGAGTAGGACGA
alr2817-43	TCGTCCTACTCCTACCCAGAAAGTTTTCAAGCCCTCAAG
alr2817-44	TTGCGGTCTACCTGGGCTAAATTC
alr2818-10	CACGTCTGGTTTCAAGGTAG
alr2818-13	CATGCTGAGCTCTGCTGGGTACAATCCTATGC
alr2818-16	TAAATACTCTTGGGCTTGTGTTCT
alr2818-17	CTAGCTTCGGAGTTTTCTTTGAGT
alr2818-19	AAAGTAGGTACCACAAGCCAAGCATGAAAGTGGT
alr2818-6	GGGGCAGTTAATAGATGAAG
alr2818-8	AAAGTAAAGCTTACAAGCCAAGCATGAAAGTGGT
alr2818-9	CCGAAGGGTCTCACGCCATTATGAATAAAATC
alr2820-1	CTTCTTGAGCTCTTACAGGAGAAGCCAGTCCAGGTT
alr2820-2	GAAAGCGATCGCTCCGATAAACTC
alr5358-1	TTATACACCTTGCGTCCCTTCCTC
alr5358-10	CTACCCAGCATGATGATTGTCAATATTGCT
alr5358-11	TCAAGCAAGTGCAGTCTACG
alr5358-12	GGGTGACGGCATCGCTAAAT
alr5358-13	TACCGTTGCCTGTTCTTTCATTGTAACCTGCTAGTC
alr5358-14	GACTAGCAGGTTACAATGAAAGAACAGGCAACGGTA
alr5358-16	GACCACTACCGCGTTCCAT
alr5358-17	TAAAGAAGCAATATTGACAATCAT
alr5358-18	AACCGGAAGCTTACGCGGTCTTGGTGTCTACATTG
alr5358-19	AAGCAAGGTCTCACGCCTGAGCGATGAGACTC
alr5358-2	CTCAACAGCTACATAGCGTGAAGCGCCGGT
alr5358-20	GTAGATGGATCCTTATACACCTTGCGTCCCTTCCTC
alr5358-21	TTTTGGCTCGAGTCCACTACACCGAACCAACGAT
alr5358-26	AACCGGGGTACCACGCGGTCTTGGTGTCTACATTG
alr5358-3	ACCGGCGCTTCACGCTATGTAGCTGTTGAG

alr5358-4	TGAAGTTCATCTCTGGCGCATTCC
alr5358-5	TGACTGGGTACGTTCCTTTGGCTA
alr5358-6	GCCTTAACCGCATTACACGTTTGG
alr5358-9	AGCAATATTGACAATCATCATGCTGGGTAG
asl2301-10	TTAGAGCGCAGTAGGGAAAAG
asl2301-11	GATGCAATTGGGTATGGCGTTGATGT
asl2301-15	TTCTTCTCCTTTACTCATAATTGCCTTCAT
asl2301-16	TTCTTCTCCTTTACTTCTACCACTACCGCG
asl2301-17	TTCTTCTCCTTTACTCATAATCTTAAAATC
asl2301-18	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTAAGTGGTGGTGGTGGTGGTGCAT AATTGCCTTCATAATCTTAAAATC
asl2301-19	CTAATGATGATGATGATGATGTCTACCACTACCGCGCTC
asl2301-32	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTAACATAATTGCCTTCGCAA
asl2301-33	CTATCTACCACTACCGCGCATAATCTTA
asl2301-34	CTATCTACCACTACCGCGCTCCATAATCTTA
asl2301-35	CTATCTACCACTACCGCGCTCATCACACATAATCTTAAAATCG
asl2301-36	CTATCTACCACTACCGCGCTC
asl2301-37	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTAACGCAATTGCC
asl2301-38	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTTGCATAAT
asl2301-39	CTATCTACCACTACCGCGCTCATCACAGAAATTCTGTAACA
asl2301-40	CTATCTACCACTACCGCGCTCATCACAGAATTCCACTA
asl2301-41	CTATCTACCACTACCGCGCTCATCACACTGATTCA
asl2301-42	CTATCTACCACTACCGCGCTCATCAGCGAAAT
asl2301-43	CTATCTACCACTACCGCGCTCTTGACAGA
asl2301-44	CTATCTACCACTACCGCGCGCATC
asl2301-45	CTATCTACCACTACCGGCCTCATCA
asl2301-46	CTATCTACCACTAGCGCGCTCA
asl2301-47	CTATCTACCAGCACCGCGCTCA
asl2301-48	CTATCTAGCACTACCGCGCTCA
asl2301-49	CTATGCACCACTACCGCGCTCA
asl2301-69	TGCGCTGCAGCTGGGCAGTTAGTACGGAAAGT
asl2301-71	TTACACTGCAGCTATCTACCACTACCGCGCTCATCACAGAAATTCACTAACATAATTTGC TTCAT
asl2301-72	TTACACTGCAGCTATCTACCACTACCGCGCTCATCACAGAAATTCACTAACATTTGTGCC TT
asl2301-74	CAGGGTCAGTATTGTTTCGG
asl2301-75	CACCATTCAATTGCACCATCAG
asl2301-76	TCGGTGAATTACTTTTCAACAG
asl2301-77	CTATCTACCACTACCGCGCTCATCCATAATCTTAAAATC
asl2301-8	CATACCCAATTGCATCTTCTCATCATCAGTCTT
asl2301-9	CTGGGCAGTTAGTACGGAAAGT
asl2301-gfp-1	ATGAAGGCAATTATGAGTAAAGGAGAAGAA
asl2301-gfp-2	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTAATTTGTATAGTTCATC
asl2301-gfp-3	CGCGGTAGTGGTAGAAGTAAAGGAGAAGAA
asl2301-gfp-4	TTATTTGTATAGTTCATCCAT
asl2301-gfp-5	GATTTTAAGATTATGAGTAAAGGAGAAGAA
asl2301-gfp-6	CTATCTACCACTACCGCGCTCATCACAGAAATTCACTAACATAATTGCCTTTTTGTATAG TTCATCCATGCC
asr2819-1	CTAAATCCCCTTTACGATATCATTCGAGAAGTTGCA
asr2819-2	TGCAACTTCTCGAATGATATCGTAAAGGGGATTTAG
asr2819-3	AACTTCAGATCTGTGAGTCTTGACACC
asr2819-4	TCTGACAGATCTCTAAATCCCCTTTAC
asr2819-5	TGCGTTTTGCGGTATTGCGA
asr2819-11	TTATTCCCGTCACTTACCACCATC
asr2819-12	TTCTAACTCAAGCACCACAAACTC

C.S3-1	GGATGACCTTTGAATGACC
C.K1-2	GGGATCTCATGCTGGAGT
GFP-4	CAAGAATTGGGACAACTCC
GFP-11	TCTGGTACCTTATTTGTATAGTTC
GFP-13	TTCTCCTTTGCTAGCACCTCCA
GFP-14	TGGAGGTGCTAGCAAAGGAGAA
nuc-nui-5	GATCTGCACCTCCTGTTAC
nuc-nui-6	AGCCAAGTTTTATCATCTAT
rnpB-4	ACTCTTGGTAAGGGTGCAAAGGTG
rnpB-5	AACCATAGTTCCTTCGGCCTTGCT

2.2.8. DNA sequencing

DNA sequencing was carried out in the Secugen S.L. Company service (Ramiro de Maeztu, 9 Madrid) using PCR products or isolated plasmids.

2.2.9. Enzymatic treatments of DNA

2.2.9.1. Restriction

Restriction digestions were conducted for 30 min to 2 h using endonucleases supplied by New England Biolabs, Roche or Amersham (currently GE HealthCare), with the manufacturer's recommended buffers and temperatures.

2.2.9.2. Ligation

The ligation of different fragments of DNA was carried out using the *T4 bacteriophage DNA ligase* (Dominion MBL). Reactions were performed at 21 °C for 1-24 h in 10 µL final volume. The vector:insert concentration ratio used was approx. 1:3, or 1:5 for ligations involving plasmids derived from pMBL-T.

2.2.9.3. Dephosphorylation

Dephosphorylation of 5' protruding extremes was made to prevent re-circularization of linearized vectors. DNA was incubated with *rAPid alkaline phosphatase* (Roche) following the manufacturer´s instructions.

2.2.10. Radioactive labelling of DNA fragments

This process was performed using the random primer method with *Ready-to-Go DNA Labelling Beads (-dCTP) kit* (GE Healthcare). Approximately 50-100 ng of DNA denatured by heating at 95 °C for 3 min were incubated for 30 min at 37 °C with the provided mix and 25 µCi [α-^{32}P] dCTP (3000 Ci mmol^{-1}) (GE Healthcare).

2.3. MANIPULATION AND ANALYSIS OF RNA

2.3.1. RNA isolation

Collected cells (see 2.1.2.3) were resuspended in 300 µL of resuspension buffer (300 mM sucrose, 10 mM sodium acetate, pH 4.5), supplemented with 100 µL of 250 mM Na_2-EDTA and 400 µL Lysis buffer (2% SDS, 10 mM sodium acetate, pH 4.5). After that, 1 mL acid phenol (pH 4.5) was added. The resulting mixture was shaken for 30 s in a vortex and incubated at 65 °C for 2 min. This process was repeated 3 times. Then it was centrifuged for 5 min at 4 °C. The resulting supernatant was transferred successively to clean Eppendorf tubes containing 1 mL acid phenol, 1 mL acid phenol:chloroform (1:1), and 1 mL chloroform. After each addition the supernatant was shaken in a vortex for 30 s and incubated at 65 °C for 2 min (both twice), and centrifuged at 4 °C for 5 min. Finally, 1 mL of isopropanol was added and the tube was frozen at -20 °C. After centrifugation at 4 °C for 30 min, the supernatant was discarded and the pellet was washed with 70% ethanol. The pellet was dried and resuspended with 200 µL of DEPC-treated water at 4 °C for at least 2 h. Water used for RNA handling is previously treated with DEPC (diethylpyrocarbonate) to inactivate RNases by covalent modification of specific residues in these enzymes. 100 µL phenol and 100 µL chloroform were added and the resulting mix was shaken in a vortex and centrifuged for 5 min. 200 µL phenol were added, shaken and centrifuged again. Then, 500 µL 100% ethanol and 50 µL 3 M sodium acetate (pH 5.2) were added and the mix was incubated for 20 min at -20 °C. After centrifugation for 30 min, the pellet was washed with 70% ethanol, dried and resuspended with 45 µL of DEPC-treated water at 4 °C for at least 2 h. For removing possible traces of DNA, 5 µL of DNAse 10x buffer (Ambion) and 1 µL (2 U) of TurboDNAse (Ambion) were added and incubated for 30 min at 37 °C. After that, another 1 µL TurboDNAse was added and the mixture was incubated for 30 additional min. 5 µL of inactivation reactive was added and then incubated for 2 min at room temperature. Then, it was centrifuged for 1.5 min at 4 °C at 10,000 x *g* and the supernatant was transferred to a clean tube.

2.3.2. Estimation of RNA concentration and purity

The concentration of RNA was determined using a *NanoDrop Spectrophotometer* ND-1000. The absorbance was measured at 260 nm (extinction coefficient ε = 0.025 $mg^{-1}{\cdot}mL{\cdot}cm^{-1}$).

2.3.3. RNA electrophoresis in agarose gels

Agarose gels were prepared with 1% agarose (w/v) in a buffer containing 40 mM MOPS, 10 mM sodium acetate, 1 mM EDTA-Na_2 (pH 7.2) and 2% (v/v) formaldehyde. The

samples were prepared in a final volume of 28 µL containing: 15 µg RNA (diluted in 7 µL); 1.4 µL MOPS buffer 10x; 4.9 µL 37% (v/v) formaldehyde; 14 µL formamide; and 0.2 µL 0.07% (v/v) ethidium bromide. The samples were then supplemented with 1/5 volumes of sample buffer (1 mM EDTA-Na_2, pH 8; 0.25% [w/v] bromophenol blue; 0.25% [w/v] xylene cyanol FF; and 50% [v/v] glycerol) and incubated at 95 °C for 3 min, cooled on ice and loaded into the gel.

2.3.4. Northern blot analysis

2.3.4.1. RNA transference to nylon membranes

Agarose gels were washed 3 times with milliQ water to wipe out the formaldehyde and then equilibrated for 10 min in SSPE buffer 10x. The RNA was transferred from the agarose gel to a positively charged Nylon membrane Hybond-N^+ (GE Healthcare) with a vacuum system (*TransVac TE80*, Pharmacia Biotech) for 1 h using 100 mM NaOH as the transference buffer. After that, the membrane was let sit for 10 min on a Whatman paper soaked with 10 mM NaOH and then with SSPE 2x and let dry at 80 °C for 15 min. SSPE buffers were made using a 20x stock solution (3 M NaCl; 0.2 M Na_2HPO_4/NaH_2PO_4, pH 7.4; 0.02M EDTA-Na_2, pH 7.4).

2.3.4.2. Northern blot analysis

A membrane containing the RNA was prehybridized by immersion for 2-4 h at 65 °C in 25 mL of hybridization solution (180 mM K_2HPO_4; 140 mM KH_2PO_4; 7% SDS; and 1 mM EDTA-Na_2) using SHELL-LAB tubes in a SHELL-LAB Hybridization Oven. After this, the labelled DNA probe previously denatured was added and the membrane was incubated at 65 °C overnight. The membranes were then washed twice by incubation 10 min with SSC 2x and 0.1 % SDS, and once 15 min with SSC 1x and 0.1 % SDS. As a control, a membrane washed for 1 h with 0.5% (w/v) SDS and hybridized with a DNA probe of the *rnpB* gene (which encodes Ribonuclease P and it is expressed constitutively; Vioque (1992)) was used.

Detection and analysis of radioactivity in the membranes were performed by electronic autoradiography with a *Cyclone Storage Phosphor System* (Packard) and the software *OptiQuant* (PerkinElmer).

2.3.5. Quantitative RT-PCR

Retrotranscription was carried out with the *QuantiTect Reverse Transcription Kit* (QUIAGEN), using 1 µg of RNA as template and 100 ng *Random Hexamer Primers* (BIOLINE) instead of those provided with the kit. Before retrotranscription, the presence of contaminating DNA on the RNA preparation was tested by PCR.

Real-time PCR was performed in an *iCycler iQ* (BioRad) with the *SensiFAST SYBR & Fluorescein Kit* (BIOLINE) using 1/10 volumes of the cDNA resulting from the retrotranscription step.

The efficiencies of the primer pairs were tested previously by running a real-time PCR with different concentrations of genomic DNA. Efficiencies were calculated from the representation of the different Ct (threshold cycle) values with the formula: Efficiency=($10^{-1/slope}$)-1 and with the *LinRegPCR* software (Ramakers *et al.*, 2003). The annealing temperature was chosen as that allowing the highest efficiency values.

The Ct values obtained were analyzed with the *REST 2009* Software (Pfaffl *et al.*, 2002).

2.4. GENETIC METHODS

2.4.1. DNA transfer to *E. coli*

2.4.1.1. Transformation

Competent cells (100 µL DH5α or 10 µL HB101) were mixed with 10 µL of ligation product or 1 µL of purified plasmidic DNA preparation. The mix was incubated on ice for 15 minutes, then at 42 °C for 90 s, and then on ice again for 5 min. After that, 1 mL LB was added and the cells were incubated at 37 °C for 1 h in a shaker. Then, the cells were collected by centrifugation and the pellet was spread atop solid LB with the appropriate antibiotics for selecting the plasmid. If needed, 40 $\mu g \cdot mL^{-1}$ X-Gal was added to the LB medium to look for colonies bearing an insert into the β-galactosidase gene in the plasmid (α complementation assay) (Sambrook & Russell, 2001).

2.4.1.2. Electroporation

Competent cells (50 µL BL21) were mixed with 1 µL of plasmidic DNA preparation and incubated for at least 1 min on ice. The mix was subjected to electroporation using a 2 mm cuvette in an *EasyjecT Optima* electroporation system (EQUIBIO) (used conditions: 2500 V, 15 µF and 335 Ω). After that, 1 mL of LB was added quickly, and the cells were incubated for 1 h at 37 °C in a shaker. Later on, the cells were spread in the same way as described in the transformation protocol.

2.4.2. DNA transfer to *Anabaena* by conjugation

Conjugation was carried out as described in Elhai & Wolk (1988). The pRL443 plasmid (from strain ED8654) that confers the ability to form the *pili* necessary for conjugation is transferred to *E. coli* HB101 strain that carries the cargo plasmid and the

helper plasmid pRL623. Plasmid pRL623 bears genes that encode methylases for AvaI, AvaII and AvaIII restriction sites (Elhai et al. 1997), which protect the cargo plasmid from digestion by the cyanobacterial endonucleases, as well as the *mob* gene. In a second step, both plasmids are transferred to *Anabaena*.

The procedure for conjugation was: 250 µL of a stationary phase culture of *E. coli* ED8654 (pRL443) and 350 µL of a stationary-phase culture of *E. coli* HB101 (pRL623, pCargo) were inoculated in 10 mL of fresh medium with antibiotics and grown at 37 °C in a shaker for 2.5 h. Exponential phase cells were collected by centrifugation, washed twice with LB to remove antibiotics and mixed together in a final volume of 500-1000 µL. They were incubated for more than 2 h without shaking. Later, an *Anabaena* cellular suspension containing 10 µg chlorophyll *a* (Chl) was added to the *E. coli* mixture, and the resulting product spread in a nitrocellulose filter (Nucleopore) set atop solid $BG11_0$ + NH_4^+ + 5 % LB in a Petri dish. The Petri dish was incubated at 30 °C with white light (10 $\mu E \cdot s^{-1} \cdot m^{-2}$) for at least 3 h and then at standard light intensity overnight. The day after, the filter was transferred to a Petri dish with $BG11_0$ + NH_4^+ medium, and one day later, to $BG11_0$ + NH_4^+ medium with proper antibiotics. After that, the filter was transferred to fresh medium with antibiotics every second day until resistant colonies appeared, usually 10-15 days after conjugation.

2.5. ANALYSIS OF PROTEINS

2.5.1. Preparation of cellular extracts from *Anabaena*

1,000 mL of bubbled cultures were used for the study of protein content in different extracts of *Anabaena* cells. Cells were grown in $BG11_0$+NH_4^+ medium and incubated in combined nitrogen-free medium for the indicated times. Then the cells were collected by filtration (see 2.1.2.3) and resuspended in 10 mL of 20 mM HEPES-NaOH (pH 8) with *Protease Inhibitors Cocktail Complete Mini EDTA-free* (Roche). The cellular suspension was broken by passage through a French Pressure cell *SLM-AMINCO* (2 x 9,000 psi) and centrifuged at 5,000 x *g* in a *Beckman Coulter AvantiJ-25* centrifuge (JA-25.50 rotor). The resulting supernatant was then centrifuged at 32,000 x *g* in the same rotor. The supernatant was then centrifuged at 120,000 x *g* for 1 h in a *Beckman XL-80* centrifuge (80Ti rotor). The supernatant was collected as the cytoplasmic soluble fraction and the pellet was resuspended in 5 mL of 20 mM HEPES-NaOH (pH 8) and centrifuged again at 120,000 x *g* for 1 h in a *Beckman XL-80* centrifuge. The resulting pellet was resuspended in 500 µL as the membrane fraction. The fractions were kept at -20 C.

2.5.3. Electrophoresis in SDS-PAGE gels

Separation of proteins by electrophoresis in polyacrylamide gels was done following the procedure of Laemmli (1970) as described in Sambrook & Russell (2001), using Miniprotean3 Cell (Bio-Rad) and Vertical Slab Gel Electrophoresis SE600 (Hoeffer) systems. The running gel was prepared with acrylamide:bisacrylamide (30:0.8) at 10-14 % in 375 mM Tris-HCl (pH 8.8). The stacking gel contained 4% acrylamide:bisacrylamide (30:0.8) in 125 mM Tris-HCl (pH 6.8). Both gels also contained 0.1 % (w/v) SDS, 0.05% APS (Ammonium persulfate) and 0.1% TEMED. Sample buffer 2x was prepared consisting of 0.125 M Tris-HCl (pH 6.8), 20% (v/v) glycerol, 4% (w/v) SDS, 10% (v/v) 2-mercaptoethanol and 0.0025% (w/v) bromophenol blue. Samples were heated for 5 min at 95 °C or 65 °C (for membrane fractions) before loading. Samples were not heated before loading to prevent total denaturation. The electrophoresis buffer was prepared with 25 mM Trizma base, 192 mM glycine and 0.1% SDS (pH 8.3). The protein size standards used were *SeeBluePlus2* Prestained Standard (Invitrogen) or *Prestained Ladder* (gTPbio).

2.5.4. Coomassie blue staining

Proteins in SDS-PAGE gels were visualized by staining with a solution containing 0.25% (w/v) Coomasie blue R-250, 10% (v/v) acetic acid and 45% (v/v) methanol in water. After staining, excess stain was removed from the gels with a solution containing the same components except the Coomasie Blue.

2.5.5. Protein quantification

Protein extracts were quantified using the method described by Bradford (1976), measuring the absorbance at 595 nm. BSA was used for calibration.

In experiments examining growth rates, whole cell protein determination was performed using the method of Lowry *et al.* (1951) with the modifications introduced by Markwell *et al.* (1978). 200-µL samples were supplemented with 50 µL of 0.5 N NaOH and 750 µL of solution C (a 100:1 mixture of Solution A and Solution B; Solution A is 2% [w/v] Na_2CO_3, 0.4% [w/v] NaOH, 0.16% [w/v] potassium sodium tartrate and 1% [w/v] SDS; Solution B is 4% [w/v] $CuSO_4 \cdot 5H_2O$). Samples were shaken with a vortex and incubated for 5 min at 37 °C. 75 µL of Folin-Ciocalteus reagent diluted 1:1 (v/v) in water were added, the samples were shaken and then incubated for 5 min at 37 °C. The absorbance at 750 nm was finally measured. Samples with different amounts of BSA were used for reference.

2.5.8. Transference of proteins to PVDF membranes and Western blot analysis

Proteins were transferred from the gel to *PVDF Hybond-P* 0.45 µm pore size membranes (GE Healthcare) with *ECL Semi-dry* blotters TE 77 and TE 70 PWR (GE-Healthcare) systems. Transference was carried out for 1.5 h at 1 $mA \cdot cm^{-2}$ using as the transference buffer 478 mM Trizma base (pH 8.0), 0.037% (w/v) SDS, 39 mM glycine and 5% (v/v) methanol. For the Western analysis, the membrane was incubated for 1 h at 30 °C or at 4 °C overnight with blocking buffer (5% [w/v] semi skimmed milk powder in TBS [15 mM Tris-HCl (pH 7.5), 200 mM NaCl]). After that, the membrane was incubated for 1.5 h at 30 °C with a primary antibody (anti-GFP 1:2,000 [Invitrogen]) diluted in blocking buffer at the indicated concentrations. The membrane was washed 3 times during 5 min in TBS-T (TBS + 0.05% [v/v] Tween-20) and incubated for 1-1.5 h at 30 °C with a secondary antibody [Anti-Rabbit IgG (whole molecule)–Peroxidase antibody produced in goat, Sigma] diluted 1:10,000 in blocking buffer. The membrane was washed 3 times (5 min each time) in TBS-T and once for 5 min in H_20. Signals in the membrane were visualized with the *ECL Plus Western Blotting Detection System* (GE-Healthcare) or *WesternBright ECL* (Advansta) and *Hyperfilm ECL* (GEHealthcare).

2.6. MICROSCOPY METHODS

2.6.1. Optic microscopy

Anabaena cells were visualized using an Olympus BX60 microscope with a UplanFI 40x/0.75Ph2 objective or a UplanFI 100x/1.30 oil objective.

2.6.2. Confocal microscopy

2.6.2.1. Detection of GFP

Samples were visualized using an immersion objective HCX PLAM-APO 63x 1.4 NA in a Leica TCS SP2 microscope or an CX PL APO lambda blue 63.0x1.40 OIL UV objective in Leica SP5 microscope. Samples were excited at 488 nm using an argon laser. Fluorescence was monitored in the range 500-540 nm (GFP imaging) or 630-700 nm (cyanobacterial autofluorescence). *Anabaena* filaments were set atop a thin slice of solidified medium in order to visualize the cells in the same plane. Images were treated with the *LAS AF Leica Software* and *ImageJ*.

2.6.2.2. Time-lapse experiments

Cells grown on $BG11_0$ Petri dishes ($BG11_0$ for this purpose was prepared with 0.5% [w/v] agar) were visualized in a confocal microscope by cutting a square of the agar and placing it in a metal support that fits the microscope sample holder. Cells in this device were placed in a 30 °C incubator and continued growing during all the experiment.

2.6.3. Fluorescence microscopy

2.6.3.1. Visualization of GFP

Samples were visualized using a 100x or 63x objective in a Leica DM6000B microscope with an ORCA-ER camera (Hamamatsu). The GFP fluorescence was detected using a FITC L5 filter with excitation window BP (band-pass) 480/40 filter and emission window BP 572/30 filter. Cyanobacterial autofluorescence was detected using a Texas Red TX2 filter with excitation window BP 560/40 filter and emission window BP 645/75 filter. Images were analysed using the *LAS AF Leica Software* and *ImageJ*.

2.6.3.2. Immunofluorescence

100 µL of pelleted cells corresponding to 1 mL of liquid cultures (containing 2-3 µg of Chl*a*·mL^{-1}) were placed atop poly-L-lysine precoated microscope slides (VWR, Sigma) and covered with a 45 µm pore-size Millipore filter. After that, the filter was removed and the slide was let to dry, incubated with 70% ethanol for 30 min at -20 °C and dried for 10 min at 80 °C. Cells were washed twice for 2 min with PBS-T buffer (137 mM NaCl, 2.7 mM KCl, 5.5 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.05% [w/v] Tween-20) and treated with blocking buffer (5% [w/v] semi-skimmed milk powder in PBS-T) for 15 min, followed by an incubation during 90 min with a primary antibody diluted in blocking buffer, 3 washes with PBS-T, an incubation with a secondary antibody during 45 min (anti-rabbit, conjugated to FITC [fluorescein isithiocyanate], Sigma, 1:500 diluted in PBS-T) followed by 3 washes with PBS-T. After that, drops of FluorSave (CarlBioChem) or Fluoroshield with PPD (Sigma) were added and the preparation was covered with a coverslip.

2.6.4. Transmission electron microscopy (TEM)

2.6.4.1. Preparation of *Anabaena* cultures for TEM

Anabaena cultures growing in $BG11_0$+NH_4^+ medium were shifted to $BG11_0$ medium. After 12 h, the cultures were centrifuged and the cells collected in Eppendorf tubes. Fixation of the samples was done treating for 2 h at room temperature with 4% glutaraldehyde (v/v) in 125 mM phosphate buffer (pH 7.4). After that, the cells were washed 3 times for 10 min with phosphate buffer, mixed with 2% (w/v) agarose

(preheated at 80 °C) and allowed to cool. Afterwards, the solidified agarose containing the cells was cut in small cubes of about 1 mm^3.

Cubes were fixed with 2% $KMnO_4$ over-night at room temperature and washed 5 times for 15 minutes with milliQ water. Cells in the cubes were dehydrated by subsequent steps of incubation with 30% to 100% (v/v) ethanol and embedded in araldite for 48 h at 60 °C in a silicone mould. Solidified blocks were trimmed to a trapezoid, which was cut in 50-70 nm sections (silver-gold colour) with an ultramicrotome (Reichert-Jung Ultracut). Glass knives were made in a LKB 7800 Knifemaker.

Ultrathin sections were placed on a copper grid, 300 mesh (Agar Scientific), let to dry, and stained by placing them in a drop of 2% (w/v) uranyl acetate for 4 min, washed with milliQ water (3 times, 10 s each), and placed in 2% (w/v) lead citrate for 4 min. Sections were visualized in a JEOL 2010 transmission electron microscope.

2.6.4.2. Immunolocalization

For immunogold localization some modifications were made from the protocol described above. Cells were fixed with 3% (v/v) paraformaldehyde in 125 mM phosphate buffer (pH 7.4), instead of glutaraldehyde. Agarose cubes were not stained with $KMnO_4$, but they were directly washed, dehydrated and embedded in the *LRWhite* resin (Agar Scientific) in gelatine capsules overnight at 55 °C.

Ultrathin sections were cut with a diamond knife and placed on nickel grids in which they were treated with 5% (v/v) H_2O_2 for 5 min, washed with PBS twice for 5 min, blocked with BSA 10% (w/v) in PBS for 1 h, treated for 1 h with a primary antibody diluted 1:50 in blocking buffer, washed with PBS (6 times, 2 min each time) and treated 1 h with a secondary antibody (Anti-rabbit IgG-Gold antibody, Sigma) diluted 1:50 in blocking buffer. Sections were washed 3 times with blocking buffer, 3 times with PBS and finally 3 times with milliQ water. After that, they were stained for 5 min with 2% uranyl acetate, washed with water, incubated with 2% lead citrate and finally washed again.

2.7. PHYSIOLOGICAL PARAMETERS MEASURED IN *ANABAENA*

2.7.1. Determination of chlorophyll *a* concentration

The concentration of chlorophyll *a* (Chl) was measured following the method described in Mackinney (1941). Pelleted cells corresponding to 1000 µL of culture were resuspended in 950-1000 µL of methanol, shaken in a vortex for 1 min and sedimented in an Eppendorf centrifuge for 2 min. After that, OD was measured at 665 nm (extinction coefficient ε=74.46 $mg^{-1} \cdot mL \cdot cm^{-1}$).

2.7.2. Determination of growth rates

Anabaena was grown for 7 days in 50 mL of BG110+NH_4^+ medium with the appropriate antibiotics in a flask. Cultures were collected by centrifugation and washed 3 times with $BG11_0$ medium. Cells, corresponding to 0.2 µg Chl·mL^{-1}, were inoculated in flasks with 25 mL of $BG11_0$+NH_4^+ or $BG11_0$ medium without antibiotics. Twice a day for 5 days, cultures were homogenized with a 1000 µL micropipette and 200 µL of the culture were withdrawn and frozen. After all the samples were taken, protein concentration was measured by the Lowry method (2.5.5). Growth rate (µ) is calculated as $\mu = \ln 2/t_d$ where t_d is the duplication time.

2.7.3. Serial dilutions spot assay of growth

Cells grown in flasks were collected and washed 3 times with $BG11_0$ medium. After measuring the Chl concentration, cell suspensions were diluted to give suspensions containing 1, 0.5, 0.25, 0.125 and 0.0625 µg Chl·mL^{-1}. Five µL of each suspension were spotted on the surface of plates containing $BG11_0$+NH_4^+ or $BG11_0$ medium without antibiotics. Thus, the spots initially contained 5, 2.5, 1.25, 0.625 and 0.3125 ng Chl, respectively. The plates were incubated under standard growth conditions for 14 days.

2.7.4. Nitrogenase activity assay

Nitrogenase activity was determined in vivo by the acetylene reduction assay described by Stewart *et al.* (1967). Cultures grown in $BG11_0C$+NH_4^+ medium were incubated in $BG11_0C$ medium for 24 or 48 h. 2 mL of culture containing an amount of cells corresponding to 10 µg Chl·mL^{-1} were placed in small flasks that were closed with a rubber stopper. 2 mL of acetylene was added to the flasks, the suspensions were incubated at 30 °C, 100 µE·s^{-1}·m^{-2} light intensity with shaking (95 rpm). Samples of 1 mL of the gas phase were taken every 10 min during 40 min or every 15 min during 60 min, and the ethylene produced was determined by gas chromatography. Nitrogenase activity is expressed as nmol ethylene produced·µg Chl^{-1}·h^{-1}.

Nitrogenase activity was also measured under anoxic conditions, which were established by adding 10 µM DCMU (dichlorophenyl dimethylurea, which blocks photosystem II preventing O_2 production) and substituting the oxygen atmosphere by sparging with argon for 3 min.

2.7.5. Staining of heterocysts with Alcian Blue

Filaments were grown in $BG11_0C$+NH_4^+ medium, harvested, washed and incubated for 24, 48 or 72 h in $BG11_0C$ medium in bubbled cultures. After these times, a sample of 1

mL of culture was withdrawn and centrifuged. The resulting pellet (approx. 100 µL) was mixed 1:10 with 1% (w/v) Alcian Blue in water and visualized under the microscope.

2.7.6. Thin-layer chromatography of lipids

Lipid isolation was carried out following the method described by Nichols & Wood (1968) with slight modifications. The cells in 25 mL of a bubbled culture grown under diazotrophic conditions for 24 or 48 h were collected by centrifugation. The pellet was resuspended in 1 mL of 2:1 (v/v) chloroform:methanol and shaken vigorously. The resulting suspension was transferred to a Pyrex glass tube, and 2 mL of 2:1 (v/v) chloroform:methanol and 5 mL of water were added. The resulting solution was shaken, let stabilize for 5 minutes and centrifuged for 5 min at 4,000 rpm at room temperature. The pellet was transferred to a clean tube and dried by sparging with N_2 at 40 °C. The dried pellet was resuspended in 100 µL chloroform.

30 µL of lipid preparation was loaded into a TLC Silica Gel 60 20x20 cm plate (Merck). Unidimensional chromatography was carried out using as a solvent a mixture of chloroform 74.75%, methanol 13.2%, acetic acid 8.79% and water 3.25%. Lipids on the plate were visualized by spraying the plate with 50% (v/v) sulfuric acid and heating at 100-150 °C in a heater plate *Stuart Hotplates & Stirrers SB-300*. Heterocyst specific glycolipids were identified according to Winkenbach *et al.* (1972).

2.7.7. Heterocyst frequency and distribution

Anabaena cells stained with Alcian Blue were visualized under the optic microscope. At least 100 hundred intervals or 300 cells were counted in each of 3 independent experiments. Dividing cells were counted as two cells.

2.8. SOFTWARE ANALYSIS OF DNA AND PROTEIN SEQUENCES

The genomic sequence of *Anabaena* sp. PCC 7120 (Kaneko *et al.*, 2001) is found in Cyanobase (http://genome.kazusa.or.jp/cyanobase/) and the Integrated Microbial Genomes of JGI (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). The *SerialCloner* and *DNAStrider* softwares were used for analyzing DNA sequences in the generation of mutants. *Amplify3x*, *AmplifX* and *PrimerSelect* were used for the design of oligodeoxynucleotide primers. CLC Sequence Viewer was used to study the RNA-seq data.

Topology predictions of proteins were made using the *MEMSAT-SVM* (http://bioinf.cs.ucl.ac.uk/psipred/), *InterPro, PredictProtein, TopPred* and *SignalP* (http://www.expasy.org/tools/) software programs.

3. STRAIN CONSTRUCTION

DNA transferred to *Anabaena* can be established in the cells as a replicative plasmid or integrated into the genome by homologous recombination. For the latter, a sequence homologous to the target DNA sequence is introduced in a suicide plasmid. If there is only one homologous sequence in the plasmid, a single cross-over event will occur resulting in the insertion of the whole plasmid into the genome and the duplication of the homologous region. If two sequences homologous to different (but nearby) regions in the genome are introduced, two cross-over events can take place leading to the substitution of the genomic region in the recipient by the one present in the plasmid, with the simultaneous loss of the suicide vector.

For the generation of mutants by single recombination, a PCR product was cloned in a high copy plasmid and then transferred to the suicide vector, which was introduced into *Anabaena* by conjugation. For the construction of mutants by double recombination, the overlapping PCR method (Ausubel *et al.*, 2006) was used to obtain two DNA sequences cloned in a high copy vector, and then the DNA was transferred to a suicide vector that bears the *sacB* gene, which is lethal when cells are exposed to 5% sucrose (Cai & Wolk, 1990).

Once exconjugant colonies were selected, clones bearing only mutant chromosomes were sought by subcloning at least 3 times, before checking the construct by PCR. In the case of double recombination mutations or difficult to segregate single recombinants, liquid cultures were sonicated in order to break down the filaments to 3-4 cells.

3.1. CONSTRUCTION OF *patS* MUTANT STRAINS

A *patS* deletion mutant (strain CSVT20) had been generated previously in our lab (Figure 3.1). Two DNA fragments, one encompassing sequences upstream of the gene and the other including sequences downstream of the gene, were amplified by PCR using DNA from *Anabaena* as template and the primer pairs asl2301-8/asl2301-9 and asl2301-10/asl2301-11. The DNA fragments were cloned in pMBL-T, sequenced and inserted together, in direct orientation, into pCSRO (Merino-Puerto *et al.*, 2010). The plasmid produced, pCSVT51, was transferred to *Anabaena* by conjugation. Exconjugants were selected by their resistance to Sm and Sp and double recombinants were then selected by their resistance to sucrose. The clone isolated, CSVT20, bore a deletion of 180 bp that includes the whole *patS* gene.

3. STRAIN CONSTRUCTION

To generate mutants producing different PatS versions, integrative plasmids bearing the *patS* promoter and the mutant *patS* version were transferred to strain CSVT20 by conjugation. In this way, the resulting mutant strains had only one copy of the *patS* gene under its native promoter in the chromosome.

To construct three different strains bearing *patS-gfp-mut2* genes encoding polypeptides in which the GFP protein was fused to different regions of the PatS peptide, overlapping PCR was performed using (i) *Anabaena* DNA as template and primer pairs asl2301-9/asl2301-15 (for strain CSL19), asl2301-9/asl2301-16 (for strain CSL20) and asl2301-9/asl2301-17 (for strain CSL21), or (ii) pCSAM135 (Flores *et al.*, 2007) as template and primer pairs asl2301-gfp-1/asl2301-gfp-2 (for strain CSL19), asl2301-gfp-3/asl2301-gfp-4 (for strain CSL20) and asl2301-gfp-5/asl2301-gfp-6 (for strain CSL21). PCR products were cloned together in pMBL-T producing plasmids pCSL29, pCSL31 and pCSL33 respectively. KpnI fragments from these plasmids bearing the *patS-gfp-mut2* constructs were inserted into the mobilizable vector pCSV3 (Valladares *et al.*, 2011) digested with the same enzyme, producing pCSL30, pCSL32 and pCSL34, respectively (Figure 3.2).

To generate fusions of PatS to a 6-His tag, a DNA fragment was amplified using primer pairs asl2301-9/asl2301-18 and asl2301-9/asl2301-19 and cloned in pMBL-T producing pCSL35 and pCSL37. KpnI fragments from these plasmids bearing the *patS*-6His fusion were inserted into pCSV3 digested with the same enzyme, producing pCSL36 and pCSL38 (Figure 3.3).

For construction of *patS* point mutants and minigenes, a PCR fragment was amplified using primer pairs asl2301-9/asl2301-32 to asl2301-49 (for strains CSL44 to CSL62), asl2301-9/asl2301-77 (for strain CSL94) or asl2301-69/asl2301-71 or asl2301-72 (for strains CSL91 and CSL92, respectively) and finally cloned into pCSV3 producing pCSL82 to pCSL95, pCSL110, pCSL104, and pCSL105, respectively. For construction of strains CSL90 and CSL93, DNA fragments from 2771000 to 2771734 (coordinates of the *Anabaena* sp. PCC 7120 genomic sequence), including nucleotide changes to produce K^2A and V^7A substitutions, respectively, were synthesized (GeneArt, Invitrogen) and cloned into pCSV3 producing plasmids pCSL103 and pCSL106, respectively (Figure 3.3).

Plasmids pCSL30, pCSL32, pCSL34, pCSL36, pCSL38, pCSL82 to pCSL95, pCSL110, pCSL104, pCSL105, pCSL103 and pCSL106 were transferred to the *patS* mutant, strain CSVT20, by conjugation (Figure 3.4). Clones resistant to Sm and Sp were selected, and their genomic structure in the *patS* region was tested by PCR with primers asl2301-74 (which lies outside of the construct, 5' from it) and asl2301-76 (which lies inside the region deleted in strain CSVT20) or asl2301-75 (which lies outside of the construct, 3'

from it) to check recombination in the correct place and segregation of the *patS*-mutant chromosomes (Figure 3.5). The corresponding *Anabaena* mutant strains were named CSL19-CSL23, CSL44-CSL62, CSL90-CSL94. Additionally, except for CSL19 and CSL21, the *patS* genomic region was sequenced in all the *Anabaena* strains generated in this work.

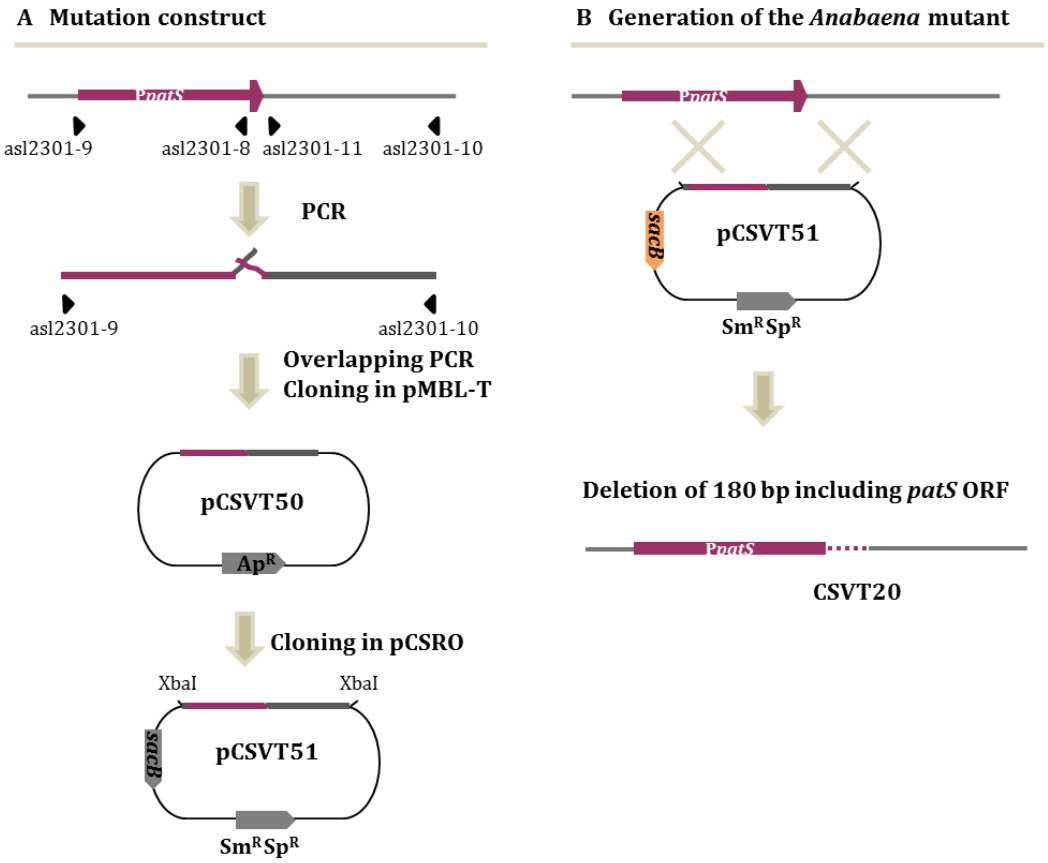


Figure 3.1. Construction of a *patS* deletion mutant. A. Generation of the genetic construct. B. Recombination of the generated construct in *Anabaena*. The *patS* promoter is represented as a thinner rectangle in the same colour as the *patS* gene (purple). The deleted region is represented as a discontinuous line. Genes in the plasmids are represented as grey pentagons. See text for details.

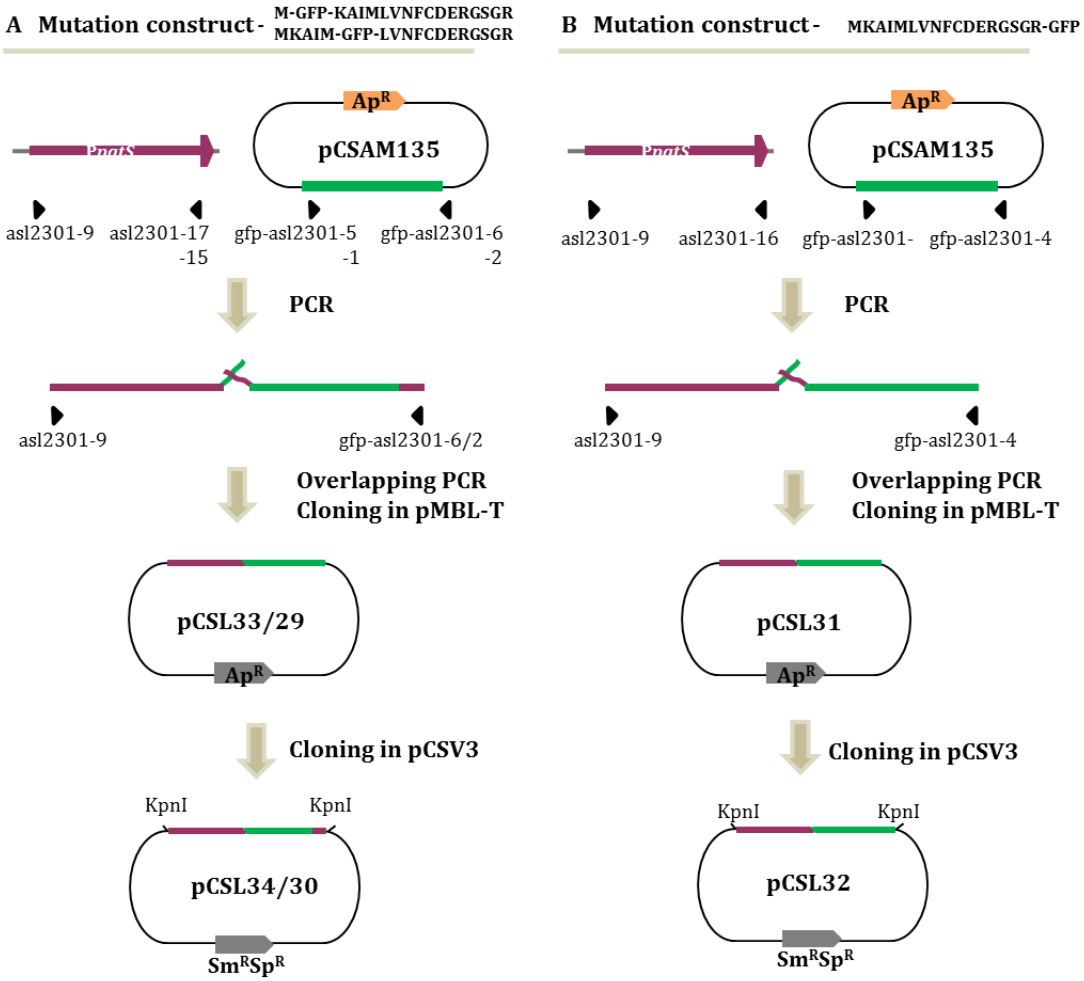


Figure 3.2. Construction of plasmids bearing *patS-gfp-mut2* fusion genes. A. Generation of two different N-terminal *patS-gfp-mut2* fusion constructs. B. Generation of the C-terminal *patS-gfp-mut2* fusion construct. The *patS* genomic region is represented in purple and the *gfp-mut2* gene is represented in green. Genes in the plasmids are represented as grey pentagons. See text for details.

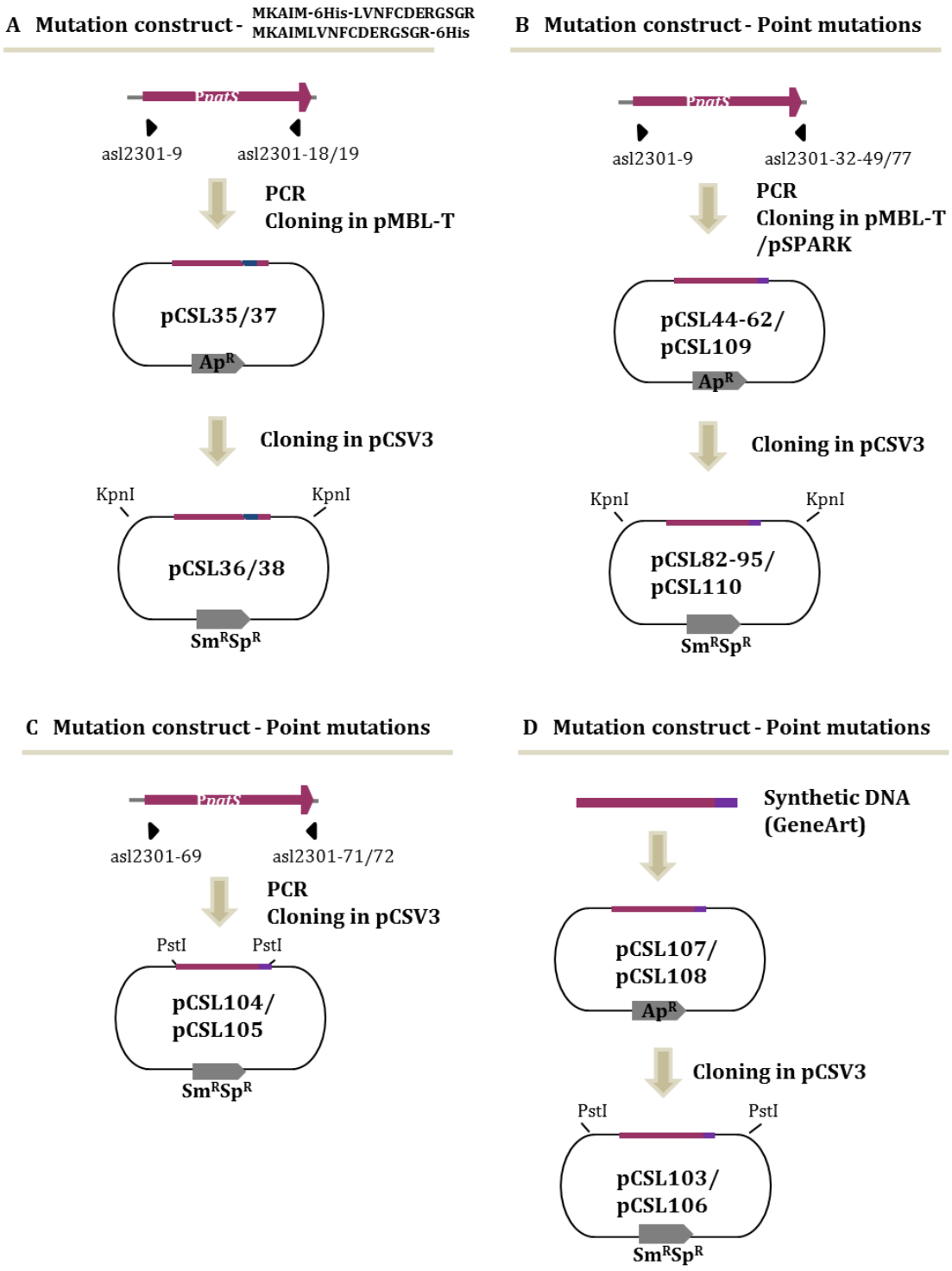


Figure 3.3. Construction of *patS* genes encoding 6His tag fused PatS polypeptides, *patS* genes with point mutations and *patS* minigenes. A. Generation of two different 6His tag fusion constructs. B,C and D. Generation of the point mutated *patS* genes and minigenes constructs. One of these three strategies was used to clone each construct. The *patS* genomic region is represented in purple, the *gfp-mut2* gene is represented in green, the 6His tag in dark blue and the point mutations in the *patS* gene in a dark purple. See text for details.

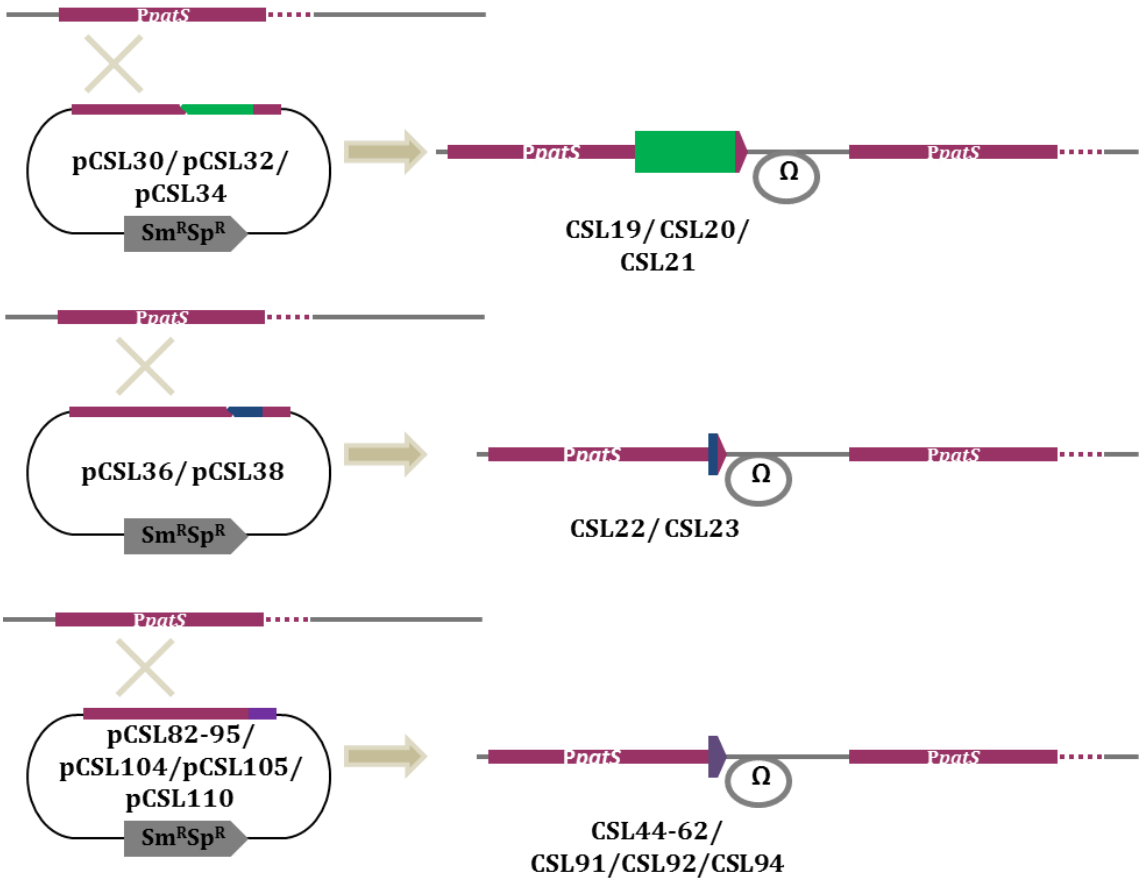


Figure 3.4. Scheme of the integration of the different *patS* versions in the *Anabaena* chromosome. The *patS* genomic region is represented in purple, the *gfp-mut2* gene is represented in green, the 6His tag in dark blue and the point mutations in the *patS* gene in dark purple. See text for details.

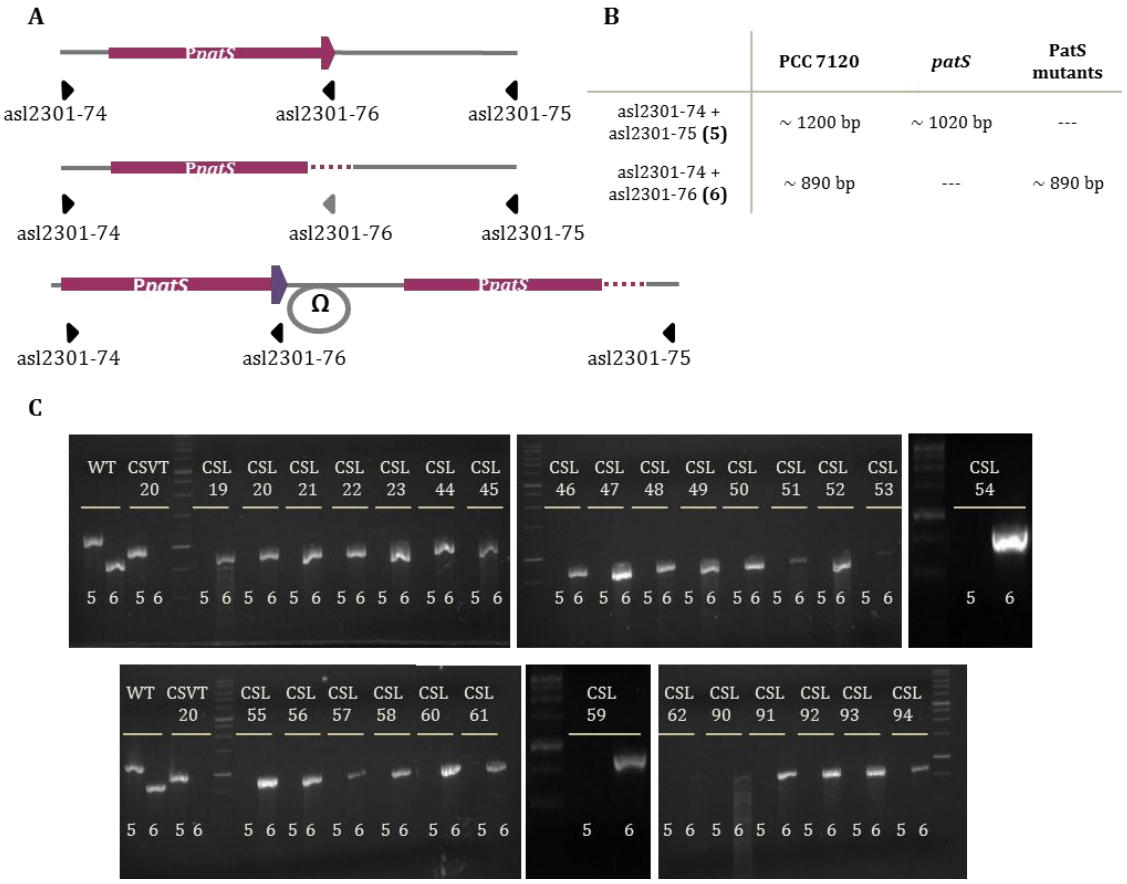


Figure 3.5. Analysis of the mutant chromosomes in the *patS* mutant strains. A. The scheme of the resulting mutant genomic region and the corresponding wild-type region are represented. B. The expected size bands in the PCR analysis are indicated. C. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, asl2301-74 and asl2301-75, and (ii) asl2301-74 and a primer that lies inside the region deleted in the *patS* mutant, asl2301-76. See text for details.

3.2. CONSTRUCTION OF *hetN* MUTANT STRAINS

A *hetN* deletion strain was constructed in which the whole ORF, except for 54 bp, was deleted (Figure 3.6). Two fragments comprising sequences upstream and downstream of the gene were amplified using primers alr5358-1/ alr5358-2 and alr5358-3/ alr5358-4 by overlapping PCR, cloned in pMBL-T (producing pCSL22), sequenced and cloned in pRL278. The resulting plasmid, pCSL23, was transferred to *Anabaena* (producing strain CSL7) and to strain CSVT20 (producing strain CSL11) and double recombinants were selected and tested for total segregation of the mutant chromosomes.

To study the role of the ERGSGR sequence, a mutant lacking only the nucleotide sequence encoding these amino acids was generated using the same technique (Figure 3.7). Overlapping PCR was performed using *Anabaena* DNA as template and primer pairs alr5358-1/ alr5358-9 and alr5358-10/ alr5358-4. The PCR fragment was cloned in pMBL-T producing pCSL41, sequenced and subsequently cloned into pRL278. The resulting plasmid, pCSL42, was transferred to wild-type *Anabaena* producing strain CSL32, and double recombinants were selected and tested for total segregation.

To study the requirement of a putative signal peptide in HetN, a strain bearing a gene lacking the sequence corresponding to amino acids 2-27 was constructed with the same technique (Figure 3.8). Two fragments comprising sequences upstream and downstream of the gene, respectively, were amplified using primers alr5358-20/alr5358-13 and alr5358-14/alr5358-21 by overlapping PCR, cloned in pRL278 and sequenced. The resulting plasmid, pCSL73, was transferred to *Anabaena* generating strain CSL73. The chromosomal region in the mutant was checked by PCR and sequenced.

For construction of a *hetN-sfgfp* fusion, the alr5358-18/alr5358-19 primer pair was used to amplify the end of *hetN* before cloning it in pCSAL39 producing pCSL71 (Figure 3.9). The pCSAL39 plasmid bears a sequence encoding a Gly-Gly-Gly-Gly linker preceded by a BsaI site, just before the *sfgfp* sequence. After that, the fused sequence was cloned in pCSV3 producing pCSL72, which was transferred to *Anabaena* (producing strain CSL71), to CSVT20 (producing strain CSL72) and to CSL73 (producing strain CSL103). Clones resistant to Sm and Sp were selected, and their genomic structure was tested by PCR.

For construction of a *hetN-gfp-mut2* fusion, overlapping PCR was performed using (i) pCSL72 as template and primer pairs alr5358-26/GFP-13, or (ii) pCSL68 as template and primer pairs GFP14/GFP-11 (Figure 3.10). The resulting DNA fragment was cloned using KpnI in pCSV3 producing plasmid pCSL124. This plasmid was transferred to *Anabaena* producing strain CSL108. Clones resistant to Sm and Sp were selected, and their genomic structure was tested by PCR (Figure 3.9).

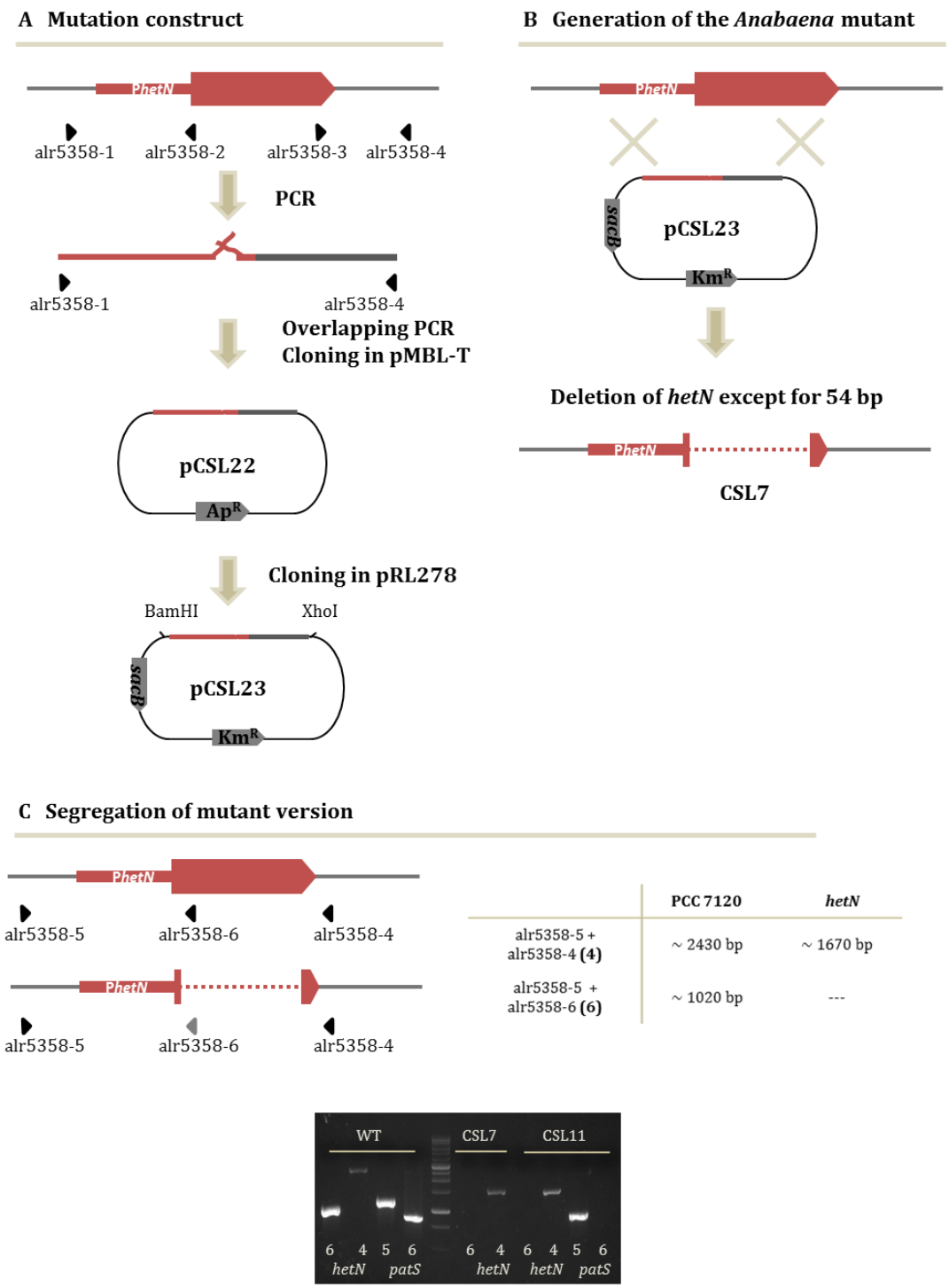


Figure 3.6. Construction and analysis of *hetN* deletion mutants. A. Generation of the genetic construct. The *hetN* gene is represented in dark red. Genes in the plasmids are represented as grey pentagons. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild-type are represented. The expected size bands in the PCR analysis are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr5358-5 and alr5358-4, and (ii) all5358-5 and a primer that lies inside the region deleted in *hetN* mutant, alr5358-6. See text for details.

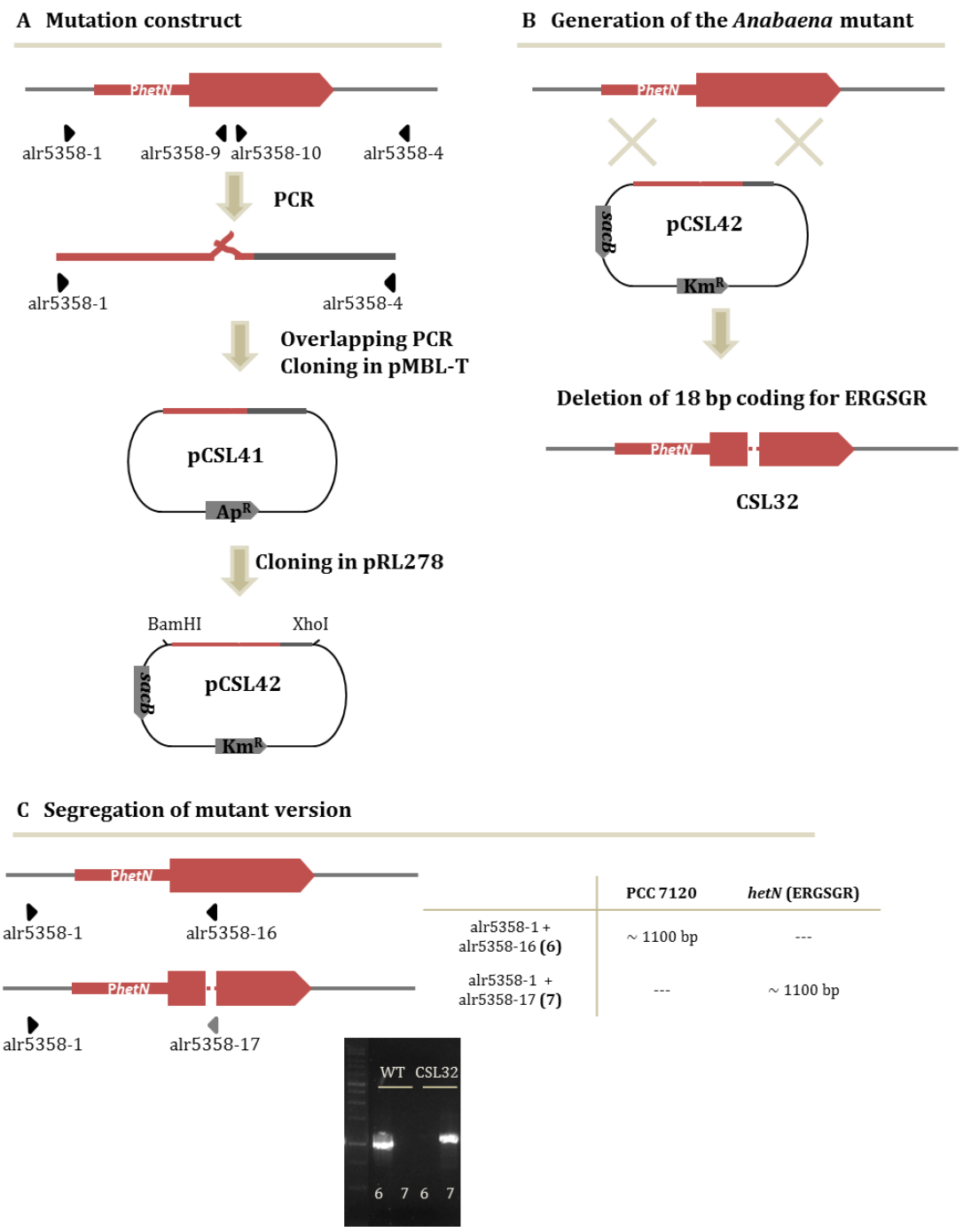


Figure 3.7. Construction and analysis of the strain bearing the deletion of the ERGSGR coding sequence in *hetN*. A. Generation of the genetic construct. The *hetN* gene is represented in dark red. Genes in the plasmids are represented as grey pentagons. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild-type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with a primer that lies outside the construct as forward primer, alr5358-1, and two reverse primers that lie in the deleted region, so that alr5358-16 only anneals in the wild-type chromosomes and alr5358-17 only anneals in the mutant ones. See text for details.

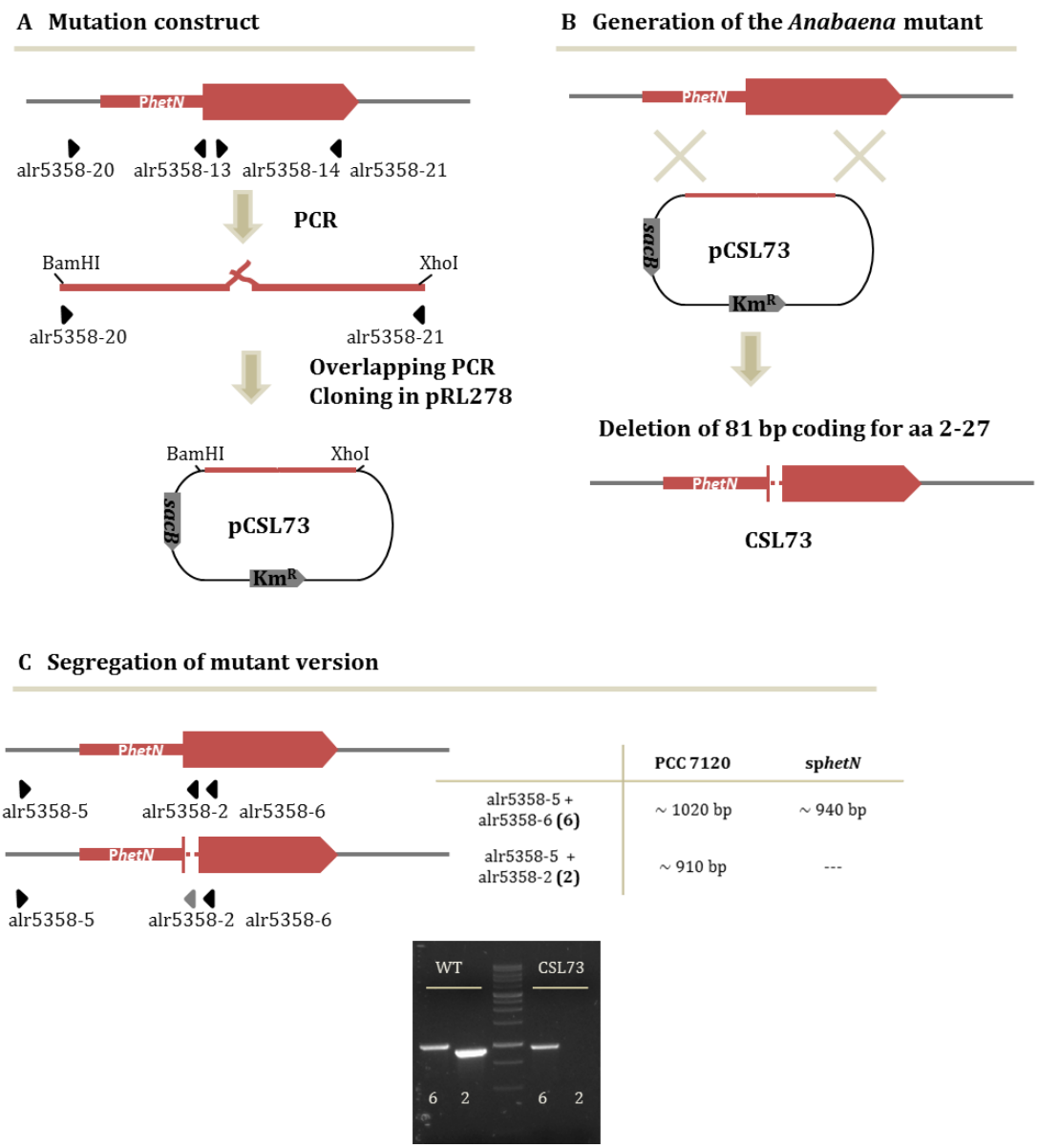
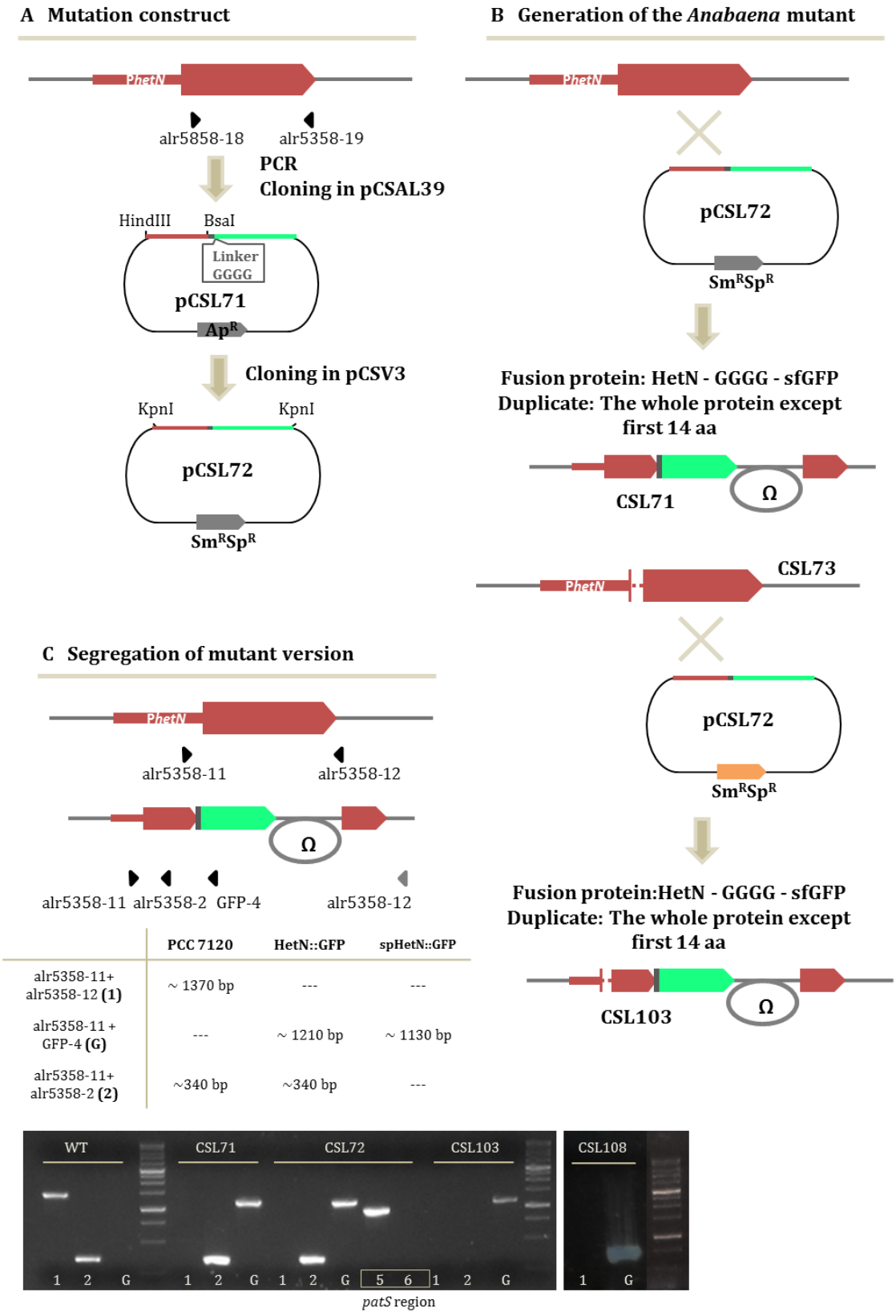


Figure 3.8. Construction and analysis of the *hetN* putative signal peptide deletion mutants. A. Generation of the genetic construct. The *hetN* gene is represented in dark red. Genes in the plasmids are represented as grey pentagons. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild-type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr5358-5 and alr5358-6, and (ii) alr5358-5 and a primer that lies inside the region deleted in sp*hetN*, alr5358-2. See text for details.



	PCC 7120	HetN::GFP	spHetN::GFP
alr5358-11+ alr5358-12 **(1)**	~ 1370 bp	---	---
alr5358-11 + GFP-4 **(G)**	---	~ 1210 bp	~ 1130 bp
alr5358-11+ alr5358-2 **(2)**	~340 bp	~340 bp	---

Figure 3.9. Construction of the *hetN-sfgfp* fusion in different genetic backgrounds. A. Generation of the genetic construct. The *hetN* gene is represented in dark red and *sfgfp* in green. Genes in the plasmids are represented as grey pentagons. B. Recombination of the generated construct in *Anabaena* sp. PCC 7120 and strain CSL73. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild-type are represented. The expected size bands in the PCR are

indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr5358-11 and alr5358-12, (ii) alr5358-11 and a primer that lies inside the region deleted in sp*hetN*, alr5358-2, and (iii) alr5358-11 and a primer that lies in the *gfp* region, GFP-4. See text for details. The analysis of strain CSL108 (See Figure 3.10) is also shown.

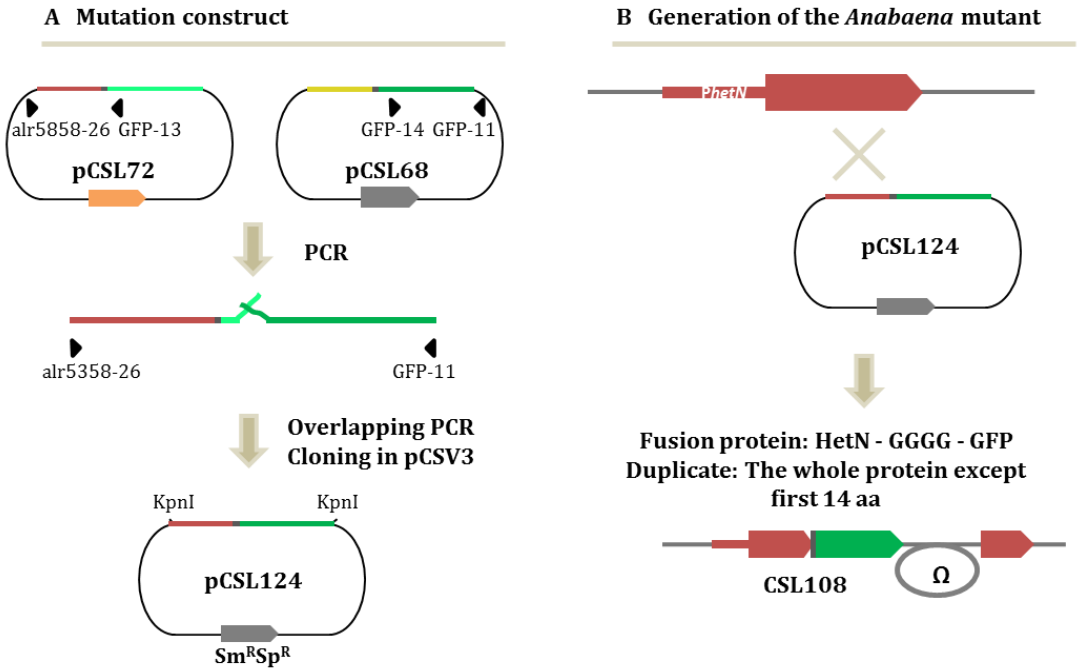


Figure 3.10. Construction of the *hetN-gfp-mut2* fusion. A. Generation of the genetic construct. The *hetN* gene is represented in dark red and gfp-mut2 in green. Genes in the plasmids are represented as grey pentagons. B. Recombination of the generated construct in *Anabaena.*

3.3. GENERATION OF STRAINS BEARING ALTERED VERSIONS OF *hetC*

A *hetC* deletion strain was constructed in which the whole ORF except for 90 bp from both extremes was substituted by the C.S3 cassette (Figure 3.11). For that, two PCR fragments comprising sequences upstream and downstream from the gene were amplified using primer pairs alr2817-34/alr2817-35 and alr2817-36/alr2817-37, cloned in the pMBL-T plasmid and sequenced, and the C.S3 gene cassette was inserted into an EcoRV site generated at the junction of the two cloned DNA fragments, resulting in plasmid pCSL19. This construct was inserted into pRL278. The produced plasmid, pCSL20, was transferred to *Anabaena* sp. PCC 7120 (producing strain CSL3), CSVT20 (producing strain CSL1), CSL7 (producing strain CSL12) and CSL11 (producing strain CSL15), and double recombinants were selected and tested for segregation of the mutation.

To delete the predicted peptidase domain of *hetC* (Figure 3.12), overlapping PCR was performed with primer pairs alr2817-38/alr2817-42 and alr2817-43/alr2817-41 and the product was cloned in plasmid pMBL-T, producing pCSL24, which was sequenced and transferred to pCSRO. The plasmid produced, pCSL25, was transferred by conjugation to *Anabaena* sp. PCC 7120 (producing strain CSL16), CSVT20 (producing strain CSL17), CSL7 (producing strain CSL30) and CSL11 (producing strain CSL31). Additionally, the plasmids pCSL80 and pCSL81 that bear the P_{patS} fused to minigenes encoding PatS-8 and PatS-17, respectively, were transferred to strain CSL17, producing strains CSL101 and CSL102.

For construction of a fusion of *hetC* to the *gfp-mut2* gene (Figure 3.13), a DNA fragment was amplified using *Anabaena* DNA as template and primer pair alr2817-13/alr2817-14. The PCR product was cloned in pMBL-T and afterwards in pCSAM135, which bears the *gfp-mut2* sequence (Flores *et al.*, 2007), and the resulting *gfp-mut2-hetC* construct was transferred to pCSV3 to produce plasmid pCSM6. This plasmid was transferred to *Anabaena* (producing strain CSM1) and to CSL16 (producing strain CSL33).

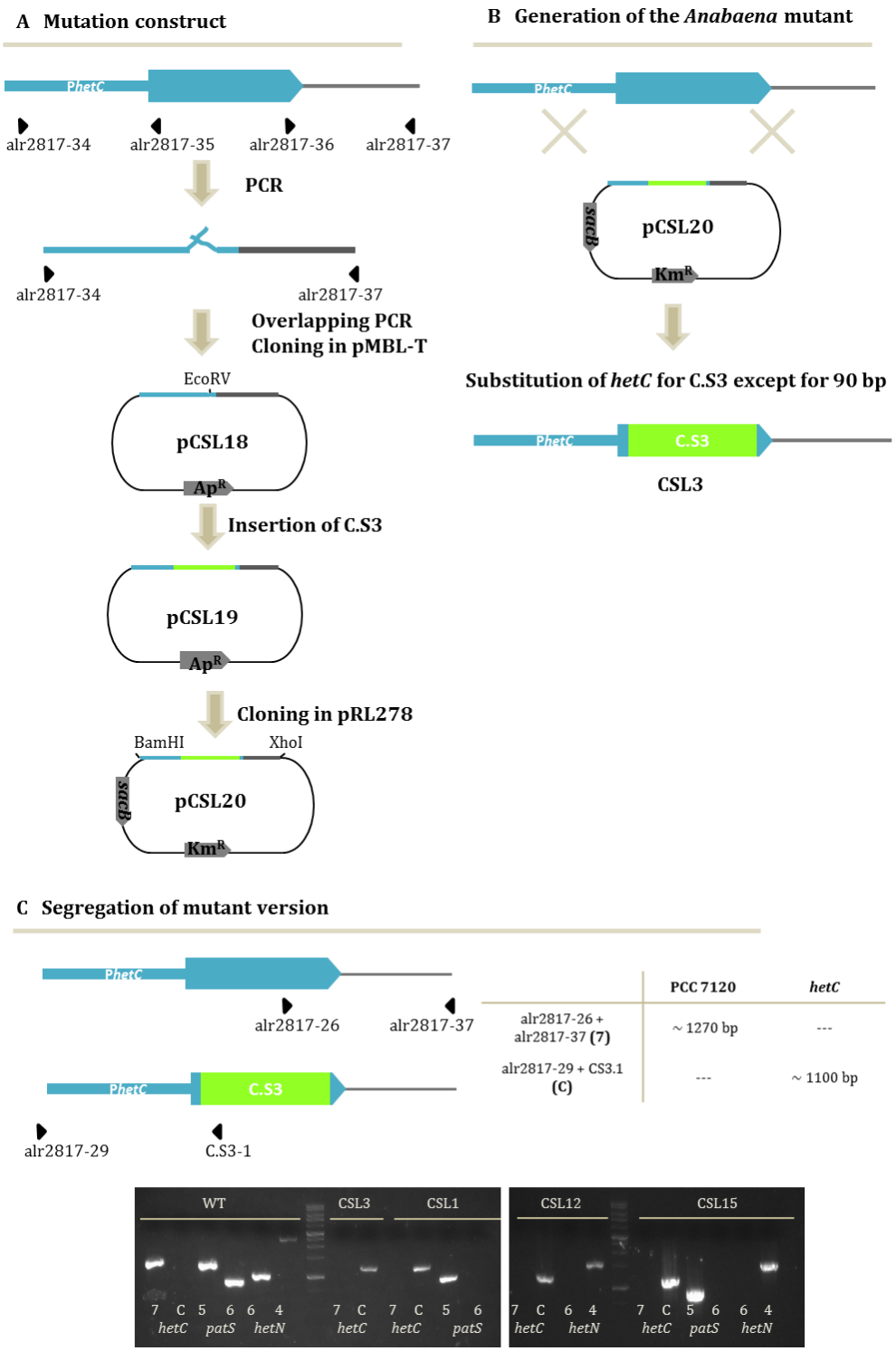


Figure 3.11. Construction and analysis of *hetC* deletion mutants. A. Generation of the genetic construct. The *hetC* gene is represented in blue, and C.S3 cassette in lime green. B. Recombination of the generated construct in *Anabaena*. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild-type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie inside and outside the construct, alr2817-26 and alr2817-37, respectively; and (ii) a pair of primers that lie outside the construct and in the C.S3 cassette, alr2817-29 and C.S3-1, respectively. See text for details.

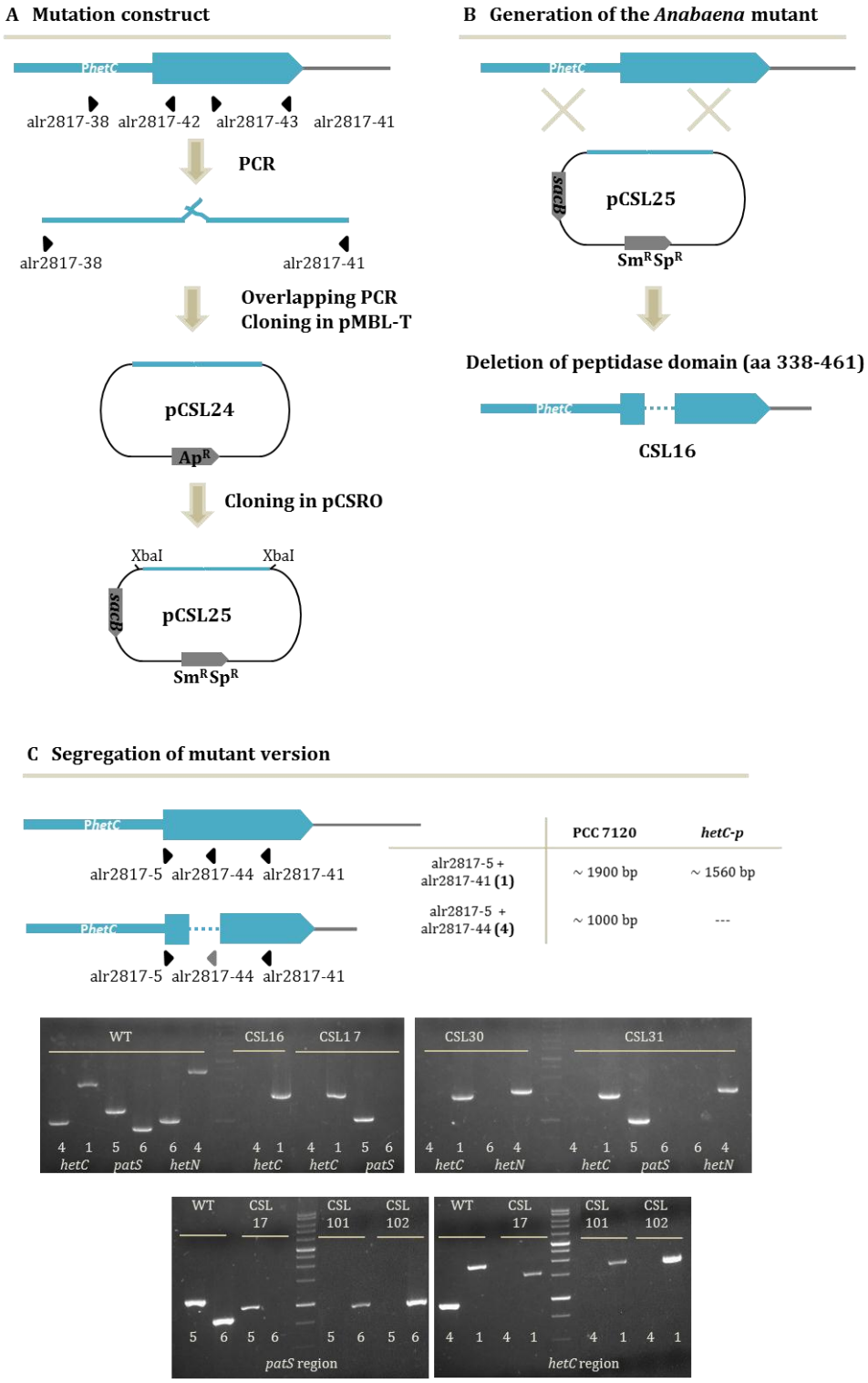


Figure 3.12. Construction and analysis of peptidase domain encoding sequence of *hetC* deletion mutants. A. Generation of the genetic construct. The *hetC* gene is represented in blue. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the deleted region, alr2817-5 and alr2817-41, and (ii) alr2817-5 and a primer that lies inside the region deleted in *hetC-p* mutant, alr2817-44. See text for details.

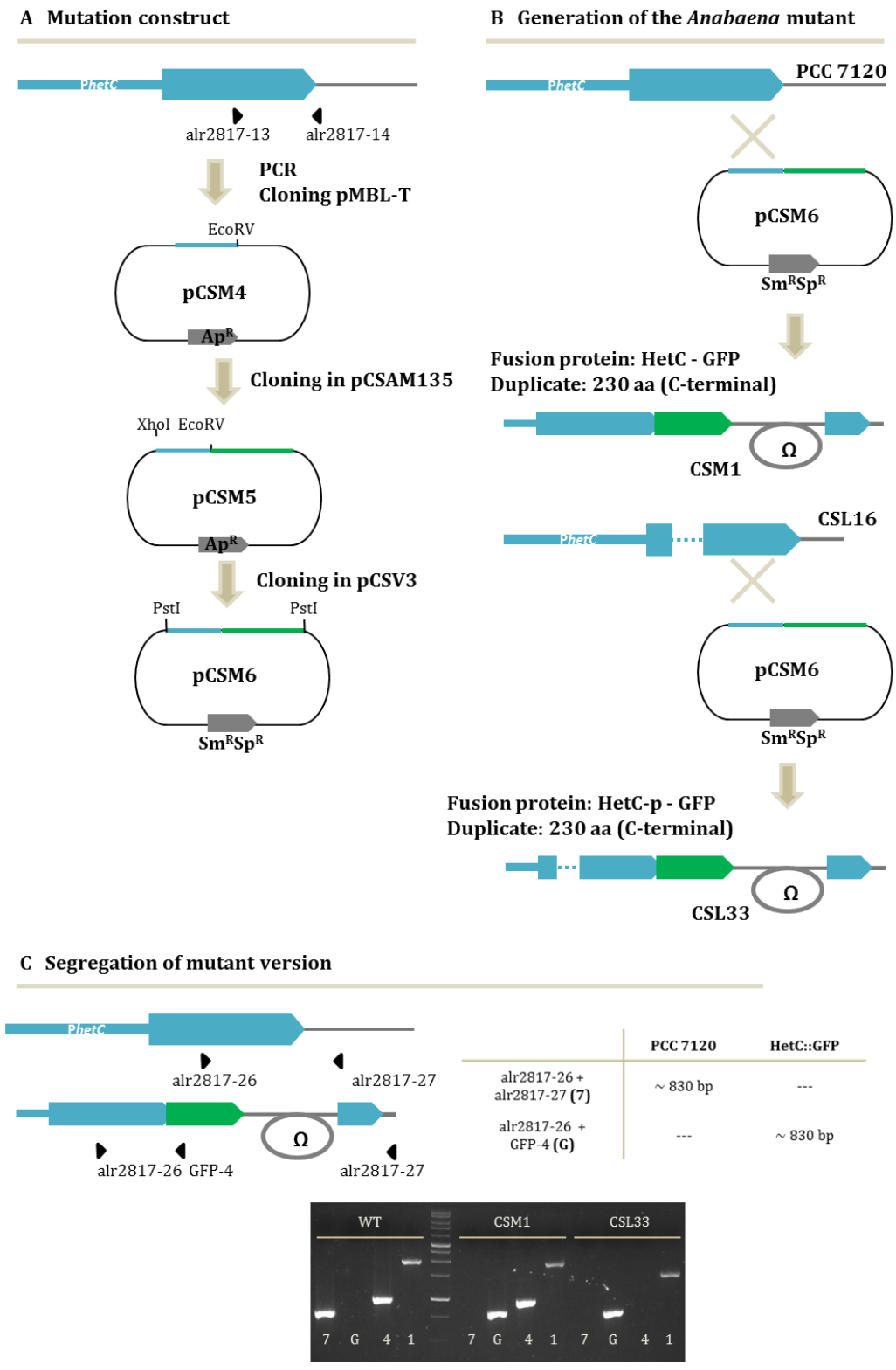


Figure 3.13. Construction and analysis of *hetC-gfp*-mut2 strains. A. Generation of the genetic construct. The *hetC* gene is represented in blue and *gfp*-mut2 in green. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr2817-26 and alr2817-27 and (ii) alr2817-27 and a primer that lies in the *gfp* region, GFP-4. See text for details.

3.4. GENERATION OF *hetP* MUTANT STRAINS

For construction of a *hetP-sfgfp* fusion, the alr2818-8/ alr2818-9 primer pair was used to amplify a fragment comprising 98 bp from upstream of the gene and the whole *hetP* ORF, which was cloned in pCSAL39 plasmid (The same plasmid used for construction of *hetN-sfgfp* described in Section 3.2). After that, the construct was cloned in pCSV3 producing pCSL70, which was transferred to *Anabaena* (producing strain CSL67), strain CSL16 (creating strain CSL68), strain CSVT20 (producing strain CSL69) and strain CSL7 (creating strain CSL70) (Figure 3.14).

For construction of a *hetP-gfp-mut2* fusion, overlapping PCR was performed using (i) pCSL70 as template and primer pairs alr2818-19/GFP-13, or (ii) pCSL68 as template and primer pairs GFP14/GFP-11, were used for an overlapping PCR (Figure 3.15). The resulting DNA fragment was cloned using KpnI in pCSV3 producing plasmid pCSL123. This plasmid was transferred to *Anabaena* producing strain CSL107. Clones resistant to Sm and Sp were selected, and their genomic structure was tested by PCR (Figure 3.14).

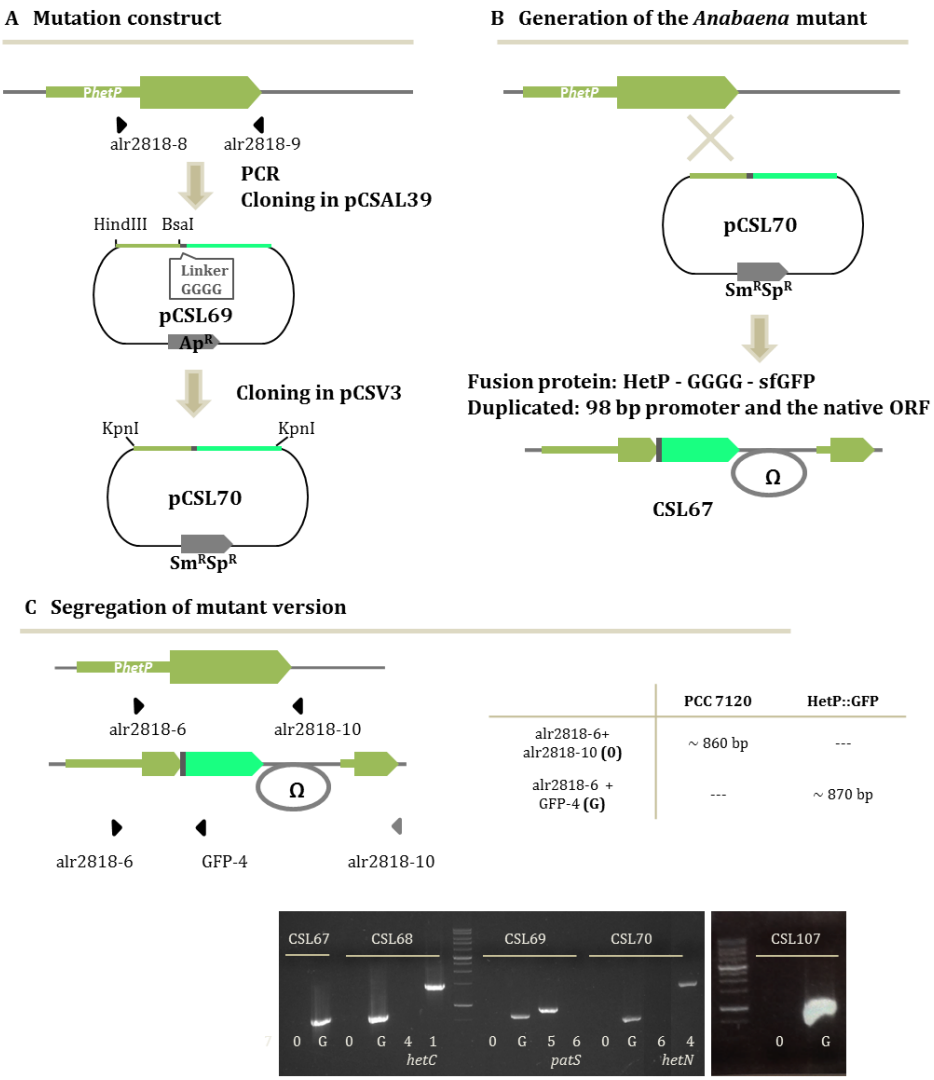


Figure 3.14. Construction and analysis of *hetP-sfgfp* strains. A. Generation of the genetic construct. The *hetP* gene is represented in olive green and *sfgfp* is green. B. Recombination of the generated construct in *Anabaena*. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with a pair of primers that lies outside the construct, alr2818-6 and alr2818-10 and a primer that lies in the *gfp* region, GFP-4. See text for details. The analysis of strain CSL107 (See Figure 3.15) is also shown.

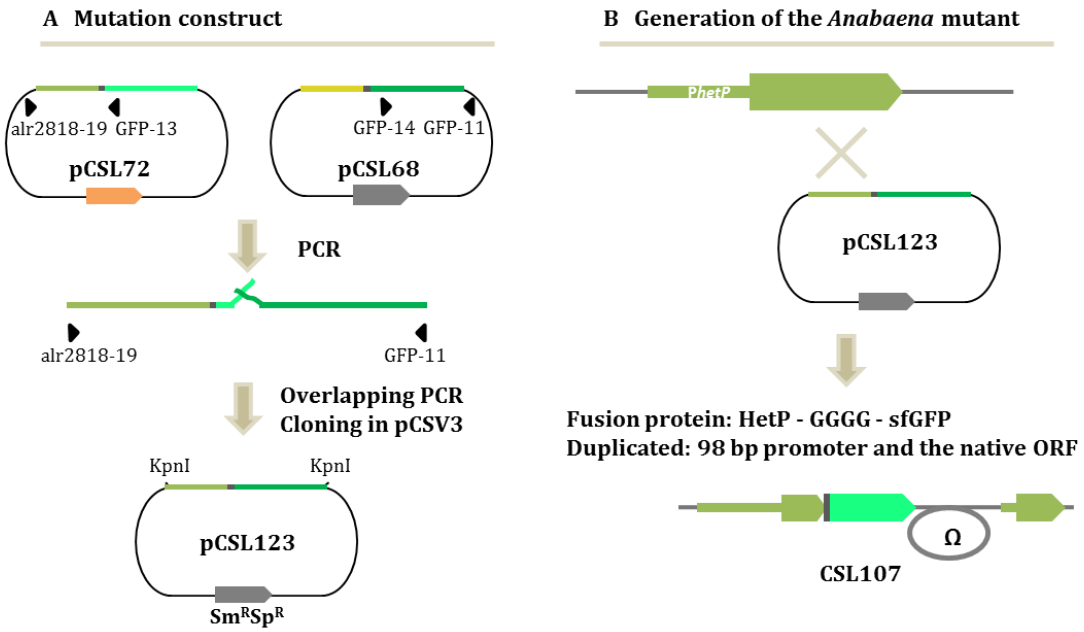


Figure 3.15. Construction of *hetP-gfp*-mut2 strain. A. Generation of the genetic construct. The *hetP* gene is represented in olive green and *gfp*-mut2 is green. B. Recombination of the generated construct in *Anabaena*.

3.5. CONSTRUCTION OF *asr2819* MUTANT STRAINS

To study if the *asr2819* ORF is actually a gene and/or is related to heterocyst differentiation, a series of mutants were generated. An *asr2819* deletion strain was constructed in which the whole ORF, except for 15 bp, was eliminated (Figure 3.16). Two PCR fragments comprising sequences upstream and downstream from the gene were amplified using primer pairs alr2818-13/asr2819-1 and asr2819-2/ alr2820-1, cloned in the pCSRO plasmid and sequenced. The resulting plasmid, pCSL111, was transferred to *Anabaena* (producing strain CSL97) and strain CSL16 (producing strain CSL98), and double recombinants were selected and tested for segregation of the mutation (Figure 3.16; Figure 3.17).

A strain in which *asr2819* was deleted and substituted by the C.S3 cassette was additionally generated (Figure 3.17). The same overlapping PCR, as explained above, was performed and the product was cloned in pCSBN1 and sequenced. The C.S3 cassette was introduced between the two fragments in an EcoRV site. The plasmid generated, pCSL112, was transferred to *Anabaena* (producing strain CSL99) and to CSL16 (producing strain CSL100), and double recombinants were selected and tested for segregation of the mutation.

A strain in which *asr2819* is overexpressed was also generated (Figure 3.18). The whole *asr2819* ORF was amplified using primer pair asr2819-3/asr2819-4 and cloned in pCSAV76 generating plasmid pCSL114. This plasmid bears the *nucA-nuiA* region of *Anabaena* followed by the C.K3 gene-cassette (Elhai & Wolk, 1988) with a BglII site in the 3' end in which the *Anabaena* DNA fragment was inserted. After sending to *Anabaena*, this construct would integrate in the *nucA-nuiA* region, which lies in the α-megaplasmid of *Anabaena*, so that the chromosomal copy of *asr2819* will remain intact. The plasmid pCSL114 was transferred to *Anabaena* (producing strain CSL95) and to CSL16 (producing strain CSL96).

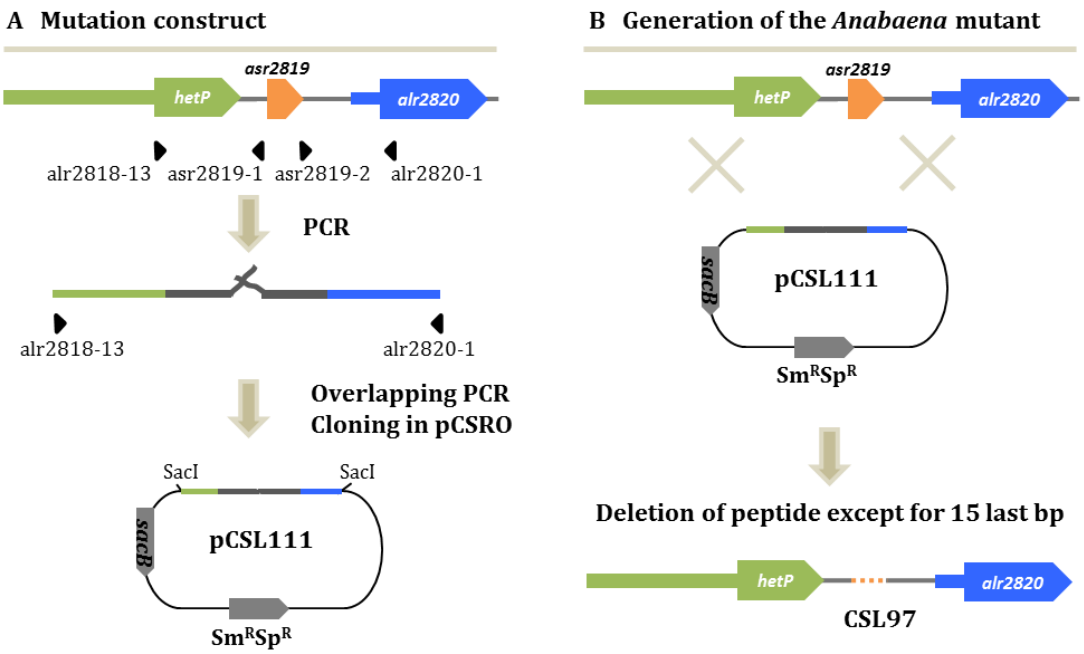


Figure 3.16. Construction of an *asr2819* deletion mutant. A. Generation of the genetic construct. The *asr2819* gene is represented in orange, the *hetP* gene in olive green and alr2820 in blue. B. Recombination of the generated construct in *Anabaena*. The deleted region is represented as a discontinuous line. See text for details.

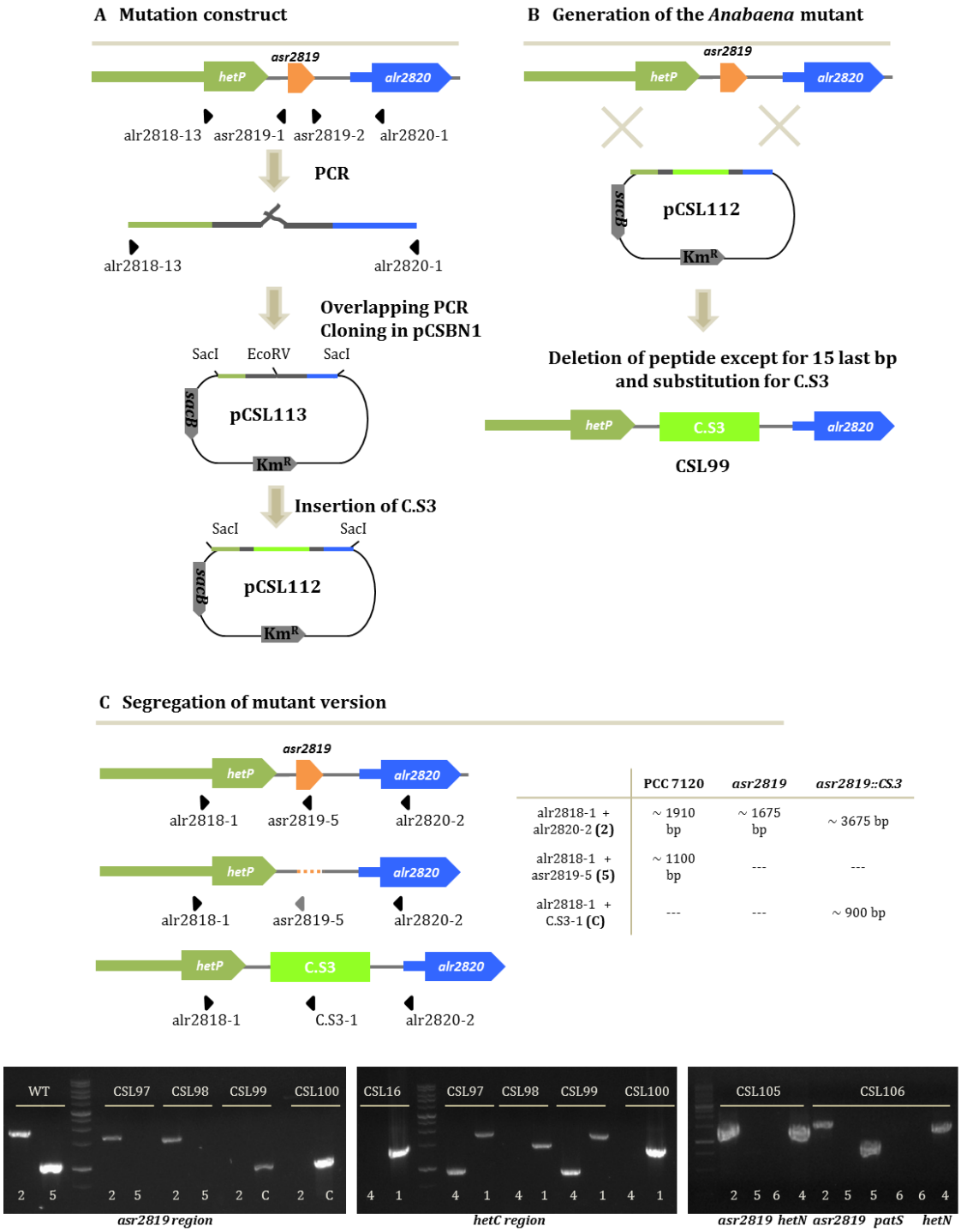
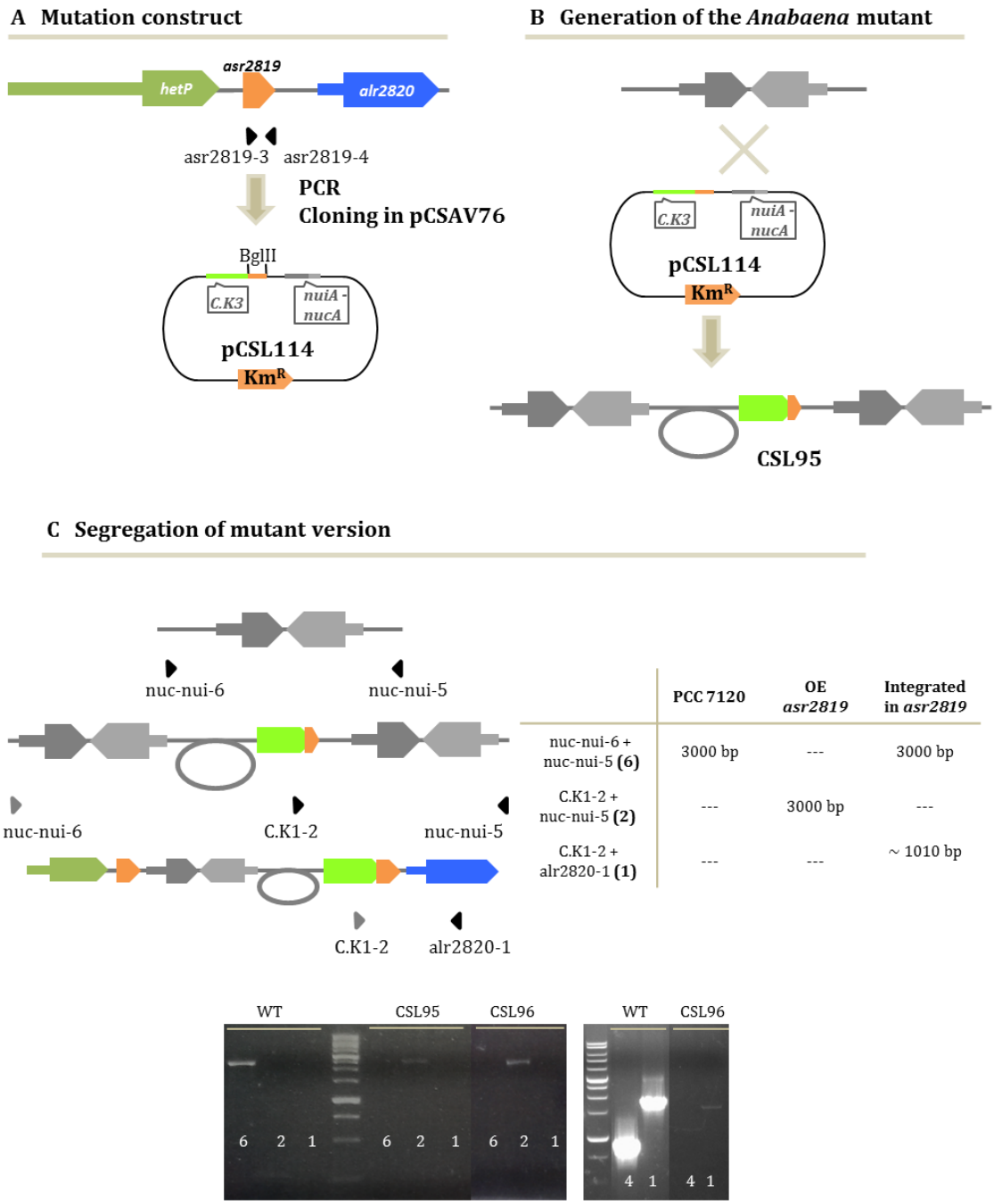


Figure 3.17. Construction and analysis of *asr2819* deletion mutants. A. Generation of the genetic construct. The *asr2819* gene is represented in orange, the *hetP* gene in olive green, alr2820 in blue and C.S3 cassette in lime green. B. Recombination of the generated construct in *Anabaena*. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr2818-1 and alr2820-2, (ii) alr2818-1 and a primer that lies in the deleted region, asr2819-5 and (iii) alr2818-1 and a primer that lies inside C.S3, C.S3-1. See text for details.



	PCC 7120	OE *asr2819*	Integrated in *asr2819*
nuc-nui-6 + nuc-nui-5 **(6)**	3000 bp	---	3000 bp
C.K1-2 + nuc-nui-5 **(2)**	---	3000 bp	---
C.K1-2 + alr2820-1 **(1)**	---	---	~ 1010 bp

Figure 3.18. Construction and analysis of an *asr2819* overexpression mutant. A. Generation of the genetic construct. The *asr2819* gene is represented in orange, the *hetP* gene in olive green, *alr2820* in blue and *nucA* and *nuiA* genes in grey. B. Recombination of the generated construct in *Anabaena*. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, nuc-nui-6 and nuc-nui-5, (ii) nuc-nui-5 and a primer that lies in C.K3, CK.1-2, and (iii) C.K1-2 and alr2820-1, a primer that lies downstream from *asr2819* gene in the genome. In this way, segregation of the mutated α-megaplasmid and a possible integration in the chromosome can be checked. See text for details.

3.6. GENERATION OF *hetR* MUTANT STRAINS

To generate a strain bearing a fusion of P_{hetR} plus the first 27 bp of the *hetR* ORF with *gfp-mut2*, PCR amplification using *Anabaena* DNA as template and primer pair alr2339-12/alr2339-36 was performed. Primer alr2339-36 lies in the *hetR* sequence and contains a sequence encoding a Gly-Gly-Gly-Gly sequence in addition to the NheI site. The amplified fragment was cloned into pCSAL33, which contains the *gfp-mut2* gene, sequenced and transferred to pCSV3. The resulting plasmid, pCSL68, which bears the P_{hetR}-(5' *hetR*)-*gfp-mut2* construct, was transferred to *Anabaena* sp. PCC 7120 (producing strain CSL64), to strain CSVT20 (producing strain CSL65), to strain CSSC2 [this strain bears a deletion of the complete *hetR* gene; Camargo *et al.* (2012)] (producing CSL86), to strain CSL17 (producing CSL87) and to strain CSL16 (producing CSL88) (Figure 3.19).

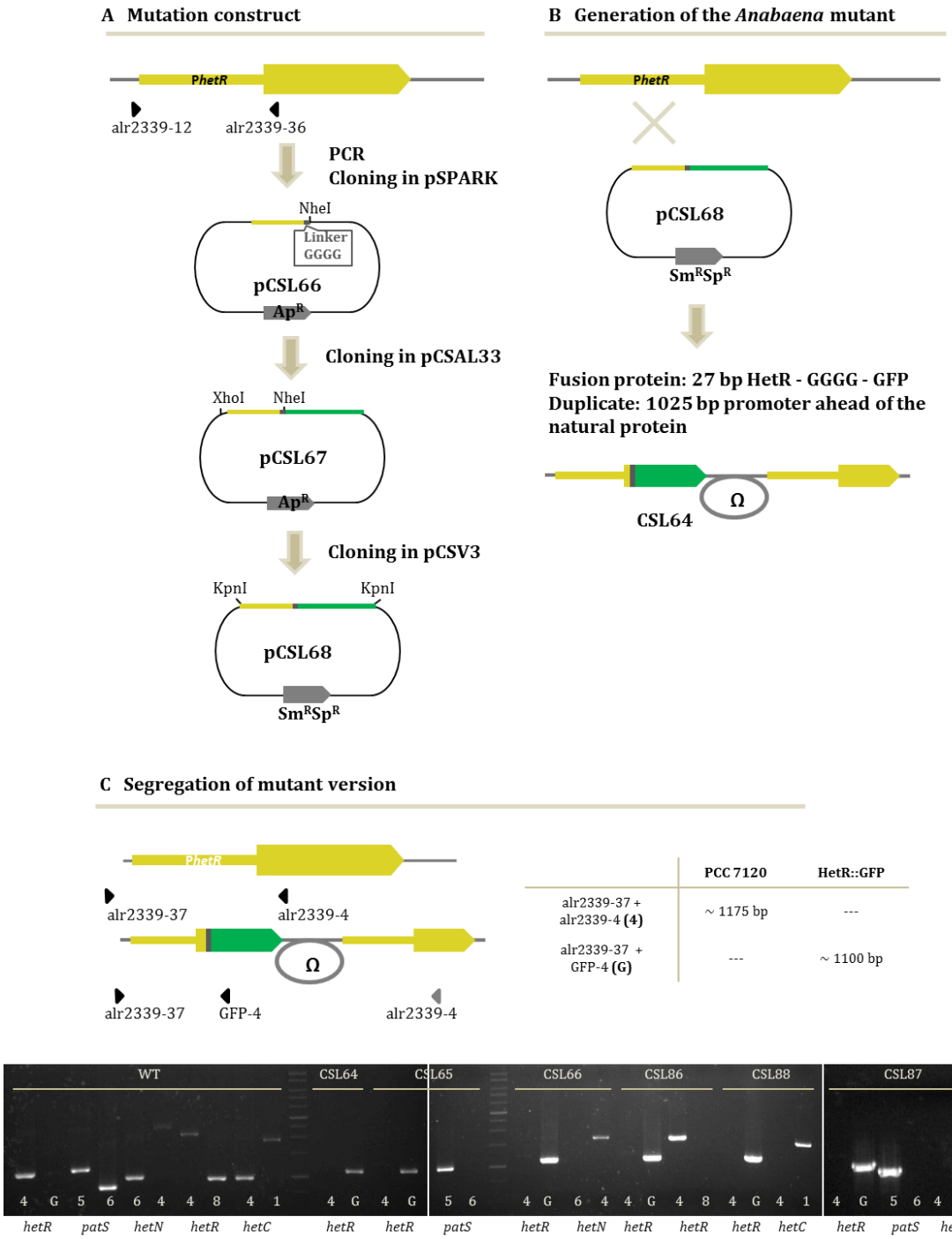


Figure 3.19. Construction of HetR-GFP reporter strains. A. Generation of the genetic construct. The *hetR* gene is represented in yellow and the *gfp*-mut2 in green. B. Recombination of the generated construct in *Anabaena*. C. Analysis of the mutant chromosomes. The scheme of the resulting mutant genomic region and the wild type are represented. The expected size bands in the PCR are indicated in the table. PCR products corresponding to wild-type or mutant bands are shown. PCR was performed with (i) a pair of primers that lie outside the construct, alr2339-37 and alr2339-4, and (ii) alr2339-37 and a primer that lies in the *gfp* region, GFP-4. The deletion of *hetR* in the CSSC2 background was checked with (i) a pair of primers that lie outside the construct, alr2339-23 and alr2339-24 (indicated as primer pair 4), and (ii) alr2339-23 and a primer that lies in the deleted region, alr2339-28 (indicated as primer pair 8). See text for details.

4. RESULTS

4.1. STUDY OF THE PatS PEPTIDE

The *patS* gene would encode a peptide of 17 or 13 amino acids (depending on which of two possible start codons is used), in which the last 5 or 6 amino acids have been demonstrated to inhibit the DNA binding activity of HetR (Yoon & Golden, 1998; Feldmann *et al.*, 2011). It is also known that PatS is expressed in early proheterocysts to inhibit the differentiation of neighbouring vegetative cells. Therefore the PatS peptide has been proposed to be transferred through the filament by unknown mechanisms. Many regulatory peptides involved in quorum-sensing in other bacteria are synthetized as pre-peptides, processed and transported to the target cells to inhibit cellular processes in them. Because of that, it seems possible that in the *Anabaena* filament the PatS peptide is processed in the producing cell.

In this work, we intended to study the function of the different amino acids in PatS by fusion of the peptide to the GFP or the His tag and mutation of each amino acid. Additionally, we studied the native localization of the peptide by immunogold labelling and immunofluorescence.

The phenotype of the *patS* mutants generated in this work can be readily investigated (Figure 4.1). If the *patS* version is functional, it will complement the deletion of CSVT20 to a wild-type phenotype. If the mutation produces a PatS peptide unable to inhibit differentiation, the strain will have an Mch phenotype. We could hypothesize that a PatS version that could not be processed or exported, if it actually happens, should stay in the producing cell and inhibit the differentiation of this cell ("autoinhibition").

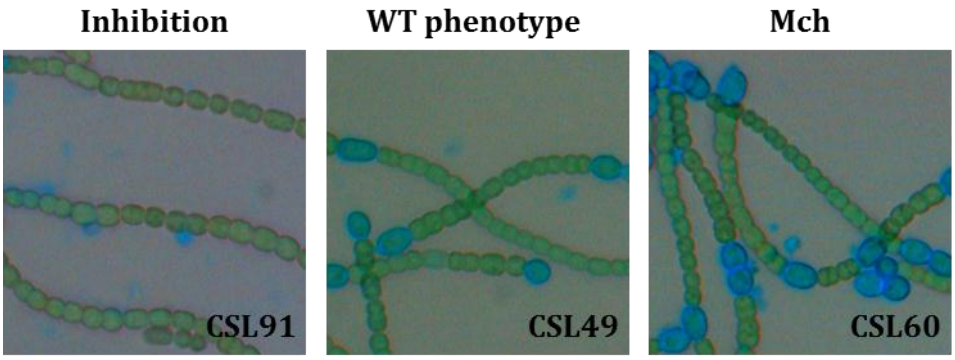


Figure 4.1. Alcian Blue staining of *patS* mutant strains. Ammonium grown filaments were incubated for 24 hours in bubbled cultures with BG110 (-N) medium. Samples from these cultures were stained with Alcian Blue 1:10 and visualized in the light microscope with a magnification of 40x. Micrographs show examples of the three different phenotypes found in the mutants.

4.1.1. Phenotype of strains expressing PatS-GFP and His tag fusion proteins

Three strains that express versions of PatS with GFP fused in frame after the first M-residue of the ORF (CSL21), the fifth residue [M] (CSL19) or in the C-terminal end (CSL20), in addition to two strains bearing 6His fusions after the fifth residue (CSL22) or in the C-terminal end (CSL23), were constructed.

The GFP fluorescence was studied in the three strains bearing GFP fusions using as a control strain CSVM17 that bears the *patS* promoter fused to the *gfp-mut2* gene inserted in a different region of the genome (the α megaplasmid) (Mariscal *et al.*, 2007). Filaments grown in ammonium-containing medium that were shifted to a medium without combined nitrogen were used for observation in a confocal microscope (Figure 4.2).

When incubated in the presence of combined nitrogen, a low level of fluorescence was observed in most cells in the four strains. The green fluorescence increased first in clusters (not shown) and after ca. 5 hours of nitrogen deprivation in specific cells. The fact that both strains CSL21 (that bears the *gfp* gene after first M-encoding codon) and CSL19 (that bears *gfp* after the second M-encoding codon) produced a similar fluorescence pattern indicates that the peptide can be translated from the first methionine.

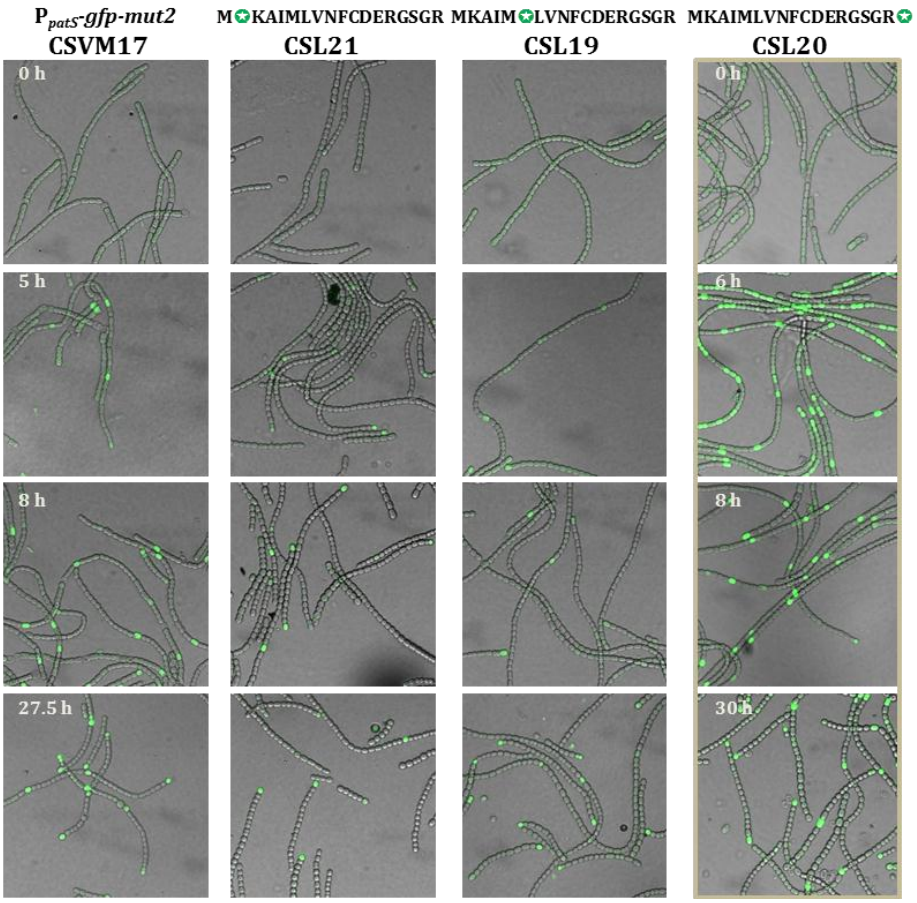


Figure 4.2. Expression of PatS-GFP fusions. Fluorescence of GFP in strains CSVM17 (P_{patS}-GFP in the α megaplasmid, used as a control), CSL21 (GFP inserted after M1), CSL19 (GFP after M5) and CSL20 (GFP in the C terminus) at the indicated times after N step-down is shown. Bright field is

merged with GFP fluorescence. Images from strain CSL20 correspond to an independent experiment, performed in the same conditions.

Additionally, some phenotypic characteristics of these strains, as well as of the 6His fusion strains, were studied (Table 4.1). Nitrogenase activity was measured under anoxic conditions at 24 hours after nitrogen deprivation. Growth rates using ammonium as a nitrogen source and under diazotrophic conditions were determined in liquid shaken cultures.

Strains CSL21, CSL19 and CSL22, all bearing fusions in the N-terminal region of the peptide, presented low nitrogenase activity levels under anoxic conditions and a diminished growth rate in comparison to the wild-type strain. Strains CSL20 and CSL23, both having a tag in the C-terminal end, presented a phenotype similar to the wild type, although the activity and growth rates were still lower than those of the wild type. In all cases the His tag had less effect on the function of the peptide than the GFP fusion.

Table 4.1. Phenotypic characteristics of mutants producing PatS-tagged peptides.

Mutant version of PatS	Strain	Nitrogenase activity nmol (μg chlorophyll *a*)$^{-1}$ h^{-1}	Growth rate (μ, day^{-1}) Ammonium	Dinitrogen
	PCC 7120	19.69 ± 0.68	0.53 ± 0.03	0.45± 0.02
Δ [180 bp]	CSVT20	12.02 ± 0.86	0.49 ± 0.07	0.36 ± 0.05
M**[GFP]**KAIMLVNFCDERGSGR	CSL21	0.63 ± 0.11	0.53 ± 0.06	0.16 ± 0.08
MKAIM**[GFP]**LVNFCDERGSGR	CSL19	0.02 ± 0.01	0.58 ± 0.08	0.09 ± 0.05
MKAIM**[H$_6$]**LVNFCDERGSGR	CSL22	2.46 ± 1.13	0.64 ± 0.03	0.28 ± 0.05
MKAIMLVNFCDERGSGR**[GFP]**	CSL20	4.86 ± 0.49	0.57 ± 0.06	0.36 ± 0.02
MKAIMLVNFCDERGSGR**[H$_6$]**	CSL23	13.59 ± 1.28	0.58 ± 0.05	0.47 ± 0.08

Data are the mean and the standard deviation of the mean of the results of two to four experiments performed with independent cultures (nitrogenase) or of four independent cultures (growth rates). The wild-type strain and *patS* deletion mutant are shown as controls.

The spatial pattern of heterocyst distribution was also studied in these strains (Table 4.2; Figure 4.3). To facilitate the counting, cells were stained with Alcian Blue, which stains the polysaccharide layer of the heterocysts, before visualization in the light microscope (see Figure 4.1). Strains with a low percentage of heterocysts are not represented in the figure since not enough intervals could be counted.

Strains CSL21 and CSL19 formed less heterocysts than the wild type, with heterocyst frequencies between 0.5% and 3%. Strain CSL22 also showed a diminished frequency of differentiated cells, but not as low as strains CSL21 or CSL19. Strain CSL20 showed a pattern similar to that of CSL22, and strain CSL23 exhibited an almost wild-type phenotype. Thus, in strains bearing insertions in the N-terminal region of the peptide, the frequency of heterocysts is more reduced than in those bearing C-terminal insertions.

4.1.2. Point mutations of the PatS peptide

In order to get insight into the function of the different peptide regions, strains bearing substitutions of each amino acid in PatS were generated. The pattern of heterocysts was studied in these strains (Table 4.2; Figure 4.3). A control strain (CSL48) with the native *patS* gene integrated in the chromosome was also generated from the CSVT20 background.

Strain CSL44 ($M^{1}A$) showed a Mch phenotype, while strain CSL49 ($M^{5}A$) showed a phenotype similar to that of the wild type (see CSL49 in Figure 4.1), indicating that M^{1} is essential for the inhibitory function of the peptide. This, in addition to the result obtained from the fusion with the green fluorescent protein in strain CSL21, indicates that the first methionine is necessary and sufficient as a translation initiation point of PatS.

The substitution $K^{2}A$, $A^{3}Q$ and $I^{4}Q$ in strains CSL90, CSL91 and CSL92, respectively, produced inhibition of differentiation (see CSL91 in Figure 4.1), whereas substitution $L^{6}Q$ (strain CSL50) had almost no effect in the heterocyst pattern. Substitution $V^{7}Q$ (in strain CSL51) produced a total inhibition of differentiation; however the $V^{7}A$ mutation in strain CSL93 produced a less drastic effect. Strain CSL62 that bears a double substitution $M^{5}V$ $V^{7}Q$ showed a heterocyst frequency ca. 0.6 that of the wild-type strain.

These results are consistent with those obtained with the fusion of tags in the N-terminal region of the peptide (strains CSL21, CSL19, CSL22). The fact that changes in this region lead to inhibition of heterocyst differentiation suggests accumulation of the gene product in the producing cells and, thus, that this region is involved in the processing and/or the export of the peptide. The V^{7} substitution produced the most severe phenotype, indicating that this residue must be critical for those functions.

Substitutions $N^{8}E$ and $F^{9}Q$ in strains CSL52 and CSL53, respectively, produced a phenotype similar to the wild-type phenotype, although the mean interval size was slightly higher and the presence of doublets (contiguous heterocysts) was rare. Thus, these amino acids may also participate in processing and/or transport of the peptide.

The substitutions $C^{10}A$ and $E^{12}A$ (in strains CSL54 and CSL56) produced an increase in the percentage of contiguous heterocysts. Substitution $D^{11}A$ (strain CSL55) had no effect in the pattern. Strains CSL57 and CSL58 bearing changes $R^{13}A$ and $S^{15}A$, respectively, exhibited phenotypes comparable to that of the *patS* deletion strain CSVT20. Substitutions $G^{14}A$ (CSL59), $G^{16}A$ (CSL60) and $R^{17}A$ (CSL61) also produced a Mch phenotype (see strain CSL60 in Figure 4.1), although weaker than that observed in strain CSVT20. In general, substitutions in the C-terminal region of the peptide lead to a higher frequency of heterocysts than in the wild type. Consequently, this part appears involved in the inhibitory activity of the peptide, consistent with previous reports (Wu *et al.*, 2004).

Table 4.2. Pattern of heterocysts in *Anabaena* mutants producing different PatS peptides.

Mutant version of PatS	Strain	Percentage heterocysts	Contiguous heterocysts	Mean interval
	PCC 7120	10.31 ± 0.64	4.36 ± 0.58	9.67 ± 0.26
Δ [180 bp]	CSVT20	19.00 ± 2.12	20.81 ± 5.13	5.69 ± 0.38
M**[GFP]**KAIMLVNFCDERGSGR	CSL21	2.94 ± 0.82	0	-
MKAIM**[GFP]**LVNFCDERGSGR	CSL19	0.54 ± 0.33	0	-
MKAIM**[H_6]**LVNFCDERGSGR	CSL22	5.31 ± 0.76	2.27 ± 2.27	13.18 ± 0.59
MKAIMLVNFCDERGSGR**[GFP]**	CSL20	5.76 ± 1.00	5.33 ± 1.28	12.98 ± 0.29
MKAIMLVNFCDERGSGR**[H_6]**	CSL23	13.21 ± 1.29	2.90 ± 1.76	8.49 ± 0.22
MRGSGR	CSL45	20.29 ± 1.58	11.30 ± 0.96	5.59 ± 0.47
MERGSGR	CSL46	13.79 ± 1.12	21.11 ± 2.92	7.40 ± 0.36
MDERGSGR	CSL94	10.01 ± 0.61	12.46 ± 3.99	11.12 ± 0.55
MCDERGSGR	CSL47	9.02 ± 0.85	4.59 ± 0.09	10.83 ± 0.51
MKAIMLVNFCDERGSGR	CSL48	12.39 ± 0.70	6.01 ± 0.66	10.40 ± 0.38
AKAIMLVNFCDERGSGR	CSL44	18.67 ± 0.69	25.55 ± 6.09	5.55 ± 0.01
M**A**AIMLVNFCDERGSGR	CSL90	4.47 ± 0.77	0	-
MK**Q**IMLVNFCDERGSGR	CSL91	2.10 ± 0.95	0	-
MKA**Q**MLVNFCDERGSGR	CSL92	1.97 ± 1.09	0	-
MKAI**A**LVNFCDERGSGR	CSL49	10.56 ± 1.30	6.73 ± 3.05	10.69 ± 0.78
MKAIM**Q**VNFCDERGSGR	CSL50	10.09 ± 0.76	3.75 ± 2.15	10.83 ± 0.29
MKAIML**A**NFCDERGSGR	CSL93	3.43 ± 0.41	0	-
MKAIML**Q**NFCDERGSGR	CSL51	0.03 ± 0.03	0	-
MKAI**V**L**Q**NFCDERGSGR	CSL62	6.01 ± 1.82	0	17.63 ± 1.68
MKAIMLV**E**FCDERGSGR	CSL52	8.35 ± 1.08	1.33 ± 1.33	10.32 ± 0.53
MKAIMLVN**Q**CDERGSGR	CSL53	7.45±1.65	0	14.01±0.97
MKAIMLVNF**A**DERGSGR	CSL54	10.24 ± 0.71	11.15 ± 3.08	9.85 ± 0.24
MKAIMLVNFC**Q**ERGSGR	CSL55	12.02 ± 2.05	4.76 ± 2.10	9.68 ± 0.77
MKAIMLVNFCD**A**RGSGR	CSL56	11.46 ± 0.92	9.08 ± 4.88	8.12 ± 0.21
MKAIMLVNFCDE**A**GSGR	CSL57	22.75 ± 2.33	25.16 ± 5.24	4.67 ± 0.12
MKAIMLVNFCDER**A**SGR	CSL59	15.30 ± 2.38	8.08 ± 1.82	7.47 ± 0.46
MKAIMLVNFCDERG**A**GR	CSL58	24.62 ± 3.79	24.27 ± 4.90	4.85 ± 0.46
MKAIMLVNFCDERGS**A**R	CSL60	16.92 ± 1.52	26.17 ± 4.03	4.98 ± 0.17
MKAIMLVNFCDERGSG**A**	CSL61	16.03 ± 1.56	9.75 ± 1.98	8.14 ± 0.21

The amino acid sequence of the PatS peptide produced in each strain is shown (the wild-type peptide sequence is MKAIMLVNFCDERGSGR). When appropriate, the location of the inserted GFP protein or 6His tag is indicated between brackets. In the case of substitution mutations, the amino acid introduced is written in bold. Heterocyst frequency (as percentage of total cells), mean size of vegetative cell intervals between heterocysts and percentage of contiguous heterocysts (interval size = 0, percentage of total intervals) are indicated for each strain.

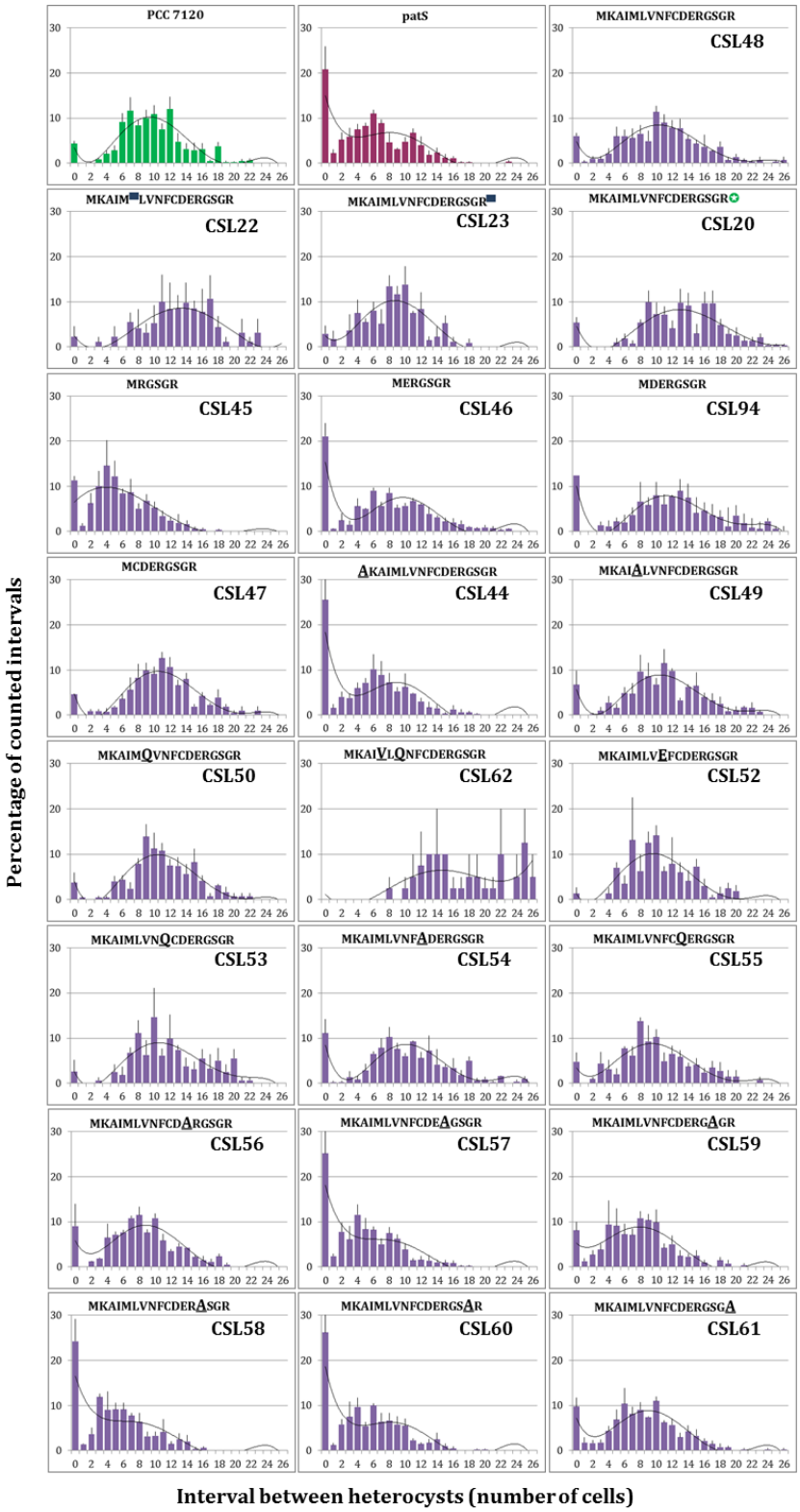


Figure 4.3. Heterocyst distribution in *Anabaena* and *patS* mutant strains. The frequencies of size intervals corresponding to the counting of Table 4.2 are shown in the graphs. For each strain, insertions (GFP as a green star or His tag as a blue square) or point substitutions (substituted amino acid indicated in bold) are highlighted.

4.1.3. PatS minigenes

To further study the parts of PatS responsible for the activity of inhibition of heterocyst differentiation, strains bearing minigenes encoding peptides with a Met residue followed by the last 5, 6, 7 or 8 amino acids of PatS were constructed (strains CSL45, CSL46, CSL94 and CSL47 respectively).

Strains CSL45 (PatS-5) and CSL46 (PatS-6) showed a Mch phenotype (Table 4.2; Figure 4.3). CSL94 (PatS-7) presented wild-type values except for an increased frequency of doublets. In contrast, strain CSL47 bearing the octapeptide showed a phenotype similar to that of the wild-type strain. Therefore, the octapeptide expressed under the *patS* native promoter reproduces the activity of the whole 17-amino acid peptide.

4.1.4. Immunolocalization of PatS

The location of the native PatS peptide was approached by immunolocalization. Antibodies against a synthetic PatS-5 peptide were generated and used to detect the peptide in the *Anabaena* filament by immunogold labelling and immunofluorescence.

Immunogold labelling was carried out in filaments grown in liquid shaken cultures in the presence of ammonium that were incubated for 12 hours in $BG11_0$ medium (Materials and Methods 2.6.4). The anti-PatS antibody at a concentration of 1:50 and a secondary IgG antibody bound to 10-nm gold particles were used. Strains PCC 7120 and CSVT20 (*patS* mutant) were studied. PCC 7120 treated without primary antibody was used as a negative control (Figure 4.4).

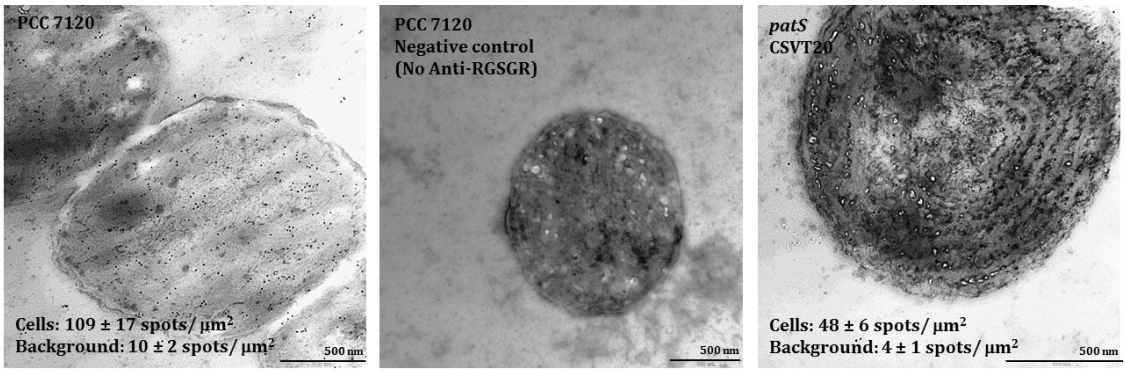


Figure 4.4. Immunogold localization of the PatS-5 peptide. Transmission electron microscopy micrographs detecting 10-nm gold particles attached to antibodies raised against the RGSGR peptide.

Black spots corresponding to PatS-5 peptide antibody deposition were observed all over the cell cytoplasm in strain PCC 7120. No spots were observed in the negative control treated without anti-PatS antibody. In the *patS* mutant the number of spots counted was less than half that in strains PCC 7120 and they are probably due to unspecific biological fixation or to fixation to other proteins that contain the PatS-5

peptidic sequence (RGSGR). Due to extensive filament fragmentation, we could not identify proheterocysts in the samples prepared for microscopy. Furthermore, we could not see the spots concentrated at any particular region of the cytoplasm or the periplasm. Thus, under the conditions used, this technique did not provide useful information about the localization of the PatS peptide.

In order to study the localization of PatS in longer filaments, an immunofluorescence approach was developed. *Anabaena* cultures were grown in ammonium-containing medium and shifted to $BG11_0C$ medium for 8 hours. After that, cells were fixed in poly-L-lysine-coated slides and incubated with the anti-PatS antibody and a secondary antibody conjugated with FITC (Materials and Methods 2.6.3.2). Samples from strains PCC 7120, CSVT20 and CSL45 were visualized by fluorescence microscopy and quantified (Figure 4.5). In CSVT20 all the cells presented the same level of basal fluorescence (96 cells studied), which was subtracted from the levels quantified for the other strains.

In the wild-type strain, different fluorescence levels were found in different cells. Proheterocysts and heterocysts (distinguished by size, granularity or an external layer) usually exhibited levels similar to those of the CSVT20 strain. In contrast, the fluorescence in cells contiguous to proheterocysts was higher. Fluorescence levels were quantified in 301 cells, showing a gradient of decreasing fluorescence away from the proheterocyst (Figure 4.5.B). This result supports the notion that the PatS peptide, which is produced in proheterocysts, is transferred to the neighbouring cells.

Strain CSL45, which bears the minigene encoding MRGSGR and shows a Mch phenotype, exhibited a totally different pattern. Proheterocysts and its close neighbours presented high levels of fluorescence whereas other vegetative cells showed lower fluorescence levels (185 cells measured). Therefore, the PatS version produced in this strain remains at a substantial level in the producing cell.

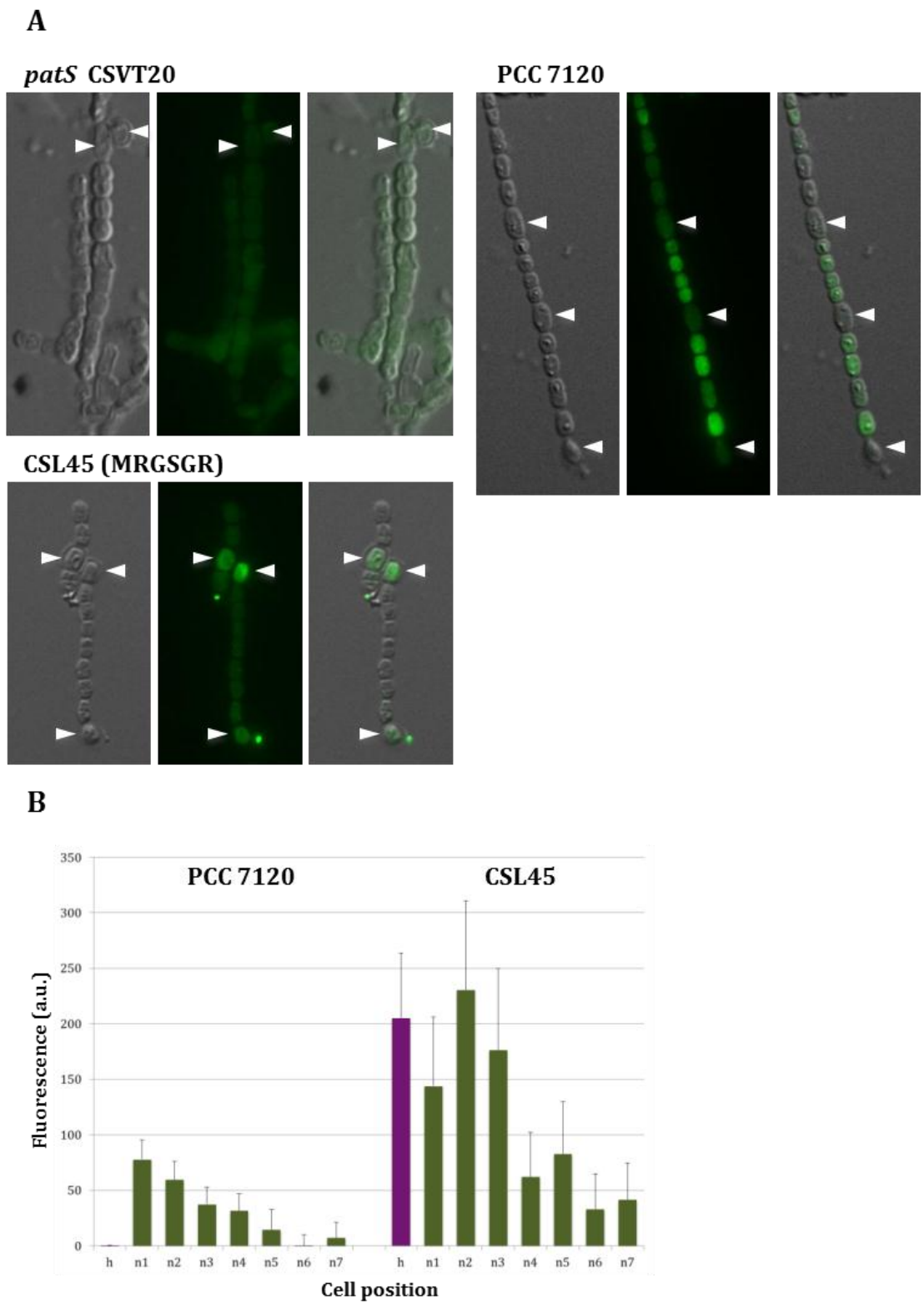


Figure 4.5. Immunofluorescence detection of the PatS-5 peptide. Filaments of the indicated strains were incubated in BG110C medium for 8 hours and tested with antibodies raised against the RGSGR peptide. A. Visualization of the filaments by fluorescence microscopy. Bright field, green fluorescence and a merged image are shown. B. Quantification of immunofluorescence of heterocysts (h) and adjacent cells (n1-n7 are neighbours 1 to 7) in PCC 7120 and strain CSL45. The values represented are the mean of the fluorescence recorded in each position less the mean of the fluorescence measured in strain CSVT20, which was used as background level. Bars indicate the standard deviations of the mean.

4.2. STUDY OF THE HetN PROTEIN

HetN is a negative regulator of the differentiation of heterocysts, but its mechanism of action has not been established yet. We studied the expression and localization of the HetN protein by creating deletions in different regions of the gene and analysing fusions to the GFP.

4.2.1. Bioinformatic analysis of HetN

HetN had been previously detected in *Anabaena* membrane fractions by western blot (Li *et al.*, 2002) and through fusions with YFP (Higa *et al.*, 2012). Because it might be a membrane-associated protein, the topology of HetN was analysed by different programs. Some of these programs predicted a transmembrane segment, but its position was variable since the probability of its occurrence is not very high (Figure 4.6.A). The possibility that HetN bore a signal peptide was studied as well. SignalP 3.0 was the only program predicting a signal peptide, from amino acid 2 to 27, with a cutting site at residue 28 (Figure 4.6.B).

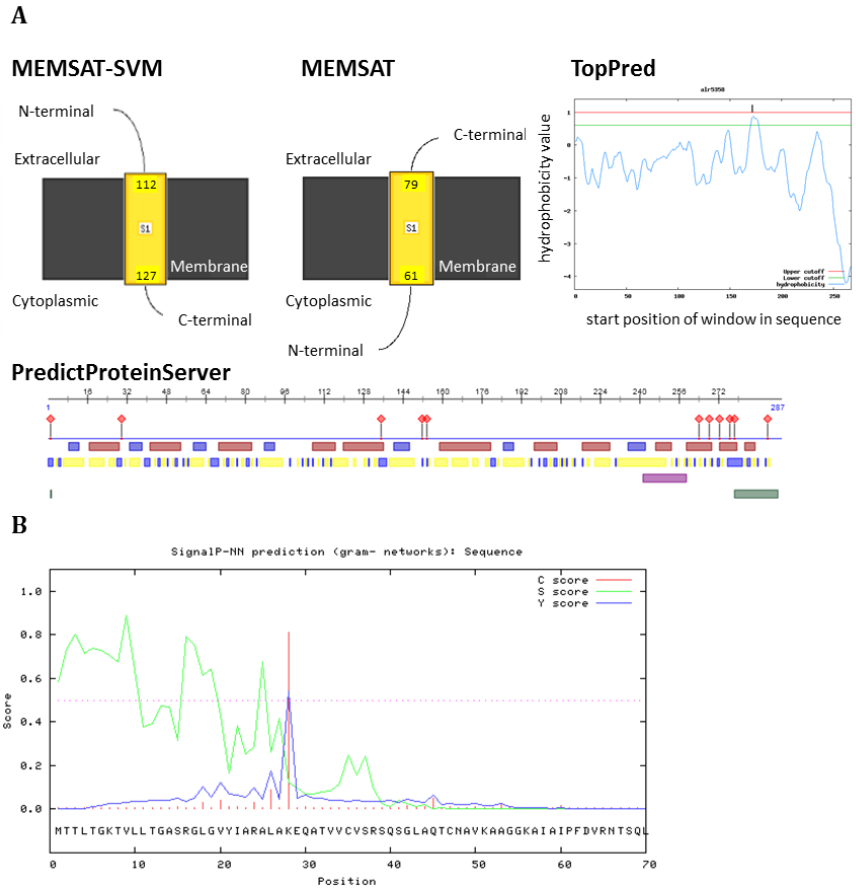


Figure 4.6. Topological predictions for the HetN protein. A. Transmembrane topology predictions with different algorithms. In the PredictProteinServer image, a predicted transmembrane domain is coloured in purple. B. SignalP 3.0 prediction of a putative signal peptide in the HetN sequence.

4.2.2. Phenotype of the *hetN* mutants

A series of strains lacking the whole *hetN* ORF except for 54 bp (*hetN*; strain CSL7 in the wild-type background and strain CSL11 in a *patS* background), lacking only 18 bp encoding the ERGSGR amino acid sequence [*hetN* (ERGSGR), strain CSL32] or lacking the putative signal peptide encompassing amino acids 2-28 (sp*hetN*, strain CSL73) were constructed. In addition, strains expressing a fusion of HetN with the sfGFP (strain CSL71 in the wild-type background, CSL72 in a *patS* background or CSL103 in the sp*hetN* background) or with the GFP (strain CSL106) were generated.

The pattern of heterocyst distribution was studied in these mutant strains. The frequency of size intervals is shown in Figure 4.7, and the mean interval size and the percentages of total and contiguous heterocysts are presented in Table 4.3.

Strain CSL7, lacking most of the *hetN* ORF, showed a Mch phenotype that could already be noticed after 24 hours of nitrogen deprivation, although it became more severe afterwards. This is in part consistent with a previous report describing a delayed Mch phenotype in a *hetN* deletion mutant (Borthakur *et al.*, 2005).

Table 4.3. Pattern of heterocysts in *Anabaena* strains bearing different *hetN* mutations.

Genotype	Strain	Time (h)	Percentage heterocysts	Contiguous heterocysts	Mean interval
	PCC 7120	24	10.39 ± 0.36	3.88 ± 0.57	10.53 ± 0.73
		48	8.62 ± 0.43	6.41 ± 0.99	13.10 ± 0.83
		72	7.53 ± 0.70	2.50 ± 1.91	13.27 ± 0.76
patS	CSVT20	24	18.63 ± 2.21	27.84 ± 8.04	5.47 ± 0.68
		48	17.23 ± 1.01	23.43 ± 6.75	6.84 ± 0.41
		72	17.62 ± 0.97	33.68 ± 2.49	6.04 ± 0.59
hetN	CSL7	24	14.30 ± 1.03	15.67 ± 5.48	6.75 ± 0.10
		48	13.92 ± 0.64	25.04 ± 3.08	6.24 ± 0.20
		72	17.66 ± 1.03	33.66 ± 8.43	5.55 ± 1.19
***hetN* (ERGSGR)**	CSL32	24	13.76 ± 2.40	9.17 ± 4.55	7.50 ± 0.98
		48	14.23 ± 3.22	36.30 ± 7.58	5.46 ± 0.46
		72	14.61± 2.84	32.67 ± 4.61	6.34 ± 0.69
sp*hetN*	CSL73	24	15.88 ± 0.87	5.28 ± 0.92	6.21 ± 0.10
		48	15.20 ± 0.82	11.61 ± 1.59	8.07 ± 0.24
		72	13.86 ± 0.27	14.63 ± 1.21	8.83 ± 0.50
hetN-sfgfp	CSL71	24	10.03 ± 0.12	5.93 ± 0.78	8.69 ± 0.30
		48	12.19 ± 1.96	12.01 ± 4.02	9.76 ± 0.61
		72	11.23 ± 1.40	12.42 ± 1.35	11.11 ± 1.21
hetN-sfgfp patS	CSL72	24	16.29 ± 0.95	20.49 ± 1.61	7.51 ± 0.22
		48	15.73 ± 0.12	23.32 ± 5.21	8.10 ± 0.58
		72	14.22 ± 0.61	16.79 ± 2.35	9.41 ± 0.41
patS hetN	CSL11	24	42.72 ± 4.83	58.16 ± 3.03	1.50 ± 0.23
		48	62.51 ± 10.01	70.63 ± 5.42	1.07 ± 0.32
		72	65.02 ± 6.72	67.23 ± 5.12	1.09 ± 0.02

Heterocyst frequency (as percentage of total cells), mean size of vegetative cell intervals between heterocysts and percentage of contiguous heterocysts (interval size = 0, percentage of total intervals) are indicated for each strain. Wild-type strain and *patS* mutant CSVT20 are shown as controls.

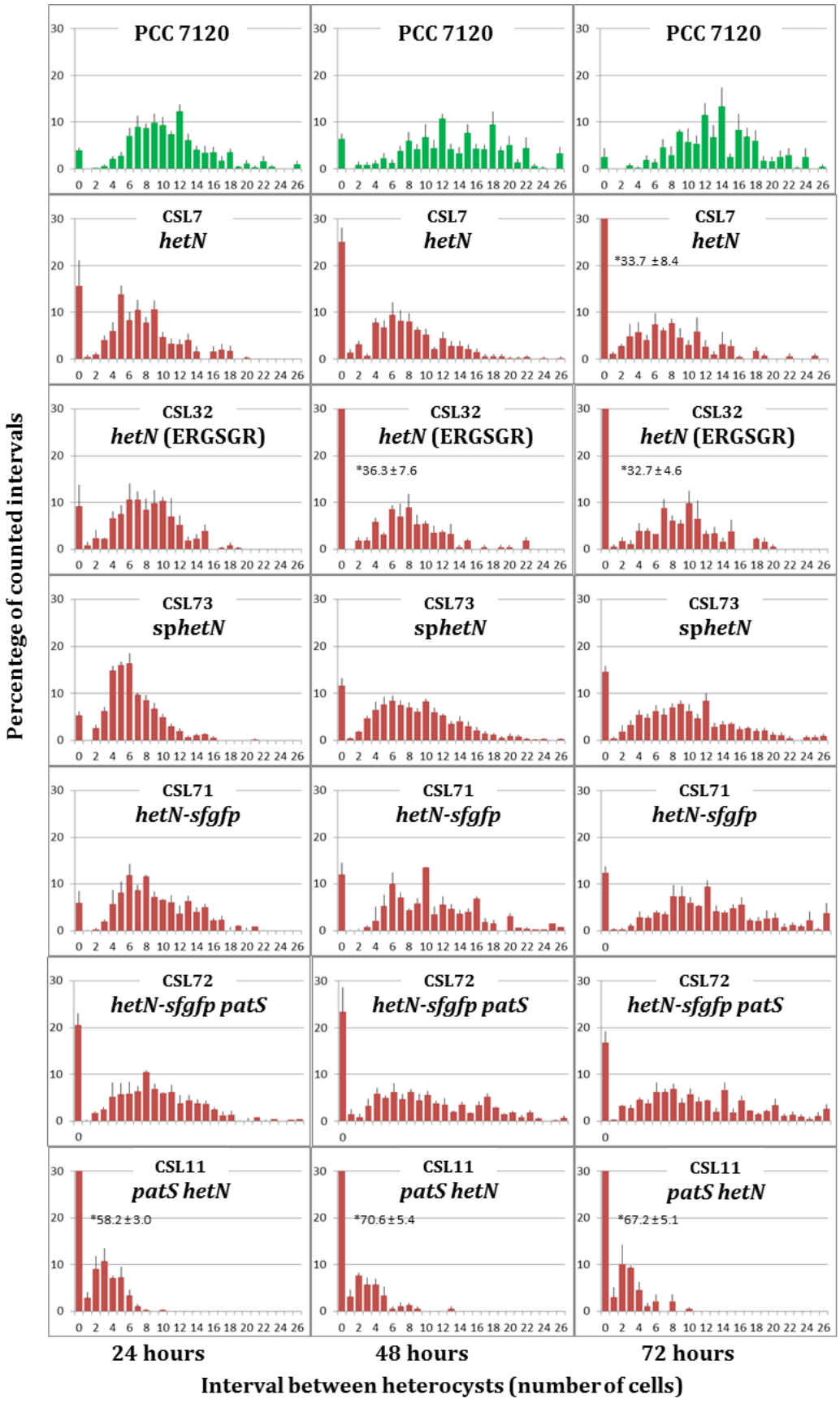


Figure 4.7. Heterocyst distribution in *Anabaena* sp. PCC 7120 and *hetN* mutant strains. The frequency of intervals at 24, 48 and 72 h of nitrogen deprivation is shown in the graphs. In the cases in which the frequency is over 30%, the exact value is indicated.

Deletion of an 18-bp fragment encoding the sequence ERGSGR from within the *hetN* gene (strain CSL32) resulted in a phenotype similar to that produced by deletion of the complete gene. Thus, this sequence is necessary for the activity of HetN at inhibition of heterocyst differentiation.

Mutant CSL73 that produces a HetN protein without the putative signal peptide showed a slight Mch phenotype, which was not as strong as that of strains CSL7 or CSL32. Therefore the predicted signal peptide may have a role, but is not essential, for the function of HetN.

The fusion of sfGFP to HetN produced a small loss of activity since strains CSL71 and CSL72 showed an increased number of contiguous heterocysts in comparison with their parental strains (wild type and *patS* deletion mutant, respectively).

The double mutant *patS hetN*, which lacks the negative regulators of the process of heterocyst differentiation, exhibited the most severe phenotype, reaching frequencies of 60-70% heterocysts after 72 hours of nitrogen step-down, consistent with previous reports (Borthakur et al., 2005).

The ability to grow under diazotrophic conditions was also studied by spot growth tests in solid medium in the different *hetN* mutants (Figure 4.8). All the strains were able to grow under the tested conditions except for the *patS hetN* double mutant. In this strain, lack of diazotrophic growth might be related to an impaired carbon-to-nitrogen balance as a consequence of the very high frequency of heterocysts compared to vegetative cells.

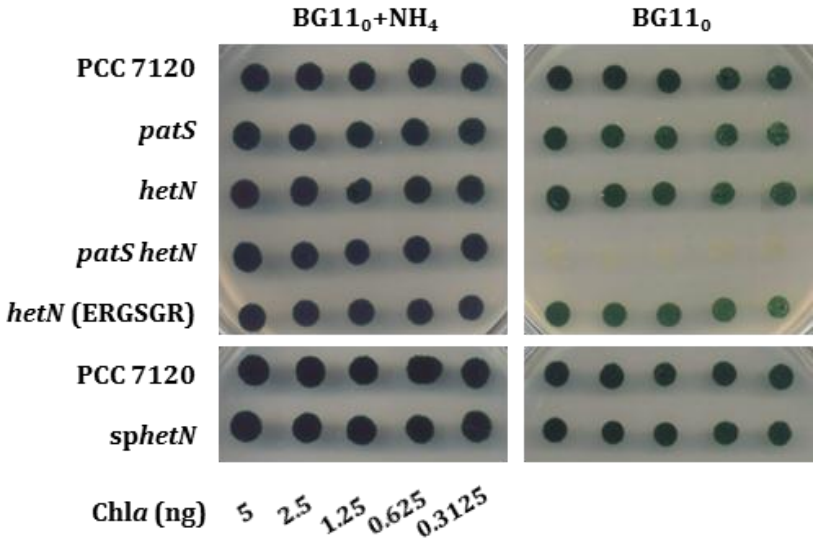


Figure 4.8. Solid growth test of *hetN* mutants. Filaments from flask cultures in ammonium-containing medium were washed with BG110 medium and dropped atop plates with or without ammonium in the medium. Drops spotted (5 µL) contained the quantities indicated of Chl (the quantities in BG110 medium were the same as in BG110+NH4+ medium). The plates were incubated for 14 days under standard growth conditions and photographed.

4.2.3. Localization of HetN-GFP fusion proteins

A HetN-YFP fusion has been previously shown to produce fluorescence located to the cytoplasmic membrane (Higa et al., 2012). However, the gene encoding this fusion protein was placed in a replicative plasmid of unpredictable copy number that had the effect of impairing heterocyst differentiation. We undertook the generation of strains bearing a *hetN-sfgfp* construct in the *hetN* locus (strains CSL71 in a wild-type background, CSL72 in a *patS* mutant background and CSL103 in a sp*hetN* mutant background). In addition, a control strain bearing a fusion of *hetN* with the *gfp-mut2* gene was generated (CSL108).

The localization of GFP fluorescence was studied by time-lapse confocal microscopy in filaments of strain CSL71 from an established diazotrophic culture (Figure 4.9.A). Fluorescence was observed first all over the cytoplasm of the proheterocysts and later in the cell poles of mature heterocysts.

The GFP signal was also observed by fluorescence microscopy. Z-scanning followed by deconvolution and 3D-imaging was done with filaments of strains CSL71, CSL72 and CSL103 (Figure 4.9.B,C,D). Strains CSL71 and CSL72 were incubated in bubbled cultures of $BG11_0$ medium for 20 hours. In both cases the GFP fluorescence was located in the heterocyst surface and focused close to the septal regions, but outside the heterocyst neck. Strain CSL103 was cultured diazotrophically in a plate, and hence heterocysts of different stages were present in the preparations. All the heterocysts observed, regardless of their maturity, presented the fluorescence in the cytoplasm. Consequently, the signal peptide is required for a proper subcellular location of the HetN protein in the heterocysts.

Strain CSL71 presented an increased number of heterocyst doublets, which seemed to be in different stages of differentiation, as the GFP was frequently found in the periphery of the cell in one of the heterocysts and all over the surface in the other (not shown).

Strain CSL108 was also cultured in a plate of $BG11_0$ medium. The only difference between this strain and CSL71 was the type of GFP fused, being the *gfp-mut2* unable to fold properly in the periplasm. The localization of the GFP fluorescence was similar to that of strain CSL71, indicating that the C-end of the HetN protein is likely located inside the cell (Figure 4.9.E).

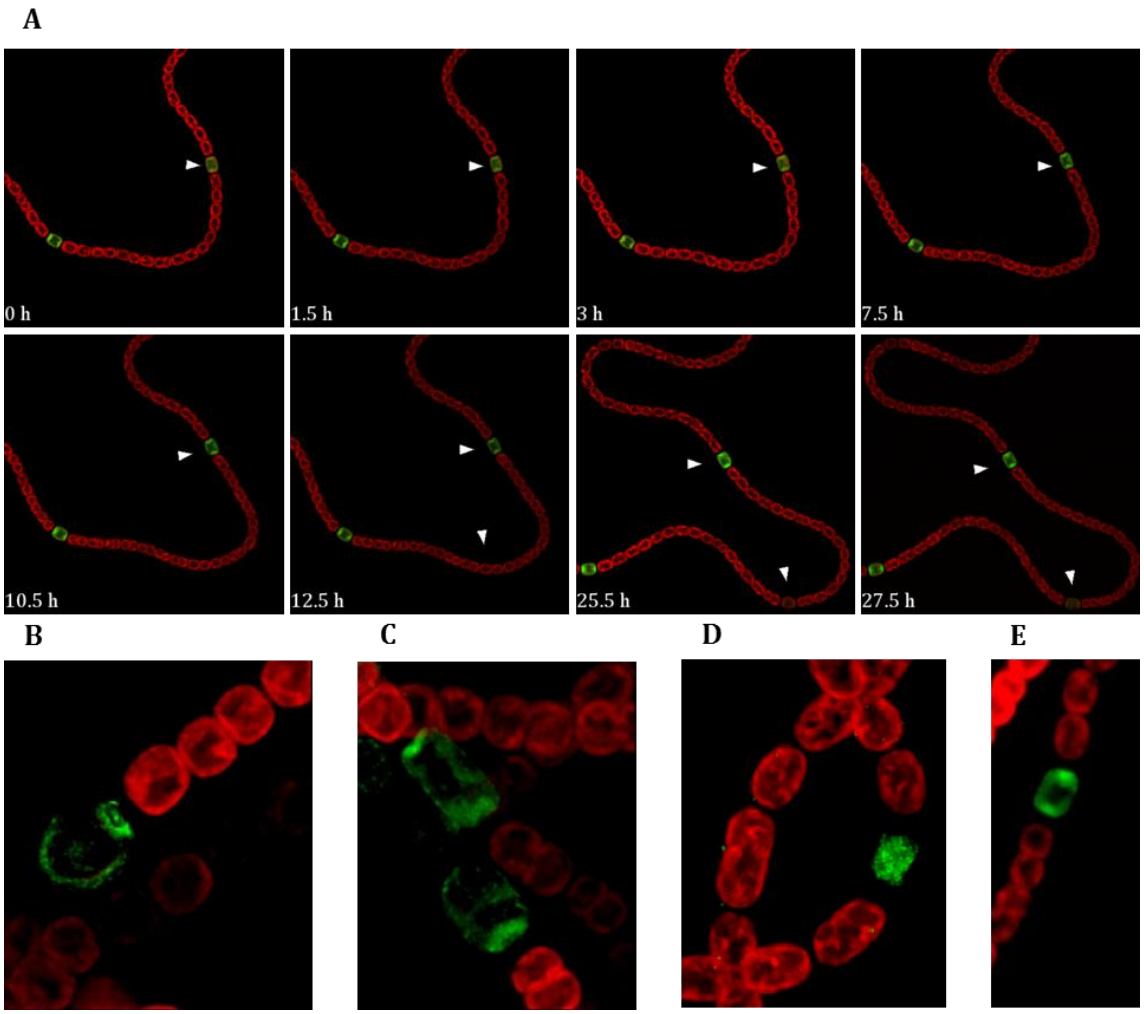


Figure 4.9. HetN-GFP expression. A. Confocal time-lapse microscopy carried out with a filament of strain CSL71 growing in a plate with BG110 medium and 0.5% agar. B. 3D deconvolved image of strain CSL71 taken by fluorescence microscopy. C. 3D deconvolved image of strain CSL72. D. Fluorescence deconvolved image of strain CSL103 growing on solid BG110 medium. E. Fluorescence image of strain CSL108 growing on solid BG110 medium.

4.3. STUDY OF *hetC*

HetC is an essential element for the differentiation of heterocysts. It has been suggested that HetC is involved in the export of a substance that regulates cell division during heterocyst differentiation (Xu & Wolk, 2001). HetC is homologous to ABC transporters with export activity and bears a putative peptidase domain that belongs to a family characterized as bacteriocin-processing peptidases (Khudyakov & Wolk, 1997; Michiels *et al.*, 2001; Rawlings *et al.*, 2012).

To study its putative substrate and its role in heterocyst differentiation, mutants lacking the whole *hetC* gene or the gene fragment encoding the peptidase domain were generated, and an epistasis analysis of these mutations and mutations in several other genes involved in heterocyst differentiation was performed. In addition, versions of the HetC protein fused to GFP were used to analyse the localization and expression of the protein.

4.3.1. Bioinformatic analysis of HetC

Topology prediction softwares identified six to eight transmembrane segments in the HetC protein, positioning the N-terminal and the C-terminal regions in the cytoplasm (Figure 4.10.A). The HetC protein would include a putative peptidase domain of the C39 family comprising residues 339-465 (InterPro, IPR005074) (Figure 4.10.B).

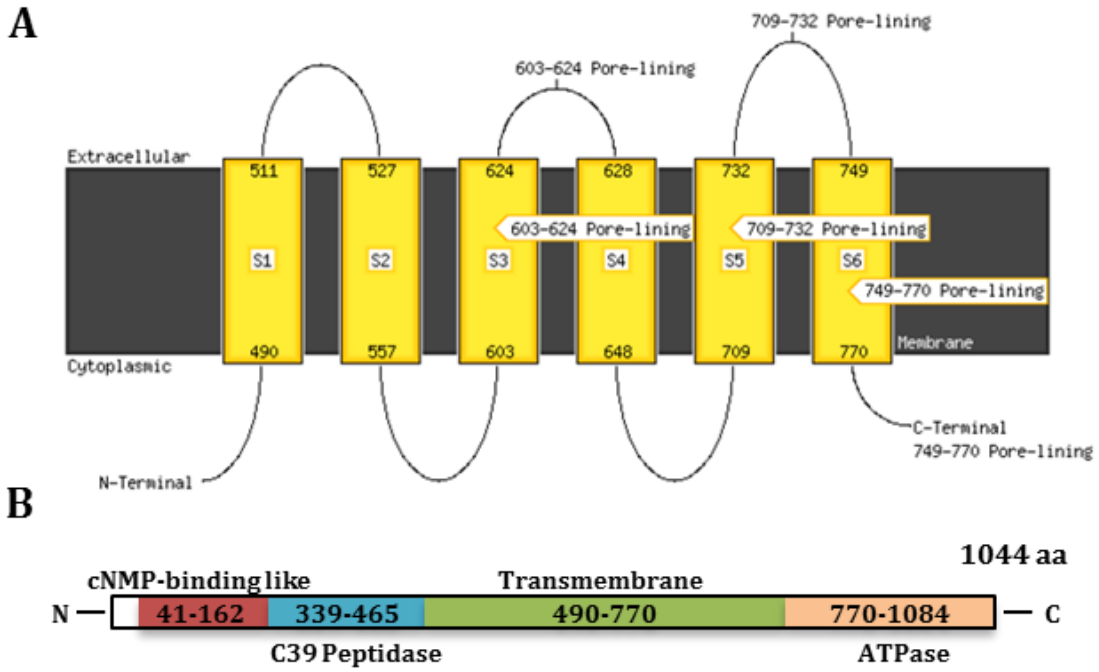


Figure 4.10. Topology prediction for the HetC protein. A. Transmembrane topology prediction of HetC with MEMSAT-SVM. B. Scheme of the predicted domains in HetC.

4.3.2. Phenotype of *hetC* mutant strains

4.3.2.1. Heterocyst pattern

Two types of deletion mutants of *hetC* were constructed, one bearing a deletion of the entire gen except for 98 bp (*hetC*) and another one lacking only the peptidase domain (*hetC-p*). To test epistatic relationships between HetC and PatS or HetN, strains bearing one, two or three of these genes mutated were generated. The spatial pattern of heterocyst distribution was studied in the generated strains (Table 4.4; Figure 4.11).

The *patS* mutant (strain CSVT20) presented a percentage of heterocysts of around 17-18% that did not vary over time after N step-down. This is in contrast to the behaviour of other *patS* deletion mutants previously described, in which the Mch phenotype was alleviated after some time under combined-nitrogen deprivation (Yoon & Golden, 2001).

Consistent with previously described mutants (Khudyakov & Wolk, 1997), strain CSL3 lacking the whole *hetC* gene showed less than 1% heterocysts at any time after N step-down. Besides, strain CSL16 (*hetC-p*) that bears a deletion of only 369 bp (comprising the putative peptidase domain) presented inhibition of differentiation as well, with percentages of heterocysts below 3% at any time studied. This means that the peptidase domain is essential for the function of HetC in heterocyst differentiation.

Double mutants CSL1 (*patS hetC*) and CSL17 (*patS hetC-p*) exhibited a high percentage of heterocysts. They not only presented many doublets, although less than those in the *patS* mutant, but also very short intervals. Indeed, the heterocyst frequency was similar to that of the *patS* mutant at 24 h, although higher than that of strain CSVT20 at 48 and 72 h after nitrogen deprivation. So, it seems that the lack of *patS* could compensate to a good extent for the lack of *hetC*.

As described in section 4.2.2., strain CSL7 (*hetN*) showed a Mch phenotype that was exacerbated over time, and strain CSL11 (*patS hetN*) presented a Mch phenotype much more severe than the corresponding single mutants. However, CSL12 (*hetC hetN*) and CSL30 (*hetC-p hetN*) showed a frequency of heterocysts below 3%. Thus, the *hetC* mutation is epistatic over that of *hetN*.

The triple mutant strains *patS hetC hetN* (CSL15) and *patS hetC-p hetN* (CSL31) presented heterocyst frequencies higher, and mean interval sizes shorter, than the double *hetC patS* mutant. Thus, the inactivation of *hetN* produces an increased differentiation capacity in the *hetC patS* background. However, the triple mutants presented less contiguous heterocysts than the *patS hetN* double mutant. Therefore the mutation of *hetC* faintly prevents the formation of doublets in the *patS hetN* background.

Table 4.4. Pattern of heterocysts in *Anabaena hetC* mutants.

Genotype	Strain	Time (h)	Percentage heterocysts	Contiguous heterocysts	Mean interval
	PCC 7120	24	10.4 ± 0.4	3.9 ± 0.6	10.5 ± 0.7
		48	8.6 ± 0.4	6.4 ± 1.0	13.1 ± 0.8
		72	7.5 ± 0.7	2.5 ± 1.9	13.3 ± 0.8
	CSL3	24	0.03 ± 0.02	-	-
hetC		48	0.00 ± 0.00	-	-
		72	0.00 ± 0.00	-	-
	CSL16	24	1.1 ± 0.7	-	-
hetC-p		48	2.8 ± 1.6	-	-
		72	1.4 ± 1.0	-	-
	CSVT20	24	18.6 ± 2.2	27.8 ± 8.0	5.5 ± 0.7
patS		48	17.2 ± 1.0	23.4 ± 6.8	6.8 ± 0.4
		72	17.6 ± 1.0	33.7 ± 2.5	6.0 ± 0.6
	CSL7	24	14.3 ± 1.0	15.7 ± 5.5	6.8 ± 0.1
hetN		48	13.9 ± 0.6	25.0 ± 3.1	6.2 ± 0.2
		72	17.7 ± 1.0	33.7 ± 8.4	5.6 ± 1.2
	CSL1	24	16.4 ± 1.5	15.2 ± 6.4	3.8 ± 0.4
patS hetC		48	29.4 ± 1.8	25.0 ± 3.2	2.3 ± 0.2
		72	33.0 ± 2.7	23.6 ± 5.7	2.7 ± 0.2
	CSL17	24	18.5 ± 3.4	5.4 ± 3.4	5.2 ± 0.7
patS hetC-p		48	25.2 ± 2.6	26.6 ± 6.0	3.0 ± 0.5
		72	24.7 ± 3.1	26.5 ± 4.0	3.1 ± 0.2
	CSL11	24	42.7 ± 4.8	58.2 ± 3.0	1.5 ± 0.2
patS hetN		48	62.5 ± 10.0	70.6 ± 5.4	1.1 ± 0.3
		72	65.0 ± 6.7	67.2 ± 5.1	1.1 ± 0.0
	CSL12	24	2.5 ± 0.8	-	-
hetC hetN		48	2.1 ± 0.8	-	-
		72	1.8 ± 0.78	-	-
	CSL30	24	2.1 ± 1.3	-	-
hetC-p hetN		48	1.2 ± 0.6	-	-
		72	0.8 ± 0.6	-	-
	CSL15	24	28.8 ± 2.0	10.5 ± 4.8	4.3 ± 1.0
patS hetC hetN		48	36.5 ± 0.4	8.9 ± 1.9	3.3 ± 0.1
		72	57.3 ± 12.7	46.4 ± 22.2	1.7 ± 0.6
	CSL31	24	20.2 ± 1.1	8.9 ± 3.8	4.5 ± 0.3
patS hetC-p hetN		48	34.8 ± 7.9	25.5 ± 3.2	2.8 ± 0.1
		72	43.5 ± 6.20	28.6 ± 1.0	2.0 ± 0.0

Heterocyst frequency (as percentage of total cells), mean size of vegetative cell intervals between heterocysts and percentage of contiguous heterocysts (interval size = 0, percentage of total intervals) are indicated for each strain. Data are the mean and standard deviation of the mean from two to six independent experiments. Data of PCC 7120, CSVT20 and CSL7 are shown as controls.

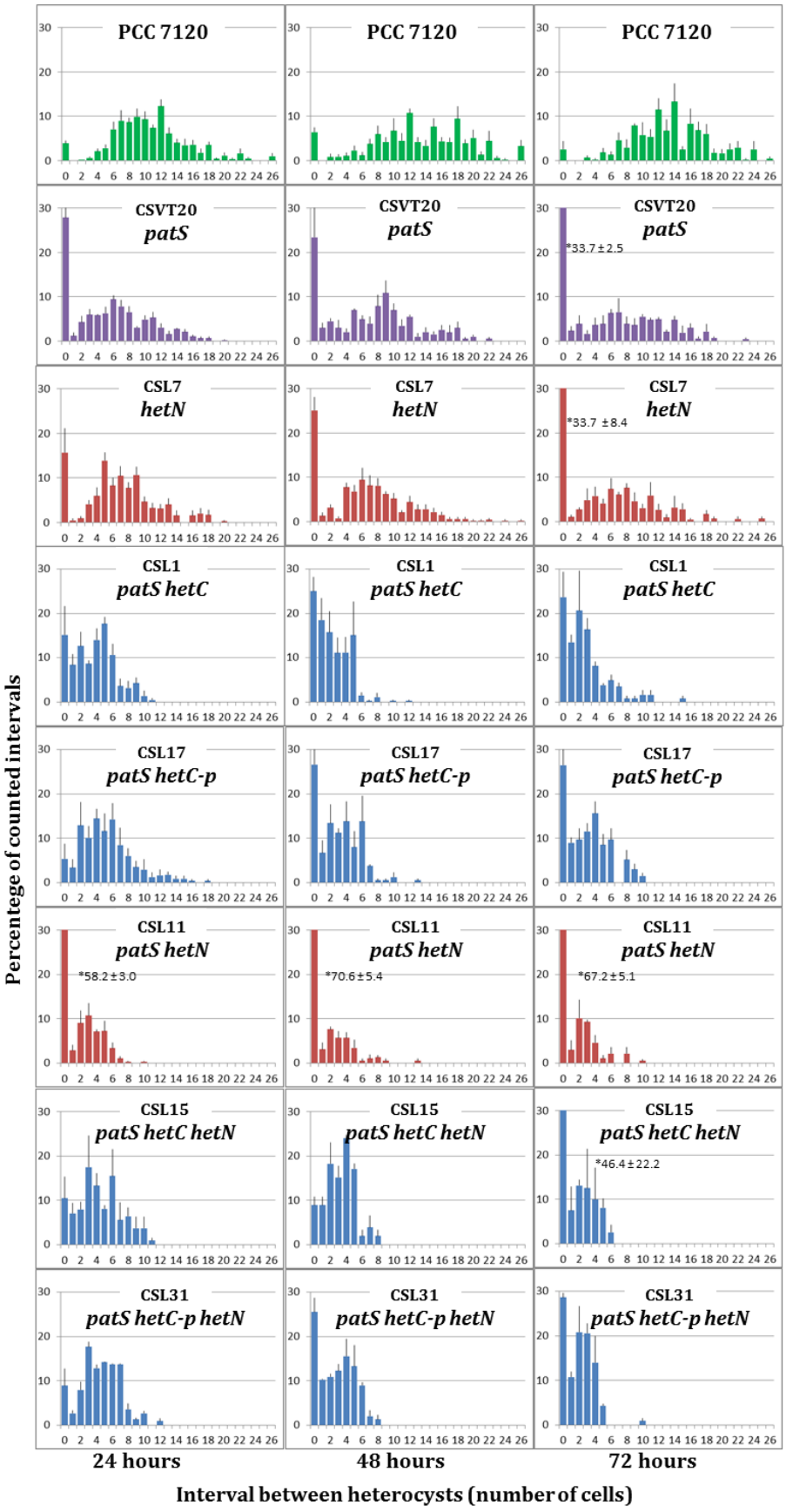


Figure 4.11. Heterocyst distribution in *Anabaena* sp. PCC 7120 and *hetC* mutant strains. The frequency of intervals at 24, 48 and 72 h of nitrogen deprivation is shown in the graphs. In the cases in which the frequency is over 30%, the exact value is indicated.

4.3.2.2. Other phenotypic characteristics

The aspect of the heterocysts was compared in strains bearing the *hetC* deletions (single mutant CSL16 and double mutants CSL1 and CSL17) and *Anabaena* sp. PCC 7120 (Figure 4.12.A.). In the wild-type strain, heterocysts showed a thick Alcian Blue-stained envelope of polysaccharides, cyanophycin granules, and a larger size and changed granularity in the cytoplasm as compared to the vegetative cells. However, the *hetC* mutant strains, despite the *patS* mutation, exhibited heterocysts that were not always larger than the vegetative cells, had not altered granularity and lacked cyanophycin granules, although they presented the Alcian Blue-stained polysaccharide envelope.

The expression of the *nifHDK* operon encoding nitrogenase was analysed in the *hetC*, *patS* and *patS hetC* mutants (Figure 4.12.B) in filaments grown with ammonium and incubated or not for 24 h in the absence of combined nitrogen. Strain CSVT20 showed levels of expression comparable to those of the wild-type strain, whereas strain CSL3 presented no expression and strain CSL1 exhibited detectable expression levels although much lower than in PCC 7120.

The presence of heterocyst-specific glycolipids was studied in some of the strains described above (Figure 4.12.C). The wild-type strain, CSL7 (*hetN*), CSL11 (*patS hetN*), CSVT20 (*patS*) and CSL1 (*patS hetC*) bore envelope glycolipids at 24 and 48 h after nitrogen deprivation. Strains CSL12 (*hetC hetN*) and CSL16 (*hetC-p*) did not show a noticeable presence of heterocyst-specific glycolipids.

The ability to grow diazotrophically and the nitrogenase activity were analysed in the *hetC* mutants. The ability to grow fixing N_2 was studied by spot growth tests in solid medium, and the nitrogenase activity was measured, under oxic and anoxic conditions, 24 hours after N step-down.

None of the mutants was able to grow well in the absence of combined nitrogen (Figure 4.13). Strains with heterocyst frequencies below 3%, *hetC*, *hetC-p*, *hetC hetN* and *hetC-p hetN* mutants, could not grow due to the lack of nitrogen fixation. However, it is feasible that the rest of the mutant strains, all of them showing high heterocyst frequencies, could not grow owing to an unbalanced carbon-to-nitrogen ratio or to an aberrant heterocyst differentiation.

Regarding nitrogenase activity, mutant strains *hetC*, *hetC-p*, *hetC hetN* and *hetC-p hetN* showed negligible nitrogenase activity levels (Table 4.5). Double mutant strains *patS hetC* and *patS hetC-p* showed appreciable levels of nitrogenase, or at least higher than those found in the single *hetC* mutants, under anoxic conditions. Nevertheless, as shown by the lack of diazotrophic growth, the deletion of *patS* does not allow the differentiation of functional heterocysts in the *hetC* mutant background.

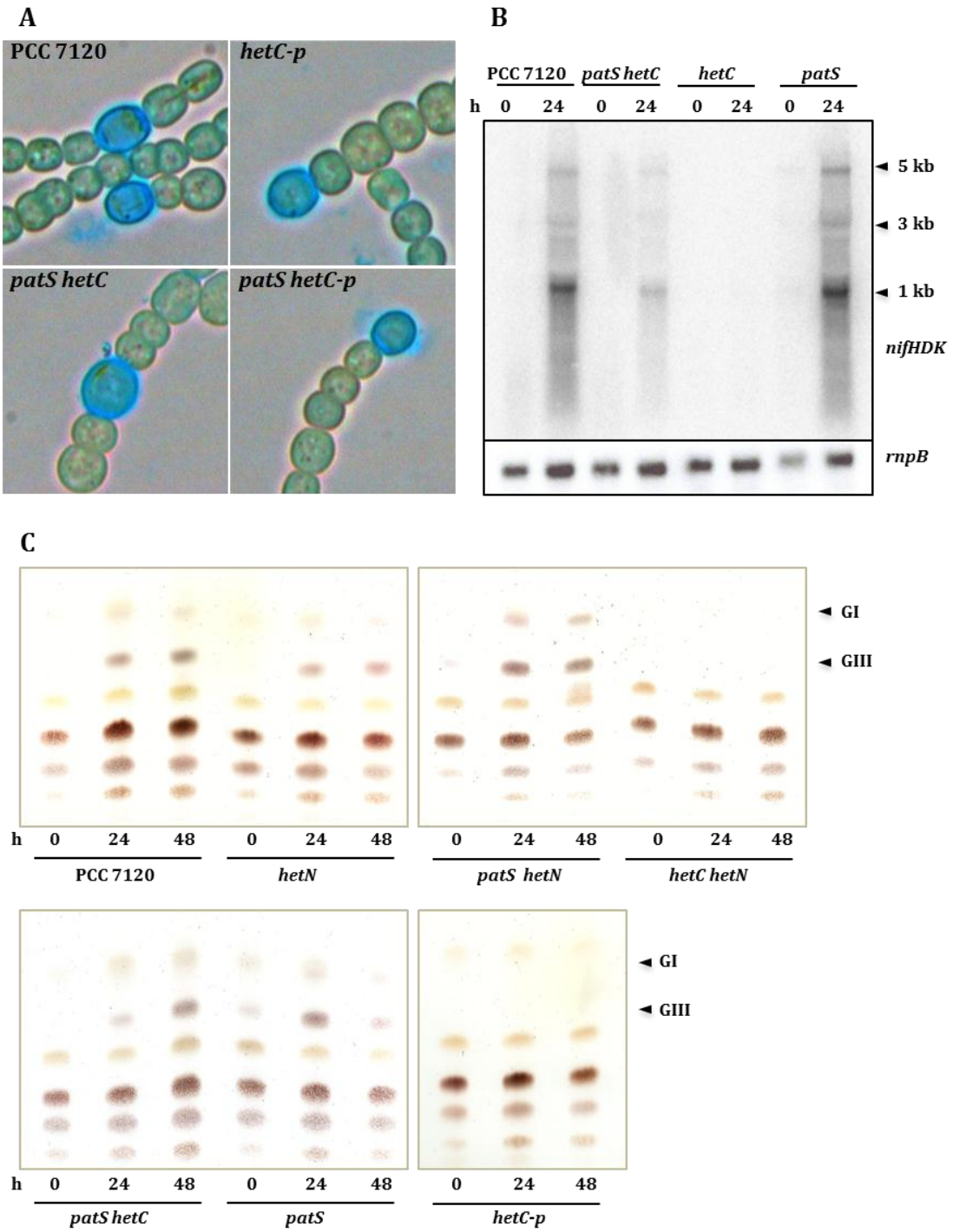


Figure 4.12. Study of different phenotypic characteristics of some *hetC* mutant strains. A. Filaments of strains PCC 7120, CSL16 (*hetC*-p), CSL1 (*patS hetC*) and CSL17 (*patS hetC*-p) stained with Alcian Blue. B. Northern blot analysis using a nifHDK probe in strains PCC 7120, CSL1 (*patS hetC*), CSL3 (*hetC*) and CSVT20 (*patS*). C. TLC analysis of glycolipids performed with the indicated strains. The two glycolipids specific of the heterocyst envelope are indicated.

Significant nitrogenase activity levels were found in strains *patS hetN* (strain CSL11) and the triple mutants, CSL15 and CSL31, although in none of them was nitrogenase comparable to that found in the wild-type strain. The fact that strain CSL15 (*patS hetC hetN*) presented higher nitrogenase activity values than CSL1 (*patS hetC*), even

showing a higher heterocyst frequency, means that the impairment of nitrogen fixation and diazotrophic growth in these mutants is due to a problem in heterocyst function and not only to the C/N balance.

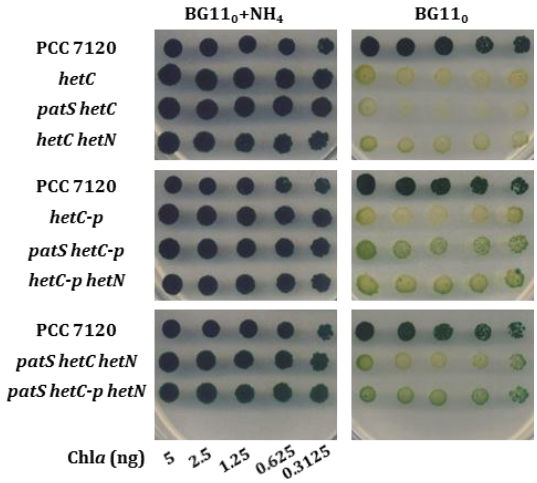


Figure 4.13. Solid growth test of *hetC* mutants. Filaments from flask cultures in ammonium-containing medium were washed with BG110 medium and dropped atop plates with or without ammonium in the medium. Drops spotted (5 µL) contained the quantities indicated of chlorophyll a (the quantities in BG110 medium were the same as in BG110+NH4+ medium). The plates were incubated for 14 days under standard growth conditions and photographed.

Table 4.5. Nitrogenase activity of *Anabaena* wild type and *hetC* mutants.

Genotype	Strain	Time (h)	Nitrogenase activity (nmol [µg chlorophyll *a*]$^{-1}$ h^{-1}) Oxic	Anoxic
	PCC 7120	24	16.0 ± 3.6 (4)	17.4 ± 1.1 (11)
		48	8.3 ± 2.4 (6)	11.7 ± 0.7 (7)
hetC	CSL3	24	0.0 ± 0.0 (3)	0.1 ± 0.0 (7)
		48	0.1 ± 0.1 (4)	0.1 ± 0.1 (3)
hetC-p	CSL16	24	0.0 ± 0.0 (2)	0.0 (1)
		48	0.1 ± 0.1 (3)	0.0 (1)
patS	CSVT20	24	3.2 (1)	12.3 ± 1.2 (5)
		48	1.1 ± 0.2 (2)	7.8 ± 0.7 (2)
hetN	CSL7	24	3.6 (1)	10.5 ± 3.7 (3)
		48	2.1 ± 0.7 (2)	6.1 ± 0.9 (3)
patS hetC	CSL1	24	0.0 ± 0.0 (2)	0.8 ± 0.5 (6)
		48	0.0 ± 0.0 (2)	0.4 ± 0.3 (4)
patS hetC-p	CSL17	24	0.1 ± 0.0 (3)	0.2 ± 0.1 (3)
		48	0.7 ± 0.4 (4)	1.3 ± 0.8 (4)
patS hetN	CSL11	24	0.0 (1)	1.7 ± 1.4 (3)
		48	0.0 ± 0.04 (2)	0.5 ± 0.4 (4)
hetC hetN	CSL12	24	0.0 (1)	0.0 (1)
		48	0.0 ± 0.07 (2)	0.1 ± 0.0 (2)
hetC-p hetN	CSL30	24	0.0 (1)	0.0 ± 0.0 (2)
		48	0.0 ± 0.0 (2)	0.1 ± 0.1 (2)
patS hetC hetN	CSL15	24	0.1 ± 0.1 (2)	0.2 ± 0.2 (3)
		48	0.1 ± 0.1 (2)	2.2 ± 1.8 (3)
patS hetC-p hetN	CSL31	24	0.1 ± 0.0 (2)	0.3 ± 0.1 (3)
		48	0.8 ± 0.4 (3)	2.4 ± 0.7 (4)

Data are the mean and standard deviation of the mean from the results of the number of independent experiments indicated in brackets, measured 24 or 48 h after N step-down.

Therefore, deletion of *patS* can partially bypass the lack of a functional HetC protein allowing the generation of cells that begin heterocyst differentiation. However, these cells still miss an essential element for heterocyst differentiation, being arrested at some point in the process.

4.3.4. PatS versions in a *hetC* mutant background

In a wild-type background, expression of PatS-17 or PatS-8 (a putative product of PatS-17 processing in the proheterocysts) as the only PatS version recreates a heterocyst pattern similar to that of the wild-type strain, whereas accumulation of PatS-17 provokes inhibition of differentiation (see Section 4.1.3). Thus, in principle, it might be expected that if the peptidase domain of HetC were involved in PatS-17 processing, expression of PatS-17, but not of PatS-8, in a *hetC-p* background would be inhibitory.

To test that possibility, we cloned these two different *patS* versions in the *patS hetC-p* background. The resulting strains, CSL101 (*patS hetC-p patS-8*) and CSL102 (*patS hetC-p patS-17*), were analysed in terms of the pattern of heterocysts (Figure 4.14). The two strains showed low heterocyst frequencies, similar to that of strain CSL16 (*hetC-p*), which means that both *patS* genes complement the *patS* mutation in the *hetC-p* background. Thus, PatS-17 seems not to be processed by HetC.

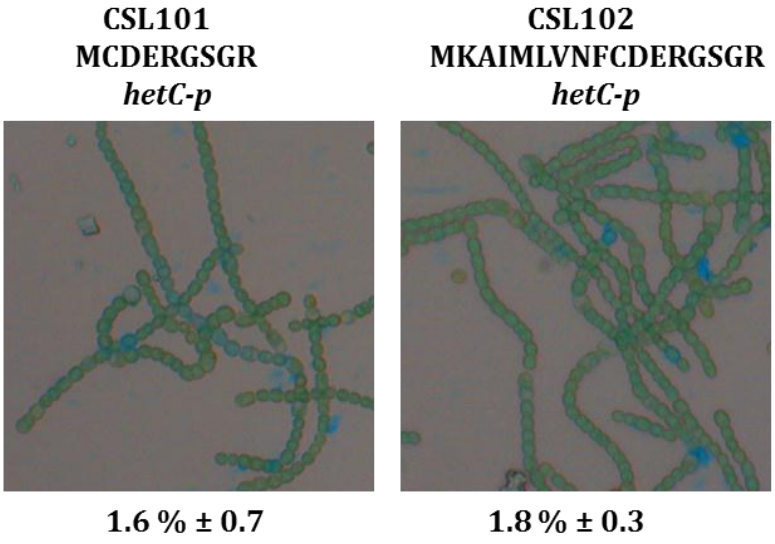


Figure 4.14. Phenotype of strains producing PatS-17 or PatS-8 in a *hetC*-p background. Filaments of bubbled cultures incubated for 24 hours in BG110C medium were stained with Alcian Blue and visualised by light microscopy (magnification 40x). Percentage of stained cells is indicated under the images.

4.3.5. Localization of the HetC protein

The localization of HetC was studied using a GFP fusion. GFP fluorescence was studied by confocal microscopy in strains expressing a HetC-GFP (CSM1) and HetC-p-GFP (CSL33) fusion. Upon N step-down, green fluorescence was observed at about 6 h (not shown), initially over the proheterocyst surface and later (after 14 h) more concentrated in the periphery of the cells, especially in the poles (Figure 4.15). No differences were observed between strain CSM1 (complete HetC protein) and CSL33 (deletion of the peptidase domain), indicating that the peptidase domain is not necessary for the localization of the protein.

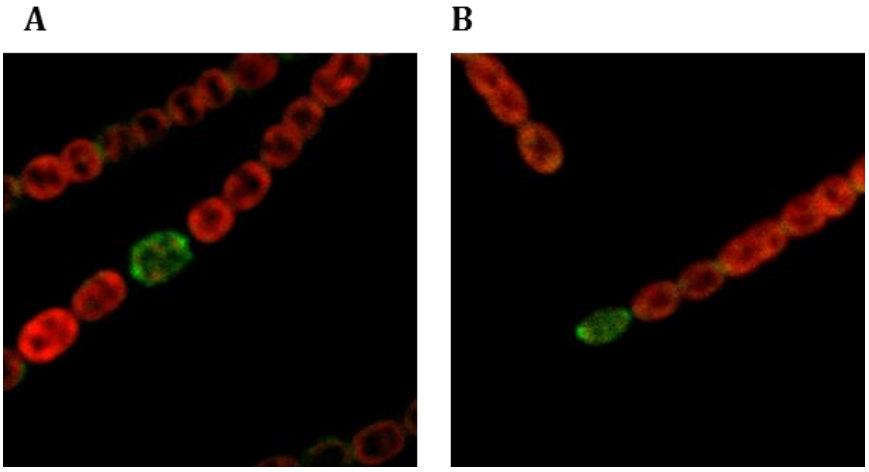


Figure 4.15. Localization of HetC. Filaments grown with ammonium and incubated in bubbled cultures with BG110C medium for 24 hours were visualised with a confocal microscope. A. Strain CSM1 producing a HetC-GFP fusion. B. Strain CSL33 producing HetC-p-GFP.

Growth of these strains was studied by a spot growth test (Figure 4.16). Strain CSM1 (*hetC-gfp-mut2*) showed a slow growth, indicating that the HetC-GFP protein is not totally functional. The CSL33 strain (*hetC-p-gfp-mut2*) did not grow under diazotrophic conditions.

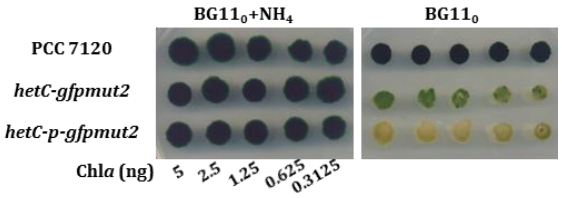


Figure 4.16. Solid growth test of strains expressing *hetC-gfp*-mut2 fusions. Filaments from flask cultures in ammonium-containing medium were washed with BG110 and dropped atop plates with or without ammonium in the medium. Drops spotted (5 µL) contained the quantities indicated of chlorophyll a (the quantities in BG110 medium were the same as in BG110+NH4+ medium). The plates were incubated for 14 days under standard growth conditions and photographed.

4.4. STUDY OF HetP

Downstream from the *hetC* gene, *hetP* is located. The deletion of the *hetP* gene leads to inhibition of the differentiation of heterocysts and its overexpression produces over-differentiation, even in a *hetR* mutant background (Fernández-Piñas *et al.*, 1994; Higa & Callahan, 2010). In order to expand the knowledge about the function of HetP, we intended to study the localization of the HetP protein.

4.4.1. Bioinformatic analysis of HetP

hetP-like genes can be found near *hetC* in most of the sequenced genomes of heterocyst-forming cyanobacteria. Moreover, there are other two *hetP*-like genes in the *Anabaena* genome.

The topology of the predicted HetP protein was analysed, and one transmembrane segment was detected (MEMSAT-SVM) (Figure 4.17). The HetP protein presents similarity only to proteins of unknown function.

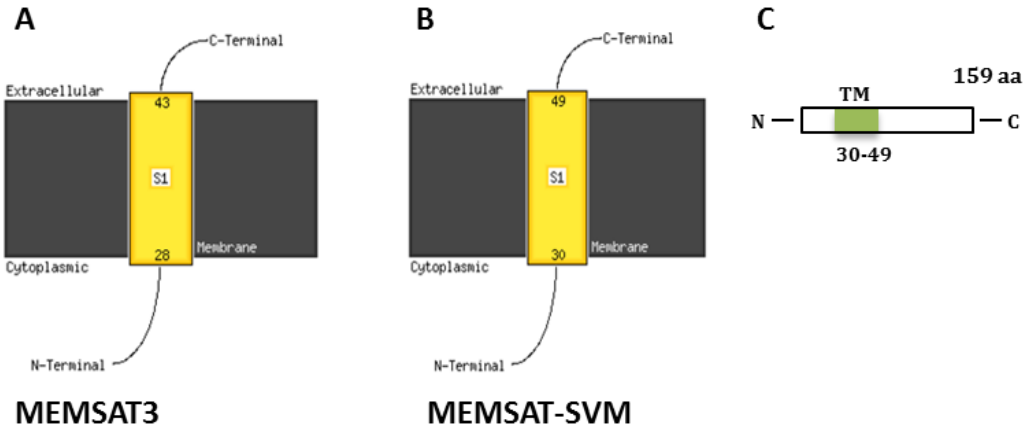


Figure 4.17. Topology prediction for HetP. One transmembrane segment is predicted for HetP with MEMSAT3 (A) and MEMSAT-SVM (B). C. A scheme showing the approximate location of the transmembrane segment in the HetP protein is shown.

4.4.2. Localization of HetP-GFP and characterization of derivative strains

Strains bearing a fusion of the protein to sfGFP (CSL67 strain in the wild-type background, CSL68 in *hetC-p*, CSL69 in *patS* and CSL70 in *hetN* backgrounds) or to the GFP (CSL107 strain in the wild-type background) were generated in order to study the localization of HetP.

The HetP-sfGFP fusion protein was observed in the cytoplasm in proheterocysts and located in the cell poles in mature heterocysts. In Figure 4.18.A, filaments of strain CSL67 from established diazotrophic cultures are shown, in which heterocysts are in different stages of differentiation, and green fluorescence is observed in these different

locations. The GFP signal was also observed in strain CSL67 by fluorescence microscopy, where Z-scanning followed by deconvolution and 3D-imaging showed a very well delimited signal in the heterocyst poles (Figure 4.18.B). The localization of HetP fusions was similar in strains CSL68, CSL69 and CSL70. Strain CSL107 bearing a HetP-GFP (*gfp-mut2*) fusion did show fluorescence in the heterocyst (Figure 4.18.C), therefore at least the C-end of HetP must be located in the cytoplasm where conventional GFP is able to fold.

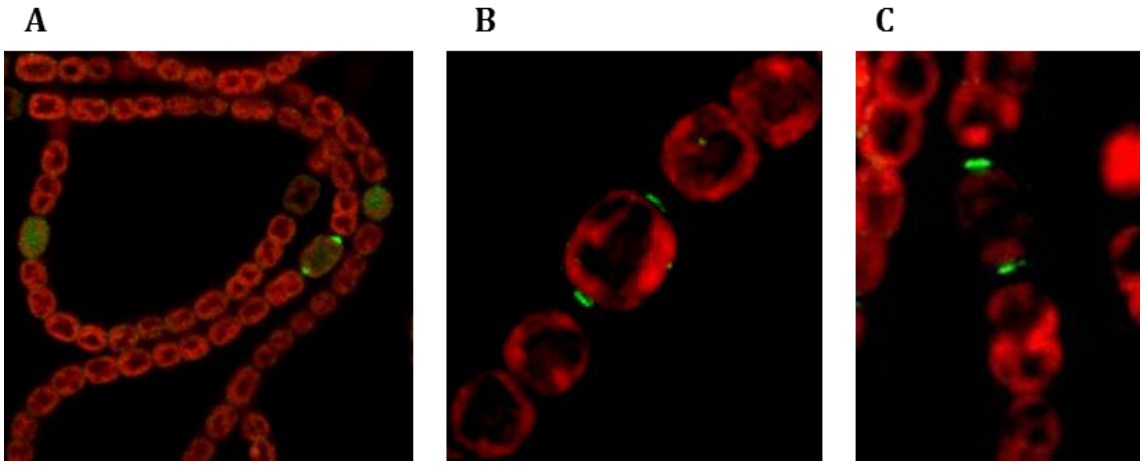


Figure 4.18. Localization of HetP-sfGFP and HetP-GFP proteins. A. GFP fluorescence in strain CSL67 that expresses a HetP-sfGFP analysed by confocal microscopy. Filaments grown in solid medium were used. B. Fluorescence microscopy deconvolved image of HetP-GFP in strain CSL67 grown for 40 hours in a bubbled culture with BG110C medium. C. Fluorescence microscopy in strain CSL107 grown in solid BG110 medium.

The localization of the HetP-GFP fusion protein was studied also by western blot analysis. Membrane and soluble fractions of wild-type and CSL67 strains at 18 hours after nitrogen deprivation were separated by ultracentrifugation, loaded into a SDS-PAGE gel and transferred into a PVDF membrane. Western blot was performed using a commercial antibody against the GFP in a 1:2000 dilution. A band with mobility corresponding to ca. 40 kDa, which could correspond to the HetP-sfGFP protein (45.3 kDa), was detected in the soluble extract fraction (Figure 4.19). In the membrane fraction, two specific bands with apparent sizes larger than ca. 75 kDa were detected. Therefore, the HetP protein is, at least in part, in the filament associated to membranes.

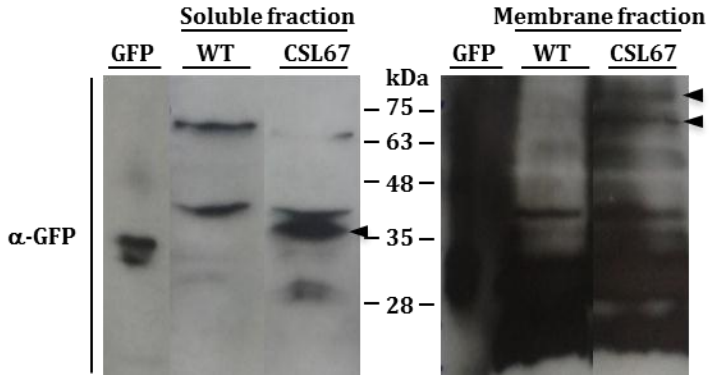


Figure 4.19. Western analysis of HetP-sfGFP strain. Soluble and membrane fractions of the wild-type and CSL67 strains were analysed using an anti-GFP antibody (Invitrogen). A whole-cell extract of a strain bearing *PrbcL-gfp-mut2* fusion was used as a control of free GFP.

Because the integration in the chromosome of *hetP-sfgfp* implies the duplication in the genome of the *hetP* gene expressed from 98 bp of its own promoter region (see Figure 3.13), the expression levels of the *hetP* gene in the different mutant strains in comparison to the wild type were studied by RT-qPCR, (Table 4.6). Two independent experiments were carried out in which total RNA was isolated 18 hours after N step-down. Quantitative PCR was performed with oligonucleotides that amplified approximately 150 bp of the *hetP* gene (alr2818-16/alr2818-17), the *rnpB* (rnpB-4/rnpB-5, as a control). The Ct values obtained were analysed with the REST 2009 software (Pfaffl *et al.*, 2002).

Table 4.6. Ratios of the expression levels in *hetP* of the indicated strains.

Sample pairs	Relative *hetP* expression	S. E. range	*P*
CSL16 vs. WT	1.354	0.451 – 3.832	0.485
CSL67 vs. WT	3.772	1.498 – 13.753	0.019
CSL68 vs. WT	3.319	0.842 – 12.508	0.032
CSL68 vs. CSL16	2.451	1.193 – 5.437	0.022
CSL68 vs. CSL67	0.880	0.357 – 1.896	0.761

Data from RT-qPCR were normalized with the expression of *rnpB* and analysed with the REST 2009 software. Two independent RNA preparations were used with the appropriate number of technical replicas.

The expression levels of *hetP* strains in CSL67 and CSL68 were on average about 3-fold higher than those found in PCC 7120 or CSL16. Thus, the construct integrated in the genome produces an increase in *hetP* expression, which leads to a functional bypass of the *hetC* mutation and an increase in the frequency of contiguous heterocysts.

Growth of these strains was studied in a spot growth test (Figure 4.20). All the strains expressing HetP-sfGFP grew diazotrophically, even CSL68 whose parental strain, CSL16, did not. Therefore, the overexpression of *hetP* allows diazotrophic growth in a *hetC-p* mutant background.

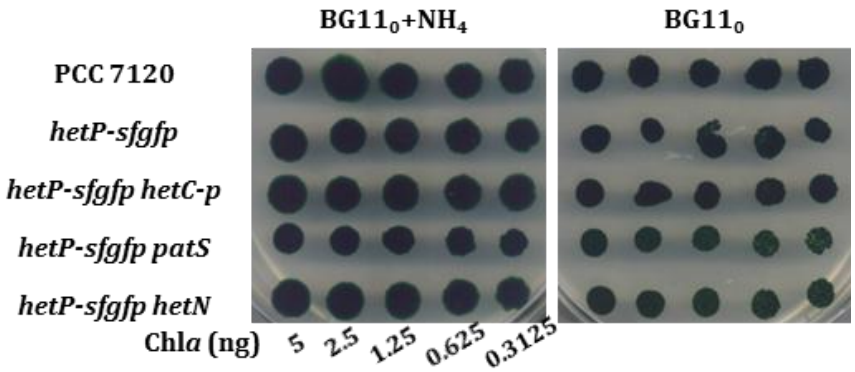


Figure 4.20. Solid growth test of *hetP-sfgfp* strains. Filaments from flask cultures in ammonium-containing medium were washed with BG110 and dropped atop plates with or without ammonium in the medium. Drops spotted (5 µL) contained the quantities indicated of Chl (the quantities in BG110 medium were the same as in BG110+NH4+ medium). The plates were incubated for 14 days under standard growth conditions and photographed.

The pattern of heterocyst distribution in the strains expressing the HetP-sfGFP fusion was studied (Table 4.7; Figure 4.21). Strains CSL67, CSL69 and CSL70 showed phenotypes similar to those of their respective parental strains, but with an increased frequency of contiguous heterocysts. However, consistent with the results of growth tests, strain CSL68 showed a pattern similar to that of strain CSL67 even though it expresses a version of HetC that lacks the peptidase domain.

In addition, the nitrogenase activity was measured in these strains to test the functionality of the heterocysts (Figure 4.8). Strains CSL67 and CSL68 showed high nitrogenase activities, comparable to that of strain PCC 7120, even under oxic conditions. However, CSL69 and CSL70 presented values similar to those of their parental strains. Thus, overexpression of *hetP* makes a functional HetC protein dispensable, but does not counteract the negative effects of excessive differentiation in nitrogenase activity.

Table 4.7. Pattern of heterocysts in the *hetP-sfgfp* mutants.

Genotype	Strain	Time (h)	Percentage heterocysts	Contiguous heterocysts	Mean interval
hetP- sfgfp	CSL67	24	11.5 ± 0.8	9.1 ± 2.5	8.6 ± 0.6
		48	11.4 ± 0.6	20.5 ± 0.4	9.8 ± 0.4
		72	11.0 ± 0.6	14.7 ± 1.2	11.1 ± 0.9
hetP-sfgfp hetC-p	CSL68	24	12.1 ± 1.8	13.4 ± 2.4	7.8 ± 0.4
		48	11.5 ± 1.2	13.7 ± 5.6	9.7 ± 0.3
		72	14.7 ± 1.8	19.4 ± 3.9	10.4 ± 0.6
hetP-sfgfp patS	CSL69	24	25.5 ± 6.3	26.3 ± 6.9	5.7 ± 1.2
		48	27.0 ± 9.6	19.8 ± 0.8	7.5 ± 0.2
		72	14.9 ± 2.7	21.2 ± 3.1	8.6 ± 1.1
hetP-sfgfp hetN	CSL70	24	17.6 ± 1.3	29.2 ± 1.0	4.8 ± 0.3
		48	22.9 ± 3.1	47.3 ± 0.0	4.4 ± 0.1
		72	16.1 ± 0.2	44.4 ± 4.8	5.2 ± 0.1

Heterocyst frequency (as percentage of total cells), mean size of vegetative cell intervals between heterocysts and percentage of contiguous heterocysts (interval size = 0, percentage of total intervals) are indicated for each strain. Data are the mean and the standard deviation of the mean from the results of two or three independent experiments.

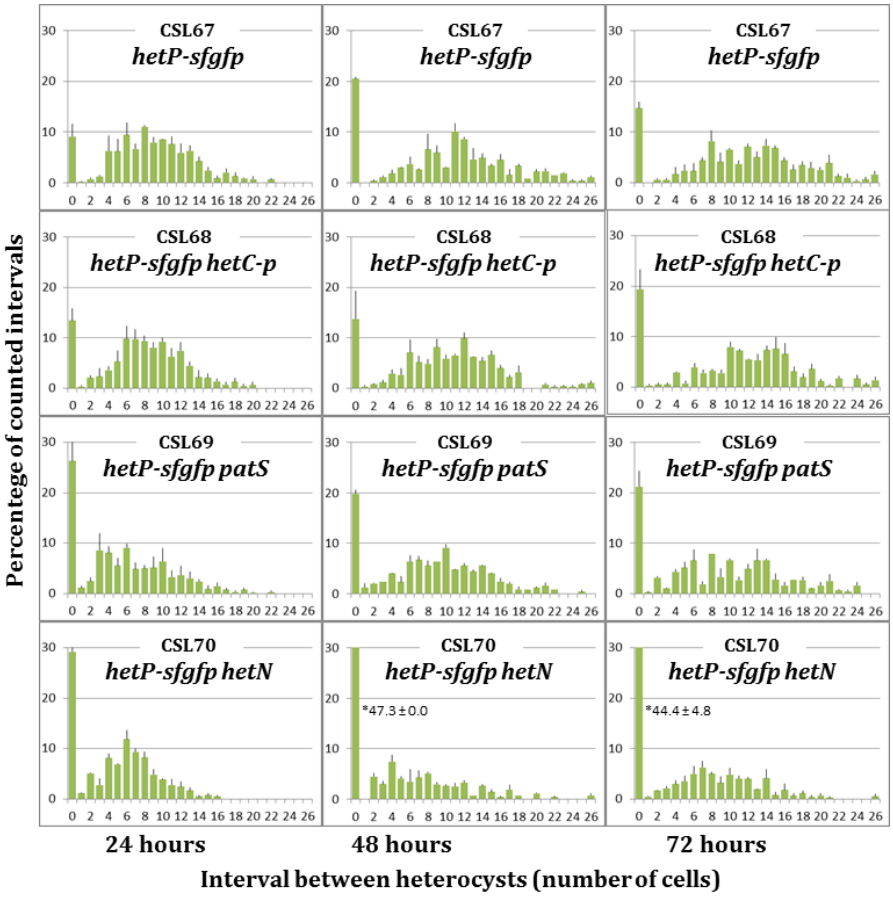


Figure 4.21. Heterocyst distribution in the *hetP-sfgfp* mutants. The frequency of intervals at 24, 48 and 72 h of nitrogen deprivation is shown in the graphs. In the cases in which the frequency is over 30%, the exact value is indicated.

Table 4.8. Nitrogenase activity of the *hetP-sfgfp* mutants.

Background	Strain	Time (h)	Nitrogenase activity (nmol [μg chlorophyll *a*]$^{-1}$ h^{-1}) Oxic	Anoxic
	PCC 7120	24	16.0 ± 3.6 (4)	17.4 ± 1.1 (11)
		48	8.3 ± 2.4 (6)	11.7 ± 0.7 (7)
hetP-sfgfp	CSL67	24	10.2 ± 3.3 (2)	22.2 (1)
		48	8.2 ± 3.0 (2)	14.9 (1)
hetP-sfgfp hetC-p	CSL68	24	12.5 (1)	14.0 (1)
		48	4.3 ± 0.0 (2)	7.9 ± 1.8 (2)
hetP-sfgfp patS	CSL69	24	1.4 (1)	7.6 (1)
		48	1.6 (1)	7.4 (1)
hetP-sfgfp hetN	CSL70	24	6.6 (1)	9.8 (1)
		48	3.8 (1)	9.2 (1)

Data are the mean and standard deviation of the mean from the results of the number of independent experiments indicated in brackets, measured 24 or 48 h after N step-down.

4.5. STUDY OF *asr2819* MUTANT STRAINS

The *asr2819* ORF is located downstream from the *hetP* gene in the *Anabaena* genome. According to the data of Mitschke *et al.* (2011) and Flaherty *et al.* (2011), no TSS has been found between *hetP* and *asr2819*. Furthermore, the RNA reads detected in this region, decreasing in number from the beginning of *hetP* gene to the end of the *asr2819* ORF and showing the same pattern of expression during heterocyst differentiation, suggest that they could be cotranscribed from a TSS located upstream of the *hetP* gene at position 3431807, [coordinates of the strain PCC 7120 genomic sequence (Kaneko *et al.*, 2001)].

4.5.1. Phenotype of *asr2819* mutant strains

To study a possible role of *asr2819* in heterocyst differentiation, a series of mutant strains that lack the *asr2819* ORF except for 15 bp (strain CSL97 in the wild-type background and strain CSL98 in the *hetC-p* background), bear an insertion of the C.S3 cassette (strain CSL99 in the wild-type background and strain CSL100 in the *hetC-p* background) or overexpress a second copy of the *asr2819* gene under the C.K3 promoter in a different region of the genome (strain CSL95 in the wild-type background and strain CSL96 in the *hetC-p* background) were generated. In addition, the *asr2819* mutation was introduced in *hetN* and *patS hetN* backgrounds to analyse possible epistatic effects (strains CSL105 and CSL106, respectively). The pattern of heterocyst distribution was studied in these mutant strains (Table 4.9; Figure 4.22). The three strains generated from the *hetC-p* background presented heterocyst frequencies and distribution similar to the CSL16 parental strain. Thus, the *hetC* mutation appears to be epistatic over the *asr2819* mutation.

Strain CSL95 (overexpressing *asr2819*) presented an increased number of contiguous heterocysts and, consequently, higher heterocyst frequencies. Therefore, Asr2819 appears to have a positive influence on heterocyst differentiation. The fact that this phenotype is the result of the expression of *asr2819* alone in a different site in the genome points out that this ORF itself has a role on its own in the heterocyst differentiation process.

Strain CSL97 (*asr2819*) showed an increased heterocyst frequency and a lower mean interval size than the wild-type strain. At 72 h, the most striking phenotypic characteristic was a relaxed pattern of heterocyst distribution (see Figure 4.22), meaning that, the distribution did not show a bell-shaped curve but instead showed apparently randomized intervals. Thus, Asr2819 seems to contribute to the establishment of the heterocyst pattern.

Strain CSL105 (*asr2819 hetN*) showed a phenotype similar to the *hetN* mutant, however strain CSL106 (*patS asr2819 hetN*) showed more than 70% heterocyst frequency even after 24 hours of nitrogen deprivation, the highest value obtained from any mutant combination studied. Exact data is not shown in the table because of the difficulty of counting heterocysts in this strain.

Table 4.9. Pattern of heterocysts in the *asr2819* mutant strains.

Genotype	Strain	Time (h)	Percentage heterocysts	Contiguous heterocysts	Mean interval
	PCC 7120	24	10.4 ± 0.4	3.9 ± 0.6	10.5 ± 0.7
		48	8.6 ± 0.4	6.4 ± 1.0	13.1 ± 0.8
		72	7.5 ± 0.7	2.5 ± 1.9	13.3 ± 0.8
C.K3-asr2819	CSL95	24	13.0 ± 0.9	8.7 ± 3.5	7.0 ± 0.2
		48	16.0 ± 0.8	16.7 ± 5.3	8.2 ± 0.7
		72	12.6 ± 0.6	10.6 ± 2.4	9.9 ± 0.7
C.K3-asr2819 hetC-p	CSL96	24	1.0 ± 0.2	-	-
		48	0.8 ± 0.4	-	-
		72	1.1 ± 0.2	-	-
asr2819	CSL97	24	13.8 ± 2.6	1.8 ± 0.9	8.2 ± 0.7
		48	12.5 ± 1.1	4.9 ± 0.9	8.1 ± 0.7
		72	15.8 ± 0.4	5.5 ± 1.3	10.5 ± 1.1
asr2819 hetC-p	CSL98	24	1.2 ± 0.1	-	-
		48	3.0 ± 1.1	-	-
		72	1.1 ± 0.4	-	-
asr2819-C.S3	CSL99	24	13.0 ± 0.7	4.9 ± 1.0	8.3 ± 0.7
		48	12.3 ± 0.6	10.4 ± 0.9	9.8 ± 0.3
		72	11.0 ± 0.4	13.0 ± 2.0	10.6 ± 0.4
asr2819-C.S3 hetC-p	CSL100	24	1.9 ± 0.4	-	-
		48	3.7 ± 1.2	-	-
		72	4.4 ± 0.9	-	-

Heterocyst frequency (as percentage of total cells), mean size of vegetative cell intervals between heterocysts and percentage of contiguous heterocysts (interval size = 0, percentage of total intervals) have are indicated for each strain. Data are the mean and standard deviation of the mean of the results from two to four independent experiments. The wild-type strain is shown as a control.

4.5.2. Study of the expression of gene *asr2819*

In order to analyse the expression of the *asr2819* gene, RT-qPCR was performed in the wild-type strain, the *hetC-p* deletion mutant (CSL16), the *asr2819* deletion mutant (CSL97) and the *asr2819* overexpression mutant strain (CSL95) (Table 4.10). Two independent experiments (three for the wild-type strain) were performed in which total RNA was isolated after 0, 3 or 24 hours of nitrogen deprivation. Quantitative PCR was performed in the same conditions as in Section 4.4.2, with oligonucleotides that amplified approximately 100 bp of the *asr2819* gene (*asr2819*-11/*asr2819*-12), using the *rnpB* gene as a control.

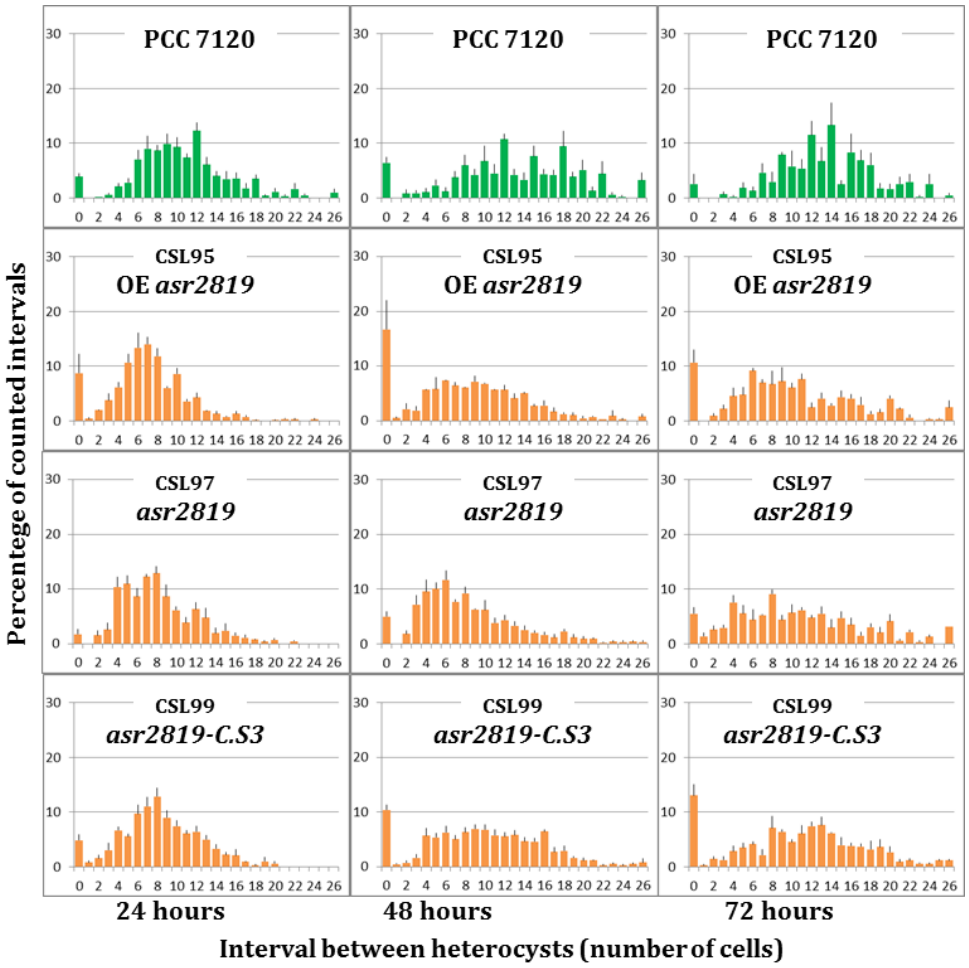


Figure 4.22. Heterocyst distribution in *asr2819* mutant strains. The frequency of intervals at 24, 48 and 72 h of nitrogen deprivation is shown in the graphs. In the cases in which the frequency is over 30%, the exact value is indicated.

The expression level of *asr2819* in the CSL95 overexpression strain was comparable to that of the *rnpB* gene, more than 1000-fold those in the wild-type or the CSL16 strain in ammonium-containing medium, and 200 to 400-fold after nitrogen deprivation. As expected, the expression level in the *asr2819* deletion mutant was negligible. The differences observed at the different times were statistically significant after 24 hours of N step-down (although not at 3 h), as assessed by the REST 2009 Software. Therefore, the expression of *asr2819* is upregulated in response to the lack of combined nitrogen.

Table 4.10. Ratios of the *asr2819* expression levels in the indicated strains.

Sample pairs	Relative *asr2819* expression	S. E. range	*P*
CSL16-0 vs. WT-0	1,033	0,051 - 22,327	0,951
CSL97-0 vs. WT-0	0,014	0,002 - 0,255	0,004
CSL95-0 vs. WT-0	1.112,836	237,296 - 4.071,805	0,000
WT-3 vs. WT-0	1,437	0,232 - 6,771	0,520
WT-24 vs. WT-0	3,403	1,481 - 9,749	0,000

Data from RT-qPCR normalized with the *rnpB* gene analysed using the REST 2009 software. Two independent RNA preparations were used.

4.5.3. Study of the connection with cell division

A possible relationship of the *hetC* and *asr2819* genes with cell division was studied by immunofluorescence using anti-FtsZ antibodies in the wild type, the *hetC-p* deletion mutant and the *asr2819* deletion strain. FtsZ is the protein that forms the Z ring during the cell division and has been suggested to be inhibited during heterocyst differentiation (Kuhn *et al.*, 2000). The experiment was performed using a primary antibody concentration of 1:100 and filaments taken 14 hours after N step-down. At this time, the culture is not growing actively and the cell division may have been synchronized in all the cells.

The FtsZ ring appeared only in independent cells at a low frequency in the wild-type strain, in contrast to the strains CSL16 (*hetC-p*) and CSL97 (*asr2819*) in which the ring could be observed at a higher frequency and in clusters of cells (
Figure **4.23**). This observation supports the connection between HetC and cell division, and suggests that Asr2819 is, directly or indirectly, also involved in this process.

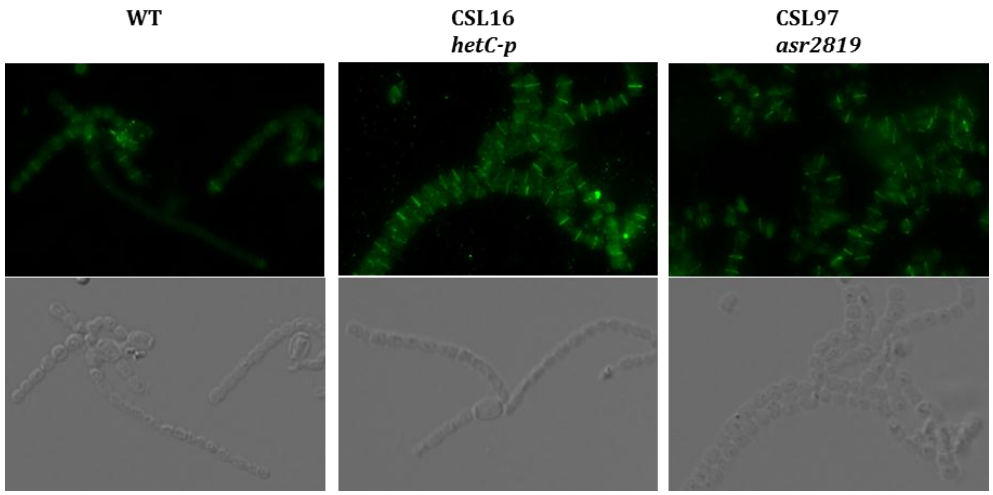


Figure 4.23. Immunofluorescence detection of FtsZ in *Anabaena* wild type and *hetC*-p and *asr2819* mutants. Filaments were incubated in $BG11_0C$ medium for 14 hours and tested with antibodies raised against the FtsZ protein. Visualization of the filaments by fluorescence microscopy; bright field and green fluorescence shown.

4.6. EXPRESSION OF *hetR* IN DIFFERENT MUTANT BACKGROUNDS

The HetR protein is a principal regulator of the heterocyst differentiation process. It presents positive autoregulation and is inhibited by PatS (Feldmann et al., 2011; Khudyakov & Golden, 2004). To study the effect of some of the mutations here investigated on *hetR* expression, expression of a *hetR-gfp* fusion was analysed in different backgrounds, wild type, *patS*, *hetR*, *hetC-p*, and *patS hetC-p*. Bubbled cultures growing in $BG11_0$+NH_4 medium were shifted to nitrogen deprivation conditions, and observed by confocal microscopy at the indicated times (Figure 4.24).

Under these conditions, *hetR-gfp* in the wild-type background was expressed more in some cells than in others, although at a low level, even under ammonium-rich conditions. At 5 hours after nitrogen deprivation, high expression was already localized to specific cells. However, in a *hetR* deletion background, the expression was very low and constant over time, consistent with the known positive autoregulation of *hetR*.

In the *patS* mutant background, expression in a culture with combined nitrogen was higher than in the wild-type background, and more cells showed green fluorescence. At 5 hours after N step-down there were still unresolved clusters of fluorescent cells. However, at 24 hours only a few doublets were found, suggesting the action of the other negative regulator HetN. This indicates that PatS is being expressed at a low level in the presence of combined nitrogen and inhibits *hetR* expression, which seems to be taking place under these conditions too.

In strain CSL88 (*hetR-gfp* in a *hetC-p* background), lower fluorescence than in the wild-type background was detected at early times. However, 5 hours after N step-down there was already fluorescence in specific cells, which was also the case at 24 hours. This is consistent with the results of a *hetR-luxA* fusion in a *hetC* background (Xu & Wolk, 2001).

Strain CSL87 (*hetR-gfp* in a *patS hetC-p* background) presented many fluorescent cells at 24 hours, in agreement with a very high heterocyst frequency (18.54%, see Table 4.4). Surprisingly, under ammonium-rich conditions there were also a high number of cells showing green fluorescence, even more than in the *patS* background. Thus, it seems that *hetC* is also expressed in the presence of combined nitrogen playing a role in the regulation of *hetR* expression.

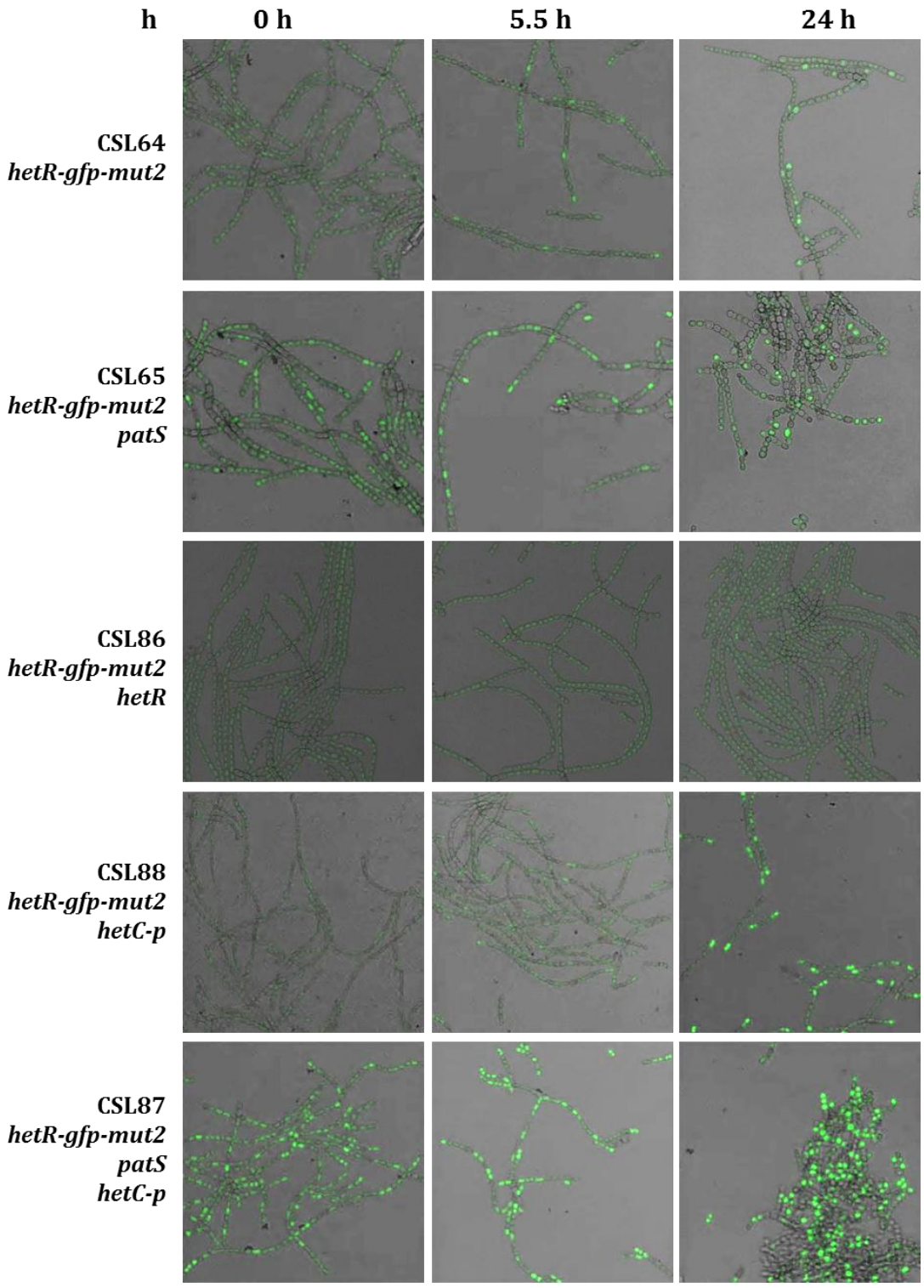


Figure 4.24. Expression of *hetR-gfp* in different genetic backgrounds. Filaments incubated in bubbled cultures with BG110 medium were visualised by confocal microscopy at the indicated times. A merged image of green fluorescence from GFP and bright field is shown. HetR-GFP in the wild-type background (strain CSL64) is shown as a control.

5. DISCUSSION

5.1. THE PatS PEPTIDE

The *patS* gene is a key player in the establishment of the spatial pattern of heterocysts in *Anabaena* sp. PCC 7120. Despite the fact that it was discovered a long time ago, little is known about the processing and transport of the peptide, since published work has focused on PatS activity. In this work, we have generated and studied a number of strains expressing, as the only version of the gene and from the native *patS* promoter in the chromosome, variants of the *patS* gene.

The *patS* ORF contains two possible methionine-encoding codons, the first and the fifth, which until now made it uncertain to predict the translation start of the *patS* gene. Two different results indicate that the first of these codons is the principal *patS* translational start. First, fluorescence from the *gfp* inserted in frame after the first M-encoding codon (in strain CSL21) or the fifth M-encoding codon (strain CSL19) is observed in cells with a specific distribution along the filament (Figure 4.1), indicating that M^1 can be used as a translational start. Second, the change M^1A (in strain CSL44) results in a phenotype of percentage and distribution of heterocysts similar to that of the *patS* mutant (strain CSVT20), whereas substitution M^5A (in strain CSL49) produces little changes compared to the wild type (Table 4.2; Figure 4.2). Therefore M^1, but not M^5, is required for PatS function and, consequently, the native PatS consists of a 17-amino acid peptide.

The PatS peptide has been compared with signalling peptides of Gram-positive bacteria, which are involved in the regulation of sporulation, quorum sensing and virulence, such as PhrA, PhrC and PhrE pentapeptides that are produced by processing of precursor molecules and secretion (Lazazzera *et al.*, 1997; Stephenson *et al.*, 2003). PhrA, PhrC, PhrE are synthesized as pre-peptide molecules with an N-terminal hydrophobic domain and a C-terminal hydrophilic region separated by a potential signal peptidase I cleavage site (which would correspond to a V, N or A residue), implying that the carboxyl half of the protein is secreted from the cell (Stephenson et al., 2003). These intermediate forms are processed again to produce the active pentapeptide molecules that are transported into the target cell by oligopeptide permeases such as SpoOK (Perego & Hoch, 1996).

A PatS-derived peptide has been proposed to move from the producing proheterocyst to the adjacent cells, and the C-terminal peptide PatS-5 has been considered the active peptide (Yoon & Golden, 1998). We have detected by immunofluorescence an accumulation of a PatS-5 containing product in the cells adjacent to proheterocysts, in

which the gene is expressed. Indeed, a gradient of concentration decreasing away from the proheterocysts is observed (Figure 4.5). This represents a direct evidence for intercellular transfer of a PatS peptide. In the case of PatS, consistent with its small size, no signal peptide could be recognized using the prediction program SignalP 4.1. Thus, PatS would be exported from the producing proheterocysts by a novel mechanism that may involve release into the periplasm, from which could be imported by neighbouring vegetative cells, or direct cell-to-cell transfer through channels located in the intercellular septa.

Results from the study of mutant strains producing variants of PatS with insertions or amino acid substitutions (see Table 4.2; Figure 4.2) lead us to consider two different regions of the 17-residue peptide in *Anabaena* sp. strain PCC 7120 (Figure 5.1). In the N-terminal region, substitutions in the second (in strain CSL90), third (in CSL91) or fourth (in CSL92) positions of PatS lead to a decrease in the heterocyst frequency. The changes in V^{7} have strong effects on differentiation: whereas its substitution by A (in strain CSL93) impairs differentiation and its substitution by Q (in CSL51) abolishes differentiation, the latter effect can be compensated in part by introducing a V residue in the 5^{th} position of the peptide (in CSL62). The insertions in the N-terminal region present in strains CSL19, CSL21 and CSL22 lead to inhibition of the differentiation. Although to a lesser extent, the changes in N^{8} (strain CSL52) and F^{9} (strain CSL53) also inhibit differentiation. These results imply that alterations in the N-terminal region inhibit the differentiation into heterocysts of the producing cells. This can be explained if the producing cell loses the self-immunity against the peptide, perhaps due to impairment in the processing or transport outside the cell of a peptide whose inhibitory activity remains intact.

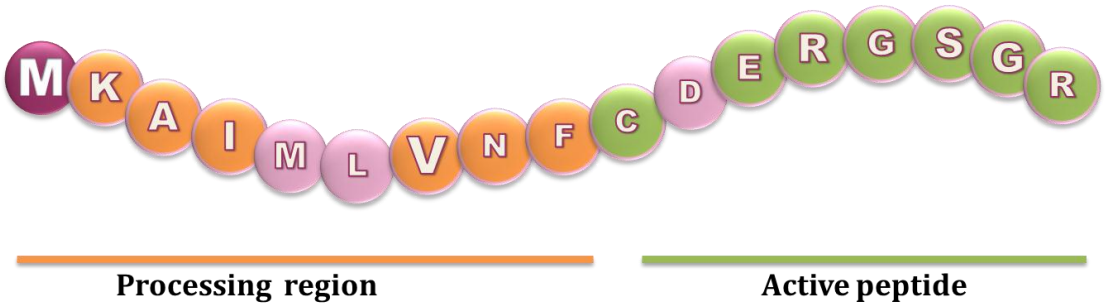


Figure 5.1. Scheme of the different regions of the PatS peptide. The *patS* gene is expressed as a 17-residue polypeptide including an N-terminal part involved in PatS processing and a C-terminal part required for the PatS regulatory function. The M-start codon is labelled in dark purple. An *Anabaena* strain carrying a mutation of a residue labelled bright purple expresses a phenotype similar to the wild-type phenotype. A strain carrying a mutation of a residue labelled orange exhibits less heterocyst differentiation than the wild type. A strain carrying a mutation of a residue labelled green exhibits more heterocyst differentiation than the wild type.

In the C-terminal region of PatS, whereas C^{10}, E^{12}, G^{14} and R^{17} appear to be important, R^{13}, S^{15} and G^{16} are essential for the inhibitory function of PatS in cells adjacent to the producing cells. Therefore, the C-end of PatS is involved in the inhibitory activity, consistent with previous reports (Wu *et al.*, 2004), although our work reveals a wider inhibitory region than that considered before. Thus, our results permit to differentiate two regions in the PatS peptide: the C-terminal region comprises the active peptide, whereas the N-terminal region would be necessary for processing and/or transport outside of the producing cell.

The processing site and the sequence of the active peptide remain unclear. However, the available data suggests that a peptide longer than 5 amino acids is necessary to be exported from the producing cell. Previously, *patS4* to *patS8* minigenes were expressed in *Anabaena* from promoters with different cell specificity (Wu et al., 2004). When *patS5* was expressed in proheterocysts in the *patS* background, from either the *patS* promoter or the (pro)heterocyst-specific *hepA* promoter, no or only a weak inhibition could be detected. These results are consistent with those obtained with strain CSL45 producing PatS-5 (Figure 4.2; Table 4.2; Figure 4.5), which exhibits a Mch phenotype but shows a confinement of the peptide in the producing cell. Therefore, PatS-5 expressed in proheterocysts is able to produce neither inhibition in the producing cells nor cell-cell signalling. In contrast, PatS-8 (in strain CSL47), which includes all the residues that our results have implicated in the inhibitory activity of PatS, appears to produce a regulation similar to that produced by wild-type PatS. In a previous report, when expressed from the copper-inducible *petE* promoter, which is active mainly in vegetative cells, *patS8*, and to a lesser extent, *patS6* showed levels of inhibition comparable to those of *patS5,* both in the wild type and in a *patS*-null mutant (Wu et al., 2004). Expression of *patS* or *patS5* from the *rbcL* promoter, which directs strong expression only in vegetative cells, completely inhibits heterocyst formation both in the wild type and the *patS* mutant (Wu et al., 2004). Additionally, PatS-5 added to the culture medium inhibits heterocyst differentiation (Yoon & Golden, 1998; Wu et al., 2004). Thus, PatS-5, which is not active in the proheterocysts and would not be exported from these cells, inhibits heterocyst differentiation when placed in the vegetative cells. It appears that, in contrast to the receptor vegetative cells, the producing proheterocysts are insensitive to the small processed peptides whereas, as deduced from the phenotype of strains expressing peptides with point substitutions in the N-terminal region, they are sensitive to the unprocessed PatS, at least when it is concentrated in the cell.

Regarding the putative processing site, our data suggest residue V^{7} as a candidate. This residue is positioned between the hydrophobic and hydrophilic regions of the

peptide, and its substitution produces a strong inhibition of differentiation, only alleviated by the introduction of another V residue in a closer position. Whether this is actually the cleavage site and the resulting C-terminal 10-residue molecule is an intermediate of the smaller mature peptide will merit further investigation. Several reports have shown in vitro activity of synthetic PatS-5 at binding to HetR and inhibition of HetR-binding to DNA (Huang *et al.*, 2004; Risser & Callahan, 2007; Feldmann *et al.*, 2011; Du *et al.*, 2012). Binding to HetR is tighter for PatS-6 (ERGSGR) than for PatS-5, and binding is tighter for PatS-5 than for PatS-7 (DERGSGR). On the other hand, no binding has been detected for PatS-8 (CDERGSGR) (Feldmann *et al.*, 2012), which is in contrast to the results obtained with strain CSL47 (PatS-8). These results could reflect differences between in vitro and in vivo experiments or could indicate a second processing of the peptide during intercellular transfer or in the receptor cells.

Due to its small size, the presence of *patS* genes in genomic sequences is difficult to predict in the absence of experimental analysis. However, searching for RGSGR in the genomes of 13 heterocyst-forming cyanobacteria other than strain PCC 7120 translated in all six reading frames (Jeff Elhai, personal communication) identified genes encoding putative PatS peptides in 10 of them (annotated only in one case, that of *Nostoc punctiforme*, although the peptide of *Anabaena variabilis* would be identical to that of PCC 7120). Besides strain PCC 7120 and *A. variabilis*, in two other strains PatS would be 17-amino acid peptides with M^1 and M^5; three strains could produce 15-amino acid peptides with only one M start codon, and four other could produce longer PatS peptides, two of them with one M aligning to M^5 of PCC 7120. Thus, a considerable variability in the N-terminal part of putative PatS peptides among heterocyst-forming cyanobacteria is apparent, with a number of strains putatively producing PatS peptides with extended N-terminal fragments. However, it is possible that the active peptide corresponding to the C-terminal part is also variable, or even strain-specific, as it occurs in the case of the Phr peptides in *Bacillus* (Lanigan-Gerdes *et al.*, 2008), making the search difficult.

The expression of *hetR*, investigated with a *gfp* reporter fusion, showed a tendency to increase in nitrogen-replete conditions in the *patS* background as compared to the wild-type background (See Figure 4.23). In the wild-type, a differential basal expression is observed in different cells of a filament but these differences are increased in the *patS* mutant. At 24 hours after N step-down, the pattern of expression is similar to that of the wild-type background, presumably because HetN is already present to resolve the clusters of cells expressing *hetR-gfp*. It can be suggested that certain level of *hetR* expression takes place, even in ammonium-containing medium, in a *patS* background, since HetR (i) is not inhibited by PatS in this mutant and (ii) has been shown to induce its own expression

(Black *et al.*, 1993). Consistently, also strain CSL88 (*hetR-gfp-mut2 hetR*) does not show appreciable induction of *hetR-gfp* expression. Therefore, cells in a filament growing in ammonium-containing medium bear certain quantities of HetR and PatS, and the balance between them could contribute to the initial response of specific cells to nitrogen deprivation.

5.2. THE HetN PROTEIN

The HetN protein has been suggested to affect the differentiation process by inhibition of the regulator HetR; however, its mechanism of action is still unknown. HetN presents homology with short-chain dehydrogenases/reductases and, additionally, bears an internal ERGSGR peptidic sequence. Most enzymes in this family have a core structure with 250-350 residues in length, frequently with N-terminal or C-terminal transmembrane domains or signal peptides, and form part of multienzyme complexes. The conserved regions cover a variable N-terminal TGX_3GXG motif as part of the nucleotide binding region, and the active site includes a triad of catalytically important Ser, Tyr, Lys residues, of which Tyr is the most conserved residue within the whole family, apart from other conserved residues all over the sequence (Oppermann *et al.*, 2003). HetN and orthologs found in other heterocyst forming cyanobacteria bear all the critical residues for the enzyme function (see scheme in Figure 5.2). In fact, HetN seems to present ATP hydrolysis activity when assayed in vitro (Liu & Chen, 2009). Proteins with mutations in S^{142} and Y^{155} expressed from a multi-copy plasmid produce inhibition of heterocyst differentiation similar to the wild-type protein. However, a mutation in K^{179} only faintly allows heterocyst differentiation while maintaining the ATP hydrolysis activity (Liu & Chen, 2009). More recently, proteins with point mutations in the same residues expressed from a replicative plasmid have been described to totally prevent heterocyst differentiation similarly to the wild-type protein (Higa *et al.*, 2012).

The deletion of the whole gene or the ERGSGR hexapeptide produced nearly the same Mch heterocyst pattern (Figure 4.7); therefore, the hexapeptide sequence is necessary for the inhibitory function of HetN. In another work, the RGSGR sequence was mutated leading to the same conclusion (Higa et al., 2012). The hexapeptide sequence, or related sequences, is conserved in some heterocyst-forming cyanobacterial genomes, but not in all of them. It is possible that, as could be also the case for PatS, the sequence would not need to be totally conserved to produce inhibition of heterocyst differentiation.

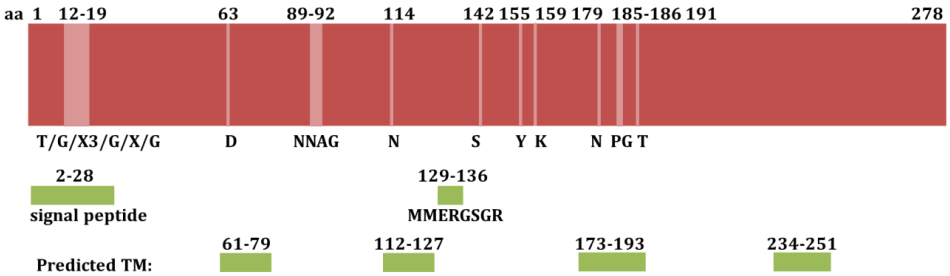


Figure 5.2. Scheme of highlighted features in the HetN protein. The HetN protein is indicated as a red rectangle. The different residues essential for function as a reductase are indicated in a brighter colour. The signal peptide and the ERGSGR peptide are indicated in green. The transmembrane (TM) segments predicted with the four different programs are also indicated in green.

Through the study of *gfp* fusions, the HetN protein has been observed to localize first in the cytoplasm and later in the periphery of the heterocysts (Figure 4.9). Post-translational translocation of proteins to the membranes seems to be commonly used by bacteria for soluble proteins, such as secretory proteins, which possess only moderately hydrophobic segments, or proteins that need to remain unfolded or loosely folded after their release from the ribosome (Rapoport, 2007). The HetN-GFP fusion protein seems to be located all over the periphery of the heterocyst but mostly in its poles, arranged as cup-shaped structures, but not inside the "heterocyst neck". It is possible that the protein is located in the "honeycomb" membranes, following a localization similar to that of the FraH protein (Merino-Puerto *et al.*, 2011). The C terminus of the protein is located in the cytoplasm, as both the HetN-sfGFP and the HetN-GFP are able to produce detectable fluorescence (Figure 4.9). In a previous study, the YFP was fused to HetN-derived polipeptides comprising the first 46, 172, 192 or 278 residues (Higa et al., 2012). Only the YFP after residue 192 or 278 showed fluorescence. Therefore, it is possible that the N-terminus of the protein lies in the periplasm or inside the thylakoid lumen. The change of localization over time during heterocyst differentiation allowed us to notice that the doublets present in strain CSL71 (*hetN-sfgfp*) contain heterocysts in different stages of the differentiation process. This would fit with a hypothesis in which the second round of differentiation would take place in a neighbour of a heterocyst.

The GFP fusion to the C-terminus of HetN leads to an increase in the heterocyst and heterocyst doublet frequency (Figure 4.7), likely due to a change in the folding or an impairment of the protein function. It was previously shown that deletion of a region comprising residues 177-195 produces a faint Mch phenotype, but that a deletion comprising residues 196-287 has almost no effect compared to the wild type (Higa et al., 2012). Finally, different insertions in residue 157 lead to an Mch phenotype (Black & Wolk, 1994). Thus, the C terminus may be important for the inhibitory function of HetN.

A strain bearing an insertion of a transposon in the N-terminal part of the protein exhibits a Het − phenotype; however, other insertions in the same place lead to a Mch phenotype (Black & Wolk, 1994). This contradictory effects were attributed to a possible influence on the expression of the neighbouring gene *hetI* (Black & Wolk, 1994). Deletion of HetN residues 2-28 in strain CSL73 produces a slight Mch phenotype (Figure 4.7) that, as shown with a *gfp* fusion in strain CSL103, can be related to a mislocalization of the protein (Figure 4.9). Therefore, as predicted by the SignalP software, HetN appears to bear a signal peptide comprising amino acids 2-28. A proper localization of HetN must be important but not essential for HetN function, as the phenotype of strain CSL73 is not as strong as that of the CSL7 mutant. It has been published that a deletion of a region comprising amino acids 2-46 did not produce any phenotypic effect in comparison with the wild type strain. In contrast, the deletion of a region comprising amino acids 47-128 produced a reduction in the heterocyst percentage (Higa et al., 2012). It is possible that the effect of this larger deletion (2-46) reflects a balance between the functions of a region that inhibits differentiation and another that promotes it. It could be also possible that the region of residues 28-128 may affect the *hetI* gene, as was interpreted in the work by Black & Wolk (1994), an effect that would hide the Mch phenotype produced by the deletion of the 2-28 residues region.

5.3. HetC, HetP AND Asr2819

HetC is an essential element in heterocyst differentiation (Khudyakov & Wolk, 1997), and it presents homology with bacteriocin ABC exporters (Michiels *et al.*, 2001). The epistatic effects between HetC, a positive element in the differentiation process, and two negative elements, PatS and HetN, have studied in this work. The lack of PatS, but not that of HetN, partially compensates for the lack of HetC (Table 4.4). The double mutant CSL30 (*hetC hetN*) presents a phenotype similar to that of strain CSL3 (*hetC*) (Table 4.4), and therefore the mutation of *hetC* is epistatic over that of *hetN*, consistent with the fact that the *hetC* gene is expressed before the *hetN* gene during the differentiation process. The double mutant strain CSL1 (*patS hetC*) forms doublets at a high frequency , but also shows short intervals between heterocysts, which resembles the Msh phenotype of the *patN* mutant (Risser *et al.*, 2012). Moreover, the heterocysts found in the double mutant produce heterocyst-specific glycolipids and express at a low level the *nifHDK* genes encoding nitrogenase (Figure 4.13), although they do not show cyanophycin granules or normal nitrogenase activity levels (Table 4.5). The fact that the mutation of *patS* is epistatic over that of *hetC* to a certain extent suggests a relationship between their

products. However, the connection may not be direct, as suggested by the fact that the complementation is partial and that the phenotype of the double mutant is different from that of any of the parental strains.

Bacteriocin ABC exporters usually bear a peptidase domain. Indeed, we have shown here that the putative peptidase domain of HetC is necessary for HetC function (Table 4.4). We have studied the possibility that the putative HetC peptidase is involved in processing of the PatS peptide. PatS-8 and PatS-17, in strains CSL101 (*patS hetC-p patS-8*) and CSL102 (*patS hetC-p patS-17*), both complement the *patS* mutation but not the *hetC-p* mutation (Figure 4.14). Therefore, the peptidase domain of HetC appears not to process the PatS-17 peptide, although the possibility that HetC could transport a PatS peptide processed by a mechanism independent of HetC cannot be discarded.

The Msh phenotype of the double-mutant strains CSL1 (*patS hetC*) and CSL17 (*patS hetC-p*) is notable (Table 4.4, Figure 4.11). It indicates that the lack of *hetC* leads to a decrease in the number of heterocyst doublets in addition to a decrease in the size of vegetative cell intervals between heterocysts. When comparing the triple mutant strains CSL15 (*patS hetC hetN*) and CSL31 (*patS hetC-p hetN*) with the double mutant CSL11 (*patS hetN*), the same effect can be noticed. This tendency can be only observed in a *patS* background that leads to increased differentiation (Figure 4.23). Consistent with this, more cells show induction of *hetR* expression, seen with a *gfp* fusion, in nitrogen-replete conditions in a *hetC-p patS* background than in a *patS* background. On the other hand, *hetR* expression in a *hetC-p* background was lower than in the wild-type background when the strain was growing in ammonium-containing medium, but 24 hours after nitrogen deprivation expression was usually seen in groups of cells, indicating that HetN cannot compensate for the lack of HetC as it does for the lack of PatS. Thus, the puzzling function of HetC in heterocyst differentiation remains uncertain, but the results obtained in this work fit with a function in heterocyst patterning independent of the PatS-HetN pathway.

The genes encoding ABC transporters similar to HetC that export toxic peptides are frequently linked to genes encoding a protein belonging to the membrane-fusion protein family that spans the periplasmic space linking the ABC exporter to an outer membrane channel, the gene encoding the substrate and a gene that encodes an immunity protein (Michiels et al., 2001). Because no homolog to membrane-fusion proteins has been detected in the *hetC* genomic region, it has been speculated that the HetC substrate could be released to the *Anabaena* periplasm rather than to the extracellular medium (Michiels et al., 2001). However, it is also possible that HetC acts in connection with a different type of protein to produce intercellular molecular transfer. Because of that, the relationship

between HetC and its neighbour HetP, which also has a positive role in heterocyst differentiation, was studied in this work.

The subcellular localization of HetC and HetP in the *Anabaena* filament was studied by making use of C-terminal GFP fusion proteins. In mature heterocysts, HetC-GFP is localized throughout the heterocyst surface and appears especially concentrated at the heterocyst poles (Figure 4.15). The fact that this protein includes at least 6 transmembrane segments predicted by different bioinformatics programs permits to consider it an integral membrane protein (Figure 4.10). The concentration at the poles could represent a specific localization of the protein or could result from the concentration of membranes found in the heterocyst neck. HetP-GFP is localized all over the cell area in proheterocysts and later almost restricted to the cell poles in mature heterocysts (Figure 4.18). Bioinformatics programs predict that HetP could bear a single membrane segment with unclear position within the protein (Figure 4.17). This segment could maintain HetP bound to membranes at the cell poles. Strains CSL67 (HetP-sfGFP) and CSL107 (HetP-GFP-mut2), which produce specific GFP proteins with different folding properties, showed fluorescence (Figure 4.18), indicating that the C-terminus of HetP is located in the cytoplasm. Western analysis showed that HetP is, in part, associated to the membranes after 18 hours of nitrogen deprivation (Figure 4.19). In the strains bearing the HetP-GFP fusion (CSL67 and CSL68), levels of expression of *hetP* mRNA are ca. 3-fold higher than in the wild type (Table 4.8). This difference in the gene expression appears to be sufficient for a total bypass of the *hetC-p* mutation, leading to restoration of the pattern and functionality of the heterocysts (Figure 4.20; Table 4.7). The co-localization of both proteins in the heterocyst poles and the full complementation by overexpression of *hetP* of the *hetC* mutation are consistent with a functional relationship between the HetC and HetP proteins.

The possibility that the product of gene *asr2819*, located downstream of *hetP* in the genome, were a putative substrate of HetC was considered (Figure 5.3). It encodes a small peptide that contains a Gly-Ala dipeptide in its N-terminal sequence, but not the rest of the processing target sequence of bacteriocins (Dirix *et al.*, 2004). Transcription of *asr2819* increases about 3-fold after 24 hours of N step-down (Table 4.10; see also Flaherty *et al.* (2011)), which would be consistent with a role in heterocyst differentiation. Regarding the effects of altering the expression levels of *asr2819*, the fact that both over-expression and deletion result in slight increases of heterocyst frequency is intriguing (Figure 4.21; Table 4.9). However, the clearest phenotypic effects are an increase in the frequency of contiguous heterocysts as a result of overexpression (in strain CSL95), the loosening of the patterned distribution of heterocysts, especially at long times after N step-down, as a

result of gene inactivation (in strain CSL97); and the massive differentiation in the triple mutant *patS hetN asr2819* (in strain CSL106). Also, the epistasis of the *hetC* mutation over any change in *asr2819* is noticeable. These results would be consistent with the hypothesis that Asr2819 exerts a positive effect on differentiation. By means of immunolocalization we have observed a different pattern of FtsZ localization in *hetC* and *asr2819* mutants as compared to the wild type. This would also be consistent with previous results showing an upregulation of *ftsZ* in a *hetC* mutant (Xu & Wolk, 2001). Therefore, a relationship, direct or indirect, of both HetC and Asr2819 with cell division can be considered. Heterocysts are terminally differentiated cells that need to turn the cell division machinery off before they are committed to differentiation. The possibility that HetC, HetP and Asr2819 act together to switch off the cell division machinery would fit with a positive role in the heterocyst differentiation. However, it could also be hypothesized that these proteins are involved in the export from a differentiating cell of a positive element that promotes differentiation in the neighbouring cells to form a cluster. Since its expression is controlled by nitrogen deprivation and its impairment has effects on heterocyst differentiation we decided to name *asr2819* the *patC* gene.

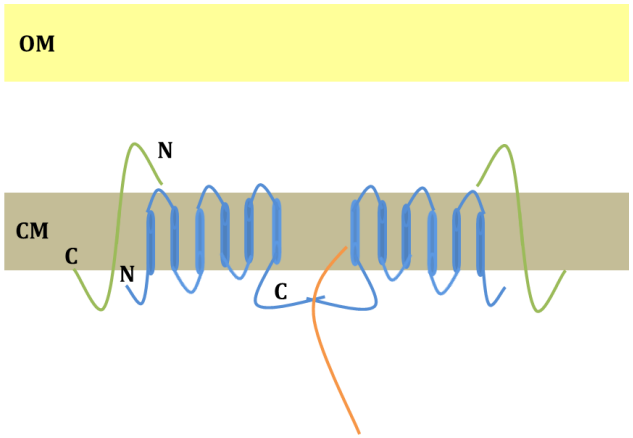


Figure 5.3. Model of the possible mechanism of export of Asr2819. The scheme is based on the model proposed for the export of the bacteriocin HlyA of *E. coli* (Gentschev et al., 2002), adapted to the data obtained for HetC, HetP and Asr2819. The outer membrane (OE), the cytoplasmic membrane (CM), and the C-terminus and N-terminus of the proteins are indicated. In this model, HetC, represented in blue, would be located in the membrane as a dimer; HetP, in green, would also act as a dimer assisting the processing or the export and Asr2819, in orange, would be transported to the periplasm or directly to the neighbour cell.

6. CONCLUSIONS

1. The *patS* gene encodes a 17-residue peptide. The native peptide needs to be processed and exported from the proheterocyst to the neighbour cells to function, forming a gradient that can be noticed by immunofluorescence.

2. The peptide has two different regions. The N-terminus of the peptide is recognised by processors and/or exporters, since its impairment lead to the inhibition of the differentiation, being the V^7 a candidate for site of cleavage. The C-terminus, comprising the last 8-10 amino acids, comprises the active peptide, which is exported from the proheterocyst.

3. The HetN protein bears an internal sequence codifying ERGSGR peptide that is essential for the inhibitory function. It is first located in the cytoplasm and later in the cell poles, in a cup-shaped structure that remembers the localization on the FraH protein. It also bears a signal peptide that is necessary for the localization, although the protein can work even when delocalized.

4. HetC bears a peptidase domain essential for its function, which is similar to those found in bacteriocin exporters. The role of this region may be to process a peptide involved in heterocyst differentiation different to PatS. Epistasis analyses confirm that the relationship between PatS and HetC is indirect.

5. HetC, and its genomic neighbour HetP, are both localized in the heterocyst poles. A 3-fold overexpression of *hetP* totally bypasses the deletion of the peptidase domain. This, together with the fact that the *hetP* mutant presented the same phenotype as the *hetC* mutant, highlights a direct relationship between HetC and HetP.

6. The *asr2819* gene, named *patC*, is induced 3-fold after 24 hours of nitrogen deprivation. It has a positive role in heterocyst differentiation, since its overexpression leads to increase of the number of doublets and its deletion leads to loss of the regulated pattern. PatC and HetC have a relationship, direct or indirect, with cell division, as FtsZ localization is upregulated in the deletion mutant.

7. There are certain quantities of HetR, PatS and HetC proteins even in nitrogen-replete conditions. The balances between the different levels of these proteins that are differentially expressed in the cells of a filament contribute to determine which cell will become a heterocyst in the event of N step-down.

7. REFERENCES

Abed, R.M.M., Dobretsov, S. & Sudesh, K. (2009) Applications of cyanobacteria in biotechnology. *J Appl Microbiol* **106**: 1-12.

Bassler, B.L. (2002) Small talk. Cell-to-cell communication in bacteria. *Cell* **109**: 421-424.

Bauer, C.C., Ramaswamy, K.S., Endley, S., Scappino, L.A., Golden, J.W. & Haselkorn, R. (1997) Suppression of heterocyst differentiation in *Anabaena* PCC 7120 by a cosmid carrying wild-type genes encoding enzymes for fatty acid synthesis. *FEMS Microbiol Lett* **151**: 23-30.

Black, T.A., Cai, Y. & Wolk, C.P. (1993) Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol Microbiol* **9**: 77-84.

Black, T.A. & Wolk, C.P. (1994) Analysis of a Het⁻ mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. *J Bacteriol* **176**: 2282-2292.

Borthakur, P.B., Orozco, C.C., Young-Robbins, S.S., Haselkorn, R. & Callahan, S.M. (2005) Inactivation of *patS* and *hetN* causes lethal levels of heterocyst differentiation in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol* **57**: 111-123.

Bryce, T.A., Welti, D., Walsby, A.E. & Nichols, B.W. (1972) Monohexoside derivatives of long-chain polyhydroxy alcohols; a novel class of glycolipid specific to heterocystous algae. *Phytochemistry* **11**: 295-302.

Buikema, W.J. & Haselkorn, R. (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes Dev* **5**: 321-330.

Burnat, M., Herrero, A. & Flores, E. (2014) Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocyst-forming cyanobacterium. *Proc Natl Acad Sci U S A*.

Busby, S. & Ebright, R.H. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199-213.

Cai, Y. & Wolk, C.P. (1997) *Anabaena* sp. strain PCC 7120 responds to nitrogen deprivation with a cascade-like sequence of transcriptional activations. *J Bacteriol* **179**: 267-271.

Callahan, S.M. & Buikema, W.J. (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Microbiol* **40**: 941-950.

Camargo, S., Valladares, A., Flores, E. & Herrero, A. (2012) Transcription activation by NtcA in the absence of consensus NtcA-binding sites in an anabaena heterocyst differentiation gene promoter. *J Bacteriol* **194**: 2939-2948.

Cardemil, L. & Wolk, C.P. (1979) The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Structure of the basic repeating unit. *J Biol Chem* **254**: 736-741.

Claessen, D., Rozen, D.E., Kuipers, O.P., Sogaard-Andersen, L. & van Wezel, G.P. (2014) Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol* **12**: 115-124.

Corrales-Guerrero, L., Mariscal, V., Flores, E. & Herrero, A. (2013) Functional dissection and evidence for intercellular transfer of the heterocyst-differentiation PatS morphogen. *Mol Microbiol* **88**: 1093-1105.

Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J. & Michiels, J. (2004) Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide in silico screening for peptides containing a double-glycine leader sequence and their cognate transporters. *Peptides* **25**: 1425-1440.

Ehira, S. & Ohmori, M. (2006a) NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **188**: 8520-8525.

Ehira, S. & Ohmori, M. (2006b) NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* **59**: 1692-1703.

Fay, P., Stewart, W.D., Walsby, A.E. & Fogg, G.E. (1968) Is the heterocyst the site of nitrogen fixation in blue-green algae? *Nature* **220**: 810-812.

Feldmann, E.A., Ni, S., Sahu, I.D., Mishler, C.H., Levengood, J.D., Kushnir, Y., McCarrick, R.M., Lorigan, G.A., Tolbert, B.S., Callahan, S.M. & Kennedy, M.A. (2012) Differential binding between PatS C-terminal peptide fragments and HetR from *Anabaena* sp. PCC 7120. *Biochemistry* **51**: 2436-2442.

Feldmann, E.A., Ni, S., Sahu, I.D., Mishler, C.H., Risser, D.D., Murakami, J.L., Tom, S.K., McCarrick, R.M., Lorigan, G.A., Tolbert, B.S., Callahan, S.M. & Kennedy, M.A. (2011) Evidence for direct binding between HetR from *Anabaena* sp. PCC 7120 and PatS-5. *Biochemistry* **50**: 9212-9224.

Fernández-Piñas, F., Leganés, F. & Wolk, C.P. (1994) A third genetic locus required for the formation of heterocysts in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **176**: 5277-5283.

Flaherty, B.L., Van Nieuwerburgh, F., Head, S.R. & Golden, J.W. (2011) Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics* **12**: 332.

Flores, E. & Herrero, A. (2010) Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* **8**: 39-50.

Flores, E., Pernil, R., Muro-Pastor, A.M., Mariscal, V., Maldener, I., Lechno-Yossef, S., Fan, Q., Wolk, C.P. & Herrero, A. (2007) Septum-localized protein required for filament integrity and diazotrophy in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **189**: 3884-3890.

Frías, J.E., Flores, E. & Herrero, A. (1994) Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol* **14**: 823-832.

Gentschev, I., Dietrich, G. & Goebel, W. (2002) The *E. coli* alpha-hemolysin secretion system and its use in vaccine development. *Trends Microbiol* **10**: 39-45.

Gerdtzen, Z.P., Salgado, J.C., Osses, A., Asenjo, J.A., Rapaport, I. & Andrews, B.A. (2009) Modeling heterocyst pattern formation in cyanobacteria. *BMC Bioinformatics* **10 Suppl 6**: S16.

Golden, J.W. & Yoon, H.S. (2003) Heterocyst development in *Anabaena*. *Curr Opin Microbiol* **6**: 557-563.

Herrero, A., Muro-Pastor, A.M. & Flores, E. (2001) Nitrogen control in cyanobacteria. *J Bacteriol* **183**: 411-425.

Herrero, A., Muro-Pastor, A.M., Valladares, A. & Flores, E. (2004) Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev* **28**: 469-487.

Higa, K.C. & Callahan, S.M. (2010) Ectopic expression of *hetP* can partially bypass the need for *hetR* in heterocyst differentiation by *Anabaena* sp. strain PCC 7120. *Mol Microbiol* **77**: 562-574.

Higa, K.C., Rajagopalan, R., Risser, D.D., Rivers, O.S., Tom, S.K., Videau, P. & Callahan, S.M. (2012) The RGSGR amino acid motif of the intercellular signalling protein, HetN, is required for patterning of heterocysts in *Anabaena* sp. strain PCC 7120. *Mol Microbiol* **83**: 682-693.

Huang, X., Dong, Y. & Zhao, J. (2004) HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. *Proc Natl Acad Sci U S A* **101**: 4848-4853.

Khudyakov, I. & Wolk, C.P. (1997) *hetC*, a gene coding for a protein similar to bacterial ABC protein exporters, is involved in early regulation of heterocyst differentiation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **179**: 6971-6978.

Khudyakov, I.Y. & Golden, J.W. (2004) Different functions of HetR, a master regulator of heterocyst differentiation in *Anabaena* sp. PCC 7120, can be separated by mutation. *Proc Natl Acad Sci U S A* **101**: 16040-16045.

Kim, Y., Joachimiak, G., Ye, Z., Binkowski, T.A., Zhang, R., Gornicki, P., Callahan, S.M., Hess, W.R., Haselkorn, R. & Joachimiak, A. (2011) Structure of transcription factor HetR required for heterocyst differentiation in cyanobacteria. *Proc Natl Acad Sci U S A* **108**: 10109-10114.

Kim, Y., Ye, Z., Joachimiak, G., Videau, P., Young, J., Hurd, K., Callahan, S.M., Gornicki, P., Zhao, J., Haselkorn, R. & Joachimiak, A. (2013) Structures of complexes comprised of Fischerella transcription factor HetR with *Anabaena* DNA targets. *Proc Natl Acad Sci U S A* **110**: E1716-1723.

Knoll, A. (2008) Cyanobacteria and earth history. In: The Cyanobacteria: Molecular Biology Genomics and Evolution. A. Herrero & E. Flores (eds). Norfolk, UK: Caister Academic Press, pp. 1-20.

Korkhov, V.M., Mireku, S.A. & Locher, K.P. (2012) Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F. *Nature* **490**: 367-372.

Kuhn, I., Peng, L., Bedu, S. & Zhang, C.C. (2000) Developmental regulation of the cell division protein FtsZ in *Anabaena* sp. strain PCC 7120, a cyanobacterium capable of terminal differentiation. *J Bacteriol* **182**: 4640-4643.

Kumar, K., Mella-Herrera, R.A. & Golden, J.W. (2010) Cyanobacterial heterocysts. *Cold Spring Harb Perspect Biol* **2**: a000315.

Lang, N.J. & Fay, P. (1971) The heterocysts of blue-green algae II. Details of ultrastructure. *Proc R Soc Lond B* **178**: 193-203.

Leganés, F., Fernández-Piñas, F. & Wolk, C.P. (1994) Two mutations that block heterocyst differentiation have different effects on akinete differentiation in Nostoc ellipsosporum. *Mol Microbiol* **12**: 679-684.

Li, B., Huang, X. & Zhao, J. (2002) Expression of *hetN* during heterocyst differentiation and its inhibition of *hetR* up-regulation in the cyanobacterium *Anabaena* sp. PCC 7120. *FEBS Lett* **517**: 87-91.

Li, J.H., Laurent, S., Konde, V., Bedu, S. & Zhang, C.C. (2003) An increase in the level of 2-oxoglutarate promotes heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Microbiology* **149**: 3257-3263.

Liang, J., Scappino, L. & Haselkorn, R. (1992) The patA gene product, which contains a region similar to CheY of Escherichia coli, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120. *Proc Natl Acad Sci U S A* **89**: 5655-5659.

Liu, D. & Golden, J.W. (2002) hetL overexpression stimulates heterocyst formation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **184**: 6873-6881.

Liu, J. & Chen, W.L. (2009) Characterization of HetN, a protein involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *FEMS Microbiol Lett* **297**: 17-23.

Liu, J. & Wolk, C.P. (2011) Mutations in genes patA and patL of *Anabaena* sp. strain PCC 7120 result in similar phenotypes, and the proteins encoded by those genes may interact. *J Bacteriol* **193**: 6070-6074.

López-Igual, R., Flores, E. & Herrero, A. (2010) Inactivation of a heterocyst-specific invertase indicates a principal role of sucrose catabolism in heterocysts of *Anabaena* sp. *J Bacteriol* **192**: 5526-5533.

Luque, I., Flores, E. & Herrero, A. (1994) Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* **13**: 5794.

Mariscal, V. & Flores, E. (2010) Multicellularity in a Heterocyst-Forming Cyanobacterium: Pathways for Intercellular Communication. In: Recent Advances in Phototrophic Prokaryotes. P.C. Hallenbeck (ed). Springer New York, pp. 123-135.

Mariscal, V., Herrero, A. & Flores, E. (2007) Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium. *Mol Microbiol* **65**: 1139-1145.

Meeks, J.C. & Elhai, J. (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev* **66**: 94-121; table of contents.

Meinhardt, H. (2008) Models of biological pattern formation: from elementary steps to the organization of embryonic axes. *Curr Top Dev Biol* **81**: 1-63.

Merino-Puerto, V., Herrero, A. & Flores, E. (2013) Cluster of genes that encode positive and negative elements influencing filament length in a heterocyst-forming cyanobacterium. *J Bacteriol* **195**: 3957-3966.

Michiels, J., Dirix, G., Vanderleyden, J. & Xi, C. (2001) Processing and export of peptide pheromones and bacteriocins in Gram-negative bacteria. *Trends Microbiol* **9**: 164-168.

Mitschke, J., Vioque, A., Haas, F., Hess, W.R. & Muro-Pastor, A.M. (2011) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc Natl Acad Sci U S A* **108**: 20130-20135.

Mullineaux, C.W., Mariscal, V., Nenninger, A., Khanum, H., Herrero, A., Flores, E. & Adams, D.G. (2008) Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. *EMBO J* **27**: 1299-1308.

Muro-Pastor, A.M., Flores, E. & Herrero, A. (2009) NtcA-regulated heterocyst differentiation genes *hetC* and devB from *Anabaena* sp. strain PCC 7120 exhibit a similar tandem promoter arrangement. *J Bacteriol* **191**: 5765-5774.

Muro-Pastor, A.M., Valladares, A., Flores, E. & Herrero, A. (1999) The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. *J Bacteriol* **181**: 6664-6669.

Muro-Pastor, A.M., Valladares, A., Flores, E. & Herrero, A. (2002) Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. *Mol Microbiol* **44**: 1377-1385.

Murry, M.A. & Wolk, C.P. (1989) Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. *Arch Microbiol* **151**: 469-474.

Nicolaisen, K., Mariscal, V., Bredemeier, R., Pernil, R., Moslavac, S., López-Igual, R., Maldener, I., Herrero, A., Schleiff, E. & Flores, E. (2009) The outer membrane of a heterocyst-forming cyanobacterium is a permeability barrier for uptake of metabolites that are exchanged between cells. *Mol Microbiol* **74**: 58-70.

Orozco, C.C., Risser, D.D. & Callahan, S.M. (2006) Epistasis analysis of four genes from *Anabaena* sp. strain PCC 7120 suggests a connection between PatA and PatS in heterocyst pattern formation. *J Bacteriol* **188**: 1808-1816.

Pernil, R., Herrero, A. & Flores, E. (2010) Catabolic function of compartmentalized alanine dehydrogenase in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **192**: 5165-5172.

Picossi, S., Flores, E. & Herrero, A. (2014) ChIP analysis unravels an exceptionally wide distribution of DNA binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. *BMC Genomics* **15**: 22.

Picossi, S., Valladares, A., Flores, E. & Herrero, A. (2004) Nitrogen-regulated genes for the metabolism of cyanophycin, a bacterial nitrogen reserve polymer: expression and mutational analysis of two cyanophycin synthetase and cyanophycinase gene clusters in heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *J Biol Chem* **279**: 11582-11592.

Rawlings, N.D., Barrett, A.J. & Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors *Nucleic Acids Research* **40**: D343-D350.

Rees, D.C., Johnson, E. & Lewinson, O. (2009) ABC transporters: the power to change. *Nat Rev Mol Cell Biol* **10**: 218-227.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M. & Stainer, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General Microbiology* **111**: 1-61.

Risser, D.D. & Callahan, S.M. (2007) Mutagenesis of *hetR* reveals amino acids necessary for HetR function in the heterocystous cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **189**: 2460-2467.

Risser, D.D. & Callahan, S.M. (2008) HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **190**: 7645-7654.

Risser, D.D., Wong, F.C. & Meeks, J.C. (2012) Biased inheritance of the protein PatN frees vegetative cells to initiate patterned heterocyst differentiation. *Proc Natl Acad Sci U S A* **109**: 15342-15347.

Shih, P.M., Wu, D., Latifi, A., Axen, S.D., Fewer, D.P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent, T., Goodwin, L., Nolan, M., Davenport, K.W., Han, C.S., Rubin, E.M., Eisen, J.A., Woyke, T., Gugger, M. & Kerfeld, C.A. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A* **110**: 1053-1058.

Stanier, R.Y. & Cohen-Bazire, G. (1977) Phototrophic prokaryotes: the cyanobacteria. *Annu Rev Microbiol* **31**: 225-274.

Valladares, A., Flores, E. & Herrero, A. (2008) Transcription activation by NtcA and 2-oxoglutarate of three genes involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **190**: 6126-6133.

Valladares, A., Herrero, A., Pils, D., Schmetterer, G. & Flores, E. (2003) Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. *Mol Microbiol* **47**: 1239-1249.

Vázquez-Bermúdez, M.F., Herrero, A. & Flores, E. (2000) Uptake of 2-oxoglutarate in Synechococcus strains transformed with the Escherichia coli kgtP gene. *J Bacteriol* **182**: 211-215.

Wang, Y. & Xu, X. (2005) Regulation by *hetC* of genes required for heterocyst differentiation and cell division in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **187**: 8489-8493.

Wilk, L., Strauss, M., Rudolf, M., Nicolaisen, K., Flores, E., Kuhlbrandt, W. & Schleiff, E. (2011) Outer membrane continuity and septosome formation between vegetative cells in the filaments of *Anabaena* sp. PCC 7120. *Cell Microbiol* **13**: 1744-1754.

Woese, C.R. (1987) Bacterial evolution. *Microbiol Rev* **51**: 221-271.

Wolk, C.P. (1967) Physiological basis of the pattern of vegetative growth of a blue-green alga. *Proc Natl Acad Sci U S A* **57**: 1246-1251.

Wolk, C.P., Ernst, A. & Elhai, J. (1994) Heterocyst Metabolism and Development. In: The Molecular Biology of Cyanobacteria. D.A. Bryant (ed). Springer Netherlands, pp. 769-823.

Wolk, C.P. & Quine, M.P. (1975) Formation of One-Dimensional Patterns by Stochastic Processes and by Filamentous Blue-Green Algae. *Dev Biol* **46**: 370-382.

Wong, F.C. & Meeks, J.C. (2001) The hetF gene product is essential to heterocyst differentiation and affects HetR function in the cyanobacterium Nostoc punctiforme. *J Bacteriol* **183**: 2654-2661.

Wu, X., Lee, D.W., Mella, R.A. & Golden, J.W. (2007) The *Anabaena* sp. strain PCC 7120 asr1734 gene encodes a negative regulator of heterocyst development. *Mol Microbiol* **64**: 782-794.

Wu, X., Liu, D., Lee, M.H. & Golden, J.W. (2004) *patS* minigenes inhibit heterocyst development of *Anabaena* sp. strain PCC 7120. *J Bacteriol* **186**: 6422-6429.

Xu, X., Elhai, J. & Wolk, C.P. (2008) Transcriptional and Developmental Responses by *Anabaena* to Deprivation of Fixed Nitrogen. In: The Cyanobacteria: Molecular Biology, Genomics and Evolution. A. Herrero & E. Flores (eds). Norfolk: Caister Academic

pp. 383–422.

Xu, X. & Wolk, C.P. (2001) Role for *hetC* in the transition to a nondividing state during heterocyst differentiation in *Anabaena* sp. *J Bacteriol* **183**: 393-396.

Yoon, H.S. & Golden, J.W. (1998) Heterocyst pattern formation controlled by a diffusible peptide. *Science* **282**: 935-938.

Yoon, H.S. & Golden, J.W. (2001) PatS and products of nitrogen fixation control heterocyst pattern. *Journal of Bacteriology* **183**: 2605-2613.

Zhang, J.Y., Chen, W.L. & Zhang, C.C. (2009) *hetR* and *patS*, two genes necessary for heterocyst pattern formation, are widespread in filamentous nonheterocyst-forming cyanobacteria. *Microbiology* **155**: 1418-1426.

Zhao, M.X., Jiang, Y.L., He, Y.X., Chen, Y.F., Teng, Y.B., Chen, Y., Zhang, C.C. & Zhou, C.Z. (2010) Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proc Natl Acad Sci U S A* **107**: 12487-12492.

Zhu, M., Callahan, S.M. & Allen, J.S. (2010) Maintenance of heterocyst patterning in a filamentous cyanobacterium. *J Biol Dyn* **4**: 621-633.